

## **Polysubstance addiction and psychiatric, somatic comorbidities among 7,989 individuals with cocaine use disorder: a latent class analysis**

Brendan Stiltner,<sup>1,2</sup> Robert H. Pietrzak,<sup>1,3</sup> Daniel S. Tylee,<sup>1,2</sup> Yaira Z. Nunez,<sup>1,2</sup> Keyrun Adhikari,<sup>1,2</sup> Henry R. Kranzler,<sup>4,5</sup> Joel Gelernter,<sup>1,2</sup> Renato Polimanti<sup>1,2</sup>

<sup>1</sup>Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA

<sup>2</sup>VA Connecticut Healthcare System, West Haven, CT, USA

<sup>3</sup>U.S. Department of Veterans Affairs National Center for Posttraumatic Stress Disorder, VA Connecticut Healthcare System, West Haven, CT, USA

<sup>4</sup>University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104, USA

<sup>5</sup>Mental Illness Research, Education, and Clinical Center, Crescenzo Veterans Affairs Medical Center, Philadelphia, PA, 19104, USA

**CORRESPONDENCE:** Renato Polimanti, PhD. Yale University School of Medicine, Department of Psychiatry. VA CT 116A2, 950 Campbell Avenue, West Haven, CT 06516, USA. Phone: +1 (203) 932-5711 x5745. Fax: +1 (203) 937-3897. E-mail: [renato.polimanti@yale.edu](mailto:renato.polimanti@yale.edu)

**RUNNING TITLE:** Comorbidities in Cocaine Use Disorder

**ACKNOWLEDGMENTS:** The authors thank the research participants enrolled in the Yale-Penn cohort.

**WORD COUNT** (excluding abstract, references, tables, and figures): 4092

**DECLARATIONS OF COMPETING INTEREST:** RP received a research grant from Alkermes. RP and JG are paid for their editorial work on the journal *Complex Psychiatry*. JG and HRK are named as inventors on PCT patent application #15/878,640 entitled: “Genotype-guided dosing of opioid agonists,” filed January 24, 2018. HRK is a member of advisory boards for Dicerna Pharmaceuticals, Sophrosyne Pharmaceuticals, and Enthion Pharmaceuticals; a consultant to Sophrosyne Pharmaceuticals; the recipient of research funding and medication supplies for an investigator-initiated study from Alkermes; a member of the American Society of Clinical Psychopharmacology’s Alcohol Clinical Trials Initiative, which for the past three years was supported by Alkermes, Ethypharm, Lundbeck, Mitsubishi, Otsuka, and Pear Therapeutics, and is paid for his editorial work on the journal *Alcohol: Clinical and Experimental Research*. The other authors have no competing interests to report.

**FUNDING:** This study was supported by the National Institutes of Health (R33 DA047527, R21 DC018098, and RF1 MH132337), One Mind, and the VISN 4 Mental Illness Research, Education and Clinical Center at the Crescenz VAMC. The Yale-Penn cohort was supported by multiple grants from the National Institutes of Health (RC2 DA028909, R01 DA12690, R01 DA12849, R01 DA18432, R01 AA11330, R01 AA017535).

**ETHICS:** Yale-Penn participants gave written informed consent as approved by the institutional review board at recruiting site.

**DATA AVAILABILITY:** All data produced in the present work are contained in the manuscript and its supplemental material.

## ABSTRACT

*Aims:* We performed a latent class analysis (LCA) in a sample ascertained for addiction phenotypes to investigate cocaine use disorder (CoUD) subgroups related to polysubstance addiction (PSA) patterns and characterized their differences with respect to psychiatric and somatic comorbidities.

*Design:* Cross-sectional study

*Setting:* United States

*Participants:* Adult participants aged 18-76, 39% female, 47% African American, 36% European American with a lifetime DSM-5 diagnosis of CoUD (N=7,989) enrolled in the Yale-Penn cohort. The control group included 2,952 Yale-Penn participants who did not meet for alcohol, cannabis, cocaine, opioid, or tobacco use disorders.

*Measurements:* Psychiatric disorders and related traits were assessed via the Semi-structured Assessment for Drug Dependence and Alcoholism. These features included substance use disorders (SUD), family history of substance use, sociodemographic information, traumatic events, suicidal behaviors, psychopathology, and medical history. LCA was conducted using diagnoses and diagnostic criteria of alcohol, cannabis, opioid, and tobacco use disorders.

*Findings:* Our LCA identified three subgroups of PSA (i.e., low, 17%; intermediate, 38%; high, 45%) among 7,989 CoUD participants. While these subgroups varied by age, sex, and racial-ethnic distribution ( $p < 0.001$ ), there was no difference on education or income ( $p > 0.05$ ). After accounting for sex, age, and race-ethnicity, the CoUD subgroup with high PSA had higher odds of antisocial personality disorder (OR=21.96 vs. 6.39, difference- $p = 8.08 \times 10^{-6}$ ), agoraphobia (OR=4.58 vs. 2.05, difference- $p = 7.04 \times 10^{-4}$ ), mixed bipolar episode (OR=10.36 vs. 2.61, difference- $p = 7.04 \times 10^{-4}$ ), posttraumatic stress disorder (OR=11.54 vs. 5.86, difference- $p = 2.67 \times 10^{-4}$ ), antidepressant medication use (OR=13.49 vs. 8.02, difference- $p = 1.42 \times 10^{-4}$ ), and sexually transmitted diseases (OR=5.92 vs. 3.38, difference- $p = 1.81 \times 10^{-5}$ ) than the low-PSA CoUD subgroup.

*Conclusions:* We found different patterns of PSA in association with psychiatric and somatic comorbidities among CoUD cases within the Yale-Penn cohort. These findings underscore the importance of modeling PSA severity and comorbidities when examining the clinical, molecular, and neuroimaging correlates of CoUD.

**KEYWORDS:** Cocaine use disorder; Polysubstance addiction; Substance use disorders; Latent class analysis; Comorbidities.

## INTRODUCTION

Substance use disorders (SUDs) represent heterogeneous patterns of behavioral, cognitive, and physiological features associated with the chronic use of a substance, despite problems that arise from its use (1). Among SUDs, cocaine use disorder (CoUD) contributes to significant morbidity, mortality, personal, and healthcare expense, posing a considerable public health problem (2). In 2019, nearly 5.5 million people in the United States reported using cocaine in the past year and 1 million met criteria for CoUD (3).

Polysubstance use and addiction are defined as the consumption and abuse of more than one substance over a short period of time or concurrently (4). These are both common and associated with adverse consequences (5). People who use and abuse multiple substances have worse treatment outcomes (6, 7), higher risk of mortality compared to people that use one substance (8), and are statistically more likely to experience an overdose, violence, and accidental injuries than those that use one substance (9). With respect to the risk of polysubstance use among cocaine users, 72.7% of cocaine-involved deaths registered in 2017 involved opioids (10).

Modeling patterns of polysubstance addiction (PSA) among substance users can help to define high-risk groups within diagnostic boundaries. Latent class analysis (LCA) is a statistical method that can be used to identify groups of people (referred to as classes) based on similar patterns of manifest response variables (11). This allows one to identify more homogenous groups within a heterogenous population (12). Studies have used LCA within samples of people who use substances or diagnosed with a SUD to evaluate patterns of association with comorbid SUDs and clinically relevant variables (13, 14). Previous studies examined nationally

representative samples (15, 16) and treatment-seeking populations (17-19). However, these types of study designs feature limited sample sizes for rare SUDs such as CoUD.

Latent variable modeling approaches have also been used to identify classes of people with certain patterns of symptoms. The approach has been used in posttraumatic stress disorder (20, 21) and individuals reporting manic episodes (22). These studies revealed that different latent symptom typologies are differentially related to meaningful clinical and neurobiological markers. However, in SUDs, these approaches have focused mainly on alcohol (23, 24) and cannabis use disorders (25-27). To date, there have been few efforts to model the criteria of other SUDs (28, 29) and, to our knowledge, none have assessed PSA patterns in individuals with CoUD. However, previous research has identified latent classes of cocaine dependence based on other approaches (30, 31).

Using data from the large Yale-Penn cohort, we investigated PSA latent classes in 7,989 participants with lifetime DSM-5 CoUD diagnoses. Within these latent classes of CoUD, we also characterized symptom profiles and tested their associations with clinically meaningful mental and physical health variables. Our results show that clusters of CoUD-affected individuals have different patterns of PSA and of psychiatric and somatic comorbidities. These findings have important implications for modeling CoUD heterogeneity in studies of addiction-ascertained cohorts (e.g., molecular, brain imaging, and treatment investigations).

## **METHODS AND MATERIALS**

### **Participants**

Participants in this study are from the Yale-Penn cohort, which was recruited to investigate the genetics of substance use disorders and their comorbid conditions (32-37).

Participants were recruited from five US sites: Yale School of Medicine (APT Foundation, New Haven, CT, USA), University of Connecticut Health Center (Farmington, CT, USA), University of Pennsylvania Perelman School of Medicine (Philadelphia, PA, USA), Medical University of South Carolina (Charleston, SC, USA), and McLean Hospital (Belmont, MA, USA). The studies were approved by the institutional review boards at each site and written informed consent was obtained from each participant.

## **Measures**

Participants were interviewed by trained personnel using the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA). The SSADDA produces lifetime diagnoses of substance use disorders and other mental illnesses (38, 39). The SSADDA includes items to diagnose substance dependence and abuse for the major substances of abuse, except for tobacco, for which there is no abuse diagnosis. In the present study, we derived lifetime DSM-5 diagnoses of SUD for alcohol, cannabis, cocaine, opioid, and tobacco. However, for DSM-5 tobacco use disorder, there are only 9 criteria available in the SSADDA. Additionally, we extracted from the SSADDA information regarding several psychiatric and behavioral traits. These included phenotypes relate to suicidality, family history of substance use, traumatic experiences, and psychopathology. Data related to socioeconomic status and physical health were also included in the analyses. Supplemental Table 1 describes the items included and how they were assessed.

## **Data analyses**

LCA was first performed including participants with a lifetime CoUD diagnosis (n=7,989) and using lifetime alcohol use disorder (AUD), cannabis use disorder (CaUD), opioid

use disorder (OUD), and tobacco use disorder (TUD) as manifest variables. We assigned participants to the resulting classes if they had posterior probability  $\geq 0.7$  as previously proposed (40, 41). Next, within each diagnosis-based latent class, we performed LCAs, in which we analyzed the specific criteria for each SUD separately (alcohol, opioid, cannabis, and tobacco). To determine the best-fitting solution, we fit models beginning with 2 classes and increased the number of classes by one at each step. We selected the best-fitting solution based on model-fit statistics, size of the smallest class, and the interpretability and applicability of the model (42). Goodness-of-fit statistics included the Bayesian information criterion (BIC) and Akaike information criterion (AIC) (42). For the criteria-based classes, we selected models that produced classes with  $\geq 5\%$  of the sample of the diagnosis-based class. When there was disagreement between these indicators, we selected models with the highest entropy (43) and most significant bootstrap likelihood ratio test (BLRT) (42).

Next, we investigated differences between each diagnosis-based latent class and Yale-Penn participants who did not meet criteria for any of the 5 SUDs of interest ( $n=2,952$ ). This analysis was performed using logistic regression models with age, sex, and self-reported race-ethnicity categories as covariates. The dependent variables included traits related to psychopathology, suicidality, traumatic experiences, family history of substance use, medical conditions, and socioeconomic status (SES). For variables on which multiple classes differed significantly from controls, we used z-tests to assess differences between the effect size of the associations observed with respect to the classes. All analyses were done in RStudio (44, 45). The poLCA package (46, 47) was used to conduct the LCAs and glca package (48) was used to test for significant differences between LCA models.

## RESULTS

### Sample Characteristics

Among 7,989 Yale-Penn participants with a lifetime diagnosis of CoUD, 39% were female, and the average age was 41 years (ranging from 18 to 76 years old, standard deviation = 9.6 years). The sample included mostly African Americans (47%) and European Americans (36%). With respect to comorbid SUDs, 82% met criteria for AUD, 57% for CaUD, 49% for OUD, and 86% for TUD. With regard to family history of substance use, 51% reported being aware of at least one family member using any drugs or alcohol before the participant was 13 years of age. Participants with CoUD had family members who used cocaine (9%), heroin (5%), abused prescription drugs (4%), or used other illegal drugs (16%), and 47% and 65% were aware of family members who drank alcohol or smoked cigarettes, respectively. Table 1 summarizes the sample characteristics.

### Diagnosis-Based Latent Class Analysis

Our initial LCA of CoUD-affected individuals was based on diagnoses of AUD, CaUD, OUD, and TUD. With only four dichotomous variables, we could identify only a two-class model (49). Considering a posterior probability  $\geq 0.70$ , we assigned 1,367 individuals (17%) to Class-1 and 3,616 individuals (45%) to Class-2. No participant had a posterior probability greater than 0.7 for both classes. Conversely, 38% of the CoUD sample ( $n=3,006$ ) had a posterior probability  $< 0.7$  with respect to both diagnosis-based latent classes. Based on these results and the prevalence of other SUDs in these three groups (Table 1), we defined three CoUD subgroups: a low-PSA subgroup (i.e., diagnosis-based Class-1), an intermediate PSA subgroup (i.e., the group of participants with posterior probabilities  $< 0.7$  for both diagnosis-based classes), and a high-PSA subgroup (i.e., diagnosis-based Class-2). Relative to the overall CoUD sample,



the low-PSA subgroup was older (42.5 vs. 40.7 years of age,  $p < 0.0001$ ), had a higher proportion of females (52% vs. 39%,  $p < 0.0001$ ) and non-Hispanic African Americans (56% vs. 47%,  $p < 0.0001$ ), and a smaller proportion of fewer non-Hispanic European Americans (26% vs. 36%,  $p < 0.001$ ). An opposite pattern was present for the high-PSA subgroup, which was slightly younger (39.2 vs. 40.7 years of age,  $p < 0.0001$ ), had a smaller proportion of females (30% vs. 39%,  $p < 0.0001$ ) and non-Hispanic African Americans (41% vs. 47%,  $p < 0.0001$ ), and a larger proportion of non-Hispanic European Americans (42% vs. 36%,  $p < 0.0001$ ) than the overall CoUD sample. The intermediate-PSA subgroup showed characteristics intermediate between the two other subgroups and more similar to the overall CoUD sample. No differences were observed between the CoUD subgroups and the overall CoUD sample on education or income ( $p > 0.05$ ).

The distribution of SUD diagnoses for the low-PSA subgroup had lower prevalence of SUD diagnoses than the overall CoUD sample: 30% vs. 82% for AUD ( $p < 0.0001$ ); 6% vs. 57% for CaUD ( $p < 0.0001$ ); 41% vs. 49% for OUD ( $p < 0.0001$ ), and 50% vs. 80% for TUD ( $p < 0.0001$ ). Conversely, the high-PSA subgroup included higher rates SUD of comorbidity, with 100% of individuals meeting criteria for AUD ( $p < 0.0001$ ), CaUD ( $p < 0.0001$ ), and TUD ( $p < 0.0001$ ). The intermediate-PSA subgroup showed a similar distribution of AUD (84% vs. 82%,  $p < 0.05$ ), OUD (46% vs. 49%,  $p < 0.05$ ), and TUD (87% vs. 86%,  $p > 0.05$ ), but a much lower proportion of CaUD cases (29% vs. 57%,  $p < 0.0001$ ) than the overall CoUD sample. Supplemental Table 2 reports the SUD distribution in African Americans and European Americans across the overall CoUD sample and three CoUD subgroups.

**Table 1.** Characteristics of the full CoUD sample and diagnosis-based LCA subgroups based on AUD, CaUD, OUD, and TUD diagnoses.

	<b>Full CoUD Sample (n=7,989)</b>	<b>Low-PSA Subgroup (n=1,367)</b>	<b>Intermediate- PSA Subgroup (n=3,006)</b>	<b>High-PSA Subgroup (n=3,616)</b>
<b>Age-years, median (IQR)</b>	41 (35-47)	42 (37-48)	42 (36-48)	40 (32-46)
<b>Sex, n (%)</b>				
Male	4879 (61)	652 (48)	1685 (56)	2542 (70)
Female	3110 (39)	715 (52)	1321 (44)	1074 (30)
<b>Race/Ethnicity, n (%)</b>				
African American, not Hispanic	3754 (47)	768 (56)	1488 (50)	1498 (41)
European American, not Hispanic	2905 (36)	360 (26)	1032 (34)	1513 (42)
African American, Hispanic	212 (3)	41 (3)	65 (2)	106 (3)
European American, Hispanic	504 (6)	82 (6)	182 (6)	240 (7)
Native American	60 (1)	18 (1)	18 (1)	24 (1)
Asian	12 (0)	2 (0)	5 (0)	5 (0)
Pacific Islander	7 (0)	0 (0)	5 (0)	2 (0)
<b>Socioeconomic Factors, n (%)</b>				
Highest education completed <sup>a</sup> , median (IQR)	5 (4-6)	5 (4-6)	5 (3-6)	5 (4-6)
High school diploma	2183 (27)	401 (29)	814 (27)	968 (27)
General Educational Development (GED)	1508 (19)	179 (13)	518 (17)	811 (22)
Currently in school	528 (7)	95 (7)	176 (6)	257 (7)
Employed past year	4957 (62)	768 (56)	1792 (60)	2397 (66)
Employed currently	2260 (28)	367 (27)	869 (29)	1024 (28)
Working full-time	1151 (14)	182 (13)	438 (15)	531 (15)
Household gross income <sup>b</sup> , median (IQR)	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)
<b>Substance use disorders, n (%)</b>				
AUD	6561 (82)	413 (30)	2532 (84)	3616 (100)
CaUD	4581 (57)	88 (6)	877 (29)	3616 (100)

	ODD	3909 (49)	560 (41)	1395 (46)	1954 (54)
	TUD	6903 (86)	684 (50)	2603 (87)	3616 (100)
<b>Family history of substance use, n (%)</b>					
	Any alcohol/drugs	4109 (51)	522 (38)	1531 (51)	2056 (57)
	Alcohol use	3736 (47)	460 (34)	1374 (46)	1902 (53)
	Cocaine use	686 (9)	64 (5)	229 (8)	393 (11)
	Heroin use	379 (5)	53 (4)	128 (4)	198 (5)
	Other illegal drugs	1313 (16)	110 (8)	423 (14)	780 (22)
	Prescription drug abuse	325 (4)	29 (2)	102 (3)	194 (5)
	Cigarette smoking	5197 (65)	775 (57)	1956 (65)	2466 (68)
<b>Suicidal Thoughts and Behaviors, n (%)</b>					
	Suicidal ideation	3391 (42)	460 (34)	1238 (41)	1693 (47)
	Suicide plan	1525 (19)	197 (14)	579 (19)	749 (21)
	Suicide attempt	1333 (17)	187 (14)	527 (18)	619 (17)
<b>Childhood Trauma, n (%)</b>					
	Any traumatic event (any age)	5519 (69)	819 (6)	1994 (66)	2706 (75)
	Witnessed violence (once)	1543 (19)	221 (16)	527 (18)	795 (22)
	Witnessed violence (2+)	920 (12)	137 (10)	285 (9)	498 (14)
	Victim of violence	514 (6)	80 (6)	182 (6)	252 (7)
	Abused sexually	1257 (16)	192 (14)	483 (16)	582 (16)
	Abused physically	786 (10)	91 (7)	295 (10)	400 (11)
<b>Psychopathology, n (%)</b>					
	ADHD	653 (8)	50 (4)	225 (7)	378 (10)
	ASPD	1314 (16)	89 (7)	383 (13)	842 (23)
	Conduct disorder	260 (3)	31 (2)	85 (3)	144 (4)
	Gambling disorder	869 (11)	88 (6)	272 (9)	509 (14)
	Agoraphobia, no Panic disorder	251 (3)	29 (2)	89 (3)	133 (4)
	Agoraphobia	511 (6)	54 (4)	189 (6)	268 (7)
	Bipolar I	393 (5)	34 (2)	134 (4)	225 (6)
	MDE 1	944 (12)	155 (11)	362 (12)	427 (12)
	MDE 2	612 (8)	75 (5)	230 (8)	307 (8)

MDD	1157 (14)	184 (13)	445 (15)	528 (15)
MDE	1355 (17)	187 (14)	210 (17)	648 (18)
Mixed episode	338 (4)	22 (2)	110 (4)	206 (6)
OCD	250 (3)	28 (2)	89 (3)	133 (4)
Panic disorder	548 (7)	66 (5)	211 (7)	271 (7)
PTSD	1397 (17)	186 (14)	511 (17)	700 (19)
Social phobia	446 (6)	44 (3)	157 (5)	245 (7)
Social phobia generalized	356 (4)	33 (2)	129 (4)	194 (5)
<b>Medical history/Physical health, n (%)</b>				
Rating of physical health <sup>c</sup> , median (IQR)	3 (2-4)	3 (2-4)	3 (2-4)	3 (2-4)
High blood pressure	2228 (28)	387 (28)	892 (30)	949 (26)
Migraine headaches	1272 (16)	215 (16)	483 (16)	574 (16)
Brain injury	1221 (15)	147 (11)	437 (15)	637 (18)
Unconscious >5 minutes	1370 (17)	182 (13)	490 (16)	698 (19)
Epilepsy	559 (7)	90 (7)	205 (7)	264 (7)
Meningitis/Encephalitis	127 (2)	22 (2)	56 (2)	49 (1)
Stroke	168 (2)	35 (3)	66 (2)	67 (2)
Heart disease	355 (4)	69 (5)	150 (5)	136 (4)
Liver disease	1543 (19)	216 (16)	623 (21)	704 (19)
Thyroid	249 (3)	54 (4)	107 (4)	88 (2)
Asthma	1746 (22)	286 (21)	670 (22)	790 (22)
Diabetes	604 (8)	102 (7)	251 (8)	251 (7)
Cancer	248 (3)	46 (3)	103 (3)	99 (3)
HIV/AIDS	557 (7)	115 (8)	233 (8)	209 (6)
STD	2121 (27)	329 (24)	793 (26)	999 (28)
Medicine for: nervousness	2311 (29)	317 (23)	902 (30)	1092 (30)
Medicine for: sleep	2647 (33)	271 (27)	1020 (34)	1256 (35)
Medicine for: depression	2988 (37)	420 (31)	1166 (39)	1402 (39)
Medicine for: headaches	732 (9)	112 (8)	285 (9)	335 (9)
Medicine for: energy	220 (3)	33 (2)	80 (3)	107 (3)
Medicine containing steroids	982 (12)	160 (12)	389 (13)	433 (12)

CoUD: Cocaine use disorder, PSA: Polysubstance addiction, LCA: Latent class analysis, AUD: Alcohol use disorder, CaUD: Cannabis use disorder, OUD: Opioid use disorder, TUD: Tobacco use disorder, ADHD: Attention deficit/hyperactivity disorder, ASPD: Antisocial personality disorder, MDE 1: One major depressive episode, MDE 2: Second major depressive episode, MDD = Major depressive disorder, MDE: Major depressive episode, OCD: Obsessive compulsive disorder, PTSD: Posttraumatic stress disorder, STD: Sexually transmitted disease. <sup>a</sup>Highest education completed: variable ranging from 1 to 8, where 1="grade 2-8 completed"; 2="grade 9 completed"; 3="grade 10 completed"; 4="grade 11 completed"; 5="grade 12 completed"; 6="technical school or 1-3 years of college"; 7="BA, BSc"; 8="MA, MS, JD, MD, PhD". <sup>b</sup>Household income: 1. \$0-\$9,999/year; 2. \$10,000-\$39,999/year; 3. \$40,000-\$150,000 or more/year. <sup>c</sup>Rating of physical health: 1. poor; 2. fair; 3. good; 4. very good; 5. excellent.

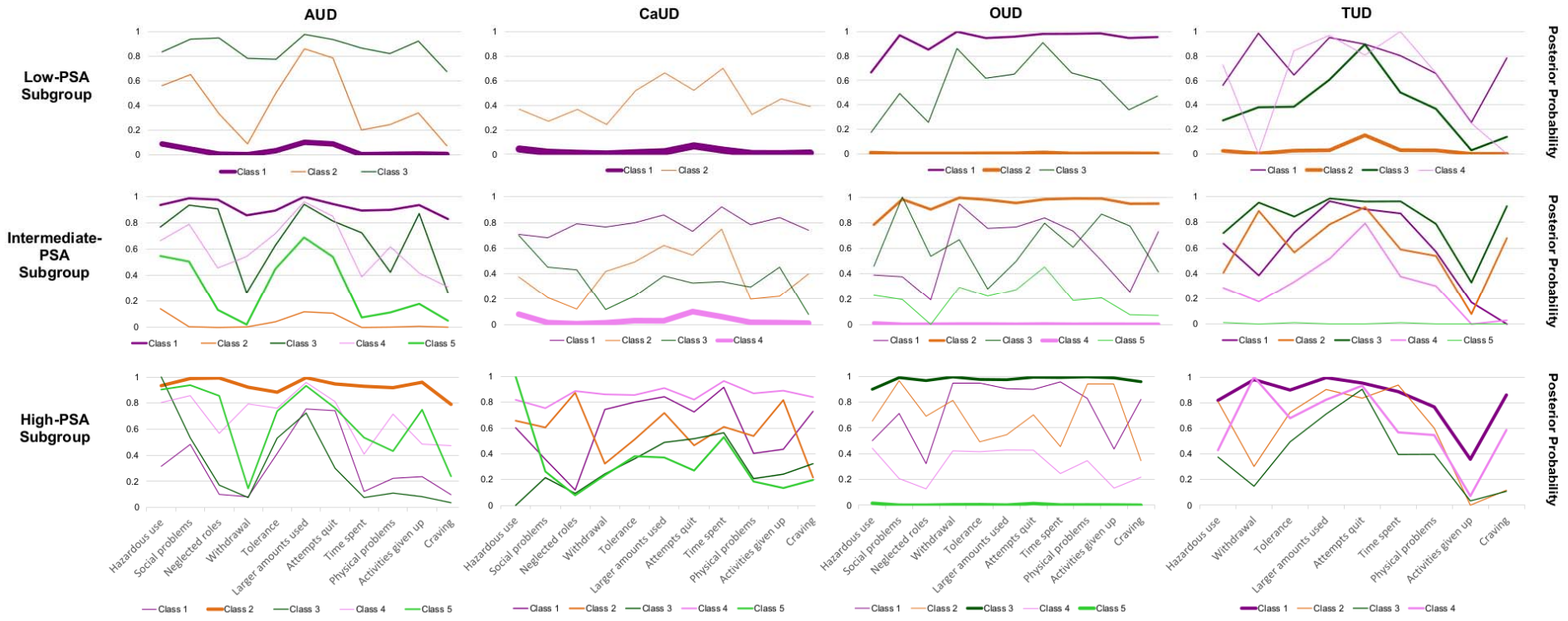
## Criteria-Based Latent Class Analysis

To further investigate PSA patterns within each of the CoUD subgroups (Table 1), we conducted additional LCAs using DSM-5 diagnostic criteria for AUD, CaUD, OUD, and TUD. The number of diagnostic-criteria latent classes in each of the SUDs considered was defined based on AIC, BIC, BLRT, and entropy statistics (Supplemental Tables 3 to 5). Figure 1 summarizes the number of diagnostic-criteria latent classes, their posterior probabilities, and predicted class membership across the three CoUD subgroups. Considering the number of diagnostic-criteria latent classes, the low-PSA subgroup showed the best fit to the data with two diagnostic-criteria latent classes for CaUD, three for AUD and OUD, and four for TUD. Conversely, four or five diagnostic-criteria latent classes across the SUDs considered were the best fits for the intermediate and high-PSA subgroups.

Considering the posterior probabilities and the predicted class memberships of the diagnostic-criteria latent classes, we observed unique PSA patterns across the CoUD subgroups. Specifically, most individuals included in the low-PSA subgroup were assigned to latent classes with <20% posterior probability of diagnostic criteria for AUD (predicted class membership=73%), CaUD (predicted class membership=95%), and TUD (predicted class membership=50%). Conversely, the high PSA subgroup showed 44% predicted class membership for a latent class with >78% posterior probability across AUD diagnostic criteria

and 43% predicted class membership for a latent class with >76% posterior probability across TUD diagnostic criteria with the exception of the TUD criterion “Important activities are reduced or given up because of use.” With respect to CaUD, the high-PSA subgroup showed diagnostic-criteria latent classes (predicted class membership ranging from 16% to 25%) with different patterns of posterior probabilities. Participants in the intermediate-PSA subgroup were evenly distributed across diagnostic-criteria latent classes with different patterns of posterior probabilities for AUD and TUD. Conversely, 75% of this subgroup was predicted to be assigned to a latent class with <10% posterior probability of CaUD diagnostic criteria. Interestingly, there was a similar finding with respect to OUD diagnostic criteria across the three CoUD subgroups. Specifically, two diagnostic-criteria latent classes were predicted to include >70% of the individuals (split almost evenly) with one having >65% posterior probability for OUD criteria and the other having <5% posterior probability for OUD criteria.

With respect to the individual CoUD diagnostic criteria, the most pronounced difference between low- and high-PSA subgroups was for “Repeated substance use in situations where it is physically hazardous” (40% vs. 70%) while the least pronounced was for “Persistent desire/unsuccessful efforts to stop using” (93% vs. 94%). The percent difference between these two subgroups ranged from 17% to 9% for the other CoUD diagnostic criteria (Supplemental Table 6). Considering the overall criterion count, 59% of the individuals in the high-PSA subgroup met ten or more CoUD criteria while this was observed in only 34% of low-PSA subgroup (Supplemental Table 7). The frequency distribution of the CoUD diagnostic criteria in the intermediate-PSA subgroup was similar to that observed in the overall sample (Supplemental Tables 6 and 7).



**Figure 1. Patterns of the posterior probabilities of the criteria-based latent classes across CoUD subgroups.** The posterior probabilities of the best-fit latent class models identified for AUD, CaUD, OUD, and TUD criteria with respect to the three CoUD subgroups. Line thickness corresponds to the proportion of the diagnosis-based subgroup that was stratified to a particular criteria-based class.

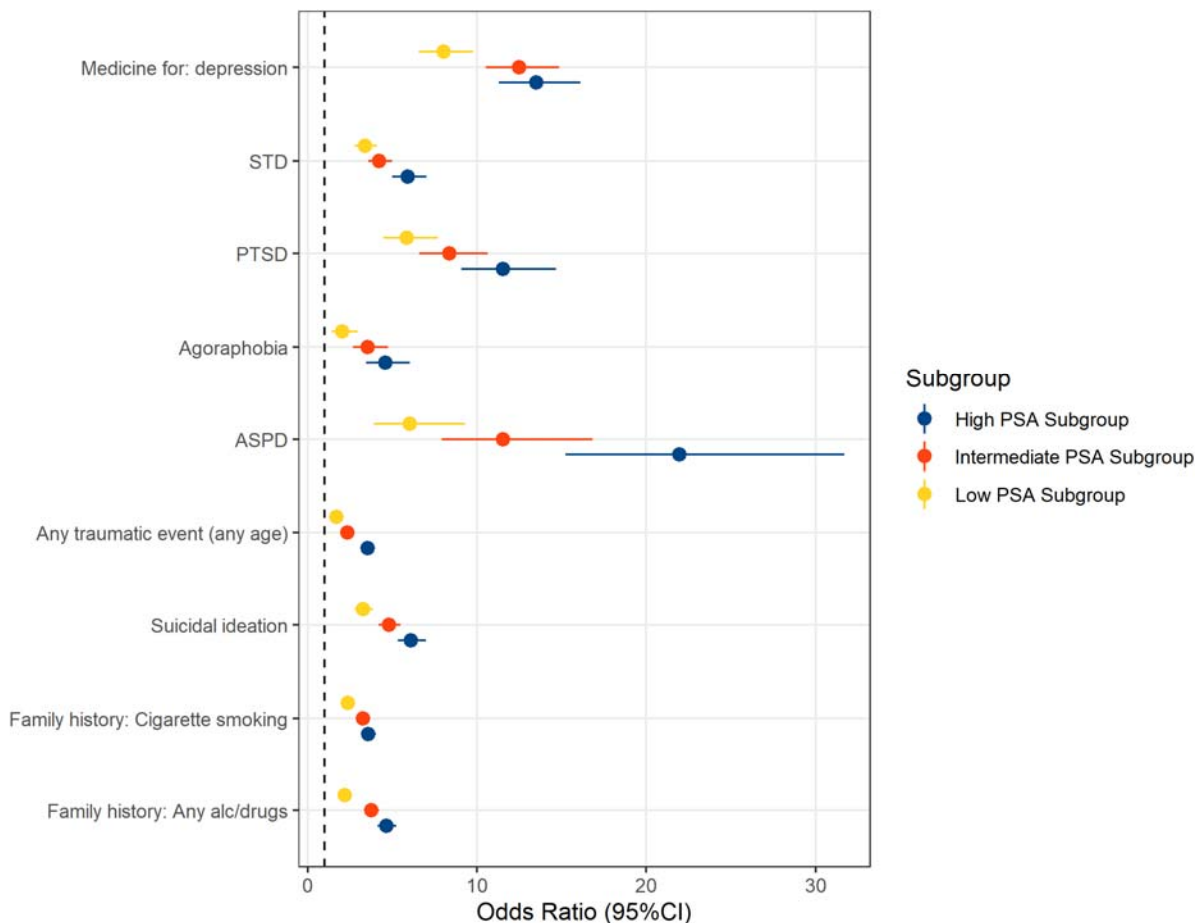
## Psychiatric and somatic associations of CoUD subgroups

The analyses above demonstrated that the diagnosis-based subgroups of Yale-Penn participants with CoUD are linked to different PSA patterns. To understand their associations with psychopathology, suicidality, traumatic experiences, family history of substance use, medical conditions, and SES, we tested the three CoUD subgroups with respect to 2,952 Yale-Penn participants who did not meet criteria for any of the 5 SUDs of interest. After accounting for age, sex, and race-ethnicity and applying Bonferroni-correction for multiple testing (corrected p-value threshold  $< 7.94 \times 10^{-4}$ ), we observed that many of the CoUD subgroups were associated with more severe adversity and comorbidities than the control group, with significant effects observed within each of the domains tested (Supplemental Table 8).

When comparing the associated effect sizes between the subgroups, no difference was observed with respect to the SES-related variables (difference- $p > 0.05$ ), but all other categories included at least one trait with Bonferroni-corrected differences in the effects detected comparing the three CoUD subgroups to controls (Figure 2). The low-PSA subgroup had lower odds to report family/household use of any substance (OR=2.18, 95%CI=1.87-2.53) compared with either the intermediate-PSA subgroup (OR=3.76, 95%CI=3.33-4.23; difference- $p = 2.64 \times 10^{-8}$ ) or the high-PSA subgroup (OR=4.66, 95%CI=4.13-5.25; difference- $p = 1.09 \times 10^{-14}$ ). A significant difference between the low- and high-PSA subgroups was also present for family/household cigarette smoking (OR=2.38 vs. 3.58, difference- $p = 8.2 \times 10^{-5}$ ). The three CoUD subgroups were associated with increased odds of family/household use of cocaine when growing up, but there was no difference in the strength of the associations (difference- $p > 0.05$ ). With respect to psychopathology, the high-PSA subgroup was associated with higher odds of antisocial personality disorder (ASPD, OR=21.96 vs. 6.03, difference- $p = 8.08 \times 10^{-6}$ ), agoraphobia (OR=4.58 vs. 2.05, difference- $p = 7.04 \times 10^{-4}$ ), and posttraumatic stress disorder (OR=11.54 vs.



5.86, difference- $p=2.67 \times 10^{-4}$ ) than the low-PSA subgroup. Significant-corrected differences in effect size were also observed among the three CoUD subgroups in relation to experiencing or witnessing a major traumatic event, where the effect size was largest for the high-PSA subgroup (OR=11.54, 95%CI=9.07-14.68), and progressively smaller for the intermediate-PSA subgroup (OR=8.38, 95%CI=6.60-10.65) and low-PSA subgroup (OR=5.86, 95%CI=4.46-7.71; low-PSA vs. high-PSA difference- $p=8.5 \times 10^{-16}$ ; low-PSA vs. intermediate-PSA difference- $p=3.28 \times 10^{-4}$ ; high-PSA vs. intermediate-PSA difference- $p=4.41 \times 10^{-7}$ ). Differences were also present in the strength of the associations of the three CoUD subgroups with suicide ideation (high-PSA OR=6.1 vs. low-PSA OR=3.27, difference- $p=5.43 \times 10^{-9}$ ; low-PSA OR=3.27 vs. intermediate-PSA OR=4.79, difference- $p=3.29 \times 10^{-4}$ ). Considering medical history, when compared to controls, the high-PSA subgroup was associated with increased odds of being diagnosed with sexually transmitted diseases (STD; OR=5.92 vs. 3.38, difference- $p=1.81 \times 10^{-5}$ ) and taking depression medication (OR=13.49 vs. 8.02, difference- $p=1.42 \times 10^{-4}$ ) than the low-PSA subgroup.



**Figure 2. Traits with statistically significant differences across CoUD subgroups in the association strength observed when compared to the control group.** Odds ratios and 95% confidence intervals (95%CI) are shown. STD: Sexually transmitted disease, PTSD: posttraumatic stress disorder. ASPD: antisocial personality disorder.

## DISCUSSION

To date, most molecular and brain imaging studies investigated CoUD-affected individuals as a singular entity, in comparison with control samples or groups of participants affected by other substance use or psychiatric disorders (50, 51). Due to the low prevalence of CoUD in the general population, few investigations of nationally representative cohorts examined differences among individuals with CoUD (52). Clinically ascertained CoUD cohorts provide a larger proportion of individuals with the disorder and offer opportunities for more in-depth clinical and phenotypic characterization. As mentioned, molecular and brain imaging

studies of CoUD have mostly investigated differences between this group and controls or individuals with other psychiatric conditions (50, 51). Understanding CoUD heterogeneity in the context of PSA and commonly comorbid psychiatric and somatic comorbidities may help inform the development of more precise treatments for this population. We investigated the phenotypic diversity of 7,989 individuals with a lifetime CoUD diagnosis recruited in the Yale-Penn cohort. Leveraging this large, deeply phenotyped sample, we found that CoUD is characterized by latent typologies of individuals with different profiles of psychiatric and somatic comorbidities.

Although the Yale-Penn cohort was ascertained based on the presence of addictive disorders and is not representative of the general population, our results highlight how CoUD heterogeneity could affect molecular, brain imaging, and treatment studies of addiction-ascertained samples.

Our initial LCA based on DSM-5 SUD diagnoses (i.e., AUD, CaUD, OUD, and TUD) distinguished three CoUD subgroups: low PSA (17%), intermediate PSA (38%), and high PSA (45%). In line with sex differences present among individuals with SUDs from epidemiologic studies (53), most of the individuals included in the high-PSA subgroup were males. With respect to race/ethnicity, the low-PSA subgroup has a higher percentage of non-Hispanic European Americans and high-PSA subgroup included more non-Hispanic African Americans than the overall CoUD sample. This trend parallels those reported by the National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III), which similarly found that non-European Americans were the majority among cocaine-only users and European Americans were the majority among cocaine-cannabis users (54). As shown in a comparison between the 2001-2002 NESARC and 2012-2013 NESARC-III, there are racial-ethnic differences among cocaine users and CoUD-affected individuals (55). Although ethnic and racial differences in CoUD are likely affected by disparities in income, education, unemployment, and discrimination (55, 56),

the relationship between cocaine addiction and SES also appears to be affected by temporal trends. Indeed, although in the 1980s cocaine use was more prevalent among highly educated individuals, during the 1990s it became more prevalent among less educated individuals, and in the 2000s it increased among higher education subgroups, but not in the lowest (55). We found no major differences among CoUD subgroups with respect to income or education. This indicates that in the Yale-Penn cohort, the differences among CoUD subgroups may be influenced by cultural factors related to sex and racial-ethnic differences rather than economic differences and education.

While the general patterns observed in the LCA based on SUD diagnoses were confirmed by the LCA based on SUD diagnostic criteria (i.e., the low-PSA subgroup corresponds to CoUD with low probability of other SUD diagnostic criteria; the intermediate PSA subgroup corresponds to intermediate probability of other SUD diagnostic criteria; the high-PSA subgroup corresponds to CoUD with high probability of other SUD diagnostic criteria), we also observed some substance-specific patterns in the criterion-based LCA. For example, most participants included in the low- and intermediate-PSA subgroups were characterized by a very low posterior probability (<10%) of CaUD diagnostic criteria. Another notable substance-specific pattern was seen with OUD diagnostic criteria, where in all CoUD subgroups participants were mostly split between a class with high posterior probabilities and a class with low posterior probabilities. Most previous studies investigating PSA in CoUD have focused on the frequency or severity of substance use (57, 58). These analyses may not have as great resolution as the analyses we performed using SUD diagnostic criteria. In the NESARC cohort, cocaine-opioid users represent about one-third of cocaine users (59). This may, at least in part, explain why a consistent number of Yale-Penn participants with low posterior probabilities of OUD diagnostic criteria were

included in the CoUD subgroup with high PSA. With respect to CoUD severity, the majority of CoUD individuals with high PSA met ten or more CoUD criteria. This is in line with evidence from NESARC-III analyses where “very high quantity use” and “daily use” of cocaine were associated with polysubstance use (60). Additionally, NESARC-III analyses focused on quantity, frequency, and duration of cocaine use and identified three cocaine use patterns (61), which appear to overlap with the CoUD subgroups we observed in the Yale-Penn cohort.

In line with the well-known association of CoUD with mental and physical health (62), the three CoUD subgroups identified in our study were all associated with increased suicidal behaviors, traumatic experience, psychopathology, and somatic comorbidities relative to controls (i.e., individuals without any SUDs). However, we additionally observed several traits where the strength of the associations was statistically different among the three CoUD subgroups. With respect to family history of substance use, individuals in the low-PSA subgroup demonstrated smaller magnitude associations with family histories of alcohol and any-drug use (compared with the intermediate- and high-PSA subgroups), cigarette smoking (compared with the high-PSA subgroup), and other illegal drugs (compared with the high-PSA subgroup). Of note, the strength of the association with family use of cocaine and heroin did not differ across CoUD subgroups. Previous studies showed that a family history of substance use is associated with polysubstance use (63-65). Our findings add to this literature and suggest that family use of multiple types of substances may be a risk factor for PSA in CoUD. Another difference among CoUD subgroups was related to suicidal behaviors. While all CoUD subgroups were associated with increased suicidal behaviors (i.e., ideation, planning, and attempt) relative to controls, suicide ideation showed smaller magnitude associations with the low-PSA subgroup, compared with the other subgroups. The association of suicidal behaviors with cocaine addiction and polysubstance use is

well established (66-68). However, to our knowledge, the current study is the first to investigate suicidality in the context of PSA among individuals with CoUD. The finding that suicide ideation may be differentially impacted by PSA can be particularly noteworthy with respect to suicide prevention among high-risk individuals affected by CoUD.

Associations between CoUD, traumatic experiences, and PTSD were previously identified in multiple studies (69, 70). We found evidence for this association in the Yale-Penn cohort and observed differences in the strength of the associations across CoUD subgroups. Specifically, individuals belonging to the low-PSA subgroup were least strongly associated with having experienced a traumatic event and had the lowest odds of having a PTSD diagnosis. However, there were no differences across CoUD subgroups in their association with different types of traumatic events experiences (e.g., violence, physical abuse, and sexual abuse). With respect to other psychiatric disorders, ASPD showed the strongest associations with the high-PSA subgroup where individuals in this group had a 22-fold increase in the odds of having an ASPD diagnosis in the Yale-Penn cohort. Although low- and intermediate-PSA subgroups were also associated with ASPD, there is a strong difference in ASPD association strength between high- and low-PSA subgroups (ASPD OR= 21.96 vs. 6.03, difference- $p=8.08\times 10^{-6}$ ). This is in line with previous studies in treatment-seeking substance users showing that CoUD and polysubstance use are both associated with ASPD with the latter having a larger effect (71, 72). Agoraphobia was the other psychiatric comorbidity showing differences across CoUD subgroups. Cocaine use and cocaine dependence have been previously associated with agoraphobia in treatment-seeking individuals and in community surveys (73, 74). We confirmed this relationship with CoUD in the Yale-Penn cohort, but also observed that CoUD with high PSA is more strongly associated with agoraphobia than CoUD with low PSA. With respect to

somatic comorbidities, we observed differences among CoUD subgroups with respect to two health outcomes, STDs and the use of depression medication. There is a vast literature supporting the association of CoUD and PSA with STD risk (75-78). In Yale-Penn, the stronger STD association with the high PSA subgroup than with the low PSA subgroup is in line with these previous findings. While all CoUD subgroups were associated with traits related to depression (e.g., major depressive disorder and major depressive episodes), there was no statistically significant difference in the strength of the associations. Conversely, depression medication was more strongly associated with the high-PSA subgroup than the low-PSA subgroup. This suggests that more severe CoUD cases such as those affected by high PSA may be more frequently treated with antidepressants, despite available evidence not supporting their efficacy in treating symptoms of CoUD (79). Alternatively, the stronger association between the high-PSA subgroup and antidepressant medication treatment could reflect greater severity or chronicity of major depression or other disorders treated with antidepressants than those in the low or intermediate PSA subgroups.

Although our study contributes to the characterization of CoUD heterogeneity in a large sample, we acknowledge four main limitations. First, Yale-Penn participants were recruited for addiction genetic studies for more than 20 years. Accordingly, the associations observed may not reflect those present in the general population. As discussed above, they may be also affected by the temporal changes in cocaine use that occurred during this time span. Second, the SSADDA was designed to obtain DSM-IV diagnoses. To investigate addiction in the context of DSM-5, we derived DSM-5 SUD diagnoses from additional information collected with the SSADDA. While all DSM-5 diagnostic criteria were adequately mapped for most SUDs, two TUD criteria (i.e., social problems and neglected roles) were not available in the SSADDA. This may have affected

our ability to investigate TUD comorbidity in the context of CoUD. Third, although the SSADDA is a previously validated diagnostic instrument (38, 39), the information collected may have been impacted by self-report bias (80). Fourth, the majority of the sample investigated had yearly household incomes below \$40,000. This may have limited our ability to investigate the impact of socioeconomic factors on PSA of individuals affected by CoUD.

Notwithstanding these limitations, we identified three CoUD subgroups in the Yale-Penn cohort, reflecting different PSA degrees that were associated with different patterns of psychiatric and somatic comorbidities. Although the results may not generalize to general population samples, our findings nevertheless highlight the need for modeling PSA when analyzing CoUD in cohorts ascertained for addiction research such as those investigated in molecular and neuroimaging studies. Further research is needed to assess how these patterns of PSA coincide or differ in a general population. Once validated, patterns of PSA in CoUD could be assessed molecularly and with neuroimaging approaches to develop our understanding of these conditions.



## REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Washington, DC2013.
2. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575-86.
3. Center for Behavioral Health Statistics and Quality. Results from the 2019 National Survey on Drug Use and Health: Detailed tables. Rockville, MD2020.
4. Roy E, Richer I, Arruda N, Vandermeerschen J, Bruneau J. Patterns of cocaine and opioid co-use and polyroutes of administration among street-based cocaine users in Montreal, Canada. *Int J Drug Policy*. 2013;24(2):142-9.
5. McCabe SE, West BT, Jutkiewicz EM, Boyd CJ. Multiple DSM-5 substance use disorders: A national study of US adults. *Human Psychopharmacology: Clinical and Experimental*. 2017;32(5):e2625.
6. Timko C, Ilgen M, Haverfield M, Shelley A, Breland JY. Polysubstance use by psychiatry inpatients with co-occurring mental health and substance use disorders. *Drug Alcohol Depend*. 2017;180:319-22.
7. Wang L, Min JE, Krebs E, Evans E, Huang D, Liu L, et al. Polydrug use and its association with drug treatment outcomes among primary heroin, methamphetamine, and cocaine users. *Int J Drug Policy*. 2017;49:32-40.
8. Gjersing L, Bretteville-Jensen AL. Patterns of substance use and mortality risk in a cohort of 'hard-to-reach' polysubstance users. *Addiction*. 2018;113(4):729-39.
9. Macdonald S, Pakula B, Martin G, Wells S, Borges G, Roth E, et al. Health profiles of clients in substance abuse treatment: a comparison of clients dependent on alcohol or cocaine with those concurrently dependent. *Subst Use Misuse*. 2014;49(14):1899-907.
10. Kariisa M SL, Wilson N, Seth P, Hoots B. Drug Overdose Deaths Involving Cocaine and Psychostimulants with Abuse Potential — United States, 2003–2017. *MMWR Morb Mortal Wkly Rep* 2019. 2019;68:388–95.
11. Vermunt JK, Magidson J. Latent class analysis. *The sage encyclopedia of social sciences research methods*. 2004;2:549-53.
12. Lubke GH, Muthen B. Investigating population heterogeneity with factor mixture models. *Psychol Methods*. 2005;10(1):21-39.
13. Garcia-Marchena N, Ladron de Guevara-Miranda D, Pedraz M, Araos PF, Rubio G, Ruiz JJ, et al. Higher Impulsivity As a Distinctive Trait of Severe Cocaine Addiction among Individuals Treated for Cocaine or Alcohol Use Disorders. *Front Psychiatry*. 2018;9:26.
14. Liu Y, Elliott AL, Serdarevic M, Leeman RF, Cottler LB. A latent class analysis of the past-30-day substance use patterns among lifetime cocaine users: Findings from a community sample in North Central Florida. *Addict Behav Rep*. 2019;9:100170.
15. Agrawal A, Lynskey MT, Madden PA, Bucholz KK, Heath AC. A latent class analysis of illicit drug abuse/dependence: results from the National Epidemiological Survey on Alcohol and Related Conditions. *Addiction*. 2007;102(1):94-104.
16. Chan G, Connor J, Hall W, Leung J. The changing patterns and correlates of population-level polysubstance use in Australian youth: a multi-group latent class analysis of nationally representative samples spanning 12 years. *Addiction*. 2020;115(1):145-55.

17. Connor JP, Gullo MJ, Chan G, Young RM, Hall WD, Feeney GF. Polysubstance use in cannabis users referred for treatment: drug use profiles, psychiatric comorbidity and cannabis-related beliefs. *Front Psychiatry*. 2013;4:79.
18. Mefodeva V, Carlyle M, Walter Z, Chan G, Hides L. Polysubstance use in young people accessing residential and day-treatment services for substance use: substance use profiles, psychiatric comorbidity and treatment completion. *Addiction*. 2022;n/a(n/a).
19. Rodriguez AS, Robinson LD, Kelly PJ, Hudson S. Polysubstance use classes and health outcomes among women attending specialist substance use treatment services. *Drug Alcohol Rev*. 2022;41(2):488-500.
20. Pietrzak RH, el-Gabalawy R, Tsai J, Sareen J, Neumeister A, Southwick SM. Typologies of posttraumatic stress disorder in the U.S. adult population. *J Affect Disord*. 2014;162:102-6.
21. Pietrzak RH, Feder A, Schechter CB, Singh R, Cancelmo L, Bromet EJ, et al. Dimensional structure and course of post-traumatic stress symptomatology in World Trade Center responders. *Psychol Med*. 2014;44(10):2085-98.
22. Arathimos R, Fabbri C, Vassos E, Davis KAS, Pain O, Gillett A, et al. Latent subtypes of manic and/or irritable episode symptoms in two population-based cohorts. *Br J Psychiatry*. 2022:1-10.
23. Bailey AJ, Ingram PF, Howe LK, Finn PR. Is lower severity alcohol use disorder qualitatively different than more severe manifestations? An evaluation of multivariate symptom clusters. *Addiction*. 2022;117(6):1598-608.
24. Watts AL, Boness CL, Loeffelman JE, Steinley D, Sher KJ. Does crude measurement contribute to observed unidimensionality of psychological constructs? A demonstration with DSM-5 alcohol use disorder. *J Abnorm Psychol*. 2021;130(5):512-24.
25. Gizer IR, Gilder DA, Lau P, Wang T, Wilhelmsen KC, Ehlers CL. Contributions of ethnicity to differential item functioning of cannabis abuse and dependence symptoms. *J Stud Alcohol Drugs*. 2013;74(2):320-8.
26. Grant JD, Scherrer JF, Neuman RJ, Todorov AA, Price RK, Bucholz KK. A comparison of the latent class structure of cannabis problems among adult men and women who have used cannabis repeatedly. *Addiction*. 2006;101(8):1133-42.
27. Howe LK, Bailey AJ, Ingram PF, Finn PR. An exploration of multivariate symptom clusters of cannabis use disorder in young adults. *Addict Behav*. 2022;135:107465.
28. Clark SL, Gillespie NA, Adkins DE, Kendler KS, Neale MC. Psychometric modeling of abuse and dependence symptoms across six illicit substances indicates novel dimensions of misuse. *Addict Behav*. 2016;53:132-40.
29. Voce A, Burns R, Castle D, Calabria B, McKetin R. Is there a discrete negative symptom syndrome in people who use methamphetamine? *Compr Psychiatry*. 2019;93:27-32.
30. Bi J, Gelernter J, Sun J, Kranzler HR. Comparing the utility of homogeneous subtypes of cocaine use and related behaviors with DSM-IV cocaine dependence as traits for genetic association analysis. *Am J Med Genet B Neuropsychiatr Genet*. 2014;165b(2):148-56.
31. Kranzler HR, Wilcox M, Weiss RD, Brady K, Hesselbrock V, Rounsaville B, et al. The validity of cocaine dependence subtypes. *Addict Behav*. 2008;33(1):41-53.
32. Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, et al. Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry*. 2014;19(1):41-9.

33. Gelernter J, Kranzler HR, Sherva R, Koesterer R, Almasy L, Zhao H, et al. Genome-wide association study of opioid dependence: multiple associations mapped to calcium and potassium pathways. *Biol Psychiatry*. 2014;76(1):66-74.
34. Gelernter J, Sherva R, Koesterer R, Almasy L, Zhao H, Kranzler HR, et al. Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol Psychiatry*. 2014;19(6):717-23.
35. Gelernter J, Kranzler HR, Sherva R, Almasy L, Herman AI, Koesterer R, et al. Genome-wide association study of nicotine dependence in American populations: identification of novel risk loci in both African-Americans and European-Americans. *Biol Psychiatry*. 2015;77(5):493-503.
36. Sherva R, Wang Q, Kranzler H, Zhao H, Koesterer R, Herman A, et al. Genome-wide Association Study of Cannabis Dependence Severity, Novel Risk Variants, and Shared Genetic Risks. *JAMA Psychiatry*. 2016;73(5):472-80.
37. Kember RL, Hartwell EE, Xu H, Rotenberg J, Almasy L, Zhou H, et al. Phenome-wide Association Analysis of Substance Use Disorders in a Deeply Phenotyped Sample. *Biol Psychiatry*. 2022.
38. Pierucci-Lagha A, Gelernter J, Feinn R, Cubells JF, Pearson D, Pollastri A, et al. Diagnostic reliability of the Semi-structured Assessment for Drug Dependence and Alcoholism (SSADDA). *Drug Alcohol Depend*. 2005;80(3):303-12.
39. Pierucci-Lagha A, Gelernter J, Chan G, Arias A, Cubells JF, Farrer L, et al. Reliability of DSM-IV diagnostic criteria using the semi-structured assessment for drug dependence and alcoholism (SSADDA). *Drug Alcohol Depend*. 2007;91(1):85-90.
40. Olaya B, Moneta MV, Caballero FF, Tyrovolas S, Bayes I, Ayuso-Mateos JL, et al. Latent class analysis of multimorbidity patterns and associated outcomes in Spanish older adults: a prospective cohort study. *BMC Geriatr*. 2017;17(1):186.
41. Simon P, Jiang Y, Buta E, Sartor CE, Krishnan-Sarin S, Gueorguieva R. Longitudinal Trajectories of Multiple Nicotine Product Use Among Youths in the Population Assessment of Tobacco and Health Study. *JAMA Netw Open*. 2022;5(3):e223549.
42. Nylund KL, Asparouhov T, Muthén BO. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*. 2007;14(4):535-69.
43. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*. 1996;13:195-212.
44. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022.
45. RStudio Team. RStudio: Integrated Development Environment for R. Boston, MA: RStudio, PBC; 2022.
46. Linzer DA, Lewis J. poLCA: Polytomous Variable Latent Class Analysis 2022 [R package version 1.6]. Available from: <https://dlinzer.github.com/poLCA>.
47. Linzer DA, Lewis J. poLCA: An R package for polytomous variable latent class analysis. *Journal of statistical software*. 2011;42(10):1-29.
48. Kim Y, Chung H. glca: An R Package for Multiple-Group Latent Class Analysis. 2021.
49. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*. 1974;61(2):215-31.
50. Gelernter J, Polimanti R. Genetics of substance use disorders in the era of big data. *Nat Rev Genet*. 2021;22(11):712-29.

51. Poireau M, Milpied T, Maillard A, Delmaire C, Volle E, Bellivier F, et al. Biomarkers of Relapse in Cocaine Use Disorder: A Narrative Review. *Brain Sci.* 2022;12(8).
52. Liu Y, Williamson V, Setlow B, Cottler LB, Knackstedt LA. The importance of considering polysubstance use: lessons from cocaine research. *Drug Alcohol Depend.* 2018;192:16-28.
53. Goodwin SR, Moskal D, Marks RM, Clark AE, Squeglia LM, Roche DJO. A Scoping Review of Gender, Sex and Sexuality Differences in Polysubstance Use in Adolescents and Adults. *Alcohol Alcohol.* 2022;57(3):292-321.
54. Liu Y, Cheong J, Setlow B, Cottler LB. Cocaine and Marijuana Polysubstance Use and Cocaine Use Disorder: Investigating Mediated Effects through Patterns of Cocaine Use. *J Dual Diagn.* 2021;17(1):23-33.
55. Kerridge BT, Chou SP, Pickering RP, Ruan WJ, Huang B, Jung J, et al. Changes in the prevalence and correlates of cocaine use and cocaine use disorder in the United States, 2001-2002 and 2012-2013. *Addict Behav.* 2019;90:250-7.
56. Palamar JJ, Davies S, Ompad DC, Cleland CM, Weitzman M. Powder cocaine and crack use in the United States: an examination of risk for arrest and socioeconomic disparities in use. *Drug Alcohol Depend.* 2015;149:108-16.
57. Aharonovich E, Scodes J, Wall MM, Hasin DS. The relationship of frequency of cocaine use to substance and psychiatric disorders in the U.S. general population. *Drug Alcohol Depend.* 2021;227:108933.
58. Stevens AK, Gunn RL, Sokolovsky AW, Colby SM, Jackson KM. Examining the heterogeneity of polysubstance use patterns in young adulthood by age and college attendance. *Exp Clin Psychopharmacol.* 2022;30(5):701-13.
59. Leeman RF, Sun Q, Bogart D, Beseler CL, Sofuoglu M. Comparisons of Cocaine-Only, Opioid-Only, and Users of Both Substances in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Subst Use Misuse.* 2016;51(5):553-64.
60. Liu Y, Vaddiparti K, Cheong J, Cottler LB. Identification of Typologies of Cocaine Use Based on Quantity, Frequency, and Duration of Use: A Latent Profile Analysis. *J Addict Med.* 2021;15(3):211-8.
61. Liu Y, Cheong J, Vaddiparti K, Cottler LB. The association between quantity, frequency and duration of cocaine use during the heaviest use period and DSM-5 cocaine use disorder. *Drug Alcohol Depend.* 2020;213:108114.
62. Kampman KM. The treatment of cocaine use disorder. *Sci Adv.* 2019;5(10):eaax1532.
63. Cho J, Stone MD, Leventhal AM. Anhedonia as a phenotypic marker of familial transmission of polysubstance use trajectories across midadolescence. *Psychology of Addictive Behaviors.* 2019;33:15-25.
64. Tomczyk S, Isensee B, Hanewinkel R. Latent classes of polysubstance use among adolescents-a systematic review. *Drug Alcohol Depend.* 2016;160:12-29.
65. Schrage SM, Kecojevic A, Silva K, Jackson Bloom J, Iverson E, Lankenau SE. Correlates and Consequences of Opioid Misuse among High-Risk Young Adults. *J Addict.* 2014;2014:156954.
66. Britton PC, Conner KR. Suicide attempts within 12 months of treatment for substance use disorders. *Suicide Life Threat Behav.* 2010;40(1):14-21.
67. Garlow SJ, Purselle D, D'Orio B. Cocaine use disorders and suicidal ideation. *Drug Alcohol Depend.* 2003;70(1):101-4.

68. Na PJ, Bommersbach TJ, Petrakis IL, Rhee TG. National trends of suicidal ideation and mental health services use among US adults with opioid use disorder, 2009-2020. *EClinicalMedicine*. 2022;54:101696.
69. Khoury L, Tang YL, Bradley B, Cubells JF, Ressler KJ. Substance use, childhood traumatic experience, and Posttraumatic Stress Disorder in an urban civilian population. *Depress Anxiety*. 2010;27(12):1077-86.
70. Najavits LM, Gastfriend DR, Barber JP, Reif S, Muenz LR, Blaine J, et al. Cocaine dependence with and without PTSD among subjects in the National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Am J Psychiatry*. 1998;155(2):214-9.
71. Kidorf M, Solazzo S, Yan H, Brooner RK. Psychiatric and Substance Use Comorbidity in Treatment-Seeking Injection Opioid Users Referred From Syringe Exchange. *J Dual Diagn*. 2018;14(4):193-200.
72. Kranzler HR, Satel S, Apter A. Personality disorders and associated features in cocaine-dependent inpatients. *Compr Psychiatry*. 1994;35(5):335-40.
73. Arias F, Szerman N, Vega P, Mesias B, Basurte I, Morant C, et al. Cocaine abuse or dependency and other psychiatric disorders. Madrid study on dual pathology. *Rev Psiquiatr Salud Ment*. 2013;6(3):121-8.
74. Sareen J, Chartier M, Paulus MP, Stein MB. Illicit drug use and anxiety disorders: findings from two community surveys. *Psychiatry Res*. 2006;142(1):11-7.
75. Butler AJ, Rehm J, Fischer B. Health outcomes associated with crack-cocaine use: Systematic review and meta-analyses. *Drug Alcohol Depend*. 2017;180:401-16.
76. Arterberry BJ, Davis AK, Walton MA, Bonar EE, Cunningham RM, Blow FC. Predictors of empirically derived substance use patterns among sexual minority groups presenting at an emergency department. *Addict Behav*. 2019;96:76-81.
77. Mayer KH, Bush T, Henry K, Overton ET, Hammer J, Richardson J, et al. Ongoing sexually transmitted disease acquisition and risk-taking behavior among US HIV-infected patients in primary care: implications for prevention interventions. *Sex Transm Dis*. 2012;39(1):1-7.
78. Mimiaga MJ, Reisner SL, Vanderwarker R, Gaucher MJ, O'Connor CA, Medeiros MS, et al. Polysubstance use and HIV/STD risk behavior among Massachusetts men who have sex with men accessing Department of Public Health mobile van services: implications for intervention development. *AIDS Patient Care STDS*. 2008;22(9):745-51.
79. Pani PP, Trogu E, Vecchi S, Amato L. Antidepressants for cocaine dependence and problematic cocaine use. *Cochrane Database Syst Rev*. 2011(12):CD002950.
80. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc*. 2016;9:211-7.