



Contents lists available at ScienceDirect

Preventive Medicine

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

Characterizing the intersection of Co-occurring risk factors for illicit drug abuse and dependence in a U.S. nationally representative sample

Allison N. Kurti, PhD^{a,b}, Diana R. Keith, PhD^{a,b}, Alyssa Noble, B.A.^{a,b}, Jeff S. Priest, PhD^d, Brian Sprague, PhD^{a,e}, Stephen T. Higgins, PhD^{a,b,c,*}

^a Vermont Center on Behavior and Health, University of Vermont, United States

^b Department of Psychiatry, University of Vermont, United States

^c Department of Psychological Science, University of Vermont, United States

^d Medical Biostatistics Unit, University of Vermont, United States

^e Department of Surgery, University of Vermont, United States

ARTICLE INFO

Article history:

Received 19 May 2016

Received in revised form 4 August 2016

Accepted 23 September 2016

Available online xxxx

Keywords:

Risk factors

Co-occurring risk factors

Adults

U.S. nationally representative sample

Illicit drug abuse and dependence

Multiple logistic regression

Classification and regression tree (CART)

Substance use disorders

ABSTRACT

Few studies have attempted to characterize how co-occurring risk factors for substance use disorders intersect. A recent study examined this question regarding cigarette smoking and demonstrated that co-occurring risk factors generally act independently. The present study examines whether that same pattern of independent intersection of risk factors extends to illicit drug abuse/dependence using a U.S. nationally representative sample (National Survey on Drug Use and Health, 2011–2013). Logistic regression and classification and regression tree (CART) modeling were used to examine risk of past-year drug abuse/dependence associated with a well-established set of risk factors for substance use (age, gender, race/ethnicity, education, poverty, smoking status, alcohol abuse/dependence, mental illness). Each of these risk factors was associated with significant increases in the odds of drug abuse/dependence in univariate logistic regressions. Each remained significant in a multivariate model examining all eight risk factors simultaneously. CART modeling of these 8 risk factors identified subpopulation risk profiles wherein drug abuse/dependence prevalence varied from <1% to >80% corresponding to differing combinations of risk factors present. Alcohol abuse/dependence and cigarette smoking had the strongest associations with drug abuse/dependence risk. These results demonstrate that co-occurring risk factors for illicit drug abuse/dependence generally intersect in the same independent manner as risk factors for cigarette smoking, underscoring further fundamental commonalities across these different types of substance use disorders. These results also underscore the fundamental importance of differences in the presence of co-occurring risk factors when considering the often strikingly different prevalence rates of illicit drug abuse/dependence in U.S. population subgroups.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Substance use disorders (SUDs) contribute to a high and preventable proportion of overall disability and premature mortality in the U.S. and other developed countries. Although the burden of SUDs and their consequences are borne out differentially across different population subgroups, the varying biological, social, and environmental risk factors for SUDs remain to be determined. While existing research has examined relations between individual risk factors and drug abuse or dependence (e.g., male sex, Compton et al., 2007; Native American/Alaska Native race, Huang et al., 2006; younger age, Han et al., 2009), these risk factors along with others inevitably co-occur. However, to our knowledge, few studies have explicitly examined the manner in

which common risk factors for substance use disorders intersect when present in these inevitable combinations.

One useful way to examine associations between co-occurring risk factors and risk of substance use (e.g., current cigarette smoking, past year alcohol or illicit drug abuse/dependence) is by using epidemiological surveys (e.g., the National Epidemiological Survey on Alcohol and Related Conditions [NESARC], the National Survey on Drug Use and Health [NSDUH]). These U.S. general national population surveys present a valid depiction of SUDs based on criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), and query respondents about numerous other factors known to be associated with SUDs (e.g., demographic and socioeconomic characteristics) in a nationally representative sample. For example, Higgins et al. (this issue) used the U.S. NSDUH (2011–2013), to characterize intersections between eight co-occurring risk factors for current cigarette smoking (age, sex, education, race/ethnicity, poverty status, past year mental illness, alcohol abuse/dependence, illicit drug abuse/dependence) in adults (≥18 years). The

* Corresponding author at: Vermont Center on Behavior and Health, University of Vermont, United States.

E-mail address: shiggins@uvm.edu (S.T. Higgins).

results indicated that when these common risk factors for cigarette smoking co-occur, they generally act independently (i.e., their effects are not conditional on the presence of the others). Higgins et al. also compared the relative strength of the risk factors while identifying sub-population risk profiles wherein smoking prevalence ranged from 11% to 74% depending on the risk-factor combinations present.

The purpose of the present study was to examine (a) whether the pattern of intersection among co-occurring risk factors for cigarette smoking has generality to illicit drug abuse/dependence, and (b) how combinations of these common risk-factors are associated with particularly low- and high-risk profiles for illicit drug abuse/dependence. In order to facilitate comparison between the current and prior study on cigarette smoking, we used the same three years of the NSDUH (2011–2013), with past-year illicit drug abuse replacing current smoking (past month) as the dependent variable in the present report, and current cigarette smoking taking the place of illicit drug abuse/dependence as an independent variable (Substance Abuse and Mental Health Services Administration, 2012, 2013, 2014). We expanded the timeframe to past year rather than past month that was used in the study on cigarette smoking because of the lower prevalence of illicit drug abuse/dependence relative to current cigarette smoking and the need for a sufficiently large sample to examine the relations of interest. Important to underscore is that the primary purpose of this study is to characterize the nature of the intersection among co-occurring risk factors using this set of eight common risk factors for substance use and not to identify a new or comprehensive profile of variables associated with illicit drug abuse/dependence.

2. Methods

2.1. Data source

The NSDUH is an annual nationally representative cross-sectional survey of the U.S. non-institutionalized population assessing prevalence and correlates of tobacco, alcohol and illicit drug use among individuals aged ≥ 12 years (Center for Behavioral Health Statistics and Quality, 2014). Only individuals aged ≥ 18 years were included in the present study to keep age of respondents consistent with the prior study on cigarette smoking.

Data from all civilian non-institutionalized respondents, including those in group homes, shelters, and college dormitories, were included in the NSDUH survey. Respondents on active military duty, in drug treatment programs, jail or homeless were excluded. The weighted interview response rates were 74.4%, 73.0% and 71.7% in 2011, 2012 and 2013, respectively. Data were weighted during analysis to adjust for the differential probability of selection and response. A detailed description of the survey procedures is provided by SAMHSA (2013).

Past year illicit drug abuse/dependence was defined based on a statistical model developed from clinical interviews that assessed disorders based on criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). The illicit drugs captured by this variable included marijuana, hallucinogens, heroin, inhalants, tranquilizers, cocaine, pain relievers, stimulants, and sedatives. With respect to the risk factors that were used as independent variables, past year alcohol abuse/dependence was based on DSM-IV criteria like drug abuse/dependence. Current cigarette smoking was defined as smoking at least part of a cigarette in the past month and ≥ 100 cigarettes lifetime. Race was defined based on the six mutually exclusive racial/ethnic categories used by the NSDUH. More specifically, respondents who identified as Hispanic may be of any race whereas those identifying as White, Black, Asian, American Indian/Alaska Native, or Other (i.e., Native Hawaiians or other Pacific Islanders and respondents of ≥ 2 races) were all non-Hispanic. Poverty status (at or below versus above the federal poverty line) was defined using thresholds determined by the U.S. Census Bureau. Past year mental illness was based on participant responses to two scales measuring psychological distress (Kessler-6) and disability

(World Health Organization Disability Assessment Schedule) where scores on both scales were used to determine the presence of any mental illness based on a statistical model derived from DSM-IV criteria.

2.2. Statistical methods

Sample-adjusted frequencies and confidence intervals (CIs) were generated across all respondents aged ≥ 18 years. We examined associations between the eight risk factors of interest (age, gender, race/ethnicity, education, poverty, smoking status, alcohol abuse/dependence, mental illness) and illicit drug abuse/dependence.

Associations of risk factors with illicit drug abuse/dependence were first examined in separate analyses. For each risk factor, weighted, univariate logistic regression analyses were conducted to identify those

Table 1

Prevalence of illicit drug abuse/dependence and results from univariate logistic regression analyses examining associations between drug abuse/dependence^a among adults (aged ≥ 18 years) and eight potential risk factors ($n = 114,426$) – National Survey on Drug Use and Health (NSDUH), United States, 2011–2013.

	Illicit Drug Abuse/Dependence		Univariate Logistic Regression	
	%	(95% CI)	OR	(95% CI)
Overall	2.5	(2.4, 2.7)		
Gender				
Male	3.4	(3.2, 3.7)	2.1***	(1.8, 2.3)
Female	1.7	(1.5, 1.8)	Ref. group	
Age group (years)				
18–25	7.7	(7.3, 8.0)	35.3***	(22.0, 56.7)
26–44	3.0	(2.7, 3.3)	13.2***	(8.3, 20.8)
45–64	1.1	(0.9, 1.3)	4.7***	(3.0, 7.3)
≥ 65	0.2	(0.1, 0.3)	Ref. group	
Race/ethnicity ^b				
White	2.4	(2.3, 2.6)	2.7***	(1.8, 3.8)
Black	3.4	(2.9, 3.9)	3.8***	(2.5, 5.6)
Hispanic	2.6	(2.3, 2.9)	2.8***	(2.0, 4.1)
American Indian/Alaska Native	5.5	(3.7, 7.4)	6.2***	(3.8, 10.1)
Asian	0.9	(0.6, 1.3)	Ref. group	
Other	3.6	(2.8, 4.4)	3.9***	(2.5, 6.1)
Education level				
<High school	3.9	(3.6, 4.3)	3.9***	(3.3, 4.5)
High school graduate	3.0	(2.7, 3.3)	2.9***	(2.5, 3.4)
Some college	2.9	(2.6, 3.2)	2.8***	(2.4, 3.3)
College graduate	1.1	(0.9, 1.2)	Ref. group	
Poverty status ^c				
Below poverty level	4.7	(4.2, 5.1)	2.3***	(2.0, 2.5)
At or above poverty level	2.1	(2.0, 2.3)	Ref. group	
Any mental illness ^d				
Yes	6.8	(6.2, 7.3)	4.5***	(4.1, 5.0)
No	1.6	(1.4, 1.7)	Ref. group	
Alcohol abuse/dependence ^a				
Yes	14.4	(13.2, 15.6)	10.0***	(9.0, 11.2)
No	1.6	(1.5, 1.8)	Ref. group	
Current cigarette smoking ^e				
Yes	7.5	(6.9, 8.1)	6.8***	(6.1, 7.6)
No	1.2	(1.1, 1.3)	Ref. group	

Notes. OR = Odds ratio, CI = Confidence interval, Ref. group = Reference group.

*** $p < 0.001$.

^a Drug and alcohol abuse and dependence criteria used in the NSDUH were defined based upon the criteria listed in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). Illicit substances included marijuana, hallucinogens, heroin, inhalants, tranquilizers, cocaine, pain relievers, stimulants, and sedatives.

^b The five racial/ethnicity categories (White, Black, Hispanic, American Indian/Alaska Native, Asian, Other) are mutually exclusive; "Other" includes Native Hawaiians or Other Pacific Islanders and persons of two or more races. Persons identified as Hispanic might be of any race.

^c Based on reported family income and poverty thresholds published by the U.S. Census Bureau.

^d Any mental illness is defined by the NSDUH as a diagnosable mental, behavioral, or emotional disorder, other than a developmental or substance use disorder, that met the criteria found in the DSM-IV. For details on the methodology, see Section B.4.3 in Appendix B of the Results from the 2011 NSDUH: Mental Health Findings.

^e Persons who reported ever smoking all or part of a cigarette in the 30 days preceding the interview AND smoked ≥ 100 cigarettes in their lifetime.

variables to include in subsequent multivariable models. PROC SURVEYLOGISTIC in SAS (SAS Institute, Cary, NC) was used to conduct the analyses, with variances estimated using Taylor series linearization. Next, multivariable logistic regression was conducted using those variables that were significantly associated with illicit drug abuse/dependence in univariate logistic regression analyses, initially with no interaction terms. Analyses were repeated examining all possible two-way interactions across the risk factors that were retained in the initial model. More specifically, we first examined associations between each interaction (e.g., age * race, poverty * any mental illness) and illicit drug abuse/dependence individually. Those two-way interactions that were significantly associated with illicit drug abuse/dependence were then entered into the multivariable model and retained in a final regression model if they remained statistically significant. SAS 9.4 software was used to conduct these analyses and statistical significance across all tests was defined as $p < 0.05$.

Initial characterization of the combined effects of co-occurring risk factors was based on the interaction results from the final logistic regression model. We identified risk factors as operating independently when there was no significant interaction between two risk factors that were each significantly associated with illicit drug abuse/dependence at the univariate level. Where significant interactions between two risk factors occurred, their combined effects were interpreted as representing a significant deviation from the multiplicative product of the odds ratios expected from combining independent risk factors. Further determination of whether deviations represented less or greater than the expected increases in the odds ratio for illicit drug abuse/dependence was determined by reviewing graphical displays of the prevalence of illicit drug abuse/dependence within each subgroup along with corresponding odds ratios (see Fig. 2).

Lastly, classification and regression tree (CART) analyses were conducted (Breiman et al., 1984) to supplement the multiple logistic regression modeling using the same eight risk factors for illicit drug abuse/dependence. CART analysis is a nonparametric procedure for dividing a population of interest (in this case the population of non-institutionalized U.S. adults aged ≥ 18 years) based on a dependent variable of interest (i.e., illicit drug abuse/dependence) (Lemon et al., 2003), and, in the process, identifying independent variables that are most strongly associated with that dependent variable. The CART analysis begins by identifying the most important independent variable for dividing the population (i.e., parent node) into two groups (i.e., child nodes) using a predetermined branching criterion, where nodes are split based on their purity using the Gini impurity function (Breiman et al., 1984). A “pure” node has no variability in the dependent variable

whereas completely “impure” nodes have a conditional probability of $p(k|t) = 0.5$, where k refers to the dependent variable and t refers to the node (Lei et al., 2015). A splitting or branching criterion “selects the split that has the largest difference between the impurity of the parent node and a weighted average of the impurity of the two child nodes” (Lemon et al., 2003, p. 174). Given that the dependent variable was binary, we used the Gini impurity function to split nodes, repeating the process recursively with every subsample until the subsample reached a minimum size or no further splits could be made. The tree was constructed using R’s rpart package (R Development Core Team, 2013; Therneau et al., 2013) using the classification method in R (as the dependent variable was binary) and with survey weights included (as the NSDUH uses a multi-stage sampling procedure). A fully saturated tree was produced initially, which was then pruned by selecting the complexity parameter that minimized cross-validation error.

3. Results

3.1. Logistic regression analyses

Overall prevalence of past year illicit drug abuse/dependence in the present sample of U.S. adults was 2.53% (Table 1, left column). Results of the univariate logistic regression analyses indicated that all eight risk factors were significantly associated with increased odds of past-year illicit drug abuse/dependence (Table 1, right column).

Each of the risk factors remained significant when entered simultaneously into the multiple logistic regression model. Of the 28 possible two-way interactions across the eight risk factors, only 5 of 28 (18%) were significant, indicating that these variables generally continued to function independently when present concurrently (Fig. 1, Table 2). The significant interactions observed involved five of the eight risk factors (i.e., gender, mental illness, alcohol abuse/dependence, age, current cigarette smoking). Alcohol abuse/dependence and mental illness were each involved in three interactions, gender was involved in two interactions, and age and current cigarette smoking status were each involved in one interaction. Race/ethnicity, education and poverty status were not involved in any interactions.

As shown in Fig. 2, these interactions were uniformly situations where the prevalence of illicit drug abuse/dependence in the riskiest subgroup was significantly less than expected if each of the respective risk factors retained their independent associations with the dependent variable when combined. For example, the interaction between past year alcohol abuse/dependence and current cigarette smoking is displayed in Fig. 2, Panel A. The intersection of these two risk factors

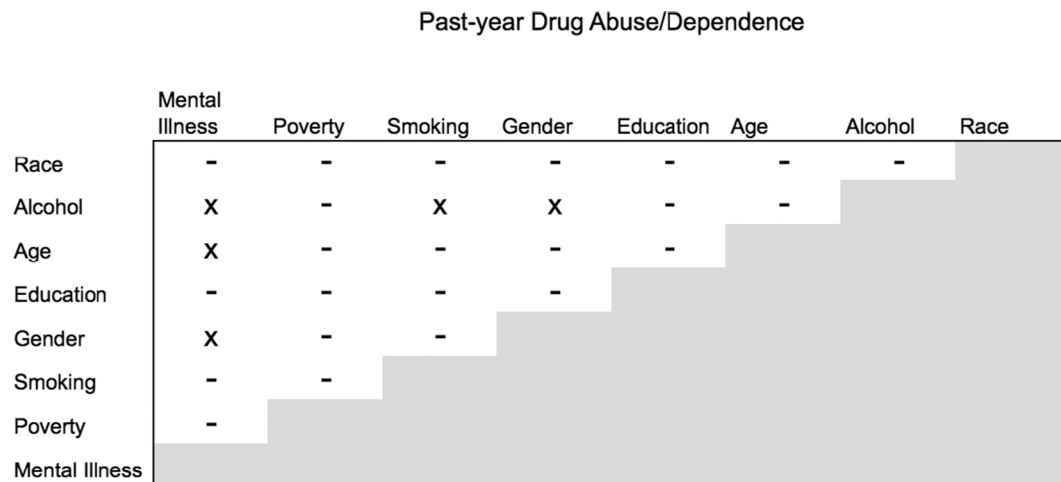


Fig. 1. Outcomes of testing all possible two-way interactions among significant risk factors for drug abuse/dependence in the multivariable logistic regression analysis; X and - symbols indicate risk-factor combinations where there was and was not a significant interaction, respectively.

Table 2

Results from multiple logistic regression examining associations between drug abuse/dependence^a and all possible two-way intersections between eight risk factors ($n = 114,426$). – National Survey on Drug Use and Health (NSDUH), United States, 2011–2013.

Risk factors	OR	95% CI
Gender	–	–
Age group (years) ^b	–	–
Race/ethnicity ^c		
American Indian/Alaska Native vs. Asian	2.1	(1.2, 3.6)
Black vs. Asian	2.5	(1.6, 3.8)
Hispanic vs. Asian	1.6	(1.0, 2.4)
Other vs. Asian	1.8	(1.1, 3.0)
White vs. Asian	1.7	(1.1, 2.5)
Education level ^d		
<High school vs. college graduate	2.0	(1.6, 2.4)
High school graduate vs. college graduate	1.7	(1.4, 2.1)
Some college vs. college graduate	1.6	(1.3, 1.9)
Poverty status ^e		
At or above poverty vs. below poverty	1.1	(1.0, 1.3)
Any mental illness ^f	–	–
Alcohol abuse/dependence ^a	–	–
Current cigarette smoking ^g	–	–
Interactions	OR	95% CI
Gender * any mental illness ^f		
Mental illness, male vs. female	1.5	(1.3, 1.8)
No mental illness, male vs. female	2.1	(1.7, 2.5)
Male, mental illness vs. no mental illness	3.5	(2.7, 4.6)
Female, mental illness vs. no mental illness	4.8	(3.3, 7.1)
Gender * alcohol abuse/dependence ^a		
Alcohol abuse/dependence, male vs. female	1.4	(1.1, 1.7)
No alcohol abuse/dependence, male vs. female	2.2	(1.9, 2.6)
Male, alcohol abuse/dependence vs. no alcohol abuse/dependence	3.7	(3.2, 4.3)
Female, alcohol abuse/dependence vs. no alcohol abuse/dependence	6.0	(5.0, 7.2)
Age group (years) ^b * any mental illness ^f		
Mental illness		
18–25 vs. 65 and older	8.7	(4.3, 17.7)
26–44 vs. 65 and older	4.0	(2.0, 8.1)
45–64 vs. 65 and older	1.7	(0.9, 3.4)
No mental illness		
18–25 vs. 65 and older	23.4	(11.5, 47.8)
26–44 vs. 65 and older	7.6	(3.7, 15.5)
45–64 vs. 65 and older	3.4	(1.7, 6.8)
Any mental illness ^f * alcohol abuse/dependence ^a		
Alcohol abuse/dependence, mental illness vs. no mental illness	3.5	(2.4, 5.0)
No alcohol abuse/dependence, mental illness vs. no mental illness	4.9	(3.6, 6.6)
Mental illness, alcohol abuse/dependence vs. no alcohol abuse/dep	4.0	(3.4, 4.7)
No mental illness, alcohol abuse/dependence vs. no alcohol abuse	5.6	(4.7, 6.7)
Current cigarette smoking ^g * alcohol abuse/dependence ^a		
Alcohol abuse/dependence, current smoker vs. non-smoker	2.4	(2.0, 2.9)
No alcohol abuse/dependence, current smoker vs. non-smoker	4.6	(3.9, 5.4)
Current smoker, alcohol abuse/dependence vs. no alcohol abuse	3.4	(2.9, 4.0)
Non-smoker, alcohol abuse/dependence vs. no alcohol abuse	6.5	(5.5, 7.7)

^a Drug and alcohol abuse and dependence criteria used in the NSDUH were defined based upon the criteria listed in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). Illicit substances included marijuana, hallucinogens, heroin, inhalants, tranquilizers, cocaine, pain relievers, stimulants, and sedatives.

^b Among persons aged ≥ 18 years (18–25, 26–44, 45–64, ≥ 65 years).

^c The five racial/ethnicity categories (White, Black, Hispanic, American Indian/Alaska Native, Asian, Other) are mutually exclusive; “Other” includes Native Hawaiians or Other Pacific Islanders and persons of two or more races. Persons identified as Hispanic might be of any race.

^d < HS, HS, some college, college graduate.

^e Based on reported family income and poverty thresholds published by the U.S. Census Bureau.

^f Any mental illness is defined by the NSDUH as a diagnosable mental, behavioral, or emotional disorder, other than a developmental or substance use disorder, that met the criteria found in the DSM-IV. For details on the methodology, see Section B.4.3 in Appendix B of the Results from the 2011 NSDUH: Mental Health Findings.

^g Persons who reported smoking of all or part of a cigarette in the 30 days preceding the interview AND smoking ≥ 100 cigarettes in their lifetime.

produces four different subgroups. In those without alcohol abuse/dependence, the prevalence of illicit drug abuse/dependence among current cigarette smokers (7.8%) is nearly five-fold higher than prevalence in the subgroup with no alcohol abuse/dependence who are not current smokers (1.6%), OR = 4.6 (95% CI: 3.9, 5.4). In contrast, in the two subgroups with alcohol abuse/dependence, the prevalence of illicit drug abuse/dependence is only two-fold higher in the subgroup of respondents with both alcohol abuse/dependence and current smoking (26.7%) relative to those with alcohol abuse/dependence who are not current smokers (12.4%), OR = 2.4 (95% CI: 2.0, 2.9). We saw no interactions where the odds ratio for illicit drug abuse/dependence associated with a risk factor combination exceeded expectations based on the product of the odds ratio for the two individual risk factors.

Using logistic regression procedures to characterize more than two-way interactions when dealing with eight risk factors often becomes impractical in terms of interpretation (e.g., [Lemon et al., 2003](#)). Thus we used the CART analysis for examining potential combinations of all eight risk factors and for comparing their relative strength.

3.2. Classification and regression tree (CART) analyses

The CART analysis identified past year alcohol abuse/dependence as having the strongest association with past year illicit drug abuse/dependence, followed by cigarette smoking status, age, mental illness, education, poverty status, race, and gender. [Fig. 3](#) shows a pruned classification tree modeling changes in the prevalence of past year illicit drug abuse/dependence associated with the various risk-factor combinations. The graphic is designed to represent an inverted tree.

The top-most rectangle in [Fig. 3](#) is referred to as the root node, which represents 100% of the U.S. adult non-institutionalized population (displayed in the bottom row of information in this node) of which 97.47% of adults had no past year illicit drug abuse/dependence and 2.53% met criteria for illicit drug abuse/dependence (displayed in top row of information in the node). The first split of the entire population was based on whether someone had past year alcohol abuse/dependence (dashed lines descending from the root node). Those without past year alcohol abuse/dependence branched leftward and downward to a terminal node (no further splitting/classification possible) where the prevalence of illicit drug abuse/dependence was 1.64% (95% CI: 1.53%, 1.76%). This terminal node classified 93.04% of the population (displayed in bottom row of information in terminal nodes). The 6.96% of the population that met criteria for past year alcohol abuse/dependence branched rightward and downward to a child node (further splitting/classification possible) where the prevalence of illicit drug abuse/dependence increased to 14.37% corresponding to the removal of those without alcohol abuse/dependence.

The second branching was based on current smoking status. Non-smokers branched leftward and downward to a terminal node representing non-smoking U.S. adults with past year alcohol abuse/dependence. The prevalence of illicit drug abuse/dependence in the subpopulation with these characteristics was 8.51% (95% CI: 7.53%, 9.50%). Smokers branched rightward and downward to a child node where prevalence of illicit drug abuse/dependence increased further to 21.70% corresponding to the removal of non-smokers.

The next branching of the subgroup of U.S. adults with comorbid alcohol abuse/dependence and current smoking was based on age. Those ≥ 26 years old branched leftward and downward to a terminal node where illicit drug abuse/dependence prevalence was 16.89% (95% CI: 14.31%, 19.47%) whereas those aged 18–25 years branched rightward and downward to a child node where drug abuse/dependence prevalence increased to 33.88%, approximately thirteen-fold higher than the 2.53% prevalence seen in the entire adult population. Further branching based on mental illness, education, race, gender, and poverty status, respectively, resulted in progressively increasing prevalence rates across the terminal nodes. Note that looking from left to right across the

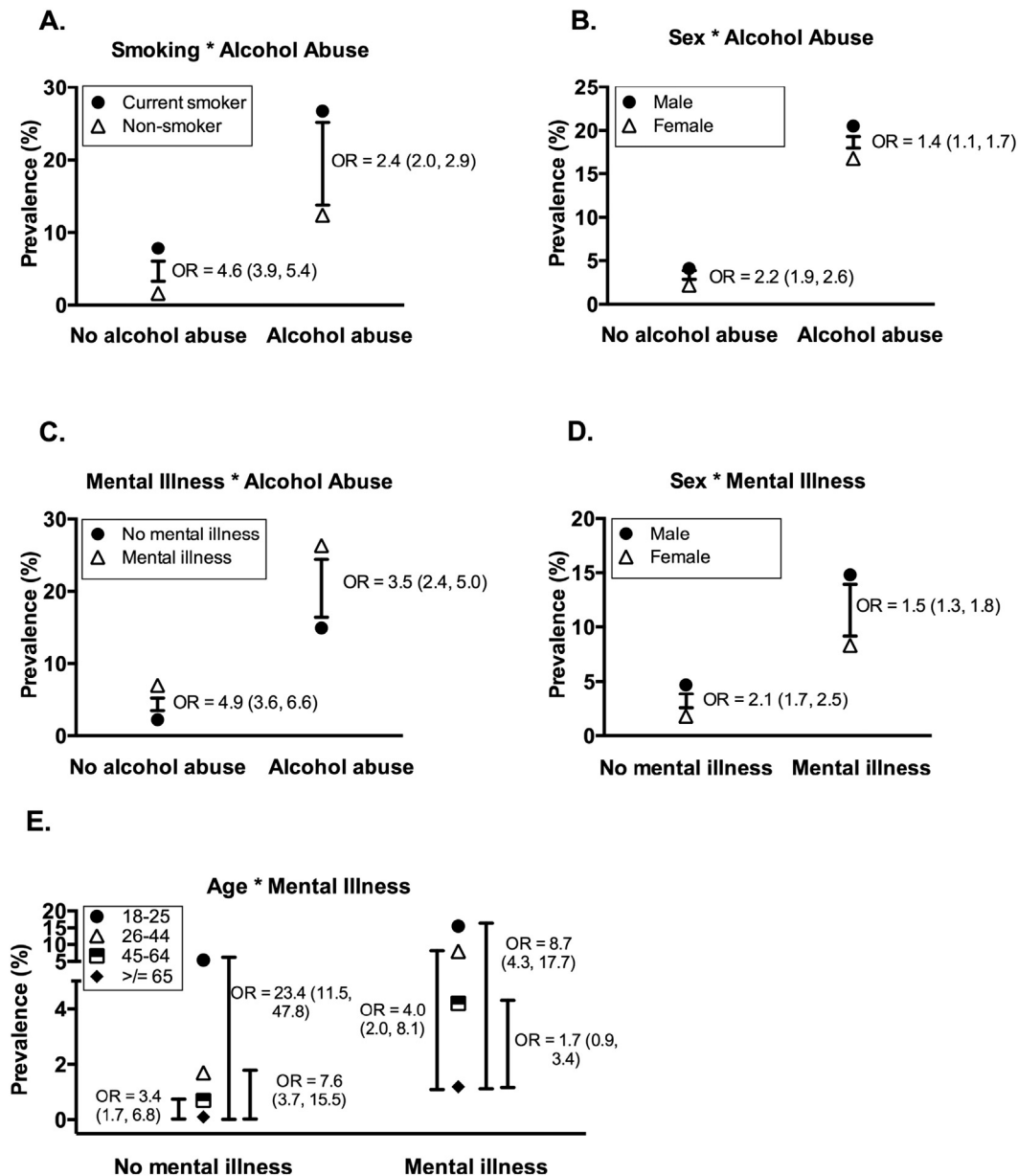


Fig. 2. Significant two-way interactions of risk factors for drug abuse/dependence; data points represent prevalence of illicit drug abuse/dependence within each subgroup along with associated odds ratios.

terminal nodes prevalence rates change as much as 35-fold corresponding to changes in the arrangement of risk factors, with risk factor profiles where prevalence is below 2% to those where the majority (bolded nodes) have past year illicit drug abuse/dependence.

Note the information presented in the middle and bottom rows of the terminal nodes (i.e., proportions of the illicit drug abusing and general population represented by a particular risk factor profile, respectively). As expected considering the low overall prevalence of illicit drug abuse/dependence, those risk-factor profiles that were associated with a higher prevalence classify a relatively small proportion of the U.S. population compared to risk profiles where rates of illicit drug abuse/dependence are low. For example the six right-most terminal nodes classify just 0.88% of the U.S. population combined whereas the three left-most nodes classify the vast majority of the population (approximately 99%). The CART analysis also illustrates the problem of comorbid substance use disorders in this population, with 40.45% of those with past year illicit drug abuse/dependence having a second comorbid

substance use disorder present (i.e., alcohol abuse/dependence) and approximately 27% having both alcohol abuse/dependence and being current cigarette smokers.

To gain more information about risk factors among those without past year alcohol abuse/dependence who nevertheless reported past year illicit drug abuse/dependence, we repeated the CART analysis on only those represented in the left-most terminal node of Fig. 1, excluding past year alcohol abuse/dependence as an independent variable (Fig. 4). Current cigarette smoking was most strongly associated with illicit drug abuse/dependence followed by age, past year mental illness, gender, race, poverty, and education, respectively. The first slice was based on smoking status, with non-smokers moving leftward and downward to a terminal node where prevalence of illicit drug abuse/dependence was reduced to 0.79% while among smokers prevalence was 5.10%. Again, depending on the particular profile of risk-factor combinations, prevalence varied from a low of 0.79% (95% CI: 0.71%, 0.87%) to a high of 82.65% (95% CI: 57.90%, 100%) across the terminal nodes.

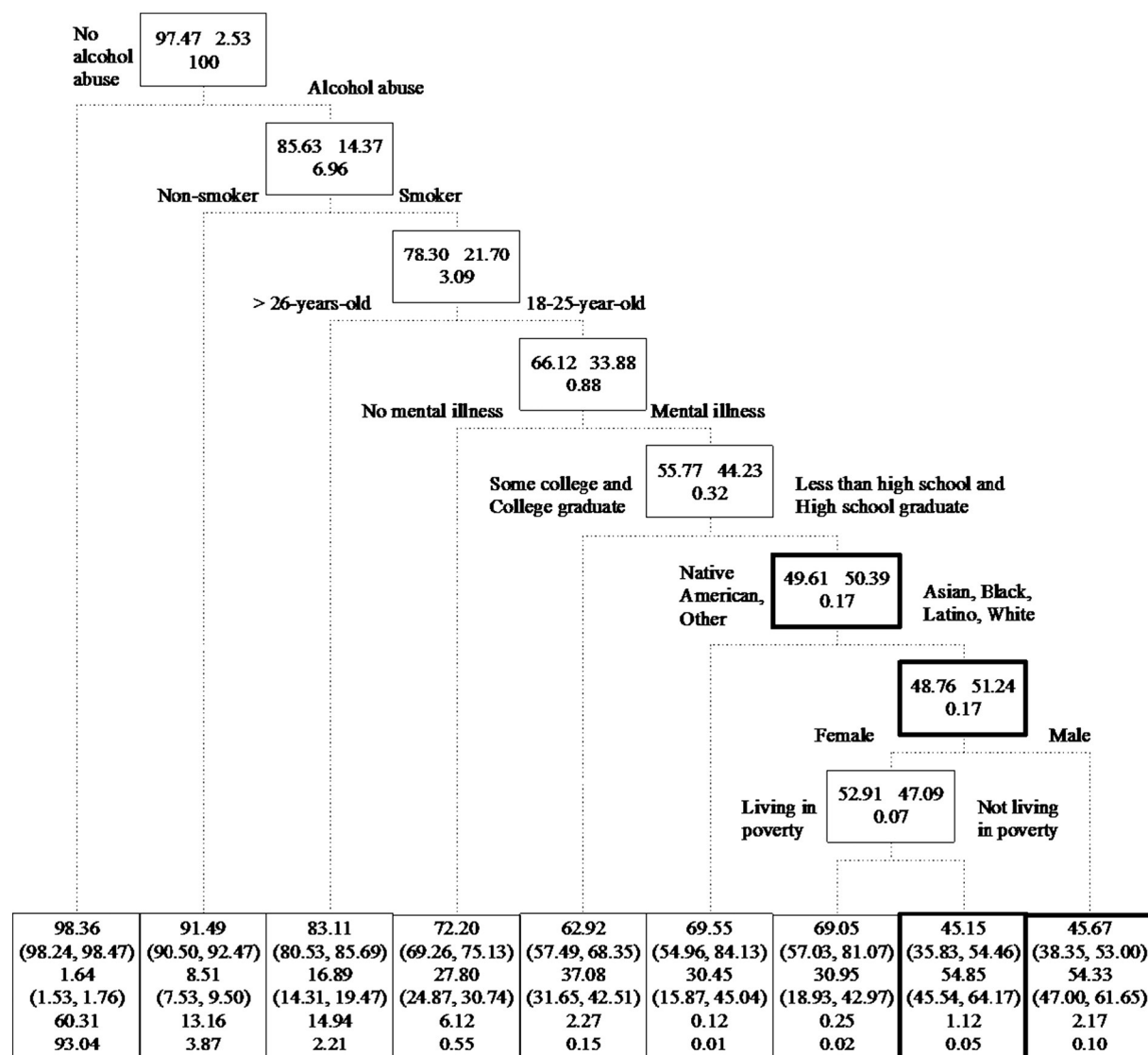


Fig. 3. A pruned, weighted classification and regression tree (CART) model of associations between past year illicit drug abuse/dependence and the following eight risk factors in the U.S. adult (≥ 18 years of age) population: educational attainment, age, race/ethnicity, current cigarette smoking, past year alcohol abuse/dependence, annual income below federal poverty level, gender, and past year mental illness. Results from a saturated model were “pruned” using CART analytic software to reduce complexity (R Development Core Team, 2008). Rectangles (nodes) represent drug abuse/dependence prevalence rates for the entire population (top-most node) or population subgroups (all others nodes). Nodes also list the proportion of the adult population represented. Using the root node as an example, 2.53% of the population met criteria for drug abuse/dependence (97.47% did not), and this node represents 100% of the U.S. non-institutionalized adult population. Lines below nodes represent the binary “yes”-“no” branching around particular risk factors and risk-factor levels, with subgroups in whom the risk factor/level is absent moving leftward and downward and those in whom it is present moving rightward and downward for further potential partitioning based on additional risk factors/levels. The bottom row comprises terminal nodes (i.e., final partitioning for a particular subgroup). Terminal nodes contain the same information as the other nodes plus the percent of all adults with drug abuse/dependence represented by that node and inclusion of 95% CIs for each. Percent of adults with drug abuse/dependence represented is calculated by the following equation: % total population represented by a node \times drug abuse/dependence prevalence in that node/drug abuse/dependence prevalence in the entire study sample $\times 100$. Tallying % adults with drug abuse/dependence represented across all terminal nodes should = 100% of adults with drug abuse/dependence in the U.S. population save possible rounding error.

4. Discussion

The present study was conducted to examine the generality of the results from the Higgins et al. (this issue) report characterizing the intersection of co-occurring risk factors for cigarette smoking to illicit drug abuse/dependence. Consistent with the results of that prior report, co-occurring risk factors generally had independent associations with risk for past year illicit drug abuse/dependence. These findings further underscore fundamental commonalities across these different types of substance use disorders. In both studies less than one-third of the 28 possible two-way interactions tested in the multivariable logistic regression analyses were significant, meaning that the typical pattern was for the risk factors to act independently. Where significant interactions between risk factors were

observed, the common pattern was less-than-expected increases with the combination of risk factors.

The CART analysis characterized the sometimes striking changes in prevalence associated with variations in co-occurring risk factor combinations across population subgroups. The consistency in the general pattern of intersection of co-occurring risk factors for cigarette smoking and illicit drug abuse/dependence suggests that the same empirical framework can be applied to (a) understanding differences in the prevalence of both cigarette smoking and illicit drug abuse/dependence observed across population subgroups, and (b) generating hypotheses about the association between novel risk-factor combinations with each type of substance use.

Although multivariable logistic regression modeling revealed commonalities in the nature of co-occurring risk factors for both cigarette

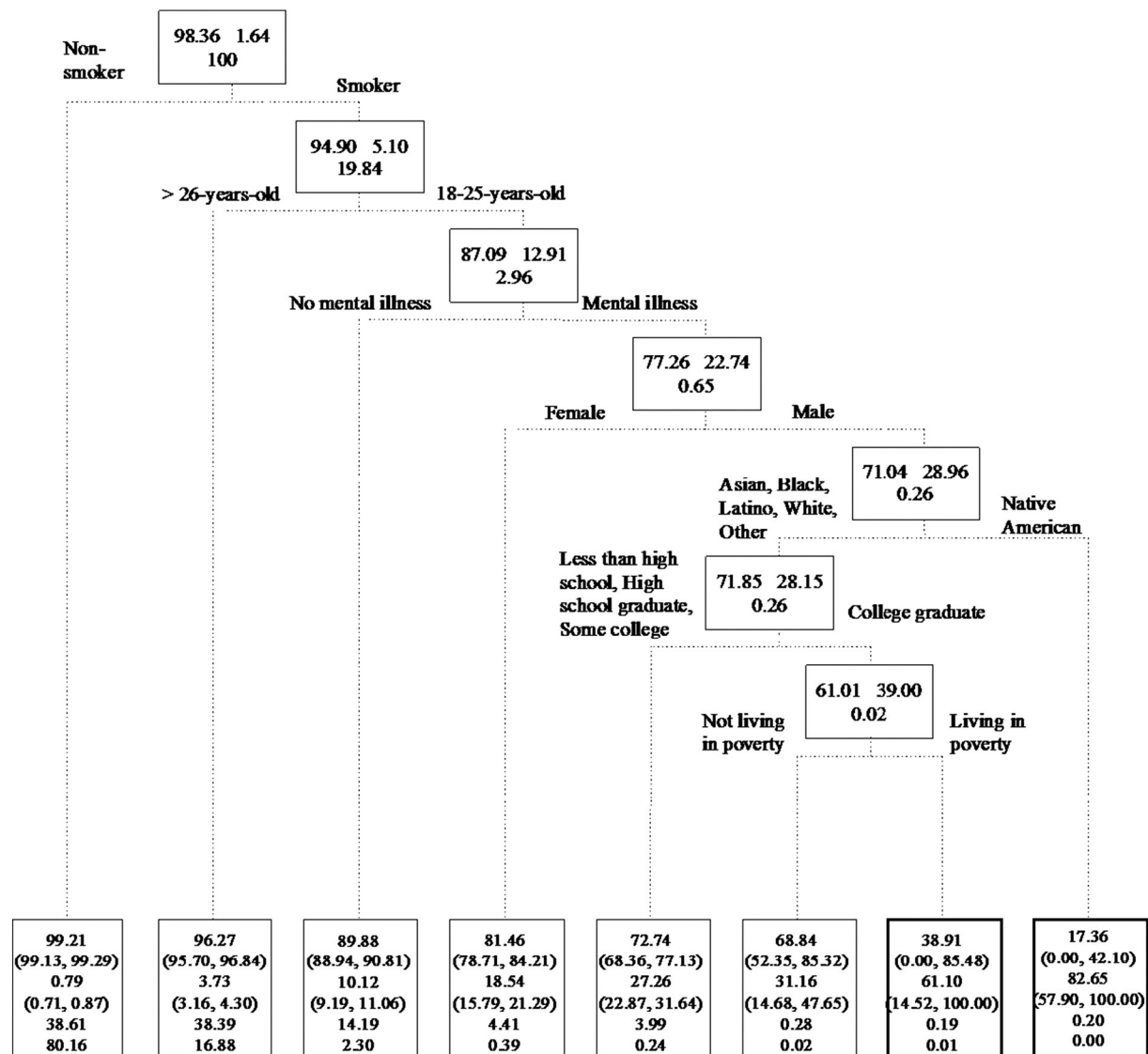


Fig. 4. A pruned, weighted classification and regression tree (CART) model of associations between past-year illicit drug abuse/dependence and seven of the eight risk factors in U.S. adults (≥ 18 years of age) without past-year alcohol abuse/dependence (all risk factors in Fig. 3 minus past-year alcohol abuse/dependence).

smoking and illicit drug abuse/dependence, the CART analysis revealed interesting differences in the relative strength of the risk factors across the respective substance use disorders. For example, educational attainment had the strongest association with cigarette smoking in the Higgins et al. report followed by age, race/ethnicity, drug abuse/dependence, alcohol/abuse dependence, poverty, mental illness, and last by gender. In contrast, past-year alcohol abuse/dependence and current cigarette smoking had the first- and second-most strongest associations with illicit drug abuse/dependence in the present report, and educational attainment ranked fifth in the analysis of the entire adult population and last among seven in the population excluding those with past year alcohol abuse/dependence. The prior and current studies were conducted using the same cross-sectional survey, thereby precluding examination of developmental trajectories of use, abuse, and dependence of cigarettes and illicit drugs. The pattern of results across studies is consistent with trajectories wherein licit substance use (i.e., tobacco, alcohol) precedes illicit drug use (Kirby & Barry, 2012; Lai et al., 2000; Mathers et al., 2006). The stronger association of education with cigarette smoking than illicit drug abuse/dependence is consistent with research in which educational attainment has been shown to mediate the relation between adolescent tobacco use and adulthood use of illicit substances (Strong et al., 2016). It should not be overlooked that some individuals

begin using cigarettes and other substances simultaneously (Biederman et al., 2012) and the possibility that early cigarette smoking may directly facilitate other drug use by modulating brain structure and function (e.g., mesolimbic dopamine activity) (Shadel et al., 2000). It will be important for future longitudinal studies to parse out the contributions of various biological and environmental risk factors for use, abuse, and dependence of different substances.

While not the primary aim of the present study, the almost 35-fold differences in the prevalence of drug abuse/dependence associated with the different risk profiles identified in the CART analysis of the entire adult population and >100-fold difference in the additional CART analysis focused on those without past-year alcohol abuse/dependence has the potential to be practically useful for purposes of targeting intervention efforts towards those at greatest risk. Certainly such targeting can be useful when making decisions about how to maximize the impact of limited funds (Lemon et al., 2003). Also important to keep in mind, however, is that risk is not zero for any of the risk profiles outlined in this report. Indeed, the largest proportion of individuals with past year illicit drug abuse/dependence in the current study was in the left-most terminal node of Fig. 3 (those without past year alcohol abuse/dependence). Prevalence was only 1.64% in that node, but because that profile categorized 93% of the general population, the percentage of

the overall population of those with illicit drug abuse and dependence in that terminal node was nevertheless relatively large (60.31%). Hence, the optimal strategy in prevention and treatment efforts would appear to be a balanced portfolio that includes targeted and general population intervention efforts. Although the particularly high-risk nodes mostly represent small proportions of the overall population, they nevertheless merit targeting because of the high likelihood of reaching those most in need. Those high-risk profile nodes also include substantial rates of co-morbid alcohol abuse/dependence and cigarette smoking, as well as poverty, low education and other risk factors that increase the likelihood of difficulties quitting substance use, serious adverse health impacts, and the potential of further exacerbating health disparities (Drobes, 2002; Guydish et al., 2011; Hser et al., 1994; Hughes, 2002; Sobell, 2002; Swendsen et al., 2010).

The present study has several limitations that merit mention. First, the NSDUH is a cross-sectional survey, which precludes examining changes in the odds of illicit drug abuse/dependence corresponding to changes in co-occurring risk factors within an individual over time. Extending this research to longitudinal surveys such as the National Epidemiological Survey on Alcohol and Related Conditions (see Grant et al., 2014) or the Population Assessment of Tobacco and Health (<http://www.pathstudyinfo.nih.gov/UI/HomeMobile.aspx>) will facilitate an examination of whether risk factors retain their independent associations with illicit drug abuse/dependence when assessed prospectively over time. Second, the NSDUH excludes several groups in which illicit drug abuse/dependence is more prevalent relative to the general population, including those in the active military, jail, or homeless. Thus the present results may not generalize to these subgroups. Adolescents were also excluded from the present sample thereby potentially limiting generality to that important subgroup. These limitations notwithstanding, the present study contributes new knowledge regarding the nature of the intersection of co-occurring risk factors for illicit drug abuse/dependence and commonalities between risk for cigarette smoking and illicit drug abuse/dependence, while also identifying particularly high-risk subpopulations in which increased efforts to prevent or reduce illicit drug use are needed.

Research funding

This research was supported by National Institute of General Medical Sciences Centers of Biomedical Research Excellence Center Award P20GM103644 and National Institute on Drug Abuse Institutional Training Award T32DA007242.

Declaration of interests

None to declare.

Transparency document

The Transparency document related to this article can be found, in the online version.

References

- Substance Abuse and Mental Health Services Administration (SAMHSA), 2012. Results from the 2011 National Survey on Drug Use and Health: Summary of the National Findings. NSDUH Series H-44 (HHS Publication No. (SMA) 12-4713), Rockville, MD.
- Biederman, J., Petty, C.R., Hammerness, P., Batchelder, H., Faraone, S.V., 2012. Cigarette smoking as a risk factor for other substance misuse: 10-year study of individuals with and without attention-deficit hyperactivity disorder. *Br. J. Psychiatry* 201, 207–214. <http://dx.doi.org/10.1192/bjp.bp.111.100339>.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. Wadsworth, Belmont, California.
- Center for Behavioral Health Statistics and Quality, 2014. 2013 National Survey on Drug Use and Health Public Use File Codebook. Substance Abuse and Mental Health Services Administration, Rockville, MD http://www.icpsr.umich.edu/cgibin/file?comp=none&study=35509&ds=1&file_id=1166336&path=SAMHDA Accessed June 16, 2015.
- Compton, W.M., Thomas, Y.F., Stinson, F.S., Grant, B.F., 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the USA: results from the National Epidemiologic Survey on alcohol and related conditions. *Arch. Gen. Psychiatry* 64, 566–576. <http://dx.doi.org/10.1001/archpsyc.64.5.566>.
- Drobes, D.J., 2002. Cue reactivity in alcohol and tobacco dependence. *Alcohol. Clin. Exp. Res.* 26, 1928–1929. <http://dx.doi.org/10.1097/0000374-200212000-00026>.
- Grant, B., Amsbary, M., Chu, A., Sigman, R., Kali, J., Sugawana, Y., et al., 2014. Source and Accuracy Statement: National Epidemiologic Survey on Alcohol and Related Conditions-III (NE-SARC-III). National Institute on Alcohol Abuse and Alcoholism, Rockville, MD, p. 2014.
- Guydish, J., Passalacqua, E., Tajima, B., Chan, M., Chun, J., Bostrom, A., 2011. Smoking prevalence in addiction treatment: a review. *Nicotine Tob. Res.* 13, 401–411. <http://dx.doi.org/10.1093/ntr/ntn048>.
- Han, B., Gfroerer, J.C., Colliver, J.D., Penne, M.A., 2009. Substance use disorder among older adults in the USA in 2020. *Addiction* 104, 88–96. <http://dx.doi.org/10.1111/j.1360-0443.2008.02411.x>.
- Hser, Y.I., McCarthy, W.J., Anglin, M.D., 1994. Tobacco use as a distal predictor of mortality among long-term narcotics addicts. *Prev. Med.* 23, 61–69. <http://dx.doi.org/10.1006/pmed.1994.1009>.
- Huang, B., Dawson, D.A., Stinson, F.S., Hasin, D.S., Ruan, W.J., Saha, T.D., Grant, B.F., 2006. Prevalence, correlates and comorbidity of nonmedical prescription drug use and drug use disorders in the USA: results of the National Epidemiologic Survey on alcohol and related conditions. *J. Clin. Psychiatry* 67, 1062–1073. <http://dx.doi.org/10.4088/jcp.v67n0708>.
- Hughes, J., 2002. Do smokers with current or past alcoholism need different or more intensive treatment? *Alcohol. Clin. Exp. Res.* 26, 1934–1935. <http://dx.doi.org/10.1097/0000374-200212000-00029>.
- Kirby, T., Barry, A.E., 2012. Alcohol as a gateway drug: a study of US 12th graders. *J. Sch. Health* 82, 371–379. <http://dx.doi.org/10.1111/j.1746-1561.2012.00712.x>.
- Lai, S., Lai, H., Page, J.B., McCoy, C.B., 2000. The association between cigarette smoking and drug abuse in the United States. *J. Addict. Dis.* 19, 11–24. http://dx.doi.org/10.1300/j069v19n04_02.
- Lei, Y., Nollen, N., Ahluwalia, J.S., Yu, Q., Mayo, M.S., 2015. An application in identifying high-risk populations in alternative tobacco product use utilizing logistic regression and CART: a heuristic comparison. *BMC Public Health* 15, 341. <http://dx.doi.org/10.1186/s12889-015-1582-z>.
- Lemon, S.C., Roy, J., Clark, M.A., Friedmann, P.D., Rakowski, W., 2003. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann. Behav. Med.* 26 (3), 172–181. http://dx.doi.org/10.1207/s15324796abm2603_02.
- Mathers, M., Toumbourou, J.W., Catalano, R.F., Williams, J., Patton, G.C., 2006. Consequences of youth tobacco use: a review of prospective behavioral studies. *Addiction* 101, 948–958. <http://dx.doi.org/10.1111/j.1360-0443.2006.01438.x>.
- Population Assessment of Tobacco and Health, d. <http://www.pathstudyinfo.nih.gov/UI/HomeMobile.aspx> Accessed 7/15/2016.
- R Development Core Team, 2008. An introduction to R. <http://www.r-project.org/> Accessed/15/2016.
- R Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria URL <http://www.R-project.org> Accessed 7/15/2016.
- Shadel, W.G., Shiffman, S., Niaura, R., Abrams, D.B., 2000. Current models of nicotine dependence: what is known and what is needed to advance understanding of tobacco etiology among young. *Drug Alcohol Depend.* 59, S9–22. [http://dx.doi.org/10.1016/S0376-8716\(99\)00162-3](http://dx.doi.org/10.1016/S0376-8716(99)00162-3).
- Sobell, M.B., 2002. Alcohol and tobacco: clinical and treatment issues. *Alcohol. Clin. Exp. Res.* 26, 1954–1955. <http://dx.doi.org/10.1097/0000374-200212000-00037>.
- Strong, C., Juon, H., Ensminger, M.E., 2016. Effect of adolescent cigarette smoking on adulthood substance use and abuse: the mediating role of educational attainment. *Subst. Use Misuse* 51, 141–154. <http://dx.doi.org/10.3109/10826084.2015.1073323>.
- Substance Abuse and Mental Health Services Administration (SAMHSA), 2013. Results from the 2012 National Survey on Drug Use and Health: Summary of the National Findings. NSDUH Series H-46 (HHS Publication No. (SMA) 13-4795), Rockville, MD.
- Substance Abuse and Mental Health Services Administration (SAMHSA), 2014. Results from the 2013 National Survey on Drug Use and Health: Summary of the National Findings. NSDUH Series H-46 (HHS Publication No. (SMA) 14-4863), Rockville, MD.
- Swendsen, J., Conway, K.P., Degenhardt, L., Glantz, M., Jin, R., Merikangas, K.R., ... Kessler, R.C., 2010. Mental disorders as risk factors for substance use, abuse, and dependence: results from the 10-year follow-up of the National Comorbidity Survey. *Addiction* 105, 1117–1128.
- Therneau, T., Atkinson, B., Ripley, B., 2013. Rpart: recursive partitioning. R package version 4.1-3. <http://CRAN.R-project.org/package=rpart>.