



ELSEVIER

Contents lists available at ScienceDirect

Preventive Medicine

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

Current cigarette smoking among U.S. college graduates

Sulamunn R.M. Coleman^{a,b,*}, Diann E. Gaalema^{a,b,c}, Tyler D. Nighbor^{a,b}, Allison A. Kurti^{a,b}, Janice Y. Bunn^d, Stephen T. Higgins^{a,b,c}^a Vermont Center on Behavior and Health, University of Vermont, 1 South Prospect Street, MS 482, Burlington, VT 05401-1419, United States of America^b Department of Psychiatry, University of Vermont, 1 South Prospect Street, MS 482, Burlington, VT 05401-1419, United States of America^c Department of Psychological Science, University of Vermont, 1 South Prospect Street, MS 482, Burlington, VT 05401-1419, United States of America^d Department of Mathematics and Statistics, University of Vermont, 1 South Prospect Street, MS 482, Burlington, VT 05401-1419, United States of America

ARTICLE INFO

Keywords:

Cigarette smoking
College graduates
Risk factors
Educational attainment
Substance use disorders
U.S. adults
Classification and Regression Tree (CART)
modeling

ABSTRACT

Although U.S. college graduates are at relatively low risk for smoking, 12–15% of U.S. smokers (~8 million people) are college graduates. Few studies have examined smoking risk among college graduates. To address that gap, the present study examined smoking risk among U.S. college graduates and those who did not graduate from college in a nationally representative sample of adults (National Survey on Drug Use and Health 2011–2017, $n = 202,137$). We examined smoking risk in association with well-established risk factors: alcohol abuse/dependence, drug abuse/dependence, mental illness, age, sex, race/ethnicity, and poverty status, using group contrasts and Classification and Regression Tree (CART) modeling. Smoking prevalence among U.S. college graduates and non-graduates was 10% and 26%, respectively. College graduates initiated any smoking and daily smoking at a later age and were lighter smokers than smokers who did not graduate college. Within college graduate and non-graduate groups, prevalence rates varied by orders of magnitude across different risk-factor profiles (ranges = 3–37% and 14–73% among graduates and non-graduates, respectively). Past year drug abuse/dependence was a robust predictor of smoking prevalence in both populations. For college graduates, past year alcohol abuse/dependence and mental illness were stronger predictors of smoking compared to those who did not graduate college, for whom race/ethnicity and age were stronger predictors. Overall, smoking risk increases to surprisingly high levels, even among college graduates, when select risk factors co-occur, particularly psychiatric conditions. Socio-demographic risk factors appear to be less robust predictors of smoking risk among college graduates relative to those who did not graduate college.

1. Introduction

U.S. smoking prevalence has decreased substantially since the landmark 1964 Surgeon General's report (Alberg et al., 2014; Levy et al., 2004). Nevertheless, smoking remains the leading cause of preventable death, contributing to more than 480,000 deaths annually in the U.S. (Ma et al., 2018). Moreover, decreases in smoking prevalence have been unevenly distributed across the general population, with considerable decreases evident among the more educated and affluent, but less so among those who are socioeconomically disadvantaged, from racial/ethnic minority groups, or those with psychiatric conditions (Gaalema et al., 2018; Higgins et al., 2016; Higgins, 2014; Hiscock et al., 2012; Schroeder and Koh, 2014). These disparities have increased interest in understanding how psychiatric and socio-demographic characteristics affect smoking risk.

Two recent studies (Gaalema et al., 2018; Higgins et al., 2016) on predictors of smoking risk in the U.S. adult population have characterized how common co-occurring risk factors (e.g., mental illness, alcohol abuse/dependence, drug abuse/dependence, age, educational attainment, sex, race/ethnicity, and poverty status) act independently and summatively in accounting for differences in smoking risk, with prevalence of smoking varying from low to strikingly high levels depending on the particular profile of risk factors present. In both studies, having a 4-year college degree (i.e., Bachelor's degree) or higher represented the only single-factor risk profile identified and was associated with the lowest smoking prevalence (9–11%). Among those who were not college graduates, smoking prevalence was as high as 74% depending on the profile of risk factors present. These studies underscore the importance of higher education in predicting low risk for current smoking. However, also important to recognize is that

* Corresponding author at: Vermont Center on Behavior and Health, University of Vermont, 1 South Prospect Street, MS 482, Burlington, VT 05401-1419, United States of America.

E-mail address: sulamunn.coleman@uvm.edu (S.R.M. Coleman).

<https://doi.org/10.1016/j.ypmed.2019.105853>

Received 2 August 2019; Received in revised form 16 September 2019; Accepted 19 September 2019

Available online 22 October 2019

0091-7435/ © 2019 Elsevier Inc. All rights reserved.

according to the 2018 U.S. Census, 76.9 million adults old enough to have completed their education (≥ 25 years) have a Bachelor's degree or higher (U.S. Census Bureau, 2018). Thus, even smoking rates as low as 9–11% indicate that between 7 and 8.5 million U.S. adults continue to smoke despite having more than a high school education.

There are relatively few studies in the literature systematically examining smoking risk among college graduates (Maralani, 2013, 2014). The purpose of the current study is to begin filling that gap by examining differences in smoking risk between U.S. college graduates and those who did not graduate college. Using a U.S. nationally-representative sample, we (1) compare psychiatric and socio-demographic risk factors for smoking as well as smoking characteristics among college graduates (i.e., U.S. adults with at least a 4-year college degree) and those who did not graduate from college, and (2) examine the intersection of these psychiatric and socio-demographic risk factors (smoking risk profiles) with the overarching goal of gaining greater understanding of smoking risk among college graduates.

2. Methods

2.1. Data source

Data for this report are from the most recent seven years (2011–17) of the National Survey on Drug Use and Health (NSDUH), a multiyear, cross-sectional survey of the U.S. non-institutionalized population that measures the prevalence of cigarette smoking (and other substance use) across various socio-demographic groups, psychiatric, and psychosocial risk factors. Detailed descriptions of survey procedures have been provided for each of the survey years (CBHSQ, 2015, 2018; SAMHSA, 2012, 2013, 2014, 2016, 2017, 2018). We pooled across years to ensure an adequate sample of smokers who held a Bachelor's degree or higher. Additionally, this report focused on adults aged ≥ 24 -years to allow sufficient time for one to have completed at least a Bachelor's degree. The overall study sample size was 202,137, including 64,299 college graduates (6913 current smokers) and 137,838 non-graduates of college (41,497 current smokers).

2.2. Measures

In addition to college graduate status, the independent variables examined in this report included seven well-established psychiatric and socio-demographic risk factors for smoking in the U.S. adult population that were also used in the two prior studies examining risk factor profiles mentioned above (Gaalema et al., 2018; Higgins et al., 2016). These other variables include (1) past-year alcohol abuse/dependence, (2) past-year drug abuse/dependence, (3) past-year mental illness, (4) age, (5) sex, (6) race/ethnicity, and (7) poverty status (based on reported family income and poverty thresholds published by the U.S. Census Bureau). Regarding the two past-year substance abuse/dependence variables, the NSDUH computer-assisted interview includes questions designed to measure alcohol and illicit drug abuse/dependence based on criteria from the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders, 4th edition. Respondents were defined as meeting criteria for substance dependence if they responded positively to three or more of the six dependence criteria (e.g., “Needed to use substance more than before to get the desired effects or noticed that using the same amount had less effect than before.”). Respondents were defined as meeting criteria for substance abuse if they responded positively to one or more of the four abuse criteria (e.g., “[...] having problems caused by substance use with family or friends and continued to use substance even though it was thought to be causing problems with family and friends.”) and were determined not to be dependent on the substance of interest (Note: drug use disorders assessed included marijuana, hallucinogens, heroin, inhalants, tranquilizers, cocaine, pain relievers, stimulants, and sedatives). Past-year mental illness was determined using a logistic

regression model developed to predict the probability of mental illness. Variables in the model included the World Health Organization Disability Assessment Schedule (Üstün et al., 2010), the worst month total score of the K6 screening instrument for psychological distress (Kessler et al., 2003), an item assessing suicidal ideation, presence of adult major depression within the past year (endorsement of at least five of the nine DSM-IV criteria for an adult major depressive episode, with at least one of the criteria reflecting depressed mood/loss of interest or pleasure in daily activities), and an age variable differentiating between adults older or younger than 30-years-old. In accordance with CBHSQ (2018) scoring procedures, respondents were determined to have past-year mental illness if the prediction probability for mental illness was ≥ 0.019 .

The dependent variable ‘current smoker status’ was defined as smoking 100 or more cigarettes in one's lifetime and smoking all or part of a cigarette in the past 30 days. Former smoker status was defined as smoking 100 or more cigarettes in one's lifetime, but not in the past 30 days. Never smoker status was defined as not meeting criteria for current or former smoker. Characteristics of college graduate and non-graduate smokers include % current daily smokers, % initiated smoking before age 18, % initiated daily smoking before age 18, mean age at 1st cigarette, mean age started daily smoking, mean number of days smoked in the past 30 days, and mean number of cigarettes smoked per day.

2.3. Statistical methods

Between-group contrasts were conducted to examine differences between college graduates and non-graduates on socio-demographic/psychiatric smoking risk factors and smoking characteristics. Contrasts of smoking characteristics focused on the subsample of current smokers within each population. Next, Classification and Regression Tree (CART) modeling was used to identify risk factor profiles (i.e., combinations of independent risk factors) for current smoking for both populations. CART modeling is a nonparametric procedure that rank orders independent variables (i.e., risk factors) according to their strength of association with the dependent variable in a given population (e.g., current smoker status among college graduates) (Breiman et al., 1984; Lemon et al., 2003). The relative ranking of independent variables is then used to divide the population into mutually exclusive subgroups (i.e., nodes), and in the process, determine smoking prevalence for individuals characterized by each node. This process facilitates the interpretation of otherwise complex interactions between multiple risk factors (Lemon et al., 2003). The “parent” node (i.e., top-most node) represents the starting point showing overall prevalence of the outcomes of interest. Descending first from the “parent” node is the risk factor that most strongly associates with the dependent variable, dividing the entire population into two subgroups (i.e., “child” nodes) (see Figs. 1 and 2). Subsequent divisions are based on the remaining risk factors with the most explanatory power in accounting for current smoker status.

Given that the dependent variable was binary, we used the Gini impurity function to split nodes by selecting “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). This process was repeated recursively with every subsample, until the subsample reached a minimum size represented as a “terminal” node or risk profile on the bottom row of the CART figure. In the current study, the minimal terminal node size was set to $\geq 1\%$ of the individuals within each population. Each node depicted in the CART figures contains two values: the top value represents the prevalence of current smokers for the node, and the bottom value represents the percent of the population characterized by the node (i.e., the particular combination of smoking risk factors). Importantly, terminal nodes towards the left side of the CART figures represent the profiles with the lower risk, and smoking risk generally increases from left to right across

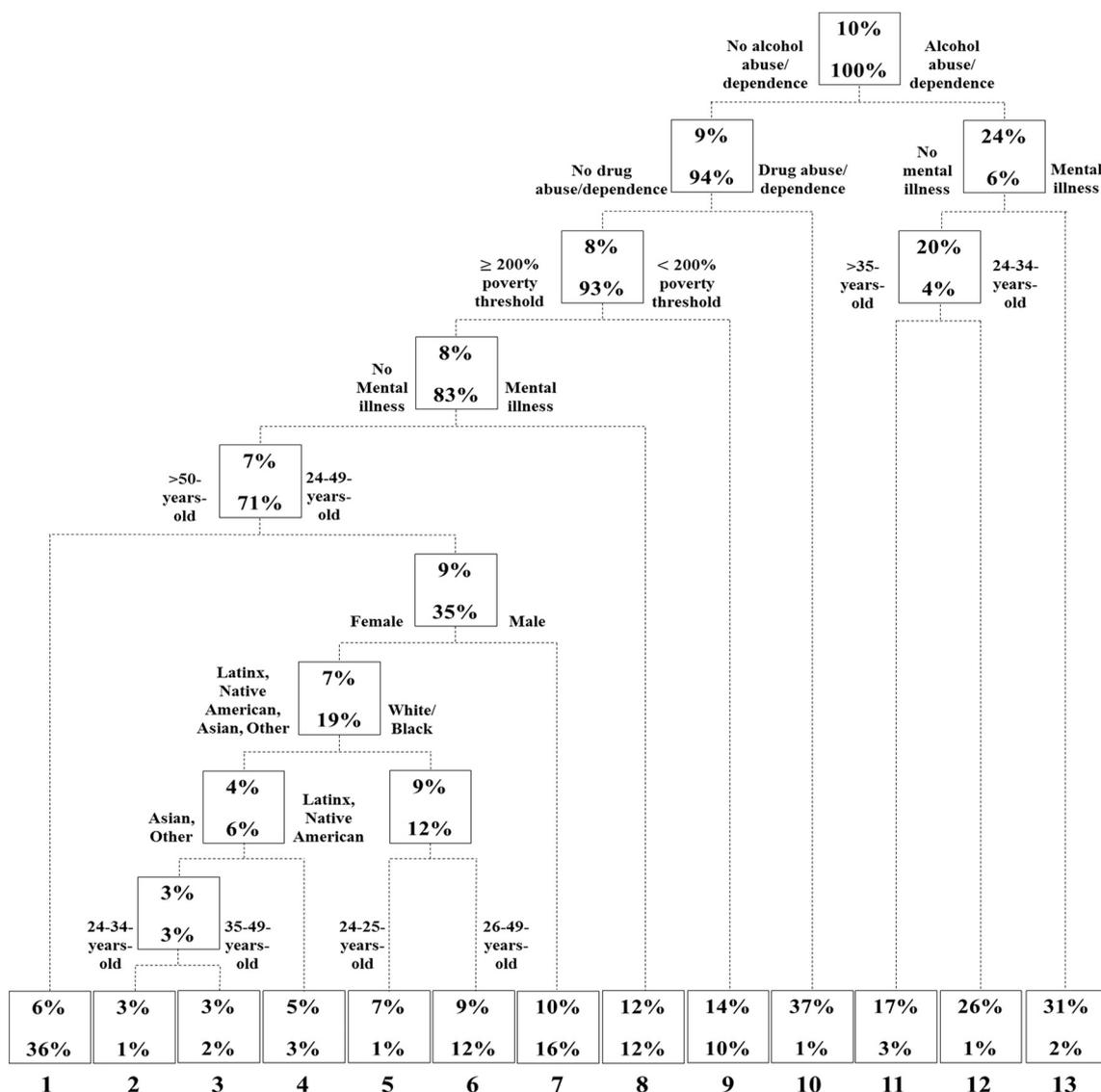


Fig. 1. A pruned, weighted classification and regression tree (CART) model of associations between current (past 30 days) smoking status and the following seven risk factors in the U.S. population of college graduates (≥ 24 years of age): age, race/ethnicity, past year drug abuse/dependence, past year alcohol abuse/dependence, annual income below federal poverty level, and past year mental illness in years 2011–2017 of the National Survey on Drug Use and Health ($N = 64,299$). Rectangles (nodes) represent the entire population (top-most node) or population subgroups (all other nodes). Within each node the top line lists the smoking rate for that population and the bottom line the percent of the U.S. college-educated population represented within that node. Using the root node as an example, this node represents 100% of the U.S. non-institutionalized adult college-educated population and 9.6% (rounded to 10%) of them are smokers. Lines below nodes represent the binary branching around particular risk factors and risk-factor levels into subgroup nodes with further potential partitioning based on additional risk factors/levels. The bottom row comprises terminal nodes or risk profiles (i.e., final partitioning for a particular subgroup, minimal terminal node size set to $\geq 1\%$ individuals). Terminal nodes contain the same information as the other nodes.

the terminal nodes such that the rightmost terminal nodes represent profiles with greater risk. Group contrasts were conducted in SAS 9.4 with survey weights to account for the multi-stage sampling procedures of the NSDUH, and the *rpart* package in R was used to produce the CART models. A fully saturated tree was produced initially and then pruned by selecting the complexity parameter that minimized cross-validation error.

It should be noted that given the rapidly changing tobacco landscape (e.g., changes in age of initiation, use of multiple tobacco products, overall prevalence of smoking), it is important to account for potential cohort effects in studies using nationally representative samples. However, CART modeling does not allow the inclusion of random variables to adjust for potential variance caused by time. We conducted additional CART analyses including survey year as an independent variable, and survey year was found to be the least important predictor

of smoking status for both college graduates and non-graduates (i.e., differences between the models with vs. without survey year as an independent variable were limited to the terminal nodes, suggesting little, if any, effect of survey year). As such, we present only the models that did not include survey year.

3. Results

3.1. Psychiatric and socio-demographic risk factors

College graduates and non-graduates differed with respect to all psychiatric and socio-demographic risk factors except past-year alcohol abuse/dependence and sex (Table 1). Notably, compared to non-graduates, a smaller percentage of college graduates reported past year drug abuse/dependence or mental illness. Additionally, a larger percentage

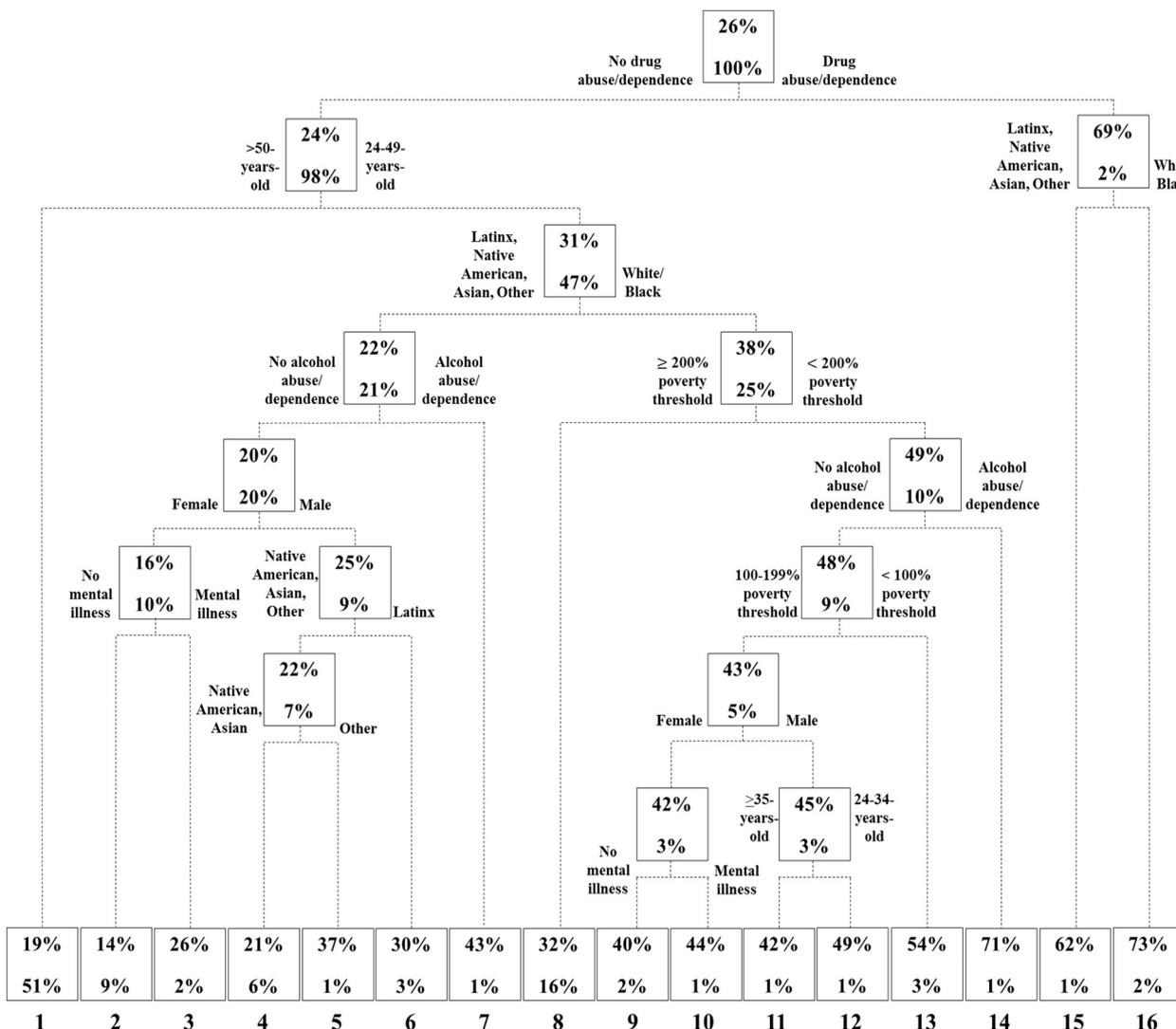


Fig. 2. A pruned, weighted classification and regression tree (CART) model of associations between current (past 30 days) smoking status and the following seven risk factors in the U.S. population of non-graduates of college (≥ 24 years of age): age, race/ethnicity, past year drug abuse/dependence, past year alcohol abuse/dependence, annual income below federal poverty level, and past year mental illness in years 2011–2017 of the National Survey on Drug Use and Health (N = 137,838). Rectangles (nodes) represent the entire population (top-most node) or population subgroups (all other nodes). Within each node the top line lists the smoking rate for that population and the bottom line the percent of the U.S. college-educated population represented within that node. Using the root node as an example, this node represents 100% of the U.S. non-institutionalized adults with less than a college education and 25.5% (rounded to 26%) of them are smokers. Lines below nodes represent the binary branching around particular risk factors and risk-factor levels into subgroup nodes with further potential partitioning based on additional risk factors/levels. The bottom row comprises terminal nodes or risk profiles (i.e., final partitioning for a particular subgroup, minimal terminal node size set to $\geq 1\%$ individuals). Terminal nodes contain the same information as the other nodes.

of college graduates were of White or Asian race/ethnicity and had annual incomes $\geq 200\%$ of the poverty threshold.

3.2. Smoking characteristics

Each of the smoking characteristics examined in this report differed significantly between college graduates and non-graduates (Table 2). A smaller percentage of college graduates reported current smoking compared to non-graduates, whereas more college graduates met criteria for former and never smoking status, especially the latter. Among those who reported current smoking, a smaller percentage of college graduates initiated any smoking or daily smoking prior to age 18, and a smaller percentage were current daily smokers. Further, college-graduate smokers, on average, reported smoking on fewer days over the past month and fewer cigarettes per day.

3.3. Classification and regression tree (CART) analyses

3.3.1. Relative strength of the risk factors

The relative strength of each of the seven risk factors within college graduates and non-graduates are shown in Table 3. The presence of a past year substance use disorder was the strongest risk factor for current smoking in both populations, with past year alcohol abuse/dependence being the strongest risk factor among college graduates and past year drug abuse/dependence the strongest among those who did not graduate college. Differences between the two populations become more discernible at the second and third ranked factors. For college graduates, two additional psychiatric conditions (past year drug abuse/dependence and mental illness, respectively) held second and third ranks as predictors of smoking risk. For non-graduates, the second and third ranks are held by two socio-demographic factors (i.e., race and age, respectively). Looking at the overall pattern of the relative rankings across the two populations suggests psychiatric conditions are relatively

Table 1
Sociodemographic and psychiatric characteristics of subpopulations.

	College graduates (n = 64,299)			Not college graduates (n = 137,838)		
	Adjusted %	95% CI		Adjusted %	95% CI	
Independent variables						
Age						
24–25	3.9%	3.7%	4.0%	4.1%	4.0%	4.2%
26–34	19.4%	18.9%	19.8%	16.9%	16.6%	17.2%
35–49	30.2%	29.8%	30.6%	27.5%	27.1%	27.9%
≥ 50	46.6%	45.9%	47.2%	51.5%	50.9%	52.1%
Sex (female)	51.7%	51.1%	52.3%	52.3%	52.0%	52.7%
Race/ethnicity						
White	73.8%	73.1%	74.4%	62.9%	62.4%	63.5%
Black	7.2%	6.9%	7.5%	13.5%	13.1%	13.9%
Latinx	7.4%	7.0%	7.7%	18.1%	17.7%	18.5%
Native American	0.2%	0.1%	0.2%	0.7%	0.6%	0.8%
Asian	10.2%	9.7%	10.7%	2.7%	2.6%	2.9%
Other	1.3%	1.2%	1.4%	2.1%	2.0%	2.2%
Poverty status						
Below threshold	3.4%	3.2%	3.6%	17.7%	17.3%	18.0%
100–199% of threshold	7.7%	7.4%	8.0%	25.1%	24.8%	25.5%
≥ 200% of threshold	88.9%	88.5%	89.3%	57.2%	56.8%	57.6%
Past-year mental illness	16.5%	16.1%	16.9%	18.9%	18.5%	19.3%
Past-year alcohol abuse/dependence	6.1%	5.8%	6.3%	5.7%	5.6%	5.9%
Past-year drug abuse/dependence	1.2%	1.1%	1.3%	2.4%	2.3%	2.5%

Note. Participants come from the National Survey on Drug Use and Health 2011–2017.

stronger contributors to smoking risk among college graduates and socio-demographics among non-graduates. These patterns are illustrated below in the CART models for college graduate and non-graduate populations.

3.3.2. CART model

As noted above, the top-most “parent” nodes in each CART model represent overall smoking-prevalence estimates (rounded) for the two populations. Beginning with college graduates (Fig. 1), the “parent” node indicates that the overall smoking prevalence was approximately 10%. The analysis resulted in a model with 13 terminal nodes, labeled from left to right across the bottom of Fig. 1. As expected based on the relative ranking of the risk factors, the first split into “child” nodes was based on the presence of past year alcohol abuse/dependence. Those with an alcohol use disorder were distributed into a “child” node to the right and downward of the “parent” node, and this “child” node represented a greater than 2-fold increase in smoking prevalence from 10% to 24%. Those without an alcohol use disorder were distributed to

Table 2
Group contrasts of smoking characteristics.

	College graduates (n = 64,299)			Not college graduates (n = 137,838)		
	Adjusted % or mean	95% CI		Adjusted % or mean	95% CI	
Current smoker	9.6%	9.2%	9.9%	25.5%	25.2%	25.8%
Former smoker	22.2%	21.6%	22.7%	24.2%	23.9%	24.6%
Never smoker	68.3%	67.7%	68.8%	50.3%	49.8%	50.7%

	College graduate smokers (n = 6913)			Not college graduate smokers (n = 41,497)		
	Adjusted % or mean	95% CI		Adjusted % or mean	95% CI	
Current daily smoker	53.1%	51.4%	54.8%	68.9%	68.2%	69.6%
Cigarette use prior to age 18	67.4%	65.8%	69.0%	75.5%	74.9%	76.2%
Daily cigarette use prior to age 18	26.7%	25.1%	28.2%	46.6%	45.8%	47.4%
Age at first cigarette	16.3	16.2	16.4	15.6	15.5	15.7
Age started smoking daily	20.0	19.8	20.2	18.2	18.1	18.3
Number of days smoked in past 30 days	21.5	21.1	21.8	24.6	24.4	24.7
Average cigarettes per day	8.6	8.2	8.9	11.7	11.6	11.9

Note. Participants come from the National Survey on Drug Use and Health 2011–2017.

Table 3
Risk factors ranked in order of importance from most to least predictive of current smoking.

Rank	College graduates	Not college graduates
1	Past-year alcohol abuse/dependence	Past-year drug abuse/dependence
2	Past-year drug abuse/dependence	Race/ethnicity
3	Past-year mental illness	Age
4	Poverty status	Poverty status
5	Age	Past-year alcohol abuse/dependence
6	Race/ethnicity	Sex
7	Sex	Past-year mental illness

Note. Participants come from the National Survey on Drug Use and Health 2011–2017.

the left and downward of the “parent” node, representing a slight decrease in prevalence from 10% to 9%.

The next critical split was based on the presence of past year drug abuse/dependence. Among college graduates without an alcohol use disorder, those endorsing a drug use disorder moved rightward and downward to a terminal node (i.e., terminal node 10), indicating that no further adjustments in risk were possible. Those without a drug use disorder were distributed into another “child” node wherein prevalence decreased slightly from 9% to 8%. Importantly, terminal node 10 (i.e., college graduates *without* alcohol use disorder, but *with* drug use disorder) represented the highest risk profile for college graduates at 37% smoking prevalence. The next split among the subgroup of college graduates with an alcohol use disorder was based on past year mental illness. Those with mental illness were distributed into a terminal node with a smoking prevalence of 31% (i.e., terminal node 13), whereas those without mental illness were distributed into a “child” node with a smoking prevalence of 20%. Among college graduates with an alcohol use disorder but not mental illness, the final division into the relatively higher-risk terminal nodes was based on age. Despite the presence of an alcohol use disorder, smoking prevalence was lower for college graduates ≥ 35 -years-old (i.e., terminal node 11; 17% prevalence), compared to college graduates between the ages of 24–34 (i.e., terminal node 12; 26% prevalence). Further divisions based on poverty status, age, race/ethnicity, and/or sex produced less dramatic differences in prevalence, and resulted in prevalence estimates ranging from 3%–14% (Fig. 1, terminal nodes 1–9). For non-graduates (Fig. 2), overall smoking prevalence was approximately 26%. The analysis resulted in a CART model with 16 terminal nodes, labeled from left to right across the bottom of Fig. 2. As expected, the first split was based on the presence of past year drug abuse/dependence. Those with a drug use disorder were distributed into a node to the right and downward of the “parent” node that represented a striking increase in smoking prevalence from 26% to 69%. Those without a drug use disorder were distributed to the left into a node representing a slight decrease in smoking prevalence from 26% to 24%.

The next critical split was based on race/ethnicity and pertained to the higher-risk subgroup of non-graduates with drug use disorders. Smoking prevalence was higher for non-graduates with drug use disorders who identified as White or Black (terminal node 16; 73% prevalence) compared to those who identified as Latinx, Native American, Asian, or Other (terminal node 15; 62% prevalence). The next split was based on age and pertained to the lower-risk subgroup of non-graduates without drug use disorders. Those of younger age were distributed to the right and downward into a node with a smoking prevalence of 31% where they faced additional splits into higher-risk nodes based on race/ethnicity, poverty, and alcohol abuse/dependence. The terminal node with the second highest prevalence (terminal node 14; 71% prevalence) was fairly idiosyncratic in that it characterized non-graduates with no drug use disorders between the ages of 24–49 who identified as White or Black, endorsed an annual income below 200% of the poverty threshold, and reported an alcohol use disorder. For all other non-graduates with no drug use disorders and between the ages of 24–49, prevalence estimates ranged from 14%–54% depending on race/ethnicity, poverty status, alcohol abuse/dependence, sex, and further age divisions (see terminal nodes 2–13). Finally, the terminal node representing the largest percentage of non-graduates was characterized by individuals with no drug use disorders who were ≥ 50 years (terminal node 1; 19% prevalence).

4. Discussion

The overarching purpose of this study was to increase our understanding of factors that put U.S. adults with a college education at risk for current smoking. As noted above, recent studies identified having a Bachelor's degree or higher as the single strongest protective factor against being a current smoker in the U.S. adult population (Gaalema

et al., 2018; Higgins et al., 2016). However, 76.9 million U.S. adults have a Bachelor's degree or higher (U.S. Census Bureau, 2018), and 10% of college graduates are current smokers (based on the estimate provided in the current study); thus, nearly 8 million people smoke despite the presence of a college education. This study demonstrated that both college graduates and non-graduates are sensitive to well-established psychiatric and socio-demographic risk factors for smoking, and that smoking prevalence among college graduates can rise to surprisingly high levels when multiple risk factors co-occur. This study also demonstrated broader patterns within each population that suggest differential sensitivity to psychiatric vs. socio-demographic risk factors. Important to underscore is that the observed patterns reflect the relative strength of each risk factor as it relates to smoking *within* each population. Therefore, when interpreting the present findings it is important to keep in mind that the influence of any one risk factor is conditional on the relative strength of the other risk factors within the same educational group. Overall, the present findings contribute to the overarching purpose of increasing understanding of smoking risk among college graduates, and at least four points warrant further discussion.

First, this study demonstrates that the presence of a substance use disorder – particularly a drug use disorder – is a robust predictor of smoking in both populations, underscoring that complications associated with substance use disorders are widespread and affect otherwise resilient populations such as college graduates. Along these lines, past year alcohol abuse/dependence emerged as the most important predictor of smoking among college graduates, although the association between alcohol use disorders and smoking risk was relatively weaker among non-graduates. It has been well established that problematic alcohol consumption (e.g., binge drinking) is prevalent among college students (Knight et al., 2002; Wechsler et al., 1994; Wechsler and Nelson, 2008; White and Hingson, 2013), and problematic alcohol consumption at an earlier age increases one's risk for developing an alcohol-use disorder (Grant, 1998). Considering this evidence along with the current findings that a higher percentage of college graduate smokers (compared to smokers with less than a Bachelor's degree) reported initiating any smoking and daily smoking at 18 years and older, it may be reasonable to assume that problematic alcohol consumption during the college years and beyond poses special risks among the college-educated. Importantly, although the prevalence of past-year alcohol abuse/dependence did not vary significantly by degree status in the present study, the relative strength of a past-year alcohol abuse/dependence differed by degree status within the context of the other risk factors. This finding warrants additional consideration for the broader patterns of how each risk factor associates with smoking status within each population.

A second notable result of the current study may help explain why past-year alcohol use disorders carried differential weight in predicting smoking for college graduates compared to non-graduates. That is, psychiatric conditions (i.e., drug abuse/dependence and mental illness) emerged as the 2nd and 3rd strongest risk factors for smoking among college graduates, whereas socio-demographic characteristics (i.e., race/ethnicity and age) emerged as the 2nd and 3rd strongest risk factor among non-graduates. Obtaining a college degree in the U.S. is associated with numerous, substantive, and wide-ranging socio-demographic and health benefits, factors that in combination likely contribute to substantial differences in risk for smoking. For example, evidence suggests that a college education can buffer against socio-demographically related stressors like racial discrimination (e.g., Allen et al., 2019; Howard and Sparks, 2015). Thus, one may speculate that a college education may help to reduce the impact of stress (and associated health consequences such as smoking) potentially stemming from experiences of racial discrimination. Turning to age, when one considers that educational attainment is associated with important age-related differences in smoking characteristics in the current study (e.g., non-graduates reported earlier age of smoking initiation and daily smoking), and more generally with multiple other life-course factors

such as greater earning potential and longevity (e.g., Cutler and Lleras-Muney, 2010; Cutler et al., 2015; Torpey, 2018), perhaps it is not surprising that within non-graduates the association of age with risk for adult smoking would be relatively stronger than it is for graduates. Put simply, a college education is associated with numerous socio-demographic outcomes that can serve to protect against smoking risk, thereby rendering socio-demographic factors relatively less important among those with a college education. With socio-demographic factors relatively less important for predicting smoking among college graduates, it is perhaps almost by default that psychiatric risk factors (particularly past-year alcohol use disorders and mental illness) would emerge to be more prominent predictors of smoking. That said, it is important not to overlook that alcohol use disorders and mental illness may develop and/or worsen during college which may also contribute to its relative ranking in the present results (Hunt and Eisenberg, 2010; Patrick et al., 2016). Of course, the present cross-sectional data limit our ability to parse out the temporal ordering of variables. However, the current results certainly underscore that psychiatric conditions need to be considered in any tobacco control and regulatory efforts to reduce smoking among college graduates.

Third, although the striking increases in smoking prevalence conditional on the presence of particular profiles of co-occurring risk factors is shared across the CART models for both populations, prevalence is shifted downward by orders of magnitude for college graduates compared to non-graduates. The present results suggest this overall difference may be largely accounted for by differences in smoking initiation, with the percentage of never smokers being 18 points higher among college graduates than non-graduates (68.3% vs. 50.3%, Table 2). There is compelling evidence that individuals who go on to obtain a college education are less likely to initiate smoking in adolescence compared to those who go on to lower levels of educational attainment, with differences in adolescent risk associated with family socioeconomic status, having a parent who ever smoked, peers who use substances, more ambitious future expectations (e.g., having plans to go to college), and school policies and experiences discouraging smoking (e.g., Maralani, 2013, 2014). The present results also reveal a small but significant difference in prevalence of former smokers among college graduates compared to non-graduates (24.2% vs. 22.2%, respectively; Table 2), suggesting that differences in successful quitting also contribute to differences in prevalence of current smokers. Those differences are consistent with college graduates initiating smoking at an older age and being lighter and non-daily smokers, which are characteristics that predict a greater likelihood of quitting smoking (Rose et al., 1996; Swayampakala et al., 2018).

Fourth, there are several limitations to the present study that merit mention. First, all of the relationships identified are correlational in nature and thus cannot support causal inferences. Second, the primary data source, NSDUH, is a cross-sectional survey, which precludes discerning the temporal ordering of observed relationships (e.g., although past year mental illness implies a recent diagnosis, it is not possible to determine whether the initial diagnosis precedes or follows becoming a current smoker). Third, all data including smoking status are based on participant self-report which increases the likelihood of underreporting of smoking status or other characteristics potentially associated with social stigma. Fourth, it is possible that the use of multiple tobacco products in addition to cigarettes may have influenced our findings; however, it was not possible to control for use of multiple tobacco products in the CART analyses used in the current study. Including use of multiple tobacco products as an additional independent variable was beyond the scope of the current study, the purpose of which was to examine the relative importance of the seven well-established psychiatric/socio-demographic smoking risk factors among college graduates and non-graduates. Future studies using CART modeling (or other machine learning techniques) may benefit from including a multiple tobacco product use variable to further examine differences in smoking risk between college graduates and non-graduates. These limitations

notwithstanding, the current study contributes new knowledge towards understanding the college-educated smoker. The results of this study can help inform future research in this domain, and may also help inform smoking prevention/intervention efforts (e.g., programs for college students struggling with substance use disorders and/or mental illness). Finally, the current results have potential implications for future research directions. For example, research is needed to better understand differences between individuals who go on to be college-graduate never-smokers vs. smokers in terms of psychiatric symptoms, knowledge/attitudes and risk perceptions about cigarette smoking and use of other substances, and how such differences vary between the two groups and also from their counterparts with lower educational attainment. Furthermore, future studies may benefit from examining how these relationships might affect responses to prevention and cessation interventions, and marketing strategies used to promote cigarette smoking and use of other substances.

Funding

This project was supported in part by Tobacco Centers of Regulatory Science (TCORS) Award [grant number U54DA036114] from the National Institute on Drug Abuse (NIDA) and Food and Drug Administration (FDA), Institutional Training Grant Award [grant number T32DA07242] from NIDA, and Centers of Biomedical Research Excellence [grant number P20GM103644] award from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Food and Drug Administration.

Declaration of competing interest

None.

References

- Alberg, A.J., Shopland, D.R., Cummings, K.M., 2014. The 2014 Surgeon General's report: commemorating the 50th Anniversary of the 1964 Report of the Advisory Committee to the US Surgeon General and updating the evidence on the health consequences of cigarette smoking. *Am. J. Epidemiol.* 179 (4), 403–412.
- Allen, A.M., Thomas, M.D., Michaels, E.K., Reeves, A.N., Okoye, U., Price, M.M., Hasson, R.E., Syme, S.L., Chae, D.H., 2019. Racial discrimination, educational attainment, and biological dysregulation among midlife African American women. *Psychoneuroendocrinology* 99, 225–235.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. Wadsworth, Belmont, CA.
- Center for Behavioral Health Statistics and Quality [CBHSQ], 2015. Behavioral Health Trends in the United States: Results From the 2014 National Survey on Drug Use and Health. NSDUH Series H-50 HHS Publication No. (SMA) 15–4927 Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Center for Behavioral Health Statistics and Quality [CBHSQ], 2018. 2017 National Survey on Drug Use and Health Final Analytic File Codebook. Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Cutler, D.M., Lleras-Muney, A., 2010. Understanding differences in health behaviors by education. *J. Health Econ.* 29 (1), 1–28.
- Cutler, D.M., Huang, W., Lleras-Muney, A., 2015. When does education matter? The protective effect of education for cohorts graduating in bad times. *Soc. Sci. Med.* 127, 63–73.
- Gaalema, D.E., Leventhal, A.M., Priest, J.S., Higgins, S.T., 2018. Understanding individual differences in vulnerability to cigarette smoking is enhanced by attention to the intersection of common risk factors. *Prev. Med.* 117, 38–42.
- Grant, B.F., 1998. The impact of a family history of alcoholism on the relationship between age at onset of alcohol use and DSM-IV alcohol dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *Alcohol Res. Health* 22 (2), 144–148.
- Higgins, S.T., 2014. Behavior change, health, and health disparities: an introduction. *Prev. Med.* 68, 1–4.
- Higgins, S.T., Kurti, A.N., Redner, R., et al., 2016. Co-occurring risk factors for current cigarette smoking in a US nationally representative sample. *Prev. Med.* 92, 110–117.
- Hiscock, R., Bauld, L., Amos, A., Fidler, J.A., Munaf, M., 2012. Socioeconomic status and smoking: a review. *Ann. N. Y. Acad. Sci.* 1248 (1), 107–123.
- Howard, J.T., Sparks, P.J., 2015. The role of education in explaining racial/ethnic allostatic load differentials in the United States. *Biodemogr. Soc. Biol.* 61 (1), 18–39.
- Hunt, J., Eisenberg, D., 2010. Mental health problems and help-seeking behavior among college students. *J. Adolesc. Health* 46 (1), 3–10.

- Kessler, R.C., Barker, P.R., Colpe, L.J., Epstein, J.F., Gfroerer, J.C., Hiripi, E., Howes, M.J., Normand, S.L., Manderscheid, R.W., Walters, E.E., Zaslavsky, A.M., 2003. Screening for serious mental illness in the general population. *Arch. Gen. Psychiatry* 60 (2), 184–189.
- Knight, J.R., Wechsler, H., Kuo, M., Seibring, M., Weitzman, E.R., Schuckitt, M.A., 2002. Alcohol abuse and dependence among US college students. *J. Stud. Alcohol* 63 (3), 263–270.
- 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.
- Levy, D.T., Chaloupka, F., Gitchell, J., 2004. The effects of tobacco control policies on smoking rates: a tobacco control scorecard. *J. Public Health Manag. Pract.* 10 (4), 338–353.
- Ma, J., Siegel, R.L., Jacobs, E.J., Jemal, A., 2018. Smoking-attributable mortality by state in 2014, US. *Am. J. Prev. Med.* 54 (5), 661–670.
- Maralani, V., 2013. Educational inequalities in smoking: the role of initiation versus quitting. *Soc. Sci. Med.* 84, 129–137.
- Maralani, V., 2014. Understanding the links between education and smoking. *Soc. Sci. Res.* 48, 20–34.
- Patrick, M.E., Schulenberg, J.E., O'Malley, P.M., 2016. High school substance use as a predictor of college attendance, completion, and dropout: a national multicohort longitudinal study. *Youth Soc.* 48 (3), 425–447.
- Rose, J.S., Chassin, L., Presson, C.C., Sherman, S.J., 1996. Prospective predictors of quit attempts and smoking cessation in young adults. *Health Psychol.* 15 (4), 261–268.
- Schroeder, S.A., Koh, H.K., 2014. Tobacco control 50 years after the 1964 surgeon general's report. *JAMA* 311 (2), 141–143.
- Substance Abuse and Mental Health Services Administration [SAMHSA], 2012. Results From the 2011 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-44, HHS Publication No. (SMA) 12–4713 Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Substance Abuse and Mental Health Services Administration [SAMHSA], 2013. Results From the 2012 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-46, HHS Publication No. (SMA) 13–4795 Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Substance Abuse and Mental Health Services Administration [SAMHSA], 2014. Results From the 2013 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-48 HHS Publication No. (SMA) 14–4863 Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Substance Abuse and Mental Health Services Administration [SAMHSA], 2016. Key Substance Use and Mental Health Indicators in the United States: Results From the 2015 National Survey on Drug Use and Health. NSDUH Series H-51 HHS Publication No. (SMA) 16–4984 Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Substance Abuse and Mental Health Services Administration [SAMHSA], 2017. Key Substance Use and Mental Health Indicators in the United States: Results From the 2016 National Survey on Drug Use and Health. NSDUH Series H-52 HHS Publication No. (SMA) 17–5044 Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Substance Abuse and Mental Health Services Administration [SAMHSA], 2018. Key Substance Use and Mental Health Indicators in the United States: Results From the 2017 National Survey on Drug Use and Health. NSDUH Series H-53 HHS Publication No. (SMA) 18–5068 Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Swayampakala, K., Thrasher, J.F., Hardin, J.W., Titus, A.R., Liu, J., Fong, G.T., Fleischer, N.L., 2018. Factors associated with changing cigarette consumption patterns among low-intensity smokers: longitudinal findings across four waves (2008–2012) of ITC Mexico Survey. *Addict. Behav. Rep.* 8, 154–163.
- Torpey, E., 2018. Measuring the value of education. In: *Career Outlook*. U.S. Bureau of Labor Statistics Available at: https://www.bls.gov/careeroutlook/2018/data-on-display/education-pays.htm?view_full, Accessed date: 28 July 2019.
- U.S. Census Bureau, 2018. Educational Attainment in the United States. Available at: <https://www.census.gov/data/tables/2018/demo/education-attainment/cps-detailed-tables.html>, Accessed date: 17 April 2019.
- Üstün, T.B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korf, M., Pull, C., 2010. Developing the World Health Organization disability assessment schedule 2.0. *Bull. World Health Organ.* 88, 815–823.
- Wechsler, H., Nelson, T.F., 2008. What we have learned from the Harvard School of Public Health College Alcohol Study: focusing attention on college student alcohol consumption and the environmental conditions that promote it. *J. Stud. Alcohol Drugs* 69 (4), 481–490.
- Wechsler, H., Davenport, A., Dowdall, G., Moeykens, B., Castillo, S., 1994. Health and behavioral consequences of binge drinking in college: a national survey of students at 140 campuses. *JAMA* 272 (21), 1672–1677.
- White, A., Hingson, R., 2013. The burden of alcohol use: excessive alcohol consumption and related consequences among college students. *Alcohol Res.* 35 (2), 201–2018.