5-2018

Principal component analysis of early alcohol, drug and tobacco use with major depressive disorder in US adults

KeSheng Wang

Ying Liu

Youssoufou Ouedraogo

Nianyang Wang

Xin Xie

See next page for additional authors

Recommended Citation

Wang, K., Liu, Y., Ouedraogo, Y., Wang, N., Xie, X., Xu, C., & Luo, X. (2018). Principal component analysis of early alcohol, drug and tobacco use with major depressive disorder in US adults. Journal of psychiatric research, 100, 113–120. https://doi.org/10.1016/j.jpsychires.2018.02.022

HHS Public Access

Author manuscript

J Psychiatr Res. Author manuscript; available in PMC 2019 January 11.

Published in final edited form as:

J Psychiatr Res. 2018 May ; 100: 113–120. doi:10.1016/j.jpsychires.2018.02.022.

Principal component analysis of early alcohol, drug and tobacco use with major depressive disorder in US adults

Kesheng Wang[a,b,*], **Ying Liu**[a], **Youssoufou Ouedraogo**[a], **Nianyang Wang**[c], **Xin Xie**[d], **Chun Xu**[e], and **Xingguang Luo**[b,f,*]

[a]Department of Biostatistics and Epidemiology, College of Public Health, East Tennessee State University, Johnson City, TN 37614, USA

[b]Center for Biological Psychiatry, Beijing HuiLongGuan Hospital, Peking University, Beijing 100096, China

[c]Division of Biostatistics, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

[d]Department of Economics and Finance, College of Business and Technology, East Tennessee State University, Johnson City, TN 37614, USA

[e]Department of Health and Biomedical Sciences, College of Health Affairs, University of Texas Rio Grande Valley, Brownsville, TX 78520, USA

[f]Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06516, USA

Abstract

Early alcohol, tobacco and drug use prior to 18 years old are comorbid and correlated. This study included 6,239 adults with major depressive disorder (MDD) in the past year and 72,010 controls from the combined data of 2013 and 2014 National Survey on Drug Use and Health (NSDUH). To deal with multicollinearity existing among 17 variables related to early alcohol, tobacco and drug use prior to 18 years old, we used principal component analysis (PCA) to infer PC scores and then use weighted multiple logistic regression analyses to estimate the associations of potential factors and PC scores with MDD. The odds ratios (ORs) with 95% confidence intervals (CIs) were estimated. The overall prevalence of MDD was 6.7%. The first four PCs could explain 57% of the total variance. Weighted multiple logistic regression showed that PC_1 (a measure of psychotherapeutic drugs and illicit drugs other than marijuana use), PC_2 (a measure of cocaine and hallucinogens), PC_3 (a measure of early alcohol, cigarettes, and marijuana use), and PC_4 (a measure of cigar, smokeless tobacco use and illicit drugs use) revealed significant associations with MDD (OR = 1.12, 95% CI = 1.08–1.16, OR = 1.08, 95% CI = 1.04–1.12, OR = 1.13, 95% CI = 1.07–1.18, and OR = 1.15, 95% CI = 1.09–1.21, respectively). In conclusion, PCA can be used to reduce the indicators in complex survey data. Early alcohol, tobacco and drug use prior to 18 years old were found to be associated with increased odds of adult MDD.

[*]Corresponding author: Department of Biostatistics and Epidemiology, College of Public Health, East Tennessee State University, PO Box 70259, Lamb Hall, Johnson City, TN 37614-1700, USA. Tel.: +1 423 439 4481; fax: +1 423 439 4606. wangk@etsu.edu (K.S. Wang) and Xingguang.luo@yale.edu (X. Luo).

Conflicts of interest:

All authors have reported no financial interests or potential conflicts of interest.

Keywords

Major depression disorder; Early use; Alcohol; Tobacco; Illicit drug; Principal component analysis

1. Introduction

Major depressive disorder (MDD), also known as depression, is one of the most common mental disorders in the United States (US). It significantly affects a person's family and personal relationships, work or school life, sleeping and eating habits, and general health. Based on the data from the National Health and Nutrition Examination Survey (NHANES) from 2009 to 2012, 7.6% of Americans ages 12 and over had depression (classified as moderate or severe depressive symptoms in the past 2 weeks) (Pratt and Brody, 2014). The National Institute of Mental Health has estimated that 16.1 million adults (6.7% of the US adult population) aged 18 years or older had reported at least one major depressive episode in the past year (NIMH, 2015); while the economic burden of depression in the US was estimated to be $210.5 billion in 2010 (Greenberg et al., 2015).

Previous studies have shown that several behavioral factors have been associated with depression including alcohol use and alcohol use disorder (Gilman and Abraham, 2001; Hasin and Grant, 2002; Sullivan et al., 2005; Boden and Fergusson, 2011; Kinyanda et al., 2011; Mugisha et al., 2015), and smoking and nicotine dependence (Glassman et al., 1990; Breslau et al., 1991; Kendler et al., 1993; Husky et al., 2008). Furthermore, MDD is highly comorbid with substance use disorders (SUDs) (Merikangas et al., 1998; Grant et al., 2016; Han et al., 2017) including alcohol and marijuana abuse/dependence (Greenbaum et al., 1991; Grant and Harford, 1995; Miguel-Hidalgo et al., 2010), opioid abuse (Brooner et al., 1997), and nicotine dependence (Cardenas et al., 2002; Chou et al., 2016; Huang et al., 2017; Martínez-Ortega et al., 2017).

Few studies have focused on the early alcohol, smoking and drug use on MDD. For example, tobacco use in late adolescence may predict MDD in the early 20s (Brook et al., 1998); while early onset of regular tobacco use is a predictor for lifetime drug use and depressive disorders (Hanna and Grant, 1999). Using a community-based longitudinal design, it has been shown that the cumulative frequency of drug use (alcohol, marijuana, and other illegal drugs) covering the period of childhood and early adolescence, middle and late adolescence increases the risk of later MDD, alcohol dependence, and SUDs; while marijuana use and the use of other illicit drugs increase the risk of later MDD, alcohol dependence, and SUDs (Brook et al., 2002). Another longitudinal data suggested that smoking increases the risk for MDD in women (Pasco et al., 2008). In addition, current smoking but not current smokeless tobacco use among adolescents and adults is associated with MDD using the NSDUH 2011 data (Redner et al., 2014). Another study showed that early-onset cannabis smokers had an increased risk of depression (Fairman and Anthony, 2012). Furthermore, adolescent alcohol use is positively associated with later depression in a population-based U.K. cohort (Edwards et al., 2014). Boden and Fergusson (2011) also concluded that increasing involvement with alcohol increases the risk of MDD. A recent study reported that the volume of drinking, age of first drink, binge drinking, and alcohol

dependence were associated with MDD using the 2010–2012 National Survey on Drug Use and Health (NSDUH) data (Jetelina et al., 2016; Wagner, 2016).

The NSDUH is a nationally representative survey that is conducted annually to assess the prevalence and correlations of drug use in the US for the population aged 12 years and older. The 2013 and 2014 NSDUH data provided 17 early alcohol, tobacco and drug use variables prior to 18 years; whereas these variables may be intercorrelated. Principal component analysis (PCA) can transform the set of correlated variables into a reduced number of uncorrelated variables known as principal components (PCs) or factors and then overcome the disturbance of the multicollinearity of the risk factors. PCA has been vastly used in social sciences, health service, and health sciences research (Lattin et al., 2002; Navarro et al. 2011; O'Rourke and Larry Hatcher, 2013; Jackson et al., 2015; Wang et al., 2017). Previous studies suggest that PCA is suitable for use with continuous and/or binary variables (Filmer and Pritchett, 2001; Howe and Hargreaves, 2008; Kolenikov and Angeles 2009; O'Rourke and Larry Hatcher, 2013; Jackson et al., 2015). To our updated knowledge, this is the first study to examine the factor structure of early alcohol, drug and tobacco use prior to 18 years of age using PCA and the associations of these factors with MDD using the NSDUH data.

2. Materials and Methods

2.1. Study population

Data were extracted from the 2013 and 2014 NSDUH data. The NSDUH is conducted by the Substance Abuse and Mental Health Services Administration of the United States Department of Health and Human Services. The NSDUH is a survey of the civilian non-institutionalized individuals (≥12 years old) in the U.S. to provide annual population estimates of substance use and health. By employing a multistage area probability sampling of the 50 states and the District of Columbia, we obtain a representative sample of all civilian non-institutionalized individuals who are ≥12 years old and response the survey items. The NSDUH measures the prevalence and correlations of drug use in the U.S. quarterly and annually. Information from the survey contains data on the use of illicit drugs (e.g., marijuana, cocaine, hallucinogens, heroin, and inhalants), alcohol and tobacco, and nonmedical use of prescription drugs (e.g., pain relievers, tranquilizers, and sedatives). The survey included questions from the Diagnostic and Statistical Manual (DSM) of Mental Disorders along with substance abuse treatment history and perceived need for treatment. Respondents in the age group of 12 −17 years were also asked of any illegal activities, neighborhood environment, and drug use by friends, etc. The background information of the respondents included gender, race/ethnicity, age, marital status, educational level, job and veteran statuses, and current household composition. Details of the survey design and data collection methods are published elsewhere (Center for Behavioral Health Statistics and Quality 2014, 2015). The total sample size of the combined 2013–2014 NSDUH data is 110,431. The current analysis was restricted to participants aged 18 years and older. There was an Institutional Review Board exemption due to secondary data analysis.

2.2. Measurements

2.2.1. Major depressive disorder—MDD in the NSDUH 2013–20414 data was determined based on the DSM-IV (American Psychiatric Association, 1994). Respondents were classified as experiencing MDD in the past year if they (a) reported meeting at least five of nine criteria within a two-week period at any point in their lifetime (i.e., depressed mood, diminished pleasure, weight gain/loss, fatigue, sleep changes, psychomotor distress, feelings of worthlessness, difficulty concentrating, or recurrent thoughts of suicide/death), one of which must have been either depressed mood or diminished pleasure, and (b) answered affirmatively to a question that queried whether they felt a depressed mood or diminished pleasure (in addition to other symptoms the respondent affirmed) for at least two weeks within the past year. A total of 6,239 adults with MDD and 72,010 controls were selected from the combined 2013 and 2014 NSDUH data.

2.2.2 Social-demographic factors—Gender was self-reported as male or female. Age was classified into four categories: 18–25 years, 26–49 years, 50–64 years and 65+ years. Race consisted of five subgroups: Whites, African American (AA), Asian, Hispanic, and Other. There were three categories of marriage status: married; widowed/divorced/separated, and never been married. The four categories of annual income were: <$20,000, $20,000-$49,999, $50,000-$74,999 and $75,000 or more. Health status had 4 levels: excellent, very good, good and fair/poor. Educational level was classified into < high School, high school, some college and college graduate. There were three levels of county type: large metro, small metro and non-metro according to the revised definitions of metropolitan statistical areas (MSAs) by the Office of Management and Budget (OMB) (Office of Management and Budget, 2003, Bulletin No. 03–04). Large metropolitan areas were defined to have a population of at least 1 million, small metropolitan areas have a population fewer than 1 million, and non-metropolitan areas fell outside MSAs.

2.2.3. Substance use disorders in the past year—Nicotine dependence was defined as the average score over 17 questions pertaining to five aspects of dependence which based on the dependence criteria according to the Nicotine Dependence Syndrome Scale (NDSS). This average score was only computed for those respondents who answered all 17 questions (Shiffman et al., 2004). Based on the NDSS score, a respondent was defined as having nicotine dependence if their average score was greater than or equal to 2.75. Illicit drug or alcohol abuse or dependence in the past year was defined by being dependent on any illicit drug or alcohol in the past year or illicit drug or alcohol abuse in the past year. Illicit drug abuse is defined as abusing any of the following substances: marijuana, hallucinogens, inhalants, tranquilizers, cocaine, heroin, pain relievers, stimulants, or sedatives. Illicit drug dependence was classified as being dependent on any of these following substances: marijuana, hallucinogens, inhalants, tranquilizers, cocaine, heroin, pain relievers, stimulants, or sedatives.

2.2.4. Early alcohol, drug and tobacco use prior to 18 years old—Seventeen variables about early alcohol, drug and tobacco use are listed in Table 2. These variables were defined by “first used the drug prior to age 18”, each of them had 2 levels (yes or no).

2.3. Statistical analysis

2.3.1. Descriptive statistics and prevalence—The PROC CROSSTAB in SAS-Callable SUDAAN 11 (Research Triangle Institute, USA), which uses Taylor Series Linearization to account for the weighting of the data and the complex survey sampling strategy, was used to weight and estimate population proportions of MDD. The overall prevalence and prevalence for potential factors were estimated. The Chi-square test was used to compare the prevalence of MDD across age groups, gender, race, and other factors.

2.3.2. Principal component analysis—Using ordinal and dichotomous indicators is a very common practice in social sciences and health sciences. Based on a previous suggestion (Muthén and Muthén, 2000), a polychoric correlation was created instead of Pearson's correlations for the categorical variable in PCA. A polychoric correlation and Pearson's correlation were calculated using PROC CORR for PCA; while PCA was performed with PROC FACTOR with SAS statistical software. An eigenvalue reflects the amount of variance captured by a given PC. The eigenvalue-one criterion (eigenvalue ≥ 1) is used to decide how many PCs to be retained (Kaiser, 1960, Stevens, 2002). A factor loading of one independent variable is considered as large if its absolute value exceeds 0.45.

2.3.3. Weighted logistic regression analysis—SAS PROC SURVEYLOGISTIC was used to estimate the odds ratios (ORs) and 95% confidence intervals (CIs) for the association between potential factors and MDD. Two models were conducted. In model one, univariate logistic regression was used to examine the role of each potential risk factor including PC scores in MDD; then multiple logistic regression was used to adjust for all potential risk factors including PC scores of MDD.

The PCA and logistic regression analyses were performed using SAS statistical software, version 9.4 (SAS Institute, Cary, NC, USA).

3. Results

3.1. Prevalence of major depression disorder

The prevalence of MDD in the past year is listed in Table 1. The overall prevalence of MDD was 6.7% (5.1% for males and 8.3% for females). The prevalence decreased with age (9.2%, 7.4%, 7.1% and 2.8% for age groups 18–25, 26–49, 50–64 and 65+, respectively). Whites (7.4%) had higher prevalence than AAs (5.0%), Asian (4.2%), and Hispanics (5.8%). The prevalence of MDD was higher in nicotine dependence and illicit drug or alcohol abuse/dependence than without these conditions (14.2% vs. 6.0% and 16.8% vs. 5.8%, respectively).

3.2. Weighted logistic regression analyses of early alcohol, tobacco and drug use

The results of univariate and multiple logistic regression analyses of early alcohol, tobacco and drug use are presented in Table 2. The univariate weighted logistic regression analysis showed that a total of 14 risk factors were positively associated with MDD ($p<0.05$). Alcohol use was negatively associated with MDD; whereas smokeless tobacco and cigarettes use did not show any association with MDD. However, after adjusted for social

demographic factors, nicotine dependence, illicit drug or alcohol abuse or dependence, as well as other early use factors (prior to age 18 years), only early cigar use, inhalants use, and marijuana use were associated with patients with MDD (OR = 1.22, 95% CI = 1.08–1.37; OR = 1.35, 95% CI = 1.11–1.65; and OR = 1.40, 95% CI = 1.27–1.55, respectively).

3.3. Principal component analysis of early alcohol, tobacco and drug use

There were 17 early alcohol, tobacco and drug use factors that had significantly positive correlations using both polychoric correlation and Pearson correlation ($p<0.0001$) (data not shown). The eigenvalues of the first 4 PCs were 5.2578, 1.7899, 1.3707 and 1.2289, respectively and the proportions of variation explained by these 4 PCs were 30.93%, 10.53%, 8.06% and 7.23%, respectively (Table 3). The first four PCs explained about 57% of total variation. The rotated factor patterns of the first 4 PCs are presented in Table 4. In terms of loading score, the PC_1 can be viewed as a measure of psychotherapeutic drugs and illicit drugs other than marijuana use (including pain relievers, tranquilizer, stimulant and any psychotherapeutics use). PC_2 can be viewed as a measure of cocaine and hallucinogens including lysergic acid diethylamide (LSD) and ecstasy. PC_3 is about a measure of early alcohol, cigarettes, marijuana. PC_4 is a measure of cigar, smokeless tobacco use and illicit drugs use.

3.4. Weighted logistic regression analyses of PC scores

The results of univariate and multiple logistic regression analyses of independent factors including the first 4 PCs are presented in Table 5. By using univariate analysis, all factors except for county type were associated with MDD ($p<0.05$). After adjusting for other factors, PC_1, PC_2, PC_3 and PC_4 revealed significant associations with MDD (OR = 1.12, 95% CI = 1.08–1.16, OR = 1.08, 95% CI = 1.04–1.12, OR = 1.13, 95% CI = 1.07–1.18, and OR = 1.15, 95% CI = 1.09–1.21, respectively).

4. Discussion

This original study was aimed to examine the factor structure of early alcohol, tobacco and drug use prior to 18 years old using PCA and the associations of these factors with MDD. The PCA analysis generated four PCs or factors which could explain 57% of total variances about 17 early alcohol, tobacco and drug use variables. Weighted multiple logistic regression showed that gender, racial groups, marital status, education, income, and general health were associated with MDD. After adjusting for other potential confounding factors, PC_1, PC_2, PC_3 and PC_4 for early alcohol, tobacco and drug use revealed significant associations with MDD.

The overall prevalence of MDD in the past year was 6.7% in U.S. adults, which is consistent with the results based on the 2009–2012 NHANES data (Pratt and Brody, 2014). This prevalence is slightly lower than the prevalence of 8.1% found by Kinyanda et al (2011); it worth noting that their study population is consisting specifically of HIV patients, which accounts for their higher value. The present results further added that the prevalence of MDD decreased with age (9.2%, 7.4%, 7.1% and 2.8% for age groups 18–25, 26–49, 50–64

and 65+, respectively), which are similar to the results that the 12-month MDD from adults aged 55 years and older were 4.1% (2.1% for 65–74 and 5.5% for 75+) (Aranda et al., 2012).

Being consistent with the findings in the previous studies, our results (Table 5) showed the adjusted OR of illicit drug or alcohol abuse or dependence with MDD was 2.32 (95% CI=2.05–2.63), suggesting that alcohol use disorder or illicit drug dependence was associated with depression (Greenbaum et al., 1991; Grant and Harford, 1995; Brooner et al., 1997; Merikangas et al., 1998; Gilman and Abraham, 2001; Hasin and Grant, 2002; Sullivan et al., 2005; Miguel-Hidalgo et al., 2010; Boden and Fergusson, 2011; Grant et al., 2016; Han et al., 2017). For example, 12-month DSM-V drug use disorder (DUD) was found to be significantly associated with MDD (OR=1.3; 95%CI=1.09–1.64) (Grant et al., 2016). Furthermore, our results showed the adjusted OR of nicotine dependence in the past year with MDD was 1.25 (95%CI=1.11–1.41), which is consistent with the relationship of smoking and nicotine dependence with depression (Glassman et al., 1990; Breslau et al., 1991; Kendler et al., 1993; Cardenas et al., 2002; Husky et al., 2008; Chou et al., 2016; Martínez-Ortega et al., 2017). Moreover, Chou et al (2016) found that the 12-month severe DSM-V nicotine use disorder was generally associated with major depressive, bipolar disorder I, bipolar disorder II, panic, generalized anxiety, posttraumatic stress, and schizotypal, borderline, and antisocial personality disorders (ORs = 1.3–2.5). In addition, it has been suggested that MDD may be a confounder for the relationship between alcohol use disorder and drinking volume (Baggio et al., 2015). In this current study, we focused on the effects of early alcohol, tobacco and drug use (prior to 18 years old) on adult MDD and we adjusted 12-month nicotine dependence and illicit drug or alcohol abuse or dependence as co-variates.

It is clear that smoking, alcohol and drug use are associated with mental problems such as MDD. It has been previously reported that alcohol and drug use, smoking and nicotine use were positively associated with mental health problems such as depression and anxiety (Boden and Fergusson, 2011; Emre et al., 2014; Brown et al., 2015; Fink et al., 2015; Steinberg et al., 2015). However, studies focused on the early alcohol, smoking and drug use on MDD are lacking. For example, previous studies reported that smoking and tobacco use, drug use in child or in adolescence increased the risks of later MDD (Brook et al., 1998; Hanna and Grant, 1999; Brook et al., 2002; Pasco et al., 2008; Fairman and Anthony, 2012; Edwards et al., 2014). Using the NSDUH 2011 data, Redner et al. (2014) found current smoking but not current smokeless tobacco use among adolescents and adults was associated with MDD. Recently, the volume of drinking, age at first drink, binge drinking, and alcohol dependence were associated with MDD using the 2010–2012 NSDUH data (Jetelina et al., 2016; Wagner, 2016). To the best of our knowledge, our present-day study is the first attempt to examine the factor structure of early alcohol, tobacco and drug use prior to 18 years old using PCA and the associations of these factors with the 12-month MDD.

Univariate logistic regression analysis showed that 14 early alcohol, tobacco and drug use factors were positively associated with MDD ($p<0.0001$), whereas alcohol use was negatively associated with MDD and smokeless tobacco and cigarettes use did not show any association with MDD (Table 2). However, after adjusted for demographic factors and other alcohol, tobacco and drug use, only early cigar use, inhalants use, and marijuana use were

associated with MDD (Table 2), which suggested that these factors may be correlated and may together influence the development of MDD. Then we performed correlation analysis and found that there were significantly positive correlations of 17 factors regarding early alcohol, tobacco and drug use using both polychoric- and Pearson- correlation ($p<0.0001$). To deal with multicollinearity of these early alcohol, tobacco and drug use variables, we conducted PCA and found the eigenvalues of the first 4 PCs were larger than unity and the first four PCs explained about 57% of total variation. PC_1 can be viewed as a measure of psychotherapeutic drugs and illicit drugs other than marijuana use (including pain relievers, tranquilizer, stimulant and any psychotherapeutics use). PC_2 can be viewed as a measure of cocaine and hallucinogens ((including LSD, also called “acid”, and ecstasy). PC_3 is about a measure of early alcohol, cigarettes, marijuana. PC_4 is a measure of cigar, smokeless tobacco use and illicit drugs use. On the one hand, our results also showed that early alcohol and drug use were comorbid; on the other hand, these factors work together on the development of adult MDD. However, early alcohol use was negatively associated with MDD in the univariate logistic regression analysis (Table 2); whereas, the effect is weaker (OR=0.88) compared with most other drug use factors (14 factors with ORs ranged from 1.40 to 2.65). Furthermore, PC3 was considered as a measure of early alcohol, cigarettes, marijuana; whereas the loading of alcohol use is low (0.48). In addition, illicit drug or alcohol abuse or dependence were positively associated with MDD (Table 5). Our results revealed that the effect of early alcohol use on MDD is weak; whereas the effect of alcohol use on MDD may depend on the alcohol use amount and intensity. Two specific studies on the use of marijuana or cannabis (Fairman and Anthony, 2012) and other illicit drugs (Han et al., 2010) showed results of higher prevalence of MDD for white females with college education, low income, and being single or widowed. The Fairman and Anthony (2012) focused more on the early onset cannabis use and also on the years of use categorized into less than 1 year, 1–10 years and more than 11 years; whereas the use of cannabis for less than 11 years did not show significant increase in OR in their study population.

There are several important strengths in this study. First, the NSDUH data is a nationally representative population-based survey, with large sample size, comprehensive information for state- and local-level data that have increased the validity of our study. In this study, we used the combined data from NSDUH years 2013–2014. We updated the national prevalence of MDD estimates using a nationally representative sample of U.S. adults; while a large sample size of subjects was widely selected at random with comprehensive information for the wide age range on drug uses characteristics, which allows us to adjust for numerous factors. Second, this is the first study to examine the factor structure of early alcohol, tobacco and drug use prior to 18 years old using PCA and found a 4 factor-solution for 17 variables. Third, we inferred PC scores (factor scores) and then used weighted logistic regression analyses to estimate the associations of potential factors and PC scores with MDD. The PCA is an effective method to reduce the dimensionality of multivariate data, which can transform the set of correlated variables into a reduced number of uncorrelated variables known as principal components (PCs). The PCs may then be used as predictors in subsequent analysis such as multiple regression (e.g., Lattin et al., 2002; O’Rourke and Larry Hatcher, 2013; Wang et al., 2017). Latent class analysis (LCA) is a statistical method for identifying unmeasured groups/latent classes of subjects by using categorical and/or

continuous observed variables (e.g., McCutcheon, 1987; Clogg, 1995) and within each latent class, the observed variables are statistically independent. The LCA is a subset of structural equation modeling (SEM), which invokes a measurement model that defines latent variables using one or more observed variables, and a structural model that imputes relationships between latent variables (e.g., Bollen, 1989; Kaplan, 2000; Yung, 2010; O'Rourke and Larry Hatcher, 2013). A number of previous studies suggested that it would be promising to combine LCA with PCA (Sotres-Alvarez et al., 2010; Costa et al. 2013); while another study found that both methods showed high agreement, but LCA is a complicated method while PCA is a simple method and seems to be an acceptable method of the analysis because of its simplicity (Sartipi et al. 2016). In addition, PCA has been used to generate PCs/factor scores, which then can be used for SEM (e.g., Voineskos et al. 2010; Yeh et al., 2010; Morris and Shakespeare-Finch, 2011; Ajayi et al. 2017). In addition, we adjusted for nicotine dependence and illicit drug or alcohol abuse or dependence in the past year as co-variates, which are comorbid with MDD.

Several limitations need to be acknowledged. First, a cross-sectional design cannot determine the causal relationships between these potential factors and MDD. However, we reported the results between early alcohol, tobacco and drug use prior to 18 years old and adult MDD, which may suggest the temporal relationship between risk factors and MDD. Second, the NSDUH data were collected by self-reported data on substance use and health conditions, making responses prone to social desirability bias and recall bias. However, previous study has shown that substance use reported by NSDUH respondents is generally valid (Harrison and Huges, 1997). Third, the information about the age at onset of MDD was unavailable in the NHDUH 2013–2014 data, thus we were not able to know whether the adult already had MDD prior to 18 years old or not.

In conclusion, using the latest combined 2013–2014 data from the NSDUH, we provided the prevalence of 12-month MDD and the correlates of early alcohol, tobacco and drug use. To the best of our knowledge, this is the first study to examine the factor structure of early alcohol, tobacco and drug use prior to 18 years old using PCA and the associations of these factors with MDD. Weighted multiple logistic regression showed that gender, racial groups, marital status, education, income, and general health were associated with MDD. After adjusting for other factors, the first 4 PCs for early alcohol, tobacco and drug use revealed significant associations with MDD. These findings highlighted the complex relationship among early alcohol, smoking and drug use factors and the complex effects of these factors on the MDD. The current findings also suggested that multivariate analysis such as PCA is one of the prospective methods to deal with correlated variables. In addition, these results highlight a need for interventions to simultaneously target alcohol, tobacco, and drug users.

Acknowledgements

The authors would like to thank Substance Abuse and Mental Health Services Administration, US for providing data from 2013 to 2014 National Survey on Drug Use and Health.

Funding sources:

No founding source is given for the present paper.

References

Ajayi OR, Matthews GB, Taylor M, Kvalsvig JD, Davidson L, Kauchali S, et al., 2017 Structural Equation Modeling of the Effects of Family, Preschool, and Stunting on the Cognitive Development of School Children. Front Nutr. 4:17. doi: 10.3389/fnut.2017.00017. [PubMed: 28555186]

American Psychiatric Association, 1994 Diagnostic and statistical manual of mental disorders: DSM-IV. Washington, DC: American Psychiatric Publishing, Inc.

Aranda MP, Chae DH, Lincoln KD, Taylor RJ, Woodward AT, Chatters L,M, 2012 Demographic correlates of DSM-IV major depressive disorder among older African Americans, Black Caribbeans, and non-Hispanic Whites: results from the National Survey of American Life. Int. J. Geriatr. Psychiatry 27, 940–7. [PubMed: 22038674]

Baggio S, Iglesias K, Studer J, Dupuis M, Daeppen J-B & Gmel G, 2015 Is the relationship between Major Depressive Disorder and self-reported Alcohol Use Disorder an artificial one? Alcohol and Alcohol., 50, 195–199.

Boden JM, Fergusson DM, 2011 Alcohol and depression. Addiction 106, 906–14. [PubMed: 21382111]

Bollen KA, 1989 Structural equations with latent variables. New York: John Wiley and Sons.

Breslau N, Kilbey M, Andreski P, 1991 Nicotine dependence, major depression, and anxiety in young adults. Arch. Gen. Psychiatry 48, 1069–74. [PubMed: 1845224]

Brook DW, Brook JS, Zhang C, Cohen P, Whiteman M, 2002 Drug use and the risk of major depressive disorder, alcohol dependence, and substance use disorders. Arch. Gen. Psychiatry 59, 1039–44. [PubMed: 12418937]

Brook JS, Cohen P, Brook DW, 1998 Longitudinal study of co-occurring psychiatric disorders and substance use. J. Am. Acad. Child Adolesc. Psychiatry 37, 322–30. [PubMed: 9519638]

Brooks Holliday S, Pedersen ER, Leventhal AM, 2016 Depression, posttraumatic stress, and alcohol misuse in young adult veterans: The transdiagnostic role of distress tolerance. Drug Alcohol Depend. 161, 348–55. [PubMed: 26948757]

Brooner RK, King VL, Kidorf M, Schmidt CW, Jr., Bigelow GE,1997 Psychiatric and substance use comorbidity among treatment-seeking opioid abusers. Arch. Gen. Psychiatry 54, 71–80. [PubMed: 9006403]

Brown CH, Medoff D, Dickerson FB, Fang LJ, Lucksted A, Goldberg RW, et al., 2015 Factors influencing implementation of smoking cessation treatment within community mental health centers. J. dual diagn. 11, 145–50. [PubMed: 25985201]

Cardenas L, Tremblay LK, Naranjo CA, Herrman N, Zack M, Busto UE, 2002 Brain reward system activity in major depression and comorbid nicotine dependence. J. Pharmacol. Exp. Ther 302, 1265–1271. [PubMed: 12183688]

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.

Center for Behavioral Health Statistics and Quality, 2015 Behavioral health trends in the United States: Results from the 2014 National Survey on Drug Use andHealth (HHS Publication No. SMA 15–4927, NSDUH Series.

Choi NG, Dinitto DM, Marti CN, 2015 Alcohol and other substance use, mental health treatment use, and perceived unmet treatment need: Comparison between baby boomers and older adults. Am. J. Addict 24, 299–307. [PubMed: 25923291]

Chou SP, Goldstein RB, Smith SM, Huang B, Ruan WJ, Zhang H, et al., 2016 The Epidemiology of DSM-5 Nicotine Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions-III. J. Clin. Psychiatry 77, 1404–1412. [PubMed: 27135834]

Clogg CC, 1995 Latent class models In Arminger G, Clogg CC, & Sobel ME (Eds.), Handbook of statistical modeling for the social and behavioral sciences (Ch. 6; pp. 311–359). New York: Plenum.

Costa PS, Santos NC, Cunha P, Palha JA, Sousa N, 2013The use of bayesian latent class cluster models to classify patterns of cognitive performance in healthy ageing. PLoS One.8(8):e71940. [PubMed: 23977183]

Edwards AC, Heron J, Dick DM, Hickman M, Lewis G, Macleod J, et al., 2014 Adolescent alcohol use is positively associated with later depression in a population-based U.K. cohort. J. Stud. Alcohol Drugs 75, 758–65. [PubMed: 25208193]

Emre N, Topal K, Bozkurt N, Topaktas E, 2014 Mental health screening and increased risk for anxiety and depression among treatment-seeking smokers. Tob. Induc. Dis 12(1), 20. [PubMed: 25745381]

Fairman BJ, Anthony JC, 2012 Are early-onset cannabis smokers at an increased risk of depression spells? J. Affect. Disord 138, 54–62. [PubMed: 22310034]

Filmer D, Pritchett LH, 2001 Estimating wealth effects without expenditure data–or tears: an application to educational enrollments in states of India. Demography 38, 115–32. [PubMed: 11227840]

Fink DS, Hu R, Cerdá M, Keyes KM, Marshall BD, Galea S, et al., 2015 Patterns of major depression and nonmedical use of prescription opioids in the United States. Drug Alcohol Depend.153, 258–64. [PubMed: 26026492]

Gilman SE, Abraham HD, 2001 A longitudinal study of the order of onset of alcohol dependence and major depression. Drug Alcohol Depend. 63, 277–86. [PubMed: 11418232]

Glassman AH, Helzer JE, Covey LS, Cottler LB, Stetner F, Tipp JE, et al., 1990 Smoking, smoking cessation, and major depression. JAMA 264, 1546–9. [PubMed: 2395194]

Grant BF, Harford TC, 1995 Comorbidity between DSM-IV alcohol use disorders and major depression: results of a national survey. Drug Alcohol Depend. 39, 197–206. [PubMed: 8556968]

Grant BF, Saha TD, Ruan WJ, Goldstein RB, Chou SP, Jung J, et al., 2016 Epidemiology of DSM-5 Drug Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions-III. JAMA Psychiatry 73, 39–47. [PubMed: 26580136]

Greenbaum PE, Prange ME, Friedman RM, Silver SE, 1991 Substance abuse prevalence and comorbidity with other psychiatric disorders among adolescents with severe emotional disturbances. J. Am. Acad. Child Adolesc. Psychiatry 30, 575–83. [PubMed: 1890091]

Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC, 2015 The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J. Clin. Psychiatry 76, 155–62. [PubMed: 25742202]

Han B, Gfroerer JC, Colliver JD, 2010 Associations between duration of illicit drug use and health conditions: results from the 2005–2007 national surveys on drug use and health. Ann. Epidemiol, 2010 20, 289–97. [PubMed: 20171900]

Han B, Olfson M, Mojtabai R, 2017 Depression care among adults with co-occurring major depressive episodes and substance use disorders in the United States. J. Psychiatr. Res 91, 47–56. [PubMed: 28314128]

Hanna EZ, Grant BF, 1999 Parallels to early onset alcohol use in the relationship of early onset smoking with drug use and DSM-IV drug and depressive disorders: findings from the National Longitudinal Epidemiologic Survey. Alcohol Clin. Exp. Res 23, 513–22. [PubMed: 10195827]

Harrison L, Huges A, eds., 1997 The validity of self-reported drug use: improving the accuracy of survey estimates. NIDA Research Monograph 167, NIH Publication No. 97–4147. Rockville (MD): National Institute on Drug Use.

Hasin DS, Grant BF, 2002 Major depression in 6050 former drinkers: association with past alcohol dependence. Arch. Gen. Psychiatry 59, 794–800. [PubMed: 12215078]

Howe LD, Hargreaves JR, 2008 Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. Emerg. Themes Epidemiol 5, 3. [PubMed: 18234082]

Huang J, Wen G, Yang W, Yao Z, Wu C, Ye X, 2017 The association between second-hand smoke exposure and depressive symptoms among pregnant women. Psychiatry Res. 256, 469–474. [PubMed: 28711818]

Husky MM, Mazure CM, Paliwal P, McKee SA, 2008 Gender differences in the comorbidity of smoking behavior and major depression. Drug Alcohol Depend. 93, 176–179. [PubMed: 17850991]

Jackson EF, Siddiqui A, Gutierrez H, Kanté AM, Austin J, Philipps JF, 2015 Estimation of indices of health service readiness with a principal component analysis of the Tanzania Service Provision Assessment Survey. BMC Health Serv. Res 15, 536. [PubMed: 26634449]

Jetelina KK, Reingle Gonzalez JM, Vaeth PA, Mills BA, Caetano R, 2016 An Investigation of the Relationship Between Alcohol Use and Major Depressive Disorder Across Hispanic National Groups. Alcohol Clin. Exp. Res 40, 536–42. [PubMed: 26887675]

Kaiser HF, 1960 The application of electronic computers to factor analysis. Educ. Psycho. Meas 20, 141–51.

Kaplan D, 2000 Structural Equation Modeling. Foundations and Extensions. Thousand Oaks London New Delhi: Sage Publications.

Kendler KS, Neale MC, MacLean CJ, Heath AC, Eaves LJ, Kessler RC, 1993 Smoking and major depression. A causal analysis. Arch. Gen. Psychiatry 50, 36–43. [PubMed: 8422220]

Kinyanda E, Hoskins S, Nakku J, Nawaz S, Patel V, 2011 Prevalence and risk factors of major depressive disorder in HIV/AIDS as seen in semi-urban Entebbe district, Uganda. BMC Psychiatry. 11, 205. [PubMed: 22208452]

Kolenikov S, Angeles G, 2009 Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? Rev. Income Wealth. 55, 128–65.

Lattin J, Carroll D, Green P, 2002 Analyzing Multivariate Data, 1st Edition, Cengage Learning.

Major Depression Among Adults. National Institute of Mental Health https://www.nimh.nih.gov/health/statistics/prevalence/major-depression-among-adults.shtml. Accessed August 15, 2017.

Martínez-Ortega JM, Franco S, Rodríguez-Fernández JM, Gutiérrez-Rojas L, Wang S, Gurpegui M, 2017 Temporal sequencing of nicotine dependence and major depressive disorder: A U.S. national study. Psychiatry Res. 250, 264–269. [PubMed: 28183022]

McCutcheon AC, 1987 Latent class analysis. Beverly Hills: Sage Publications.

Merikangas KR, Mehta RL, Molnar BE, Walters EE, Swendsen JD, Aguilar-Gaziola S, et al., 1998 Comorbidity of substance use disorders with mood and anxiety disorders: results of the International Consortium in Psychiatric Epidemiology. Addict. Behav 23, 893–907. [PubMed: 9801724]

Miguel-Hidalgo JJ, Waltzer R, Whittom AA, Austin MC, Rajkowska G, Stockmeier CA, 2010 Glial and glutamatergic markers in depression, alcoholism, and their comorbidity. J. Affect. Disord 127, 230–240. [PubMed: 20580095]

Morris BA, Shakespeare-Finch J, 2011 Rumination, post-traumatic growth, and distress: structural equation modelling with cancer survivors. Psychooncology. 20(11):1176–83. [PubMed: 20731009]

Mugisha J, Muyinda H, Malamba S, Kinyanda E, 2015 Major depressive disorder seven years after the conflict in northern Uganda: burden, risk factors and impact on outcomes (The Wayo-Nero Study). BMC Psychiatry. 15, 48. [PubMed: 25881310]

Muthén B, Muthén LK, 2000 Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. Alcohol Clin. Exp. Res 24, 882–91. [PubMed: 10888079]

Navarro Silvera SA, Mayne ST, Risch HA, Gammon MD, Vaughan T, Chow WH, et al., 2011 Principal component analysis of dietary and lifestyle patterns in relation to risk of subtypes of esophageal and gastric cancer. Ann. Epidemiol 21, 543–50. [PubMed: 21435900]

O'Rourke N, Larry Hatcher L, 2013 A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling, Second Edition SAS Institute.

Pasco JA, Williams LJ, Jacka FN, Ng F, Henry MJ, Nicholson GC, et al., 2008 Tobacco smoking as a risk factor for major depressive disorder: population-based study. Br. J. Psychiatry 193, 322–6. [PubMed: 18827296]

Pratt LA, Brody DJ, 2014 Depression in the U.S. Household Population, 2009–2012. NCHS Data Brief No. 172, December 2014. https://www.cdc.gov/nchs/data/databriefs/db172.pdf

Redner R, White TJ, Harder VS, Higgins ST, 2014 Examining vulnerability to smokeless tobacco use among adolescents and adults meeting diagnostic criteria for major depressive disorder. Exp. Clin. Psychopharmacol 22, 316–22. [PubMed: 24978349]

Sartipi M, Nedjat S, Mansournia MA, Baigi V, Fotouhi A, 2016 Assets as a Socioeconomic Status Index: Categorical Principal Components Analysis vs. Latent Class Analysis. Arch Iran Med. 19(11):791–796. [PubMed: 27845549]

Shiffman S, Waters AJ, Hickcox M, 2004 The Nicotine Dependence Syndrome Scale: A multi-dimensional measure of nicotine dependence. Nicotine Tob. Res 6, 327–348. [PubMed: 15203807]

Sotres-Alvarez D, Herring AH, Siega-Riz AM, 2010 Latent class analysis is useful to classify pregnant women into dietary patterns. J Nutr. 140(12):2253–9. [PubMed: 20962151]

Steinberg ML, Williams JM, Li Y, 2015 Poor Mental Health and Reduced Decline in Smoking Prevalence. Am. J. Prev. Med 49, 362–9. [PubMed: 26071864]

Stevens J, 2002 Applied multivariate statistics for the social sciences. Hillsdale, NJ: L. Erlbaum.

Sullivan LE, Fiellin DA, O'Connor PG 2005 The prevalence and impact of alcohol problems in major depression: a systematic review. Am. J. Med 118, 330–41. [PubMed: 15808128]

Voineskos AN, Rajji TK, Lobaugh NJ, Miranda D, Shenton ME, Kennedy JL, et al., 2012 Age-related decline in white matter tract integrity and cognitive performance: a DTI tractography and structural equation modeling study. Neurobiol Aging. 33(1), 21–34. [PubMed: 20363050]

Wagner EF, 2016 Commentary on "An Investigation of the Relationship Between Alcohol Use and Major Depressive Disorder Across Hispanic National Groups" by Jetelina and Colleagues. Alcohol Clin. Exp. Res 40, 1226–7. [PubMed: 27122471]

Wang KS, Liu Y, Xie X, Gong S, Xu C, Sha Z, 2017 Principal component regression analysis of nutrition factors and physical activity with diabetes. Journal of Biometrics and Biostatistics 8(4): 364.

Yeh PH, Zhu H, Nicoletti MA, Hatch JP, Brambilla P, Soares JC,2010Structural equation modeling and principal component analysis of graymatter volumes in major depressive and bipolar disorders: differencesin latent volumetric structure. Psychiatry Res. 184(3),177–85. [PubMed: 21051206]

Yung YF,2010 Introduction to Structural Equation Modeling Using the CALIS Procedure in SAS/STAT® Software Computer technology workshop presented at the Joint Statistical Meeting on August 4, 2010, Vancouver, Canada.

Table 1.

Prevalence of major depressive disorder (%)

Variable	Total (N)	MDD	Prevalence (%)	95%CI	p-value
Gender					
Male	36213	2058	5.1	4.7–5.4	<0.0001
Female	42036	4181	8.3	7.9–8.7	
Age group					
18–25 years	30841	2844	9.2	8.8–9.7	<0.0001
26–49 years	32269	2531	7.4	7.0–7.9	
50–64 years	9202	689	7.1	6.5–7.6	
65+ years	5937	175	2.8	2.3–3.4	
Race					
White	44238	4219	7.4	7.1–7.7	<0.0001
AA	9640	579	5.0	4.4–5.6	
Asian	3222	194	4.2	3.2–5.2	
Hispanic	12728	846	5.8	5.2–6.6	
Other	4008	401	9.3	8.0–10.6	
Marital status					
Married	30405	1599	4.5	4.2–4.9	<0.0001
Widowed/Divorced/Separated	10399	1115	9.1	8.3–9.9	
Never Married	37445	3525	9.2	8.7–9.7	
Income					
<$20,000	18717	2014	10.2	9.6–10.9	<0.0001
20,000–49,999	25373	2031	6.8	6.4–7.3	
50,000–74,999	12512	878	6.4	5.8–7.0	
75,000+	21647	1316	4.9	4.6–5.3	
Education					
Less than high school	11119	879	6.7	5.9–7.4	<0.0001
High school graduate	24102	1854	6.5	5.9–6.9	
Some college	23143	2126	7.9	7.4–8.4	
College graduate	19885	1380	5.6	5.5–6.5	
County type					
Non-metro	16542	1348	7.3	6.7–8.01	0.1052
Small metro	26972	2205	6.7	6.3–7.2	
Large metro	32049	2686	6.6	6.2–6.9	
Health					
Excellent	18798	825	3.7	3.3–4.1	<0.0001
Very good	30119	1995	5.1	4.7–5.5	
Good	20988	1975	7.3	6.7–7.9	
Fair/poor	7501	1228	13.6	12.5–14.8	
Nicotine dependence					
No	70464	5093	6.0	5.8–6.3	<0.0001

Variable	Total (N)	MDD	Prevalence (%)	95%CI	p-value
Yes	7785	1146	14.2	13.1–15.3	
Illicit drug or Alcohol abuse or dependence					
No	69175	4675	5.8	5.5–6.1	<0.0001
Yes	9074	1564	16.8	15.5–18.1	
Overall	78249	6239	6.7	6.5–7.0	

Abbreviations: MDD = Major depressive disorder; AA=African American; CI= Confidence interval; p-value is based on χ^2 test.

Table 2.

Univariate and multiple logistic regression analyses for the relationship between early use factors (prior to age 18yrs) and major depressive disorder.

Variable	Crude OR	95%CI	p-value	Adjusted OR	95%CI	p-value
Cigars use						
Yes	1.51	1.35–1.67	<0.0001	1.22	1.08–1.37	0.0015
Smokeless tobacco						
Yes	1.12	0.96–1.31	0.1553	0.87	0.73–1.04	0.1160
Cocaine use						
Yes	2.49	2.15–2.91	<0.0001	1.19	0.96–1.48	0.1180
Hallucinogens use						
Yes	2.35	2.06–2.69	<0.0001	1.20	0.90–1.60	0.2142
Lysergic acid diethylamide use						
Yes	2.25	1.90–2.66	<0.0001	0.90	0.68–1.18	0.4368
Ecstasy use						
Yes	2.41	2.01–2.88	<0.0001	0.77	0.58–1.03	0.0761
Inhalants use						
Yes	2.59	2.19–3.07	<0.0001	1.35	1.11–1.65	0.0030
Pain relievers use						
Yes	2.51	2.17–2.89	<0.0001	0.87	0.66–1.16	0.3425
Tranquilizers use						
Yes	2.56	2.19–2.99	<0.0001	0.84	0.64–1.11	0.226
Stimulants use						
Yes	2.72	2.28–3.24	<0.0001	0.96	0.71–1.28	0.759
Psychotherapeutics use						
Yes	2.65	2.36–2.98	<0.0001	1.43	0.98–2.09	0.0639
Illicit drugs use						
Yes	2.21	2.01–2.44	<0.0001	1.13	0.98–1.30	0.106
Illicit drugs other than marijuana						
Yes	2.50	2.26–2.77	<0.0001	1.07	0.85–1.35	0.578
Cigarettes use						
Yes	1.07	0.98–1.17	0.1124	0.97	0.88–1.07	0.5164
Cigarettes daily use						
Yes	1.40	1.28–1.53	<0.0001	1.01	0.90–1.12	0.9844
Alcohol use						
Yes	0.88	0.80–0.97	0.0077	1.01	0.90–1.12	0.9245
Marijuana use						
Yes	1.70	1.57–1.83	<0.0001	1.40	1.27–1.55	<0.0001

Abbreviations: OR= Odds ratio; CI= Confidence interval.

Table 3.

Eigenvalues of first 7 principal components.

PC	Eigenvalue	Difference	Proportion	Cumulative
1	5.2578	3.4679	0.3093	0.3093
2	1.7899	0.4192	0.1053	0.4146
3	1.3707	0.1418	0.0806	0.4952
4	1.2289	0.2966	0.0723	0.5675
5	0.9323	0.0778	0.0548	0.6223
6	0.8545	0.0466	0.0503	0.6726
7	0.8078		0.0475	0.7201

PC: Principal component.

Table 4.

Rotated factor pattern of 4 principal components.

Variable	PC_1	PC_2	PC_3	PC_4
Cigars use	14	7	4	70*
Smokeless tobacco	4	7	5	72*
Cocaine use	27	58*	−5	12
Hallucinogens use	20	88*	4	11
Lysergic acid diethylamide use	7	83*	1	6
Ecstasy use	24	60*	−1	0
Inhalants use	22	42	0	31
Pain relievers use	85*	10	0	13
Tranquilizers use	69*	22	−3	5
Stimulants use	56*	30	−2	7
Psychotherapeutics use	93*	19	2	13
Illicit drugs use	36	30	32	49*
Other illicit drugs use	62*	55*	8	27
Cigarettes use	−2	−6	75*	−11
Cigarettes daily use	4	16	62*	13
Alcohol use	−6	−16	48*	−42
Marijuana use	0	−4	69*	20

PC: Principal component.

*A factor loading of one independent variable is considered as large if its absolute value exceeds 0.45.

Table 5.

Univariate and multiple logistic regression analyses of potential factors and principal components.

Variable	Crude OR	95% CI	p-value	Adjusted OR	95% CI	p-value
Gender (ref = Male)						
Female	1.69	1.53–1.85	<0.0001	2.02	1.82–2.24	<0.0001
Age group (ref = 65+ years)						
18–25 years	3.87	3.07–4.88	<0.0001	3.84	3.01–4.89	<0.0001
26–49 years	2.97	2.35–3.76	<0.0001	3.32	2.61–4.21	<0.0001
50–64 years	2.68	2.07–3.48	<0.0001	2.80	2.15–3.65	<0.0001
Race (ref = Whites)						
AA	0.67	0.59–0.77	<0.0001	0.52	0.45–0.60	<0.0001
Asian	0.53	0.42–0.67	<0.0001	0.61	0.47–0.79	0.0002
Hispanic	0.79	0.69–0.90	0.0003	0.70	0.60–0.82	<0.0001
Other	1.25	1.06–1.48	0.0072	0.91	0.75–1.09	0.2987
Marital status (ref = Married)						
Widowed/Divorced/Separated	2.04	1.80–2.33	<0.0001	1.67	1.44–1.93	<0.0001
Never	2.19	1.99–2.42	<0.0001	1.50	1.34–1.69	<0.0001
Education (ref = less than high school)						
High school graduate	0.97	0.84–1.13	0.7122	1.11	0.95–1.29	0.1779
Some college	1.23	1.05–1.42	0.0082	1.45	1.22–1.71	<0.0001
College graduate	0.92	0.79–1.07	0.2836	1.62	1.35–1.93	<0.0001
Income (ref = > 75,000)						
<20,000	2.05	1.84–2.29	<0.0001	1.40	1.22–1.62	<0.0001
20,000–49,999	1.38	1.24–1.53	<0.0001	1.16	1.02–1.32	0.0272
50,000–74,999	1.32	1.16–1.49	<0.0001	1.22	1.06–1.39	0.0052
County (ref = Large metro)						
Small metro	0.99	0.91–1.09	0.9508	0.92	0.83–1.02	0.0985
Non-metro	1.11	0.99–1.25	0.0585	0.99	0.88–1.12	0.8700
Health (ref = Excellent)						
Very good	1.40	1.22–1.61	<0.0001	1.34	1.17–1.54	<0.0001
Good	2.04	1.77–2.35	<0.0001	2.11	1.82–2.46	<0.0001
fair/poor	4.09	3.46–4.83	<0.0001	4.86	4.09–5.79	<0.0001
Nicotine dependence (ref = No)						
Yes	2.53	2.27–2.81	<0.0001	1.25	1.11–1.41	0.0003
Illicit drug or alcohol use (ref = No)						
Yes	3.40	3.04–3.80	<0.0001	2.32	2.05–2.63	<0.0001
PC1	1.27	1.23–1.32	<0.0001	1.12	1.08–1.16	<0.0001
PC2	1.22	1.18–1.26	<0.0001	1.08	1.04–1.12	<0.0001
PC3	1.22	1.17–1.28	<0.0001	1.13	1.07–1.18	<0.0001
PC4	1.17	1.12–1.22	<0.0001	1.15	1.09–1.21	<0.0001

Abbreviations: AA = African American; PC: Principal component; OR = Odds ratio; CI = Confidence interval.