Published in final edited form as:

Perspect Public Health. 2014 November ; 134(6): 331–338. doi:10.1177/1757913913495780.

Clustering of modifiable biobehavioral risk factors for chronic disease in US adults: a latent class analysis

Adam M Leventhal,

Departments of Preventive Medicine and Psychology, Keck School of Medicine, University of Southern California, 2250 Alcazar Street, CSC 240, Los Angeles, CA 90033, USA

Jimi Huh, and Genevieve F Dunton

Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

Adam M Leventhal: adam.leventhal@usc.edu

Abstract

Aims—Examining the co-occurrence patterns of modifiable biobehavioral risk factors for deadly chronic diseases (e.g. cancer, cardiovascular disease, diabetes) can elucidate the etiology of risk factors and guide disease-prevention programming. The aims of this study were to (1) identify latent classes based on the clustering of five key biobehavioral risk factors among US adults who reported at least one risk factor and (2) explore the demographic correlates of the identified latent classes.

Methods—Participants were respondents of the National Epidemiologic Survey of Alcohol and Related Conditions (2004–2005) with at least one of the following disease risk factors in the past year ($N = 22,789$), which were also the latent class indicators: (1) alcohol abuse/dependence, (2) drug abuse/dependence, (3) nicotine dependence, (4) obesity, and (5) physical inactivity. Housing sample units were selected to match the US National Census in location and demographic characteristics, with young adults oversampled. Participants were administered surveys by trained interviewers.

Results—Five latent classes were yielded: ‘obese, active non-substance abusers’ (23%); ‘nicotine-dependent, active, and non-obese’ (19%); ‘active, non-obese alcohol abusers’ (6%); ‘inactive, non-substance abusers’ (50%); and ‘active, polysubstance abusers’ (3.7%). Four classes were characterized by a 100% likelihood of having one risk factor coupled with a low or moderate likelihood of having the other four risk factors. The five classes exhibited unique demographic profiles.

Conclusions—Risk factors may cluster together in a non-monotonic fashion, with the majority of the at-risk population of US adults expected to have a high likelihood of endorsing only one of these five risk factors.

Copyright © Royal Society for Public Health 2013

Correspondence to: Adam M Leventhal, adam.leventhal@usc.edu.

DECLARATION OF CONFLICTING INTERESTS

The authors report no conflicts of interest related to the submission of this manuscript.

Keywords

multiple risk factors; obesity; physical activity; substance-use disorders; latent class analysis

INTRODUCTION

A promising method for studying the etiology of chronic diseases that plague the population, such as cardiovascular disease, cancer, and diabetes, involves investigating the modifiable biobehavioral risk factors that give rise to these diseases. Of the various modifiable biobehavioral risk factors for chronic disease, harmful alcohol and drug use, habitual cigarette smoking, obesity, and physical inactivity are important to study because they are highly prevalent in the US population.^{1–5} Furthermore, these particular risks are clear causes of numerous negative health outcomes, account for considerable economic and social burden, and are modifiable via intervention.^{6,7} These risk factors are also notable because they tend to co-occur in some individuals,^{8–19} and the clustering of these risk factors may have synergistic effects that disproportionately increase the likelihood of negative health outcomes.²⁰

Examining the co-occurrence patterns of risk factors is important as such information could shed light on their etiology. For instance, if two risk factors cluster together with each other more than the other factors, this may suggest that they could be influenced by a common source. Furthermore, if certain cluster configurations are more common in certain subgroups (e.g. lower income bracket), characteristics common in that subgroup (e.g. reduced access to quality health care) may perhaps give rise to the co-occurrence of those risk factors.

Information on clustering may also guide designing disease-prevention programs. Intervention approaches that target multiple modifiable risk factors in a single program are gaining favor because they are potentially more efficient, cost-effective, and may have a greater public health impact than single-risk-factor approaches.^{21–23} Thus, data on the prevalence of subgroups with particular configurations of various risk factors and types of individuals represented in each of these subgroups would inform practitioners and policymakers regarding (1) the overall need for multi-risk factor interventions to prevent chronic disease, (2) the particular risk factors that should be grouped together in such interventions, and (3) the segments of the population that may benefit most from multi-factor intervention approaches.

These modifiable biobehavioral risk factors for chronic disease are unlikely to occur in an entirely uniform manner in population. That is, it is doubtful whether people can be accurately characterized into two groups – those that lead healthy versus those that lead unhealthy lifestyles.⁸ For instance, certain individuals who are physically active tend to use alcohol *more* frequently than their inactive counterparts (e.g. young adults).²⁴ Also, tobacco use is associated with *lower* rates of obesity among some individuals (e.g. older adults).²⁵ Thus, alcohol-, drug-, and tobacco-use disorders; physical inactivity; and obesity do not appear to co-occur in a monotonic fashion;⁸ rather, there are qualitative differences in profiles of heterogeneity, which may pose problems for designing disease-prevention programs. Indeed, a common challenge raised by public health professionals is that program

development for multiple risk factor interventions is difficult due to the ‘infinite combinations of risk factors with which patients may present’.²⁶

Latent class analysis (LCA) is a useful method for identifying homogeneous subgroups within a heterogeneous population with regard to the manifestations of a set of categorical characteristics.^{27–29} Thus, instead of subgrouping every possible profile, LCA helps to reduce the data into the most parsimonious set of classes while accounting for measurement error. Despite the fact that there are a number of possible risk factors for chronic disease (e.g. family histories of cancer, heart disease, stroke, blood pressure, or blood cholesterol levels), focusing on modifiable factors that can be addressed via non-biologic treatment is important for public health.

There have only been a few prior studies that have applied LCA to characterize the clustering of modifiable behavioral risk factors for disease,^{10–13} which leaves several important gaps in the literature. First, to our knowledge, no study has used LCA to specifically investigate the co-occurrence of alcohol-, drug-, and tobacco-use disorder; physical inactivity; and obesity. Nonetheless, these specific five biobehavioral risk factors are critically important to study as a unique set of indicators using LCA because they tend to co-occur in a non-uniform fashion and are among the most common targets in behavioral interventions.^{1–8,24,30} Thus, subjecting them to LCA would be of great value for advancing the literature. Second, prior LCA studies of substance use have often relied on non-diagnostic indicators of substance use, such as use status, frequency, or duration.^{10,11,13} However, clinically significant manifestations of substance use are most aptly identified by substance-use disorder diagnoses,³¹ which may be important predictors of disease, disability, and mortality risk and may be independent of one’s frequency or duration of substance use,^{32–35} for which the available studies using LCA with other biobehavioral risk factors are currently lacking. Third, there has been little prior research applying LCA to population-based samples, with the majority of prior work focusing on clinic-based samples or restricted community-based samples. By examining clustering of biobehavioral risk factors for chronic disease in a national population-based sample, the resultant findings may be relevant for guiding health promotion programs at the national community level.

This study identified latent classes based on the clustering of five key modifiable biobehavioral risk factors for chronic disease in a representative sample of US adults. We aimed to clarify (1) number of classes that could parsimoniously describe patterns of risk factor co-occurrence, (2) configurations of risk factor clustering represented in each class, and (3) the expected prevalence of each class. Once we established latent classes, we explored the demographic correlates of class membership to describe the makeup of each class, with the idea that this information could inform how prevention efforts are to be directed toward demographic cross-sections of the population with particular risk profiles. Given the potential for non-uniformity in patterns of co-occurrence, we did not hypothesize a particular number of classes or the risk factor configuration that would constitute each class. Rather, as in prior LCA work, we utilized an empirical data-driven approach for LCA.

METHODS

Sample and procedure

Participants were respondents in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC).³⁶ The NESARC was selected because, to the best of our knowledge, the NESARC represents the most recent nationally representative study that yields clinical diagnoses of substance-use disorder as well as physical activity. The US Census Bureau and the US Office of Management reviewed and approved all consent procedures, research measures, and protocol.

For NESARC Wave 1 (2001–2002), the housing unit sampling frame was based to provide a systematic random sampling stratified by location and demographic factors to match the US Census 2000. All NESARC participants were civilian, non-institutionalized individuals aged 18 years and above. The following groups were oversampled because the NESARC aimed to inform public health substance-use trends in high-priority populations: African-Americans (1.55 oversampling ratio), Hispanic-Americans (1.59 over-sampling ratio), and young adults aged 18–24 years (2.25 oversampling ratio). In each household, one adult was selected, and face-to-face interviews were conducted by lay interviewers or clinicians in respondents' homes following informed consent ($N = 43,093$). A second wave of assessment was conducted 3 years later (Wave 2, 2004–2005). The overall response rate for Wave 1 was 81.0%, and the subsequent response rate among those eligible for Wave 2 was 86.7% ($N = 34,653$). The current analyses utilize the Wave 2 sample only because physical activity was measured only at Wave 2. The sample was further restricted to those who reported having at least one of the risk factors ($N = 22,789$), which served as the five binary latent class indicators for the study: (1) past-year alcohol abuse/dependence, (2) past-year drug abuse/dependence, (3) past-year nicotine dependence, (4) current obesity, and (5) failed to meet federal government recommended weekly physical activity levels in the past year (see assessment section below for more information about the indicators). Requiring the presence of one or more risk factors to be included in the analysis increases the likelihood that a reasonable number of latent classes will be yielded with adequate distributions of individuals in each class. Because pregnancy can inflate body mass index (BMI), we excluded those who were currently pregnant ($n = 348$, 2.7%). Finally, participants younger than 21 years old were excluded because illegality of drinking may have impacted the accuracy of their reports ($n = 24$, 0.07%), leaving a final sample of 22,427 participants for analyses.

Assessment

The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS-IV³⁷) was used to collect information on demographics, physical activity, weight/height, *Diagnostic and Statistical Manual for Mental Disorders* (4th ed.; *DSM-IV*)³¹ criteria for substance-use disorders, and other characteristics. Reliability and validity estimates for substance-use disorder diagnoses and other characteristics in this sample were adequate and available in a previous report.³⁷

Substance-use disorders

Participants who met criteria for either abuse or dependence were combined into an abuse/dependence classification for both alcohol and drug indicators. We combined abuse and dependence because (1) both abuse and dependence are associated with negative health outcomes,³⁸ thus any substance-use disorder may indicate increased disease risk; (2) abuse and dependence represent different severity levels of the same risk factor condition (i.e. substance-use disorder); and (3) it is likely that abuse and dependence will be combined into a single category in *Diagnostic and Statistical Manual for Mental Disorders* (5th ed.; *DSM-V*),³⁹ thus the present findings may have more relevance to future diagnostic classifications. In concordance with extant reports,^{1–3,40} all illicit and prescription drug-use disorders were analyzed as a combined drug abuse/dependence category to increase prevalence and reduce the number of indicators.

Obesity

BMI was computed based on the respondents' self-reported body weight and height (kg/m²) because objective assessments were not collected in NESARC. Individuals with BMIs ≥ 30 kg/m² were classified as obese. We chose to focus on the distinction between obese versus not obese (rather than other cutoffs, such as overweight ≥ 25 kg/m²) based on evidence suggesting that obesity (but not overweight) is associated with increased mortality from cardiovascular disease and obesity-related cancers.⁴¹

Physical activity

Participants were asked to report their levels of physical activity over the past 12 months. Moderate physical activity was defined as 'activities that caused only LIGHT sweating or SLIGHT to MODERATE increases in your breathing or heart rate' and vigorous physical activity was defined as 'activities that caused you to sweat HEAVILY or caused LARGE increases in your breathing or heart rate'. Participants were asked about the frequency (e.g. number of times per week/month/year) and average duration of a typical period of activity (answer in minutes). The US federal government currently recommends that for substantial health benefits, adults should engage in at least 150 minutes a week of moderate-intensity aerobic physical activity, 75 minutes a week of vigorous-intensity aerobic physical activity, or an equivalent combination of moderate- and vigorous-intensity aerobic activity.⁴² As in previous work,⁴³ our physical activity indicator was whether individuals met some combination of this weekly recommendation on average over the past year.

Demographic characteristics

Participants reported their age in years, gender, marital status, race/ethnicity, education level, and annual household income. Marital status was recoded to a binary variable (married vs other). Education was an ordinal variable (less than high school, high school diploma or equivalent, or post-high school). Income was also a continuous variable (21 levels with increments ranging from US\$3,000 to US\$200,000+).

ANALYSIS

A series of latent class models were iteratively conducted, and the smallest set of latent classes that best fit the data was determined. To do so, we evaluated various model fit indices, including Pearson χ^2 , likelihood ratio χ^2 , Akaike information criterion (AIC) and Bayesian information criterion (BIC), Lo–Mendell–Rubin likelihood ratio (LMRLR; at likelihood ratio test for statistically significant differences in fit across models with different numbers of classes), and entropy values. Thereafter, the latent classes identified were then subjected to multinomial logistic regression to assess the association between demographic characteristics and latent classes. All analyses were conducted using Mplus Version 6.0 software program⁴⁴ and Stata Version 11. The complex survey design of the NESARC was accounted for, and Wave 2 sampling weights were applied to approximate the demographics of US Census 2000 population. A two-sided .01 significance level was applied.

RESULTS

Unweighted distribution of risk factors

In the sample, 14.0%, 20.0%, and 3.3% met criteria for past-year alcohol-, nicotine-, and drug-use disorder, respectively. Approximately, 45% of the sample was currently obese, and 57.8% of the sample failed to meet recommended physical activity levels over the past year.

Determining the best-fitting model

As illustrated in Table 1, as the number of classes increased, AIC and BIC decreased, indicating improved model fit. The fit improvement was evidenced only up to the 5-class solution. LMRLR compares n versus $n - 1$ class models (i.e. reject the null hypothesis that $n - 1$ class fits the data better than n class model, if $p < .01$).⁴⁵ LMRLR test showed that the 5-class model was the best-fitting model and yielded highest value of entropy.

Prevalence and characterization of five latent classes

For the 5-class solution, conditional item-response probabilities are presented in Table 2. A total of 23% of the sample was expected to belong to Class 1, which had a response pattern characterized as ‘obese, active non-substance abusers’. Class 2 (19%) was characterized as ‘nicotine-dependent, active, and non-obese’. Class 3 (6%) was characterized as ‘active, non-obese alcohol abusers’. Class 4 constituted the largest latent class (50%) and was characterized as ‘inactive, non-substance users’. Respondents expected to belong to Class 5 were characterized as ‘active, polysubstance abusers’ (3.7%).

Association between demographic characteristics and latent class membership

After the latent classes were identified, we examined the demographic correlates of class membership (results reported in Table 3); latent class was not contingent upon the demographic variables. Relative to the most prevalent Class 4 (inactive, non-substance abusers), older participants were less likely to belong to all other latent classes (see Table 3). Relative to Class 4, males were more likely to belong to all other classes than females were. Married respondents were less likely to belong to Classes 2 (nicotine-dependent, active, and non-obese), 3 (active, non-obese alcohol abusers), and 5 (active polysubstance abusers),

relative to Class 4, than non-married respondents. The married participants, however, were 11% more likely to be classified as Class 1 ('obese, active non-substance abusers') versus Class 4. Relative to Class 4, respondents with a non-White racial background were less likely to belong to all the other classes except for the individuals with an American Indian/Alaska Native racial background. Also, African-Americans were 30% more likely to belong to Class 1 (obese, active, non-substance abusers) than White respondents. Relative to Class 4, participants with higher income were more likely to be classified as Classes 1 (obese, active, non-substance abusers) and 3 (active, non-obese alcohol abusers) but less likely to belong to Class 2 (nicotine-dependent, active, and non-obese). Compared with participants who did not attain a high school diploma, those with a high school diploma/General Educational Development (GED) were more likely to be classified as Classes 1 (obese, active, non-substance abusers) and 3 (active, non-obese alcohol abusers), relative to Class 4. Furthermore, respondents with higher levels of income were more likely to belong to Classes 1 (obese, active non-substance abusers) and 3 (active, non-obese alcohol abusers) relative to Class 4, yet less likely to belong to Class 2 (nicotine-dependent, active, and non-obese) relative to Class 4. Finally, those with some post-high school education were more likely to belong to Classes 1 (obese, active non-substance abusers) and 3 (active, non-obese alcohol abusers), but less likely to belong to Class 2 (nicotine-dependent, active, and non-obese).

DISCUSSION

In a population-based sample of US adults, five distinct latent subgroups adequately accounted for variation in co-occurrence patterns of five clinically important modifiable biobehavioral risk factors for chronic disease. Each of these classes exhibited unique risk factor configurations and demographic profiles from one another. Thus, chronic disease-prevention programming distributed to the at-risk population may be beneficial if tailored and targeted to those particular population subgroups and corresponding risk factor profiles illustrated in this analysis (see Tables 2 and 3).^{46,47}

Among the US adults who endorsed at least one of the index risk factors studied in this report, four of the five classes were characterized by a 100% likelihood of having one of the risk factors coupled with a low or moderate likelihood of having the other four risk factors. Latent classes with 1.00 or 0.00 probabilities for indicators have been reported before in the literature⁴⁸ and would be expected in studies using samples saturated for risk factor endorsement such as in this study, suggesting highly discriminative classes. Each of these four classes was typified by the presence of a different risk factor. Collectively, these four classes represented 96% of the overall sample. This finding has implications for chronic disease-prevention programming for these particular biobehavioral factors among those who endorsed at least one of these risk factors. One perspective is that programs focused primarily on single risk factor may be most efficient and practical in circumstances in which financial and temporal resources are limited.²⁶ Another perspective is that targeting multiple health behaviors may nonetheless be useful, particularly when a single intervention model can be readily applied to multiple risk factors.²¹ Regardless, our findings suggest that for these five risk factors known to account for substantial public health burden,¹⁻⁷ most

members of the at-risk community are typified by a high likelihood of only a single risk factor.

Several high-prevalence classes were identified in this sample. About half of the respondents were expected to belong to a class of inactive individuals who did not have concurrent substance-use problems and had slightly lower probability of obesity in comparison to the overall sample. Thus, physical inactivity appears to be a highly prevalent factor, which underscores the potential utility of population-based physical activity programming.⁴⁹ Almost a quarter of the sample comprised a class of obese individuals who were nevertheless highly likely to currently meet recommendations for weekly physical activity. Accordingly, members of this class may benefit most from obesity interventions that do not solely focus on exercise,⁵⁰ or perhaps require a higher dose of physical activity (above the *Centers for Disease Control* and Prevention (CDC) recommendations) to reduce their obesity risk.⁵¹ About one-fifth of the sample was expected to belong to a class typified by nicotine dependence and lower likelihood of physical inactivity, drug-use disorder, and obesity than the overall sample. Given the poor outcomes associated with tobacco users who are nicotine-dependent, intensive pharmacotherapy–counseling combination interventions may be indicated for this class of sizeable prevalence.⁵²

A class of active, non-obese individuals with alcohol-use disorders predicted to represent 6% of the sample was identified. The profile of this class is consistent with recent work illustrating a counterintuitive positive association between alcohol use and physical activity.^{24,53,54} Recent evidence suggests that relation between physical activity and substance use appears to be specific to alcohol but not other substances even after controlling for the effects of other substance-use disorders,⁵⁵ which is concordant with the current results illustrating that members of this class had little or no likelihood of comorbid tobacco- and drug-use disorder. Furthermore, extant data suggest that physical activity and alcohol use co-occurrence is more common among younger (vs older) and male (vs female) individuals.^{24,53,54} Consistent with past findings, demographic analyses of this class also suggest that physically active alcohol users tend to be younger and males. We also found that members of this class are more likely to be non-married, White, of a higher income bracket, or more educated. Thus, it may be important to assess for alcohol-use disorder among active adults from this demographic bracket, who may otherwise appear healthy and potentially overlooked in alcohol programming.

A class of polysubstance abusers expected to comprise 4% of the sample was yielded. This class was defined by meeting diagnostic criteria for an illicit drug-use disorder combined with a high likelihood of comorbid alcohol- and tobacco-use disorder, but slightly lower risk of obesity and physical inactivity than the overall sample. Relative to the largest class, this class was more likely to be younger, male, non-married, and White. The characteristics of this class are consistent with notions that the co-occurrence among multiple substance-use disorders may reflect an underlying syndrome of externalizing behavior that is more common in young men and may be associated with antisocial tendencies that could interfere with successful marriage.^{56,57} Clinically, these findings suggest that a small, but sizeable, portion of individuals may require multi-substance addiction interventions.^{58–60}

This study had limitations that impact interpretation of these findings. The goal of this study was to assess clustering of five putatively important modifiable biobehavioral risk factors for chronic disease. Thus, the findings primarily generalize to considerations of groupings of individuals with regard to these indicators, and the prevalence and profiles of classes may be different if other risk factors (e.g. diet) are considered in concert with the five we focused on. The intention of this study was to describe patterns of co-occurrence, and therefore, we used a cross-sectional design. However, this design does not elucidate causal and temporal mechanisms that may link risk factors to one another. The sample included only respondents who were positive for one of the five risk factors, which may have influenced the prevalence and configuration of the classes yielded and impacted generalizability considerations. Nevertheless, this restriction was important for yielding an interpretable number of classes, and allows the data to generalize to the population of adults who may be seen in settings such as primary care and who present with at least one apparent risk factor. Additionally, all measurement was based on self-report binary indicators. Although self-report measures that rely on cutoffs are useful for screening purposes and classifying those who respond to interventions, objective and biological assessments would have been more accurate. Furthermore, binary indicators have limited ability to characterize individuals across the continuum of functioning or who have subclinical levels of risk. Also, this study utilizes LCA, which is a data-driven a posteriori approach. Thus, the classes identified here are generalizable to the US population primarily, and may not replicate in all other populations. Finally, these data were drawn from 2004–2005 because this was the most recently conducted US population-based national study that we were aware of that collected detailed clinical information on substance-use disorder *DSM-IV* diagnoses coupled with BMI and physical activity. Hence, because these data are 8 years old, they may not entirely reflect trends that could be occurring today. However, studies that have examined the prevalence of substance-use disorders, obesity, and physical activity in separate databases have shown that these factors have not dramatically changed in prevalence.^{61–63} Thus, these findings are likely generalizable to today's public health trends in the most part. This study also had several offsetting strengths, such as the use of a nationally representative sample, application of a current analytic methodology, and use of valid clinical interview assessments to generate *DSM-IV* substance-use disorder diagnoses.

CONCLUSION

To our knowledge, this was first LCA study of these key modifiable biobehavioral risk factors, which were purposefully selected because of their public health significance and their tendency to co-occur in non-uniform configurations.^{1–7,24} Results suggest that risk factors cluster together in a unique fashion, with the majority of the at-risk population of US adults expected to have a high probability of endorsing only one of these five biobehavioral risk factors. Accordingly, health promotion programming for substance-use disorder, physical activity, and obesity that targets the population that endorses at least one of these risk factors should perhaps consider the utility of single risk factor interventions.

Acknowledgments

This research was supported by US National Institutes of Health grant DA025041.

References

1. Grant BF, Hasin DS, Chou SP, Stinson FS, Dawson DA. Nicotine dependence and psychiatric disorders in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*. 2004; 61(11):1107–15. [PubMed: 15520358]
2. Hasin DS, Stinson FS, Ogburn E, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*. 2007; 64(7):830–42. [PubMed: 17606817]
3. Compton WM, Thomas YF, Stinson FS, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*. 2007; 64(5):566–76. [PubMed: 17485608]
4. CDC. Prevalence of regular physical activity among adults – United States, 2001 and 2005. *MMWR: Morbidity and Mortality Weekly Report*. 2007; 56(46):1209–12. [PubMed: 18030281]
5. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA*. 2010; 303(3):235–41. [PubMed: 20071471]
6. Lopez, AD.; Mathers, CD.; Ezzati, M.; Jamison, DT.; Murray, CJL. Measuring the global burden of disease and risk factors. In: Lopez, AD.; Mathers, CD.; Ezzati, M.; Jamison, DT.; Murray, CJL., editors. *Global Burden of Disease and Risk*. Washington, DC: The World Bank; 2006.
7. Danaei G, Ding EL, Mozaffarian D, Taylor B, Rehm J, Murray CJ, et al. The preventable causes of death in the United States: Comparative risk assessment of dietary, lifestyle, and metabolic risk factors. *PLoS Medicine*. 2009; 6(4):e1000058. [PubMed: 19399161]
8. Pronk NP, Anderson LH, Crain AL, Martinson BC, O'Connor PJ, Sherwood NE, et al. Meeting recommendations for multiple healthy lifestyle factors. Prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. *American Journal of Preventive Medicine*. 2004; 27(2 Suppl):25–33. [PubMed: 15275671]
9. Fine LJ, Philogene GS, Gramling R, Coups EJ, Sinha S. Prevalence of multiple chronic disease risk factors. 2001 National Health Interview Survey. *American Journal of Preventive Medicine*. 2004; 27(2 Suppl):18–24. [PubMed: 15275670]
10. Funderburk JS, Maisto SA, Sugarman DE, Wade M. The covariation of multiple risk factors in primary care: A latent class analysis. *Journal of Behavioral Medicine*. 2008; 31(6):525–35. [PubMed: 18800242]
11. Cleveland MJ, Collins LM, Lanza ST, Greenberg MT, Feinberg ME. Does individual risk moderate the effect of contextual-level protective factors? A latent class analysis of substance use. *Journal of Prevention & Intervention in the Community*. 2010; 38(3):213–28. [PubMed: 20603758]
12. Huh J, Riggs NR, Spruijt-Metz D, Chou CP, Huang Z, Pentz M. Identifying patterns of eating and physical activity in children: A latent class analysis of obesity risk. *Obesity*. 2011; 19(3):652–8. [PubMed: 20930718]
13. Laska MN, Pasch KE, Lust K, Story M, Ehlinger E. Latent class analysis of lifestyle characteristics and health risk behaviors among college youth. *Prevention Science*. 2009; 10(4):376–86. [PubMed: 19499339]
14. Paavola M, Vartiainen E, Haukka A. Smoking, alcohol use, and physical activity: A 13-year longitudinal study ranging from adolescence into adulthood. *The Journal of Adolescent Health*. 2004; 35(3):238–44. [PubMed: 15313507]
15. Chioloro A, Wietlisbach V, Ruffieux C, Paccaud F, Cornuz J. Clustering of risk behaviors with cigarette consumption: A population-based survey. *Preventive Medicine*. 2006; 42(5):348–53. [PubMed: 16504277]
16. Johnson MF, Nichols JF, Sallis JF, Calfas KJ, Hovell MF. Interrelationships between physical activity and other health behaviors among university women and men. *Preventive Medicine*. 1998; 27(4):536–44. [PubMed: 9672947]

17. Schneider S, Huy C, Schuessler M, Diehl K, Schwarz S. Optimising lifestyle interventions: Identification of health behaviour patterns by cluster analysis in a German 50+ survey. *European Journal of Public Health*. 2009; 19(3):271–7. [PubMed: 19164433]
18. Chou KL. The prevalence and clustering of four major lifestyle risk factors in Hong Kong Chinese older adults. *Journal of Aging and Health*. 2008; 20(7):788–803. [PubMed: 18562762]
19. Poortinga W. The prevalence and clustering of four major lifestyle risk factors in an English adult population. *Preventive Medicine*. 2007; 44(2):124–8. [PubMed: 17157369]
20. Kannel WB. Contributions of the Framingham Study to the conquest of coronary artery disease. *The American Journal of Cardiology*. 1988; 62(16):1109–12. [PubMed: 3189175]
21. Goldstein MG, Whitlock EP, DePue J. Multiple behavioral risk factor interventions in primary care. Summary of research evidence. *American Journal of Preventive Medicine*. 2004; 27(2 Suppl):61–79. [PubMed: 15275675]
22. Pronk NP, Peek CJ, Goldstein MG. Addressing multiple behavioral risk factors in primary care. A synthesis of current knowledge and stakeholder dialogue sessions. *American Journal of Preventive Medicine*. 2004; 27(2 Suppl):4–17. [PubMed: 15275669]
23. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: An introduction and overview. *Preventive Medicine*. 2008; 46(3):181–8. [PubMed: 18319098]
24. Lisha NE, Martens M, Leventhal AM. Age and gender as moderators of the relationship between physical activity and alcohol use. *Addictive Behaviors*. 2011; 36(9):933–6. [PubMed: 21561719]
25. Akbartartoori M, Lean ME, Hankey CR. Relationships between cigarette smoking, body size and body shape. *International Journal of Obesity*. 2005; 29(2):236–43. [PubMed: 15505632]
26. Prochaska JJ, Nigg CR, Spring B, Velicer WF, Prochaska JO. The benefits and challenges of multiple health behavior change in research and in practice. *Preventive Medicine*. 2010; 50(1–2): 26–9. [PubMed: 19948184]
27. Auerbach KJ, Collins LM. A multidimensional developmental model of alcohol use during emerging adulthood. *Journal of Studies on Alcohol*. 2006; 67(6):917–25. [PubMed: 17061010]
28. Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS procedure for latent class analysis. *Structural Equation Modeling*. 2007; 14(4):671–94. [PubMed: 19953201]
29. Magidson J, Vermunt JK. Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research*. 2002; 20:37–44.
30. Leventhal AM, Mickens L, Dunton GF, Sussman S, Riggs NR, Pentz MA. Tobacco use moderates the association between major depression and obesity. *Health Psychology*. 2010; 29(5):521–8. [PubMed: 20836607]
31. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4. Washington, DC: American Psychiatric Association; 2000. Text Revision
32. Berlin I, Singleton EG. Nicotine dependence and urge to smoke predict negative health symptoms in smokers. *Preventive Medicine*. 2008; 47(4):447–51. [PubMed: 18602945]
33. Muscat JE, Ahn K, Richie JP Jr, Stellman SD. Nicotine dependence phenotype and lung cancer risk. *Cancer*. 2011; 117:5370–6. [PubMed: 21826644]
34. Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *The Lancet*. 2009; 373(9682):2223–33.
35. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The burden of disease and injury in Australia. *Bulletin of the World Health Organization*. 2001; 79(11):1076–84. [PubMed: 11731817]
36. Grant, BF.; Moore, TC.; Shepard, J.; Kaplan, K. *Source and Accuracy Statement: Wave 1 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2003.
37. Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): Reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence*. 2003; 71(1):7–16. [PubMed: 12821201]
38. Samokhvalov AV, Popova S, Room R, Ramonas M, Rehm J. Disability associated with alcohol abuse and dependence. *Alcoholism: Clinical and Experimental Research*. 2010; 34(11):1871–8.

39. American Psychiatric Association. [Last accessed June 2013] Proposed Revisions to the Substance Use Disorder Diagnoses in DSM-V. Available online at: <http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=431>
40. Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RM, Petukhova M, et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Archives of General Psychiatry*. 2007; 64(5):543–52. [PubMed: 17485606]
41. Flegal KM, Graubard BI, Williamson DF, Gail MH. Cause-specific excess deaths associated with underweight, overweight, and obesity. *JAMA*. 2007; 298(17):2028–37. [PubMed: 17986696]
42. Services UDoHaH. [Last accessed June 2013] 2008 Physical Activity Guidelines for Americans: Summary. 2008. Available online at: <http://www.health.gov/paguidelines/guidelines/summary.aspx>
43. Mukamal K, Ding E, Djoussé L. Alcohol consumption, physical activity, and chronic disease risk factors: A population-based cross-sectional survey. *BMC Public Health*. 2006; 6(1):118. [PubMed: 16670030]
44. Muthén, LK.; Muthén, BO. [Last accessed June 2013] Mplus User's Guide Version 6. 2010. Available online at: <http://www.statmodel.com>
45. Lo Y, Mendell N, Rubin D. Testing the number of components in a normal mixture. *Biometrika*. 2001; 88(3):767–78.
46. Rakowski W. The potential variances of tailoring in health behavior interventions. *Annals of Behavioral Medicine*. 1999; 21(4):284–9. [PubMed: 10721434]
47. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. *Health Education Research*. 2008; 23(3):454–66. [PubMed: 18349033]
48. Lanza ST, Savage JS, Birch LL. Identification and prediction of latent classes of weight-loss strategies among women. *Obesity*. 2010; 18(4):833–40. [PubMed: 19696754]
49. World Health Organization. A Guide for Population-Based Approaches to Increasing Levels of Physical Activity. Geneva: WHO Library Cataloguing; 2007.
50. Avenell A, Broom J, Brown TJ, Poobalan A, Aucott L, Stearns SC, et al. Systematic review of the long-term effects and economic consequences of treatments for obesity and implications for health improvement. *Health Technology Assessment*. May; 2004 8(21):iii–iv. 1–182. [PubMed: 15147610]
51. Ohkawara K, Tanaka S, Miyachi M, Ishikawa-Takata K, Tabata I. A dose-response relation between aerobic exercise and visceral fat reduction: Systematic review of clinical trials. *International Journal of Obesity*. 2007; 31(12):1786–97. [PubMed: 17637702]
52. Fiore MC. US public health service clinical practice guideline: Treating tobacco use and dependence. *Respiratory Care*. 2000; 45(10):1200–62. [PubMed: 11054899]
53. French MT, Popovici I, Maclean JC. Do alcohol consumers exercise more? Findings from a national survey. *American Journal of Health Promotion*. 2009; 24(1):2–10. [PubMed: 19750956]
54. Pate RR, Heath GW, Dowda M, Trost SG. Associations between physical activity and other health behaviors in a representative sample of US adolescents. *American Journal of Public Health*. 1996; 86(11):1577–81. [PubMed: 8916523]
55. Lisha NE, Sussman S. Relationship of high school and college sports participation with alcohol, tobacco, and illicit drug use: A review. *Addictive Behaviors*. 2010; 35(5):399–407. [PubMed: 20100638]
56. Eaton NR, Keyes KM, Krueger RF, Balsis S, Skodol AE, Markon KE, et al. An invariant dimensional liability model of gender differences in mental disorder prevalence: Evidence from a national sample. *Journal of Abnormal Psychology*. 2012; 121:282–8. [PubMed: 21842958]
57. Krueger RF, Hicks BM, Patrick CJ, Carlson SR, Iacono WG, McGue M. Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology*. 2002; 111(3):411–24. [PubMed: 12150417]
58. Prochaska JJ, Delucchi K, Hall SM. A meta-analysis of smoking cessation interventions with individuals in substance abuse treatment or recovery. *Journal of Consulting and Clinical Psychology*. 2004; 72(6):1144–56. [PubMed: 15612860]

59. Baca CT, Yahne CE. Smoking cessation during substance abuse treatment: What you need to know. *Journal of Substance Abuse Treatment*. 2009; 36(2):205–19. [PubMed: 18715746]
60. Kahler CW, Metrik J, LaChance HR, Ramsey SE, Abrams DB, Monti PM, et al. Addressing heavy drinking in smoking cessation treatment: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*. 2008; 76(5):852–62. [PubMed: 18837602]
61. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA*. 2012; 307(5):491–7. [PubMed: 22253363]
62. Carlson SA, Densmore D, Fulton JE, Yore MM, Kohl HW 3rd. Differences in physical activity prevalence and trends from 3 US surveillance systems: NHIS, NHANES, and BRFSS. *Journal of Physical Activity & Health*. 2009; 6(Suppl 1):S18–27. [PubMed: 19998846]
63. Substance Abuse and Mental Health Services Administration. Results from the 2011 National Survey on Drug Use and Health: Summary of National Findings. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2012.

Table 1

Model-fit indices for latent class models.

	Number of classes					
	2	3	4	5	6	
Pearson χ^2	478.360	114.643	22.927	1.235	3.475	
LR χ^2	550.124	140.977	19.503	1.250	4.215	
df	20	14	8	2	N/A	
No. of parameters	11	17	23	29	35	
Log likelihood	-52,570.834	-49,855.539	-48,753.798	-48,515.066	-48,503.588	
AIC	105,163.668	99,745.079	97,553.596	97,088.132	97,077.176	
BIC	105,251.866	99,881.385	97,738.010	97,320.654	97,357.806	
LMR testing the null hypothesis	Class 1 vs Class 2	Class 2 vs Class 3	Class 3 vs Class 4	Class 4 vs Class 5	Class 5 vs Class 6	
LMR probability	<.001	<.001	<.001	<.001	.433	
Entropy	.791	.791	.871	.928	.902	

LMR: Lo–Mendell–Rubin; AIC: Akaike information criterion; BIC: Bayesian information criterion; N/A: Not Applicable.

Results from model fit indices. LMR illustrates that model fit improves up to 5 classes and then does not significantly improve fit beyond 5 classes. Hence, we selected the 5-class model as our final model.

Table 2

Prevalence of latent classes and conditional risk factor probabilities within each latent class.

		Latent classes, %, N, and description				
		Class 1 (22.5%) (n = 5,042)	Class 2 (18.7%) (n = 4,183)	Class 3 (5.6%) (n = 1,245)	Class 4 (49.6%) (n = 11,129)	Class 5 (3.7%) (n = 828)
	Total sample (N = 22,427)	Active, obese non-substance abusers	Nicotine-dependent active non-obese	Active, non- obese alcohol abusers	Inactive non-substance abusers	Active, polysubstance abusers
Alcohol abuse/dependence	.15	.08	.21	1.00	.04	.58
Nicotine dependence	.22	.05	1.00	.00	.06	.51
Drug abuse/dependence	.04	.00	.00	.02	.00	1.00
Obesity (BMI ≥30.0)	.43	1.00	.21	.00	.28	.23
Below physical activity cutoff	.55	.06	.21	.00	1.00	.20

Proportion of total sample (N = 22,427) and probabilities of positive classifications on five latent class analysis indicators. Perfect prediction of class membership by indicator is highlighted in bold.

Table 3

Multinomial logistic regression: demographic characteristics predicting latent classes.

	Class 1	Class 2	Class 3	Class 5
	<i>RRR^a</i> (SE)	<i>RRR^a</i> (SE)	<i>RRR^a</i> (SE)	<i>RRR^a</i> (SE)
Age	0.97 (0.01) **	0.96 (0.01) **	0.94 (0.01) **	0.92 (0.01) **
Male	1.59 (0.06) **	1.99 (0.08) **	3.94 (0.17) **	3.93 (0.33) **
Married ^b	1.11 (0.04) *	0.70 (0.03) **	0.51 (0.04) **	0.33 (0.03) **
Race/ethnicity ^c				
Black, non-Hispanic	1.30 (0.06) **	0.50 (0.03) **	0.46 (0.05) **	0.47 (0.05) **
American Indian/Alaska Native, non-Hispanic	1.39 (0.19)	1.47 (0.19) *	1.11 (0.28)	1.44 (0.37)
Asian/Pacific Islander, non-Hispanic	0.27 (0.04) **	0.29 (0.04) **	0.27 (0.06) **	0.17 (0.06) **
Hispanic, any race	0.86 (0.04) *	0.27 (0.02) **	0.48 (0.04) **	0.34 (0.04) **
Income	1.02 (0.01) **	0.98 (0.01) **	1.07 (0.01) **	0.99 (0.01)
Education ^d				
High school diploma/GED	1.39 (0.08) **	0.95 (0.06)	1.72 (0.24) **	1.15 (0.15)
Post-high school education	1.33 (0.07) **	0.69 (0.04) **	2.07 (0.27) **	0.98 (0.12)

Class descriptors: Class 1 – active, obese non-substance abusers; Class 2 – nicotine-dependent active non-obese; Class 3 – active, non-obese alcohol abusers; Class 4 – inactive non-substance abusers; Class 5 – active, polysubstance abusers; *RRR*: relative risk ratio; GED: General Educational Development; SE: standard error.

^aRelative to Class 4, the most prevalent class.

^bCompared with 'non-married', including cohabitation, widowed, divorced, separated, and never married.

^cReference category was non-Hispanic White.

^dReference category was those who did not attain high school diploma or GED.

* $p < .01$;

** $p < .001$.