

# The Effects of Heavy Drinking, Smoking, Anxiety and Depression on Stroke Odds: A Cross-Sectional Study

**Robert Wardrup**  
**Elizabeth Murphy**  
**Natalie Requinez**  
**Joshua Averitt**

---

Though numerous studies have investigated the relationships between lifestyle risk factors and stroke outcome, few have done so while controlling as age and race/ethnicity range. Utilizing quasi-binomial logistic regression, we show that the odds of having had a stroke increase with being a daily smoker, days of anxiety, and having a depression diagnosis.

*Keywords:* Stroke, Alcohol, Smoking, Anxiety, Depression

---

## Introduction

Strokes nearly affect 800,000 people in the United States every year and are the fifth leading cause of death in the United States ([American Stroke Association, 2017a](#)). Despite increased efforts on education and general public knowledge about stroke and other cardiovascular risk diseases, the prevalence rates have shown little improvement within the past ten years ([Mozaffarian et al., 2015](#)). Previous research has shown that a variety of controllable risk factors are strongly associated with stroke outcome. Although behaviors such as smoking and poor diet have been well-known risk factors, other factors such as stress, lack of sleep, depression, and anxiety still have unclear associations [(Risk Factors Stroke Association, ([American Stroke Association, 2017b](#))).

Stress, lack of sleep, and mental disorders such as depression and anxiety are issues prominent throughout the United States. Anxiety is the most common mental illness and approximately 18% of the U.S population is affected while depression affects nearly 7% of adults ([Anxiety and Depression Association of America, 2017](#)). Those who are diagnosed with anxiety also have a 50% chance of being diagnosed with depression and vice versa ([Anxiety and Depression Association of America, 2017](#)). Furthermore, approximately, 30% of Americans suffer from lack of sleep ([Centers for Disease Control and Prevention, 2015](#)). Although average stress levels have decreased since the recession, the average stress scores for American adults are still at moderate levels, but is increased in lower household incomes ([American Psychological Association, 2015](#)). Each of these factors has a considerable significance to public health as issues on their

own. If someone is suffering from multiple risk factors, the risk of stroke could be significant and possibly result in death.

The goal of this study is to test if there is an association between the different risk factors, particularly with stress, lack of sleep, depression, and anxiety, with strokes using data from the Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS collects data from 50 states, D.C., and the even throughout the territories via landline and cellphones. It is in partnership with the CDC and has been useful for providing large quantities of health related data ([Centers for Disease Control and Prevention, 2014](#)).

## Methods

The BRFSS survey data were weighted using the raking method, which is a two-part methodology to help insure unbiased results by accounting for noncoverage and nonresponse bias and forcing the total number of cases to equal the population estimate of each state in the United States ([Centers for Disease Control and Prevention, 2007](#)). Raking works by repeatedly adjusting weight across a set of selected variables until the weights converge and the survey population totals are equal to the census population totals for each selected variable ([Fricker and Andersen, 1993](#)).

### *Dependent and Independent Variables*

The dependent variable, stroke outcome, was recoded to a binary variable for use in a logistic regression model. Depression diagnosis and heavy drinker were also recoded into a binary variable. For each variable used in the analysis, *don't know / unsure* and *refused* responses were dropped. Table 1 lists the variables included in the model.

Table 1: Dependent and Independent Variables Used in Logistic Regression Model

Variable Name	Variable Type
Stroke Diagnosis	Dichotomous (DV)
Heavy Alcohol Drinker	Dichotomous
Depression Diagnosis	Dichotomous
Sex	Dichotomous

Variable Name	Variable Type
Days Anxious	Continuous
Daily Sleep Hours	Continuous
Smoker Status	Categorical
Race/Ethnicity	Categorical
Age Group	Categorical

A subject was designated as a heavy drinker if they were either an adult male who reported consuming more than 14 drinks per week or an adult female who reported consuming more than 7 drinks per week. This value was calculated after asking the subjects, “during the past 30 days, on the days when you drank, about how many drinks did you drink on average?” and “during the past 30 days, how many days per week or month did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage or liquor?” (?).

### *Logistic Regression Model*

The first model used in calculating the logistic regression utilized the standard binomial function to predict the log odds of binary outcome  $k$ . The binomial formula, is given as

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k},$$

where  $k$  represents a stroke outcome,  $n$  represents the sample size, and  $p$  is the probability that a stroke will occur. The results of the logistic regression model were then exponentiated to produce the odds ratios (OR) reported within this paper.

## **Results**

The weighted number of strokes in the United States in 2016 was 8,020,080, representing approximately 2.5% of the total 2016 United States population. Out of these, 3.20% were female and 3.13% were male.

Table 2: Proportion of Stroke Outcomes in Males and Females

Diagnosis	Male	Female
No Stroke	96.87%	96.80%
Stroke	3.13%	3.20%

The proportion of subjects who have suffered from a stroke by smoker status is shown in Table 3, and visually in Figure 2. The percentage of subjects who have had a stroke is highest for the subjects who smoke daily, followed by subjects who are former smokers, followed by subjects who smoke some days. A possible explanation for the lower percentage of stroke in subjects who smoke some days than subjects who are former smokers is that stroke victims may simply have ceased smoking after having a stroke.

Table 3: Proportion of Stroke Outcomes and Smoker Status

Smoker Status	No Stroke	Stroke
Never Smoked	97.8%	2.2%
Former Smoker	95.4%	4.6%
Smokes Some Days	96.2%	3.8%
Smokes Daily	95.0%	5.0%

The proportions shown in Table 3 are depicted as a mosaic plot in Figure 2. The mosaic plot is a graphical depiction the proportions within each table cell, shaded by the difference from the expected observation. Blue depicts a higher than expected number of observations for that cell, and red depicts a lower than expected number of observations for that cell. As shown in Figure 2, there were a greater than expected number of non-smokers that have never had a stroke, and a greater than expected number of smokers and former smokers who have had a stroke.

A chi-squared test of independence revealed that there is statistically significant evidence that stroke outcome and smoking status were not independent from one another,  $\chi^2(3, N=486,303) = 2352.5$ ,  $p < 0.001$ .

The mean hours of sleep for those who had a stroke was 6.99 hours and 6.98 for those who did not have

# Proportions of Stroke Diagnoses by Smoker Status

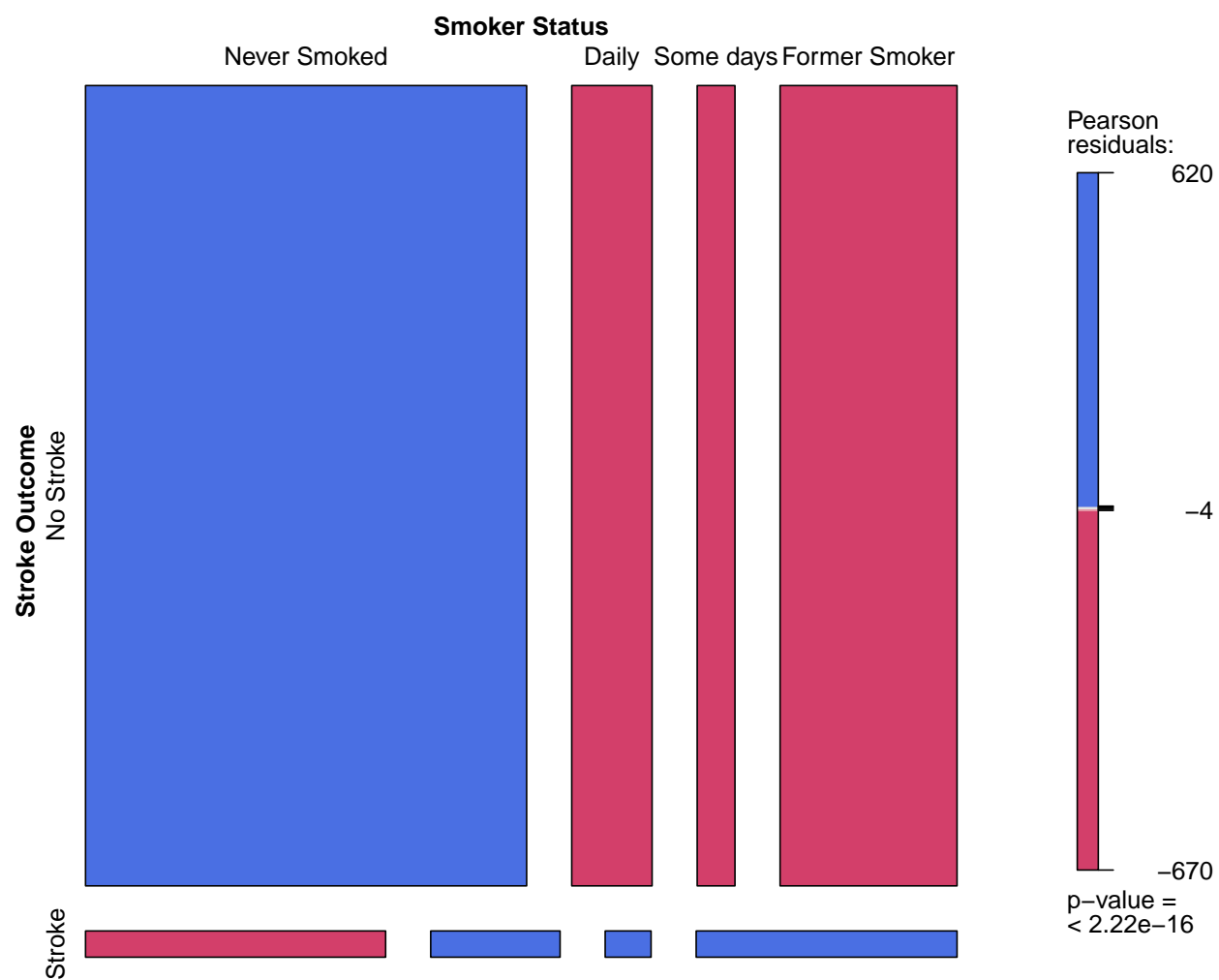


Figure 1: Proportion of Stroke Outcomes and Smoker Status

a stroke. Though the difference in average hours of sleep was small, approximately 6 seconds, the difference nearly achieved statistical significance at  $\alpha = 0.05$  with a p-value of 0.06.

Using the Tukey method, where observations greater or less than the sum of the median plus or minus 1.5 times the IQR marked as outliers, we noted that there were a large number of outliers in the reported number of alcoholic drinks weekly. There were 69,934 observations that were marked as outliers, amounting to 18.1% of the pre-weighted data. The mean of the outliers was 1,465.8 drinks weekly. The mean number of alcoholic drinks consumed weekly, including outliers, was 1,465.8 drinks weekly, and 72.83 drinks weekly with the outliers removed. It is unclear if the CDC accounted for these outliers when calculating the *heavy drinker* variable.

Table 4: Mean Values of Selected Risk Factors

Factor	Outcome	Mean
Days Anxious	No Stroke	4.79
	Stroke	8.05
Nightly Sleep Hours	No Stroke	6.98
	Stroke	6.99
Number Drinks Weekly, With Outliers	No Stroke	314.12
	Stroke	267.97
Number Drinks Weekly, Outliers Removed	No Stroke	79.07
	Stroke	40.31

### *Logistic Regression Model*

The binomial logistic regression model resulted in a  $\phi$  value of 2.52, indicating that there is overdispersion in the estimates and the quasi-binomial function should be used instead.

The results of the quasi-binomial logistic regression (Table 4) indicate that, controlling for age, race/ethnicity, and sex, Several of the tested risk factors were significantly correlated with stroke outcome. Being a daily smoker, Monthly days with anxiety, and diagnosed depression were all positively associated with greater odds of stroke.

Daily cigarette smokers had 1.68 times the odds of having had a stroke than those who have never

smoked cigarettes. For each additional day heavily affected by anxiety, the odds of having had a stroke were 1.03 times higher, and for those who have been diagnosed with depression the odds of stroke were 2.12 times higher. Of the control variables, age of 25 years old and above were the only factors significantly associated with stroke outcome, with an odds ratio that increased with age.

Table 5: Logistic Regression Model Estimating Effects of Risk Factors & Demographic Variables on Stroke Outcome Odds.

Variable	OR (95% CI)	P-Value
<b>Risk Factors</b>		
<i>Drinking</i>		
Non-Heavy Drinker	1 (Baseline Factor)	
Heavy Drinker	0.989 (0.41 - 2.38)	0.98
<i>Smoking</i>		
Non-Smoker	1 (Baseline Factor)	
Former Smoker	1.24 (0.91 - 1.70)	0.18
Smoker (Some days)	1.65 (0.95 - 2.88)	0.08
Smoker (Daily)	1.68 (1.12 - 2.53)	0.01
<i>Mental Health</i>		
Monthly Days Anxious	1.03 (1.01 - 1.04)	<0.01
No Depression Diagnosis	1 (Baseline Factor)	
Depression Diagnosis	2.12 (1.48 - 3.04)	<0.01
Nightly Sleep Hours	0.91 (0.82 - 1.00)	0.06
<i>Sex</i>		
Male	1 (Baseline Factor)	
Female	0.77 (0.58 - 1.02)	0.07
<b>Race/Ethnicity</b>		
Other Race / Non-Hispanic	1 (Baseline Factor)	
Hispanic	1.42 (0.43 - 4.70)	0.57
Black Only / Non-Hispanic	2.41 (0.88 - 6.62)	0.09

Variable	OR (95% CI)	P-Value
White Only / Non-Hispanic	1.87 (0.73 - 4.83)	0.20
Multiracial / Non-Hispanic	2.92 (0.90 - 9.50)	0.08
<b>Age Group</b>		
Age 18 to 24	1 (Baseline Factor)	
Age 25 to 34	3.92 (0.80 - 19.66)	0.10
Age 35 to 44	9.33 (2.00 - 43.51)	<0.01
Age 45 to 54	12.35 (2.79 - 54.71)	<0.01
Age 55 to 64	20.93 (4.80 - 91.34)	<0.01
Age 65 or Older	44.87 (10.36 - 194.43)	<0.01

The effect of each significant risk factor, stratified by depression diagnosis, on the probability of having had a stroke is shown in Figure 2. The probability of a person having a stroke increases with the number of days a person is significantly affected by anxiety. This probability is multiplied if a person has been diagnosed with depression. For non-smokers with no depression diagnosis who did not significantly suffer from anxiety, the probability of having had a stroke approaches zero. If the same subject was diagnosed with depression, the probability of having had a stroke nears 5%.

The probability of having had a stroke is multiplied in daily smokers, beginning at approximately 5% for daily smokers who have not been diagnosed with depression nor have suffered from significant anxiety. With a depression diagnosis, the probability originates at approximately 7% for subjects with zero days of anxiety and increases to 15% for subjects who experienced thirty days of significant anxiety. This relationship is similar for subjects who currently smoke some days.

## Conclusion

Our study assessed the association between heavy drinking, anxiety, depression on stroke odds. Using the 2016 Behavioral Risk Factor Surveillance System dataset from the CDC, we determined that being a daily smoker or having a depression diagnosis are associated with increased odds of having had a stroke. Additionally, we found that the number of days a person experienced significant anxiety was associated with increased odds of having had a stroke. Neither the number of hours per night a subject slept, nor being a



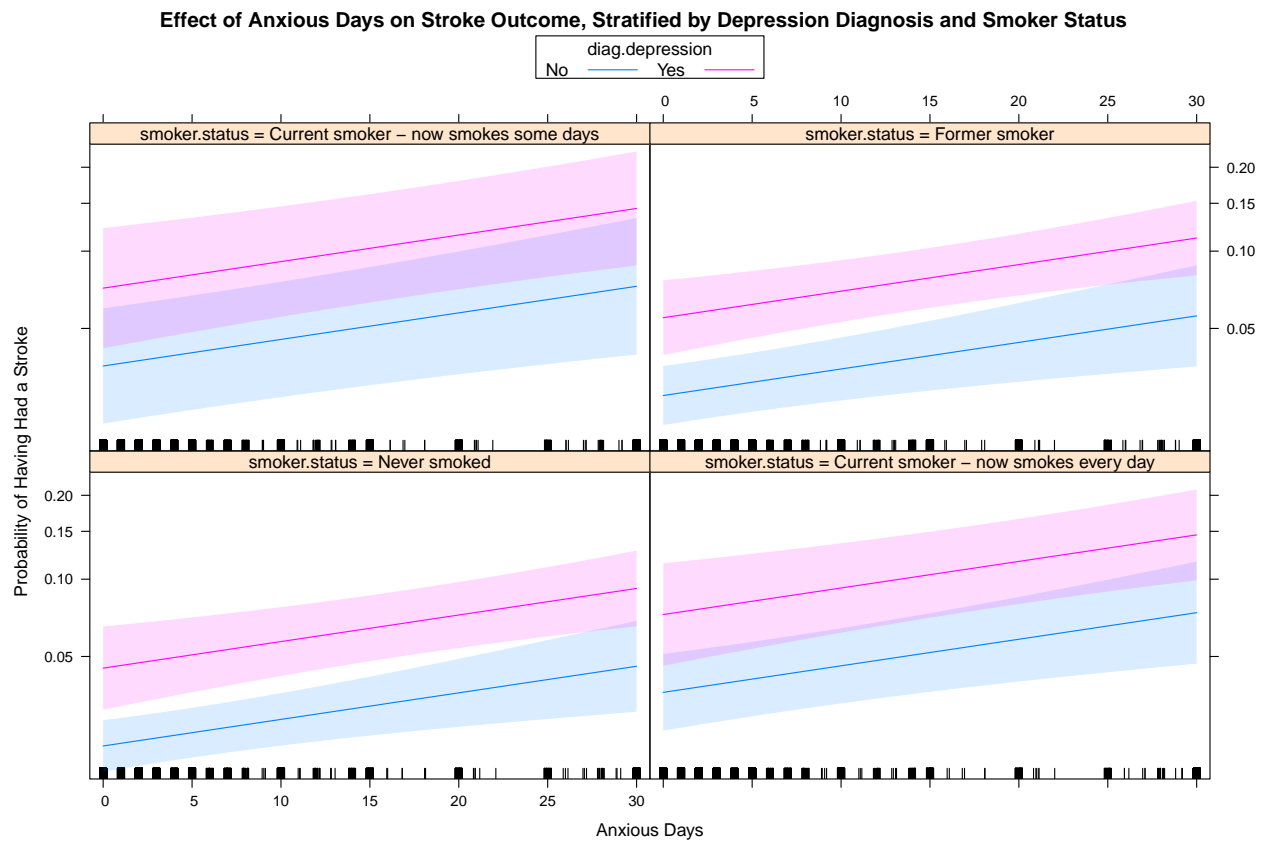


Figure 2: Effects of Anxious Days on Stroke Outcome, Stratified by Depression Diagnosis and Smoker Status

heavy drinker was associated with the odds of having had a stroke.

One limitation of this study is its cross-sectional design. Because stroke prevalence was used, causality and risk cannot be inferred from our data. Furthermore, we made the assumption that there was no pattern behind the *unsure* and *refused* survey responses, when there could have been an underlying reason for these answers. Although removing these responses still afforded the authors adequate statistical power, the use of multiple imputation to recover lost data could have resulted in more robust results.

Another limitation in our study is the possibility of biases inherent in any prevalence study. Subjects may not always be forthcoming when answering questions about unhealthy behaviors such as, “during the past 30 days, how many days per week or per month did you have at least one drink of any alcoholic beverage. . .” or “. . . During the past 30 days, on the days when you drank, how many drinks did you drink on the average?” Questions involving mental health may also invoke in the subject the same desire to “save face.”

Finally, there were clearly outliers in the reported weekly number of alcohol drinks. It is unclear whether or not the CDC accounted for outliers in their calculation of heavy drinker status, but

Reducing the incidence of stroke is an important goal of public health that would result in a vast reduction of the health care burden. Our findings provide researchers and policymakers a set of variables that should be further investigated in order to elucidate causative relationships that may occur between these risk factors and stroke outcome. Future studies should look further at the effect of mental health diagnosis on stroke odds, as well as how smoking can increase an individual’s odds of a stroke.

## References

- American Psychological Association (2015). Paying with our health. *Stress in America*, pages 1–23.
- American Stroke Association (2017a). Impact of Stroke (Stroke statistics).
- American Stroke Association (2017b). Understanding Stroke Risk.
- Anxiety and Depression Association of America (2017). Facts & Statistics | Anxiety and Depression Association of America, ADAA.
- Centers for Disease Control and Prevention (2007). Behavioral Risk Factor Surveillance System Survey (BRFSS) data. 2009(9/8/2008).
- Centers for Disease Control and Prevention (2014). CDC - About BRFSS.
- Centers for Disease Control and Prevention (2015). Insufficient Sleep Is a Public Health Problem | Features | CDC.
- Fricker, R. D. and Andersen, L. (1993). Raking: An important and often overlooked survey tool. Technical report.
- Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., De Ferranti, S., Després, J. P., Fullerton, H. J., Howard, V. J., Huffman, M. D., Judd, S. E., Kissela, B. M., Lackland, D. T., Lichtman, J. H., Lisabeth, L. D., Liu, S., Mackey, R. H., Matchar, D. B., McGuire, D. K., Mohler, E. R., Moy, C. S., Muntner, P., Mussolino, M. E., Nasir, K., Neumar, R. W., Nichol, G., Palaniappan, L., Pandey, D. K., Reeves, M. J., Rodriguez, C. J., Sorlie, P. D., Stein, J., Towfighi, A., Turan, T. N., Virani, S. S., Willey, J. Z., Woo, D., Yeh, R. W., and Turner, M. B. (2015). *Heart disease and stroke statistics-2015 update : A report from the American Heart Association*, volume 131.

## Code