

The influencing factors for depression in US adults: Data from the NHANES(2009–2018)

Chong Xu

Wenzhou Kean University, Wenzhou, China, 310000,

xucho@kean.edu

Abstract. In today's world, as society puts more and more pressure on people, the incidence of depression is also increasing. Depression can have a number of serious harms, not only affecting an individual's emotional and psychological state, but also having a serious impact on quality of life, work and personal relationships. To find the causes of depression, this paper will analyze the contributing factors of depression by analyzing data from U.S. adults in NHANES(2009-2018), a series of cross-sectional, complex, multi-stage surveys conducted by the CDC(Centers for Disease Control and Prevention). Some variables were obtained from the questionnaire -- sleep status and smoking status. The results included age, sex, education, marital status, smoking status, obesity, and some of the diseases associated with CRD. This study found that sleep disorders had the largest OR value, Education level and smoking status have the least relationship with CRD because of the lowest OR value.

Keywords: depression, pressure, sleep, Arthritis, coronary heart disease.

1. Introduction

Depression is more than just sadness; it is a mood disorder that can cause lasting feelings of sadness and loss of interest [1], accompanied by changes in somatic and cognitive aspects that significantly impact individuals' functioning [2]. In recent years, the prevalence of depression has been increasing globally, including in the United States. Understanding the factors that contribute to the development of depression is crucial for effective prevention and treatment strategies.

In the US, the prevalence of depression has reached alarming levels. From 2015 to 2020, there has been a rising trend in depression rates, with a prevalence of over 9% in 2020. This poses a significant public health concern, as depression not only causes personal distress but also has wide-ranging societal implications.

Depression has a profound impact that extends beyond individual suffering. It affects relationships, work productivity, and overall quality of life. Individuals with depression often experience difficulties in maintaining relationships, impairments in occupational functioning, increased healthcare utilization, and higher rates of mortality. Additionally, depression is associated with an increased risk of developing other physical and mental health conditions, such as anxiety disorders, substance abuse, and cardiovascular diseases.

Contrary to popular belief, depression is not solely caused by a chemical imbalance in the brain. Research has shown that it is influenced by a combination of factors, including mood regulation, genetic

vulnerability, and life events [3]. Similarly, the relationship between anxiety, depression, and sleep quality is not solely due to the feelings of depression and anxiety but also has a genetic component [4].

The purpose of this paper is to systematically review and analyze the various factors that contribute to the development of depression in adults in the United States. By doing so, we aim to contribute to the growing body of knowledge on depression and provide insights into effective prevention and intervention strategies.

2. Methods

The data is from the NHANES, a series of cross-sectional, complex, multi-stage surveys conducted by the CDC(Centers for Disease Control and Prevention)[5]. The NHANES survey combines interviews and physical examinations. All the data are from 2009 to 2018, and the research sample is aged from 20 to 79 years old. R 4.2.2, and Excel 2016 are used to analyze our data. The variables involved are age, gender, and BMI, which is conducted by weight divided by height square. Education level is divided into three categories: college graduate or above, high school or below. It was divided in this level because people from different educational levels have different kinds of pressure. Some variables were obtained from the questionnaire: Sleeping status, Smoking status. As these two are related with daily performance in both work and daily life. Sleeping status was measured by asking the behind questions in the questionnaire: Ever told a doctor had trouble sleeping? Smoking status was asked by Have you smoked at least 100 cigarettes in your entire life? Arthritis and Coronary heart disease are measured by doctors' diagnoses. And some disease information, in our project, we only explore two diseases, Arthritis, and coronary heart disease. The mathematical method we used is the Chi-square test, t-test, and adjusted logistic regression analysis.

3. Results

In logistic regression, three models are set up: Model 1, Adjusted for gender and age; Model 2, Adjusted for all the collected factors; Model 3, Adjusted for age, gender, Education level, Marital status, Smoking status, BMI, and Arthritis. And Depression was measured using the Patient Health Questionnaire (PHQ-9), a nine-item screening instrument that asked questions about the frequency of symptoms of depression over the past 2 weeks. Response categories for the nine-item instrument "not at all," "several days," "more than half the days" and "nearly every day" were given a point ranging from 0 to 3. According to the answer to each question, we can calculate the whole score of this questionnaire. A total score ≥ 10 was considered to be clinically relevant depression (CRD).

Table 1. Baseline table.

Characteristic	Having sleeping trouble					
	level	overall	No	Yes	p-value	test
a		24804	18159	6638		
age(mean(SD))		49.8(17.7)	48.6(18.0)	53.1(16.3)	<0.01	t-test
education(N)	collage	7592(30.6)	5341(29.4)	2250(33.9)	<0.01	Chi-squared test
	graduate or above	5972(24.1)	4521(24.9)	1448(21.8)		
	High school or below	11218(45.3)	8278(45.6)	2938(44.3)		
Marital(h)	77	15(0.1)	11(0.1)	4(0.1)	<0.01	Chi-squared test
	99	1(0.0)	0(0.0)	1(0.0)		
	coupled	14596(58.8)	10993(60.5)	3601(54.2)		
	Single or separated	10192(41.1)	7155(39.4)	3032(45.7)		
gender(N)	Male	12194(49.2)	9416(51.9)	2775(41.8)	<0.01	Chi-squared test
	Female	12610(50.8)	8743(48.1)	3863(58.2)		

Table 1. (continued).

Medium score(mean(SD))		3.4(4.7)	2.5(3.9)	5.7(5.7)	<0.01	t-test
Smoking status(N)	No	13870(55.9)	10775(59.4)	3090(46.6)	<0.01	Chi-squared test
	Yes	10920(44.1)	7374(40.6)	3545(53.4)		
BMI(mean(SD))		29.4(7.1)	28.9(6.7)	30.8(8.0)	<0.01	t-test
Arthritis(N)	NA	9113(36.8)	6834(37.7)	2278(34.4)	<0.01	Chi-squared test
	No	11275(45.5)	8930(49.2)	2343(35.4)		
	Yes	4382(17.7)	2374(13.1)	2004(30.2)		
Coronary heart disease(N)	NA	9113(36.8)	6834(37.7)	2278(34.4)	<0.01	Chi-squared test
	No	14971(60.5)	10923(60.3)	4043(61.1)		
	Yes	668(2.7)	369(2.0)	298(4.5)		
Race(N)	Mexican American	2636(13.5)	2097(14.8)	539(10.2)	<0.01	Chi-squared test
	Hispanic	2003(10.3)	1500(10.6)	503(9.5)		
	White	7397(37.9)	4899(34.5)	2497(47.1)		
	Black	4432(22.7)	3269(23.0)	1160(21.9)		
	Asian	2296(11.8)	1956(13.8)	339(6.4)		
	Other	730(3.7)	470(3.3)	259(4.9)		

This is the baseline table. It is found that of all Of the 24804 subjects (49.2% males and 50.8% females, mean (SD) age 49.8 [17.7] years), the p-value of different Age, Education level, Marital status, gender, Smoking status, BMI, Race, and disease groups are less than 0.05, which shows that they are significantly related to the Sleeping disorder.

3 models are designed for logistic regression. The result is shown in Tables 2-4.

Table 2. The result of logistic regression of Model 1.

Var	OR	p-value	OR1	OR2	OR mean
Trouble sleeping	4.54(4.16~4.97)	<0.0001	4.156367965	4.96814356	4.543327991
Age	1(0.99~1)	0.003	0.993613365	0.998728191	0.996169243
Gender	1.5(1.38~1.65)	<0.0001	1.375374458	1.646930581	1.504687968

Table 3. The result of logistic regression of Model 2.

Var	OR	p-value	OR1	OR2	OR mean
Trouble sleeping	4.24(3.75~4.8)	<0.001	3.753297522	4.795376581	4.240895346
Age	0.99(0.99~1)	<0.001	0.987443405	0.995181901	0.99131623
Educational level	0.74(0.7~0.78)	<0.001	0.70256007	0.775195789	0.738015144
Marital status	1.04(1.02~1.06)	<0.001	1.020754063	1.057893579	1.039468537
Gender	1.57(1.39~1.78)	<0.001	1.393657078	1.781328587	1.574997958
Smoking status	0.61(0.54~0.68)	<0.001	0.536213553	0.682885507	0.605299766
Race	1(0.96~1.04)	0.868	0.957214624	1.037177098	0.996601535
BMI	1.02(1.01~1.02)	<0.001	1.009172704	1.024012526	1.016602483
Coronary heart disease	0.98(0.86~1.1)	0.741	0.859821311	1.096823652	0.979865967
Arthritis	0.65(0.57~0.75)	<0.001	0.571773244	0.745065927	0.653733091

Table 4.The result of logistic regression of Model 3.

Var	OR	p-value	OR1	OR2
Trouble sleeping	4.23(3.75~4.77)	<0.001	3.753750984	4.773417895
Age	0.99(0.99~0.99)	<0.001	0.987069512	0.99465139
Education	0.74(0.7~0.77)	<0.001	0.701860059	0.770532637
Marital	1.04(1.02~1.06)	<0.001	1.022175482	1.059048497
Gender	1.58(1.4~1.78)	<0.001	1.397945478	1.778391305
Smoking status	0.61(0.54~0.69)	<0.001	0.541177175	0.685988057
BMI	1.02(1.01~1.02)	<0.001	1.009262954	1.023840267
Arthritis	0.65(0.57~0.74)	<0.001	0.572713798	0.742980384

Model 1: is adjusted for age and gender, (OR : 4.54 (4.16~4.97)). Model 2 is adjusted for age, gender, Education level, Marital status, Smoking status, Race, BMI, and two diseases (ACD and Arthritis) (OR : 4.24 (3.75~4.8)). Model 3 removed variables in model 2 where p-value is greater than 0.05, Adjusted for age, gender, Education level, Marital status, Smoking status, BMI, and Arthritis (OR : 4.23 (3.75~4.77)).

Since in the background information, it is learned that depression is significantly related to gender and age differences, the first model only takes gender and age as covariates to examine the relationship between depression and sleeping trouble. The OR value (confidence intervals) of sleeping trouble is 4.54 (4.16-4.97), age is 1 (0.99-1), gender is 1.5 (1.38-1.65), and all the p-values are less than 0.05. In the baseline table, it is learned that all variables had significant differences in the presence or absence of sleep trouble, so all variables are tried to set as covariates and consider their relationship with depression. In the second model, the OR value (confidence interval) of trouble sleeping and depression has dropped to 4.24 (3.75-4.8). Nearly all the variables' p-value is less than 0.001, except race and Coronary heart disease. Lastly, due to the variables' p-value being larger than 0.05, these two variables are removed, not setting them as covariates, and again examine the relationship between the remaining eight variables and the depression. The OR value (confidence interval) of trouble sleeping has dropped to 4.23 (3.75-4.77).

Then R's forest plot package is used to draw a forest plot. The numbers show the OR value and its confidence interval. And the most right is the p-value. As shown in fig4, in the age- and gender-adjusted model (model 1), sleeping trouble (OR 4.54, 95% CI: 4.16-4.97) was substantially related to CRD. In model 2 (OR 4.24, 95% CI: 3.75-4.8) and model 3 (OR 4.23, 95% CI: 3.75-4.77), the sleeping trouble and CRD remained significant after adjusting for potential confounding factors.

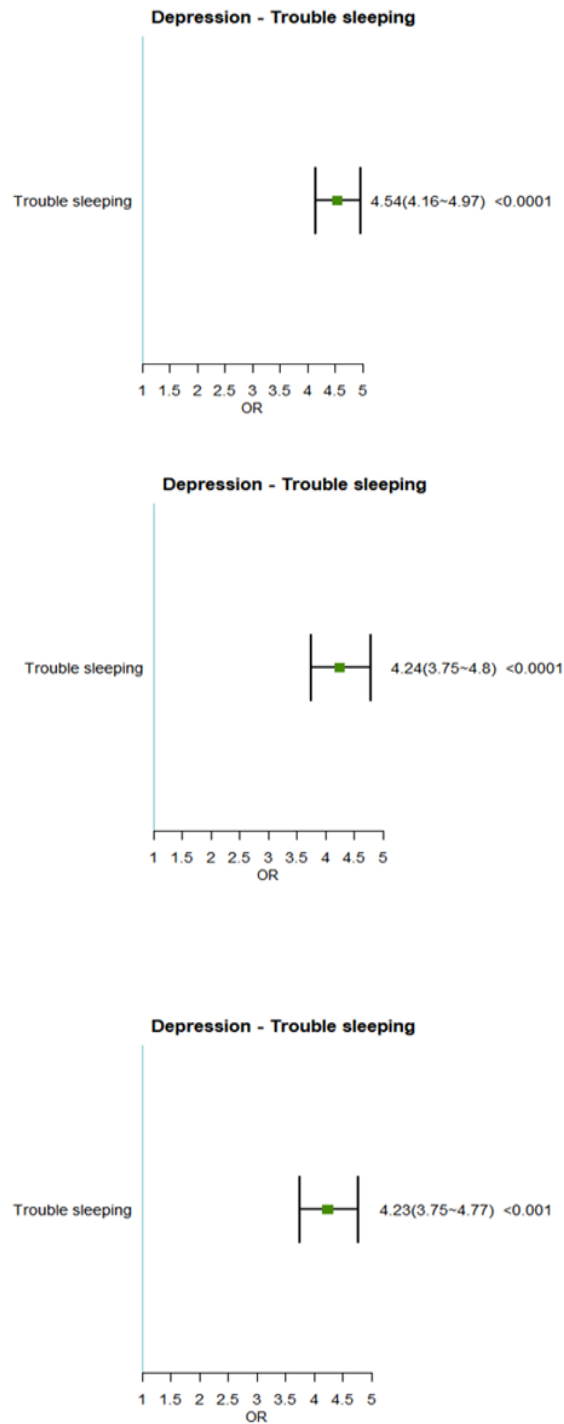


Figure 1. Forest plot.

4. Discussion

In the research data is searched from enhances, maybe more data can be found from different resources, and we can also minimize the scale of the sample. 49000 data are collected after the detection of NA, there are only 25000. More data from different resources can be found and more samples can be collected. In the aspect of the sample, a more specific sample range can be decided, for example, all the students

in a university. Dinis et al did their research on depression and the quality of sleep in college students [6], a suitable group can also be found to improve the research.

We drew the baseline table of our data and use logistics to analyze it. To improve our analysis, more advanced methods can be used, for example, hypothesis testing and a better-adjusted logistic model. More disease information and better measurement of sleeping trouble/sleeping status will make our study stricter. Other studies have shown that depression is associated with many other diseases, more data can be collected from other diseases besides the two diseases we analyzed. Also, more specific criteria for sleeping trouble can be found, the question mentioned in NHANES is very subjective, if an objective one can be found then our results will be more referential. Baglioni et al did their research on the relationship between insomnia and depression, and they collected their data from the previous study done by others [7]. In that case, more diseases that are suitable for research can be found and researched and data from previous studies can be found. There is a study which did research on the association between depression and neurological changes. It concluded that the depression influence the neuro to be less active and get down the quality of sleep [8].

5. Conclusion

The self-reported trouble sleeping was significantly related to CRD (OR: 4.23, 95% CI: 3.75-4.77). There are a lot of factors related to CRD, in the study, age, gender, education level, marital status, smoking status, obesity degree and also some diseases are confirmed to have an association with CRD. However, of all the factors that are analyzed, Sleeping disorder has the biggest OR value, which means it is most related to CRD. The more sleeping disorder the patients have, the more CRD patients have. Age, martial status and gender have relationship with CRD but the relationship is related to different groups in these two conditions. Educational level and smoking status have the least relationship with CRD because of the lowest OR value.

References

- [1] Kanter et al. (2008). The Nature of Clinical Depression: Symptoms, Syndromes, and Behavior Analysis. *Behav Anal*, pp. 1-21.
- [2] Chand et al. (2022). Depression. <https://www.ncbi.nlm.nih.gov/books/NBK430847/>.
- [3] Harvard Medical School. (2022). What causes depression. <https://www.health.harvard.edu/mind-and-mood/what-causes-depression>.
- [4] Gregory et al. (2011). Associations between sleep quality and anxiety and depression symptoms in a sample of young adult twins and siblings. *J Psychosom Res*, pp.250-255
- [5] Li Chunnan, Shang Shaomei, Liang Wannian. (2022). The association between sleep and depressive symptoms in US adults: data from the NHANES (2007-2014). *Epidemiol Psychiatr Sci*.
- [6] Dinis et al. (2018). Quality of Sleep and Depression in College Students: A Systematic Review. *Sleep Sci*, pp.290-301.
- [7] Baglioni et al. (2011). Clinical implications of the causal relationship between insomnia and depression: how individually tailored treatment of sleeping difficulties could prevent the onset of depression. *EPMA J*, pp. 287-293.
- [8] Adrien, J. (2002). Neurobiological bases for the relation between sleep and depression. *Sleep medicine reviews*, 6(5), 341-351.