Prevalence of Depression Among U.S. Adults ≥ 20 Years Old Across Different Education Levels, Data from the NHANES (2017-Prepandemic March 2020)

School of Global Public Health, New York University, New York, NY, USA

Spring 2024, GPH-GU 2450 Intermediate Epidemiology

Helen Liang

Target Population: U.S. adults ≥ 20 years old

Research Question: To use the 2017-prepandemic March 2020 NHANES data to analyze the association between education level and depression among U.S. adults ≥ 20 years old while controlling for age, race, and health insurance coverage, and whether this relationship is modified by gender.

Hypothesis: In the target population, individuals with lower education level will be more likely to experience depressive symptomology, controlling for age, race, and health insurance coverage. In addition, we expect the relationship between education level and depressive symptomology will be modified by gender, with females with lower education level more likely to experience depressive symptomology compared to males with lower education level.

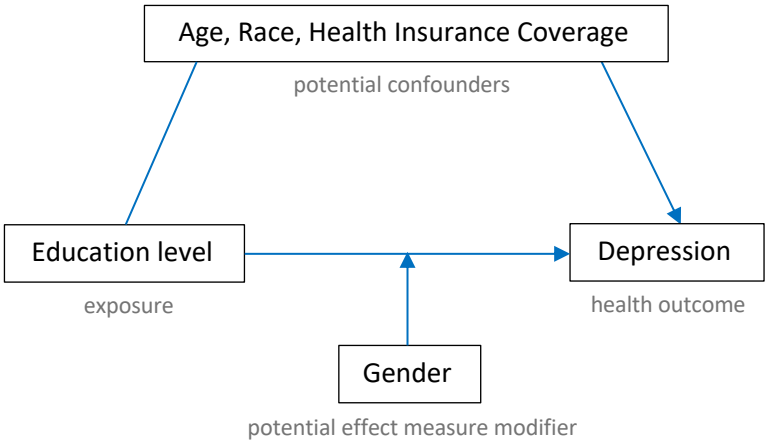


FIGURE 1. Directed acyclic graph of the study.

Abstract

Introduction: Depression is a prevalent psychiatric disorder affecting approximately 20% of U.S. adults, posing significant public health and economic burdens.

Objective: This study aims to analyze the association between education level and depression among U.S. adults ≥ 20 years old, using the NHANES 2017-prepandemic March 2020 data.

Methods: A cross-sectional design was utilized with N=7,812. Depression was assessed using the Patient Health Questionnaire (PHQ-9). Univariable, bivariable, and multivariable logistic and linear regression analyses were used to examine the association between education level and depression, controlling for confounders such as age, race/ethnicity, and health insurance coverage, and exploring the effect measure modification by gender.

Results: Higher education levels were significantly associated with lower odds and scores of depression, particularly for college graduates or above (OR ≈ 0.40-0.45, β ≈ -0.88 to -0.79). Lower education levels, lower income, and certain demographic characteristics (e.g., being female, younger age, being single, being a part of the Other Race – Including Multi-Racial racial/ethnic group) significantly increased the odds and scores of depression. Health insurance coverage did not have a confounding effect.

Conclusion: There is an inverse relationship between education level and depression among U.S. adults ≥ 20 years old. These findings highlight the importance of education in mitigating depression and suggest that policies aimed at increasing access to education should be implemented to reduce the burden of depression, thus improving overall public health. Future longitudinal studies are recommended to confirm these findings and explore the causal relationship between education level and depression.

Keywords: depression, depression scale, PHQ-9, education level, NHANES, bias, validity

Introduction/Background

Major depressive disorder (MDD), or depression, is one of the most common and serious psychiatric disorders of the 21st century affecting the health of approximately one in five U.S. adults both physically and psychologically and poses a significant public health concern both globally and in the United States due to its increasing prevalence, resulting in increased

disability, morbidity, mortality, and socioeconomic burden (Hao et al., 2019; Li et al., 2022; Patel et al., 2019). According to Hao et al. (2019), it is said that "a mild depression episode manifests as sadness, anhedonia, and a feeling of worthlessness, whereas the severe condition is classified by a recurring intention to commit suicide" and has contributed to 13.3% to 17.1% of the incidence rate in the United States. As the incidence rate of depression increases, research has been carried out investigating its pathological mechanism, therapeutic effect, and relationship with other factors (Hao et al., 2019).

The 2017-prepandemic March 2020 National Health and Nutrition Examination Survey (NHANES) administered by the Centers for Disease Control and Prevention (CDC) provided data showing the prevalence of depression from analyzing the national, state-level, and county-level prevalence of U.S. adults ≥ 18 years old who self-reported depression or experiencing depressive symptomology, which will be used to analyze the association between education level and depression among U.S. adults ≥ 20 years old while controlling for age, race, and health insurance coverage, and whether this relationship is modified by gender (Lee et al., 2023).

The NHANES data collected information from 392,746 individuals and categorized them into several groups according to age, gender, race and ethnicity, education level, etc. Point estimates and 95% confidence intervals (CI) are used to estimate the prevalence of depression in each group. Both national and state-level data were directly estimated and collected from the weighted BRFSS 2020 data whereas multilevel logistic regression and post-stratification were used for county-level estimates. The analysis revealed an overall estimated prevalence of depression of 18.4% among the entire sample size according to the 95% CI. In addition, it revealed a stratified burden of depression, with higher prevalence in younger age groups (21.5% for the 18-24 years group, 19.9% for the 25-44 years group, 18.4% for the 45-64 years group, and 14.2% for the ≥ 65 years group) and women (13.1% for men and 23.4% for women), regardless of race and ethnicity (23.3% for AI/AN (American Indian or Alaska Native), non-Hispanic, 7.6% for Asian, non-Hispanic, 16.1% for Black or African American, non-Hispanic, 15.1% for NH/OPI (Native Hawaiian or other Pacific Islander), non-Hispanic, 20.6% for White, non-Hispanic, 14.6% for Hispanic or Latino, 28.5% for Multiracial, non-Hispanic, and 19.3% for Other, non-Hispanic). The data further demonstrated an inverse relationship between

education level and depression (21.0% for those whose education level was identified as “Less than high school”, 18.1% for “High school or equivalent”, 21.0% for those identified as “Technical college or some college”, and 14.9% for those identified as “College degree or higher”), which led to the question that higher education level could act as a protective factor or buffer against depression (Lee et al., 2023; Taple et al., 2020).

Studies have been done on examining the association between education level and depression while adjusting for covariates (e.g., confounders, effect measure modifiers, etc.) that could affect the relationship, which revealed a decreased prevalence of depression as education level increases (Samudio-Cruz et al., 2023). According to Kondirolli et al. (2022), “an extra year of education led to a lower likelihood of reporting any symptoms related to depression (11.3%)” and that “more educated people also suffered less severe symptoms – depression (6.1%)”. Another study has shown that increased education level has a higher impact on decreased depression for women compared to men. In addition, women who experience depression are more likely to drop out of high school, not enroll in college, and are less likely to enroll in a 4-year college (Taple et al., 2020; Li et al., 2022). Furthermore, individuals with lower levels of education face socioeconomic challenges, including poorer access to healthcare such as lack of access to screening and treatment for depression, resulting in increased vulnerability to experiencing depression or depressive symptoms (Taple et al., 2020).

The objective of this study is to analyze the association between education level and depression among U.S. adults ≥ 20 years old using the NHANES data while referencing existing studies, identify additional risk factors associated with depression, and ultimately develop and inform the public with targeted interventions that highlight the influence of education level on physical and psychological health. We hypothesize that individuals with lower education levels will be more likely to experience depressive symptomology, controlling for age, race, and health insurance coverage (Figure 1). In addition, we expect the relationship between education level and depressive symptomology will be modified by gender, with females with lower education levels more likely to experience depressive symptomology compared to males with lower education levels.

Methods

The 2017-prepandemic March 2020 National Health and Nutrition Examination Survey (NHANES) administered by the Centers for Disease Control and Prevention (CDC) aims "to provide prevalence data on selected diseases and risk factors, to monitor trends in selected diseases, behaviors, and environmental exposures, to explore emerging public health needs, and to maintain a national probability sample of baseline information and nutritional status" for the period 2017-prepandemic March 2020 on noninstitutionalized civilian U.S. population. NHANES utilizes a cross-sectional, multistage and multiyear, stratified, clustered probability-based sampling design with its sample selection categorized into four stages: (1) primary sampling units (PSU) such as counties, (2) segments within PSUs that contain a cluster of households, (3) households within segments, and (4) individuals within a household. Stratification is used in sample selection to ensure the sample is evenly distributed among characteristics such as overall health index, geography, urban-rural distribution, and population demographics (National Center for Health Statistics, 2013).

The purpose of the NHANES survey is to collect data on the health, nutritional status, and health behaviors of the noninstitutionalized civilian resident population of the United States. Data is collected through a combination of personal interviews with standardized physical examinations and laboratory tests. Based on the NHANES data provided, the sample size or the number of individuals who completed the survey is N=15,560. The target population for this study is those who are ≥ 20 years old, which is N=9,232. However, after conducting an initial overview of the dataset, we have determined that the final analytic sample size is N=7,812, by dropping all missing values ("Refused" and "Don't know" responses) from exposure, covariates, and sociodemographic variables.

According to the NHANES survey dataset, depression or depressive symptoms are measured using the Patient Health Questionnaire (PHQ)-9, a self-administered questionnaire evaluating depression severity that is demonstrated to be "acceptable to use in major U.S. sociodemographic groups and allows for meaningful comparisons in total, cognitive/affective, and somatic depressive symptoms across these groups" (Patel et al., 2019). Participants are assessed by how often they have experienced 9 depressive symptoms over the last 2 weeks at

the time of the survey. These 9 symptoms are having “little interest in doing things”, “feeling down, depressed, or hopeless”, “trouble sleeping or sleeping too much”, “feeling tired or having little energy”, having a “poor appetite or overeating”, “feeling bad about yourself”, having “trouble concentrating on things”, “moving or speaking slowly or too fast”, and having “thoughts you would be better off dead” and are labeled as “DPQ010” – “DPQ090” in the NHANES dataset. Each of these 9 items is scored from 0-3, with 0 representing “not at all”, 1 representing “several days”, 2 representing “more than half the days”, and 3 representing “nearly every day”. “Refused” (coded as “7”) and “Don’t know” (coded as “9”) responses are recoded as missing values (coded as “.”) and excluded from the analysis. The total score of the depression questionnaire ranges from 0-27, with 0-9 indicating “no depression” and 10-27 indicating “depression” (Kroenke et al., 2001; Ettman et al., 2020). The same scale will be employed when performing statistical analysis to explore the association between education level and depression while controlling for confounders and effect measure modifiers.

As stated previously, the final analytic sample size is determined based on the results of the variable “DMDEDU2”, representing the education level of adults ≥ 20 years old, while excluding all missing values from variables that were previously stated. Participants are asked “What is the highest grade or level of school you have completed or the highest degree you have received?” and are given 5 answer choices to choose from. These answer choices are “less than 9^{th} grade” (coded as “1”), “9-11^{th} grade (includes 12^{th} grade with no diploma)” (coded as “2”), “high school graduate/GED or equivalent” (coded as “3”), “some college or AA degree” (coded as “4”), and “college graduate or above” (coded as “5”). “Refused” (coded as “7”) and “Don’t know” (coded as “9”) responses are recoded as missing values (coded as “.”) and excluded from the analysis.

The confounders in this study are age, race, and health insurance coverage and the effect measure modifier is gender. In the NHANES dataset, age in years at screening (labeled as “RIDAGEYR”) is initially coded from 0-79 indicating the participant’s age at the time of the screening, and 80 indicating those who are 80 years old and above. However, age will be recoded with “1” indicating “20-29 (years)”, “2” indicating “30-39 (years)”, “3” indicating “40-49 (years)”, “4” indicating “50-59 (years)”, “5” indicating “60-69 (years)”, “6” indicating “70-79

(years)", and "7" indicating "≥ 80 (years)" in this study. Race is initially defined by two variables, race/Hispanic origin (labeled as "RIDRETH1") and race/Hispanic origin w/ NH Asian (labeled as "RIDRETH3"). Both race variables have identical codes and code descriptions for "1" indicating "Mexican American", "2" indicating "Other Hispanic", "3" indicating "Non-Hispanic White", and "4" indicating "Non-Hispanic Black". The ones that differ are "5" indicating "Other Race – Including Multi-Racial" for race/Hispanic origin, "6" indicating "Non-Hispanic Asian", and "7" indicating "Other Race – Including Multi-Racial" for race/Hispanic origin w/ NH Asian. Both race variables will be merged into a single variable and recoded with "1" indicating "Mexican American", "2" indicating "Other Hispanic", "3" indicating "Non-Hispanic White", "4" indicating "Non-Hispanic Black", "5" indicating "Non-Hispanic Asian", and "6" indicating "Other Race – Including Multi-Racial" in this study.

Health insurance coverage (labeled as "HIQ011") is a categorical variable and is coded as "1" indicating "yes" and "2" indicating "no". "Refused" (coded as "7") and "Don't know" (coded as "9") responses are recoded as missing values (coded as ".") and excluded from the analysis. Gender (labeled as "RIAGENDR") is also a categorical variable and is coded as "1" indicating "male" and "2" indicating "female".

StataSE Version 18.0 will be used to perform statistical analysis on the data. For univariable analysis, frequencies and percentages will be computed for categorical variables such as race, gender, and health insurance coverage. Mean, standard deviation, median, and IQR or range will be computed for continuous/ordinal variables such as education level, depression, and age. For bivariable analysis, odds ratios will be computed to evaluate the associations between the exposure and outcome. In addition, data will be stratified by gender to determine effect measure modification, and crude and adjusted ORs will be used to determine confounding. ANOVA will be used to compare mean depression scores across different levels of education, utilizing the 5% significance level ($p < 0.05$) to determine the statistical significance between and within groups. For multivariable analysis, both a univariate and multivariate logistic regression will be used to assess the association between education level and depression with adjustments for covariates. For all variables with "Refused" and

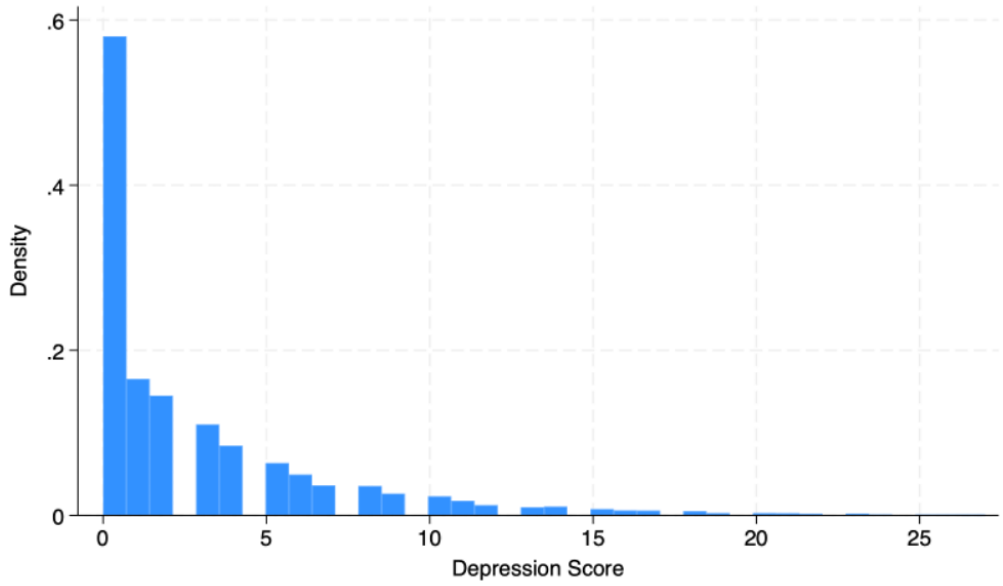
"Don't know" responses, these will be recoded as missing values and excluded from the data analysis.

Results

Univariable Analysis

Table 1 displayed the univariable analysis of the association between education level and depression among U.S. adults ≥ 20 years old. The analytic sample size consists of 48.39% male and 51.61% female participants. Age is distributed across seven age categories, with 14.90% of participants between 20-29 years old, 15.53% between 30-39 years old, 15,77% between 40-49 years old, 16.95% between 50-59 years old, 18.89% between 60-69 years old, 10.69% between 70-79 years old, and 7.27% of participants 80 years old and above. Race/ethnicity consists of 10.94% of participants identified as Mexican Americans, 9.66% identified as Other Hispanic, 36.67% identified as Non-Hispanic White, 26.02% identified as Non-Hispanic Black, 11.80% identified as Non-Hispanic Asian, and 4.89% identified as Other Race – Including Multi-Racial. Education level consists of 6.87% of participants with an education level less than 9^{th} grade, 10.77% with 9-11^{th} grade (includes 12^{th} grade with no diploma, 24.05% with a high school graduate/GED or equivalent, 33.21% with some college or AA degree, and 25.10% with college graduate or above. Health insurance coverage consists of 84.91% of participants with health insurance and 15.09% without. Some additional sociodemographic characteristics, such as marital status and the ratio of family income to poverty, are included in Table 1 as a part of this univariable analysis to better describe the study sample but are not covariates of this study. Depression is presented in the form of both a continuous and a categorical variable. Depression score ranges from 0-27 with a mean (SD) of 2.88 (4.12) and median (IQR) of 1.00 (0,4) (Figure 2). The categorical results consist of 7.95% of participants with and 92.05% without depression or depressive symptoms.

Figure 2. Distribution of depression score among adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.



Bivariable Analyses

Table 2 presented bivariable associations between education level and depression, key sociodemographic characteristics, and other covariates. There is a significant association between the main exposure and outcome with $p < 0.001$, where individuals with higher education levels tend to have a lower prevalence of experiencing depression or depressive symptoms. For example, 9.30% of individuals with an education level less than 9th grade reported depression compared to only 3.98% of those with a college degree or higher. Furthermore, Table 2 shows that education level is strongly associated with all other variables – gender, age, race/ethnicity, marital status, the ratio of family income to poverty, and health insurance coverage, all showing $p < 0.001$, suggesting there may be additional confounders associated with this study that were not mentioned previously, such as marital status and income ratios.

Tables 3a and 3b presented bivariable associations between depression as both a binary and continuous variable and education level, key sociodemographic characteristics, and other covariates. Results of both Tables 3a and 3b demonstrated a significant association between education level and depression with $p < 0.001$. In addition, Table 3a showed a higher

prevalence of depression among females with a lower education level (62.48%) than males (37.52%) and Table 3b showed a mean (SD) depression score of 2.88 (4.12), with females scoring higher on average (3.29 (4.36)) compared to males (2.45 (3.81)), suggesting that gender may be a potential effect measure modifier. Furthermore, the analyses from Tables 3a and 3b demonstrated that depression is strongly associated with other variables such as gender, age, marital status, the ratio of family income to poverty, and health insurance coverage, all showing $p < 0.05$, suggesting there may be additional confounders associated with this study that were not mentioned previously, such as marital status and income ratios. It should also be noted that race/ethnicity is significantly associated with depression as a categorical variable but not as a continuous variable, whereas health insurance coverage is significantly associated with depression as a continuous variable but not as a categorical variable, which could be due to differences in the statistical methods used to assess these associations.

Multivariable Analysis

Table 4a presented the results of a multivariable logistic regression model examining education level and depression as a categorical variable. The results demonstrated that higher education levels are associated with lower odds of depression, particularly for college graduates or above ($OR \approx 0.40\text{-}0.45$). In addition, females have higher odds for depression compared to males ($OR \approx 1.55\text{-}1.62$). Depression decreases with age, significantly for those aged ≥ 80 ($OR \approx 0.45\text{-}0.61$). Non-Hispanic Asians have lower odds ($OR \approx 0.44\text{-}0.66$), while Other Race – Including Multi-Racial have higher odds ($OR \approx 2.15\text{-}2.50$) compared to Mexican Americans. Widowed/Divorced/Separated and Never Married groups have higher odds compared to Married/Living with a Partner ($OR \approx 1.48\text{-}1.83$). Higher income (> 1) is associated with lower odds of depression ($OR \approx 0.47\text{-}0.63$). However, no significant difference is observed between those with and without health insurance.

Table 4b presented the results of a multivariable linear regression models examining education level and depression as a continuous variable. Similar to Table 4a, the results demonstrated that higher education levels are associated with lower scores of depression, particularly for college graduates or above ($\beta \approx -0.88$ to -0.79). In addition, females have higher scores for depression compared to males ($\beta \approx 0.77\text{-}0.86$). Depression decreases with age,

significantly for those aged ≥ 80 (β ≈ -0.96 to -1.35). Non-Hispanic Asians have lower scores (β ≈ -1.02 to -0.47), while Other Race – Including Multi-Racial have higher scores (β ≈ 1.31-1.47) compared to Mexican Americans. Widowed/Divorced/Separated and Never Married groups have higher scores compared to Married/Living with a Partner (β ≈ 0.77-0.96). Higher income (> 1) is associated with lower scores of depression (β ≈ -1.22 to -0.79). However, no significant difference is observed between those with and without health insurance.

Both Tables 4a and 4b demonstrated that lower education levels, lower income, and certain demographic characteristics (e.g., being female, younger age, being single, being a part of the Other Race – Including Multi-Racial racial/ethnic group) significantly increased the odds and scores of depression, indicating confounding effects from age and race/ethnicity, and effect measure modification from gender. However, health insurance coverage does not have a confounding effect in the study.

Discussion

Main Findings

This study analyzed the association between education level and depression among U.S. adults ≥ 20 years old using the NHANES 2017-prepandemic March 2020 data. The findings indicated that higher education levels are significantly associated with lower odds and scores of depression, particularly for those with a college degree or above. This inverse relationship suggests that increased education acts as a protective factor against depression. Furthermore, females, younger age groups, individuals that are single, those with lower education levels, and having lower income exhibited higher odds and scores of depression compared to males, older age groups, individuals that are married/living with a partner, those with higher education levels, and having higher income. Notably, there was no significant difference in depression between those with and without health insurance. These findings are consistent with previous studies that have demonstrated an inverse relationship between education level and depression. For example, Li et al. (2022) and Taple et al. (2020) found that increased education levels reduce the prevalence of depression, particularly among women. In addition, this study supports the existing literature indicating that females are more likely to experience depression than males, and that depression decreases with age. The findings regarding the higher

prevalence of depression among individuals of Other Race – Including Multi-Racial, and lower prevalence among Non-Hispanic Asians, are consistent with the results of Hao et al. (2019) and Lee et al. (2023).

Strengths

The strengths of this study include the use of a large, nationally representative sample from the NHANES survey which improves generalizability. The multistage, stratified, clustered probability-based sampling design of NHANES ensures a diverse and well-distributed sample across various demographic characteristics. Furthermore, the use of PHQ-9, a validated tool for assessing depression, and detailed sociodemographic data allow for robust analysis and control for confounding variables.

Limitations

Some limitations of this study are that a causal relationship between education level and depression cannot be established since it uses a cross-sectional design, the reliance on self-reported data may introduce information bias due to overreporting or underreporting of depressive symptoms, and the exclusion of participants with missing data could introduce selection bias, potentially affecting the internal validity. Selection and information bias may limit both the internal and external validity of this study and affecting the measures of association such as overestimating or underestimating the true relationship between education level and depression. To minimize the impact of these biases, we controlled for confounders (e.g., age, race/ethnicity, and health insurance coverage) and performed stratification by gender to assess effect measure modification.

Conclusion

In conclusion, there is an inverse relationship between education level and depression among U.S. adults ≥ 20 years old. These findings highlight the importance of education in mitigating depression and suggest that policies aimed at increasing access to education should be implemented to reduce the burden of depression, thus improving overall public health. Future longitudinal studies are recommended to confirm these findings and explore the causal relationship between education level and depression.

References

1. Hao Y., Ge H., Sun M., Gao Y. (2019). Selecting an Appropriate Animal Model of Depression. *International Journal of Molecular Sciences*. 2019;20(19):4827. https://doi.org/10.3390/ijms20194827
2. Li L., Sun W., Luo J., Huang H. (2022). Associations between education levels and prevalence of depressive symptoms: NHANES (2005–2018). *Journal of Affective Disorders*. 2022;301:360-367. https://doi.org/10.1016/j.jad.2022.01.010
3. Patel J. S., Oh Y., Rand K. L., Wu W., Cyders M.A., Kroenke K., Stewart J. C. (2019). Measurement invariance of the patient health questionnaire-9 (PHQ-9) depression screener in U.S. adults across sex, race/ethnicity, and education level: NHANES 2005–2016. *Depression and Anxiety.* 2019;36(9):813-823. https://doi.org/10.1002/da.22940
4. Lee B., Wang Y., Carlson S. A., Greenlund K. J., Lu H., Liu Y., Croft J. B., Eke P. I., Town M., Thomas C. W. (2023). National, State-Level, and County-Level Prevalence Estimates of Adults Aged ≥18 Years Self-Reporting a Lifetime Diagnosis of Depression - United States. *MMWR Morbidity and Mortality Weekly Report*. 2023;72(24). https://doi.org/10.15585/mmwr.mm7224a1
5. Taple B. J., Chapman R., Schalet B. D., Brower R., Griffith J. W. (2020). The Impact of Education on Depression Assessment: Differential Item Functioning Analysis. *Assessment*. https://doi.org/10.1177/1073191120971357
6. Samudio-Cruz M. A., Toussaint-González P., Estrada-Cortés B., Martínez-Cortéz J. A., Rodríguez-Barragán M. A., Hernández-Arenas C., Quinzaños-Fresnedo J., Carrillo-Mora P. (2023). Education Level Modulates the Presence of Poststroke Depression and Anxiety, But It Depends on Age. *The Journal of Nervous and Mental Disease.* 2023;211(8):585-591. https://doi.org/10.1097/NMD.0000000000001663
7. Kondirolli F., Sunder N. Mental health effects of education. *Health Economics.* 2022;31(Suppl 2):22-39. https://doi.org/10.1002/hec.4565

8. National Center for Health Statistics. (2013). *Vital and Health Statistics Report Series 1, Number 56 August 2013.* https://www.cdc.gov/nchs/data/series/sr_01/sr01_056.pdf
9. Kroenke K., Spitzer R. L., Williams J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine.* 2001;16(9):606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x
10. Ettman C. K., Abdalla S. M., Cohen G. H., Sampson L., Vivier P. M., Galea S. Prevalence of Depression Symptoms in US Adults Before and During the COVID-19 Pandemic. *JAMA Network Open.* 2020;3(9):e2019686. https://doi.org/10.1001/jamanetworkopen.2020.19686

Table 1. Univariable analysis of the association between education level and depression among U.S. adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.

Sociodemographic Characteristics	Total N (%)
Total	7812 (100.00)
Gender	
Male	3780 (48.39)
Female	4032 (51.61)
Age *(years)*	
20-29	1164 (14.90)
30-39	1212 (15.53)
40-49	1232 (15.77)
50-59	1324 (16.95)
60-69	1476 (18.89)
70-79	835 (10.69)
≥ 80	568 (7.27)
Race/Ethnicity	
Mexican American	855 (10.94)
Other Hispanic	755 (9.66)
Non-Hispanic White	2865 (36.67)
Non-Hispanic Black	2033 (26.02)
Non-Hispanic Asian	922 (11.80)
Other Race – Including Multi-Racial	382 (4.89)
Education Level[1]	
Less than 9th grade	537 (6.87)
9-11th grade (Includes 12th grade with no diploma)	841 (10.77)
High school graduate/GED or equivalent	1879 (24.05)
Some college or AA degree	2594 (33.21)
College graduate or above	1961 (25.10)
Marital Status[2]	
Married/Living with Partner	4528 (57.96)
Widowed/Divorced/Separated	1811 (23.18)
Never Married	1473 (18.86)
Ratio of Family Income to Poverty[3]	
≤ 1	1517 (19.42)
> 1	6295 (80.58)
Main Outcome	
Depression Score, Continuous	
Mean (SD)	2.88 (4.12)
Median (IQR)	1 (0,4)
Depression, Dichotomous	
No	7191 (92.05)
Yes	621 (7.95)
Covariate	
Health Insurance Coverage[4]	
Yes	6633 (84.91)
No	1179 (15.09)

1 missing data for "Education Level" (missing, n=15) due to incomplete questionnaire, answering "don't know" or "refused"
2 missing data for "Marital Status" (missing, n=10) due to incomplete questionnaire, answering "don't know" or "refused"
3 missing data for "Ratio of Family Income to Poverty" (missing, n=1,404) due to incomplete questionnaire, answering "don't know" or "refused"
4 missing data for "Health Insurance Coverage" (missing, n=22) due to incomplete questionnaire, answering "don't know" or "refused"

Acronyms: SD (standard deviation); IQR (interquartile range)

Table 2. Bivariable associations between education level and depression and key sociodemographic characteristics and other covariates among U.S. adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.

		Education Level[1]					
Sociodemographic Characterstics	**Total**	**Less than 9th grade**	**9-11th grade (Includes 12th grade with no diploma)**	**High school graduate/GED or equivalent**	**Some college or AA degree**	**College graduate or above**	***p*-value[5]**
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	
Total	7812 (100.00)	537 (6.87)	841 (10.77)	1879 (24.05)	2594 (33.21)	1961 (25.10)	
Gender							
Male	3780 (48.39)	280 (52.14)	436 (51.84)	957 (50.93)	1173 (45.22)	934 (47.63)	< 0.001
Female	4032 (51.61)	257 (47.86)	405 (48.16)	922 (49.07)	1421 (54.78)	1027 (52.37)	
Age *(years)*							
20-29	1164 (14.90)	20 (3.72)	100 (11.89)	332 (17.67)	473 (18.23)	239 (12.19)	
30-39	1212 (15.53)	66 (12.29)	132 (15.70)	232 (12.35)	418 (16.11)	365 (18.61)	
40-49	1232 (15.77)	75 (13.97)	140 (16.65)	238 (12.67)	411 (15.84)	368 (18.77)	< 0.001
50-59	1324 (16.95)	93 (17.32)	124 (14.74)	330 (17.56)	439 (16.92)	338 (17.24)	
60-69	1476 (18.89)	153 (28.49)	164 (19.50)	369 (19.64)	446 (17.19)	344 (17.54)	
70-79	835 (10.69)	69 (12.85)	96 (11.41)	225 (11.97)	264 (10.18)	181 (9.23)	
≥ 80	568 (7.27)	61 (11.36)	85 (10.11)	153 (8.14)	143 (5.51)	126 (6.43)	
Race/Ethnicity							
Mexican American	855 (10.94)	204 (37.99)	137 (16.29)	204 (10.86)	229 (8.83)	81 (4.13)	
Other Hispanic	755 (9.66)	136 (25.33)	102 (12.13)	162 (8.62)	221 (8.52)	134 (6.83)	
Non-Hispanic White	2865 (36.67)	71 (13.22)	250 (29.73)	735 (39.12)	1052 (40.56)	757 (38.60)	< 0.001
Non-Hispanic Black	2033 (26.02)	45 (8.38)	254 (30.20)	577 (30.71)	756 (29.14)	401 (20.45)	
Non-Hispanic Asian	922 (11.80)	71 (13.22)	62 (7.37)	113 (6.01)	171 (6.59)	505 (25.75)	
Other Race – Including Multi-Racial	382 (4.89)	10 (1.86)	36 (4.28)	88 (4.68)	165 (6.36)	83 (4.23)	
Marital Status[2]							
Married/Living with Partner	4528 (57.96)	350 (65.18)	462 (54.93)	1016 (54.07)	1404 (54.12)	1296 (66.09)	< 0.001
Widowed/Divorced/Separated	1811 (23.18)	143 (26.63)	225 (26.75)	484 (25.76)	614 (23.67)	345 (17.59)	
Never Married	1473 (18.86)	44 (8.19)	154 (18.31)	379 (20.17)	576 (22.21)	320 (16.32)	
Ratio of Family Income to Poverty[3]							
≤ 1	1517 (19.42)	208 (38.73)	311 (36.98)	479 (25.49)	406 (15.65)	113 (5.76)	< 0.001
> 1	6295 (80.58)	329 (61.27)	530 (63.02)	1400 (74.51)	2188 (84.35)	1848 (94.24)	
Main Outcome							
Depression Score, Continuous							
Mean (SD)	2.88 (4.12)	2.94 (4.63)	3.55 (4.77)	3.19 (4.44)	3.06 (4.07)	2.06 (3.21)	< 0.001[6]
Median (IQR)	1 (0,4)	0 (0,4)	2 (0,5)	1 (0,5)	2 (0,4)	1 (0,3)	
Depression, Categorical							
No	7191 (92.05)	487 (90.69)	743 (88.35)	1702 (90.58)	2376 (91.60)	1883 (96.02)	< 0.001
Yes	621 (7.95)	50 (9.31)	98 (11.65)	177 (9.42)	218 (8.40)	78 (3.98)	
Covariate							
Health Insurance Coverage[4]							
Yes	6633 (84.91)	396 (73.74)	645 (76.69)	1544 (82.17)	2199 (84.77)	1849 (94.29)	< 0.001
No	1179 (15.09)	141 (26.26)	196 (23.31)	335 (17.83)	395 (15.23)	112 (5.71)	

1 missing data for "Education Level" (missing, n=15) due to incomplete questionnaire, answering "don't know" or "refused"
2 missing data for "Marital Status" (missing, n=10) due to incomplete questionnaire, answering "don't know" or "refused"
3 missing data for "Ratio of Family Income to Poverty" (missing, n=1,404) due to incomplete questionnaire, answering "don't know" or "refused"
4 missing data for "Health Insurance Coverage" (missing, n=22) due to incomplete questionnaire, answering "don't know" or "refused"
5 all the *p*-values are obtained via the chi-square test
6 *p*-value for ANOVA

Acronyms: SD (standard deviation); IQR (interquartile range)

Table 3a. Bivariable associations between depression (categorical) and education level, key socioidemographic characteristics and other covariates among U.S. adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.

		Depression, Categorical		
Sociodemographic Characterstics	**Total** N (%)	**No** N (%)	**Yes** N (%)	***p*-value[5]**
Total	7812 (100.00)	7191 (92.05)	621 (7.95)	
Gender				
Male	3780 (48.39)	3547 (49.33)	233 (37.52)	< 0.001
Female	4032 (51.61)	3644 (50.67)	388 (62.48)	
Age *(years)*				
20-29	1164 (14.90)	1070 (14.88)	94 (15.14)	
30-39	1212 (15.53)	1119 (15.56)	94 (15.14)	
40-49	1232 (15.77)	1136 (15.80)	96 (15.46)	0.018
50-59	1324 (16.95)	1195 (16.62)	129 (20.77)	
60-69	1476 (18.89)	1351 (18.79)	125 (20.13)	
70-79	835 (10.69)	781 (10.86)	54 (8.70)	
≥ 80	568 (7.27)	539 (7.50)	29 (4.67)	
Race/Ethnicity				
Mexican American	855 (10.94)	793 (11.03)	62 (9.98)	
Other Hispanic	755 (9.66)	681 (9.47)	74 (11.92)	
Non-Hispanic White	2865 (36.67)	2617 (36.39)	248 (39.94)	< 0.001
Non-Hispanic Black	2033 (26.02)	1882 (26.17)	151 (24.32)	
Non-Hispanic Asian	922 (11.80)	891 (12.39)	31 (4.99)	
Other Race – Including Multi-Racial	382 (4.89)	327 (4.55)	55 (8.86)	
Education Level[1]				
Less than 9th grade	537 (6.87)	487 (6.77)	50 (8.05)	
9-11th grade (Includes 12th grade with no diploma)	841 (10.77)	743 (10.33)	98 (15.78)	< 0.001
High school graduate/GED or equivalent	1879 (24.05)	1702 (23.67)	177 (28.50)	
Some college or AA degree	2594 (33.21)	2376 (33.04)	218 (35.10)	
College graduate or above	1961 (25.10)	1883 (26.19)	78 (12.56)	
Marital Status[2]				
Married/Living with Partner	4528 (57.96)	4247 (59.06)	281 (45.25)	< 0.001
Widowed/Divorced/Separated	1811 (23.18)	1615 (22.46)	196 (31.56)	
Never Married	1473 (18.86)	1329 (18.48)	144 (23.19)	
Ratio of Family Income to Poverty[3]				
≤ 1	1517 (19.42)	1318 (18.33)	199 (32.05)	< 0.001
> 1	6295 (80.58)	5873 (81.67)	422 (67.95)	
Covariate				
Health Insurance Coverage[4]				
Yes	6633 (84.91)	6115 (85.04)	518 (83.41)	0.278
No	1179 (15.09)	1076 (14.96)	103 (16.59)	

1 missing data for "Education Level" (missing, n=15) due to incomplete questionnaire, answering "don't know" or "refused"
2 missing data for "Marital Status" (missing, n=10) due to incomplete questionnaire, answering "don't know" or "refused"
3 missing data for "Ratio of Family Income to Poverty" (missing, n=1,404) due to incomplete questionnaire, answering "don't know" or "refused"
4 missing data for "Health Insurance Coverage" (missing, n=22) due to incomplete questionnaire, answering "don't know" or "refused"
5 all the *p*-values are obtained via the chi-square test

Acronyms: SD (standard deviation); IQR (interquartile range)

Table 3b. Bivariable associations between depression (continuous) and education level, key socioidemographic characteristics and other covariates among U.S. adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.

	Depression Score, Continuous		
Sociodemographic Characterstics	**Total** N (%)	**Mean (SD)**	***p*-value[5]**
Total	7812 (100.00)	2.88 (4.12)	
Gender			
Male	3780 (48.39)	2.45 (3.81)	< 0.001
Female	4032 (51.61)	3.29 (4.36)	
Age *(years)*			
20-29	1164 (14.90)	3.10 (4.07)	
30-39	1212 (15.53)	2.76 (3.92)	
40-49	1232 (15.77)	2.94 (4.24)	0.0001
50-59	1324 (16.95)	3.22 (4.46)	
60-69	1476 (18.89)	2.88 (4.27)	
70-79	835 (10.69)	2.66 (3.86)	
≥ 80	568 (7.27)	2.13 (3.35)	
Race/Ethnicity			
Mexican American	855 (10.94)	2.78 (3.89)	
Other Hispanic	755 (9.66)	3.20 (4.54)	
Non-Hispanic White	2865 (36.67)	3.12 (4.29)	0.3534
Non-Hispanic Black	2033 (26.02)	2.77 (3.93)	
Non-Hispanic Asian	922 (11.80)	1.75 (3.11)	
Other Race – Including Multi-Racial	382 (4.89)	4.09 (4.92)	
Education Level[1]			
Less than 9th grade	537 (6.87)	2.94 (4.63)	
9-11th grade (Includes 12th grade with no diploma)	841 (10.77)	3.55 (4.77)	< 0.001
High school graduate/GED or equivalent	1879 (24.05)	3.19 (4.44)	
Some college or AA degree	2594 (33.21)	3.06 (4.07)	
College graduate or above	1961 (25.10)	2.06 (3.21)	
Marital Status[2]			
Married/Living with Partner	4528 (57.96)	2.50 (3.75)	< 0.001
Widowed/Divorced/Separated	1811 (23.18)	3.47 (4.60)	
Never Married	1473 (18.86)	3.34 (4.44)	
Ratio of Family Income to Poverty[3]			
≤ 1	1517 (19.42)	3.87 (5.04)	< 0.001
> 1	6295 (80.58)	2.65 (3.83)	
Covariate			
Health Insurance Coverage[4]			
Yes	6633 (84.91)	2.84 (4.08)	0.0114
No	1179 (15.09)	3.15 (4.32)	

1 missing data for "Education Level" (missing, n=15) due to incomplete questionnaire, answering "don't know" or "refused"
2 missing data for "Marital Status" (missing, n=10) due to incomplete questionnaire, answering "don't know" or "refused"
3 missing data for "Ratio of Family Income to Poverty" (missing, n=1,404) due to incomplete questionnaire, answering "don't know" or "refused"
4 missing data for "Health Insurance Coverage" (missing, n=22) due to incomplete questionnaire, answering "don't know" or "refused"
5 *p*-value for ANOVA

Acronyms: SD (standard deviation); IQR (interquartile range)

Table 4a. Multivariable logistic regression model examining education level and depression (categorical) among U.S. adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.

	Crude OR (95% CI)	Full model 1 Adjusted OR (95% CI)	Full model 2 Adjusted OR (95% CI)
Gender			
Male	Referent Group (1.00)	Referent Group (1.00)	Referent Group (1.00)
Female	1.62 (1.36-1.91)	1.58 (1.32-1.88)	1.55 (0.85-2.82)
Age *(years)*			
20-29	Referent Group (1.00)	0.91 (0.66-1.25)	Referent Group (1.00)
30-39	0.95 (0.70-1.28)	0.95 (0.70-1.29)	1.04 (0.77-1.42)
40-49	0.96 (0.71-1.29)	Referent Group (1.00)	1.09 (0.79-1.49)
50-59	1.22 (0.93-1.62)	1.21 (0.91-1.61)	1.32 (0.98-1.79)
60-69	1.05 (0.79-1.39)	0.98 (0.73-1.31)	1.07 (0.78-1.46)
70-79	0.78 (0.55-1.11)	0.67 (0.47-0.96)	0.73 (0.49-1.08)
≥ 80	0.61 (0.39-0.94)	0.45 (0.28-0.71)	0.48 (0.30-0.78)
Race/Ethnicity			
Mexican American	Referent Group (1.00)	Referent Group (1.00)	Referent Group (1.00)
Other Hispanic	1.38 (0.97-1.97)	1.43 (0.99-2.04)	1.43 (1.00-2.06)
Non-Hispanic White	1.21 (0.90-1.61)	1.62 (1.19-2.22)	1.63 (1.20-2.23)
Non-Hispanic Black	1.02 (0.75-1.39)	1.02 (0.73-1.41)	1.02 (0.74-1.42)
Non-Hispanic Asian	0.44 (0.28-0.69)	0.66 (0.42-1.05)	0.66 (0.42-1.05)
Other Race – Including Multi-Racial	2.15 (1.46-3.16)	2.50 (1.67-3.75)	2.50 (1.67-3.75)
Education Level[1]			
Less than 9th grade	Referent Group (1.00)	Referent Group (1.00)	Referent Group (1.00)
9-11th grade (Includes 12th grade with no diploma)	1.28 (0.89-1.84)	1.13 (0.77-1.65)	1.24 (0.71-2.18)
High school graduate/GED or equivalent	1.01 (0.72-1.40)	0.88 (0.61-1.25)	0.77 (0.45-1.31)
Some college or AA degree	0.89 (0.64-1.23)	0.75 (0.53-1.08)	0.75 (0.44-1.27)
College graduate or above	0.40 (0.27-0.58)	0.41 (0.27-0.61)	0.45 (0.24-0.82)
Marital Status[2]			
Married/Living with Partner	Referent Group (1.00)	Referent Group (1.00)	Referent Group (1.00)
Widowed/Divorced/Separated	1.83 (1.51-2.22)	1.62 (1.32-1.99)	1.62 (1.32-2.00)
Never Married	1.63 (1.32-2.02)	1.48 (1.17-1.87)	1.48 (1.18-1.87)
Ratio of Family Income to Poverty[3]			
≤ 1	Referent Group (1.00)	Referent Group (1.00)	Referent Group (1.00)
> 1	0.47 (0.39-0.56)	0.63 (0.52-0.76)	0.63 (0.51-0.76)
Covariate			
Health Insurance Coverage[4]			
Yes	Referent Group (1.00)	Referent Group (1.00)	Referent Group (1.00)
No	1.13 (0.90-1.40)	0.92 (0.72-1.16)	0.92 (0.73-1.17)
Education Level x Gender			
Less than 9th grade, male	NA	NA	Referent Group (1.00)
9-11th grade..., female	NA	NA	0.84 (0.40-1.75)
High school graduate..., female	NA	NA	1.24 (0.63-2.45)
Some college or AA degree, female	NA	NA	1.00 (0.51-1.96)
College graduate or above, female	NA	NA	0.86 (0.40-1.85)

1 missing data for "Education Level" (missing, n=15) due to incomplete questionnaire, answering "don't know" or "refused"
2 missing data for "Marital Status" (missing, n=10) due to incomplete questionnaire, answering "don't know" or "refused"
3 missing data for "Ratio of Family Income to Poverty" (missing, n=1,404) due to incomplete questionnaire, answering "don't know" or "refused"
4 missing data for "Health Insurance Coverage" (missing, n=22) due to incomplete questionnaire, answering "don't know" or "refused"

Acronyms: OR (odds ratio); CI (confidence interval)

Table 4b. Multivariable linear regression model examining education level and depression (continuous) among U.S. adults ≥ 20 years old (N=7,812), National Health and Nutrition Examination Survey (NHANES), United States, 2017-prepandemic March 2020.

	Unjusted model		Full model 1		Full model 2	
	beta (SE)	*p*-value	beta (SE)	*p*-value	beta (SE)	*p*-value
Gender						
Male	Referent Group (1.00)	NA	Referent Group (1.00)	NA	Referent Group (1.00)	NA
Female	0.84 (0.09)	< 0.001	0.77 (0.09)	< 0.001	0.86 (0.34)	0.013
Age *(years)*						
20-29	Referent Group (1.00)	NA	-0.12 (0.17)	0.478	Referent Group (1.00)	NA
30-39	-0.33 (0.16)	0.045	-0.26 (0.16)	0.105	-0.14 (0.16)	0.386
40-49	-0.15 (0.16)	0.357	Referent Group (1.00)	NA	0.11 (0.17)	0.522
50-59	0.11 (0.16)	0.472	0.18 (0.15)	0.247	0.29 (0.17)	0.082
60-69	-0.21 (0.16)	0.186	-0.21 (0.15)	0.18	-0.09 (0.17)	0.589
70-79	-0.44 (0.18)	0.018	-0.58 (0.18)	0.002	-0.47 (0.20)	0.018
≥ 80	-0.96 (0.21)	< 0.001	-1.35 (0.21)	< 0.001	-1.25 (0.22)	< 0.001
Race/Ethnicity						
Mexican American	Referent Group (1.00)	NA	Referent Group (1.00)	NA	Referent Group (1.00)	NA
Other Hispanic	0.41 (0.20)	0.041	0.45 (0.20)	0.024	0.45 (0.20)	0.024
Non-Hispanic White	0.33 (0.15)	0.035	0.77 (0.16)	< 0.001	0.78 (0.16)	< 0.001
Non-Hispanic Black	-0.01 (0.16)	0.937	-0.05 (0.17)	0.764	-0.03 (0.17)	0.818
Non-Hispanic Asian	-1.02 (0.19)	< 0.001	-0.47 (0.19)	0.017	-0.47 (0.19)	0.016
Other Race – Including Multi-Racial	1.31 (0.25)	< 0.001	1.46 (0.25)	< 0.001	1.47 (0.25)	< 0.001
Education Level[1]						
Less than 9th grade	Referent Group (1.00)	NA	Referent Group (1.00)	NA	Referent Group (1.00)	NA
9-11th grade (Includes 12th grade with no diploma)	0.60 (0.22)	0.007	0.41 (0.22)	0.069	0.39 (0.31)	0.210
High school graduate/GED or equivalent	0.24 (0.20)	0.214	0.03 (0.20)	0.85	-0.04 (0.28)	0.866
Some college or AA degree	0.11 (0.19)	0.544	-0.12 (0.20)	0.537	-0.11 (0.27)	0.669
College graduate or above	-0.88 (0.19)	< 0.001	-0.79 (0.21)	< 0.001	-0.56 (0.28)	0.048
Marital Status[2]						
Married/Living with Partner	Referent Group (1.00)	NA	Referent Group (1.00)	NA	Referent Group (1.00)	NA
Widowed/Divorced/Separated	0.96 (0.11)	< 0.001	0.77 (0.11)	< 0.001	0.77 (0.11)	< 0.001
Never Married	0.83 (0.12)	< 0.001	0.62 (0.13)	< 0.001	0.62 (0.13)	< 0.001
Ratio of Family Income to Poverty[3]						
≤ 1	Referent Group (1.00)	NA	Referent Group (1.00)	NA	Referent Group (1.00)	NA
> 1	-1.22 (0.11)	< 0.001	-0.80 (0.12)	< 0.001	-0.79 (0.13)	< 0.001
Covariate						
Health Insurance Coverage[4]						
Yes	Referent Group (1.00)	NA	Referent Group (1.00)	NA	Referent Group (1.00)	NA
No	0.30 (0.13)	0.018	-0.01 (0.13)	0.899	-0.01 (0.13)	0.958
Education Level x Gender						
Less than 9th grade, male	NA	NA	NA	NA	Referent Group (1.00)	NA
9-11th grade..., female	NA	NA	NA	NA	0.03 (0.44)	0.941
High school graduate..., female	NA	NA	NA	NA	0.15 (0.39)	0.692
Some college or AA degree, female	NA	NA	NA	NA	-0.04 (0.38)	0.904
College graduate or above, female	NA	NA	NA	NA	-0.46 (0.39)	0.233

1 missing data for "Education Level" (missing, n=15) due to incomplete questionnaire, answering "don't know" or "refused"
2 missing data for "Marital Status" (missing, n=10) due to incomplete questionnaire, answering "don't know" or "refused"
3 missing data for "Ratio of Family Income to Poverty" (missing, n=1,404) due to incomplete questionnaire, answering "don't know" or "refused"
4 missing data for "Health Insurance Coverage" (missing, n=22) due to incomplete questionnaire, answering "don't know" or "refused"

Acronyms: SE (standard error)