# Relationship between Sleep and Depression

What is the relationship between sleep duration and depression severity among US adults, and is it non-linear (quadratic)?

## Literature Review and Theoretical Background

Sleep has long been understood to be linked to depression. Bojarska et al. (2024) review evidence showing that sleep disturbance is one of the most common and disabling features of depressive disorders, affecting around 90% of patients with major depression. The authors explain that depression is linked to multiple types of sleep disruption, which include difficulties falling asleep, frequent awakenings, shortened total sleep time, and changes in REM sleep. As a result, daytime functioning and quality of life are worsened, which play key roles in depression.

Nutt et al. (2008) identify sleep disorders as core symptoms of depression; indeed, changes in sleep are embedded in the diagnostic criteria for major depressive disorder. Additionally, the authors explain that depression is associated to both *too little* or *too much* sleep; both insomnia (reduced or fragmented sleep, early morning awakening) and hypersomnia (excessive sleep and daytime sleepiness) are possible. Nutt et al.'s study this bimodality mechanistically, focusing on how neurotransmitter systems and circadian timing play a role.

Overall, the bidirectional causal link between sleep and depression has long been established based on neuroscience. In contrast, as my study will explore, this link has also been studied at the level of statistical association. For instance, Ding et al. (2022) use survey data from 1,429 older Chinese women to study how self-reported sleep duration and sleep quality are associated with depressive symptoms; depressive symptoms were used as the dependent variables, and sleep duration and quality were used as the independent variables. They use logistic regression models controlling for demographic, socioeconomic, lifestyle, and health factors. They find a U-shaped relationship between sleep duration and depressive symptoms: both short sleep and long sleep are associated with higher odds of depression compared with sleeping 6-8 hours. Other findings support a quadratic association between these variables (figure 1).

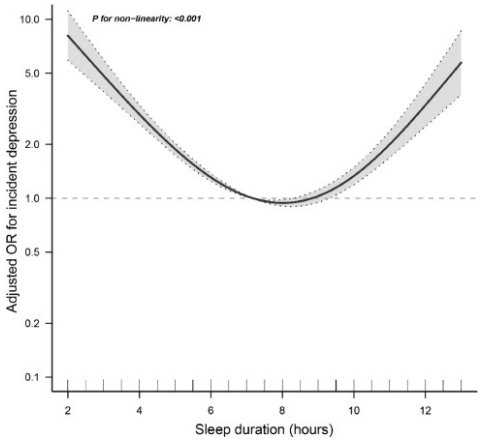


Figure 1. Adjusted odds ratio for depressive symptoms across sleep duration.
(Dong et al., 2022)

There is therefore a strong empirical motivation for modeling sleep with a quadratic term. With this consideration, my paper will estimate the non-linear association between sleep duration and depression severity using PHQ-9 scores (see below) while controlling for key variables. This paper will thereby extend the existing literature on quadratic shaped sleep-depression association.

### PHQ-9

PHQ-9 is a standard depression questionnaire used in medicine and research. They are questions that map onto the main diagnostic criteria to diagnose depression in practise using DSM (Diagnostic and Statistical Manual of Mental Disorders). Each PHQ-9 question is about a core symptom of depression (sad mood, anhedonia, sleep, appetite, energy, concentration, etc.) over the last two weeks and is answered on a scale of 0 – 3. PHQ-9 is calculated by summing the total scores of the nine responses. Obtaining a PHQ-9 score of ten or greater indicates a strong likelihood of having Major Depressive Disorder.

## Data

I am using the NHANES dataset (National Center for Health Statistics). Their survey collects a wide variety of data on the health of the United States population. NHANES does not use a simple random sample to collect participants. Rather, it uses oversampling to ensure that certain subgroups (such as specific racial/ethnic groups) are selected at higher rates to improve estimate precision. It provides weights for each participant that represent how many people in the U.S. population that one participant represents.

For each component (for example *sleep*, or *diet*) in the survey, there are several cycles which correspond to the years that the data was collected. For this project, I am using the following components in the D cycle (2005, 2006):

- Depression Screener (DPQ_D) – **Dependent Variable**
- Sleep Disorders (SLQ_D) – **Independent Variables**
- Physical Activity (PAQ_D) - Control
- Demographic Variables & Sample Weights (DEMO_D) – Controls and weights
- Dietary Interview - Individual Foods, First Day (DR1IFF_D) – Control

Each component contains its many questions within it. Some questions and demographics are categorical. For these, codes are used to refer to participant responses. For example, income responses are coded from 1 – 13, where each one refers to a possible income range.

Components are merged into one dataset using their SEQN (participant ID). Only relevant variables (which are responses within the questionnaires) are kept from each component. From this, the prepared data set contains precisely the following variables:

**Dependent Variable:** Depression severity = DPQ010 + DPQ020 + … + DPQ090

**Independent Variables**:

- Sleeping hours = SLD010H
- Sleeping hours squared = $(SLD010H)^2$

**Controls:**

- Physical activity = PAQ180,
- Diet = DBQ700,
- Age = RIDAGEYR,
- Gender = RIAGENDR,
- Race = RIDRETH1,
- Income = INDFMINC

## Strength of Dataset

For the purposes of this paper, the NHANES data are a very strong fit because they provide relatively large sample of US adults with information on both sleep and depression in the same individuals; the constructed data set that is analyzed has 5,334 observations. The Depression Screener (DPQ_D) uses the PHQ-9, which, as discussed, is a widely used questionnaire with theoretical backing. Hence, the dependent variable is measured in a way that's comparable to the clinical and econometric literature. The Sleep Disorders component (SLQ_D) provides a quantitative measure of sleep duration, which can be modeled with a quadratic term to capture non-linearity.

NHANES also includes a wide variety of covariates, allowing to control for key confounders which helps mitigate omitted variable bias in a regression. Finally, the relatively large sample size means the Central Limit Theorem justifies using the regression results (i.e. standard errors, confidence intervals and hypothesis tests) to interpret the estimated association between sleep and depression.

## Limitations of Dataset

Because sleep and depression are measured at one time, the regression can only estimate associations, not clean causal effects. Reverse causality (depression affecting sleep) is established (Nutt et al., 2008), and omitted variable bias remains plausible.

Additionally, NHANES does not select individuals using a simple random sample; instead, it uses oversampling, meaning it intentionally selects more people from certain groups than their real share in the population, so those groups have enough data to analyze accurately. Thus, in order to make population-level inferences, a standard unweighted OLS regression does not suffice. In some sense, this feature can be seen a strength of the dataset, however it introduces additional complexity into the regression.

# Methodology

## Model

The multiple linear regression model used is described by the following equation:

$$\text{Dep\_severity}_i = \beta_0 + \beta_1 * \text{Sleep}_i + \beta_2 * (\text{Sleep}_i)^2 + \vec{\beta}_c \cdot \overrightarrow{controls}_i + e_i$$

- $\beta_0$ : the intercept
- $\beta_1$ : the causal effect of sleeping duration for individual *i*
- $\beta_2$ : the causal effect of sleeping duration squared for individual *i*
- $\vec{\beta}_c$ : a vector containing the coefficient of each control, i.e. each's causal effect
- $\overrightarrow{controls}$ : a vector containing the values of control variables for individual *i*

$\vec{\beta}_c \cdot \overrightarrow{controls}_i$ is a dot product resulting in the sum of all control coefficients multiplied by their corresponding control variable's value.

This equation represents the *theoretical* model of the population, where each coefficient is related to the population (and not the sample). For the actual sample regression, weights must be accounted for.

The residual for each participant is given by the following (by isolating the error term)

$$\hat{u}_i = \text{Dep\_severity}_i - \beta_0 - \hat{\beta}_1 * \text{Sleep}_i - \hat{\beta}_2 * (\text{Sleep}_i)^2 - \text{sample control weights} * \text{controls}$$

The model then tries to minimize the sum of squared residuals, multiplied by their weights:

$$\sum_i w_i * (\hat{u}_i)^2$$

Intuitively, respondents who represent more people in the target US population receive larger weights, so the resulting estimates can be interpreted as associations at the population level for rather than averages over the unweighted sample.

## Methods used to Obtain the Dataset

After importing the components (.xpt files) into R, I merge them into one tibble (df), with inner_join using participant id's (SEQN). I then calculate each participant's phq9 score by summing their responses to all questions in the depression screener, stored as dep_severity. Because NHANES uses specific codes to mean that the response is unusable in certain questions, I replace them with NA. Finally, I discard all residual variables from components that are irrelevant, by using the select() method to keep the relevant variables. In this step, I also rename the variables to more direct names instead of their question codes. For example, sleeping

duration from the sleeping component is initially given as SLD010H, but I rename it to sleeping_hours. The resulting tibble serves as my dataset that is ready for the regression.
I then estimate the quadratic regression using the lm() function, specifying the PHQ-9 score as the dependent variable and including sleep hours, sleep hours squared, and the chosen controls. The regression is run with NHANES given weights passed through the weights argument. Plots are generated using ggplot2.

# Results and Analysis

The model predicts a quadratic (U-shaped) relationship between sleeping duration and depression severity (figure 2). The minimum depression severity is around 8 hours, which is widely considered healthy, even optimal. This suggests that, holding constant physical activity, diet, demographics and income, depressive symptoms are lowest among adults who sleep around 8 hours. The parabolic relationship suggests that both short sleep and long sleep are associated with higher depression severity scores, as predicted by previous literature.

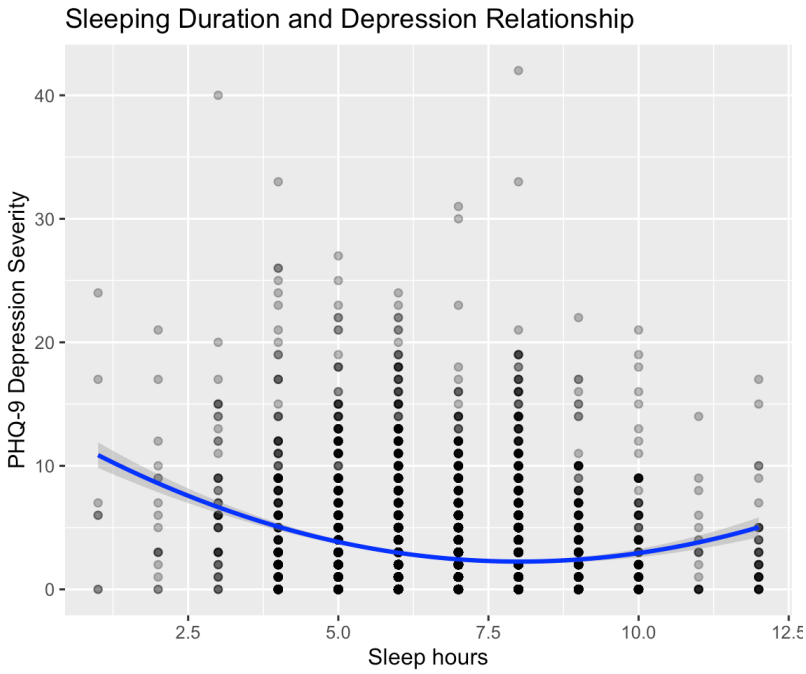


Figure 2

Table 1 reports the coefficients and standard errors of the statistically significant variables (p-value < 0.05). All variables in this table have a p-value less than 0.001 except income 4 and income 12. Certain variables (such as income) are numbered because they are categorical values, i.e. income 1 corresponds to the estimate for the income bracket given by code 1 (in this case $0 to $4999).

**The key coefficients, sleeping hours and sleeping hours squared are −2.02 and 0.13 respectively.** Their standard errors for these variables imply that the sleeping hours estimate

(–2.017) would vary by +- 0.21 hours from the *theoretical* population parameter while sleeping hours squared estimate (0.12) would vary by roughly +- 0.02 hours. The adjusted $R^2$ is around 0.1394, meaning that the regressors used in the model explain about 14% of the variation in depression severity across all participants.

To give a sense of magnitude, the coefficients imply that, compared with 8 hours of sleep, sleeping about 5 hours is associated with roughly a 1-point higher PHQ-9 score, while sleeping about 10 hours is associated with roughly a 0.5-point higher score (all else equal).

**Table 1**

| | Coefficient | Std. Error |
|---|---|---|
| **sleeping hours** | -2.017404 | 0.212443 |
| **sleeping hours$^2$** | 0.127665 | 0.015310 |
| **phys activity 2** | -0.792501 | 0.133080 |
| **phys activity 3** | -0.912700 | 0.161819 |
| **phys activity 4** | -0.851306 | 0.216064 |
| **diet 4** | 1.341508 | 0.214481 |
| **diet 5** | 3.155126 | 0.279470 |
| **female** | 0.768499 | 0.107148 |
| **income 4** | -1.323475 | 0.415737 |
| **income 6** | -1.559616 | 0.383779 |
| **income 7** | -1.927675 | 0.389182 |
| **income 8** | -1.916387 | 0.388153 |
| **income 9** | -2.168003 | 0.403656 |
| **income 10** | -1.997361 | 0.407677 |
| **income 11** | -2.165477 | 0.368009 |
| **income 12** | -1.841781 | 0.596714 |

## Interpretation

Overall, the results show evidence for a non-linear association between sleep duration and depression severity in US adults; both short and long sleep are associated with higher PHQ-9 scores. The pattern is consistent with prior work finding quadratic relationships between sleep and depressive symptoms.

However, the relatively low adjusted $R^2$ highlights that most of the variation in depression severity remains unexplained by the model's included regressors. However, this is not surprising given the complexity and multifactorial nature of depression. Moreover, the simultaneous measurement of sleep and depression in NHANES means that these results can only be interpreted as *associations*, not causal effects. Reverse causality (depression affecting sleep), measurement error in self-reported sleep, and omitted variables may all influence the estimated coefficients.

# References

Bojarska, W., Bury, K., Januszczak, R., Burda, B., & Pawęzka, J. (2024). Role of sleep in depressive disorders and the potential therapeutic role of short-term sleep deprivation and light therapy: A review. *Medical Science Monitor, 30*, e945319. https://doi.org/10.12659/MSM.945319 PubMed

Nutt, D., Wilson, S., & Paterson, L. (2008). Sleep disorders as core symptoms of depression. *Dialogues in Clinical Neuroscience, 10*(3), 329–336. https://doi.org/10.31887/DCNS.2008.10.3.dnutt PubMed

Ding, L., Zhang, L., Cui, Y., Gong, Q., Ma, J., Wang, Y., & Sang, H. (2022). The association of sleep duration and quality with depressive symptoms in older Chinese women. *PLOS ONE, 17*(3), e0262331. https://doi.org/10.1371/journal.pone.0262331 PLOS

Dong, L., Xie, Y., & Zou, X. (2022). Association between sleep duration and depression in US adults: A cross-sectional study. *Journal of Affective Disorders, 296*, 183–188. https://doi.org/10.1016/j.jad.2021.09.075 PubMed

Liu, J., Magielski, J., Glenn, A., & Raine, A. (2022). The bidirectional relationship between sleep and externalizing behavior: A systematic review. *Sleep Epidemiology, 2*, 100039. https://doi.org/10.1016/j.sleepe.2022.100039 PubMed