Exploring a model between sleep and BMI, with mediation by worry frequency

Rose Porta and Clara Li

4/30/2020

Abstract

Despite its well-studied effect on physical health, sleep in relation to weight management has been relatively little researched. Mental health, particularly depression, has also shown possible predictive power for weight, yet remains largely unexplored in relation to sleep quality and weight. A cross-sectional analysis of data (2018, n = 19,861) sourced from the IPUMS National Health Interview Survey was performed by fitting a multiple regression model using sleep, anxiety, and an interaction variable as predictors of BMI. Sleep was found to be a significant predictor ($p < 0.05$) for BMI, while anxiety frequency and the interaction between anxiety and sleep fell short of the significance threshold ($p = 0.0680$, $p = 0.0535$ respectively). We also fit a logistic regression model using the same three predictors, with a binary response variable indicating obesity (BMI >= 30) or not. Using this model, all three predictors yeilded significant p-values (< 0.05).Our findings corroborate the current literature on sleep and weight; despite not being significant in the linear model, the small p-values for anxiety frequency and its interaction with sleep nevertheless also suggest further research into both variables as predictors of BMI is warranted.

Background and Significance

Over the past four decades, excess weight has become an increasingly critical public health concern. Tied to a plethora of health issues, from diabetes, cardiac arrest, and some kinds of cancer, excess weight and obesity were found in 71.6% of adults 20 years and older in a 2015-16 survey by the CDC. In order to combat this epidemic of overweightness, research regarding exercise and dietary habits in relation to BMI have continually emerged and disseminated.

Despite the substantive literature behind exercise and nutrition, sleep has been comparatively little studied. When people think of weight management, few people even consider the effect of their sleep habits or their mental health. Growing evidence, however, has linked lack of sleep to poorer health outcomes; numerous recent studies, for instance, have linked poor sleep hygiene to higher rates of obesity and lower rates of success in weight loss programs. However, due to the multiplicity of factors influencing both sleep and BMI, more extensive research is necessary to establish a direct association between sleep and weight (Chaput et. al, Coughlin, Lyytikäinen et al, Romney et. al). Furthermore, relatively little research has explored the interaction between sleep and anxiety in relation to maintaining a healthy weight. A few studies have linked poor quality sleep and obesity to poor mental health outcomes, but this relationship has not been analyzed using weight as the outcome, with sleep and mental health as the primary predictors (Romney et. al, Sawyer et. al). Because anxiety is another factor known to be associated with poor physical health and might confound the relationship between sleep and weight (e.g. anxious people sleep more poorly, leading to weight gain), it was included as a second predictor of interest in our model.

Our analysis aims to model the relationship between the number of hours of sleep per night and BMI and how the additional factor of anxiety interacts with this association. We also controlled for a variety of factors already known to influence weight, namely sex, age, activity level, race, income level, and overall health. Dietary data was not provided by the data source we used and was therefore omitted from the analysis.

Our primary hypothesis is that there will be a negative correlation between hours of sleep per night and BMI: that is, as individuals get fewer hours of sleep, their predicted BMI increases. Our secondary hypothesis is that this negative association between lack of sleep and weight will be stronger for individuals who regularly experience anxiety.

Methods

Data

This study used data from IPUMS-NHIS, the National Health Interview Survey from 2018. The data was collected by obtaining a random sample of 35,000 U.S. households and randomly selecting one adult and one child (if any were present) from each. Data analysis was restricted to those 18 years and older, to avoid bias from adolescents undergoing puberty and possibly sleeping more.

An observational unit comprised of an individual adult in the U.S. surveyed by the CDC's National Center for Health Statistics. The population we intend to generalize our analysis to is the adult population of the United States (roughly 300 million individuals total).

Variables

The variables in our analysis include BMI ("BMI", response), hours of sleep ("HRSLEEP", primary explanatory), worry frequency (from "WORFREQ", secondary explanatory), age ("AGE", covariate), sex ("SEX", covariate), race (from "RACEA", covariate), activity level score (calculated from "STRONGFWK", "MOD10FWK", and "VIG10FWK", covariate), income level (from "INCFAM07ON", covariate), and overall health (from "HEALTH", covariate).

BMI is a continuous numerical variable indicating body mass index. HRSLEEP is an integer between 1 and 24 indicating self-reported number of hours of sleep per night an individual gets. We are analyzing anxiety using a binary variable worfreq_binary, with the two categories "worries daily" and "worries less frequently than daily". We created this variable by collapsing the WORFREQ variable, which originally had 5 categories. We collapsed it in this way because after analyzing a scatterplot of hours of sleep versus BMI faceted by the 5 categories, the association between sleep and BMI for the "daily" group appeared to be different from the other groups. AGE is an integer indicating age in years. We are controlling for activity level using the numerical variable activity_score, which we created using the Oncology Nursing Society formula Weekly leisure activity score = (9 × Strenuous) + (5 × Moderate) + (3 × Light) from the Godin Leisure-Time Exercise Questionnaire. We substituted strength exercise for light exercise due to the availability of the IPUMS data, and we think strength training is analogous to light exercise for the purposes of our analysis. SEX is a binary variable indicating male or female. In order to control for race, we created 3 binary variables, race_asian (Asian versus not Asian), race_black (black/African American versus not black/African American), and race_na (Native American or Alaska native or not), with white race as the reference group. We collapsed the RACEA variable in this way based on the race categories available, analysis of a faceted scatterplot of sleep versus BMI, as well as rational in order to avoid over complicating our model. In order to control for income level, we created two binary variables, income_high (annual family income over \$75,000 or not), and income_middle (annual family income between \$35,000 and \$75,000 or not), with low income (annual family income below \$35,000) as the reference group. We collapsed the NCFAM07ON variable in this way based on the sample sizes of each group as well as logistic reasons to avoid overcomplicating our analysis. We are controlling for overall health using the binary variable health_binary, with two categories "good" and "poor". We collapsed the HEALTH variable, which originally had 5 categories, in this way based on the sample sizes of each group as well as logistic reasons to avoid overcomplicating our analysis.

Since our data is based on survey responses, all values are self-reported. Additionally, we did not include in our analysis any values for any of the variables that corresponded to non-response, inavailability, or extreme values that had been top or bottom coded.

Analytic Methods

In order to analyze our primary hypothesis, we fit a multiple linear regression model with the form $log(BMI) = \beta_0 + \beta_1 HRSLEEP + \beta' X$, where $\beta' X$ represents our vector of covariates outlines above. A log transformation on BMI was performed after a normal Q-Q plot and a histogram of residuals revealed a severe violation of the normality assumption for linear regression. The assumption became reasonable with the transformation applied.

In order to analyze our secondary hypothesis, we fit a multiple linear regression model with the form

$log(BMI) = \beta_0 + \beta_1 \cdot HRSLEEP + \beta_2 \cdot worfreq\_binary + \beta_3 \cdot (HRSLEEP \cdot worfreq\_binary) + \beta' X$

As well as a multiple logistic regression model with the form

$log(odds) = \beta_0 + \beta_1 \cdot HRSLEEP + \beta_2 \cdot worfreq\_binary + \beta_3 \cdot (HRSLEEP \cdot worfreq\_binary) + \beta' X$

With the binary response variable BMI_binary, with the two categories "obese" (BMI >= 30) and "not obese" (BMI < 30).

In order to analyze the significance of the worry frequency variable, we performed a nested F-test comparing the linear model with the worfreq_binary variable and the interaction $HRSLEEP \cdot worfreq\_binary$ to the model without these two terms.

Results

After performing data wrangling and filtering out observations not applicable to our analysis (described above), we fit our regression models using a sample size of n = 19,861, with n = 9,234 males and 10,627 females and a median age of 52 years. The median BMI was 27.05, and the median number of hours of sleep is 7. There were n = 7,798 individuals classified as high income, n = 5,695 as middle income, and n = 6,368 as low income. The majority of the individuals sampled identified as white (n = 16, 179), with n = 2348 identifying as black or African American, n = 1055 identifying as Asian, and n = 279 identifying as Native American or Alaska native. The majority of individuals sampled reported being in good, very good, or excellent overall health (n = 17,379), with n = 2,482 reporting fair or poor health. The median activity score was 889. The majority of individuals reported worrying less frequently than daily (n = 17,925), with n = 1,936 reporting worrying daily.

In order to analyze our primary hypothesis based on our nested model $log(BMI) = \beta_0 + \beta_1 \cdot HRSLEEP + \beta' X$, we performed a t-test for the HRSLEEP coefficient with the hypotheses

$H_0 : \beta_1 = 0$

$H_a : \beta_1 \neq 0$

We performed all of our significance tests at the 0.05 significance level. The coefficient for HRSLEEP was significant at the 0.05 level (p < 0.001), indicating a significant relationship between hours of sleep and BMI.

Our secondary hypothesis based on our full model $log(BMI) = \beta_0 + \beta_1 \cdot HRSLEEP + \beta_2 \cdot worfreq\_binary + \beta_3 \cdot (HRSLEEP \cdot worfreq\_binary) + \beta' X$ was examined using three t-tests on the coefficients for HRSLEEP, worfreq_binary and the interaction HRSLEEP*worfreq_binary, using the following hypotheses, for i = 1, 2, and 3:

$H_0 : \beta_i = 0$

$H_a : \beta_i \neq 0$

The coefficient for HRSLEEP remained significant for the full model (p < 0.05). However, neither the coefficient for worfreq_binary or the interaction HRSLEEP*worfreq_binary was found to be significant (p < 0.10 but > 0.05 for both variables).

The results of our full linear regression model are summarized in table 1 below. The table displays the average predicted BMIs and corresponding 95% prediction intervals by hours of sleep (5-10) and whether or not an individual worries daily (1 = worries daily, 0 = worries less frequently).

Table 1

	worfreq_binary	HRSLEEP	avg_predicted_BMI	avg_lci	avg_uci
1	0	5	27.88	27.87	27.89
2	0	6	27.53	27.52	27.53
3	0	7	27.12	27.11	27.13
4	0	8	27.12	27.12	27.13
5	0	9	27.05	27.04	27.06
6	0	10	27.24	27.23	27.25
7	1	5	27.88	27.87	27.90
8	1	6	27.67	27.65	27.68
9	1	7	27.36	27.35	27.37
10	1	8	27.53	27.51	27.54
11	1	9	27.66	27.64	27.67
12	1	10	27.88	27.86	27.90

To determine whether the worry frequency and interaction term significantly improved the model, a nested F-test was performed on the full and nested models above. The test indicated that the full model including worry frequency and its interaction with sleep did not explain significantly more variation in BMI than the nested model without either variable. This is consistent with the similar R^2 values between the full and nested models (R^2 = 0.04862 vs. 0.04854).

To give additional insight into the relationship between sleep and BMI, we performed similar significance tests on our logistic model using obese vs. not obese as the response. Three z-tests on the coefficients for HRSLEEP, worfreq_binary and the interaction HRSLEEP*worfreq_binary, using the following hypotheses, for i = 1, 2, and 3:

$H_0 : \beta_i = 0$

$H_a : \beta_i \neq 0$

All coefficients were significant ($p < 0.05$). A likelihood ratio test was also performed, yielding a significant p-value ($p < 0.05$). The concordance value was 0.617, indicating our model accurately predicted 61.7% of outcomes in the model.

The results of our logistic regression model are summarized in table 2 below. The table displays the average predicted probabilities that an individual is classified as obese, and corresponding 95% prediction intervals, by hours of sleep (5-10) and whether or not an individual worries daily (1 = worries daily, 0 = worries less frequently).

Table 2

	worfreq_binary	HRSLEEP	prob	avg_lci	avg_uci
1	0	5	0.34	0.32	0.37
2	0	6	0.31	0.29	0.33
3	0	7	0.28	0.26	0.30
4	0	8	0.29	0.27	0.30
5	0	9	0.28	0.26	0.30
6	0	10	0.30	0.27	0.32
7	1	5	0.35	0.32	0.39
8	1	6	0.34	0.31	0.37
9	1	7	0.32	0.29	0.35
10	1	8	0.34	0.31	0.37
11	1	9	0.34	0.31	0.38
12	1	10	0.38	0.33	0.42

Discussion

Based on our results, it seems clear that there is a significant negative correlation between sleep and BMI, but it is less clear whether or not anxiety is a significant confounder of this relationship. These findings are consistent with the current literature, yet there are numerous limitations within our analysis as well as across the studies conducted previously in relation to this topic, so more research on this topic is necessary. Within our analysis, the validity of our conclusions were limited by the fact that we were unable to control for several covariates that may have been significant, the most notable being nutrition and specific health conditions such as sleep disorders and cardiovascular disease. This limitation was due to lack of availability of relevant variables from the ipums database. Another limitation is that all of our data was self-reported, which could have led to misestimation and loss of precision, specifically with regard to our primary explanatory variable, hours of sleep per night. However, due to the difficulty of measuring sleep directly over the course of long-term studies, much of the accredited research on sleep has been conducted using self-reported measures of sleep. Thus, this is a potentially a limitation that applies to the broader field of research on sleep. Furthermore, healthy weight management is an extremely complex issue influenced by so many factors that are impossible to control for such as genetics and the subtle differences in the metabolic processes of each individual. Thus, it is very difficult to make generalizable conclusions about which factors are most important to maintaining a healthy BMI, and this is an ongoing limitation that applies to all health science research. Overall, although no causal inferences can be made and further research is needed to validate our finding, our analysis showed a clear association indicating that as the average number of hours of sleep an individual gets per night increases, their expected average BMI and their odds of being classified as obese both decrease. Thus, our findings support the growing body of research indicating that sleep may be an extremely important factor to consider in relation to healthy weight management and maintenance of overall health and wellbeing.

References

Chaput, J-P, et al. "Longer Sleep Duration Associates with Lower Adiposity Gain in Adult Short Sleepers." International Journal of Obesity, vol. 36, no. 5, 2011, pp. 752–756., doi:10.1038/ijo.2011.110.

Coughlin, Janelle W., and Michael T. Smith. "Sleep, Obesity, and Weight Loss in Adults: Is There a Rationale for Providing Sleep Interventions in the Treatment of Obesity?" International Review of Psychiatry, vol. 26, no. 2, Apr. 2014, pp. 177–188., doi:10.3109/09540261.2014.911150.

Lyytikäinen, P, et al. "Sleep Problems and Major Weight Gain: a Follow-up Study." International Journal of Obesity, vol. 35, no. 1, 2011, pp. 109–114., doi:10.1038/ijo.2010.113.

Mitchell, N. S., Catenacci, V. A., Wyatt, H. R., & Hill, J. O. (2011, December). Obesity: Overview of an epidemic. Psychiatric Clinics of North America. NIH Public Access. https://doi.org/10.1016/j.psc.2011.08.005

Romney, Lora, et al. "Reduced Sleep Acutely Influences Sedentary Behavior and Mood But Not Total Energy Intake in Normal-Weight and Obese Women." Behavioral Sleep Medicine, vol. 14, no. 5, 2015, pp. 528–538., doi:10.1080/15402002.2015.1036272.

Sawyer, A., Fisher, A., Llewellyn, C. et al. Self-reported sleep quality, weight status and depression in young adult twins and siblings. BMC Obes 2, 50 (2015). https://doi.org/10.1186/s40608-015-0079-8

2015-16 survey by the CDC https://www.cdc.gov/nchs/fastats/obesity-overweight.htm