Evan Edelstein

Cross-sectional Assessment of the Association Between Sleep and Health

**Introduction**

Numerous publications suggest sleep to be a necessary and beneficial, component of life.1,6 Its conservation throughout most “living” species and analogous states of rest and function observed in bacteria and plants, implies the basic need for sleep. Both, the fields of neurology and psychology converge on the notion that sleep is required for brain development and repair, a necessity for memory formation and learning.1 Known to many are the immediate short-term consequences of sleep deprivation such as lethargy, mood changes, and drowsiness which can be felt for days after sleep loss.2 Long-term consequences of sleep deprivation and poor sleep are reflected by various biomarkers. Poor sleep is associated with decreased cardiovascular function and cognitive impairment as a result of disruption of the circadian rhythm. 2 Immunologically, it has been reported that sleep disorders are associated with weak immune responses and even weaker vaccine uptake.3 Sleep is also associated with metabolic dysfunction, usually manifesting in the endocrine system. This is the result of the body locked into the sympathetic nervous system, which leads to elevated glucose levels leading to glucose tolerance, hypertension and insulin imbalance.4 Increased sleep-debt, or the difference in sleep gotten verse sleep required, is also associated with high levels of stress, anxiety, and depression.5 Overall sleep-debt, poor sleep, and sleep loss leads to an imbalance of homeostasis that under chronic conditions can lead to negative health as identified by numerous biomarkers.

The association between sleep and health was investigated using the NHanes survey data from years 1999 - 2004. The hypothesis formed from contingency analysis was that poor trouble sleeping and decreased sleep duration correlate to poor health, increased Body-Mass-Index (BMI), and/or unhealthy weight class. Hypothesis testing, including t-test, chi-square, and ANOVA confirmed the hypothesis. Subgroup analysis allowed for a fine-tune expression of the hypothesis between differing gender, race, socioeconomic status, and age groups. Finally, regression models were constructed to leverage this hypothesis for screening and treating patients.

**Methods**

The NHanes data set was loaded into R and analyzed. The data set was filtered to only include adults and those who filled out general health data. The summary statistics for BMI of all the patients, those reporting sleep trouble, and those who do not report sleep trouble were calculated. The gender, race, age, and employment composition of the survey were visualized. The distribution of BMI and hours of sleep per night was graphed.

The healthiness of the patients was converted to a binary variable, with healthy individuals being labeled as those who self-report health at least “fair” and the unhealthy individuals grouped as those who self-report healthiness as “poor”. The health of patients was also stratified by BMI, by grouping patients into weight categories: obese, overweight, underweight, and healthy classes. The mean BMI of patients who do and do not report sleep trouble was analyzed via a t-test. A chi-square test was used to examine the association between sleep trouble and general health, and sleep and weight category. ANOVA tests was performed to compare hours slept amongst the health group and weight categories.

The survey members were further broken down into groups based on gender, race, employment, and age. A t-test and chi-square test was performed on each group to study the association between sleep and health in the groups.

A logistic regression model was trained on half the data set and tested on the other half. The regression was trained using hours slept and sleep trouble as independent variables, to test the binary dependent variable, healthy or not. A p-value and odds ratio were reported for the regression. An ROC curve was plotted, and a corresponding AUC was calculated.

**Results and Discussion**

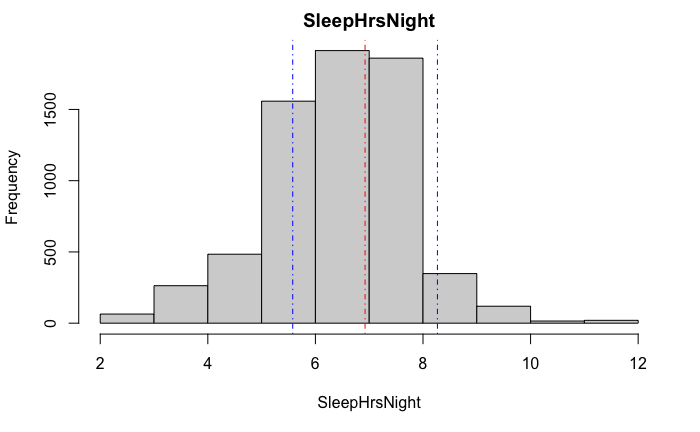
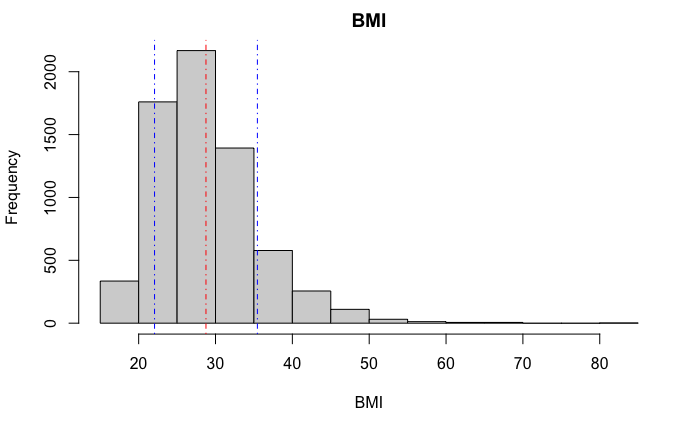
The composition and summary statistics of the dataset are described in table 1. The exposure, sleep trouble, was compared to the outcome BMI. The distribution of BMI and hours slept per night is visualized in figure 1. The ratio of sleep trouble, health, weight, gender, age, and race are shown in figure 2. The contingency tables for the data set are shown in figure 3.

The results of the t-test comparing sleep trouble with low health are shown in Table 2. This test was used to compare the null hypothesis, that the mean BMI of the sleep trouble group and not sleep trouble group is statistically equivalent, to the alternative hypothesis, that the mean BMI between those with and without sleep trouble is statistically different. Chi-square tests, table 3, show the comparison between sleep trouble and self-reported health and the comparison between sleep trouble and weight category. Table 4 compares the two nominal variables in the chi-square test, health, and weight categories, to hours slept, using Anova. A t-test and chi-square were performed on subgroups of the patients. The contingency tables for the subgroups are shown in figure 4. The t-test and chi-square data are shown in tables 5A and 5B. The regression analysis data is tabulated in table 6. The ROC curve and AUC for the regression model is shown in figure 5.

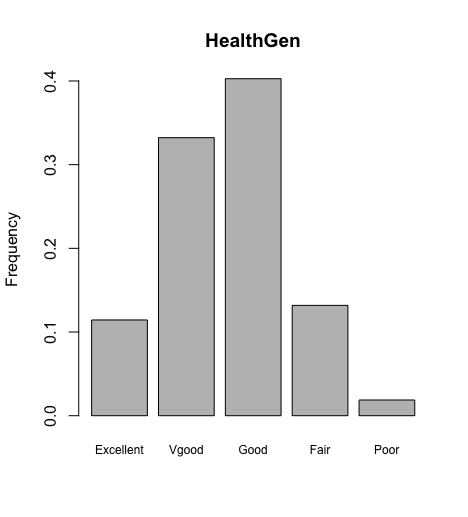
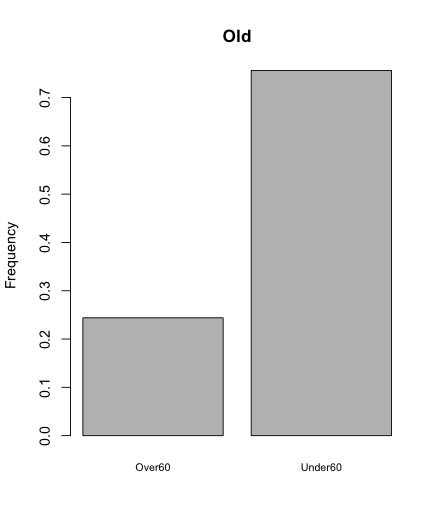
**Table 1:** Summary statistics of BMI as compared in all, no sleep trouble and sleep trouble groups.

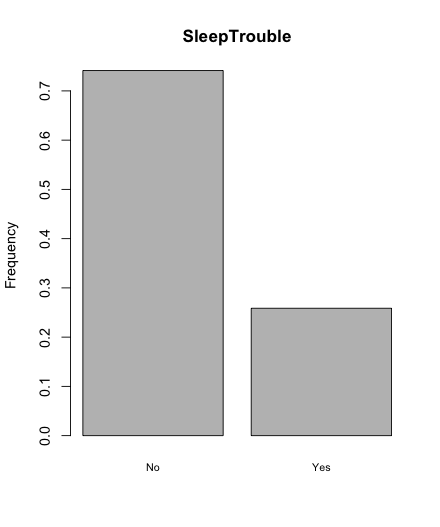
|  |  |  |  |
| --- | --- | --- | --- |
|  | **All** | **No sleep trouble** | **Sleep trouble** |
| sample size (patients) | n = 6905 | n = 5157 | n= 1748 |
| **BMI** |  |  |  |
| Mean (95% CI) | 28.67 (28.59 - 28.75) | 28.41 (28.32 - 28.49 ) | 29.43 (29.26 - 29.6) |
| Median | 27.6 | 27.4 | 28.3 |
| SD | 6.70 | 6.46 | 7.28 |
| SE | 0.078 | 0.087 | 0.166 |

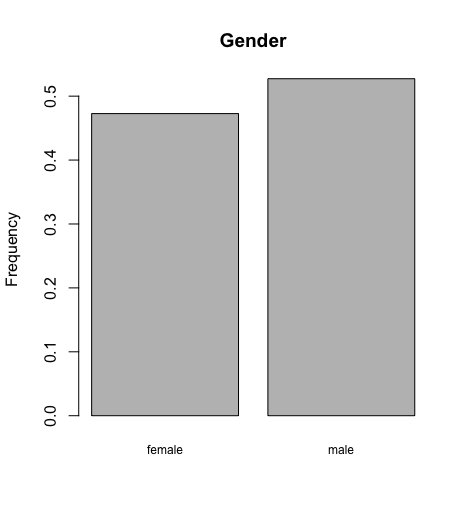
**Figure 1:** Histogram of BMI and Hours slept per night, with mean (red) and 1 SD (blue).

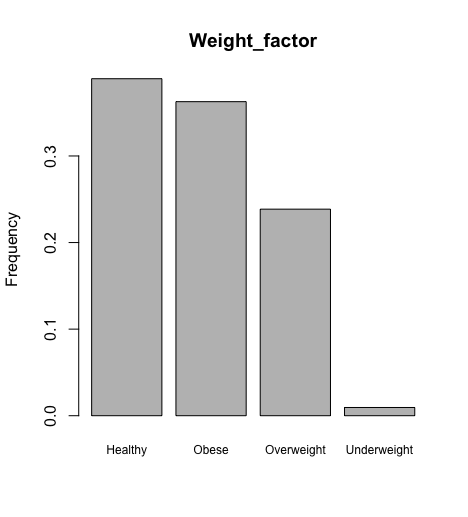


**Figure 2:** Box Plots visualizing the proportion of variables in the dataset.









Chart, bar chart, histogram

Description automatically generated

**Figure 3:** Contingency tables for health and weight class.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sleep trouble/**  **Weight class** | **No** | **Yes** | **Sum** |
| **Healthy** | 0.29 | 0.09 | 0.38 |
| **Obese** | 0.25 | 0.10 | 0.36 |
| **Overweight** | 0.19 | 0.06 | 0.25 |
| **Underweight** | 0.01 | 0.00 | 0.01 |
| **Sum** | 0.74 | 0.26 | 1.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Sleep trouble/**  **Health class** | **No** | **Yes** | **Sum** |
| **Excellent** | 0.097 | 0.019 | 0.116 |
| **Very good** | 0.257 | 0.066 | 0.323 |
| **Good** | 0.281 | 0.113 | 0.395 |
| **Fair** | 0.092 | 0.048 | 0.140 |
| **Poor** | 0.012 | 0.014 | 0.026 |
| **Sum** | 0.740 | 0.260 | 1.000 |
|  |  |  |

**Table 2:** T-test comparing BMI across no sleep trouble and sleep trouble group.

|  |  |  |  |
| --- | --- | --- | --- |
| **Student's t.test (BMI as CONT. Var)** | **T** | **p-value** | **Mean BMI x; Mean BMI y** |
| No sleep trouble vs. Sleep trouble | -5.4635 | < 0.05 | 28.4; 29.4 |

**Table 3:** Chi-square comparing sleep trouble to nominal variables.

|  |  |  |
| --- | --- | --- |
| **X^2 test** | **x^2** | **p-value** |
| Sleep trouble vs Health | 185.96 | < 0.05 |
| Sleep trouble Vs Weight | 23.764 | < 0.05 |

**Table 4:** Anova test between hours slept and classifiers health and weight.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Anova** | **df** | **Sum sq** | **Mean sq** | **F value** | **Pr(>F)** |
| Hours slept vs Health | 4 | 255 | 63.77 | 35.9 | <2e-16 |
| Hours slept vs Weight | 3 | 27 | 9.034 | 5.032 | 0.00175 |

**Figure 4:** Contingency table for subgroups.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sleep trouble** | **No** | **Yes** | **Sum** |
| **not-white** | 0.271 | 0.063 | 0.334 |
| **White** | 0.471 | 0.195 | 0.666 |
| **Sum** | 0.741 | 0.259 | 1.000 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Sleep trouble** | **No** | **Yes** | **Sum** |
| **Employed** | 0.48 | 0.13 | 0.61 |
| **Unemployed** | 0.26 | 0.12 | 0.39 |
| **Sum** | 0.74 | 0.26 | 1.00 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Sleep trouble** | **No** | **Yes** | **Sum** |
| **female** | 0.355 | 0.152 | 0.507 |
| **male** | 0.386 | 0.107 | 0.493 |
| **Sum** | 0.741 | 0.259 | 1.000 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Sleep trouble** | **No** | **Yes** | **Sum** |
| **Over60** | 0.171 | 0.073 | 0.244 |
| **Under60** | 0.570 | 0.186 | 0.756 |
| **Sum** | 0.741 | 0.259 | 1.000 |

**Table 5A:** Subgroup T-test.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **t.test**  **(BMI as CONT. variable)** | **Group** | **T** | **P\_value** | **mean MBI x ; Mean BMI y** | **Diff\_in\_means (95% CI)** |
| **No sleep trouble vs. Sleep trouble** | Male  n = 3646 | -1.744461 | > 0.05 | 28.59 ; 29 | -0.4 (-0.9 to 0.1) |
|  | Female  n =3751 | -5.486465 | < 0.05 | 28.21 ; 29.73 | -1.5 (-2.1 to -1) |
|  | White  n = 4927 | -3.673735 | < 0.05 | 28.23 ; 29 | -0.8 (-1.2 to -0.4) |
|  | Not-white  n = 2470 | -4.989505 | < 0.05 | 28.72 ; 30.75 | -2 (-2.8 to -1.2) |
|  | Employed  n = 4519 | -4.154251 | < 0.05 | 28.36 ; 29.37 | -1 (-1.5 to -0.5) |
|  | Unemployed  n = 2877 | -3.427661 | < 0.05 | 28.48 ; 29.49 | -1 (-1.6 to -0.4) |
|  | Under60  n = 5592 | -4.803083 | < 0.05 | 28.31 ; 29.37 | -1.1 (-1.5 to -0.6) |
|  | Over60  n = 1805 | -2.412213 | < 0.05 | 28.74 ; 29.58 | -0.8 (-1.5 to -0.2) |

**Table 5B:** Subgroup chi-square test comparing sleep trouble and health.

|  |  |  |  |
| --- | --- | --- | --- |
| **X^2** | **Group** | **T** | **P\_value** |
| **Sleep trouble vs Health** | Male  n = 3646 | 120.11653 | < 0.05 |
|  | Female  n =3751 | 82.89521 | < 0.05 |
|  | White  n = 4927 | 179.95061 | < 0.05 |
|  | Not-white  n = 2470 | 72.30486 | < 0.05 |
|  | Employed  n = 4519 | 36.46979 | < 0.05 |
|  | Unemployed  n = 2877 | 118.15108 | < 0.05 |
|  | Under60  n = 5592 | 162.0077 | < 0.05 |
|  | Over60  n = 1805 | 44.57444 | < 0.05 |

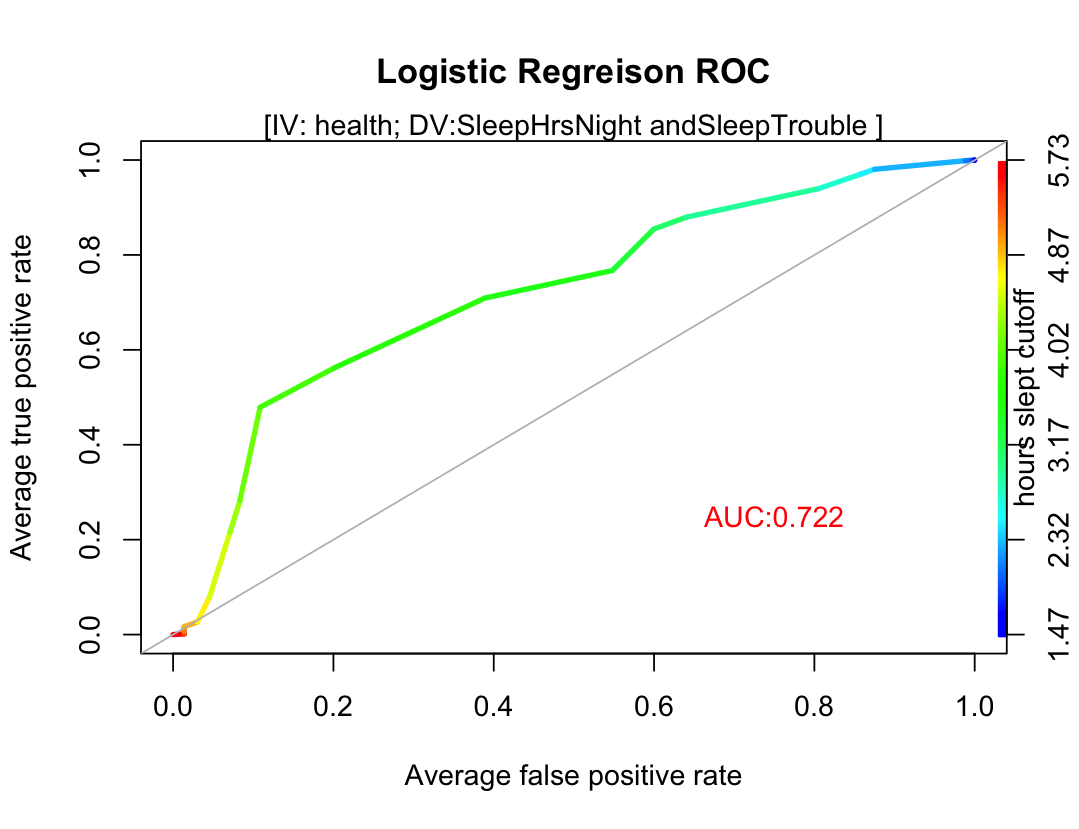
**Table 5C:** SubgroupChi-square test comparing sleep trouble and weight class.

|  |  |  |  |
| --- | --- | --- | --- |
| **X^2** | **group** | **T** | **P\_value** |
| **Sleep trouble Vs Weight** | Male  n = 3646 | 3.313102 | > 0.05 |
|  | Female  n =3751 | 29.639271 | < 0.05 |
|  | White  n = 4927 | 13.579394 | < 0.05 |
|  | Not-white  n = 2470 | 20.005554 | < 0.05 |
|  | Employed  n = 4519 | 15.307257 | < 0.05 |
|  | Unemployed  n = 2877 | 7.780497 | > 0.05 |
|  | Under60  n = 5592 | 17.79746 | < 0.05 |
|  | Over60  n = 1805 | 7.382057 | > 0.05 |

**Table 6:** Regression data, the dependent variables are hours slept per night and binary classifier, sleep trouble or not. The independent variable is a binary classifier based on patients’ self-reported health data.

|  |  |  |
| --- | --- | --- |
| **Characteristic** | **OR (95% CI)** | **P** |
| **SleepHrsNight** | 1.46 (1.327 - 1.63) | < 0.05 |
| **Sleep Trouble** | 0.389 ( 0.261 - 0.491) | < 0.05 |

**Figure 5:** ROC curve for logistic regression model. Figure includes the AUC (red) and random performance (gray, y = x).

****

**Conclusion**

The compositional analysis of the NHanes data set used shows an increase in BMI in the patients who report trouble sleeping as compared to those who don't report sleep trouble. This allows fits with the hypothesis that exposure to sleep trouble results in lower health and increased BMI. This conclusion is supported by a t-test (Table 2), which confirms a significant difference in BMI between the sleep trouble and normal group, with a p-value less than 0.05. The results of the chi-square analysis in table 3, indicate a strong correlation with lower self-reported health and sleep trouble, as well as a strong correlation between unhealthy weight categories and sleep trouble. An ANOVA test was used to compare hours slept amongst differing health and weight groups. The null hypothesis, that the average hours slept per night is equivalent amongst health and weight groups were rejected, as indicated by a p-value less than 0.05. This result leads to an acceptance of the alternative hypothesis that fewer hours of sleep per night is associated with poor health or unhealthy weight groups.

The survey participants were further sub-stratified by gender, race, employment, and age to examine which groups are associated with sleep trouble. The null hypothesis suggests that the mean hours slept is equivalent between those with and without sleep trouble amongst patients in the subgroup. The results of a t-test (table 5A) indicate the rejection of the null hypothesis in all subgroups but males. This concludes that in all subgroups, excluding males, the alternative hypothesis, that sleep trouble is associated with increased BMI, is maintained. The result of a chi-square test (table 5B) indicated that sleep trouble is correlated with poor self-reported health amongst all subgroups. The correlation between sleep trouble and unhealthy weight category amongst subgroups (table 5C) was less definitive with, male, unemployed and over 60 years of age groups indicating no association between unhealthy weight category and sleep trouble whole the remaining groups did maintain a correlation between the exposure and outcome.

The regression analysis shows the predictive ability of hours slept per night and sleep trouble in determining the health of a patient. The odds ratio with 95% CI for the hours slept per night and sleep trouble was 1.46 (1.327 - 1.63) and 0.389 (0.261 - 0.491), respectively. This result, combined with an AUC of 0.722, indicates the predictive power of sleep in determining the health of patients.

Overall, the data analysis confirms the hypothesis that sleep trouble is associated with poor health. This can be understood in light of the numerous links between sleep and physiological markers. In the future, comprehensive longitudinal studies can pinpoint the physiological pathways sleep deprivation disrupts. Furthermore, research in how to sleep better can help remedy the malice associated with poor sleep. Lastly, sleep data, which is relatively easy to collect and self-report, can be used to train better machine learning systems to help identify high-risk patients. Interestingly, the reverse hypothesis, that poor health leads to decreased sleep, could yield further information about the link between sleep and health.

**References**

1. Benington, J. H., & Craig Heller, H. (1995). Restoration of brain energy metabolism as the function of sleep. Progress in Neurobiology, 45(4), 347–360. doi:10.1016/0301-0082(94)00057-o
2. Czeisler CA. Impact of sleepiness and sleep deficiency on public health--utility of biomarkers. *J Clin Sleep Med*. 2011;7(5 Suppl):S6-S8. doi:10.5664/JCSM.1340
3. Spiegel K, Sheridan JF, Van Cauter E. Effect of Sleep Deprivation on Response to Immunizaton. *JAMA.* 2002;288(12):1471–1472. doi:10.1001/jama.288.12.1469
4. Karine Spiegel, Rachel Leproult, Eve Van Cauter,Impact of sleep debt on metabolic and endocrine function,The Lancet,Volume 354, Issue 9188,1999,Pages 1435-1439,ISSN 0140-6736,https://doi.org/10.1016/S0140-6736(99)01376-8.
5. Nollet, M., Wisden, W., & Franks, N. P. (2020). Sleep deprivation and stress: a reciprocal relationship. *Interface focus*, *10*(3), 20190092. <https://doi.org/10.1098/rsfs.2019.0092>
6. Buysse DJ. Sleep health: can we define it? Does it matter?. *Sleep*. 2014;37(1):9-17. Published 2014 Jan 1. doi:10.5665/sleep.3298