Observational Study

Investigating the effects of sleep disorder on sleep quality

STA305: Ziyi Zhang 1005282720

INTRODUCTION

Sleep quality plays a vital role in maintaining individuals' overall health and well-being. Research have found that adequate and restful sleep is essential for physical, cognitive, and emotional functioning[1]. Numerous lifestyle factors such as sleep duration, physical activity, stress[2], diet, body mass index, and occupation[3] all could contribute a significant influence in sleep quality, which could potentially lead to sleep-related disorders like insomnia and sleep apnea[6]. In order to develop effective methods and promote healthy sleep habits, it is crucial for individuals to understand the impact of these lifestyle factors on sleep quality.

The purpose of this study aims to investigate the effect of sleep disorder[6] on sleep quality, and to examine the associations among sleep duration, physical activity level, stress, heart rate, blood pressure, and BMI category. 400 individuals were asked whether they had a type of sleep disorder or not. One group was formed if they claim to have a type of sleep disorder, and another group was formed if they claim to not have. These two groups were matched by age, gender, and BMI category, and interviewed about their past sleeping habits, and lifestyle habits. The findings will contribute to the existing knowledge and research on sleep health and aims to provide comprehensive analysis to shed light on the complex relationship between lifestyle factors and sleep quality.

METHODS

Dataset URL: kaggle datasets download -d uom190346a/sleep-health-and-lifestyle-dataset

Study population

The Sleep Health Lifestyle Dataset contains information about a group of test subjects and their sleep patterns. This analysis was conducted using data collected by surveying individuals in their late twenties to late fifties. Like any form of self-report measures, we need to keep in mind that these data collected are susceptible to various type of biases[2]. The author of this dataset did not disclose the survey instruments used to conduct the data collecting. The limitations, bias and potential uncertainties associated with the dataset is acknowledged.

Measures of Interest

Below are all the variables of interest of this study. Sleep quality, sleep duration, physical activity level, stress level, sleep disorder was all collected through self-assessment surveys. Age, gender, blood pressure, heart rate, and BMI category variables were measured and recorded down by respondents. Treatment factor: lifestyle habits; outcome/response: Sleep quality.

Gender & Age

Every respondent's gender and age were recorded down.

Sleep Quality
The Primary exposures of interest in this study was Sleep Quality. Respondents were asked to give a subjective rating of the quality of sleep, ranging from 1 to 10. This method relies on individuals' self-perception and self-reporting of their own sleep quality.

Sleep Duration
Sleep duration was calculated by the number of hours the person sleeps per day, which was self-reported by the respondents is another common approach to assess sleep habits and patterns in research studies. Various methods that researchers could use to collect the data are questionnaires, interviews, or sleep diaries[1].

Physical Activity Level
Physical Activity Level was reported daily by the respondents. Individuals have to self-report the number of minutes that they engaged in physical activity every day. It wasn't made clear if definitions and examples of what counts as physical activity were provided to the respondents.

Stress Level
The data for Stress Level variable is collected through a survey where respondents gave a subjective rating of the stress level, they had experienced, ranging from 1 to 10. It was not disclosed how the stress level assessment was tested to measure the individual's perceived stress levels.

Heart Rate (bpm)
Individuals were also asked to report their rest heart rate in beats per minute. The methods of measuring bpm were not disclosed in this dataset. Some common techniques used include electrocardiography (ECG), photoplethysmography (PPG), or wearable devices[3].

BMI Category
Body Mass Index (BMI) are usually calculated using measured individuals' height and weight. It is then classified into body weight status (Normal, Normal weight, Obese, and Overweight). Not enough data found under "Obese" and "Normal weight" thus it was added into "Overweight" and "Normal" respectively. To use a categorical data in regression analysis, BMI Category variable was transformed to numerical data, with "Normal" = 0, and "Overweight" = 1.

Daily Steps
Daily Steps are usually tracked by a digital device or a wearable tracker[3].

Sleep Disorder
The data on sleep disorders can be collected through self-reported surveys and questionnaires that include questions about their sleep patterns, quality of sleep, and presence of sleep-related symptoms or disorders[6].

Variable Selection
Data from a total of 374 individuals were to be used in analysis. No apparent outliers can be detected in the data. All parameters seem intuitively related to sleep quality. Lifestyle and health factors such as age, gender, sleep duration, physical activity, stress level, heart rate, daily steps,

and BMI Category were used as matching factors when individuals were grouped into one group with sleeping disorder, and another group without sleeping disorder.

Statistical Analysis

After the data set for analysis was defined, a comprehensive approach was implemented to address bias and confounding in this observational study. Propensity scores and various techniques for propensity score adjustment were utilized to control for potential confounders and to further improve the validity of the study's findings[5]. Initially, the propensity scores were estimated, which are probabilities of individuals having sleep disorders based on their covariates (age, gender, BMI category, lifestyle factors, etc.). This step helps create a balancing score that summarizes the covariates' influence on the likelihood of having a sleep disorder[4]. Then propensity score matching is utilized to create matched groups of individuals with and without sleep disorders who have similar or "balanced" covariate distributions. This step reduces bias by ensuring that the groups being compared are comparable in terms of the covariates. After matching, the balance of covariates between the two groups was assessed to ensure that the matching process was successful. A balanced distribution of covariates indicates that the groups were comparable[4]. The analysis was further refined by subclassifying or stratifying the study population into subgroups based on their propensity scores[5]. This approach allowed me to examine the effects of sleep disorders and lifestyle habits within more homogeneous subgroups. Lastly, by using the propensity score itself as a covariate in a multivariate analysis, it accounts for the propensity of individuals having sleep disorders as a covariate in the regression models, ensuring that the effect of sleep disorders is estimated while controlling for potential confounding variables[4]. Through these steps of analysis, I have taken a thorough approach to minimize bias and confounding, thus making the study result more robust and reliable.

RESULTS & DISCUSSION

The main objective of this observational study is to investigate the effects of sleep disorder on sleep quality, this was done so by estimating the propensity scores **[Table 1]**, conducting propensity score matching, subclassification and covariance adjustment to control for potential biases.

[Table 1]

	T1 <dbl>	T0 <dbl>
years	46.63	39.04
gender	33.55	62.56
sleep.duration	6.81	7.36
physical.activity	60.90	57.95
stress.level	5.77	5.11
BMI.category	7.74	83.56
heart.rate	71.79	69.02
daily.steps	6765.81	6852.97

[Result 1]

```
Estimate...  -0.43011
AI SE......  0.32345
T-stat.....  -1.3297
p.val......  0.1836

Original number of observations..............  374
Original number of treated obs...............  155
Matched number of observations...............  155
Matched number of observations  (unweighted).  278
```

[Result 3]

```
        Two Sample t-test

data:  data2$Quality.of.Sleep[as.factor(data2$Sleep.Disorder) == 0] and
data2$Quality.of.Sleep[as.factor(data2$Sleep.Disorder) == 1]
t = 6.311, df = 372, p-value = 7.892e-10
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.5194851 0.9897209
sample estimates:
mean of x mean of y
 7.625571  6.870968
```

After matching on covariates the treatment effect (difference in sleep quality between the group that has sleeping disorder and the group that did not have sleeping disorder) is -0.43011, with p-value 0.1836 **[Result 1]**. Then checking for covariate balance, Gender has an absolute standardized difference shift from 61.24 to -28.511, the absolute standardized difference of age has shifted from 96.893 to -37.082, the absolute standardized difference for physical activity level is 14.282 to -36.113. Then when compared to not adjusting for imbalance, the unadjusted treatment effect is -0.754603, with p-value 0 **[Result 2]**. Then through stratification, the 5 quantiles produced treatment effects. The overall treatment effect is -0.3307, which can be obtained by averaging the estimates within each stratum **[Result 3]**.

Overall analysis suggests that there may be a negative association between having sleep disorder and sleep quality, but this association is not statistically significant after matching covariates. Additionally, not accounting for covariate imbalance can lead to biased treatment effect estimates, highlighting the importance of covariate adjustment in observational studies. Stratification offers a middle-ground approach that considers the effects within different groups of the population.

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          7.45333    0.08663  86.040   <2e-16 ***
Sleep.Disorder[strat1] -0.20333    0.38498  -0.528    0.599
```

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          8.26087    0.05175  159.63   <2e-16 ***
Sleep.Disorder[strat2] -0.26087    0.19908   -1.31    0.194
```

```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)          7.3000     0.1626  44.885  < 2e-16 ***
Sleep.Disorder[strat3]  -1.2091     0.4132  -2.926  0.00464 **
```

```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)          7.0000     0.5207  13.442   <2e-16 ***
Sleep.Disorder[strat4]   0.2429     0.5497   0.442     0.66
```

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          6.7143     0.3878  17.316   <2e-16 ***
Sleep.Disorder[strat5]  -0.2220     0.4081  -0.544    0.588
```

[Result 3]

CONCLUSIOIN

The treatment effect of -0.43011 suggests that, on average, individuals with sleep disorders have slightly lower sleep quality compared to those without sleep disorders. However, this difference is not statistically significant. Then after propensity score matching, covariate balance has improved substantially, as indicated by the reduced absolute standardized differences. This indicates that matching has successfully controlled for covariate imbalances. The unadjusted treatment effect of 0.754603 suggests a larger negative impact of sleep disorders on sleep quality when not considered covariate imbalances. But this result may be misleading because it does not account for differences in covariates between the two groups. Lastly the overall treatment effect obtained through stratification is -0.3307, which falls between the values of Result 1 and Result 2.

References

1. Henrich, Liv C., et al. "Sleep Quality in Students: Associations with Psychological and Lifestyle Factors." *Current Psychology (New Brunswick, N.J.)*, vol. 42, no. 6, 2023, pp. 4601–08, https://doi.org/10.1007/s12144-021-01801-9.
2. Brown, V., and A. Arikawa. "The Relationship Between Caffeine Intake, Lifestyle Factors, Sleep Quality and Perceived Stress in College Students." *Journal of the Academy of Nutrition and Dietetics*, vol. 121, no. 9, 2021, pp. A35–A35, https://doi.org/10.1016/j.jand.2021.06.100.
3. Bukowska, Agnieszka, et al. "Rotating Night Shift Work, Sleep Quality, Selected Lifestyle Factors and Prolactin Concentration in Nurses and Midwives." *Chronobiology International*, vol. 32, no. 3, 2015, pp. 318–26, https://doi.org/10.3109/07420528.2014.975353.
4. Singh, M (2023). Lecture slides. "Module 7: Casual Inference – Observational Studies."
5. Taback, N. (2022). "Design and Analysis of Experiments and Observational Studies using R: A Volume in the Chapman & Hall/CRC Texts in Statistical Science Series." http://designexptr.org/index.html
6. Zwart, Tom C., et al. "Long-Term Melatonin Therapy for Adolescents and Young Adults with Chronic Sleep Onset Insomnia and Late Melatonin Onset : Evaluation of Sleep Quality, Chronotype, and Lifestyle Factors Compared to Age-Related Randomly Selected Population Cohorts." *Healthcare (Basel)*, vol. 6, no. 1, 2018, p. 23–, https://doi.org/10.3390/healthcare6010023.