Statistical Learning for the Analysis of Multimodal Sleep in Older Men

by

Jianhui Zhu

MD, Nantong University, China, 1984

MMed, Fudan University, China, 1990

PhD, The Chinese University of Hong Kong, China, 2000

Submitted to the Graduate Faculty of

the Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Jianhui Zhu

It was defended on

July 25, 2019

and approved by

Thesis Advisor:

Robert Krafty, PhD
Associate Professor, Department of Biostatistics
Graduate School of Public Health, University of Pittsburgh

Committee Members:

Stephen Smagula, PhD
Assistant Professor, Department of Psychiatry
University of Pittsburgh

Chung-Chou H. Chang, PhD
Professor of Medicine, Biostatistics, and
Clinical & Translational Science
School of Medicine and Graduate School of Public Health
University of Pittsburgh

Copyright © by Jianhui Zhu

2019

Robert Krafty, PhD

Statistical Learning for the Analysis of Multimodal Sleep in Older Men

Jianhui Zhu, MS

University of Pittsburgh, 2019

Abstract

Introduction: Sleep is essential for human development and maintaining physical and mental health. Sleep disturbances have long been known to be associated with mental illness, metabolic, neurological or other systems diseases. Knowing what factors are associated with sleep quality and sleep-wake homeostasis is important for the study of sleep disorders and may potentially inform new treatment strategies to preserve patients' normal sleep-wake cycle. The present study aims to identify what actigraphic measures, self-reported sleep variables, and other chronic diseases, medications are related to the percentage of slow-wave sleep and delta power spectra in older men.

Method: Categorical variables are summarized using frequencies and percentages. For continuous variables, means and standard deviations are computed, and distributions are displayed in histograms. Possible correlations among variables are examined by a matrix of scatterplots and Pearson correlation coefficients. The LASSO is used for feature selection in multiple linear regression models and multiple imputation used to overcome missing data.

Results: The past month sleep hours (β=0.0896, $p<0.05$), kidney diseases (β=0.161, $p<0.05$) and oral corticosteroids (β=0.148, $p<0.05$) are significantly positively associated with percentage of deep sleep, while sleep apnea severity (β=-0.0043, $p<0.001$), age (= -0.0042, $p<0.01$), Benzodiazepine use (-0.155, $p<0.001$), NSAIDS use (β=-0.0418, $p<0.05$), and race(β=-

0.0476, $p<0.01$) are negatively associated when controlling other variables' effect. Cognitive function ($\beta=0.0015$, $p<0.001$), and oral corticosteroids ($\beta=0.0733$, $p<0.01$) are positively related to delta power, while sleep apnea severity ($\beta=-0.0011$, $p<0.001$), age (= -0.0013, $p<0.05$), mean sleep minutes (-0.0002, $p<0.001$) , BMI (-0.031, $p<0.001$), Diabetes ($\beta=-0.0404$, $p<0.001$), Benzodiazepine use (-0.061, $p<0.001$), and the consumption of alcoholic beverages ($\beta=-0.0125$, $p<0.05$) are negatively related to delta power when controlling other covariates.

Conclusions: Our study suggested several factors are either positively or negatively associated with the percentage of deep sleep and delta power. Most of the factors affect the percentage of slow-wave sleep and delta power in the same direction.

Public Health Significance: These analyses may provide important messages for future study and potential medical interventions application.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Sleep plays an important role in maintaining the body's circadian rhythm, physical and mental health [1-4]. Sleep cycles consist of non-rapid eye movement (NREM) sleep, which is subdivided into the three stages N1, N2, and N3, and rapid eye movement sleep (REM). Polysomnography (PSG), or the comprehensive recording of electrophysiological changes that occur during sleep, is used by clinicians and researchers for evaluating sleep. Included are electroencephalography (EEG), which records brain activity, and is used to define sleep-wake cycles, sleep stages, and sleep waves. **Figure 1**[5] displays typical EEG signals during different sleep stages. In particular, N3 stage sleep is characterized by low-frequency waves in the 0.5–4.5 Hz range and is often referred to as delta wave, slow-wave sleep (SWS), or deep sleep. This stage is important for health and functioning as it is during N3 sleep when body restoration, growth, and development occurs, immune function is boosted, and energy builds up [6]. A comprehensive sleep profile includes sleep duration and sleep quality [7]. Sleep duration is simply defined by the amount of sleep during the night. The normal sleep time is age dependent. Based on the National Sleep Foundation's recommendations [8], the normal sleep duration range for older adults (> 65 yr.) is 7-8 hours. Short sleepers are generally defined as those who sleep less than 7 hours/night. Sleep quality can be defined in different ways. One of the common assessments for sleep quality and disturbance is the Pittsburgh Sleep Quality Index (PSQI), a self-rated questionnaire [9]. This assessment is simple, easy, and cost-effective, but it is subjective and may not be correlated with objective measures of sleep [10]. Objective measures of sleep by PSG can characterize sleep quality by the total sleep time (TST), time spent awake after sleep onset (WASO), sleep efficiency (the ratio of TST to time in bed), and the amount of

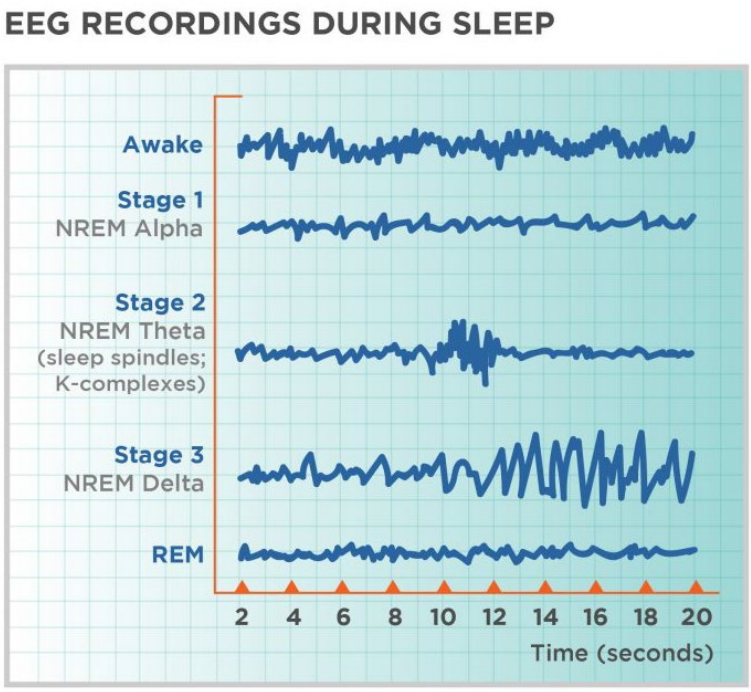

Figure 1 Representative EEG recording during sleep

Slow-wave sleep (SWS, stage 3) and rapid eye movement (REM) sleep [11]. Since SWS sleep plays important roles in cerebral restoration, the maintenance of sleep, and memory consolidation in humans, [12], the percentage of total sleep time that is in stage 3 and 4 is a critical measure of deep sleep. Another useful measure computed from the EEG channel of PSG is the amount of delta power during the night. Delta power is correlated with the change of sleep duration and intensity, which are considered as a critical measure for deep sleep [13]. It is also related to prior sleep quality and wakefulness and acted as an index for sleep-wake homeostasis. Sleep deprivation will evoke an increase of delta power, while excess sleep leads to decrease of delta power [13]. Low levels of delta power are related to reduced sleep quality in aging patients and patients with diseases such as diabetes [1]. Delta wave analysis is used to study the homeostatic process in the two-process model of sleep regulation [14].

Disturbance of sleep such as abnormal sleep duration and poor sleep quality is correlated with cardiovascular diseases, metabolic abnormalities, neuronal function, diabetes, depression, falls, accidents, impaired cognition, a poor quality of life, and even mortality risk for some diseases [1 7 15-18]. On the other hand, many diseases such as cancer, neurodegenerative diseases including Parkinson's and Alzheimer's, medications, cardiovascular diseases, metabolic diseases, diet, cocaine, or even lifestyle changes and environmental changes could compromise sleep patterns and sleep quality [1 19-25].

Knowing participants' slow-wave sleep could provide important information related to the patient's general health status, physical or mental stress, and even disease progress. Sleep parameters from PSG are considered definitive, "gold standard" measures [26]. However, PSG is expensive and requires professionals to perform. Actigraphy is one type of wearable technology that uses accelerometer-based transformations to quantify physical activity [27]. They are noninvasive, objective (provides an accurate representation of sleep/wake history), continuous monitoring (for day and night, weeks/months), low cost, and very convenient. If actigraphic sleep-wake activities and other covariates including self-reported sleep information, medication, and chronic disease history can be used to estimate patients' sleep quality, it could not only serve as a screening test for sleep disturbance, but also be used as a monitoring system to evaluate the effectiveness of medical or psychological intervention for some diseases [28].

In this study, we will use statistical learning to analyze multimodal sleep in older men. We aim to identify which actigraphy, self-report, clinical and demographic factors are related to slow-wave sleep and delta power spectra. The results from this study could provide important information for the future sleep study and for potential physical/pharmacological intervention to preserve SWS/delta power in aging.

2.0 Materials and Methods

2.1 Participants

Data are from an NIH-funded study in the Osteoporotic Fractures in Men (MrOS) conducted between December 2003 and March 2005 at six clinical centers in the United States (Birmingham, Alabama; Minneapolis, Minnesota; Palo Alto, California; Monongahela Valley, Pennsylvania; Portland, Oregon; and San Diego, California). 5994 community-dwelling men aged 65 or older were enrolled at 6 clinical centers in a baseline examination between 2000 and 2002. After excluding participants who had used overnight nocturnal oxygen therapy or positive pressure/oral appliances for the treatment of sleep apnea or other sleep problems, 3135 participants were recruited to the Sleep Study. These participants underwent full unattended PSG and 3 to 5-day actigraphy studies between December 2003 and March 2005. However, our original dataset included 5994 observations and 44 variables.

2.2 Measures

2.2.1 Outcome Measures

There are two outcome variables that objectively measure "deep sleep". The first one is the percentage of total sleep time spent in stage 3 and stage 4 sleep, which was calculated as the participants' slow-wave sleep time in NREM, N3, divided by total sleep time per night. The

secondary outcome variable is delta power, which was computed from the EEG delta wave signal by fast Fourier transform with Welch's method after artifact removal with automated pipeline [29]. The average absolute delta spectral power density in log10(uV^2/Hz) was calculated for each sleep cycle within the delta band of frequencies between 1.25-4 Hz.

2.2.2 Rest-activity Circadian Rhythm (RAR) Variables

The RAR is measured by actigraphy, which are wearable devices to record acceleration and provide indirect measures of physical activity over time. Participants were trained to use the Actiwatch 2 device. Their daily regular sleep-wake hours and rest-activity were recorded. To avoid masking of the rest-activity cycle, they are advised to avoid alcohol and caffeine during the data collection week. Sleep and activity levels were primarily calculated using the Respironics Actiware 5.70.1 software (Philips Respironics®, Andover, MA, USA), and recordings per minute for six complete days and seven nights were used for the analysis [27].

A common parametric approach to analyzing RARs includes fitting a 5-parameter extended cosine model [30]. A new approach referred to as the residual circadian spectrum (RCS) has been reported recently [31]. The RAR measures included in this study are 1. Mean minutes – sleep in bed; 2. Mean sleep efficiency – calculated as the percentage of sleep time divided by time in bed; 3. Amplitude – peak-nadir difference; 4. MESOR – middle of modeled rhythm; 5. Phi – the time of day of the average peak activity over the week with later peak times indicating later timing of high activity levels. 6. Pseudo F – the extended cosine model also provided for a measure of how well the observed activity data are fitted by the 24-hour rhythm model. For the pseudo F, lower values indicate poorer model fit, which suggests that the rhythm may be erratic, and/or variable [27]. Some of these measures are illustrated in **Figure 2** cited from the published paper [25].

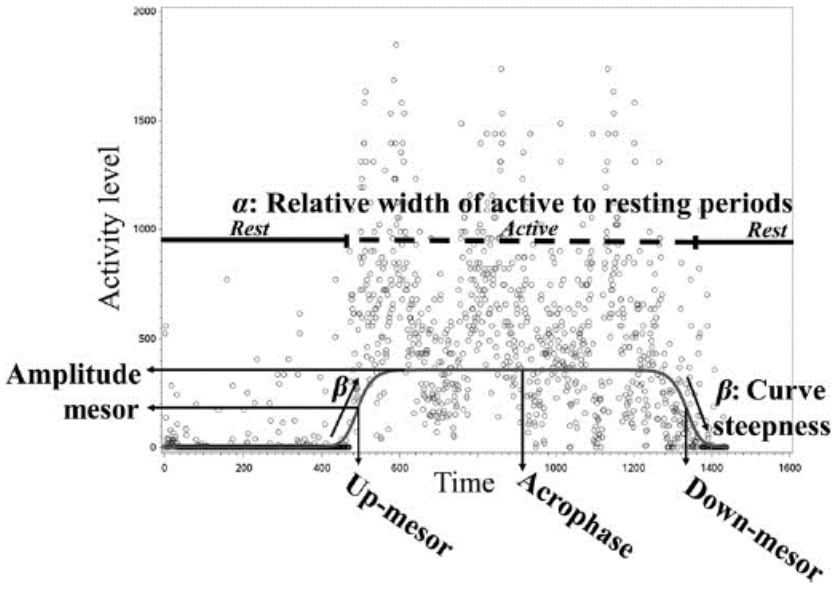

Figure 2 Illustration of 5 parameter extended cosine model

2.2.3 Self-reported Sleep Variables

Self-reported sleep variables include participants' last month's daily sleep hours (during the past month, how many hours of actual sleep did you get each night?) and sleep efficiency (sleep time divided by total time in bed).

2.2.4 Chronic Diseases, Medications, and Other Covariates

The covariates we considered in this study include demographic/lifestyle factors (age, race, education, weekly alcohol consumption, smoking status, body mass index (BMI), self-reported physical activity, mental health characteristics (Teng 3MS score, range from 0 to 100), depression (geriatric depression score, GDS, >5 suggesting, ≥10 depression), anxiety symptoms (Goldberg Anxiety Scale score, ranges from 0 to 9, the higher the score, the more likely anxiety), chronic

diseases (Peripheral vascular disease, Osteoarthritis, Rheumatoid Arthritis, Hypertension, Stroke, Angina, Congestive Heart Failure, Myocardial Infarction, Diabetes, COPD, Parkinson's, Renal disease, Cataracts, Liver disease), current medication use (antidepressant such as non-benzodiazepine non-barbiturate sedatives /hypnotics and benzodiazepines use), nonsteroidal anti-inflammatory drugs (NSAID) use, and corticosteroid use (oral or nasal/inhaled). The severity of sleep-disordered breathing was assessed with obstructive apnea-hypopnea index, defined as the number of respiratory events with oxygen desaturation ≥ 3% per hour.

2.3 Missing Data

The overall data inspection was analyzed by the Amelia II program [32]. Participants who have missing data for most of the measures will be excluded from the dataset. After that, all remaining missing data were multiply imputed via the Amelia II program which use Expectation-Maximization Bootstrap-based algorithm (EMB) with consumption of multivariate normal distribution. Five imputed datasets will be generated, and all the data sets were used for modeling. The mean of the parameters was calculated and reported.

2.4 Statistical Analysis

The data contains two continuous outcomes variables and over 40 predictors and covariates. In exploratory analyses, categorical variables are summarized using frequencies and percentages. Continuous variables are computed for their mean and standard deviation; their

distributions are displayed in histograms. The possible correlation between continuous variables is examined by a matrix of scatterplots. Pearson correlation coefficients are calculated. A multiple linear regression model will be fitted as follows:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i$$

For $i \in \{1 - n\}$, where

y_i is the real-valued response for the i-th observation

β_0 is the regression intercept

β_j is the j-th predictor's regression slope

x_{ij} is the j-th predictor for the i-th observation

ε_i iid N (0, σ^2) is a Gaussian error term

The regression model is fit using the LASSO, which is a method for shrinkage and selection for regression and generalized regression problems [33] to perform multiple variables regression. LASSO was computed using glmnet package for R statistical software, version 1.1.14 as per the following equation:

$$\min_{\beta_0 \beta} \left(\frac{1}{2N} \sum_{i=1}^{N} \left(y_i - \beta_0 - x_i^T \beta \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right)$$

where

N is the number of observations

yi is the response at observation *i*

xi is data, a vector of *p* values at observation *i*

λ is a nonnegative regularization parameter corresponding to one value of Lambda

The parameters $\beta 0$ and β are scalar and *p*-vector, respectively

The LASSO problem involves the $L1$ norm of β. As λ increases, the number of nonzero components of β decreases. The best penalty tuning parameter of the LASSO λ was chosen by the program based on 10-fold cross validation that generated the smallest average MSE. The model per imputed data set that corresponded to the optimal penalty was referred to as "best". The multiple linear models were run after LASSO performed feature selection with computation of inferences. The results from multiple imputations are combined using the Zelig package [34].

To check the overall goodness of fit, I tested for normality of residuals by plotting the observed values against the predicted values, accumulated residuals distribution, and Q-Q normality plot.

2.5 Software

All analyses were implemented using the RStudio statistical software, version 1.1.141. The Amelia II package [32] was used to performing missing data mapping and imputation. Variable selection and model fitting was performed using the glmnet [35] and Plotmo package. Multiple linear regression is fitted using ordinary least squares after LASSO feature selection. The results from multiple imputations are combined using the Zelig package[34].

3.0 Results

3.1 Data Inspection and Manipulation

Our raw dataset has 5994 observations and 44 variables. When checking the data using the Amelia II package, there is a total of 44% of missing values (**Figure 3**). Clearly, most of these participants who are missing in the outcome variable percentage of slow-wave sleep also have missing data in most of the other variables. These observations were excluded from the dataset. After removing these observations, there was 1% missing values (**Figure 4**). In the variable level, data was missing in the range of 0 to 14.7% among the predictors and the outcome variable (delta power). All remaining missing values were multiply imputed via Amelia II program as showed in **Figure 5**. A total of 5 datasets were generated for further analysis. For sleep hour data, raw data had hour ranges from 0 to 12 hours. According to the National Sleep Foundation's new recommendation [8], I categorized the sleep hours into three groups – less than 5 hours, 5 to 9 hours, and more than 9 hours).

3.2 Descriptive Analysis of Variables

First, the distribution of the outcome measures, delta power spectra, and the percentage of slow-wave sleep were characterized (**Figure 6 and Table 1**). Delta power is approximately normally distributed (**Figure 6A**), but the percentage of slow-wave sleep distribution was highly skewed (**Figure 6B**). The log transformation (Log (percentage of slow-wave sleep +1)) seemed to

improve the distribution pattern (**Figure 6C**). However, the Shapiro-Wilk test of normality suggests the outcome may not be normally distributed (W=0.93, P-value < 0.001). In a very large

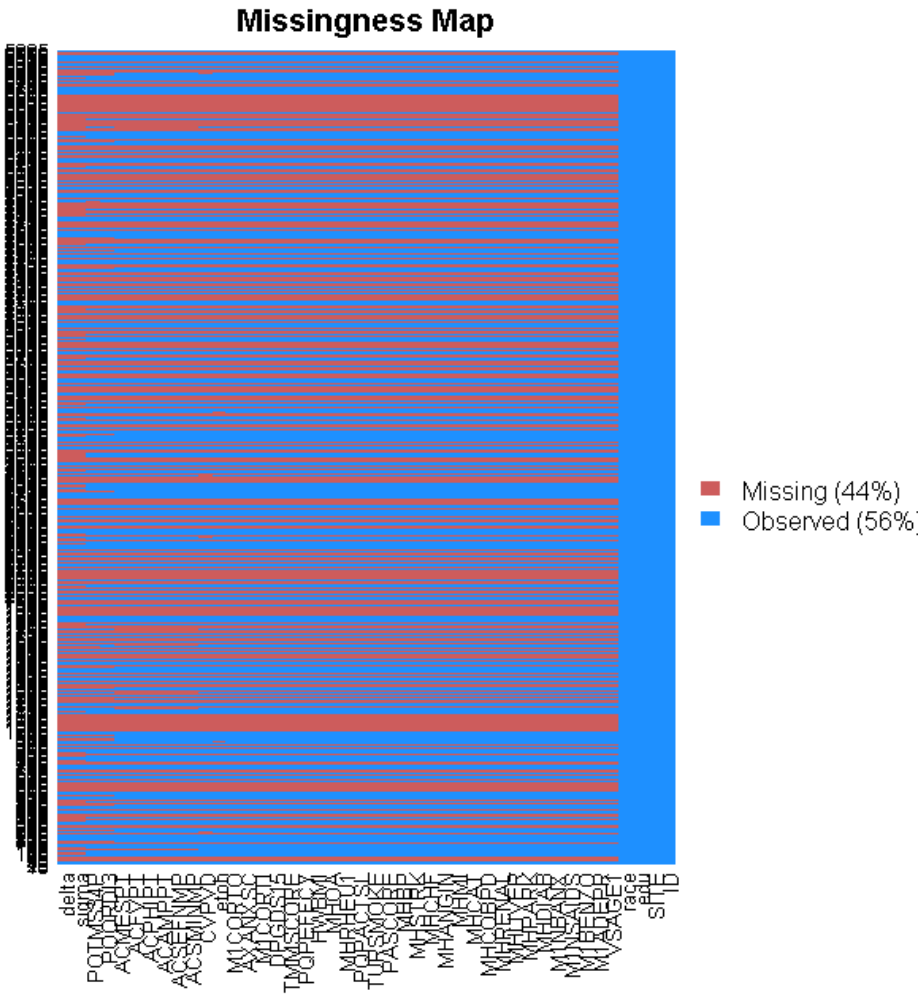

Figure 3 Missingness Map for raw data

sample size, this should be not a concern for linear regression analysis, but log transformation did improve our regression model fitting (see results), therefore we will choose the log-transformed percentage of slow-wave sleep for further analysis.

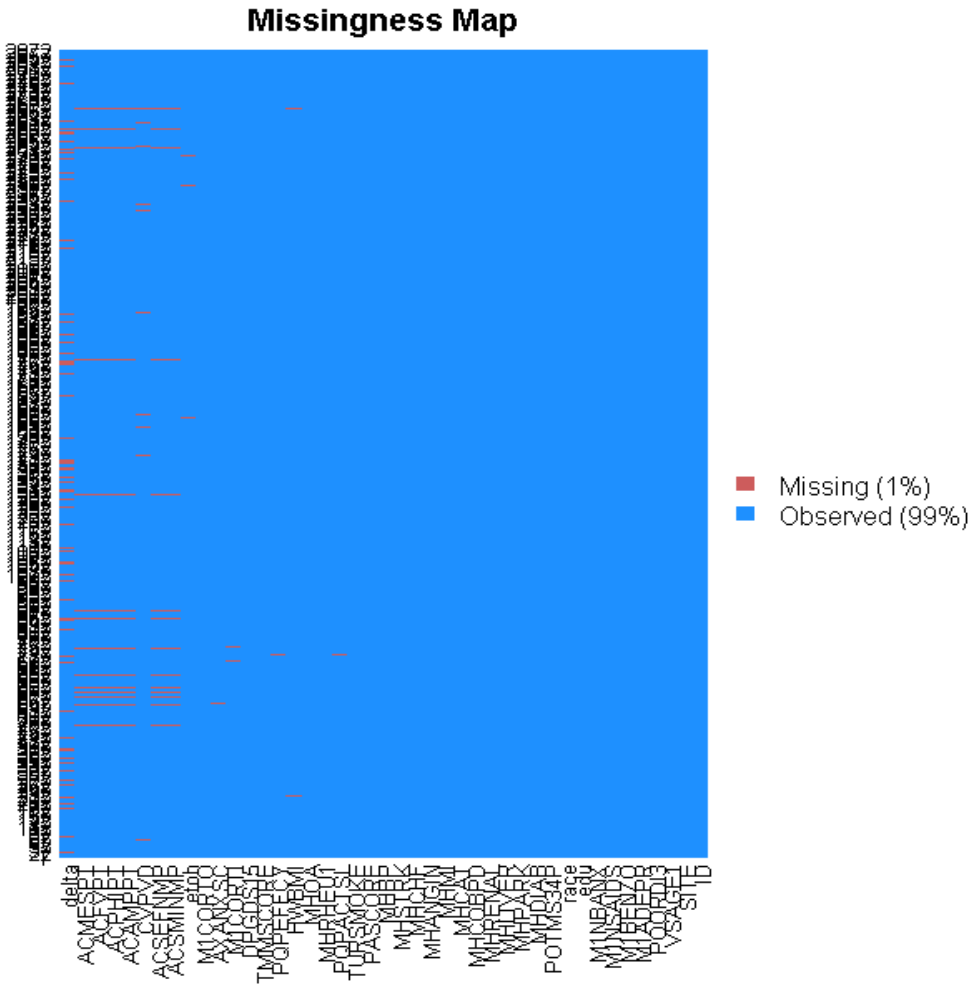

Figure 4 Missingness map after excluded missing value in variables for the percentage of slow-wave sleep

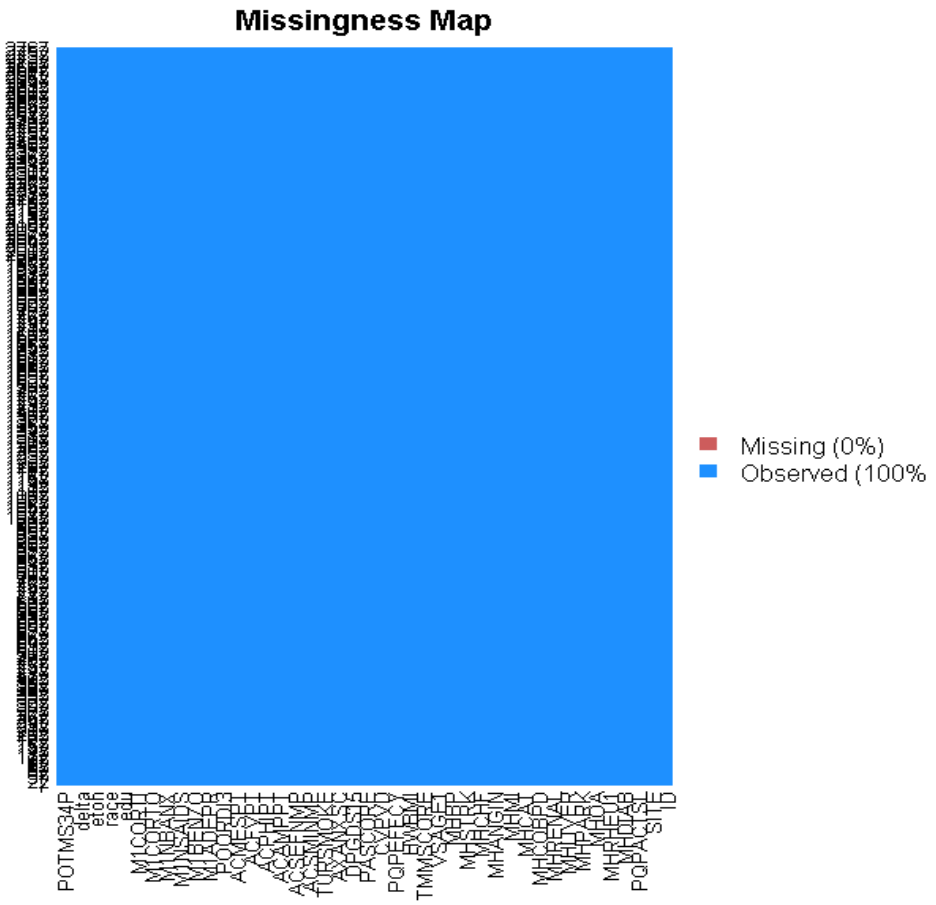

Figure 5 Missingness map after imputation via Amelia II

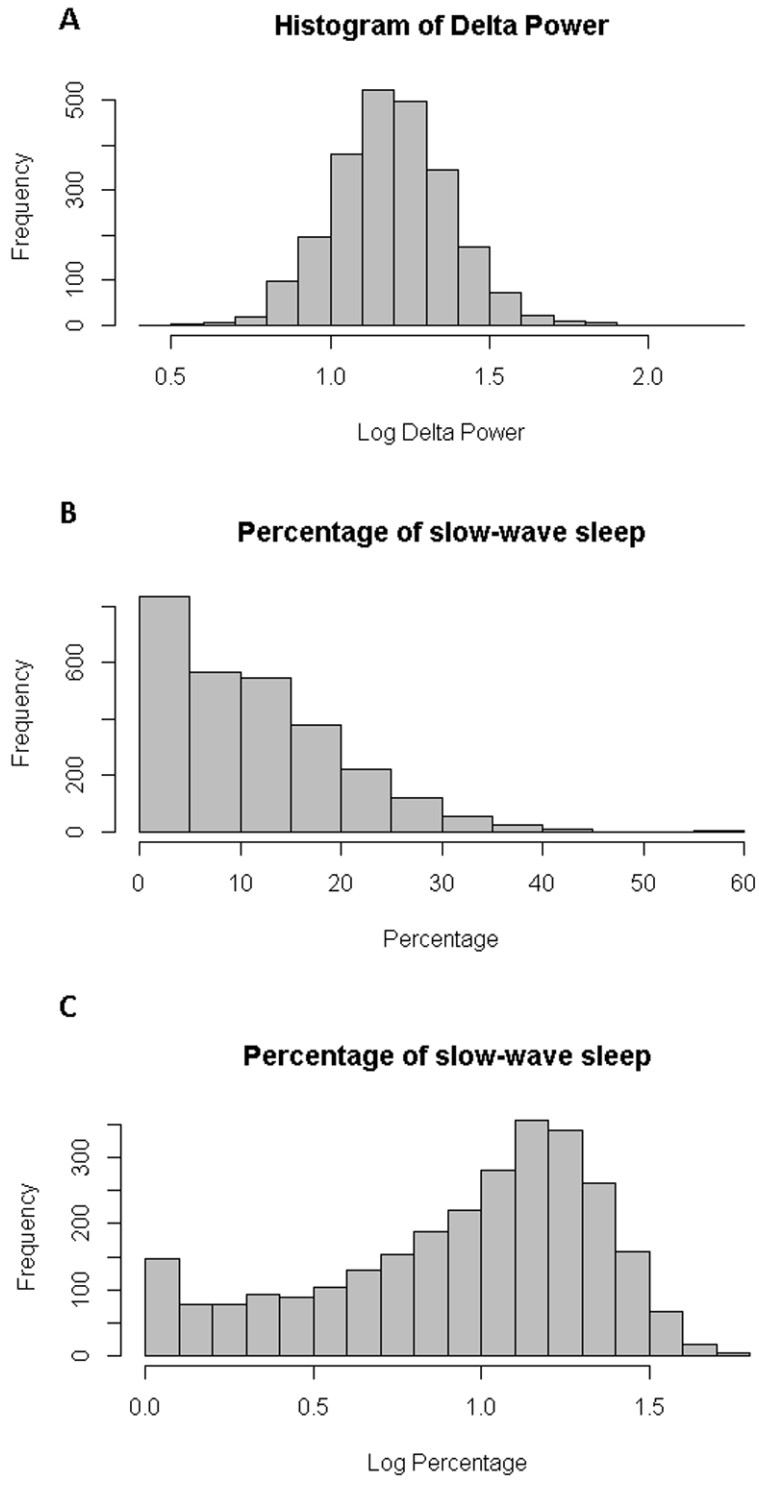

Figure 6 Distribution of delta power and percentage of slow-wave sleep

The Modified Mini-Mental State Test (3MS) scores are highly skewed (**Figure 7a**), The mean 3MS Score is 92.6 with a standard deviation of 6.4. Lower 3MS scores indicate cognitive impairment [27], so we decided to exclude observations with 3MS scores lower than 80 [27] (3.7% of the participants were excluded) to avoid possible interference with our analysis (**Figure 7b**).

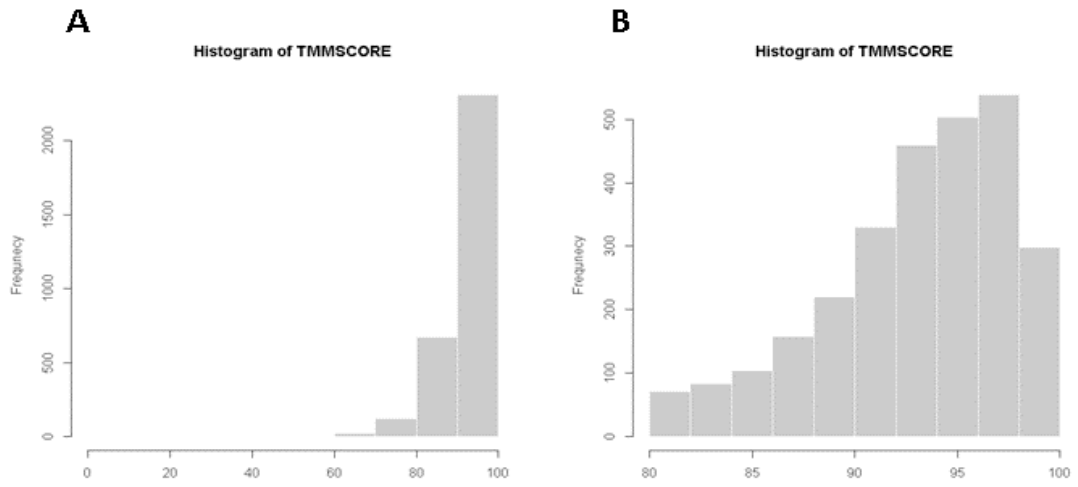

Figure 7 Distribution of Participants' Mental State Test Score before (A) and after (B) excluded those samples with scores lower than 80

Table 1 Rest-activity Rhythm and Outcomes Variables

Outcomes variables	Mean (SD)
Delta Power Spectrum (Log level)	1.196 (0.186)
Percentage of slow-wave sleep (Log level)	0.934 (0.411)
Rest-Activity rhythm variables	
Mean Minutes Scored as Sleep INBD	385 (72.85)
Mean SLP EFF (%) INBD	78.34 (11.87)
Amplitude Antilogistic	3623 (1089)
PHI Antilogistic	14.26 (1.19)
Pseudo F Antilogistic	1047 (505.6)
MESOR Antilogistic	2173 (506.8)

The final dataset contains 2767 observations. The mean age is 75.24 years old. 90.67% of participants were white; black and other ethnicities amounted to 3.38% and 5.95% of the

sample, respectively (**Table 2**). 96.09% of last month's daily sleep hours fell in the 5 to 9 hours group, which is considered a normal sleep time for those over 65 years old. [8]. Alcohol consumption and smoking information are also characterized in **Table 2**. The mean score of the Teng mini-mental state examination after excluding those with scores less than 80 is 93.47 with a standard deviation of 4.46. Using this cut-off could prevent possible biased self-reported information used in the analysis by some older participants who have some mental diseases. The mean value, standard deviation, and histograms of these self-reported variables are summarized in **Table 2** and Supplemental **Figure S1**.

The RAR variables provide objective measures for sleep-wake variables. The mean minutes of sleep and mean sleep efficiency are 385 (± 72.85, SD) minutes and 78.34% (± 11.87, SD), respectively (**Table 1**). Phi, an index for the peak of activity, higher values of which may indicate a more delayed rhythm [36], has a mean of 14.26 and a standard deviation of 1.19. The Pseudo F, provided by the extended cosine model, has a mean of 1057 with a standard deviation of 505.6. The higher the value of Pseudo F, the more robust the rhythm and better cosine model fitting [27]. Amplitude and MESOR have means of 3623 (±1089) and 2173 (±506.8), respectively.

All RAR measures included in this study are shown in **Table 1** and Supplemental **Figure S2** (histogram). The participants' chronic diseases and current medication use information are summarized in **Table 3**.

Table 2 Sample Characteristics

Age (yr.)	76.24(5.46)
Body Mass Index	27.19(3.78)
Education (n, %)	
< High school	126(5.22)
High school	441(16.30)
> High school	2200(78.48)
Race (n, %)	
White	2532(90.67)
Black	79(3.38)
Other	156(5.95)
Alcohol drinks (n, %)	
≤1 drink / week	1267(46.02)
2-13 drinks/week	1336(48.53)
≥14 drinks / week	150(5.45)
Smoke (n,%)	
Never	1101(39.80)
Ever	1610(58.21)
Current	55(1.99)
Past month sleep hour (n, %)	
< 5	106(3.83)
5-9	2659(96.09)
>9	2(0.07)
Sleep efficiency	86.23(12.17)
Physical activity score	146.5(71.04)
Geriatric depression score	1.71(2.11)
Goldberg anxiety scale score	0.90(1.90)
Mental state examination score	93.47(4.46)
Obstructive apnea-hypopnea index	16.86(14.54)

Mean (standard deviation) shown unless otherwise noted

Table 3 Chronic Diseases and Medication Use

Chronic Diseases	n	NO (%)	YES (%)
Diabetes	2766	86.98	13.02
Osteoarthritis	2765	92.08	8.92
Rheumatoid Arthritis	2765	75.73	24.27
Parkinson Disease	2766	98.95	1.05
Liver Disease	2766	97.87	2.13
Kidney Disease / Failure	2766	98.99	1.01
COPD / Emphysema	2766	94.87	5.13
Cataract	2766	53.62	46.38
Heart Attack	2766	82.79	17.21
Angina	2766	85.14	15.86
Congestive Heart Failure	2766	94.03	5.97
Stroke	2766	96.28	3.72
Hypertension	2766	50.25	49.75
Peripheral Arterial Disease	2723	90.12	9.88
Medications			
Antidepressant	2767	92.48	7.52
Benzodiazepine	2767	95.59	4.41
NSAID	2767	79.40	20.60
Nonbenzo Nonbarbituate Sedative Hypnotic	2767	97.98	2.02
Oral Corticosteroids	2758	98.19	1.81
Inhaled / Nasal Corticosteroids	2762	92.34	7.66

Multicollinearity of continuous variables is examined by a matrix of the scatterplot. **Figure 8** displays the correlation between actigraphic variables and outcome variables. **Figure 9** displays the correlation between outcome variables and other continuous variables. The Pearson correlation coefficients were computed (Supplemental **Table S1**). As expected, there are positive correlations between delta power and percentage of slow-wave sleep (ρ =0.63), between mean minutes scored as sleep and mean sleep efficiency (ρ=0.76), between amplitude and MESOR (ρ=0.80), and between Pseudo F and amplitude (ρ=0.55), because some of them were computed by use of the same raw/signal parameters, such as amplitude using peak – nadir activity difference, while MESOR is mean activity.

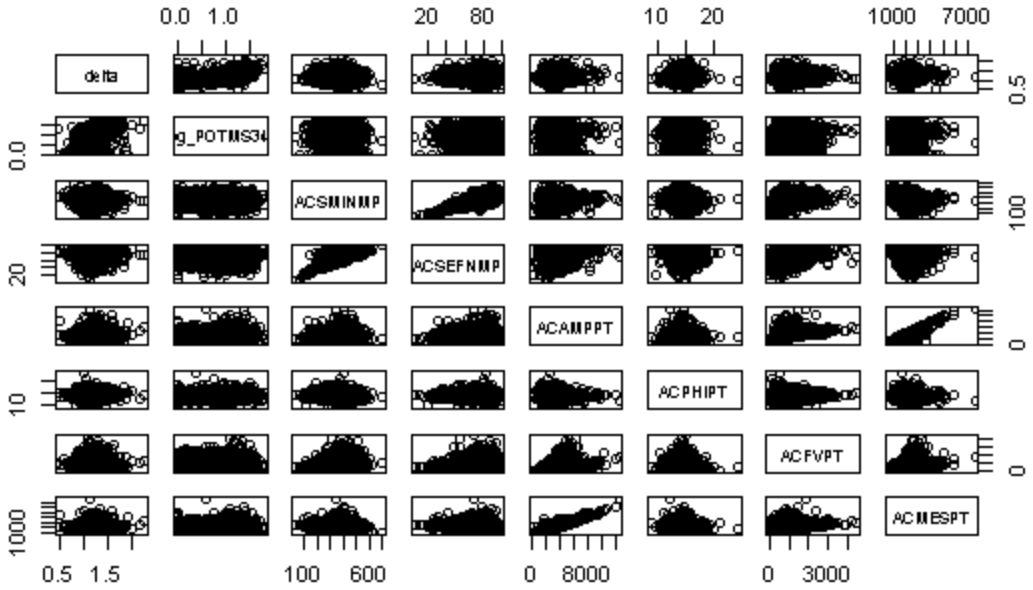

Figure 8 Correlation among actigraphic variables and outcome variables

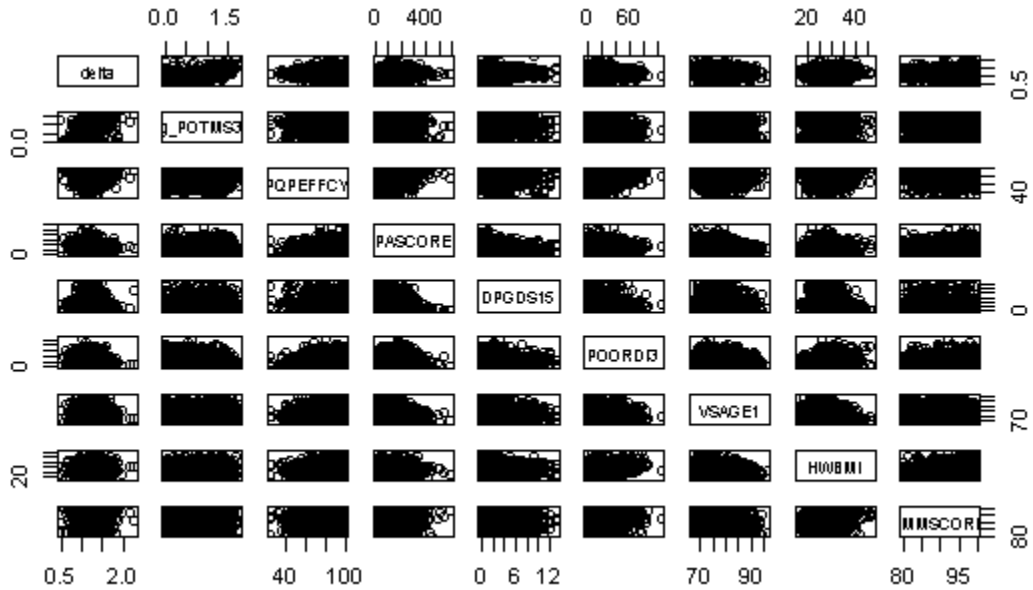

Figure 9 Correlation among outcome variables and other continuous variables

3.3 LASSO regression for the percentage of slow-wave sleep

The multiple variants regression analysis was performed by using the glmnet package with default settings (standardize = True, program decide the range and tuning parameter λ). The best value of λ was chosen by 10-fold cross-validation based on minimum mean squared error. **Figure 10** provides an illustration of 10-fold cross-validation applied to the LASSO fits of our dataset. The left vertical dashed line indicates a λ that gives the minimum mean cross-validated error, and the right vertical dashed line indicates a λ that gives a model such that the error is within one standard error of the minimum.

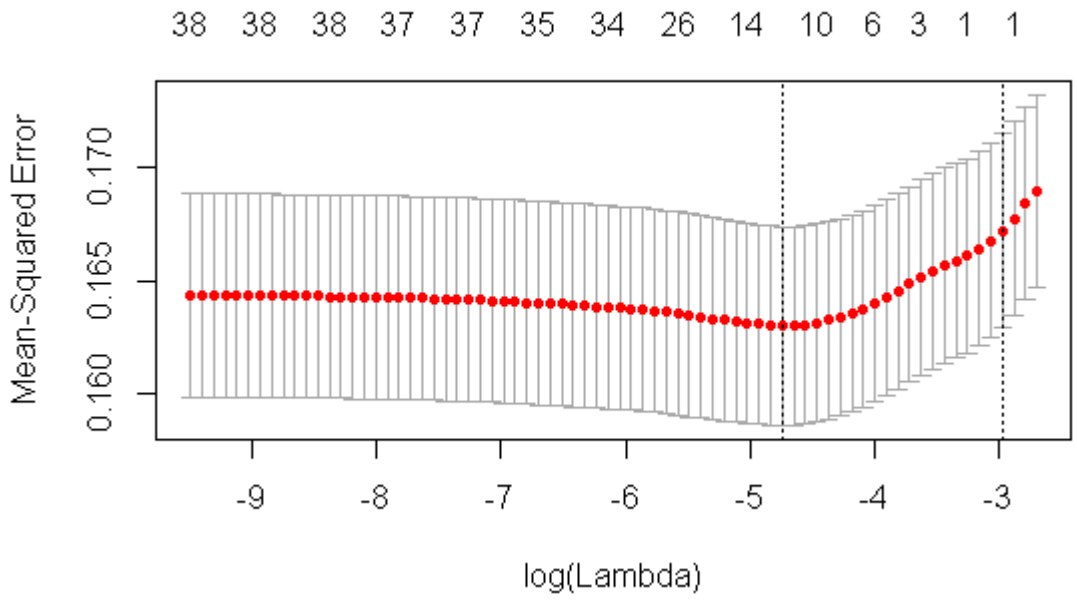

Figure 10 Ten-fold cross-validation MSE for the LASSO, applied to the dataset to select the lambda for the percentage of slow-wave sleep

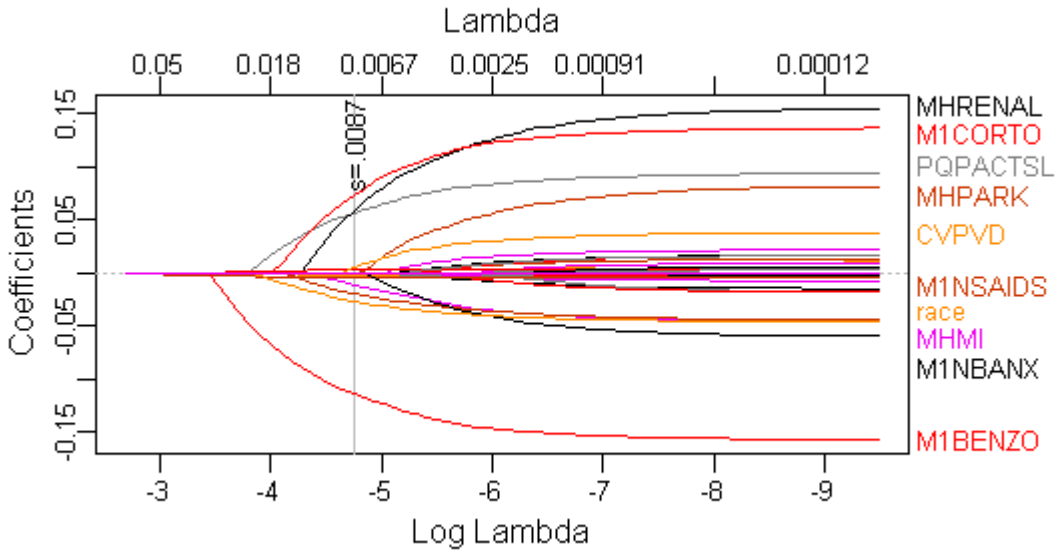

Figure 11 The corresponding LASSO coefficient estimates for the percentage of slow-wave sleep

Figure 11 displays the coefficient estimates after variant selection and dimension shrinkage applied to the LASSO model. The vertical line indicates the best λ value for this model. Each line represents one of the explanatory variables and its role in the model. In the plot, we can see when each variable entered the model and to what extent they influenced the response variable. Entering the model early and steadily with a large coefficient plays a big role in the model. In this model, the benzodiazepine, commonly used for depression and anxiety, is the first variable to enter the model. It has a negative effect on deep sleep, while oral corticosteroids, kidney disease, and past month sleep hours have a positive effect on predicting the percentage of deep sleep. We use the same strategy to repeat all five multiple imputed data. The averaged results are shown in **Table 4**. We also ran a multiple linear regression model and its estimation using ordinary least squares (OLS) for variables selected by the LASSO process to compute the p-values. The results from multiply imputed datasets are combined to reflect the average change using Zelig package [34]. The

results and comparison between multiple linear regression and LASSO regression are shown in **Table 4.** After controlling for all other covariates, the past month sleep hours are positively associated with the percentage of slow-wave sleep (β_{LASSO}=0.05686, p=0.024), while the obstructive apnea-hypopnea index is negatively associated with the percentage of slow-wave sleep, conditional on controlling other covariates (β_{LASSO} =-0.0039, p<0.001). Detailed results are shown in **Table 4**. Compared with multiple linear regression, LASSO has much smaller coefficients because the nature of L1 penalty shrinks coefficients toward the null.

Table 4 Models for the Percentage of Slow-wave sleep

	LASSO	Multiple linear regression		
Covariates	Parameter	Parameters	SE	P value
Intercept	0.8606	0.968	0.25	0.00015***
Coefficients				
Past Month Sleep Hours	0.0569	0.0896	0.039	0.024*
Obstructive Apnea Hypopnea Id.	-0.0039	-0.0043	0.0005	0.0000***
Mental State Test Score	0.0028	0.0035	0.0017	0.051
Age	-0.0024	-0.0042	0.0015	0.0057**
Body Mass Index	-0.0004	-0.0026	0.0022	0.234
Kidney Diseases	0.0583	0.161	0.0766	0.0349*
Heart Attack	-0.011	-0.033	0.0204	0.099
Benzodiazepine	-0.114	-0.155	0.0374	0.0000***
NSAID	-0.0199	-0.0418	0.019	0.028*
Oral Corticosteroids	0.0796	0.148	0.057	0.0101**
Race	-0.0266	-0.0476	0.0162	0.0033**

3.4 LASSO Regression for Delta Power Spectra

Similarly, we fitted a LASSO model for delta power spectra. The illustrations of 10-fold cross-validation and selection of parameters are displayed in **Figure 12** and **Figure 13**. The tuning parameter λ for minimum mean squared error in this model is 0.0031 (**Figure 13**). The selected variables based on the λ_{min} are summarized in **Table 5**. The computed p-values after multiple linear regression are also showed in **Table 5**. Our results showed that except a few variables such as diabetes, benzodiazepine, BMI, obstructive apnea-hypopnea index, the majority of variables has very small, non-significant coefficients. Since the dependent variable delta power has 14.7% missing; it may potentially add needless noise when running a regression model using imputed values for delta [37] and could affect the model fitting. Therefore, we created a second dataset (dataset2) by removing missing delta, then performed multiple imputations for other missed predictors. The results compared with those using dataset1 are displayed in **Table 6**. As shown in the table, the majority of coefficients between the two models are very close suggesting that even dependent variable has about 15% of missing data, the mean coefficients calculated from LASSO models using generated multiple imputations still quite reliable. However, the variable selection in this model using the λ_{min} is not very powerful. From **Figure 13**, many covariates around 0 could be just noise, In **Figure 12**, within the λ_{min} to one standard error of the minimum λ range (between two vertical dashed lines), the log (lambda) level increase from -6.1 to about -5.3 does not substantially change the mean squared error, but significantly increases the penalty power. We remodeled the LASSO linear regressing using $\lambda = \exp^{-5.3}$ and the result is shown in **Table 7**. Conditional on controlling other variables, age, BMI, mean sleep minutes, diabetes, benzodiazepine, nonsteroidal anti-inflammatory drugs, and alcohol drinks are negatively associated the delta power, while Parkinson disease and mini-mental state examination score are

positively associated with the delta power. The p-values are also displayed in **Table 7**. Similarly, a combination of regression results from multiply imputed datasets is analyzed using the Zelig. Results are displayed in **Table 7.**

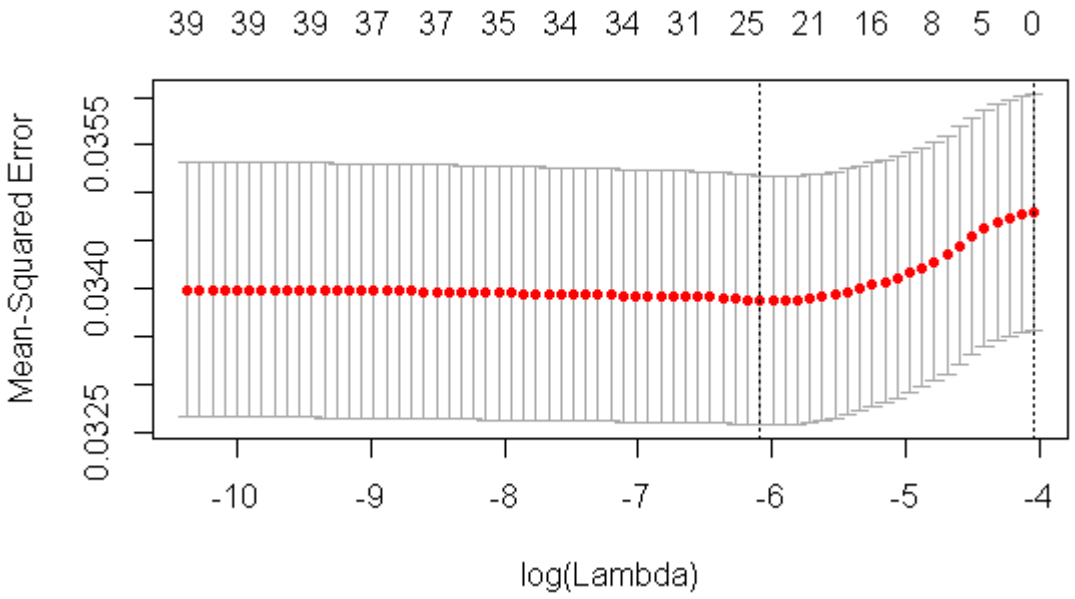

Figure 12 Ten-fold cross-validation MSE for the LASSO, applied to the dataset to predict the Delta Power Spectra

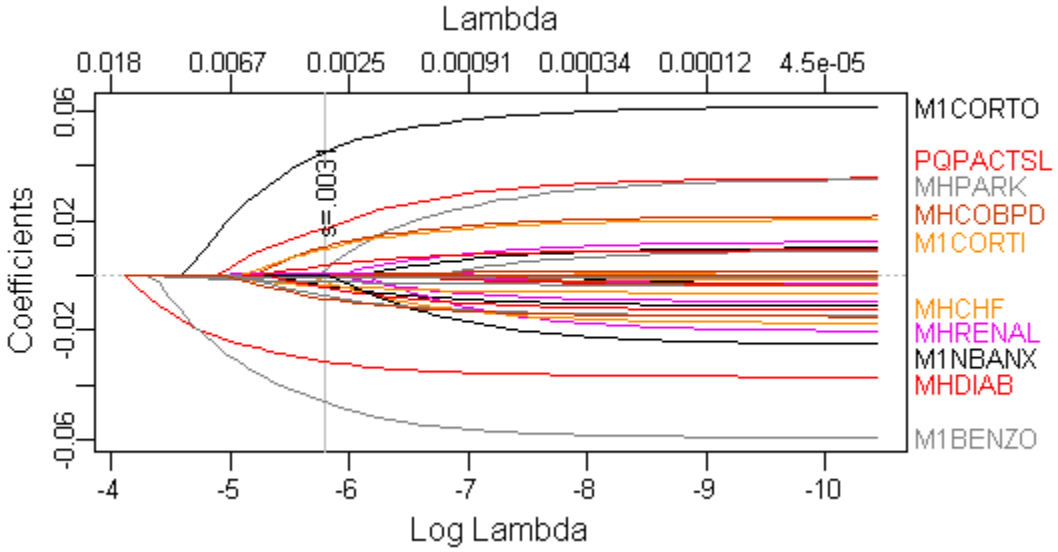

Figure 13 The corresponding LASSO coefficient estimaes for Delta Power

Table 5 Model for Delta Power Spectra using λ_{min}

Covariates	Best Model parameters	P Value
Intercept	1.31147	0.000**
Coefficients		
Past Month Sleep Hours	0.011532	0.099
Obstructive Apnea Hypopnea Index	-0.000938	0.000**
Mental State Test Score	0.000978	0.212
Age	-0.000814	0.0923
Body Mass Index	-0.002471	0.002**
Physical Activity	-0.000032	0.218
Mean Minutes Scored as Sleep INBD	-0.00014	0.127
Mean SLP EFF (%) INBD	-0.000187	0.158
Amplitude Antilogistic	0.000003	0.0506
Diabetes	-0.033098	0.000**
Osteoarthritis	-0.002618	0.222
Parkinson Disease	0.022895	0.2914
COPD / Emphysema	0.010604	0.161
Congestive Heart Failure	-0.005931	0.306
Hypertension	-0.005282	0.432
Benzodiazepine	-0.047638	0.000**
NSAID	-0.008208	0.211
Nonbenzo Nonbarbituate Sedative Hypnotic	-0.013197	0.430
Oral Corticosteroids	0.054667	0.0163*
Inhaled / Nasal Corticosteroids	0.009406	0.145
Race	-0.006050	0.045*
Education	0.003584	0.209
Alcohol Drinks	-0.008557	0.0105*

Table 6 Comparison of Models for Delta Power Spectra Using Two Different Datasets

Covariates	**Parameters in Dataset1**	**Parameters in Dataset2**
Intercept	1.31147	1.2936
Coefficients		
Past Month Sleep Hours	0.011532	0.01257
Obstructive Apnea Hypopnea Index	-0.000938	-0.0009246
Mental State Test Score	0.000978	0.0009483
Age	-0.000814	-0.000906
Body Mass Index	-0.002471	-0.00237
Physical Activity	-0.000032	-0.0000428
Mean Minutes Scored as Sleep INBD	-0.00014	-0.000114
Mean SLP EFF (%) INBD	-0.000187	-0.0000831
Amplitude Antilogistic	0.000003	0.0000017
Diabetes	-0.033098	-0.03277
Osteoarthritis	-0.002618	-0.00374
Parkinson Disease	0.022895	0.0225
COPD / Emphysema	0.010604	0.023028
Congestive Heart Failure	-0.005931	
Hypertension	-0.005282	-0.007239
Benzodiazepine	-0.047638	-0.04633
NSAID	-0.008208	-0.00773
Nonbenzo Nonbarbituate Sedative Hypnotic	-0.013197	-0.0211
Oral Corticosteroids	0.054667	0.04618
Inhaled / Nasal Corticosteroids	0.009406	0.0127
Race	-0.006050	-0.00526
Alcohol Drinks	-0.008557	-0.0073578
EDU	0.003584	0.003989

Dataset1 has missing delta (14.7%) with imputation
Dataset2 excludes the missing delta, then imputed using Amelia II

Table 7 Models for Delta Power Spectra

	LASSO λ_{best}	Multiple linear regression		
Covariates	Parameters	Parameters	SE	P Value
Intercept	1.251	1.351	0.113	0.000***
Coefficients				
Obstructive Apnea Hypopnea Id.	-0.0008	-0.0011	0.0003	0.000***
Mental State Test Score	0.00068	0.0015	0.0008	0.061
Age	-0.0003	-0.0013	0.0007	0.067
Body Mass Index	-0.0017	-0.0031	0.001	0.003**
Mean Minutes Sleep INBD	-0.0001	-0.0002	0.00005	0.000***
Diabetes	-0.0283	-0.0404	0.0109	0.000***
Parkinson Disease	0.00393	0.046	0.0376	0.217
Benzodiazepine	-0.0354	-0.061	0.0183	0.000***
Hypertension	-0.0038	-0.0087	0.0077	0.255
NSAID	-0.0021	-0.0149	0.0091	0.10
Oral Corticosteroids	0.0374	0.0733	0.027	0.007**
Alcohol Drinks	-0.0028	-0.0125	0.0061	0.038*

3.5 LASSO Models Diagnosis

To validate the predicted best model, a residual analysis was conducted. As shown in **Figure 14** for the percentage of slow-wave sleep model, the residuals vs. fitted plot showing the average value of the residuals at each value of the fitted value (red line) is overall flat. There is no clear distribution pattern and the residuals appear to be equally variable across the entire range of fitted values, except for a few values in which their observation IDs are numbered. Considering the very large sample size, this is completely acceptable. Although in the QQ plot, both the upper and lower tales are a little bit off, the overall fit is great. The accumulative residual distribution looks normal, too.

Similarly, the validation of the delta power model is displayed in **Figure 15**. From the residuals vs. fitted plot, QQ plot, and accumulative residual distribution plot, there is no indication

of non-linearity, heteroscedasticity, or substantial deviation from normality. This suggests that this LASSO model fits very well. The variance inflation factor (VIF) values were also examined based on best fit models. VIFs for all included variables are less than 2 (data not shown), indicating there is no concern for multicollinearity in these two models.

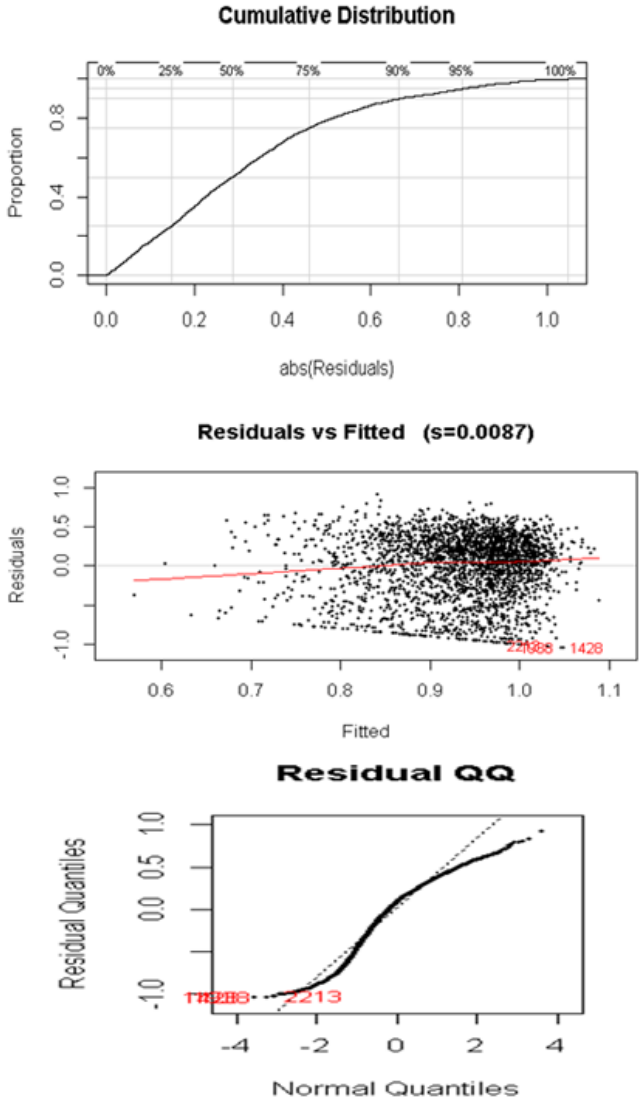

Figure 14 Residual analysis for the percentage of slow-wave sleep LASSO model

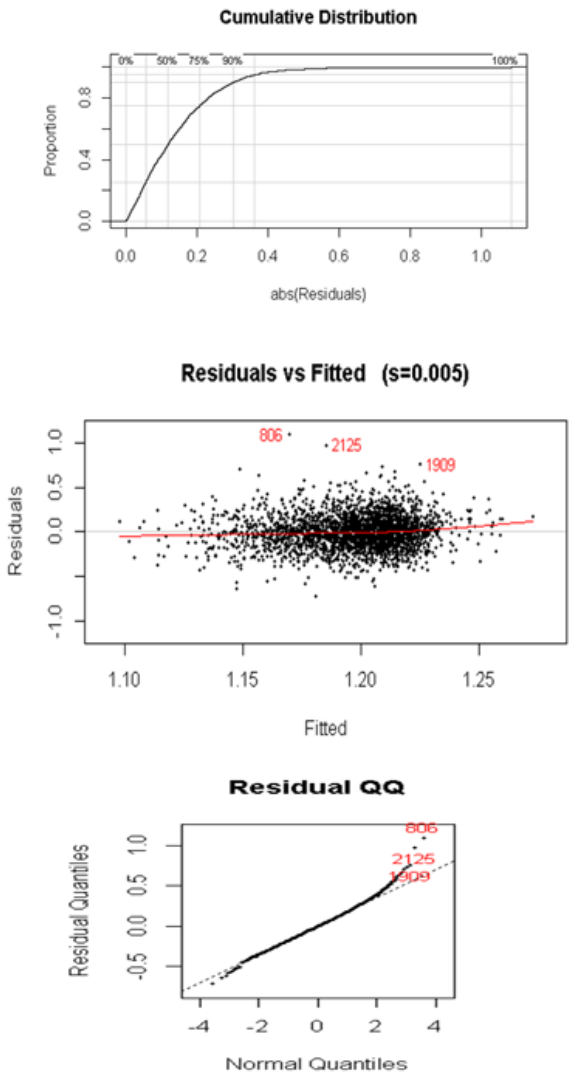

Figure 15 Residual analysis for Delta Power LASSO model

4.0 Discussion and Conclusion

Our final manipulated dataset contains 2767 observations and 44 variables, including 2 outcomes. To identify the relationship between predictor variables and response variables, different multiple regression methods can be used to fit the models, such as ordinary least squares and ridge regression. In order to increase the interpretability and avoid overfitting, it is necessary to reduce the variable dimension and select an appropriate model based on different criteria such as Akaike's Information Criterion (AIC), Schwarz's BIC, R^2, and adjusted R^2 [38]. LASSO is one of the most commonly used statistical learning approaches to perform shrinkage and selection via the L1 regularization method [33]. LASSO selection and shrinkage is much easier and faster computationally than other automatic variable selections such as the forward, backward and stepwise method. The prediction accuracy is high because shrinking and removing coefficients can reduce variance without a substantial increase in bias. Moreover, eliminating irrelevant variables that are not associated with the response variable helps to increase the model interpretability.

In this study, we fit LASSO regression models using optimal tuning parameter λ based on 10 folds cross-validation results to estimate what variables can explain variance in delta power spectra and percentage of slow-wave sleep. The past month sleep hours are positively associated with the percentage of slow-wave sleep. This indicates that people with longer sleep time have more "deep sleep." Although last month's sleep hours are not directly related to current sleep quality, a positive association indicates that it is a useful predictor in the near future for people. The longer the sleep time, the better and higher quality of sleep they will have. Among several actigraphic wake-sleep variables, none of them are associated with the percentage of slow-wave sleep. The amplitude is also a measure for physical activity positive. In general, the actigraphic

measures are well correlated with PSG data, but it may not be very reliable for more specific measures such as sleep efficiency [11 39]. The MESOR is an averaged activity of 24 hours and may not be a clinically meaningful measure because the MESOR value is reflected in both day and night activity [36]. For delta power analysis, the mean sleep minutes is negatively associated with delta power (**Table 7**). Delta power is a measure of slow-wave sleep intensity and related to prior sleep and wakefulness [40]. Prior deprivation of sleep will enhance the delta power, while too much sleep will cause a decrease of delta power [13]. In our study, the negative relationship may simply be an indicator of too much sleep. However, it is noted that delta power can be regulated by non-sleep related effectors such as Benzodiazepine, which can inhibit EEG delta power [13]. Some actigraphic parameters are not included in this study, which include shape parameter alpha (the relative width of the activity peak compared with rest periods), beta (the slope of the curve, with higher values suggesting a sharper transition between rest and activity periods), and down MESOR (approximating the time to 'settle down' for the night) [25]. Fitting these parameters into the models may provide useful information to predict sleep quality.

Aging is usually considered a common factor associated with changes in sleep patterns [41 42]. Our results reveal that an increase in age is negatively associated with both percentage of slow-wave sleep and delta power, which is consistent with previous findings [41 42]. The obstructive apnea-hypopnea index is negatively associated with both the percentage of slow-wave sleep and delta power. The more severe patients' obstructive apnea-hypopnea is, the lower percentage of slow-wave sleep they have, which has an in other previous studies [43].

There is increasing evidence showing a relationship between slow-wave sleep and diabetes. Suppression of slow-wave sleep increases the risk of developing type II diabetes [1]. In our study, diabetes has a positive effect on both delta power and percentage of slow-wave sleep. The

relationship between diabetes and sleep could be bidirectional; a similar phenomenon was described between exercise and sleep [44].

An increase in BMI indicates overweight/obesity, which is associated with an increased risk of many diseases including diabetes and cardiovascular diseases. Our analysis showed an increase in BMI attributed to both the lower level of percentage of slow-wave sleep and delta power. It was suggested that the link between overweight/obesity and altered sleep quality is due to compromised non-rapid eye movement sleep [45].

Interestedly, the race has a negative correlation with delta power and percentage of deep sleep. Compared to white people, black or other ethnic groups displayed lower delta power and a lower percentage of slow-wave sleep, which could be related to lower or unstable socioeconomic status [46]. Alcohol and many medications could interfere with sleep patterns and affect sleep quality [42 47-50]. Our results showed that corticosteroids can increase the deep sleep duration and delta power, which is not consistent with some other studies [24 49], while benzodiazepine and NSAIDS have the opposite effect. The effect of corticosteroids on sleep may not be due to the medication itself but may be related to its treatment and improvement for some underlying diseases. Parkinson's disease is usually related to poor sleep quality [51], but a recent study indicated that a higher accumulated power of slow waves was associated with slower motor progression, particularly of axial motor symptoms in Parkinson's disease. [22]. In our dataset, 1% of participants have Parkinson's disease. It would be interesting to know if slow waves are related to their disease progress and motor function in this group of Parkinson's disease cases. Many diseases are related to sleep disorders including kidney disease [52], heart failure [18], and Alzheimer's disease [53]. Our results indicate a positive association between kidney diseases and percentage of deep sleep. Kidney diseases are a large group of different disorders. Different disorders and different stages

of kidney function could have different clinical phenotypes. Higher Blood urea nitrogen (BUN) is related to daytime sleepiness in chronic kidney disease[54]. It is not clear if the BUN level also attributed the change in slow-wave sleep.

Physical activity has long been associated with better sleep. Exercise is thought to be a nonpharmacologic intervention for disturbed sleep [55-57]. However, some studies report that this correlation is weak [58]. Our data showed physical activity, as well as amplitude, has no significant association with delta power or percentage of deep sleep (**Table 4, Table 7**), This inconsistency could be associated with the time of physical activity. The significant interaction between physical activity and the time spent outdoors was reported [59], which showed increasing time outdoors in the afternoon (versus morning) predicted lower sleep efficiency.

The incompleteness of data is a very common problem for data analysis. Removing missing variables will lose some important information. Multiple imputations (MI) is now widely used to handle missing data. Although imputed data, particularly in the dependent variable, may add some noise for analysis, performing assay using multiple imputations and calculating the mean is quite reliable **(Table 6).**

Although LASSO regression is a very powerful approach to reduce the high dimensionality within a dataset and to identify important variables, LASSO does not evaluate whether you have chosen the correct form of the relationship between the independent and dependent variable(s). The tuning parameter, λ, is an important parameter to determine the model selection. If we choose λ for minimum mean squared error, in the model for the percentage of slow-wave sleep, the λ_{min} selected by the model seems the best according to **Figure 10.** However, in the model for delta power, the penalty power from λ_{min} is weak and many noise variables are included in the initial model (**Figure 13)**. The best λ range is between λ_{min} and one standard error of the λ_{min} or λ_{1se}

(between two vertical dashed lines). In our model (**Figure 12**), the log lambda level increase from -6.1 to -5.3 does not substantially change the mean squared error but significantly increases the penalty power. We remodeled the linear regressing using $\lambda = \exp^{-5.3}$ and the variables selection is much improved and many noise variables including COPD, hypertension, congestive heart failure, and physical activity were removed (**Table 5, Table 7)**. The model fits well, looks simple, and is easy to interpret.

In conclusion, we used the LASSO linear regression to model the correlation between the two outcome variables, percentage of slow-wave sleep and delta power, and 40 other predictors including actigraphic rest-activity rhythm variables, subjective sleep variables, chronic diseases, medications, and other demographic/lifestyle factors. The tuning parameters were selected based on 10-fold cross-validations. The model fitting is appropriate. Our study showed that longer past month sleep hours, good mental performance score, and corticosteroids are positively associated with a higher percentage of slow-wave sleep and high delta power. Age, body mass index, obstructive apnea-hypopnea index, benzodiazepine, race, alcohol drinks, and mean minutes scored as sleep are negatively associated with either percentage of slow-wave sleep, delta power or both. These multimodal regression models help the researcher to identify what factors are associated with participants' two important sleep parameters related to sleep quality and sleep homeostasis and provide important information for future sleep study and potential development of some treatment strategies to preserve SWS/delta power in aging. However, some actigraphic sleep variables such as alpha, beta, and down-Mesor are not included in this study, but they may play an important role to study sleep quality in the future. Possible interactions between variables could exist, such as between medication and some diseases, or between physical activity and actigraphic measures. We did not consider the interaction term in this study but could include this in future

analysis. LASSO performs automatic selection and shrinking, but sometimes this may not be optimal. If there are strong correlations among variables, LASSO may only select one as a parameter group, in this situation, we may consider using sparse group LASSO or elastic net regularization [60]. Scientific background and manual checks could improve variable selection and make the result more meaningful and easier to interpret.

Appendix Supplemental Figures and Tables

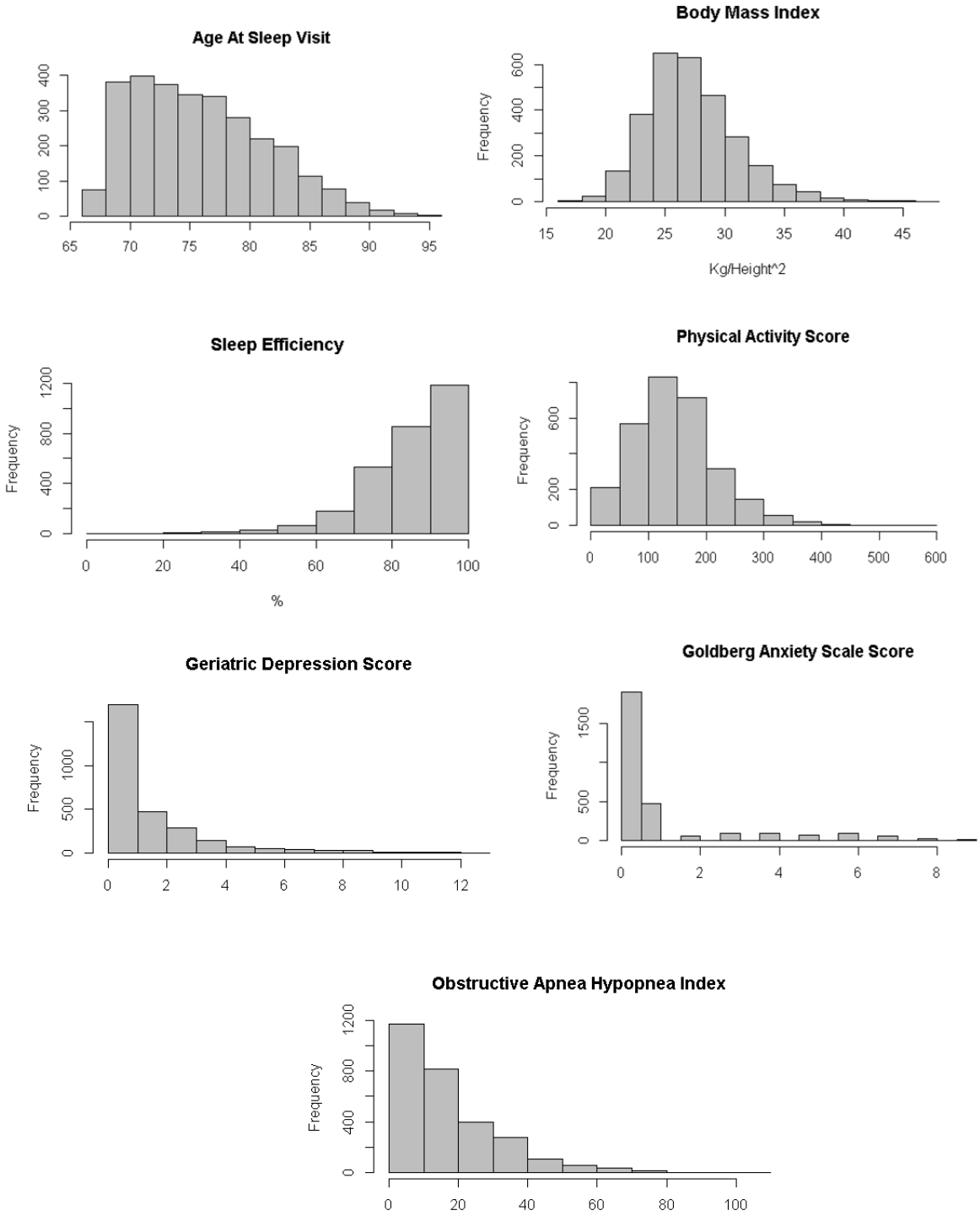

Figure S1 Histogram of Variable Distribution

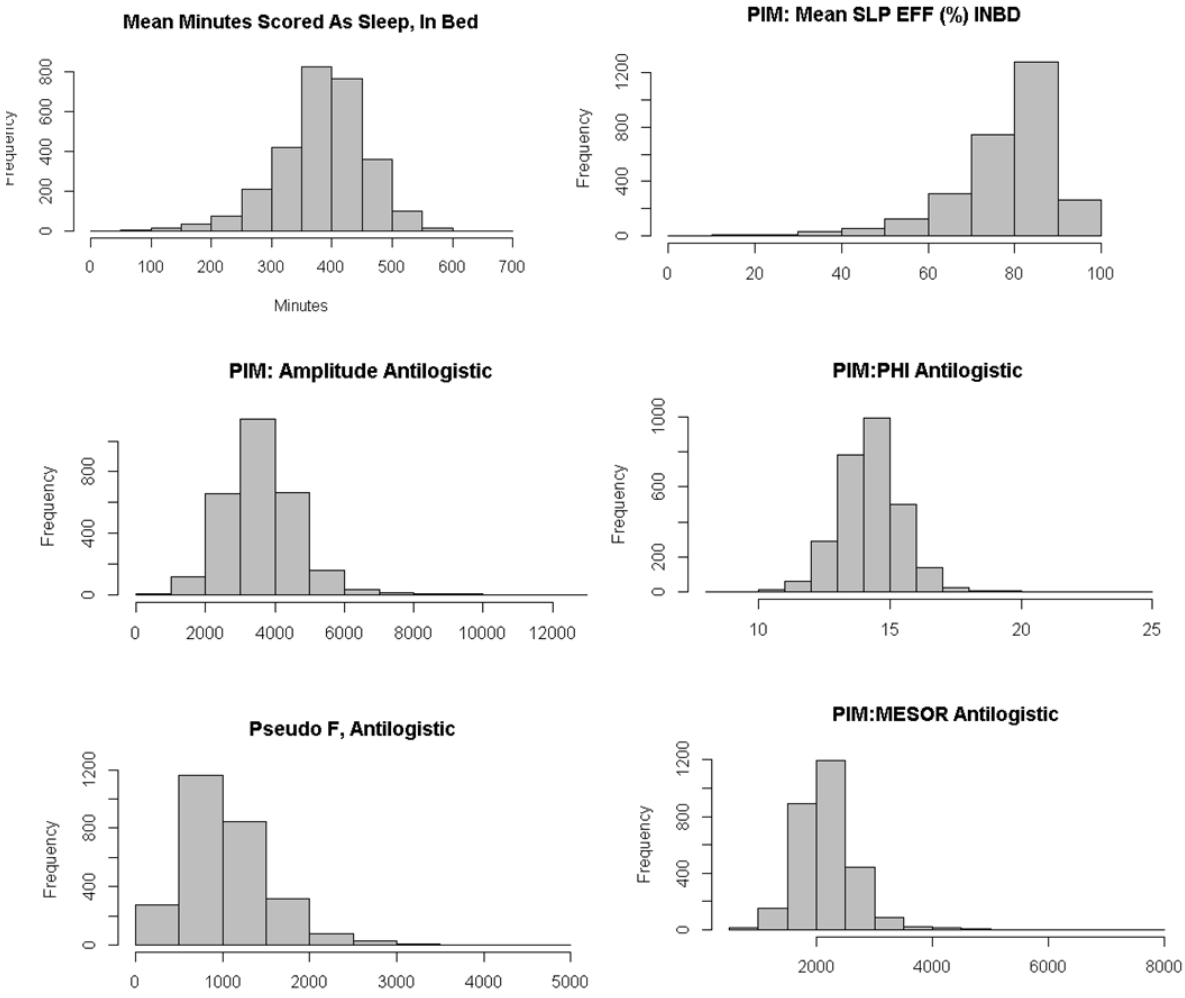

Figure S2 Distribution of Rest-activity Rhythm Variables

Table S1 Pearson Correlation Coefficients

	PQPACTSL	MHDIAB	MHRHEU1	MHOA	MHPARK	MHLIVER	MHRENAL	MHCOBPD	MHCAT	MHMI	DPGDS15	AXANXSC	TURSMOKE	ACSMINMP	ACSEFNMP	ACAMPPT	ACPHIPT
PQPACTSL	1.00	-0.06	-0.04	-0.02	-0.09	0.00	-0.02	-0.01	-0.02	-0.05	-0.11	-0.09	-0.01	0.02	0.06	0.04	0.03
MHDIAB	-0.06	1.00	0.04	-0.03	-0.01	-0.03	0.07	-0.03	0.06	0.07	0.04	0.03	0.02	-0.03	-0.08	-0.06	0.04
MHRHEU1	-0.04	0.04	1.00	0.04	0.04	0.01	0.04	0.05	0.03	0.04	0.12	0.09	0.04	-0.02	-0.04	-0.04	0.03
MHOA	-0.02	-0.03	0.04	1.00	0.02	0.03	0.02	0.03	0.08	0.01	0.15	0.15	0.04	-0.06	-0.08	-0.02	0.05
MHPARK	-0.09	-0.01	0.04	0.02	1.00	0.01	0.03	0.01	0.03	-0.01	0.05	0.06	0.00	0.02	0.00	-0.09	0.01
MHLIVER	0.00	-0.03	0.01	0.03	0.01	1.00	0.06	0.00	-0.01	0.01	0.00	0.03	0.01	0.04	0.03	0.01	-0.01
MHRENAL	-0.02	0.07	0.04	0.02	0.03	0.06	1.00	0.04	0.00	0.07	0.00	0.01	-0.04	0.02	0.01	-0.03	-0.01
MHCOBPD	-0.01	-0.03	0.05	0.03	0.01	0.00	0.04	1.00	0.02	0.02	0.11	0.06	0.12	-0.03	-0.05	-0.05	0.01
MHCAT	-0.02	0.06	0.03	0.08	0.03	-0.01	0.00	0.02	1.00	0.06	0.08	0.04	0.01	0.01	0.00	-0.08	0.01
MHMI	-0.05	0.07	0.04	0.01	-0.01	0.01	0.07	0.02	0.06	1.00	0.09	0.05	0.03	0.00	-0.04	-0.06	0.01
MHANGIN	-0.06	0.08	0.05	0.07	0.02	0.02	0.02	0.01	0.07	0.37	0.11	0.09	0.04	-0.04	-0.05	-0.07	0.01
MHCHF	-0.04	0.07	0.04	0.02	-0.01	0.03	0.05	0.06	0.03	0.21	0.09	0.08	0.04	-0.08	-0.11	-0.04	0.01
MHSTRK	-0.02	-0.01	0.03	0.04	0.04	0.01	0.00	0.02	0.00	0.05	0.11	0.03	0.02	0.03	0.02	-0.03	0.00
MHBP	-0.02	0.16	0.04	0.06	-0.02	-0.01	0.05	0.01	0.06	0.10	0.09	0.09	0.06	-0.03	-0.06	-0.08	-0.04
VSAGE1	-0.03	-0.03	0.04	-0.02	0.01	-0.04	0.03	0.00	0.24	0.09	0.12	-0.02	-0.08	0.04	-0.05	-0.20	0.01
TMMSCORE	0.10	-0.06	-0.10	0.02	-0.07	0.03	-0.02	-0.02	0.01	-0.05	-0.14	-0.10	0.00	-0.04	0.06	0.07	0.02
HWBMI	-0.06	0.15	0.05	0.03	0.00	0.00	0.04	0.01	-0.01	-0.01	0.03	0.05	0.08	-0.21	-0.24	-0.10	-0.02
PQPEFFCY	0.26	-0.05	-0.05	-0.06	-0.01	0.03	-0.02	-0.06	-0.05	-0.07	-0.18	-0.23	-0.03	-0.05	0.14	0.04	0.00
CVPVD	-0.04	0.05	0.03	0.02	0.02	0.00	0.06	0.06	0.04	0.17	0.14	0.09	0.07	0.01	-0.03	-0.04	0.00
PASCORE	0.06	-0.09	-0.03	-0.07	-0.05	-0.01	-0.03	-0.04	-0.11	-0.07	-0.23	-0.05	0.00	-0.02	0.06	0.25	-0.09
DPGDS15	-0.11	0.04	0.12	0.15	0.05	0.00	0.00	0.11	0.08	0.09	1.00	0.49	0.05	-0.02	-0.10	-0.15	0.06
AXANXSC	-0.09	0.03	0.09	0.15	0.06	0.03	0.01	0.06	0.04	0.05	0.49	1.00	0.05	-0.04	-0.07	-0.02	0.05
TURSMOKE	-0.01	0.02	0.04	0.04	0.00	0.01	-0.04	0.12	0.01	0.03	0.05	0.05	1.00	-0.06	-0.10	0.01	-0.01
ACSMINMP	0.02	-0.03	-0.02	-0.06	0.02	0.04	0.02	-0.03	0.01	0.00	-0.02	-0.04	-0.06	1.00	0.76	0.02	0.00
ACSEFNMP	0.06	-0.08	-0.04	-0.08	0.00	0.03	0.01	-0.05	0.00	-0.04	-0.10	-0.07	-0.10	0.76	1.00	0.09	-0.05
ACAMPPT	0.04	-0.06	-0.04	-0.02	-0.09	0.01	-0.03	-0.05	-0.08	-0.06	-0.15	-0.02	0.01	0.02	0.09	1.00	-0.03
ACPHIPT	0.03	0.04	0.03	0.05	0.01	-0.01	-0.01	0.01	0.01	0.01	0.06	0.05	-0.01	0.00	-0.05	-0.03	1.00
ACFVPT	0.05	-0.07	-0.02	-0.05	-0.04	0.02	-0.03	-0.05	-0.06	-0.07	-0.15	-0.05	0.02	0.20	0.17	0.55	-0.02
ACMESPT	0.01	-0.03	-0.02	0.00	-0.08	0.00	-0.02	-0.03	-0.05	-0.05	-0.11	0.00	0.03	-0.22	-0.20	0.82	-0.14
POORDI3	-0.01	0.06	0.02	0.04	-0.01	-0.03	-0.01	0.01	0.01	0.02	0.08	0.03	0.01	-0.20	-0.27	-0.09	0.01
M1ADEPR	-0.03	0.02	0.02	0.09	0.11	0.00	0.07	0.05	0.05	0.05	0.19	0.16	0.00	0.04	-0.02	-0.05	0.08
M1BENZO	-0.08	0.01	0.01	0.08	0.06	0.04	0.05	0.04	0.03	0.05	0.12	0.15	0.03	0.02	-0.03	-0.06	0.08
M1NSAIDS	0.00	-0.05	0.07	0.26	-0.02	0.02	-0.02	0.06	0.01	-0.04	0.05	0.09	0.03	-0.09	-0.06	0.01	0.02
M1NBANX	-0.01	0.01	0.02	0.02	0.01	0.01	0.01	-0.02	0.02	-0.01	0.04	0.06	-0.02	-0.02	-0.01	-0.05	0.02
M1CORTO	-0.03	-0.01	0.12	0.04	-0.01	0.02	0.01	0.05	0.02	0.02	0.06	0.05	0.02	-0.07	-0.07	-0.01	0.02
M1CORTI	0.01	-0.02	0.00	0.05	-0.02	-0.02	-0.02	0.20	0.03	-0.03	0.02	0.00	0.03	-0.01	-0.04	-0.04	0.04
edu	0.10	-0.06	-0.05	-0.01	0.04	0.02	-0.02	-0.10	0.00	-0.05	-0.09	-0.06	-0.10	0.08	0.12	0.01	0.04
race	-0.06	0.05	0.01	-0.04	0.01	-0.01	0.02	-0.01	-0.05	-0.06	-0.01	0.01	0.00	0.00	0.01	0.03	0.03
etoh	0.05	-0.06	-0.05	0.01	-0.01	0.02	-0.06	-0.02	-0.01	-0.08	-0.07	-0.05	0.12	0.03	0.01	0.07	-0.01
delta	0.03	-0.09	0.00	-0.02	0.03	0.01	0.00	0.02	-0.02	-0.02	-0.03	-0.02	-0.01	-0.05	-0.02	0.04	-0.01
POTMS34P	0.03	-0.02	0.01	0.00	0.02	-0.01	0.05	0.01	0.01	-0.02	0.00	-0.03	0.02	0.01	0.04	0.02	-0.03
Log_POTMS34P	0.06	-0.03	0.00	-0.01	0.01	0.00	0.03	0.01	0.00	-0.04	-0.02	-0.03	0.00	0.02	0.06	0.03	-0.02

Table S1 Continued (Pearson Correlation Coefficients)

	ACFVPT	ACMESPT	POORDI3	M1ADEPR	M1BENZO	M1NSAIDS	M1NBANX	M1CORTO	M1CORTI	edu	race	etoh	delta	POTMS34P	Log_POTMS34P
PQPACTSL	0.05	0.01	-0.01	-0.03	-0.08	0.00	-0.01	-0.03	0.01	0.10	-0.06	0.05	0.03	0.03	0.06
MHDIAB	-0.07	-0.03	0.06	0.02	0.01	-0.05	0.01	-0.01	-0.02	-0.06	0.05	-0.06	-0.09	-0.02	-0.03
MHRHEU1	-0.02	-0.02	0.02	0.02	0.01	0.07	0.02	0.12	0.00	-0.05	0.01	-0.05	0.00	0.01	0.00
MHOA	-0.05	0.00	0.04	0.09	0.08	0.26	0.02	0.04	0.05	-0.01	-0.04	0.01	-0.02	0.00	-0.01
MHPARK	-0.04	-0.08	-0.01	0.11	0.06	-0.02	0.01	-0.01	-0.02	0.04	0.01	-0.01	0.03	0.02	0.01
MHLIVER	0.02	0.00	-0.03	0.00	0.04	0.02	0.01	0.02	-0.02	0.02	-0.01	0.02	0.01	-0.01	0.00
MHRENAL	-0.03	-0.02	-0.01	0.07	0.05	-0.02	0.01	0.01	-0.02	-0.02	0.02	-0.06	0.00	0.05	0.03
MHCOBPD	-0.05	-0.03	0.01	0.05	0.04	0.06	-0.02	0.05	0.20	-0.10	-0.01	-0.02	0.02	0.01	0.01
MHCAT	-0.06	-0.05	0.01	0.05	0.03	0.01	0.02	0.02	0.03	0.00	-0.05	-0.01	-0.02	0.01	0.00
MHMI	-0.07	-0.05	0.02	0.05	0.05	-0.04	-0.01	0.02	-0.03	-0.05	-0.06	-0.08	-0.02	-0.02	-0.04
MHANGIN	-0.07	-0.05	0.02	0.09	0.06	-0.01	-0.01	0.03	0.02	-0.03	-0.02	-0.06	-0.03	0.01	0.00
MHCHF	-0.08	0.00	0.08	0.04	0.03	-0.04	-0.03	0.01	0.02	-0.03	-0.03	-0.03	-0.03	-0.01	-0.03
MHSTRK	-0.05	-0.04	-0.02	0.01	0.03	-0.02	0.04	0.02	0.00	-0.02	-0.01	-0.03	0.00	0.02	0.02
MHBP	-0.11	-0.03	0.10	0.05	0.08	0.00	0.03	0.00	0.00	-0.06	0.04	0.00	-0.05	0.00	-0.01
VSAGE1	-0.15	-0.13	0.06	-0.02	-0.01	-0.06	0.02	0.00	-0.02	-0.06	-0.07	-0.04	-0.04	-0.02	-0.07
TMMSCORE	0.07	0.02	-0.04	-0.01	-0.02	0.03	0.00	0.04	0.04	0.21	-0.16	0.10	0.06	0.05	0.08
HWBMI	-0.16	-0.04	0.29	-0.01	0.01	0.10	-0.03	-0.02	0.01	-0.12	-0.07	-0.05	-0.08	-0.06	-0.06
PQPEFFCY	0.03	-0.01	-0.01	-0.04	-0.10	-0.01	-0.07	-0.04	-0.03	0.11	-0.06	0.03	0.00	0.02	0.03
CVPVD	-0.06	-0.04	0.09	0.08	0.03	-0.01	0.01	-0.02	-0.01	-0.07	-0.02	-0.03	-0.01	0.02	0.00
PASCORE	0.24	0.20	-0.04	-0.06	-0.09	0.01	-0.04	-0.05	-0.03	0.04	0.01	0.03	-0.01	-0.02	0.01
DPGDS15	-0.15	-0.11	0.08	0.19	0.12	0.05	0.04	0.06	0.02	-0.09	-0.01	-0.07	-0.03	0.00	-0.02
AXANXSC	-0.05	0.00	0.03	0.16	0.15	0.09	0.06	0.05	0.00	-0.06	0.01	-0.05	-0.02	-0.03	-0.03
TURSMOKE	0.02	0.03	0.01	0.00	0.03	0.03	-0.02	0.02	0.03	-0.10	0.00	0.12	-0.01	0.02	0.00
ACSMINMP	0.20	-0.22	-0.20	0.04	0.02	-0.09	-0.02	-0.07	-0.01	0.08	0.00	0.03	-0.05	0.01	0.02
ACSEFNMP	0.17	-0.20	-0.27	-0.02	-0.03	-0.06	-0.01	-0.07	-0.04	0.12	0.01	0.01	-0.02	0.04	0.06
ACAMPPT	0.55	0.82	-0.09	-0.05	-0.06	0.01	-0.05	-0.01	-0.04	0.01	0.03	0.07	0.04	0.02	0.03
ACPHIPT	-0.02	-0.14	0.01	0.08	0.08	0.02	0.02	0.02	0.04	0.04	0.03	-0.01	-0.01	-0.03	-0.02
ACFVPT	1.00	0.44	-0.11	-0.04	-0.04	0.01	-0.03	-0.03	-0.02	0.02	-0.02	0.12	0.02	0.00	0.02
ACMESPT	0.44	1.00	0.00	-0.05	-0.03	0.00	-0.05	0.02	-0.03	-0.03	0.04	0.04	0.02	-0.01	-0.01
POORDI3	-0.11	0.00	1.00	0.02	-0.04	0.04	-0.03	0.01	-0.02	-0.10	-0.01	0.00	-0.11	-0.15	-0.16
M1ADEPR	-0.04	-0.05	0.02	1.00	0.17	0.06	0.01	0.05	0.02	-0.01	-0.04	-0.05	-0.03	-0.01	0.00
M1BENZO	-0.04	-0.03	-0.04	0.17	1.00	0.06	0.09	0.04	0.03	-0.03	-0.02	-0.01	-0.07	-0.04	-0.07
M1NSAIDS	0.01	0.00	0.04	0.06	0.06	1.00	0.02	0.04	0.05	-0.05	-0.05	0.02	-0.03	-0.04	-0.04
M1NBANX	-0.03	-0.05	-0.03	0.01	0.09	0.02	1.00	0.06	0.02	0.03	0.05	0.00	-0.02	-0.02	-0.02
M1CORTO	-0.03	0.02	0.01	0.05	0.04	0.04	0.06	1.00	0.04	-0.01	-0.02	-0.03	0.05	0.04	0.04
M1CORTI	-0.02	-0.03	-0.02	0.02	0.03	0.05	0.02	0.04	1.00	-0.01	0.04	0.00	0.03	0.01	0.00
edu	0.02	-0.03	-0.10	-0.01	-0.03	-0.05	0.03	-0.01	-0.01	1.00	0.02	0.14	0.03	0.03	0.03
race	-0.02	0.04	-0.01	-0.04	-0.02	-0.05	0.05	-0.02	0.04	0.02	1.00	-0.03	-0.03	-0.06	-0.05
etoh	0.12	0.04	0.00	-0.05	-0.01	0.02	0.00	-0.03	0.00	0.14	-0.03	1.00	-0.03	0.00	0.01
delta	0.02	0.02	-0.11	-0.03	-0.07	-0.03	-0.02	0.05	0.03	0.03	-0.03	-0.03	1.00	0.64	0.63
POTMS34P	0.00	-0.01	-0.15	-0.01	-0.04	-0.04	-0.02	0.04	0.01	0.03	-0.06	0.00	0.64	1.00	0.89
Log_POTMS34	0.02	-0.01	-0.16	0.00	-0.07	-0.04	-0.02	0.04	0.00	0.03	-0.05	0.01	0.63	0.89	1.00

Bibliography

1. Tasali E, Leproult R, Ehrmann DA, et al. Slow-wave sleep and the risk of type 2 diabetes in humans. *Proc Natl Acad Sci U S A* 2008;105(3):1044-9. doi: 10.1073/pnas.0706446105 [published Online First: 2008/01/04]

2. Blagestad T, Pallesen S, Lunde LH, et al. Sleep in older chronic pain patients: a comparative polysomnographic study. *Clin J Pain* 2012;28(4):277-83. doi: 10.1097/AJP.0b013e3182313899 [published Online First: 2012/02/15]

3. Lallukka T, Sivertsen B, Kronholm E, et al. Association of sleep duration and sleep quality with the physical, social, and emotional functioning among Australian adults. *Sleep Health* 2018;4(2):194-200. doi: 10.1016/j.sleh.2017.11.006 [published Online First: 2018/03/21]

4. Liu K, Ma Q, Wang M. Comparison of Quantitative Electroencephalogram During Sleep in Depressed and Non-Depressed Patients with Parkinson's Disease. *Med Sci Monit* 2019;25:1046-52. doi: 10.12659/MSM.913931 [published Online First: 2019/02/08]

5. Learning l. Brainwave activity changes dramatically across the different stages of sleep. *https://courseslumenlearningcom/wmopen-psychology/chapter/outcome-sleep-and-dreams/*

6. Roth T. Slow wave sleep: does it matter? *J Clin Sleep Med* 2009;5(2 Suppl):S4-S5.

7. Lao XQ, Liu X, Deng HB, et al. Sleep Quality, Sleep Duration, and the Risk of Coronary Heart Disease: A Prospective Cohort Study With 60,586 Adults. *J Clin Sleep Med* 2018;14(1):109-17. doi: 10.5664/jcsm.6894 [published Online First: 2017/12/05]

8. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health: Journal of the National Sleep Foundation* 2015;1(1):40-43. doi: 10.1016/j.sleh.2014.12.010

9. Buysse DJ, Reynolds CF, 3rd, Monk TH, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28(2):193-213. [published Online First: 1989/05/01]

10. Grandner MA, Kripke DF, Yoon I-Y, et al. Criterion validity of the Pittsburgh Sleep Quality Index: Investigation in a non-clinical sample. *Sleep and biological rhythms* 2006;4(2):129-39. doi: 10.1111/j.1479-8425.2006.00207.x [published Online First: 2006/06/09]

11. Blackwell T, Redline S, Ancoli-Israel S, et al. Comparison of sleep parameters from actigraphy and polysomnography in older women: the SOF study. *Sleep* 2008;31(2):283-91. doi: 10.1093/sleep/31.2.283

12. Walker MP. The role of slow wave sleep in memory processing. *J Clin Sleep Med* 2009;5(2 Suppl):S20-6. [published Online First: 2009/12/17]

13. Davis CJ, Clinton JM, Jewett KA, et al. Delta wave power: an independent sleep phenotype or epiphenomenon? *J Clin Sleep Med* 2011;7(5 Suppl):S16-S18. doi: 10.5664/JCSM.1346

14. Borbely AA, Daan S, Wirz-Justice A, et al. The two-process model of sleep regulation: a reappraisal. *J Sleep Res* 2016;25(2):131-43. doi: 10.1111/jsr.12371 [published Online First: 2016/01/15]

15. Amorim L, Magalhaes R, Coelho A, et al. Poor Sleep Quality Associates With Decreased Functional and Structural Brain Connectivity in Normative Aging: A MRI Multimodal Approach. *Front Aging Neurosci* 2018;10:375. doi: 10.3389/fnagi.2018.00375 [published Online First: 2018/12/14]

16. Bu L, Wang D, Huo C, et al. Effects of poor sleep quality on brain functional connectivity revealed by wavelet-based coherence analysis using NIRS methods in elderly subjects. *Neurosci Lett* 2018;668:108-14. doi: 10.1016/j.neulet.2018.01.026 [published Online First: 2018/01/22]

17. Elder SJ, Pisoni RL, Akizawa T, et al. Sleep quality predicts quality of life and mortality risk in haemodialysis patients: results from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Nephrol Dial Transplant* 2008;23(3):998-1004. doi: 10.1093/ndt/gfm630 [published Online First: 2007/10/04]

18. Turoff A, Thiem U, Fox H, et al. Sleep duration and quality in heart failure patients. *Sleep Breath* 2017;21(4):919-27. doi: 10.1007/s11325-017-1501-x [published Online First: 2017/04/09]

19. Ma CL, Chang WP, Lin CC. Rest/activity rhythm is related to the coexistence of pain and sleep disturbance among advanced cancer patients with pain. *Support Care Cancer* 2014;22(1):87-94. doi: 10.1007/s00520-013-1918-0 [published Online First: 2013/09/03]

20. St-Onge MP, Mikic A, Pietrolungo CE. Effects of Diet on Sleep Quality. *Adv Nutr* 2016;7(5):938-49. doi: 10.3945/an.116.012336 [published Online First: 2016/09/17]

21. Umigai N, Takeda R, Mori A. Effect of crocetin on quality of sleep: A randomized, double-blind, placebo-controlled, crossover study. *Complement Ther Med* 2018;41:47-51. doi: 10.1016/j.ctim.2018.09.003 [published Online First: 2018/11/28]

22. Schreiner SJ, Imbach LL, Werth E, et al. Slow-wave sleep and motor progression in Parkinson disease. *Ann Neurol* 2019;85(5):765-70. doi: 10.1002/ana.25459 [published Online First: 2019/03/20]

23. Minakawa EN, Wada K, Nagai Y. Sleep Disturbance as a Potential Modifiable Risk Factor for Alzheimer's Disease. *Int J Mol Sci* 2019;20(4) doi: 10.3390/ijms20040803 [published Online First: 2019/02/20]

24. Daniel LC, Li Y, Kloss JD, et al. The impact of dexamethasone and prednisone on sleep in children with acute lymphoblastic leukemia. *Supportive care in cancer* 2016;24(9):3897-906. doi: 10.1007/s00520-016-3234-y [published Online First: 2016/04/23]

25. Smagula SF, Krafty RT, Taylor BJ, et al. Rest-activity rhythm and sleep characteristics associated with depression symptom severity in strained dementia caregivers. *Journal of sleep research* 2017;26(6):718-25. doi: 10.1111/jsr.12549 [published Online First: 2017/05/10]

26. Ancoli-Israel S, Cole R, Alessi C, et al. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 2003;26(3):342-92. doi: 10.1093/sleep/26.3.342 [published Online First: 2003/05/17]

27. Smagula SF, Ancoli-Israel S, Blackwell T, et al. Circadian rest-activity rhythms predict future increases in depressive symptoms among community-dwelling older men. *Am J Geriatr Psychiatry* 2015;23(5):495-505. doi: 10.1016/j.jagp.2014.06.007 [published Online First: 2014/06/26]

28. Noor ZM, Smith AJ, Smith SS, et al. A feasibility study: Use of actigraph to monitor and follow-up sleep/wake patterns in individuals attending community pharmacy with sleeping disorders. *J Pharm Bioallied Sci* 2016;8(3):173-80. doi: 10.4103/0975-7406.171739

29. Mariani S, Tarokh L, Djonlagic I, et al. Evaluation of an automated pipeline for large-scale EEG spectral analysis: the National Sleep Research Resource. *Sleep Med* 2018;47:126-36. doi: 10.1016/j.sleep.2017.11.1128 [published Online First: 2018/05/29]

30. Marler MR, Gehrman P, Martin JL, et al. The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes. *Stat Med* 2006;25(22):3893-904. doi: 10.1002/sim.2466 [published Online First: 2005/12/29]

31. Krafty RT, Fu H, Graves JL, et al. Measuring Variability in Rest-Activity Rhythms from Actigraphy with Application to Characterizing Symptoms of Depression. *Statistics in Biosciences* 2019 doi: 10.1007/s12561-018-09230-2

32. Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *2011* 2011;45(7):47. doi: 10.18637/jss.v045.i07 [published Online First: 2011-12-12]

33. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B (Methodological)* 1996;58(1):267-88.

34. Choirat C HJ, Imai K, King G, Lau O Zelig: Everyone's Statistical Software, Version 5.16.1. *http://zeligprojectorg/* 2018

35. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *2010* 2010;33(1):22. doi: 10.18637/jss.v033.i01 [published Online First: 2010-02-02]

36. Neikrug AB, Donaldson G, Iacob E, et al. Activity rhythms and clinical correlates in fibromyalgia. *Pain* 2017;158(8):1417-29. doi: 10.1097/j.pain.0000000000000906 [published Online First: 2017/03/23]

37. von Hippel PT. Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data. *Sociol Methodol* 2007;37:83-117. doi: DOI 10.1111/j.1467-9531.2007.00180.x

38. Kuk D, Varadhan R. Model selection in competing risks regression. *Stat Med* 2013;32(18):3077-88. doi: 10.1002/sim.5762 [published Online First: 2013/02/26]

39. Reid K, Dawson D. Correlation between wrist activity monitor and electrophysiological measures of sleep in a simulated shiftwork environment for younger and older subjects. *Sleep* 1999;22(3):378-85. doi: 10.1093/sleep/22.3.378 [published Online First: 1999/05/26]

40. Franken P, Chollet D, Tafti M. The homeostatic regulation of sleep need is under genetic control. *J Neurosci* 2001;21(8):2610-21. [published Online First: 2001/04/18]

41. Landolt HP, Borbely AA. Age-dependent changes in sleep EEG topography. *Clin Neurophysiol* 2001;112(2):369-77. [published Online First: 2001/02/13]

42. Van Cauter E, Leproult R, Plat L. Age-related changes in slow wave sleep and REM sleep and relationship with growth hormone and cortisol levels in healthy men. *JAMA* 2000;284(7):861-8. doi: 10.1001/jama.284.7.861 [published Online First: 2000/08/11]

43. Chervin RD, Burns JW, Ruzicka DL. Electroencephalographic Changes during Respiratory Cycles Predict Sleepiness in Sleep Apnea. *American Journal of Respiratory and Critical Care Medicine* 2005;171(6):652-58. doi: 10.1164/rccm.200408-1056OC

44. Kline CE. The bidirectional relationship between exercise and sleep: Implications for exercise adherence and sleep improvement. *Am J Lifestyle Med* 2014;8(6):375-79. doi: 10.1177/1559827614544437

45. Pacheco SR, Miranda AM, Coelho R, et al. Overweight in youth and sleep quality: is there a link? *Arch Endocrinol Metab* 2017;61(4):367-73. doi: 10.1590/2359-3997000000265 [published Online First: 2017/06/29]

46. Mezick EJ, Matthews KA, Hall M, et al. Influence of race and socioeconomic status on sleep: Pittsburgh SleepSCORE project. *Psychosom Med* 2008;70(4):410-16. doi: 10.1097/PSY.0b013e31816fdf21

47. Thakkar MM, Sharma R, Sahota P. Alcohol disrupts sleep homeostasis. *Alcohol* 2015;49(4):299-310. doi: 10.1016/j.alcohol.2014.07.019 [published Online First: 2014/12/17]

48. Bourgeois J, Elseviers MM, Van Bortel L, et al. One-year evolution of sleep quality in older users of benzodiazepines: a longitudinal cohort study in belgian nursing home residents. *Drugs Aging* 2014;31(9):677-82. doi: 10.1007/s40266-014-0203-3 [published Online First: 2014/08/26]

49. Moser NJ, Phillips BA, Guthrie G, et al. Effects of Dexamethasone on Sleep. *Pharmacology & Toxicology* 1996;79(2):100-02. doi: 10.1111/j.1600-0773.1996.tb00249.x

50. Grabovac I, Haider S, Berner C, et al. Sleep Quality in Patients with Rheumatoid Arthritis and Associations with Pain, Disability, Disease Duration, and Activity. *J Clin Med* 2018;7(10) doi: 10.3390/jcm7100336 [published Online First: 2018/10/12]

51. Junho BT, Kummer A, Cardoso FE, et al. Sleep quality is associated with the severity of clinical symptoms in Parkinson's disease. *Acta Neurol Belg* 2018;118(1):85-91. doi: 10.1007/s13760-017-0868-6 [published Online First: 2017/12/07]

52. Nigam G, Camacho M, Chang ET, et al. Exploring sleep disorders in patients with chronic kidney disease. *Nat Sci Sleep* 2018;10:35-43. doi: 10.2147/NSS.S125839

53. Kent BA, Michalik M, Marchant EG, et al. Delayed daily activity and reduced NREM slow-wave power in the APPswe/PS1dE9 mouse model of Alzheimer's disease. *Neurobiol Aging* 2019;78:74-86. doi: 10.1016/j.neurobiolaging.2019.01.010 [published Online First: 2019/03/19]

54. Pierratos A, Hanly PJ. Sleep disorders over the full range of chronic kidney disease. *Blood Purif* 2011;31(1-3):146-50. doi: 10.1159/000321859 [published Online First: 2011/01/14]

55. Hedlund ER, Villard L, Lundell B, et al. Physical exercise may improve sleep quality in children and adolescents with Fontan circulation. *Cardiol Young* 2019:1-8. doi: 10.1017/S1047951119001136 [published Online First: 2019/06/21]

56. Paparrigopoulos T, Tzavara C, Theleritis C, et al. Physical activity may promote sleep in cardiac patients suffering from insomnia. *Int J Cardiol* 2010;143(2):209-11. doi: 10.1016/j.ijcard.2008.11.178 [published Online First: 2008/12/26]

57. Lang C, Brand S, Feldmeth AK, et al. Increased self-reported and objectively assessed physical activity predict sleep quality among adolescents. *Physiol Behav* 2013;120:46-53. doi: 10.1016/j.physbeh.2013.07.001 [published Online First: 2013/07/16]

58. Dolezal BA, Neufeld EV, Boland DM, et al. Interrelationship between Sleep and Exercise: A Systematic Review. *Adv Prev Med* 2017;2017:1364387-87. doi: 10.1155/2017/1364387 [published Online First: 2017/03/26]

59. Murray K, Godbole S, Natarajan L, et al. The relations between sleep, time of physical activity, and time outdoors among adult women. *PloS one* 2017;12(9):e0182013-e13. doi: 10.1371/journal.pone.0182013

60. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2005;67(2):301-20. doi: 10.1111/j.1467-9868.2005.00503.x