HHS Public Access

Author manuscript

*Biol Psychol*. Author manuscript; available in PMC 2018 February 01.

Published in final edited form as:

*Biol Psychol*. 2017 February ; 123: 37–46. doi:10.1016/j.biopsycho.2016.11.010.

# When a gold standard isn't so golden: Lack of prediction of subjective sleep quality from sleep polysomnography

**Katherine A. Kaplan, Ph.D.**[1], **Jason Hirshman**[2], **Beatriz Hernandez, B.A.**[1,3], **Marcia L. Stefanick, Ph.D.**[4], **Andrew R. Hoffman, M.D.**[4], **Susan Redline, M.D. M.P.H.**[5], **Sonia Ancoli-Israel, Ph.D.**[6], **Katie Stone, Ph.D.**[7], **Leah Friedman, Ph.D.**[1,3], **Jamie M. Zeitzer, Ph.D.**[1,3,*], and **Osteoporotic Fractures in Men (MrOS), Study of Osteoporotic Fractures (SOF) Research Groups**

[1]Department of Psychiatry and Behavioral Sciences, Stanford Center for Sleep Sciences and Medicine, Stanford University, Stanford CA 94305, USA.

[2]Department of Mathematics, Stanford University, Stanford CA 94305, USA.

[3]Mental Illness Research Education and Clinical Center, VA Palo Alto Health Care System, Palo Alto CA 94304, USA.

[4]Department of Medicine, Stanford University, Stanford CA 94305, USA.

[5]Departments of Medicine, Brigham and Women's Hospital and Beth Israel Deaconess Medical Center, Harvard Medical School, Boston MA 02115, USA.

[6]Departments of Psychiatry and Medicine, University of California, San Diego, San Diego CA 92093, USA.

[7]California Pacific Medical Center Research Institute, San Francisco CA 94107, USA.

## Abstract

**Background—**Reports of subjective sleep quality are frequently collected in research and clinical practice. It is unclear, however, how well polysomnographic measures of sleep correlate with subjective reports of prior-night sleep quality in elderly men and women. Furthermore, the relative importance of various polysomnographic, demographic and clinical characteristics in

[*]to whom correspondence should be addressed: Jamie M. Zeitzer, Ph.D., 3801 Miranda Avenue (151Y), Palo Alto CA 94304 USA, 650-493-5000 x1-1-62410 (ph), 650-852-3297 (fax), jzeitzer@stanford.edu.



**Authors Contributions:** Each author participated sufficiently in the work and analysis of data, as well as the writing of the manuscript.

**Disclosure:** There was no off-label or investigational use of any medication.

**Authorship Responsibility:** Each author participated sufficiently in the work and analysis of data, as well as the writing of the manuscript.

**Conflict of Interest:** The authors report no conflict of interest associated with this manuscript.

**Consent for Publication:** Not applicable.

**Competing Interests.** The authors report no conflict of interest associated with this manuscript.

predicting subjective sleep quality is not known. We sought to determine the correlates of subjective sleep quality in in older adults using more recently developed machine learning algorithms that are suitable for selecting and ranking important variables.

**Methods—**Community-dwelling older men (n=1024) and women (n=459), a subset of those participating in the Osteoporotic Fractures in Men study and the Study of Osteoporotic Fractures study, respectively, completed a single night of at-home polysomnographic recording of sleep followed by a set of morning questions concerning the prior night's sleep quality. Questionnaires concerning demographics and psychological characteristics were also collected prior to the overnight recording and entered into multivariable models. Two machine learning algorithms, lasso penalized regression and random forests, determined variable selection and the ordering of variable importance separately for men and women.

**Results—**Thirty-eight sleep, demographic and clinical correlates of sleep quality were considered. Together, these multivariable models explained only 11-17% of the variance in predicting subjective sleep quality. Objective sleep efficiency emerged as the strongest correlate of subjective sleep quality across all models, and across both sexes. Greater total sleep time and sleep stage transitions were also significant objective correlates of subjective sleep quality. The amount of slow wave sleep obtained was not determined to be important.

**Conclusions—**Overall, the commonly obtained measures of polysomnographically-defined sleep contributed little to subjective ratings of prior-night sleep quality. Though they explained relatively little of the variance, sleep efficiency, total sleep time and sleep stage transitions were among the most important objective correlates.

### Keywords

Sleep quality; machine learning; polysomnography; aging; sex differences

## Introduction

Although the issue of subjective “sleep quality” is frequently encountered in research and treatment, the correlates that may contribute to these subjective estimates remain poorly understood. Clinically, subjective sleep quality is fundamental to complaints of insomnia and non-restorative sleep, two conditions associated with considerable morbidity and impairment, and relevant to sleep health in general. Dissatisfaction with “sleep quantity or quality” is a central feature of Insomnia Disorder in the DSM-5, and it is well known that a number of individuals with insomnia will report poor sleep quality even when objectively-measured sleep appears relatively normal [1, 2]. Subjective sleep quality is also central to complaints of non-restorative sleep, which was included as a subtype of insomnia until publication in the DSM-5. Research by Roth and colleagues has shown that complaints of non-restorative sleep are stable and separable from those of difficulty initiating or maintaining sleep, though still associated with comparable levels of functional impairment [3]. In the general population, non-restorative sleep complaints are quite common [4] and associated with a variety of physical and mental conditions [5].

Given the clinical significance of subjective sleep quality, understanding the biological and psychological parameters that influence perceived sleep quality is important to inform

interventions. In both young and older individuals, a variety of objective correlates of better sleep quality have been posited, including increased duration of polysomnographically-determined stage N2 [6-8] and N3 sleep [8-10], decreased N1 sleep [6, 10, 11], decreased wakefulness at night [7, 9, 11], increased sleep efficiency [12], and fewer transitions from sleep to wake [13]. It is important to note, however, that many of these studies are small and limited to good-sleeping adults or to clinical populations such as those with insomnia. This study improves upon prior work by examining sleep quality in a very large sample of community-dwelling older adults who have both normal and disturbed sleep.

Previous research attempting to understand sleep quality is also limited by investigating only a limited number of correlates in predictive models. As recognized by Buysse and colleagues some 25 years earlier, sleep quality is likely to be a multifaceted construct that is difficult to characterize by any single correlate [14]. Though there have been calls to examine sleep quality using multivariable analyses with a wider range of predictors [15], no studies have yet been completed. The present research takes an important step towards clarifying subjective sleep quality by considering a large number of sleep, demographic and clinical correlates together using recently developed analytical techniques for multivariable analysis.

Mounting evidence suggests that estimates of subjective sleep quality must be interpreted in light of age and sex differences. There are age-dependent changes in sleep that predispose older individuals to poor objective sleep quality, including increased fragmentation and wakefulness at night [16]. Moreover, multiple studies suggest that the prevalence of sleep disorders—including insomnia, periodic limb movements, and sleep disordered breathing—increase with age [17-20], though there may be some leveling off or improvement by the seventh decade [21, 22]. Interestingly, although objective sleep deteriorates and rates of sleep disturbances increase, some studies indicate that older individuals rate their sleep quality similarly to or better than younger cohorts [14, 23]. It is therefore important to understand the objective correlates of perceived sleep quality in older individuals. This question must also be considered in light of sex differences in sleep quality reports. Women report higher rates of insomnia [24] and non-restorative sleep [4] compared to men. Women are also more likely to report worse subjective sleep parameters relative to objective sleep parameters [25, 26]. Finally, sleep architecture itself appears to differ between men and women, with men obtaining more light sleep and less slow wave sleep, a discrepancy that widens with increasing age [27].

Subjective sleep quality may be assessed through a variety of means, often via retrospective self-report inventories such as the Pittsburgh Sleep Quality Index [28] or via ordinal or visual analog scales included on prospective sleep diaries [29]. The present research focused on ordinal sleep diary ratings of two dimensions of subjective sleep quality from the prior night: Sleep Depth (light to deep) and Sleep Restfulness (restless to restful). Research suggests both dimensions are important. Åkerstedt and colleagues reported that sleep restfulness was strongly associated with subjective sleep quality [30] and, in a subsequent factor analysis, that sleep restfulness and sleep quality loaded strongly onto the same factor. [31] In validating a morning questionnaire assessing the prior night's sleep, Ellis and

colleagues determined that sleep depth was also relevant.[32] Hence, Sleep Depth and Sleep Restfulness were used as two non-exclusive measures of sleep quality.

The overall aim of this investigation was to determine the correlates of subjective sleep quality in older men and women, as well as to evaluate the relative importance of each correlate in predicting subjective sleep quality. Given the large number of potential variables that may explain subjective sleep quality, we used modern machine learning methods to determine which variables were important in predicting subjective sleep quality, as well as the ordering of this importance, using statistical techniques designed to reduce bias. We focused our analyses on a sample of community-dwelling older adults, and we ran all analyses separately and in parallel between men and women as we expected sex differences to emerge.

## Methods

### Study Population

Male participants in these analyses were enrolled in the Osteoporotic Fractures in Men Study (MrOS). At baseline, a total of 5994 men aged 65 or older were recruited between 2000 and 2002 from six areas of the United States: Birmingham, Alabama; Minneapolis, Minnesota; Palo Alto, California; Pittsburgh, Pennsylvania; Portland, Oregon; and San Diego, California. Assessment of sleep occurred at two separate time points. These analyses used data collected at the Sleep Visit 2, which took place from 2009 through 2012. A total of 1055 men participated across all six sites, of whom 1026 (97%) had polysomnographic data available and were included in final analyses.

Female participants in these analyses were enrolled in the Study of Osteoporotic Fractures (SOF). At baseline, a total of 9704 Caucasian women aged 65 or older were recruited between 1986 and 1988 from four areas of the United States: Baltimore, Maryland; Minneapolis, Minnesota; Pittsburgh, Pennsylvania; and Portland, Oregon. A cohort of 662 elderly African American women with similar baseline characteristics was recruited between 1997 and 1998. Assessment of sleep occurred at the eighth visit in the Caucasian cohort, which took place from 2002 through 2004. A total of 494 women (both Caucasian and African American) underwent polysomnographic monitoring at the Minneapolis and Pittsburgh sites, of whom 461 (93%) had usable polysomnographic data and were included in final analyses.

Both the MrOS and SOF protocols were approved by institutional review boards at all participating sites. All participants provided written informed consent. Study protocols adhered to the principles described in the Declaration of Helsinki. Complete descriptions of the MrOS [33, 34] and SOF [35] cohorts have already been published.

### Polysomnographic Monitoring

Sleep studies were completed in-home and unattended using polysomnography (PSG; Siesta unit in SOF, Safiro unit in MrOS; Compumedics, Inc., Melbourne, Australia). Recording montages between the two studies were identical and included C3/A2 and C4/A1 electroencephalograms (EEG), bilateral electrooculograms (EOG), and a chin

electromyogram (EMG) to gauge sleep; thoracic and abdominal respiratory inductance plethysmography to determine respiratory effort; nasal-oral thermocouple and nasal pressure cannula to detect airflow; finger pulse oximetry; electrocardiogram; and bilateral anterior tibialis piezoelectric sensors to determine leg movements. Variables of interest included those that would be traditionally derived from an overnight sleep study: polysomnography-determined total sleep time, wake after sleep onset, sleep efficiency (total sleep time divided by total time in bed), percentage of time spent in each sleep stage (with NREM Stage 3 and Stage 4 collapsed), latency to REM sleep, the number of sleep to wake shifts per hour, the number of deep (Stages 3 and 4) to lighter (Stages 1 and 2) stage shifts per hour, lights off and lights on time. Apnea-hypopnea index, defined as the number of respiratory events with oxygen desaturation ≥ 4% per hour, was also determined. Periodic limb movements and PSG-determined sleep onset latency were not included in either cohort due to measure unreliability and missing data (>20%), respectively, in MrOS. All PSG data were scored by the Central Sleep Reading Center (Case Western Reserve University; Cleveland OH) by trained polysomnologists. Data were scored in 30-second epochs according to standard criteria [36-38].

### Sleep Quality Ratings

All participants completed a standard sleep diary on the morning following their polysomnographic monitoring. This diary asked participants to rate the quality of their prior night's sleep using 5-point Likert-type scales, with higher scores indicating higher quality, along two dimensions: “Sleep Depth” (Light to Deep) and “Sleep Restfulness” (Restless to Restful). These sleep quality ratings have been used in prior research [13].

### Self-reported Sleep Variables

Self-reported global sleep quality in the past month was assessed using the Pittsburgh Sleep Quality Index (PSQI) [28], with scores considered both continuously and dichotomously with PSQI scores greater than 5 indicating poor sleep quality. The PSQI was included to relativize the prior-night's sleep quality ratings to the individual's previous month of sleep quality. Daytime sleepiness was assessed using the Epworth Sleepiness Scale [39], with subjects dichotomized into those with scores greater than 10 indicating excessive daytime sleepiness, and those with scores 10 or lower indicating absence of daytime sleepiness. Participants were also asked to report in hours their habitual sleep duration over the previous month.

### Demographic and Clinical Variables

We chose to include variables related to demographics and health (physical, cognitive, mental), selecting variables that are commonly investigated in these and similar cohorts. At the baseline visit, all participants reported level of education attained (grouped as <12 years, 12-16 years, >16 years). At the sleep visit, depression and anxiety were measured with the Geriatric Depression Scale [40] and the Goldberg Anxiety Scale [41] respectively. Additional information collected at the sleep visit and considered in subsequent models included history of diabetes, stroke, or heart disease; caffeine use (in milligrams per day); alcohol use (in drinks per week); smoking status (current, former, never); self-reported ratings of health on a 1-5 Likert-type scale (excellent to very poor, with poor and very poor

collapsed to improve variable distribution); and body mass index and waist circumference (in centimeters). Medication use was assessed by having participants bring to their sleep visit all prescription and non-prescription medications used within the prior 30 days. All medications were entered into an electronic database, verified by pill bottle examination, and each medication was matched to its ingredient(s) based on the Iowa Drug Information Service (IDIS) Drug Vocabulary (College of Pharmacy, University of Iowa, Iowa City, IA) [42]. The present analyses considered three medication categories as correlates: benzodiazepines, antidepressants, and "any reported medication taken for sleep" (which may overlap with the former two categories). All participants also completed a mental status exam at the time of the sleep visit; men completed the Teng Modified Mini Mental State Exam [43] and women completed the Mini Mental State Exam [44]. Several additional variables were available in each cohort separately. Men reported on socioeconomic status relative to their community (1-10 ladder). Women reported whether they lived alone (yes, no).

## Data Analytic Strategy

Table 1 lists all variables included in multivariable analyses. These variables were selected on the basis of previous empirical work examining sleep and sleep quality in these and similar cohorts [13, 45, 46].

Potential predictor variables with greater than 10 percent missing data were excluded from analyses under the assumption that such variables may not be missing at random and imputation may introduce bias [48]. Individuals missing all outcome sleep quality variables (2 in MrOS, 2 in SOF) were excluded from analyses, leaving 1024 men and 459 women. All remaining missing data were imputed using the Expectation Maximization algorithm embedded in version 1.7.3 of the "Amelia II" package in R [49]. There were no missing data for total sleep time, sleep efficiency, wake after sleep onset, sleep-wake shifts, apnea-hypopnea index, bed time, wake time, medication use, race, or education for either cohort. All remaining variables had fewer than 3% missing data, with the exception of smoking status, which was missing in 8% of women.

Given that traditional stepwise regression methods may overfit training data and perform poorly in predicting new data [50, 51], we used two machine learning methods designed to improve upon stepwise regression in selecting and ordering the importance of variables: least absolute shrinkage and selection operator (lasso) penalized regressions [52] and random forests [53].

We used the lasso penalized regression to determine variable selection and importance [52]. Lasso penalized regression is a multivariable regression method useful for predicting an outcome in the presence of a sizeable number of variables. Unlike ordinary least squares regression, lasso regression penalizes models that have many large coefficients under the assumption that such models are likely to be overfit. Lasso regression is useful for handling correlated predictor variables and for performing variable selection as the penalty imposed shrinks some variable coefficients to zero. In lasso penalized regression, the size of the penalty coefficient ($\boldsymbol{\lambda}$) is chosen to minimize model predictive error using ten-fold cross-validation. This optimal $\boldsymbol{\lambda}$ is subsequently used to perform variable selection by shrinking

all coefficients (some to zero). To make the model more robust to randomness in the cross-validation procedure, and following recommendations from the package authors [50], we used the stringent one standard error from the lambda minimum as the penalty in all lasso models. To estimate a ranking of variable importance, the maximal value of lambda at which the variable first entered the model was determined. Separate lasso regression models were evaluated for each of our measures of sleep quality (Sleep Depth and Restfulness) in both men and women. All lasso models were calculated using version 1.9-8 of the "glmnet" package in R [54].

To evaluate whether variables selected by the lasso had statistically significant effect sizes, along with the ordering of these selected variables, we ran an ordinary least squares regression on variables initially selected via the lasso. We randomly split half of the data into a training set and used the remaining half as a test set. We ran the lasso regression on the training set and included all nonzero variables determined by the lasso as ordinary least squares predictors in our test set. This analysis was not completed for women given their reduced sample size. To address multiple comparisons within these ordinary least squares regressions, we used the Benjamini & Hochberg [47] method to control the false discovery rate. All $p$ values reported are two-tailed.

We also used a second nonparametric machine learning method, random forests, to rank the importance of correlates by explanatory relevance [53]. A random forest model is a regression-based procedure averaging across a series of traditional decision trees (in the present sample, 2000) and using a subset of variables (in the present sample, 13) to grow an individual tree, which has the effect of reducing variance associated with any individual tree and reducing correlations between the trees [53]. Random forests rank the importance of all correlates by increases in mean standard errors when a given variable omitted from the model. The relative importance of each variable is derived by scaling the most important correlate to one; we examined all correlates with importance scores ≥ 0.10 relative to the most important correlate. Separate random forest models were evaluated for each of our measures of sleep quality (Sleep Depth and Restfulness). In addition, partial plots of random forest models, showing the impact of a given correlate on sleep quality while simultaneously accounting for all other variables in the model, were computed. All random forest models were calculated using the "randomForestSRC" package version 1.6.1 in R [55].

The two measures of sleep quality (Sleep Depth and Restfulness) were treated as continuous in regression modeling for several key reasons. First, nonparametric random forest models are robust to violations of non-normality, making them a suitable fit for modeling ordinal data. Second, lasso penalized regressions were used for variable selection and not to derive model-specific coefficients, yielding them less sensitive to violations of non-normality. However, we cross-checked all lasso variable selection results against those from a package specifically designed for ordinal response modeling ("lassocr" version 1.0.2.), with all variable selection results confirmed by this second package.

## Results

### Demographic Characteristics and Sleep Quality Ratings

Table 1 presents baseline demographic, clinical and sleep characteristics of our sample. Considering ratings of sleep quality, women reported significantly higher ratings compared to men on Sleep Depth (3.5 ± 1.5 vs. 3.1 ± 1.1, W=187375, $p < 0.001$, Wilcoxon Rank Sum Test), and Restfulness (3.7 ± 1.4 vs. 3.2 ± 1.2, W=180749.5, $p < 0.001$, Wilcoxon). A histogram of Sleep Depth and Sleep Restfulness for men and women is displayed in Figure 1.

### Lasso models

Table 2 lists all non-zero predictor variables returned in lasso models, ranked by order of entry, for Sleep Depth and Restfulness in men and women. Higher sleep efficiency, greater total sleep time, study site and global PSQI score emerged as some of the strongest predictors of sleep quality in both men and women. Sleep stage transitions (both sleep to wake and deep to light sleep) were also important to the models. WASO was not selected in any lasso model, likely due to its very high correlation with sleep efficiency ($r = -0.88$, Pearson). Indeed, when sleep efficiency was removed from lasso models, WASO was selected and emerged as the most important variable in lasso models (results not shown). $R^2$ estimates from lasso models predicting Sleep Depth and Restfulness accounted for only 11-17% of the variance in men and 12-15% of the variance in women.

### Ordinary least squares models

After separating MrOS data into training and test sets, we determined non-zero coefficients from an initial lasso training model and entered these into an OLS regression using the test set. This methodological step is recommended to reduce bias and generate accurate *p* values [50]. Table 3 lists *p* values associated with beta coefficients in this model. Adjusted $R^2$ estimates suggested OLS models accounted for 15-16% of the variance, consistent with the $R^2$ estimates of lasso and random forest models. We also evaluated higher-order polynomials (quadratic, cubic) in the regression analysis but did not find these to improve model fit.

### Random forest models

Figure 2 displays the relative contribution of all correlates with relative importance scores of 10% or greater as determined by random forest models. Higher sleep efficiency, greater total sleep time, study site and higher global PSQI scores again emerged as strong correlates in both men and women. Unlike the lasso, a sparse methodology that excludes highly correlated variables, random forests include correlated variables and, as such, WASO was here a strong correlate in all random forest models for men and women. Higher out-of-bag $R^2$ estimates from random forest models predicting Sleep Depth and Restfulness accounted for only 13-15% of the variance in men and 12-15% of the variance in women.

In addition to describing the relative importance of the predictor variables, random forest models can also be used to examine the relationship between an individual predictor variable and the outcome measure while allowing all other variables to simultaneously enter the model (partial dependence plot, Figure 3). For nearly all continuous predictor variables,

there was a linear relationship between predictor and outcome, coupled with a relative plateau at the lower and upper ranges of the predictor variable (*i.e.,* a sigmoidal relationship). The one exception to this was an inverted U-shaped relationship between habitual sleep duration and Restfulness in women, where both long and short habitual sleep durations were associated with reduced sleep quality reports. Categorical variables in random forest models were also explored. Plots of study site revealed Birmingham was associated with worse sleep quality relative to other sites. The relationship between Sleep Restfulness and education in women was a stepwise decline, with higher education associated with less Restfulness.

PSQI, added as a measure of global dissatisfaction with sleep during the past month (*i.e.,* the historical context of sleep), made a small contribution to each of the random forest models. Given that the PSQI was also returned as a prominent correlate in lasso models, we examined the interaction between PSQI and sleep efficiency in predicting Sleep Restfulness using partial dependence surface plots. In men (Figure 4, top), there was an interaction such that low PSQI scores and low sleep efficiency resulted in slightly better Restfulness than those with high PSQI scores (*i.e.,* those with past-month reports of poor sleep) and low sleep efficiency. The primary relationship between sleep efficiency and Restfulness, however, was independent of PSQI at elevated sleep efficiencies. In women (Figure 4, bottom), however, those with low sleep efficiency had low Restfulness regardless of PSQI. Those women with higher sleep efficiency had higher Restfulness when they had lower PSQI scores.

## Correlate exploration

To better understand the contribution of sleep efficiency and site, two of the strongest correlates that emerged across models and outcomes, we ran an additional set of exploratory analyses and lasso penalized regressions in men. We did not run parallel analyses for women given the reduced sample size.

To explore sleep efficiency, we stratified our data into high, medium and low sleep efficiency groups and reran lasso regression models predicting sleep quality in each sleep efficiency group. For those with low sleep efficiency (<70.2%), sleep efficiency remained a correlate in the lasso model, along with total sleep time, age, PSQI, and mental status. Thus, despite capping its variance by stratifying, sleep efficiency was still a relevant predictor for this subset, and this suggests that even small differences in sleep efficiency can explain variation in sleep quality. For those with moderate sleep efficiency (70.2%-81.4%), study site and PSQI were the only correlates of sleep quality in this group, consistent with previous models (presented above) that these two variables are of high importance. At high sleep efficiency (>81.4%), no non-zero variables were returned, suggesting that sleep efficiency alone was a sufficient correlate of sleep quality. In other words, if sleep efficiency is above a certain threshold, no observed variation in the other variables can reliably predict even a small change in sleep quality.

To explore site, a Kruskal-Wallis test revealed significant differences in sleep quality across site ($\mathbf{X}^2 = 60.7$, $p<0.0001$), with Wilcoxon follow-up tests indicating individuals in Birmingham endorsed lower ratings on each of the two sleep quality measures ($p<0.001$ for all), and individuals in Palo Alto endorsing higher ratings of Sleep Depth relative to

Birmingham, Minneapolis and Portland ($p<0.05$). To further explore the characteristics that might vary by site, we ran an additional lasso penalized regression with site as the response variable (using a "multinomial" distribution to reflect the categorical response) and all other variables listed in Table 1 as predictor variables. As determined by the non-zero coefficients returned, along with their direction, sleep quality at the Birmingham site appeared most strongly related to increased N1%, decreased N3%, earlier bedtimes, less alcohol consumption and including a greater number of African Americans. Sleep quality in Palo Alto was most strongly related to older age, more Asians and those of Hispanic origin, greater college education, fewer ratings of poor health, more alcohol consumption and smaller waist circumferences.

## Discussion

We sought to determine the objective correlates of subjective sleep quality in older men and women using machine learning methods. Perhaps the most notable finding to emerge from this investigation was precisely how little the metrics traditionally derived from PSG contribute to an individual's perception of the quality of their previous night of sleep. Though we chose our variables in keeping with prior research, it may be the case that other metrics derivable from PSG—for example, variables extracted from non-REM power spectral analysis or the cyclic alternating pattern of non-REM sleep [15]—may yield more sensitive correlates of subjective sleep quality accounting for more variance.

Of the correlates examined, PSG-derived sleep efficiency explained the most variance in prior night sleep quality. Sleep efficiency was among the top variables in all models, across both dimensions of sleep quality (Sleep Depth and Restfulness), and across both sexes. Indeed, when sleep efficiencies surpassed approximately 80%, none of the 37 other variables in our lasso penalized regression model were deemed important. Consistent with sleep efficiency as the sole explanatory variable at efficiencies above 80%, random forest partial dependence plots revealed that sleep quality ratings improved only modestly as sleep efficiencies surpassed this 80% range. Though sleep efficiency was clearly superior, our machine learning algorithms also determined sleep stage transitions, total sleep time, PSQI scores, and study site were most informative in explaining subjective sleep quality ratings.

Our findings are consistent with some previous accounts and discrepant with others. Consistent with our finding on sleep efficiency, Åkerstedt and colleagues reported that PSG-determined sleep efficiency was among the best correlates of sleep quality in a young adult sample followed over a series of irregular nights (N=8) [12] and in an adult sample examined on a single night (N=37) [31]. Laffan and colleagues [13] highlighted the importance of sleep to wake transitions in the Sleep Heart Health Study, a middle-aged and older community-based U.S. cohort (N=5,684), using sleep quality ratings that were identical to the ones included here. We also found that sleep to wake transitions were important in our elderly cohort, which, together with our sleep efficiency findings, highlights the importance of sleep and wake continuity during the sleep period in self-reports of sleep quality. However, some important discrepancies between our findings and those of previous researchers merit mention. Overall, sleep staging variables were a weak correlate of subjective sleep quality. In women only, we observed some evidence that increased N1 was

related to poorer sleep quality and, to a lesser extent, increased N2 related to greater sleep quality. Although prior research supports a relationship between slow wave sleep and subjective sleep quality [8-10], we did not observe this association here. One might speculate that the reduced quantity of N3 in an elderly sample may have attenuated this finding. However, women in our sample spent 20% of the night in N3 sleep, which would likely be sufficient for such an N3-sleep quality relationship to emerge. It has, however, been proposed that slow wave sleep may vary in its association with cognition in older as compared to younger populations; slow wave sleep may also be a weaker correlate of sleep quality in older compared to younger individuals [56].

Given that women experience higher rates of insomnia [24] and that women are more likely to report worse subjective sleep estimates relative to objective findings [25, 26], we expected that women would rate their sleep quality less favorably overall. This was not the case. Women endorsed higher ratings on both measures of subjective sleep quality compared to men. Interestingly, women also reported significantly higher PSQI scores indicating poorer global sleep quality. As illustrated in Figure 3, the interplay between these variables indicates that women endorse comparable ratings of sleep quality even as PSQI scores are higher and sleep efficiency scores are lower relative to men, suggesting the interpretation of sleep quality differs between the sexes. Since the study's outcomes were designed to characterize the prior night's sleep during monitoring, these findings suggest that subjective assessment of sleep quality over a single night of monitoring differs from global indices of sleep quality over the prior month. It is possible that women, who generally seek more healthcare than men, may also perceive monitoring to be less intrusive than do men. It is important to note that that the male and female data used here were taken from separate cohorts, from multiple geographic locations across the United States, and with enrollment commencing seven years part. It is possible that sex differences may be explained in part or in full by these factors. Future research should continue to include sex as a moderator or examine the sexes in parallel as was done here. Future research should also further examine acute as compared to chronic sleep perceptions.

Our study, in the context of these previous studies, also brings up an important concept: any investigation of subjective sleep quality is necessarily dependent on the measure used to assess it. The two subjective measures we captured in these data (depth, restfulness) are two of a myriad of psychological aspects that a person might define as “quality,” and there may be other measures of subjective sleep quality that would better correspond to our measures of objective sleep [57]. For example, the Consensus Sleep Diary [58], which was published after these data were collected, recommends rating the “quality of your sleep” on a 5-point scale ranging from “very poor” to “very good.” We did not have this question available for analysis. Instead, we can conclude that these two particular questions, presented in the manner in which they were, very poorly reflect objective metrics derived from standard sleep variables on a single night of study. These two questions, however, were related to an overlapping set of variables, and sleep efficiency was a main contributor in all of the models. This confluence of the models onto a single objective measure (PSG-determined sleep efficiency) implies that it is unlikely that a different subjective measure would be related to variables very different from the ones we observed here. We note that there may not be an “average” when it comes to ratings of prior-night sleep quality, and little understanding of

how these ratings may change with age, though this is an important area for future research. It is also worth noting that these subjective measures concern psychological aspects of sleep and do not necessarily represent any of the biological functions of sleep. That someone reports having a “good” night of sleep does not necessitate that the sleep sufficiently meets all of its biological functions. Indeed, the weak relationship between polysomnography summary characteristics and subjective sleep quality reports here might suggest that the biological recovery processes in sleep do not reflect the subjective phenomenology of sleep quality. There may be additional features across multiple domains of analysis worth considering: subjective processes (e.g., expectations about sleep quality); biological processes (e.g., terminal sleep stage prior to awakening; regional brain metabolism); or sleep inertia severity (e.g., decrements in objective performance).

There are several reasons why site may have emerged as a prominent correlate in our multivariable models. Individuals in Birmingham reported poorer sleep quality than at all other sites, and this reduction in sleep quality was not attributable to race, socioeconomic or health variables alone. It is possible that variation in sleep quality across sites reflects a composite of correlates and that sleep quality in Birmingham or Palo Alto had a different relationship to various correlates than sleep quality in other sites. For example, individuals in Birmingham went to bed earlier, drank less, more were African-American, and had more light stage sleep, and it may be shared variance between these correlates influences sleep quality via a unique combination not present in subjects at other sites. Alternatively, though procedures were standardized across study sites, we are unable to discount the possibility that protocol variation in polysomnography or subjective data collection may have influenced differences across sites.

Several limitations and strengths of the present research merit mention. Our focus was on elderly men and women and generalizability of these results to other age groups is uncertain. Though we included sleep, medication use, medical history, and health variables in our models, it is likely that additional variables with explanatory relevance may have been omitted from the present models. For example, quantitative EEG was not available at these study visits, but this remains an important area for future exploration. We also were unable to examine the impact of sleep onset latency or limb movement in relation to sleep quality given unreliability of these measures in the present sample. Our study also has three considerable strengths over previous investigations. First, we use machine learning techniques that reduce bias and prediction error relative to traditional methods (e.g. stepwise variable selection). Second, rather than focusing on a limited set of correlates, we considered a wide range of demographic, clinical and sleep variables together to explore the importance of each correlate, along with the associations between correlates of various types. Finally, we use a large sample of elderly men and women who underwent polysomnography and completed next-morning sleep quality ratings in our investigation. This is the largest investigation of sleep quality using these newly-developed analytical techniques.

## Conclusions

Our analyses determined that objective sleep efficiency was the most prominent correlate of subjective sleep quality in older men and women. Total sleep time, sleep stage transitions,

PSQI, and study site were also important predictors of sleep quality. Improving upon previous work, we used machine learning algorithms to model multiple sleep, demographic and clinical correlates together. However, the thirty-eight variables examined still explained little of the overall variance in subjective sleep quality. Sleep quality remains elusive, and it will be the important task of future research to consider novel correlates to better understand it.

## Acknowledgments

The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Advancing Translational Sciences (NCATS), and NIH Roadmap for Medical Research under the following grant numbers: U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, and UL1 TR000128. The National Heart, Lung, and Blood Institute (NHLBI) provides funding for the MrOS Sleep ancillary study "Outcomes of Sleep Disorders in Older Men" under the following grant numbers: R01 HL071194, R01 HL070848, R01 HL070847, R01 HL070842, R01 HL070841, R01 HL070837, R01 HL070838, and R01 HL070839. The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576. Dr. Zeitzer was supported by the Sierra-Pacific Mental Illness Research, Education, and Clinical Center.

## List of Abbreviations

| | |
|---|---|
| **EEG** | Electroencephalogram |
| **EMG** | Electromyogram |
| **EOG** | Electrooculogram |
| **IDIS** | Iowa Drug Information Service |
| **Lasso** | Least Absolute Shrinkage and Selection Operator |
| **MrOS** | Osteoporotic Fractures in Men Study |
| **OLS** | Ordinary Least Squares |
| **PSQI** | Pittsburgh Sleep Quality Index |
| **SE** | Sleep Efficiency |
| **SOF** | Study of Osteoporotic Fractures |
| **TST** | Total Sleep Time |
| **WASO** | Wake After Sleep Onset |
| **PSG** | Polysomnography |

## References

1. Edinger JD, Fins AI, Glenn DM, Sullivan RJ Jr. Bastian LA, Marsh GR, Dailey D, Hope TV, Young M, Shaw E, et al. Insomnia and the eye of the beholder: are there clinical markers of objective sleep disturbances among adults with and without insomnia complaints? J Consult Clin Psychol. 2000; 68(4):586–593. [PubMed: 10965634]

2. Salin-Pascual RJ, Roehrs TA, Merlotti LA, Zorick F, Roth T. Long-term study of the sleep of insomnia patients with sleep state misperception and other insomnia patients. Am J Psychiatry. 1992; 149(7):904–908. [PubMed: 1609869]

3. Roth T, Zammit G, Lankford A, Mayleben D, Stern T, Pitman V, Clark D, Werth JL. Nonrestorative sleep as a distinct component of insomnia. Sleep. 2010; 33(4):449–458. [PubMed: 20394313]

4. Ohayon MM. Prevalence and correlates of nonrestorative sleep complaints. Arch Intern Med. 2005; 165(1):35–41. [PubMed: 15642872]

5. Walsh JK, Coulouvrat C, Hajak G, Lakoma MD, Petukhova M, Roth T, Sampson NA, Shahly V, Shillington A, Stephenson JJ, et al. Nighttime insomnia symptoms and perceived health in the America Insomnia Survey (AIS). Sleep. 2011; 34(8):997–1011. [PubMed: 21804662]

6. O'Donnell D, Silva EJ, Munch M, Ronda JM, Wang W, Duffy JF. Comparison of subjective and objective assessments of sleep in healthy older subjects without sleep complaints. J Sleep Res. 2009; 18(2):254–263. [PubMed: 19645969]

7. Baekeland F, Hoy P. Reported vs recorded sleep characteristics. Arch Gen Psychiatry. 1971; 24(6): 548–551. [PubMed: 4325303]

8. Westerlund A, Lagerros YT, Kecklund G, Axelsson J, Akerstedt T. Relationships between questionnaire ratings of sleep quality and polysomnography in healthy adults. Behav Sleep Med. 2014:1–15. [PubMed: 23390921]

9. Hoch CC, Reynolds CF 3rd, Kupfer DJ, Berman SR, Houck PR, Stack JA. Empirical note: self-report versus recorded sleep in healthy seniors. Psychophysiology. 1987; 24(3):293–299. [PubMed: 3602285]

10. Riedel BW, Lichstein KL. Objective sleep measures and subjective sleep satisfaction: how do older adults with insomnia define a good night's sleep? Psychol Aging. 1998; 13(1):159–163. [PubMed: 9533198]

11. Bonnet MH, Johnson LC. Relationship of arousal threshold to sleep stage distribution and subjective estimates of depth and quality of sleep. Sleep. 1978; 1(2):161–168. [PubMed: 227030]

12. Akerstedt T, Hume K, Minors D, Waterhouse J. The meaning of good sleep: a longitudinal study of polysomnography and subjective sleep quality. J Sleep Res. 1994; 3(3):152–158. [PubMed: 10607120]

13. Laffan A, Caffo B, Swihart BJ, Punjabi NM. Utility of sleep stage transitions in assessing sleep continuity. Sleep. 2010; 33(12):1681–1686. [PubMed: 21120130]

14. Buysse DJ, Reynolds CF 3rd, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women using the Pittsburgh Sleep Quality Index (PSQI). Sleep. 1991; 14(4):331–338. [PubMed: 1947597]

15. Krystal AD, Edinger JD. Measuring sleep quality. Sleep Med. 2008; 9(Suppl 1):S10–17. [PubMed: 18929313]

16. Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep. 2004; 27(7):1255–1273. [PubMed: 15586779]

17. Ancoli-Israel S, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Periodic limb movements in sleep in community-dwelling elderly. Sleep. 1991; 14(6):496–500. [PubMed: 1798881]

18. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev. 2002; 6(2):97–111. [PubMed: 12531146]

19. Bloom HG, Ahmed I, Alessi CA, Ancoli-Israel S, Buysse DJ, Kryger MH, Phillips BA, Thorpy MJ, Vitiello MV, Zee PC. Evidence-based recommendations for the assessment and management of sleep disorders in older persons. J Am Geriatr Soc. 2009; 57(5):761–789. [PubMed: 19484833]

20. Ancoli-Israel S, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Sleep-disordered breathing in community-dwelling elderly. Sleep. 1991; 14(6):486–495. [PubMed: 1798880]

21. Young T, Shahar E, Nieto FJ, Redline S, Newman AB, Gottlieb DJ, Walsleben JA, Finn L, Enright P, Samet JM, et al. Predictors of sleep-disordered breathing in community-dwelling adults: the Sleep Heart Health Study. Arch Intern Med. 2002; 162(8):893–900. [PubMed: 11966340]

22. Grandner MA, Martin JL, Patel NP, Jackson NJ, Gehrman PR, Pien G, Perlis ML, Xie D, Sha D, Weaver T, et al. Age and sleep disturbances among American men and women: data from the U.S. Behavioral Risk Factor Surveillance System. Sleep. 2012; 35(3):395–406. [PubMed: 22379246]

23. Zilli I, Ficca G, Salzarulo P. Factors involved in sleep satisfaction in the elderly. Sleep Med. 2009; 10(2):233–239. [PubMed: 18387848]
24. Zhang B, Wing YK. Sex differences in insomnia: a meta-analysis. Sleep. 2006; 29(1):85–93. [PubMed: 16453985]
25. van den Berg JF, Miedema HM, Tulen JH, Hofman A, Neven AK, Tiemeier H. Sex differences in subjective and actigraphic sleep measures: a population-based study of elderly persons. Sleep. 2009; 32(10):1367–1375. [PubMed: 19848365]
26. Vitiello MV, Larsen LH, Moe KE. Age-related sleep change: gender and estrogen effects on the subjective-objective sleep quality relationships of healthy, noncomplaining older men and women. J Psychosom Res. 2004; 56(5):503–510. [PubMed: 15172206]
27. Redline S, Kirchner HL, Quan SF, Gottlieb DJ, Kapur V, Newman A. The effects of age, sex, ethnicity, and sleep-disordered breathing on sleep architecture. Arch Intern Med. 2004; 164(4): 406–418. [PubMed: 14980992]
28. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. 1989; 28(2):193–213. [PubMed: 2748771]
29. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. Sleep. 2006; 29:1155–1173. [PubMed: 17040003]
30. Akerstedt T, Hume K, Minors D, Waterhouse J. The subjective meaning of good sleep, an intraindividual approach using the Karolinska Sleep Diary. Percept Mot Skills. 1994; 79(1 Pt 1): 287–296. [PubMed: 7991323]
31. Keklund G, Akerstedt T. Objective components of individual differences in subjective sleep quality. J Sleep Res. 1997; 6(4):217–220. [PubMed: 9493520]
32. Ellis BW, Johns MW, Lancaster R, Raptopoulos P, Angelopoulos N, Priest RG. The St. Mary's Hospital sleep questionnaire: a study of reliability. Sleep. 1981; 4(1):93–97. [PubMed: 7232974]
33. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. Contemp Clin Trials. 2005; 26(5):569–585. [PubMed: 16084776]
34. Blank JB, Cawthon PM, Carrion-Petersen ML, Harper L, Johnson JP, Mitson E, Delay RR. Overview of recruitment for the osteoporotic fractures in men study (MrOS). Contemp Clin Trials. 2005; 26(5):557–568. [PubMed: 16085466]
35. Cummings SR, Black DM, Nevitt MC, Browner WS, Cauley JA, Genant HK, Mascioli SR, Scott JC, Seeley DG, Steiger P, et al. Appendicular bone density and age predict hip fracture in women. The Study of Osteoporotic Fractures Research Group. JAMA. 1990; 263(5):665–668. [PubMed: 2404146]
36. Redline S, Sanders MH, Lind BK, Quan SF, Iber C, Gottlieb DJ, Bonekat WH, Rapoport DM, Smith PL, Kiley JP. Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep Heart Health Research Group. Sleep. 1998; 21(7):759–767. [PubMed: 11300121]
37. Association. ASD: EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep. 1992; 15(2):173–184. [PubMed: 11032543]
38. Rechtschaffen, A., Kales, A. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. US Department of Health, Education, and Welfare, Government Printing Office; Washington D.C.: 1968.
39. Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep. 1991; 14:540–545. [PubMed: 1798888]
40. Sheikh J, Yesavage J. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clin Gerontol. 1986; 5:165–173.
41. Goldberg D, Bridges K, Duncan-Jones P, Grayson D. Detecting anxiety and depression in general medical settings. BMJ. 1988; 297(6653):897–899. [PubMed: 3140969]
42. Pahor M, Chrischilles EA, Guralnik JM, Brown SL, Wallace RB, Carbonin P. Drug data coding and analysis in epidemiologic studies. Eur J Epidemiol. 1994; 10(4):405–411. [PubMed: 7843344]

43. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. J Clin Psychiatry. 1987; 48(8):314–318. [PubMed: 3611032]
44. Folstein MF, Robins LN, Helzer JE. The Mini-Mental State Examination. Arch Gen Psychiatry. 1983; 40(7):812. [PubMed: 6860082]
45. Patel SR, Blackwell T, Redline S, Ancoli-Israel S, Cauley JA, Hillier TA, Lewis CE, Orwoll ES, Stefanick ML, Taylor BC, et al. The association between sleep duration and obesity in older adults. Int J Obes (Lond). 2008; 32(12):1825–1834. [PubMed: 18936766]
46. Blackwell T, Yaffe K, Laffan A, Ancoli-Israel S, Redline S, Ensrud KE, Song Y, Stone KL. Osteoporotic Fractures in Men Study G: Associations of objectively and subjectively measured sleep quality with subsequent cognitive decline in older community-dwelling men: the MrOS sleep study. Sleep. 2014; 37(4):655–663. [PubMed: 24899757]
47. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc B (Methodological). 1995:289–300.
48. Allison, PD. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-136. Sage; Thousand Oaks, CA: 2001. Missing data..
49. Honaker J, King G, Blackwell M, Amelia. A program for missing data. J Stat Software. 2011; 45(7):1–47.
50. Hastie T, Tibshirani R, Friedman J, Franklin J. The elements of statistical learning: data mining, inference and prediction. Math Intelligencer. 2005; 27(2):83–85.
51. Harrell F, Lee KL, Mark DB. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996; 15:361–387. [PubMed: 8668867]
52. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Stat Soc B (Methodological). 1996:267–288.
53. Breiman L. Random forests. Machine learning. 2001; 45(1):5–32.
54. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Software. 2010; 33(1):1.
55. Ishwaran H, Kogalur U. randomForestSRC: Random Forests for Survival, Regression and Classification (RF-SRC). R package version. 2014; 1(0)
56. Scullin MK. Sleep, memory, and aging: the link between slow-wave sleep and episodic memory changes from younger to older adults. Psychol Aging. 2013; 28(1):105–114. [PubMed: 22708533]
57. Harvey AG, Stinson K, Whitaker KL, Moskovitz D, Virk H. The subjective meaning of sleep quality: a comparison of individuals with and without insomnia. Sleep. 2008; 31(3):383–393. [PubMed: 18363315]
58. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, Morin CM. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep. 2012; 35(2):287–302. [PubMed: 22294820]

## Highlights

- We explored correlates of subjective sleep quality using machine learning techniques
- Data from a large cohort of older, community-dwelling adults were examined
- Polysomnography, clinical, and demographic measures did not explain sleep quality well
- Sleep efficiency was the most important of the objective correlates of sleep quality

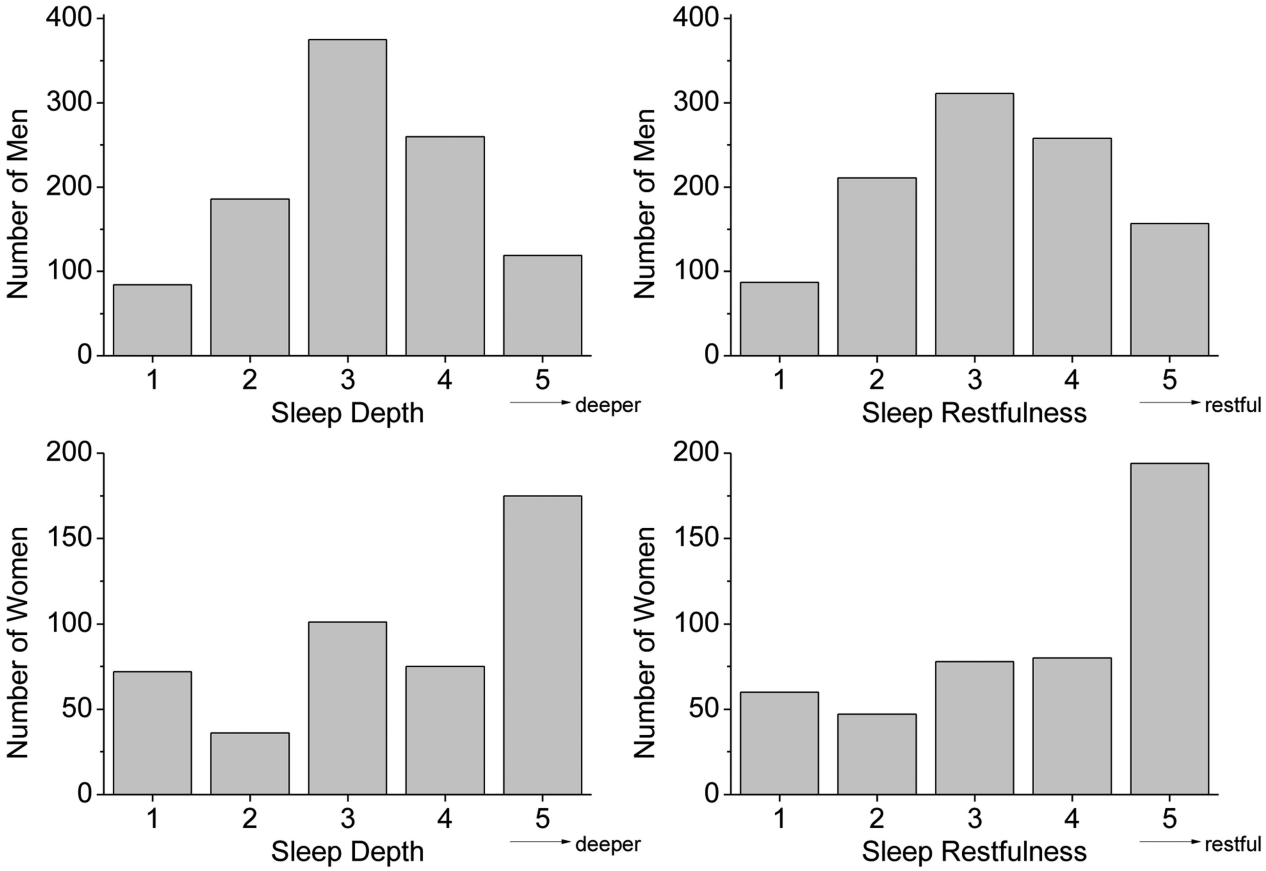


**Figure 1.**
Histograms of sleep quality outcome variables in men (top) and women (bottom)

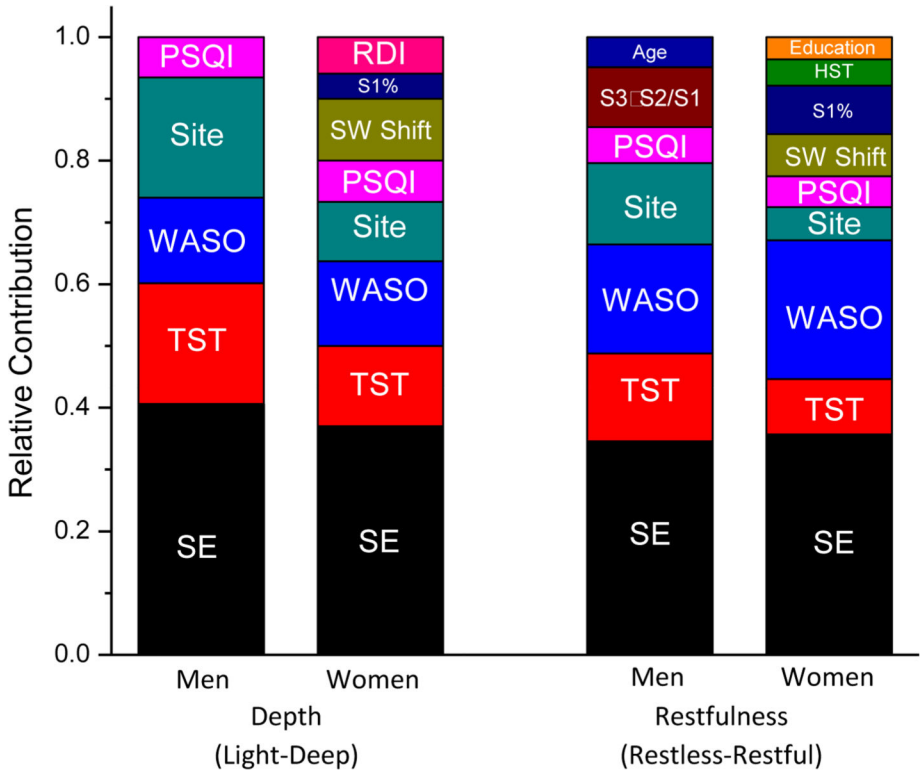


**Figure 2.**

Top predictors of sleep quality by random forest models. See text for details on each measure. Education, level of education; HST, habitual sleep time; S1%, percent of sleep spent in NREM stage 1; S3->S2/S1, number of transitions between stage 3 and either stage 2 or 1 of NREM; PSQI, Pittsburgh Sleep Quality Index; RDI, Respiratory Disturbance Index; SE, sleep efficiency; Site, city of data collection; SW Shift, number of transitions between sleep and wake; TST, total sleep time; WASO, wake after sleep onset

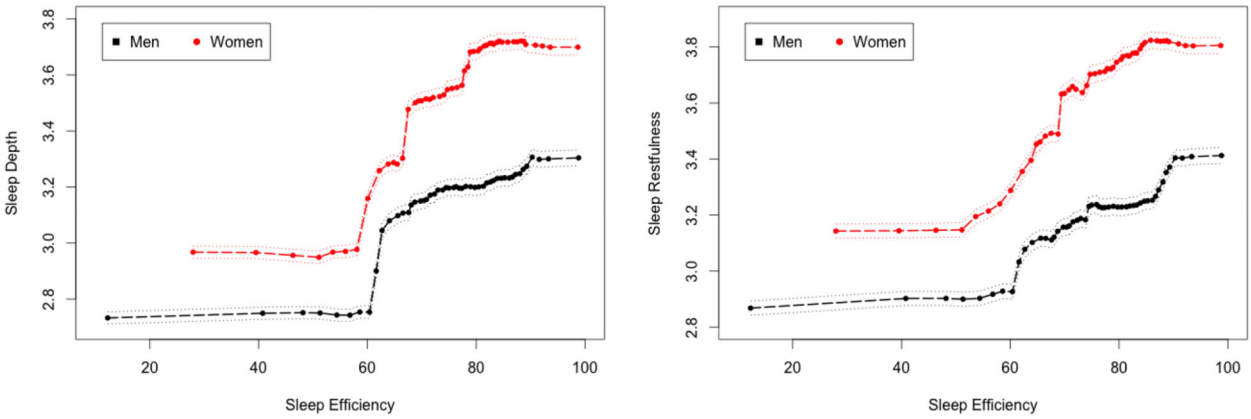


**Figure 3.**
Partial dependence plots of sleep efficiency predicting sleep depth and sleep restfulness in men and women.

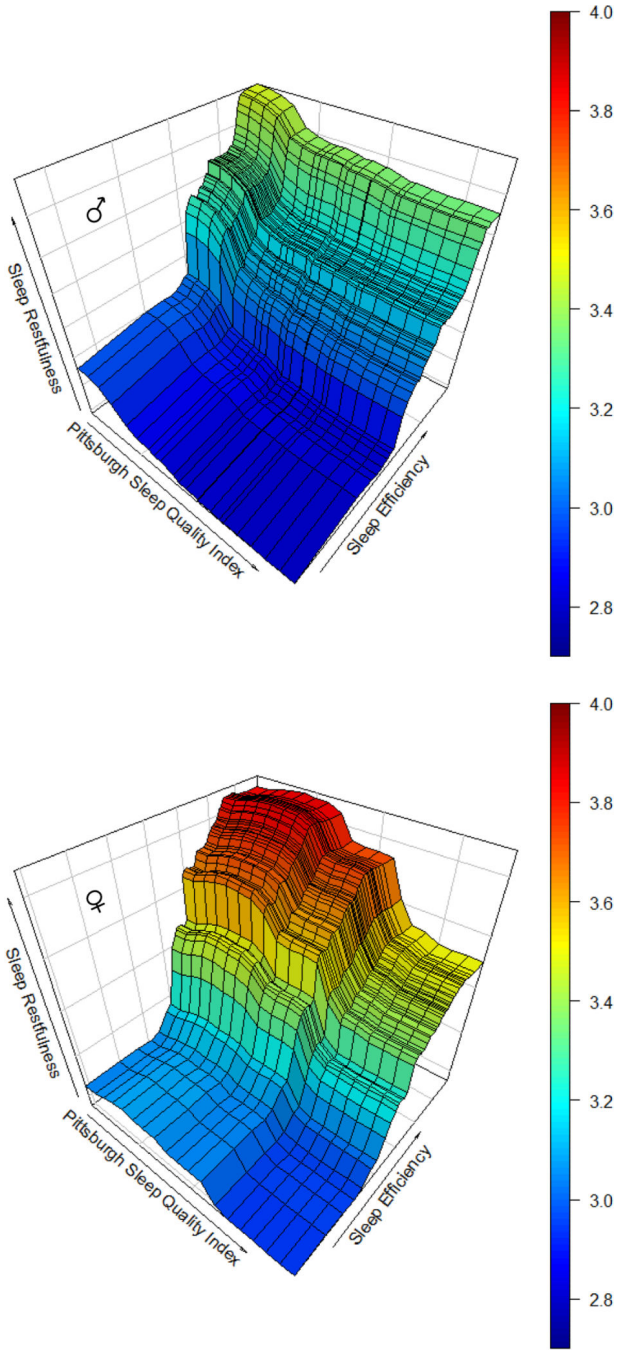


**Figure 4.**
Partial dependence surface plot of sleep efficiency and Pittsburgh Sleep Quality Index predicting sleep restfulness in men (top) and women (bottom)

## Table 1

Demographic, clinical and sleep variables considered in machine learning models.

| Variable | Men (n = 1024) | Women (n = 459) |
|---|---|---|
| Age, M ± SD | 81.1 ± 4.6 | 82.9 ± 3.3 |
| Race, No. (%)[a] | | |
| White | 892 (87.1) | 421 (91.7) |
| African American | 40 (3.9) | 38 (8.3) |
| Asian | 48 (4.7) | 0 |
| Hispanic | 28 (2.7) | 0 |
| Other | 16 (1.6) | 0 |
| Education, No. (%)[a] | | |
| ≤ 12 years | 173 (16.9) | 321 (69.9) |
| 12-16 years | 417 (40.7) | 104 (22.7) |
| > 16 years | 434 (42.4) | 34 (7.4) |
| Health, No. (%) | | |
| Excellent | 324 (31.7) | 100 (21.8) |
| Good | 540 (52.8) | 253 (55.1) |
| Fair | 142 (13.9) | 96 (20.9) |
| Poor – Very Poor | 18 (1.8) | 10 (2.2) |
| Geriatric Depression Scale, M ± SD | 1.8 ± 2.2 | 2.4 ± 2.7 |
| Goldberg Anxiety Scale, M ± SD | 0.74 ± 1.6 | 1.6 ± 2.4 |
| Pittsburg Sleep Quality Index, M ± SD | 5.5 ± 3.1 | 6.7 ± 3.8 |
| Epworth Sleepiness Scale, M ± SD | 7.0 ± 3.9 | 5.8 ± 3.7 |
| Habitual Sleep Duration (hours), M ± SD | 7.2 ± 1.2 | 6.9 ± 1.4 |
| Body Mass Index, M ± SD | 26.9 ± 3.8 | 27.7 ± 4.6 |
| Waist Circumference (cm), M ± SD | 101 ± 12.3 | 88.5 ± 12.5 |
| Poor Sleep Quality, No. (%) | 445 (44.0) | 259 (56.4) |
| Antidepressant Use, No. (%) | 108 (10.5) | 51 (11.1) |
| Benzodiazepine Use, No. (%) | 39 (3.8) | 37 (8.8) |
| Sleep Medication Use, No. (%) | 142 (13.9) | 81 (17.6) |
| Drinks Per Week, M ± SD | 1.8 ± 1.7 | 1.0 ± 2.9 |
| Caffeine Per Day (mg), M ± SD | 246 ± 247 | 154 ± 155 |
| Smoking Status, No. (%) | | |
| Current | 13 (1.3) | 4 (0.0) |
| Former | 526 (51.4) | 143 (34.0) |
| Never | 484 (47.3) | 274 (65.1) |
| Cerebrovascular History, No. (%) | 135 (13.2) | 64 (13.9) |
| Cardiovascular History, No. (%) | 319 (31.4) | 149 (32.5) |
| Diabetes, No. (%) | 167 (16.3) | 62 (13.5) |

| Variable | Men (n = 1024) | Women (n = 459) |
|---|---|---|
| Mini-Mental State Exam | -- | 28.3 ± 1.6 |
| Modified Mini-Mental State Exam | 92.5 ± 6.3 | -- |
| Socioeconomic ladder (1-10), M ± SD[a,b] | 7.2 ± 1.6 | -- |
| Currently living alone, No. (%)[c] | -- | 282 (61.4) |
| Apnea-Hypopnea Index, M ± SD | 12.7 ± 14.3 | 10.0 ± 12.4 |
| Total Sleep Time, M ± SD | 342 ± 78.2 | 347 ± 77.1 |
| Wake After Sleep Onset, M ± SD | 123 ± 69.3 | 98.6 ± 66.5 |
| Sleep Efficiency, M ± SD | 73.9 ± 13.2 | 74.3 ± 13.2 |
| Stage 1 Percent, M ± SD | 12.4 ± 8.9 | 5.3 ± 3.8 |
| Stage 2 Percent, M ± SD | 61.6 ± 10.7 | 55.4 ± 12.3 |
| Stage 3/4 Percent, M ± SD | 6.7 ± 7.0 | 20.8 ± 12.5 |
| REM Percent, M ± SD | 19.3 ± 7.1 | 18.5 ± 7.2 |
| REM Latency, M ± SD | 94.0 ± 69.5 | 111 ± 82.6 |
| Sleep to Wake Shifts per hour, M ± SD | 6.1 ± 3.4 | 3.8 ± 2.0 |
| Deep to Light Stage Shifts per hour, M ± SD | 1.7 ± 1.4 | 3.7 ± 2.3 |
| Lights Off Time | 22:28 ± 1:07 | 22:48 ± 1:06 |
| Lights On Time | 6:20 ± 1:09 | 6:41 ± 1:07 |
| Study Site | -- | -- |

Note:

[a]Variables assessed at baseline visit.

[b]Socioeconomic status available for men only.

[c]Living alone status available for women only.

## Table 2

Top predictors of sleep quality by lasso regression in men and women.

| **Sleep Depth (Light – Deep)** | | **Sleep Restfulness (Restless – Restful)** | |
|---|---|---|---|
| **Men** | **Women** | **Men** | **Women** |
| SE | SE | SE | SE |
| TST | Sleep to Wake Shifts | Site | Site |
| Site | PSQI | TST | Education |
| PSQI | Site | Deep to Light Stage Shifts | GAS |
| | | PSQI | Sleep to Wake Shifts |
| | | Age | |

Variables are listed in descending order of importance. Abbrev.: GAS, Goldberg Anxiety Scale; PSQI, Pittsburgh Sleep Quality Index; SE, sleep efficiency; TST, total sleep time.

## Table 3

Top predictors of sleep quality by ordinary least squares regression in men.

| **Sleep Depth (Light – Deep)** | | **Sleep Restfulness (Restless – Restful)** | |
|---|---|---|---|
| **Men** | ***p* value** | **Men** | ***p* value** |
| SE | <0.001 | PSQI | <0.01 |
| PSQI | <0.05 | SE | <0.01 |
| Site: Birmingham | <0.05 | Site: Palo Alto | <0.05 |
| Site: Palo Alto | <0.05 | | |

Variables entered into above models reflect non-zero coefficients derived from a Lasso training model. Variables are displayed in descending order of standardized coefficients. Only variables with significant *p* values (determined via false discovery rate) are displayed. PSQI, Pittsburgh Sleep Quality Index; SE, sleep efficiency; TST, total sleep time.