Computers in Biology and Medicine 110 (2019) 276–288

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Sleep quality prediction in caregivers using physiological signals

Reza Sadeghi[a,*], Tanvi Banerjee[a], Jennifer C. Hughes[b], Larry W. Lawhorne[c]

[a] Department of Computer Science and Engineering, Kno.e.sis Research Center, Wright State University, Dayton, OH, USA
[b] Department of Social Work, Wright State University, Dayton, OH, USA
[c] Department of Geriatrics, Boonshoft School of Medicine, Wright State University, Dayton, OH, USA

ARTICLE INFO

Keywords:
Caregivers of people with dementia (CPWD)
Physiological signals
Sleep quality
Restfulness
E4 wristband

ABSTRACT

Most caregivers of people with dementia (CPWD) experience a high degree of stress due to the demands of providing care, especially when addressing unpredictable behavioral and psychological symptoms of dementia. Such challenging responsibilities make caregivers susceptible to poor sleep quality with detrimental effects on their overall health. Hence, monitoring caregivers' sleep quality can provide important CPWD stress assessment. Most current sleep studies are based on polysomnography, which is expensive and potentially disrupts the caregiving routine. To address these issues, we propose a clinical decision support system to predict sleep quality based on trends of physiological signals in the deep sleep stage. This system utilizes four raw physiological signals using a wearable device (E4 wristband): heart rate variability, electrodermal activity, body movement, and skin temperature. To evaluate the performance of the proposed method, analyses were conducted on a two-week period of sleep monitored on eight CPWD. The best performance is achieved using the random forest classifier with an accuracy of 75% for sleep quality, and 73% for restfulness, respectively. We found that the most important features to detect these measures are sleep efficiency (ratio of amount of time asleep to the amount of time in bed) and skin temperature. The results from our sleep analysis system demonstrate the capability of using wearable sensors to measure sleep quality and restfulness in CPWD.

1. Introduction

Extensive studies indicate pivotal effects of sleep in job performance [1], memory [2], fatigue recovery [3], and both mental [4] and physical health [5]. The discharge of anabolic hormones (e.g. prolactin, testosterone, luteinizing hormone) during sleep along with physical restoration lead to feeling refreshed after sleep [6]. When individuals experience difficulty sleeping, the detrimental effects range from daytime sleepiness, performance reduction, to lack of attention [7]. Therefore, assessing sleep performance, especially for individuals with demanding tasks such as caregiving, is beneficial to both caregivers and loved ones with dementia-related illness.

Caregivers of people suffering dementia-related illnesses are more susceptible to sleep problems than the rest of the caregiver community [8]. At least, two-thirds of these caregivers reported sleep disturbance during the course of their care relationship [9]. As most primary caregivers are unpaid spouse or other family members of the patient [10], we focus on predicting sleep quality of primary CPWD in this paper.

Family CPWD face physical and psychological challenges due to the high level of stress [11,12]. The potential for long-standing worry coupled with caregiving challenges contributes to chronic sleep deficiency. Poor sleep quality increases the difficulty of providing quality care, and increases risk of persistent tiredness, heart disease, and premature death [13]. Sleep quality monitoring can be beneficial for caregivers to identify potential concerns and provide potential interventions before reaching a detrimental level. To the best of our knowledge, this is the first work trying to predict sleep quality specifically among family CPWD who provide unpaid services with a cost savings to society of more than $470 billion annually [10].

A well-founded sleep quality assessment requires a comprehensive understanding of underlying sleep mechanisms. The American Academy of Sleep Medicine (AASM) suggests using polysomnography (PSG) as a standard guideline to evaluate sleep architecture [14]. PSG enables physicians to assess sleep and sleep disturbance by continuous monitoring of multiple physiological parameters during sleep. However, PSG is expensive and intrusive as a sleep study which is conducted by attaching numerous electrodes to the participants.

To inexpensively evaluate sleep quality of different subjects sleeping in their usual home environments, we propose a clinical decision

* Corresponding author.
E-mail addresses: sadeghi.2@wright.edu (R. Sadeghi), tanvi.banerjee@wright.edu (T. Banerjee), jennifer.hughes@wright.edu (J.C. Hughes), larry.lawhorne@wright.edu (L.W. Lawhorne).

https://doi.org/10.1016/j.compbiomed.2019.05.010
Received 13 February 2019; Received in revised form 10 May 2019; Accepted 11 May 2019

support system utilizing four physiological signals: (i) heart rate variability, (ii) electrodermal activity, (iii) body movement, and (iv) skin temperature. Our sleep quality prediction system tracks the physiological trends by processing the raw time series physiological signals measureable with a cost-effective and easy-to-use wearable device: the E4 wristband [15]. More importantly, our method is applicable for different subjects who sleep in either their own homes or alternative care facilities such as assisted living as opposed to sleep laboratories.

The paper is organized as follows: A general outline about the sleep studies is provided in Section 2. Section 3 describes the proposed method comprising signal preprocessing, feature extraction, feature selection, and classification modules. The performance of the proposed method is considered via different experiments in Section 4. Section 5 compares the proposed method and experimental results with published works in sleep quality analysis. Finally, Section 6 concludes the paper and illustrates the future work.

2. Literature review

Sleep quality has remarkable effects on personal well-being, as sleep encompasses roughly one-third of human life [16]. A wide range of studies explore the relationship between sleep and health issues such as obesity [17], weakened immune system [18], Parkinson's disease [19,20], mortality [5], heart failure [21], and cancer [22]. Additionally, the relationship between sleep and performance of students [23], athletes [24], and workers [25] have been correlated. In this section, we review some of these massive studies from three different aspects: caregivers' sleep problems, sleep architecture, and its measurement tools.

Studies specific to caregiver sleep disturbance have shown that the act of providing care to loved ones with dementia-related illness results in chronic stress, sleep hygiene difficulties, and sleep disturbance [26]. CPWD can experience sleep difficulties because of both overwhelming caregiving responsibilities [9,27] and unpredictable dementia patient behaviors [28]. Inadequate sleep makes it challenging for caregivers in their role of providing care to their loved one with dementia-related illness [29].

To evaluate caregivers' sleep, it is necessary to scrutinize the architecture of sleep (sleep stages) [30], fluctuations of physiological signals during sleep [16], and sleep disorders [31]. Sleep architecture comprises two broad segments: rapid eye movement (REM) and non-REM (NREM) [16]. During NREM, most of the physiological processes in terms of brain activity, heart rate, blood pressure, sympathetic nerve activity, respiration, and body temperature decrease from their usual amount in wakefulness. In contrast, these physiological signals show an increasing pattern in REM rather than in NREM [6]. Since about 80% of sleep time of an adult is NREM sleep [7], most of the sleep studies focus on NREM sleep. NREM sleep can be further divided into four stages (stages 1–4) according to the R-K scoring manual [32]. Due to the co-occurrence and similarity of NREM stage 3 and stage 4, they are considered as one stage of slow wave sleep (SWS) based on the current American Academy of Sleep Medicine (AASM) scoring manual [14], primarily on the basis of electroencephalogram (EEG) criteria.

Among these sleep stages, SWS or deep sleep has special characteristics relevant to sleep quality. For instance, human growth hormone that corresponds to tissue repair is released during the first SWS episode [33,34]. Also, it has been shown that taking sleeping pills is accompanied with increasing length of deep sleep in patients suffering insomnia [35]. Furthermore, patients suffering sleep disorders like obstructive sleep apnea (OSA), periodic limb movement syndrome (PLMS), and insomnia experience less SWS than healthy subjects [36]. Moreover, the reduction of sleep quality in the elderly is accompanied with the loss of deep sleep [37]. Therefore, sleep quality can be evaluated by monitoring long-term SWS.

Various methods have been proposed to scrutinize sleep by classifying the sleep state of participants, i.e., sleep detection [38]. The simplest binary classification is Awake vs. Sleep classification. Common methods for Awake vs. Sleep classification are based on Cole's function [39] or analyzing survey information [40] using wrist actigraphy and mobile applications, respectively. In most cases, researchers need more detailed information to help patients suffering from sleep disorders. As a result, many methods are proposed to investigate the whole sleep stages as awake, stage 1, stage 2, SWS, and REM. The polysomnography (PSG) [41] is the most accurate medical procedure to investigate sleep stages. A modern PSG utilizes an exhaustive list of tests such as EEG, respiratory monitoring, and the recording of sounds/videos. PSG is an expensive and intrusive sleep study, which requires attaching many electrodes to participants under medical surveillance. Other research such as ours seeks to address sleep detection via cheaper and easier-to-use devices. The first option is to reduce interactions between measurement gadgets and participants. Some studies employ contactless measurement tools like microphones [42], video cameras [43], or pillow or mattress accelerometers [44]. To enhance the performance and reliability of measurements, small contact-based devices (e.g. Empatica E4 wristband [15]) can be used [45]. As stated in Refs. [46–48], this easy-to-use device enables us to monitor important physiological signals during sleep while it does not interrupt regular sleep pattern. Also, the E4 wristband provides access to raw time series of physiological signals in contrast to the other popular wristbands (e.g. Fitbit), which offer processed data like step counts and sleep stages.

As mentioned above, there is a strong relation between the length and transition of different sleep stages, especially SWS and the quality of sleep [33–37]. This relation has been widely studied using PSG, for instance Refs. [49–52], demonstrate distinguishability of sleep stages based on the physiological signals. However, using PSG is cost prohibited and intrusively effects sleep quality of participants. Hence, there is a need to explore the sleep quality using cost-effective and unobtrusive methods such as the one proposed in this paper. In the following section, we describe our proposed method to scrutinize sleep quality via the trends of physiological signals which are effortlessly measured by devices like the E4 wristband.

3. The proposed method

Sleep as a mental procedure is highly correlated to physiological signals. This fact is proven under experiments, which involve monitoring sleep stages using PSG [33–37,45,49–66]. Therefore, we propose a novel clinical decision support system to predict the sleep quality based on trends of physiological signals in deep sleep stage. This methodology enables us to estimate sleep quality according to physiological signals which are accessible by wearable devices like E4 wristbands [15]. As shown in Fig. 1, this sleep analysis strategy reuses the sleep investigations which have been mainly conducted using the expensive and accurate PSG. Then, it provides transparent outcomes based on the proven medical evidence, and it is applicable for a wide range of users, especially for individuals experiencing stress or burden like CPWD.

Our computer-assisted clinical decision support system evaluates the effectiveness of sleep according to various physiological signals including heart rate variability (HRV), electrodermal activity (EDA), body movement, and temperature. The proposed system utilizes signals in a time-series manner instead of focusing on certain statistical measurements such as the mean of vital signals. Studies using features such as the mean of heart rate for sleep stage detection [67] can be subjective since the resting heart rate can vary between 60 and 100 beats per minute (BPM). Our system, applicable for a wide variety of subjects, tracks trends of physiological signals, in contrast to considering fixed statistical measurements. Moreover, using time-domain features makes it possible to provide a transparent and explainable clinical decision support system.

As shown in Fig. 2, our sleep analysis system processes the input physiological signals through a four-state procedure, which includes

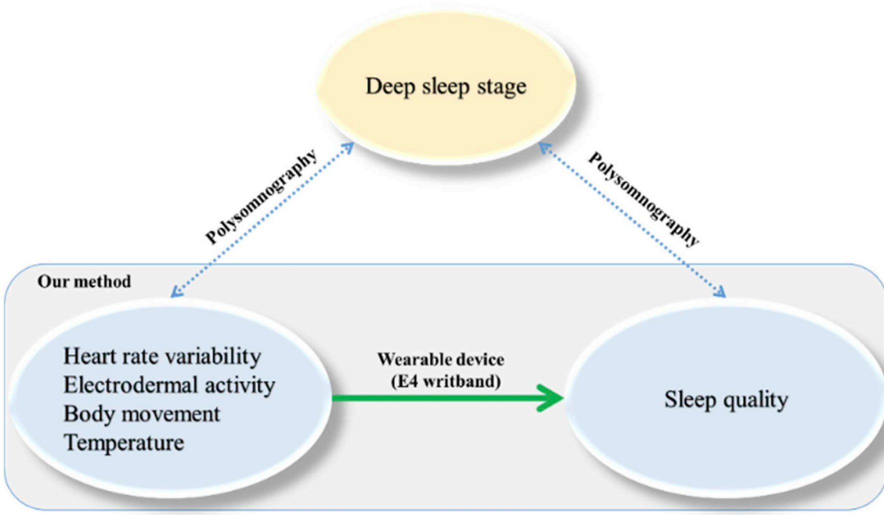


Fig. 1. Assessing the sleep quality using easy-to-use and cost-effective wearable devices based on the firm relation among physiological signal trends in deep sleep and sleep quality.

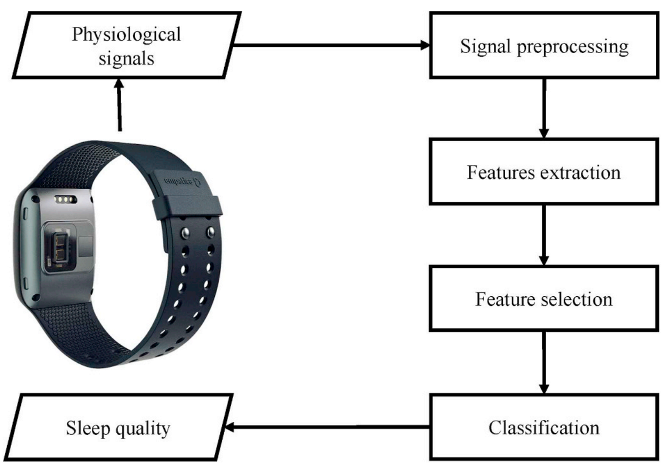


Fig. 2. The outline of the proposed sleep analysis system.

signal preprocessing, feature extraction, feature selection, and classification. The input physiological signals are collected using the E4 wristband during caregiver's sleep period. This FDA-approved device is easy-to-use and unobtrusive and has been recently used in many sleep assessment studies like [46–48]. To investigate the physiological trends in the deep sleep stage, the collected signals need to be preprocessed as described in the following section.

3.1. Signal preprocessing

Through the signal preprocessing state, two important tasks are completed. Since the recorded physiological signals are generally accompanied with noise, removing noises using filtering techniques is essential before gathering any information from the original signals. As distinct signals have different characteristics in SWS each signal is processed in a different way as explained below.

3.1.1. Heart rate variability

As sleep goes from stage 1 to SWS, parasympathetic activity increases, which leads to a continuous reduction in cardiovascular output [56]. This decreasing pattern continues such that heart rate reaches its lowest point at SWS [54]. As heart rate variability (HRV) provides more details than heart rate, we investigate HRV with the aim of predicting the deep sleep stages during sleep.

HRV describes the irregularity among two successive heart beats by measuring the variation in the beat-to-beat interval known as RR interval or inter-beat interval (IBI). HRV can be assessed using different procedures, such as electrocardiography or Photoplethysmography (PPG). Since E4 wristbands provide PPG (also known as blood volume pulse) based on pulse oximetry [68], we focus on calculating IBI from the PPG. As shown in Fig. 3, the IBI in PPG is computed by taking the

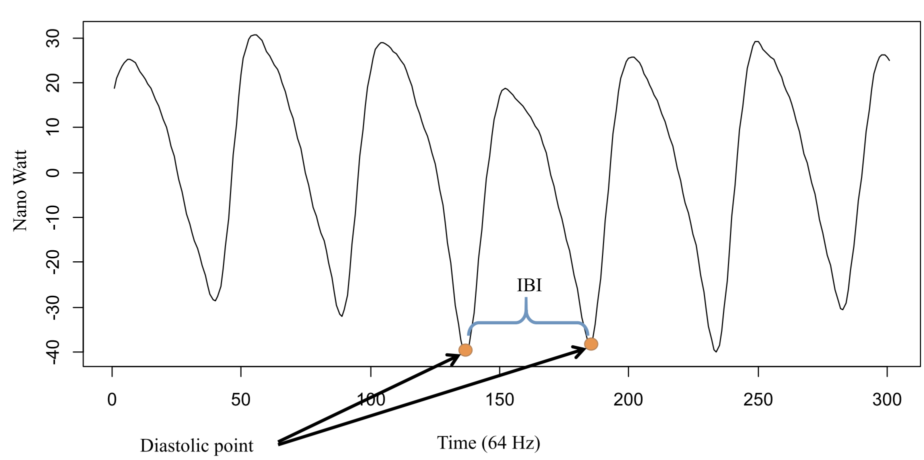


Fig. 3. Computing inter-beat interval (IBI) from Photoplethysmography.

difference between two consecutive diastolic points, which indicates the lowest blood pressure in arteries when heart rests among beats.

After computing HRV as a time series of IBIs from the PPG, HRV is transferred to a new scale to detect the SWS during the sleep. One of the best methods that has been evaluated by PSG is Poincare plot [69]. In Poincare plot, the rRR (correlation coefficient of consecutive RR intervals) is computed in the form of (1) over a window with a 5-min length and its value is assigned to the 20-s interval [69]. Then, the next 20-s window slide and the rRR will be calculated again. The rRR of HRV in time t over a series of RR intervals in IBI can be computed as shown in Equation (1):

$$rRR_t = \frac{\mathrm{E}[(RR_t - \overline{RR}_T)(RR_{t+1} - \overline{RR}_T)]}{\sqrt{\mathrm{E}[(RR_t - \overline{RR}_T)^2]\mathrm{E}[(RR_{t+1} - \overline{RR}_T)^2]}} \tag{1}$$

where T is a sliding window ranging from $1+(20(\lceil \frac{t}{20} \rceil\text{-}1) \times$ (HRV sampling rate)) to $301+(20(\lceil \frac{t}{20} \rceil\text{-}1) \times$ (HRV sampling rate)). Since each rRR value represents 5 × 60 × (HRV sampling rate) number of original RR data points, the rRR time series will be a transformation of HRV, which allows us to track the trend of HRV without the interference of noise data. This transformation marks an interval as SWS if their rRR value is less than 0.1 units below the mean rRR of the initial 4 h for at least 10 min.

3.1.2. Electrodermal activity

The occurrence of intensive fluctuations in brain activities at stage 2 called K-complex leads to a burst of sympathetic activities in NREM [16]. The sympathetic activity can be described by electrical changes of skin surface, called electrodermal activity (EDA) [57]. EDA recorded in overnight PSG has proven that EDA is strongly associated with deep sleep rather than other stages [37]. In fact, various EDA studies have shown that people in SWS experience the highest level of both EDA values [59] and number of local EDA peaks [60]. Also [60], exhibits the stability of EDA properties in SWS by testing different places for wearable device attachment and different threshold selections for EDA peak definition; However, each individual has a different pattern and varied magnitude of EDA during sleep.

We utilize the EDA transformation provided in Ref. [60] to isolate the tendency of EDA from different personal basic level of EDA. In this procedure, the validated algorithm of Cole's function [39] over the accelerometer data has been used to select the EDA signal for further process according to the sleep time. Next, the low-pass finite impulse response filtering with cutoff frequency 0.4 Hz and 32nd order is applied over these EDA time series to eliminate the possible artifacts [70,71]. Following that, the first derivative of EDA provides a map of EDA fluctuations during sleep stages. The higher the intensity level of fluctuations in this map, the deeper the sleep is. The intensity level of EDA fluctuation is defined in terms of the occurrence of peaks which have values higher than 0.01 in a 30 s interval.

3.1.3. Body movement and temperature

Body movement is another easy-to-measure factor that has a strong relation with sleep stages [61]. Short movements appear over all sleep stages, however the frequency of their occurrence in SWS is significantly lower [62]. Also, body motility has a direct relationship with sleep quality since it can disrupt sleep. In severe cases like periodic leg movements, medical assessment is needed to relieve the prolonged tiredness during daytime [63]. Body movement has a more destructive effect on sleep of elderly people who have severe difficulty to maintain a stable sleep with lower SWS as a result of aging [62].

Like body movement and length of SWS, body temperature is also influenced by aging. Indeed, thermoregulation of the newborn remains stable in both SWS and REM, while adults experience a reduction of temperature during their deep sleep [64]. The speed and amount of this temperature reduction has a strong relationship with sleep quality [65]. Hence, subjects who sleep on warm mattresses lose more heat along with experiencing longer SWS as an index of quality of sleep [66].

As explained above, accelerometer and skin temperature are other measurable signals via the E4 wristband relevant to a deep sleep stage and are important factors for sleep quality. Since these two signals are affected by different environmental and mental conditions, there is no direct transformation relating to a deep sleep like HRV and EDA signals. However, they have similar characteristics in that decreasing pattern in these signals correspond to the deeper sleep stages.

To investigate the descending trend for these two signals, we apply two transformation and filtering methods in processing the signals. In the transformation method, we utilize the Poincare plot, identical to the signal processing applied to HRV. This method highlights the part of sleep showing a declining pattern. Also, it focuses on the first half of the sleep period corresponding to NREM in contrast to the second half which respondents REM. For the filtering method, a central moving average filter with the windows length of 1 min is conducted over the accelerometer and temperature signals. Then, the positive/negative sign of slope indicates the increasing/decreasing trend of the time series.

3.2. Feature extraction

In order to predict sleep quality, quantitative features are computed based on processed signals. Since different signals have different characteristics, they are processed using distinct methods. In this section, we describe extracting varied features from processed signals using Poincare plot, filtering methods, Cole's function, and first

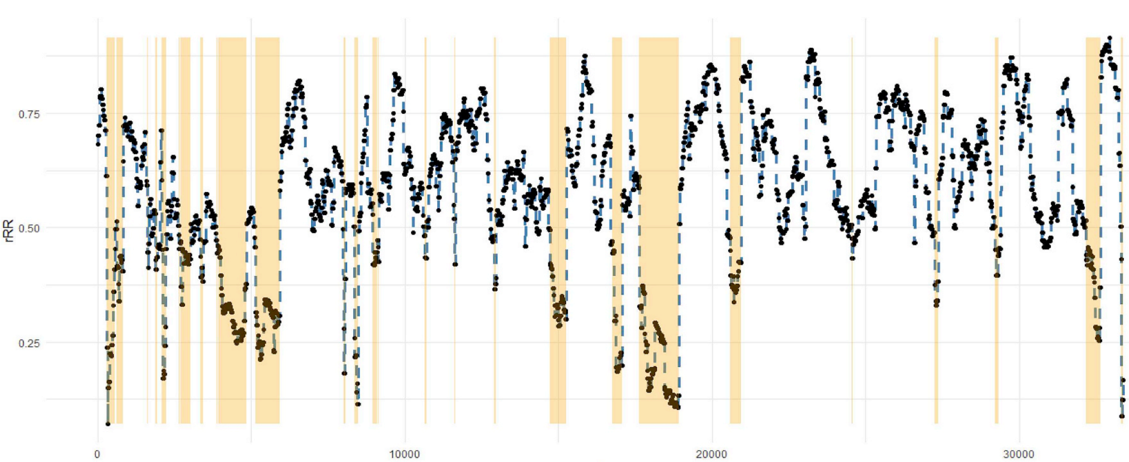


Fig. 4. Prediction of SWS occurrence based on the Poincare plot.

derivative.

As shown in Fig. 4, Poincare plot of rRR contains some labeled segments assigned as SWS. The number of occurrences of these segments is considered as the SWS time feature. Also, the accumulated sum of the length of these segments forms the SWS length feature. The SWS time and length feature are calculated for HRV, body movement, and temperature processed signals.

Body movement and temperature are also processed through the filtering method. The result of this processed signals highlights portions representing descending trend of temperature and body movements. Since people experience different length of sleep each night, we consider the percentage of sleep time when the time series of temperature and body movements have decreasing trends.

According to Cole's function, we can segregate periods when people are awake from when they sleep. These labels enable us to access sufficient features in deep sleep. In other words, sleep architecture is a cyclic procedure and staying asleep for a longer time enhances the chance of experiencing more and longer deep sleep stages. On the contrary, the more a person wakes up during the sleep, the less the person experiences deep sleep. Therefore, we use amount of sleep, amount of awake, and times of awoken as features. Furthermore, we calculate the sleep efficiency (SE) feature as the proportion of sleep time to the time a person tries to sleep [72]. The SE considers the side effect of spending too much time in bed trying to sleep.

The first derivative is another signal processing method widely used to track the fluctuation of EDA signals. Fluctuations are reported in terms of EDA peak per epoch. An epoch is defined as a 30-s section of EDA having at least one EDA peak. The highest number of EDA peaks occurs in SWS rather than other sleep stages [60]. Then, we compute the number of epochs during a sleep period as the Epoch peak counter feature. Moreover, higher level of EDA in the deep sleep stage [59] results in higher number of peaks in each epoch. This trend is shown by the Epoch peak feature which provides the mean of number of peaks over all epochs. Additionally, we introduce the Epoch capacity feature as the ratio of the number of epochs in a sleep period to the total possible epochs. High Epoch capacity illustrates high level of EDA activity during the sleep and a long deep sleep period.

Another well-known term in EDA study relevant to sleep quality is EDA storm, the consecutive EDA peak epochs. Since EDA storms appear most during deep sleep stage [73], we report EDA storm by five features: (i) storm peak, (ii) largest storm, (iii) mean storm, (iv) times storm, and (v) total length of storm. The storm peak expresses the percentage of peaks which occur in the storms as continuous active EDA periods. Since long storms with a large number of EDA peaks are more likely to occur in SWS [60], the largest storm and mean storm features are calculated in this study. These two features represent the number of epochs constructing the largest storm and the average number of epochs comprising each EDA storm, respectively. Finally, the more occurrences of EDA storms lead to a higher probability of experiencing deep sleep and it is worth mentioning that most of the night sleeps comprise at least two storms [73].

3.3. Feature selection and classification

In feature extraction, we obtain twenty features from different physiological signals as listed in appendix A. Selecting the right subset of these features have several advantages, such as faster training of the classification algorithm, reducing the complexity of models, enhancing the interpretability of models, and improving the accuracy of final results [74].

Among feature selection techniques, filtering methods like Pearson' correlation or analysis of variance are computationally simple and fast; however, they neglect feature dependencies since they consider each feature separately [75]. To address this, we employ a wrapper method called the recursive feature elimination (RFE) [76] which selects the optimal combination of features while maximizing the performance of the chosen classifiers. To handle nonlinear relation between features, as well as reducing the risk of overfitting, we use a version of RFE called RFECV model which incorporates resampling of k-fold cross-validation strategy [77].

In RFECV, records are arbitrarily divided into ten separate sections. In each iteration, nine sections form a group of records used to train classifiers, and the remaining is utilized to test the learning process. During the training step for each iteration, the selected classifier is trained using all features. Then, features are ranked by their importance to the classifier. Following this, the least important feature is removed and the classifier is fitted to training data with remaining features. And, again the importance of each feature is calculated. This process is repeated until no further feature remains. The average of learning rates over the models trained using the same number of features illustrates the model performance corresponding to different number of features. After the optimal subset size is determined by the highest performance, classifiers which are trained with optimal subset size of features are considered to select final features. Then, the consensus ranking over the list of feature importance describes the final features. Finally, the trained model will be fitted using the optimal subset of features in the training set.

After feature selection, we try to select the best classifier in terms of performance, transparency, and computational complexity. In essence, choosing a proper classifier for a special framework is a tradeoff between performance and computational complexity. Usually more complex classifiers lead to higher performances since they can investigate more complicated relationships among features and conversely, simple classifiers tend to perform poorer than their complex counterparts [78]. It is necessary to choose a classifier which has acceptable complexity with regard to performance. Hence, we utilize three classifiers including naïve Bayes, random forest, and bagged tree with varying complexity.

Naïve Bayes is a simple conditional probabilistic model. It tries to predict the probability of different states of the outcome variable using their posterior probabilities. Since it does not model dependencies between input parameters, it is highly scalable, fast, and easy to implement. However, it is possible that predictors have interdependencies. Thus, we use bagged tree and random forest classifiers as well. These two ensemble classifiers are able to provide more accurate prediction by making decisions based on multiple decision trees. In the bagged tree method, a combined decision tree model is obtained from a single training data set by repeatedly using several bootstrapped subsets of the data and averaging models [79]. Random forest enhances the bagging strategy by bootstrap sampling the predictor variables at each splitting criterion as well [80].

4. Results

In this IRB-approved study, we invited eight caregivers to participate in our experiments for a period of two weeks each. However, using the wearable device E4, we were only able to record 100 nights of sleep (as opposed to 112 nights). The main reason for the missing nights was that some of the caregivers accidently forgot to wear the E4 device, or failed to recall filling the corresponding morning survey which meant we did not have the labels to go with the collected E4 data. The cohort included six female and two male caregivers. Two of the caregivers were adult children of dementia patients and six were patient spouses. Dementia-related illness results in the patients suffering from multiple symptoms simultaneously. As described in Table 1, most caregivers reported that their loved one experienced depression, communication difficulty, or agitation. Loved ones Dementia-related symptoms caused the majority of caregivers' to experience median and high level of caregiver.

To administer the study, a team consisting of a social worker and a computer scientist conducted three home visits, distributed over the two weeks. During each visit, the social worker would conduct a

Table 1
Background information of CPWD.

# Participants	1	2	3	4	5	6	7	8
Gender (Female, Male)	F	F	M	F	F	F	M	F
Relationship (Children, Spouse)	C	S	S	S	S	C	S	S
Burden level (1–10)	6	4	4	6	8	6	7	3
Level of dementia in loved one (Mild, Moderate, Severe)	Mi	Mo	Mo	Mo	Mo	S	Mo	Mo
Comorbidity symptoms in loved one								
Wandering		*		*	*		*	
Incontinence				*	*	*		
Inappropriate social behavior		*		*	*	*	*	
Inappropriate sexual behavior		*		*			*	
Aggression		*		*	*	*	*	
Communication difficulties	*	*		*	*	*		
Depression	*	*	*	*	*	*	*	
Agitation		*	*	*	*	*	*	

qualitative interview with the caregivers. The computer scientist would describe the wearable device use and the Android tablet application, which contained the daily use caregiver sleep survey (DUCSS) [81] (Appendix B). The app on the mobile tablet uses the SQLite database to store the responses to the sleep survey questions. The app did not save any data with respect to the caregivers' identity to address privacy concerns. The caregivers were requested to wear the wristband approximately 15 min before sleep and remove the device immediately after waking up.

Other available sleep surveys [82–85] are not intended for frequent use or caregivers' sleep assessment. To proactively recognize sleep difficulties in CPWD, DUCSS was developed using a mixed method design utilizing 4 available sleep surveys: the Pittsburgh sleep quality index (PSQI) [82], functional outcomes of sleep questionnaire (FOSQ) [83], Calgary sleep apnea quality of life index (SAQLI) [84], and the RAND 36-item health survey (SF-36) [85]. A qualitative focus group of geriatric experts refined and streamlined multiple existing sleep surveys questions [82–85], while adding additional items. The DUCSS was validated using Rasch statistical model [86] on a cohort of (N = 24) participants.

In the following section, the CPWD sleep data set are analyzed from two aspects: (i) overall sleep quality (Question 9 from survey) and (ii) feeling refreshed after waking (Question 10 from survey) in Subsection 3.1. The efficiency of our proposed decision support system as well as the feature selection process has been described in Subsection 3.2.

4.1. Data set description

To avoid bias in measurement of sleep quality, the sleep survey evaluates sleep from different aspects. As shown in the following table, questions 9 and 10 of the survey attempt to appraise the effectiveness of sleep through "overall quality of sleep" and "feeling rested", respectively.

These two categorical variables represent different aspects of the sleep quality of caregivers. The "sleep quality" illustrates the overall satisfaction of caregivers from their sleep which has a direct effect on providing care for dementia patients [13]. However, the stressful task of providing care for dementia patients can have a destructive effect in the process of fatigue recovery [87]. Therefore, we investigate fatigue recovery through the "feeling rested" variable. Studying the interactions among them can provide a holistic view about the sleep quality of CPWD. The interaction can be expressed using the contingency table in Table 3 where each cell counts joint co-occurrence of two variables in certain circumstances.

Using an efficient visualization technique for the contingency table called the mosaic plot [88] can provide a more comprehensive outline. The mosaic plot in Fig. 5 denotes each cell of the contingency table in a rectangular shape such that the area of each rectangle represents the value in the corresponding cell. For instance, the rectangular area corresponding to the caregiver sleep with 'bad quality' and 'feeling rested' is zero which is equal to the corresponding value in the contingency table. Also, applying Pearson's chi-squared test over the contingency table measures the departure of each two random variables from independence in different conditions. Pearson residuals depict this dependency and are presented in the mosaic plot via different colors. The blue color illustrates there are adequate observations to reject the null hypothesis of chi-square test (independence). In contrast, red

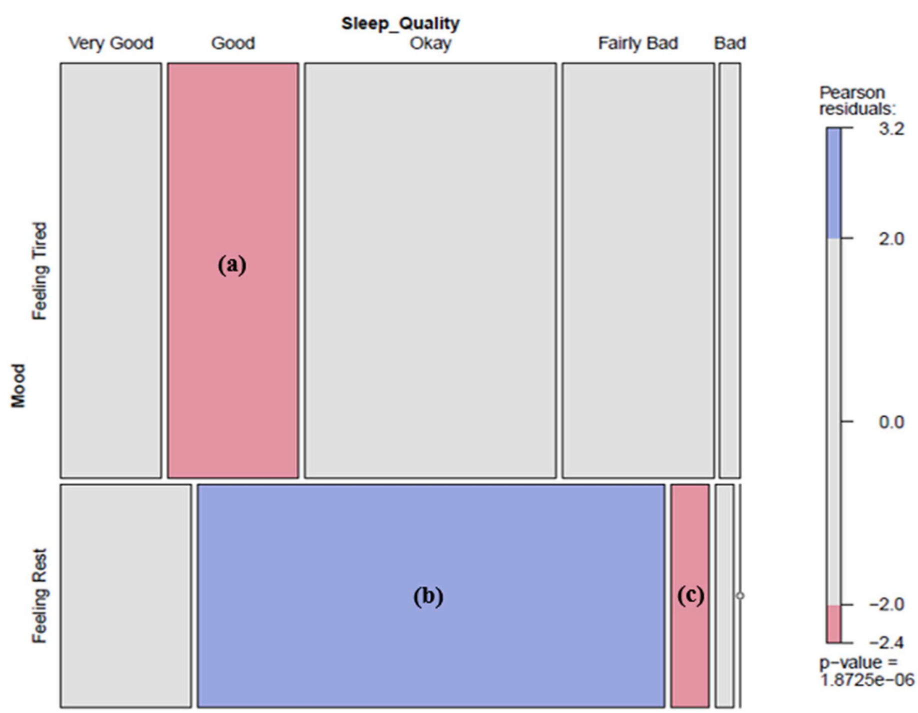


Fig. 5. The relation of Sleep Quality and the mood of caregivers using a mosaic plot. The bar on the right indicates the Pearson residuals by color.

Table 2
Questions 9 and 10 of sleep survey.

#	Question
9	Please select one. I would rate the overall quality of my sleep as a.) Very good b.) Good c.) Okay d.) Fairly bad e.) bad
10	Please select one. Did you wake up feeling rested? a.) No b.) Yes

means there are fewer observations to reject the null hypothesis; the two random variables are independent in that condition.

The above mosaic plot shows the distribution of caregivers' mood versus their sleep quality. It also shows the marginal distribution of sleep quality (the width of the bars). Similarly, the length of cells with respect to mood attributes indicates 65% of caregivers still feel tired after sleep. Almost all of the caregivers who felt rested after sleep reported good or very good sleep quality. This denotes the strong relation among sleep quality and feeling refreshed. On the other hand, 17% of caregiver sleep records are associated with feeling tired along with bad and fairly bad quality of sleep.

Also as explained before, the mosaic plot provides inferences about the population defined by Pearson residuals in terms of different colors which measure the departure of each cell from independence. For example, the blue region (b) in Fig. 5 indicates that there is a high probability for a caregiver to both feel rested, as well as have an "Ok" sleep the night before. In contrast, the red regions (a) indicates a strong lack of probability of the caregiver feeling rested after sleep while they report "Ok" sleep quality. The gray rectangles convey there is not enough evidence that caregivers who feel tired after sleep experienced bad or fairly bad quality of sleep.

Table 3 and its mosaic plot show that the data set labeled using binary caregiver tiredness perspective leads to a roughly balanced classification problem. However, using a coarser labeling method like sleep quality results in an imbalanced data set. The distribution of sleep quality displayed in Table 2 demonstrates the dataset is highly biased. Only 18 records report the fairly bad and bad sleep quality. While over 1 out of 4 records are assigned to "Ok", the "Good" and "Very Good" categories comprise 38% and 17% of the data, respectively.

In addition to general sleep characteristics, the sleep dataset also contains some special attributes. For instance, caregivers are usually deprived from sleep caused by interruption. Various conditions can cause these interruptions. Special needs of dementia patients to use the bathroom, or their behaviors, like shouting, coughing, or wandering during the night are some examples. Also, family caregivers are usually older adults themselves and more sensitive to sounds and temperature changes. This is a common pattern among caregivers' sleep since 98% of the records reported a sleep break. These sleep interruptions can last longer than 1 h. This reduces the chance of having deep sleep among this population. It may also be the reason why 65% of records reported feeling tired after sleep. However, 82% of them expressed the "Ok", "Good", or "Very Good" sleep quality, possibly due to the fact that the presence of disrupted sleep is a routine among caregivers [9].

Table 3
Contingency table of Sleep quality and rest feeling after sleep in CPWD.

		Sleep quality					
		Very good	Good	Okay	Fairly bad	Bad	Sum
Mood	Feeling tired	10	13	25	15	2	65
	Feeling rested	7	25	2	1	0	35
	Sum	17	38	27	16	2	100

Table 3 indicates only two records are assigned to bad sleep quality. This low number of bad sleep quality can be related to caregivers becoming accustomed to the difficulties of dealing with dementia patients [9]. A low number of samples reduces the ability of any machine learning method to learn the patterns of physiological signals. Therefore, we combine two categories of bad and fairly bad as a single category of bad sleep quality for further processing. However, the issue of an imbalanced data set remains since the number of records in the class of good is more than two times the number of records in bad and very good categories. To address this problem, various techniques such as resampling [89] or cost sensitive classifiers [90] have been proposed. Resampling methods are less sensitive to outliers than other techniques. Furthermore, they make no assumptions about the distribution of records which enable them to be used for any classification problem. To balance the data set in terms of sleep quality, we apply the resampling method which randomly samples by replacing the minority classes to be the same size as the majority class.

4.2. Performance evaluation

Every decision support system should be interpretable, transparent, and utilize medical background knowledge in their analysis. Thus, we evaluate the performance of the proposed method from different aspects including selecting predictors, reliability in the decision making, and integration of background knowledge.

Referring to the selected predictors, distinct classifiers interpret the relationship among training samples differently. The recursive feature elimination (RFE) selection methodology leads to the highest performance of classifiers since it supplies each classifier with the most efficient features. However, this methodology should incorporate the feature selection variations using a proper resampling method like 10-fold cross-validation [77]. Therefore, the 10-fold RFE cross-validation (RFECV) enables us to find useful predictors in different models.

The overall accuracy of classifiers is highly sensitive to the number of input features. This sensitivity may be the result of nonlinear relations among predictors and outcomes. As shown in Fig. 6, the classifiers' performance does not follow a linear increasing trend with the growth of number of features selected by RFECV. RFECV utilizes the importance of features in each classifier to reduce the state space from 20! (20 factorial) possible subsets of features to only 20 cases. In such a way, regarding the random forest and bagged tree gained the highest performance for feeling refreshed and sleep quality, respectively. The best model for feeling refreshed utilizes 12 features while the sleep quality one employs 10 features.

To evaluate reliability of best predicted models, we need to scrutinize performance of classifiers in terms of sensitivity, specificity, precision, and accuracy. These metrics provide valuable information about precise patients' diagnosis in medical assistant systems. Sensitivity computes the probability of detecting the desired condition in medical applications, whereas specificity is the ability of the assistant system to correctly identify those who are not affected by that condition. Precision is another important metric which reports the fraction of truly detected records among all detected records for the desired condition. This metric is important since faulty detections impose series of medical observations.

By evaluating sleep records referring to feeling "refreshed" or "tired" after sleep, we consider it as a binary class problem. On the other hand, distinguishing sleep qualities can be a multi-class problem. In this case, we employ one-vs-all evaluation methodology [91]. Finally, the average of all values per metric is reported as shown in Table 4.

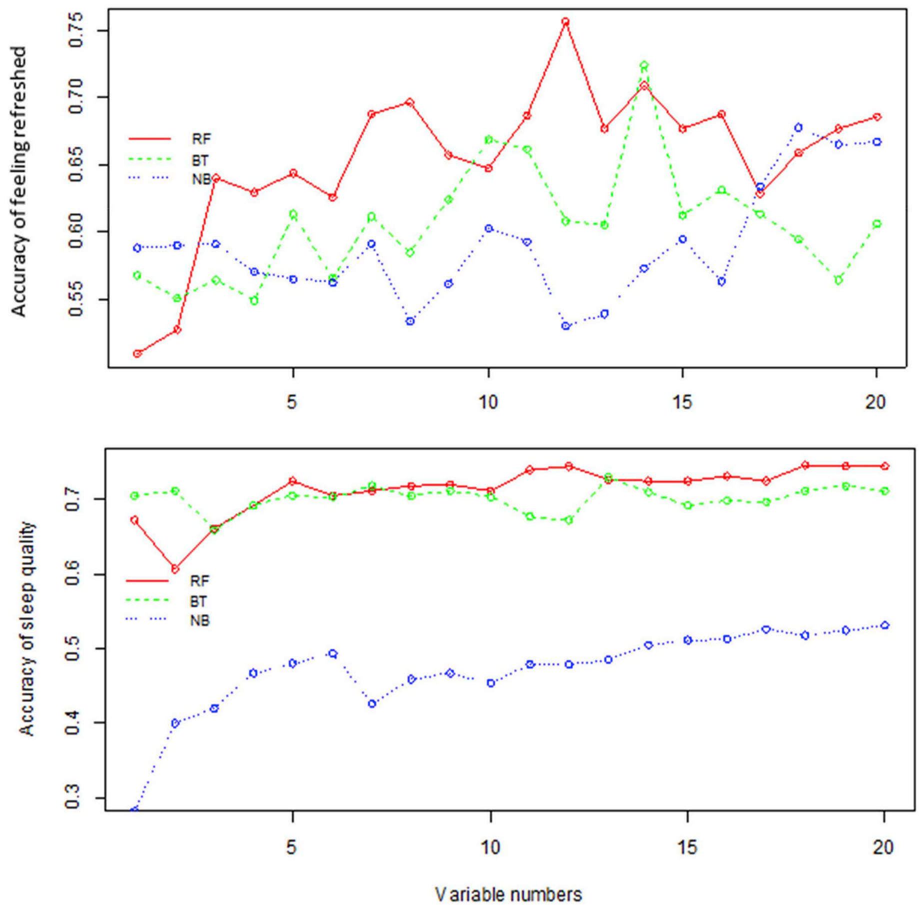


Fig. 6. The resampling results for the candidate subset sizes evaluated during the recursive feature elimination (RFE) process over random forest (RF), Bagged tree (BT), and naïve bays (NB) classifiers.

Table 4 demonstrates that random forest outperforms in both labeling approaches. Also, the bagged tree and the naïve Bayes are the second and third best classifiers according to performance, respectively. The order of classifiers' performance reveals there is a nonlinear relation among features such that more complex classifiers like random forest gain higher performances. Also, random forest obtains high specificity which means the proposed method can detect records of tiredness in caregivers with high probability. As a result, our sleep analysis system can be a reliable system to alarm the caregivers about the potential increase in caregiving-related stress or burnout.

The selected features by RFECV play different roles in the created random forest models. The tree-based methods try to make a model by repeatedly splitting records based on the provided features. The order of these features is dedicated by the Gini impurity criterion (2) such that the two descendent groups should have the least possible impurity.

Table 4
Comparison of the performance of classifiers with the best possible features.

Labeling	Metric	Classifier		
		Random forest	Bagged tree	Naïve Bayes
Feeling refreshed	Accuracy	**0.73**	0.69	0.61
	Precision	**0.75**	0.58	0.42
	Sensitivity	0.34	**0.42**	0.34
	Specificity	**0.94**	0.83	0.75
Sleep quality	Accuracy	**0.75**	0.73	0.52
	Precision	**0.75**	0.74	0.52
	Sensitivity	**0.75**	0.70	0.52
	Specificity	**0.92**	0.91	0.84

$$Gini\ Index = 1 - \sum_{i} (p(i))^2 \tag{2}$$

where $P(i)$ is the observed fraction of samples in the node belonging to the class i. The Gini index is an important impurity criterion which can be employed to investigate what extent each feature contributes to the classification process [92]. In this process, the importance for each feature computes by mean decrease in Gini coefficient. It is a fraction with a numerator which is the sum over the number of splits across all trees that include the feature and the denominator of the number of samples it splits.

Fig. 7 illustrates feeling refreshed model selects a lower number of features for making decision rather than the sleep quality model. This fact is as a result of, the feeling refreshed model having lower variability of possible cases rather than sleep quality. Also, sleep efficiency is an important factor in both the sleep quality and feeling refreshed models. These results coincide with previous results relevant to CPWD sleep characteristics showing they experienced fewer total sleep time while taking longer to fall asleep [27].

Moreover, features relevant to temperature have substantial effect in making decision in both models, especially in sleep quality. In fact, the nocturnal temperature dysregulation as an age-related sleep disturbance contributes to fragmentation of sleep which is a common predisposing factor for sleep complaints in caregivers [9]. Since importance features provided by models agree with well-established CPWD sleep studies, the proposed models can be applicable for CPWD [93].

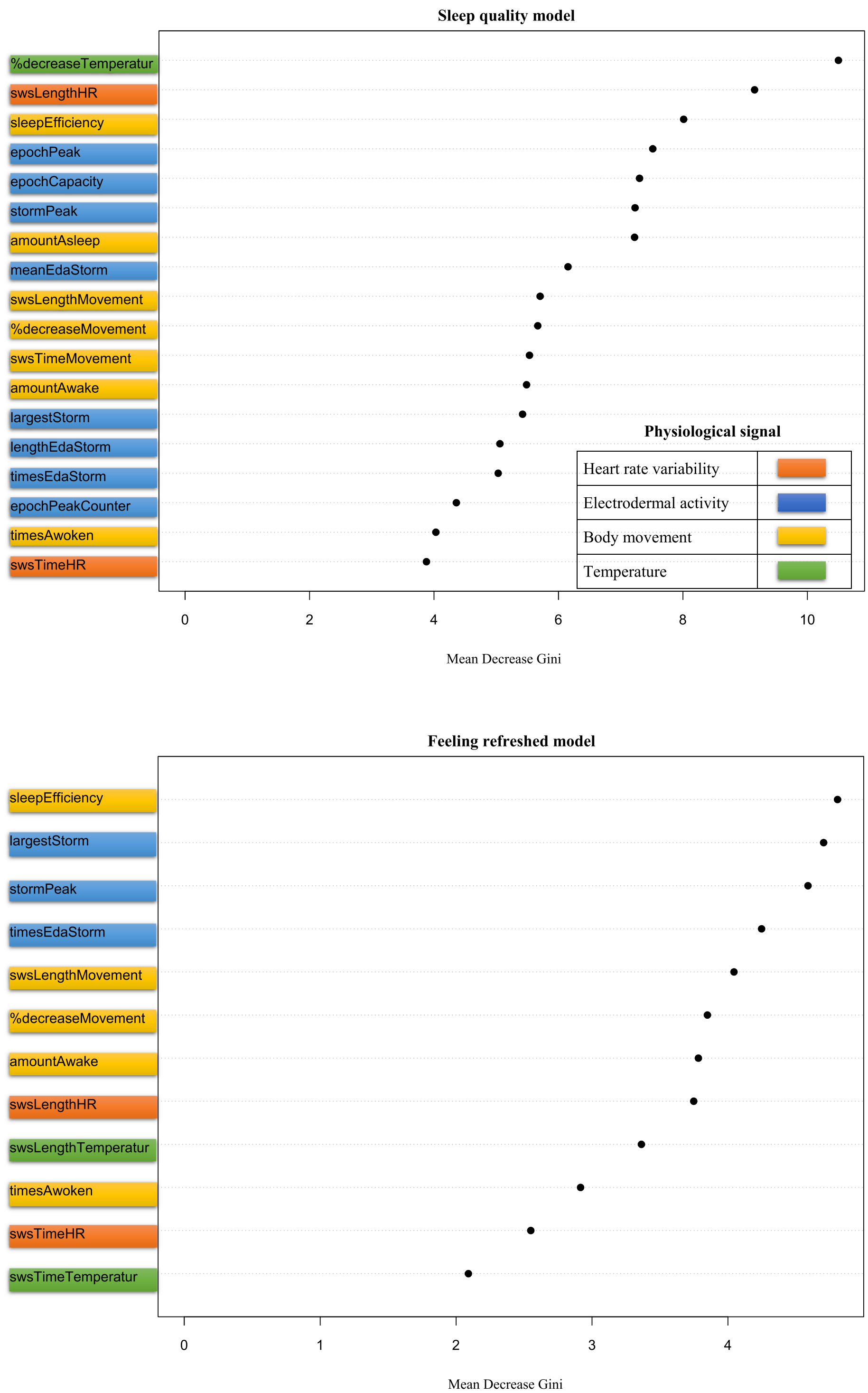


Fig. 7. Feature Importance in the best models.

5. Discussion

A review in Ref. [94] evaluated more than 100 sleep quality analysis papers published in scientific conferences and journals such as, *Sleep Medicine*, *Sensors*, and *IEEE Transactions on Biomedical Engineering Sleep*. This review specified the need for a robust system that applies ensemble machine learning methods to multiple sensors for sleep quality measurement. We tried to bridge this verified gap by proposing a methodology that makes decision based on four raw physiological signals measured from different sensors within an E4 wristband. Our proposed method utilizes ensemble methods, like random forest and bagged tree, to process the extracted features from heart rate variability, electrodermal activity, body movement (accelerometer), and skin temperature.

In some previous sleep studies [26–28], both subjective and objective sleep quality parameters, e.g. total sleep time, sleep efficiency, and mean range of sleep stages, in caregivers' sleep and its difference from the sleep of non-caregivers have been investigated. However, none of them endeavor to predict the sleep quality of caregivers. To the best of our knowledge, this is the first work that predicts the sleep quality specifically in caregivers.

The majority of previous works [23–28] investigated sleep quality using a monthly survey of PSQI [82]. However, in this study we used the daily survey of DUCSS [81] for assessing each individual's sleep quality every night as a step toward early detection and early intervention. Studies, such as [95], used sleep measurements, like sleep efficiency, as a proxy for sleep quality without actually measuring sleep quality through self-reported means. Our evaluation discovered that sleep efficiency is one of the most important factors in predicting sleep quality, which is consistent with the previous studies [6,7].

Regarding predicting sleep quality using wearable devices, two works of [23,95] studied estimation of sleep quality using ActiGraph Gt3X+ and Q-sensor, respectively. These studies predicted sleep quality of high school [23] and undergraduate students [95], respectively, by asking them to wear the device for a week [23] and one month [95] consecutive days and nights, respectively. However, we asked our participants to wear the device only during the night. As a result, our methodology builds a prediction based on the data collected only at night, while they [23,95] utilized the work burden (daily activity) of the participants as an input for their systems. Both studies reported accuracy scores of 89% without discussing the role of individual features.

Our methodology provides interpretation for the utilized physiological signals, the extracted features, and their importance in sleep quality prediction. The evaluation results demonstrate that %decreaseTemperature (Temperature), swsLengthHR (HRV), and sleep efficiency (Body movement) play the most important roles in predicting sleep quality. On the other hand, sleep efficiency (Body movement), largest storm (Electrodermal activity), and storm peaks (Electrodermal activity) are the most efficient features in predicting feeling refreshed/tired after sleep. These outcomes coincide with previous findings that reveal sleep efficiency and skin temperature are two common predisposing factors for sleep complains in caregivers [9,27].

6. Conclusion and future work

Evaluation of sleep quality and feeling refreshed for agents with important tasks such as caregivers are necessary. This paper proposes a new method using the easy-to-use E4 wristband instead of the expensive, cumbersome PSG for sleep evaluation. It provides a reliable clinical decision system which predicts sleep quality and feeling refreshed by tracking heart rate variability, EDA, skin temperature, and body movement signals. The accuracy of 75% in detecting the quality of sleep demonstrates the capability of the proposed method.

This work can be extended using various approaches. Investigating the quality of daily short sleep can provide complementary information for fatigue recovery in CPWD. Also, using new signal processing methods can provides more informative features from the physiological signals. Finally, we also plan on creating a personalized sleep evaluation systems to enhance the performance of sleep evaluation by embedding personal behaviors in the system.

Conflicts of interest

None Declared.

Acknowledgment

The authors thank JoAnna Evans, Morgan Freeman, and Garrett Goodman for their valuable work as the social work and computer scientist teams involved in data collection. We also thank Alzheimer's Association Miami Valley Chapter for allowing us to showcase our study, and present our research findings. Lastly, we are grateful to all the caregivers who participated in our study.

This paper is based on work supported by the National Institutes of Health under Grant no K01 LM012439. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NIH.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.05.010.

Appendix A. The list of features

Number	Abbreviation Name	Feature description	Physiological signal
1	swsLengthHR	The predicted fraction of sleep belonging to the deep sleep portion based on the heart signal	Heart rate variability
2	swsTimeHR	The number of transitions to the deep sleep stage according to the heart signal	
3	epochCapacity	The ratio of the number of the epochs in a sleep to the total possible epochs	Electrodermal activity
4	epochPeak	The mean number of peaks in all epochs	
5	epochPeakCounter	The number of epochs during a sleep	
6	stormPeak	The percentage of peaks which occur in the storms	
7	largestStorm	The number of epochs that construct the largest storm	
8	timesEdaStorm	The number of distinct storms	
9	meanEdaStorm	The average number of epochs comprising each EDA storm	
10	lengthEdaStorm	The number of whole epochs shaping the storms	
11	timesAwoken	The number of times people awake from sleep	Body movement
12	sleepEfficiency	The proportion of sleep time to the time a person tries to sleep	
13	amountAwake	The length of night time subjects are awake	
14	amountAsleep	The length of time subjects are asleep	
15	swsTimeMovement	The predicted fraction of sleep belonging to the deep sleep portion based on body movement	
16	swsLengthMovement	The number of transitions to the deep sleep stage according to the body movement	
17	%decreaseMovement	The percentage of sleep time in which the body movement has a decreasing pattern	
18	swsLengthTemperature	The predicted fraction of sleep belonging to the deep sleep portion based on the skin temperature	Temperature
19	swsTimeTemperature	The number of transitions to the deep sleep stage according to the skin temperature	
20	%decreaseTemperature	The percentage of sleep time which participants experienced temperature reduction	

Appendix B. Sleep quality survey

1.) What time did you go to bed last night? E.g. 11:30
2.) Please select one. Approximately how long did it take you to fall asleep?
  a.) Less than 30 min
  b.) Between thirty and 60 min
  c.) More than 60 min
3.) What time did you get up this morning? E.g. 9:00
4.) How many hours of sleep did you get last night? E.g. 8
5.) How many hours total were you in bed? E.g. 9.5
6.) Please select one. Did you take a sleep aid last night?
  a.) No
  b.) Yes, an over the counter sleep aid
  c.) Yes, a prescription sleep aid
7.) Please list any occurrences that you can recall from last night that disturbed your sleep. Please also include the time when they happened. Examples include feeling too hot, or too cold, coughing, having nightmares, getting up to use the restroom, etc.
8.) Please select one. If you woke up during the night, approximately how long did it take you to fall back to sleep?
  a.) Less than 30 min
  b.) Between thirty and 60 min
  c.) More than 60 min
  d.) Not applicable
  9.) Please select one. I would rate the overall quality of my sleep as
  a.) Very good
  b.) Fairly good
  c.) Okay
  d.) Fairly bad
  e.) Bad
10.) Please select one. Did you wake up feeling rested?
  a.) No
  b.) Yes
11.) Please rate how much sleepiness interfered with your ability to complete tasks from one (not at all) to five (very much).
  1, 2, 3, 4, 5
12.) Please select one. Did lack of sleep interfere with your personal relationships?
  a.) No
  b.) Yes
13.) Please select one. Did you find it more difficult to bathe, dress, or eat because you felt tired?
  a.) No
  b.) Yes
14.) Please select one. Did sleepiness interfere with your ability to do housework, arrange transportation, or talk on the telephone?
  a.) No
  b.) Yes
15.) Please select one. Please rate your overall quality of your day yesterday.
  a.) Very bad
  b.) Bad
  c.) Okay
  d.) Good
  e.) Very good
16.) Please select one. Did you have difficulties at work, school, or at a volunteer position because of your sleepiness?
  a.) No
  b.) Yes
17.) Please select one. Did you have difficulty performing your caregiving duties because you felt tired?
  a.) No
  b.) Yes
18.) Please check all that apply. Did you feel more tired in:
  a.) The morning
  b.) Afternoon
  c.) Evening
19.) Please rate how tired you felt on a scale from one to five, where one is not at all tired and five is extremely tired.
20.) Please select one. How would you rate your energy level yesterday?
  1, 2, 3, 4, 5
21.) Please select one. Did you find yourself feeling sad, angry, or overwhelmed yesterday because you were tired?
  a.) No
  b.) Yes
22.) Please select one. Did you experience any generalized aching, joint pain, or headaches because you felt tired?
  a.) No
  b.) Yes

References

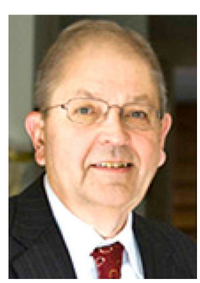
[1] J.S. Ruggiero, N.S. Redeker, Effects of napping on sleepiness and sleep-related performance deficits in night-shift workers: a systematic review, Biol. Res. Nurs. 16 (2014) 134–142.
[2] A.J. Phillips, W.M. Clerx, C.S. O'Brien, A. Sano, L.K. Barger, R.W. Picard, S.W. Lockley, E.B. Klerman, C.A. Czeisler, Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing, Sci. Rep. 7 (2017) 3216.
[3] H. Yoon, S.H. Hwang, J.-W. Choi, Y.J. Lee, D.-U. Jeong, K.S. Park, Slow-wave sleep estimation for healthy subjects and OSA patients using R–R intervals, IEEE J. Biomed. Health Informatic. 22 (2018) 119–128.
[4] S. Dobing, N. Frolova, F. McAlister, J. Ringrose, Sleep quality and factors influencing self-reported sleep duration and quality in the general internal medicine inpatient population, PLoS One 11 (2016) e0156735.
[5] H.A. Beydoun, M.A. Beydoun, X. Chen, J.J. Chang, A.A. Gamaldo, S.M. Eid, A.B. Zonderman, Sex and age differences in the associations between sleep behaviors and all-cause mortality in older adults: results from the National Health and Nutrition Examination Surveys, Sleep Med. 36 (2017) 141–151.
[6] P. Montagna, S. Chokroverty, Sleep Disorders, Elsevier, 2012.
[7] S. Chokroverty, Overview of normal sleep, Sleep Disorders Medicine, Springer, 2017, pp. 5–27.
[8] R. Schulz, J. Eden, National Academies of Sciences, Medicine, Family Caregiving Roles and Impacts, (2016).
[9] S.M. McCurry, R.G. Logsdon, L. Teri, M.V. Vitiello, Sleep disturbances in caregivers of persons with dementia: contributing factors and treatment implications, Sleep Med. Rev. 11 (2007) 143–153.
[10] Caregiver Statistics, Demographics | Family Caregiver Alliance (n.d.), https://www.caregiver.org/caregiver-statistics-demographics , Accessed date: 27 August 2018.
[11] M.B. Wasilewski, J.N. Stinson, J.I. Cameron, Web-based health interventions for family caregivers of elderly individuals: a scoping review, Int. J. Med. Inform. 103 (2017) 109–138.
[12] S.H. Belle, L. Burgio, R. Burns, D. Coon, S.J. Czaja, D. Gallagher-Thompson, L.N. Gitlin, J. Klinger, K.M. Koepke, C.C. Lee, Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: a randomized, controlled trial, Ann. Intern. Med. 145 (2006) 727–738.
[13] S.M. McCurry, L.E. Gibbons, R.G. Logsdon, M.V. Vitiello, L. Teri, Insomnia in caregivers of persons with dementia: who is at risk and what can be done about it? Sleep Med. Clin. 4 (2009) 519–526.
[14] I. Conrad, The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, American Academy of Sleep Medicine, Westchester, IL, 2007.
[15] Real-time physiological signals | E4 EDA/GSR sensor, Empatica (n.d.), https://www.empatica.com/research/e4 , Accessed date: 7 December 2017.
[16] B.M. Altevogt, H.R. Colten, Sleep Disorders and Sleep Deprivation: an Unmet Public Health Problem, National Academies Press, 2006.
[17] S. Reutrakul, E. Van Cauter, Sleep influences on obesity, insulin resistance, and risk of type 2 diabetes, Metabolism 84 (2018) 56–66.
[18] N. Asif, R. Iqbal, C.F. Nazir, Human immune system during sleep, Am. J. Clin. Exp. Immunol. 6 (2017) 92.
[19] R. Prashanth, S.D. Roy, P.K. Mandal, S. Ghosh, High-accuracy detection of early Parkinson's disease through multimodal features and machine learning, Int. J. Med. Inform. 90 (2016) 13–21.
[20] R. Prashanth, S.D. Roy, Early detection of Parkinson's disease through patient questionnaire and predictive modelling, Int. J. Med. Inform. 119 (2018) 75–87.
[21] C.L. Gardner, M.C. Flanagan, C. Franklin, C. John-Swayers, S. Walsh-Pouch, F.J. Bryant, C.A. Romano, S. Gibbons, M. De Jong, A. Hoang, Electronic physiologic and subjective data acquisition in home-dwelling heart failure patients: an assessment of patient use and perception of usability, Int. J. Med. Inform. 93 (2016) 42–48.
[22] L. Wu, A. Amidi, H. Valdimarsdottir, S. Ancoli-Israel, L. Liu, G. Winkel, E. Byrne, A.V. Sefair, A. Vega, K. Bovbjerg, The effect of systematic light exposure on sleep and sleep quality in a mixed group of fatigued cancer survivors, J. Clin. Sleep Med 14 (01) (2018) 31–39.
[23] A. Sano, A.J. Phillips, Z.Y. Amy, A.W. McHill, S. Taylor, N. Jaques, C.A. Czeisler, E.B. Klerman, R.W. Picard, Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones, Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on, IEEE, 2015, pp. 1–6.
[24] N.S. Simpson, E.L. Gibbs, G.O. Matheson, Optimizing sleep to maximize performance: implications and recommendations for elite athletes, Scand. J. Med. Sci. Sports 27 (2017) 266–274.
[25] T. Ichiba, S. Gorlova, Y. Nagashima, M. Aitake, E. Hori, H. Nishijo, Association between sleep quality and subjective fatigue in night-shift nurses with good and poor sleep, Sleep Vigilance 2 (2018) 63–69.
[26] C.L. McKibbin, S. Ancoli-Israel, J. Dimsdale, C. Archuleta, R. von Kanel, P. Mills, T.L. Patterson, I. Grant, Sleep in spousal caregivers of people with Alzheimer's disease, Sleep 28 (2005) 1245–1250.
[27] M.A. Rowe, C.S. McCrae, J.M. Campbell, A.P. Benito, J. Cheng, Sleep pattern differences between older adult dementia caregivers and older adult noncaregivers using objective and subjective measures, J. Clin. Sleep Med. 4 (2008) 362–369.
[28] C. Simpson, P. Carter, Dementia behavioural and psychiatric symptoms: effect on caregiver's sleep, J. Clin. Nurs. 22 (2013) 3042–3052.
[29] M.E. de Vugt, F. Stevens, P. Aalten, R. Lousberg, N. Jaspers, F.R. Verhey, A prospective study of the effects of behavioral symptoms on the institutionalization of patients with dementia, Int. Psychogeriatr. 17 (2005) 577–589.
[30] B.D. Yetton, E.A. McDevitt, N. Cellini, C. Shelton, S.C. Mednick, Quantifying sleep architecture dynamics and individual differences using big data and Bayesian networks, PLoS One 13 (2018) e0194604.
[31] K. Crowley, Sleep and sleep disorders in older adults, Neuropsychol. Rev. 21 (2011) 41–53.
[32] A. Rechtschaffen, A Manual for Standardized Terminology, Techniques and Scoring System for Sleep Stages in Human Subjects, Brain Information Service, 1968.
[33] J.F. Sassin, D.C. Parker, J.W. Mace, R.W. Gotlin, L.C. Johnson, L.G. Rossman, Human growth hormone release: relation to slow-wave sleep and sleep-waking cycles, Science 165 (1969) 513–515.
[34] G. Brandenberger, L. Weibel, The 24-h growth hormone rhythm in men: sleep and circadian influences questioned, J. Sleep Res. 13 (2004) 251–255.
[35] D. Schneider-Helmert, Why low-dose benzodiazepine-dependent insomniacs can't escape their sleeping pills, Acta Psychiatr. Scand. 78 (1988) 706–711.
[36] J.K. Walsh, Enhancement of slow wave sleep: implications for insomnia, J. Clin. Sleep Med.: JCSM: Off. Pub. Am. Acad. Sleep Med. 5 (2009) S27.
[37] D.-J. Dijk, J.A. Groeger, N. Stanley, S. Deacon, Age-related reduction in daytime sleep propensity and nocturnal slow wave sleep, Sleep 33 (2010) 211–223.
[38] V. Ibáñez, J. Silva, O. Cauli, A survey on sleep assessment methods, PeerJ 6 (2018) e4849.
[39] R.J. Cole, D.F. Kripke, W. Gruen, D.J. Mullaney, J.C. Gillin, Automatic sleep/wake identification from wrist activity, Sleep 15 (1992) 461–469.
[40] A. Cuttone, P. Baekgaard, V. Sekara, H. Jonsson, J.E. Larsen, S. Lehmann, Sensiblesleep: a bayesian model for learning sleep patterns from smartphone events, PLoS One 12 (2017) e0169901.
[41] N.J. Douglas, S. Thomas, M.A. Jan, Clinical value of polysomnography, The Lancet 339 (1992) 347–350.
[42] H. Nakano, K. Hirayama, Y. Sadamitsu, A. Toshimitsu, H. Fujita, S. Shin, T. Tanigawa, Monitoring sound to quantify snoring and sleep apnea severity using a smartphone: proof of concept, J. Clin. Sleep Med. 10 (2014) 73–78.
[43] J. Lee, M. Hong, S. Ryu, Sleep monitoring system using kinect sensor, Int. J. Distributed Sens. Netw. 11 (2015) 875371.
[44] W. Chen, X. Zhu, T. Nemoto, Y. Kanemitsu, K. Kitamura, K. Yamakoshi, Unconstrained detection of respiration rhythm and pulse rate with one under-pillow sensor during sleep, Med. Biol. Eng. Comput. 43 (2005) 306–312.
[45] H. Mönttinen, H. Koskimäki, P. Siirtola, J. Röning, Electrodermal activity asymmetry in sleep-a case study for migraine detection, EMBEC & NBC 2017, Springer, 2017, pp. 835–838.
[46] A. Kratz, S. Ansari, M. Duda, X. Chen, J. Murphy, E. Bay, D. Williams, D. Clauw, K. Najarian, (315) Ambulatory heart rate variability identifies differences in sleep between people with fibromyalgia and matched controls, J. Pain 18 (2017) S54.
[47] A. Kratz, S. Ansari, M. Duda, X. Chen, J. Murphy, D. Williams, K. Najarian, D. Clauw, (317) Feasibility of the E4 Wristband to assess sleep in adults with and without fibromyalgia, J. Pain 18 (2017) S54.
[48] W. Romine, T. Banerjee, G. Goodman, Toward sensor-based sleep monitoring with electrodermal activity measures, Sensors 19 (2019) 1417.
[49] T. Penzel, J.W. Kantelhardt, R.P. Bartsch, M. Riedl, J.F. Kraemer, N. Wessel, C. Garcia, M. Glos, I. Fietze, C. Schöbel, Modulations of heart rate, ECG, and cardio-respiratory coupling observed in polysomnography, Front. Physiol. 7 (2016) 460.
[50] Q. Li, Q. Li, C. Liu, S.P. Shashikumar, S. Nemati, G.D. Clifford, Deep learning in the cross-time frequency domain for sleep staging from a single-lead electrocardiogram, Physiol. Meas. 39 (2018) 124005.
[51] X. Zhang, W. Kou, I. Eric, C. Chang, H. Gao, Y. Fan, Y. Xu, Sleep stage classification based on multi-level feature learning and recurrent neural networks via wearable device, Comput. Biol. Med. 103 (2018) 71–81.
[52] H.J. Lee, S.H. Hwang, H.N. Yoon, W.K. Lee, K.S. Park, Heart rate variability monitoring during sleep based on capacitively coupled textile electrodes on a bed, Sensors 15 (2015) 11295–11311.
[53] T. Penzel, J.W. Kantelhardt, L. Grote, J.-H. Peter, A. Bunde, Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 50 (2003) 1143–1151.
[54] M.G. Miglis, Autonomic dysfunction in primary sleep disorders, Sleep Med. 19 (2016) 40–49.
[55] C.A.G. Martínez, A.O. Quintana, X.A. Vila, M.J.L. Touriño, L. Rodríguez-Liñares, J.M.R. Presedo, A.J.M. Penín, Heart Rate Variability Analysis with the R Package RHRV, Springer, 2017.
[56] P.A. Lanfranchi, V.K. Somers, Cardiovascular physiology: autonomic control in health and in sleep disorders, Principles and Practice of Sleep Medicine, fifth ed., Elsevier Inc., 2010.
[57] A.S.R.W. Picard, Understanding Ambulatory and Wearable Data for Health and Wellness, in: (n.d).
[58] J.A. Onton, D.Y. Kang, T.P. Coleman, Visualization of whole-night sleep EEG from 2-channel mobile recording device reveals distinct deep sleep stages with differential electrodermal activity, Front. Hum. Neurosci. 10 (2016) 605.
[59] J.A. Onton, S.C. Matthews, D.Y. Kang, T.P. Coleman, In-home sleep recordings in military veterans with posttraumatic stress disorder reveal less REM and deep Sleep < 1 hz, Front. Hum. Neurosci. 12 (2018).
[60] A. Sano, R.W. Picard, R. Stickgold, Quantitative analysis of wrist electrodermal activity during sleep, Int. J. Psychophysiol. 94 (2014) 382–389.
[61] J. Wilde-Frenz, H. Schulz, Rate and distribution of body movements during sleep in humans, Percept. Mot. Skills 56 (1983) 275–283.
[62] F. Giganti, G. Ficca, S. Gori, P. Salzarulo, Body movements during night sleep and

their relationship with sleep stages are further modified in very old subjects, Brain Res. Bull. 75 (2008) 66–69.
[63] R.K. Bogan, Effects of restless legs syndrome (RLS) on sleep, Neuropsychiatric Dis. Treat. 2 (2006) 513.
[64] J.P. Libert, Thermal regulation during sleep, Rev. Neurol. 159 (2003) 6S30–4.
[65] R.J. Raymann, D.F. Swaab, E.J. Van Someren, Skin deep: enhanced sleep depth by cutaneous temperature manipulation, Brain 131 (2008) 500–513.
[66] K. Kräuchi, E. Fattori, A. Giordano, M. Falbo, A. Iadarola, F. Aglì, A. Tribolo, R. Mutani, A. Cicolin, Sleep on a high heat capacity mattress increases conductive body heat loss and slow wave sleep, Physiol. Behav. 185 (2018) 23–30.
[67] A. Tataraidze, L. Korostovtseva, L. Anishchenko, M. Bochkarev, Y. Sviryaev, Sleep architecture measurement based on cardiorespiratory parameters, Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the, IEEE, 2016, pp. 3478–3481.
[68] J.E. Sinex, Pulse oximetry: principles and limitations, Am. J. Emerg. Med. 17 (1999) 59–66.
[69] D. Herzig, P. Eser, X. Omlin, R. Riener, M. Wilhelm, P. Achermann, Reproducibility of heart rate variability is parameter and sleep stage dependent, Front. Physiol. 8 (2017) 1100.
[70] W. Boucsein, Electrodermal Activity, Springer Science & Business Media, 2012.
[71] J. Shukla, M. Barreda-Ángeles, J. Oliver, D. Puig, MuDERI: multimodal database for emotion recognition among intellectually disabled individuals, International Conference on Social Robotics, Springer, 2016, pp. 264–273.
[72] B. Bei, J.C. Ong, S.M. Rajaratnam, R. Manber, Chronotype and improved sleep efficiency independently predict depressive symptom reduction after group cognitive behavioral therapy for insomnia, J. Clin. Sleep Med. 11 (2015) 1021–1027.
[73] A. Sano, R.W. Picard, Toward a Taxonomy of Autonomic Sleep Patterns with Electrodermal Activity, (2011).
[74] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40 (2014) 16–28.
[75] Y. Saeys, I. Inza, P. Larrañaga, A review of feature selection techniques in bioinformatics, Bioinformatics 23 (2007) 2507–2517.
[76] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Mach. Learn. 46 (2002) 389–422.
[77] M. Kuhn, Variable selection using the caret package, http://Cran.Cermin.Lipi.Go.Id/Web/Packages/Caret/Vignettes/CaretSelection.Pdf, (2012).
[78] E. Alpaydin, Introduction to Machine Learning, MIT press, 2014.
[79] L. Zhang, J. Xiong, H. Zhao, H. Hong, X. Zhu, C. Li, Sleep stages classification by CW Doppler radar using bagged trees algorithm, Radar Conference (RadarConf), 2017 IEEE, IEEE, 2017, pp. 0788–0791.
[80] A. Kassambara, Machine Learning Essentials: Practical Guide in R, STHDA, (2018).
[81] J. Hughes, T. Banerjee, W. Romine, G. Goodman, A. Edwards, L. Lawhorne, Interdisciplinary Development of a Daily Use Caregiver Sleep Survey (DUCSS): A Mixed Methods Design, SocArXiv, (2019), https://doi.org/10.31235/osf.io/6v2sr.
[82] D.J. Buysse, C.F. Reynolds III, T.H. Monk, S.R. Berman, D.J. Kupfer, The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research, Psychiatr. Res. 28 (1989) 193–213.
[83] T.E. Weaver, A.M. Laizner, L.K. Evans, G. Maislin, D.K. Chugh, K. Lyon, P.L. Smith, A.R. Schwartz, S. Redline, A.I. Pack, An instrument to measure functional status outcomes for disorders of excessive sleepiness, Sleep 20 (1997) 835–843.
[84] W. Ward Flemons, M.A. Reimer, Development of a disease-specific health-related quality of life questionnaire for sleep apnea, Am. J. Respir. Crit. Care Med. 158 (1998) 494–503.
[85] R.D. Hays, C.D. Sherbourne, R.M. Mazel, The rand 36-item health survey 1.0, Health Econ. 2 (1993) 217–227.
[86] G.N. Masters, A Rasch model for partial credit scoring, Psychometrika 47 (1982) 149–174.
[87] A. Dahlgren, G. Kecklund, T. Akerstedt, Different levels of work-related stress and the effects on sleep, fatigue and cortisol, Scand. J. Work. Environ. Health (2005) 277–285.
[88] A. Zeileis, D. Meyer, K. Hornik, Residual-based shadings for visualizing (conditional) independence, J. Comput. Graph. Stat. 16 (2007) 507–525.
[89] A. Awad, M. Bader-El-Den, J. McNicholas, J. Briggs, Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach, Int. J. Med. Inform. 108 (2017) 185–195.
[90] T. Perry, M. Bader-El-Den, S. Cooper, Imbalanced classification using genetically optimized cost sensitive classifiers, Evolutionary Computation (CEC), 2015 IEEE Congress on, IEEE, 2015, pp. 680–687.
[91] R. Rifkin, A. Klautau, In defense of one-vs-all classification, J. Mach. Learn. Res. 5 (2004) 101–141.
[92] G. Louppe, L. Wehenkel, A. Sutera, P. Geurts, Understanding variable importances in forests of randomized trees, Advances in Neural Information Processing Systems, 2013, pp. 431–439.
[93] J. Wang, J. Oh, H. Wang, J. Wiens, Learning credible models, Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, 2018, pp. 2417–2426.
[94] F. Mendonça, S.S. Mostafa, F. Morgado-Dias, A.G. Ravelo-García, T. Penzel, A review of approaches for sleep quality analysis, IEEE Access 7 (2019) 24527–24546.
[95] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Ofli, J. Srivastava, A. Elmagarmid, T. Arora, S. Taheri, Sleep quality prediction from wearable data using deep learning, JMIR MHealth UHealth 4 (2016) e125.

Reza Sadeghi is a PhD candidate of computer science and graduate research assistant at Wright State University. He focuses on investigating applied machine learning in the fields of healthcare, especially in elder care. His research leads to creating varied clinical decision support systems (CDSSs), such as predicting mortality and alcohol withdrawal in ICUs, and sleep quality prediction using both wearable devices and polysomnography. He tries to provide easy-to-use, cost-effective, and interpretable CDSSs based on big multimodal data to aide physicians, patients, and healthcare managements.

Dr. Banerjee is an Assistant Professor in the Department of Computer Science and Engineering, with a secondary appointment at the Department of Geriatrics, Boonshoft School of Medicine at Wright State University. Her current research interests are in healthcare applications with a focus on issues within the geriatric population, including caregiver stress assessment for patients with dementia (featured in local media sources, Alzheimer's Association, as well as by Research Features in the special issue of Women in Science, 2018). Her areas of expertise include multimodal data fusion and machine learning to analyze clinical problems, such as in the areas of gerontology and sickle cell disease.

Jennifer C Hughes received her PhD in social work from the University of Utah (2012), Her Master of Social Work Degree (MSW) from the Ohio State University (1995) and her Bachelors of Science in Art Therapy from Bowling Green State University (1989). She is currently an Assistant Professor of Social Work and program director of the Bachelors of social work program at Wright State University in Dayton, Ohio. Dr. Hughes has extensive clinical social work experience specializing in caregiver issues.



Dr. Larry Lawhorne is Professor and former Chairman of the Department of Geriatrics at Wright State University Boonshoft School of Medicine in Dayton, Ohio. Prior to coming to the medical school in 2006, he was Professor of Family Practice and Director of the Geriatric Center of Michigan at Michigan State University's College of Human Medicine in East Lansing.

Hereceived his medical degree from the University of Virginia in Charlottesville and completed residency at the University of Iowa in Iowa City. He is board certified in family medicine and in geriatrics.

He has received a number of teaching awards, including teacher of the year at Iowa, Missouri, and Michigan State and the national James Pattee Award presented by the American Medical Directors Association. Since coming to Wright State, he has worked with the curriculum committee to incorporate the principles of geriatric medicine into the medical school's courses and has led efforts to create a geriatric medicine fellowship to train a new generation of geriatricians.

His clinical and research efforts focus on dementia and falls.