ORIGINAL RESEARCH

Sleep/Wakefulness Detection Using Tracheal Sounds and Movements

Nasim Montazeri
Ghahjaverestan [1,2]
Sina Akbarian[1,2]
Maziar Hafezi[1,2]
Shumit Saha[1,2]
Kaiyin Zhu[1]
Bojan Gavrilovic[1]
Babak Taati [1–3],*
Azadeh Yadollahi [1,2],*

[1]Kite - Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada; [2]Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada; [3]Computer Science, University of Toronto, Toronto, ON, Canada

*These authors contributed equally to this work

Purpose: The current gold standard to detect sleep/wakefulness is based on electroencephalogram, which is inconvenient if included in portable sleep screening devices. Therefore, a challenge in the portable devices is sleeping time estimation. Without sleeping time, sleep parameters such as apnea/hypopnea index (AHI), an index for quantifying sleep apnea severity, can be underestimated. Recent studies have used tracheal sounds and movements for sleep screening and calculating AHI without considering sleeping time. In this study, we investigated the detection of sleep/wakefulness states and estimation of sleep parameters using tracheal sounds and movements.

Materials and Methods: Participants with suspected sleep apnea who were referred for sleep screening were included in this study. Simultaneously with polysomnography, tracheal sounds and movements were recorded with a small wearable device, called the Patch, attached over the trachea. Each 30-second epoch of tracheal data was scored as sleep or wakefulness using an automatic classification algorithm. The performance of the algorithm was compared to the sleep/wakefulness scored blindly based on the polysomnography.

Results: Eighty-eight subjects were included in this study. The accuracy of sleep/wakefulness detection was 82.3±8.66% with a sensitivity of 87.8±10.8 % (sleep), specificity of 71.4 ±18.5% (awake), F1 of 88.1±9.3% and Cohen's kappa of 0.54. The correlations between the estimated and polysomnography-based measures for total sleep time and sleep efficiency were 0.78 ($p<0.001$) and 0.70 ($p<0.001$), respectively.

Conclusion: Sleep/wakefulness periods can be detected using tracheal sound and movements. The results of this study combined with our previous studies on screening sleep apnea with tracheal sounds provide strong evidence that respiratory sounds analysis can be used to develop robust, convenient and cost-effective portable devices for sleep apnea monitoring.

Keywords: sleep apnea, apnea/hypopnea index, principal component analysis, classification, imbalanced data

Introduction

Sleep apnea is a chronic disorder, which is associated with intermittent reductions (hypopnea) or pauses (apnea) in breathing during sleep. The severity of sleep apnea is commonly quantified based on apnea/hypopnea index (AHI) defined as the average number of apneas/hypopneas per hour of sleeping time. In clinical practice, sleep is assessed by in-laboratory polysomnography (PSG), which requires attachment of various sensors to the body. In PSG, sleeping time is measured through attaching about 8–10 electrodes on the individual's head to record electroencephalogram (EEG), followed by manual inspection of the EEG recordings for sleep scoring. Thus, PSG is inconvenient, expensive and may not be representative of

Correspondence: Azadeh Yadollahi
Kite - Toronto Rehabilitation Institute, University Health Network, Room 12-106, 550 University Avenue, Toronto, ON M5G 2A2, Canada
Tel +1 416 597 3422 Ext 7936
Fax +1 416 597 8959
Email Azadeh.Yadollahi@uhn.ca

natural sleep, since it is performed for a single night in an unfamiliar environment. Alternatively, sleep can be assessed in the home setting using portable sleep screening devices with fewer numbers of sensors. While portable devices address some challenges of PSG, their limited number of integrated sensors can reduce their accuracy. The main challenge of most portable devices is that since they do not record EEG due to the inconvenience, they do not have an estimate of sleeping time. Thus, recording time is usually used in the portable devices to estimate AHI, which leads to lower accuracy and underestimation of AHI. To increase the accuracy of portable devices, it is important to develop algorithms to estimate sleeping time without compromising the convenience of the portable devices.

To detect sleeping time without EEG, most of the portable sleep screening devices have used an accelerometer to detect body movements, such as actigraphy-based methods.[1,2] Actigraphy assumes that wakefulness is associated with spontaneous body motions, while sleep is associated with no or less movements. Although, actigraphy is simple and sensitive to detect sleep periods, its specificity to detect wakefulness is low. The same drawback exists in similar methods based on analyzing the electromyogram of tibialis (leg) or submentalis (chin) muscles.[3,4] Alternatively, Dafna et al[5,6] proposed the analysis of breathing sounds in patients with respiratory sleep disorders for detecting the sleeping periods. In this approach, sleep/wakefulness periods were detected by analyzing the pattern of breathing sounds recorded by an ambient microphone in the room. Despite the high overall sensitivity of 92.2%, the specificity to detect wakefulness was low (56.6%) and the algorithm can be sensitive to the ambient noises.

Previously, Soltanzadeh and Moussavi[7] used tracheal sounds to detect sleep/wakefulness during sleep. Through performing higher order statistics analysis, they were able to differentiate sleep stages from wakefulness. However, they only selected 10 breaths from each state for each subject. Although they achieved 100% accuracy over the data of 12 individuals, they never validated their algorithm on overnight-length data or in larger populations such as sleep patients. Especially in sleep apnea patients, fluctuations in the pattern of respiration and body motions can affect the performance of sleep detection algorithms. Therefore, there is a need to develop and validate robust algorithms to detect sleep and wakefulness from full overnight sleep data in sleep apnea patients using tracheal signals.

Tracheal signals have been extensively used to monitor respiration during sleep and assess the severity of sleep apnea by estimating AHI.[8–17] Tracheal sounds can be conveniently recorded using a microphone embedded in a portable device attached over the suprasternal notch. Recently, we have developed a device called "The Patch" to record tracheal respiratory related sounds and movements. As a portable device, The Patch does not include EEG; thus, by detecting sleep/wakefulness intervals from tracheal signals, it can be used to estimate total sleep time and AHI with higher accuracy. The goal of this study was to detect sleep/wakefulness by analyzing tracheal sounds and movements recorded by the Patch. The detected sleep/wakefulness intervals were used to estimate total sleep time and other sleep quality parameters. For validation, the detected scores and sleep quality parameters were compared to those derived based on EEG and actigraphy.

Materials and Methods
Study Participants and Protocol

Ninety participants aged 18 years and above with suspected sleep apnea, who were referred to the sleep laboratory of Toronto Rehabilitation Institute for sleep diagnosis were included in this study. We excluded individuals with history of tracheostomy or allergy to adhesive medical tapes. The study was conducted based on the protocol approved by the Research Ethics Board of the University Health Network (IRB #: 15-8967). Each individual was informed about the purpose of the study and gave written consent before participating in the study which was conducted in accordance with the Declaration of Helsinki.

Data Collection

Overnight in-laboratory sleep monitoring using polysomnography was conducted. As part of the PSG, EEG and electrooculogram were recorded and used to detect sleep and wakefulness periods by technicians who were blinded to our study. According to the American Academy of Sleep Medicine criteria,[18] each 30-second epoch of data was annotated as sleep (NREM: non-rapid eye movement and REM: rapid eye movement) or wakefulness and used as the reference labels. Simultaneously with PSG, a wearable device, The Patch, which was developed in our laboratory,[13] was attached over the suprasternal notch using double sided tape that none of the participants had

Figure I The Patch.

any challenge regarding discomfort of The Patch (Figure 1). The Patch records tracheal sounds with a one-directional microphone (sampling rate = 15 kHz) and tracheal-related movements with a 3-dimensional accelerometer (sampling rate = 60 Hz). All the sensors are in special housing to minimize recording ambient sounds. The Patch data recording was synchronized by a micro-controller with those of PSG. The microcontroller embedded in The Patch restores a pulse every 10 minutes, while at the same time, sends a pulse to one of the PSG channels.

Data Analysis

The following analyses were implemented in Matlab (2016b, MathWorks, Natick, MA, USA) software.

Preprocessing

For preprocessing, recorded tracheal movements in the X, Y, and Z directions were filtered using a 5th order zero-phase Butterworth band-pass filter (0.1–0.35 Hz) to extract the movements related to respiration (0.2–0.33 Hz[19]). To remove baseline swings and high frequency noises, tracheal sound was band-pass filtered by 5th order zero-phase Butterworth filter with 70–2000 Hz bandwidth. Subsequently, the filtered tracheal movements and sounds were segmented using a moving window of 10 seconds with 50% overlap. From each segment of data, four features from each accelerometer dimension and four features from sound (16 features in total) were extracted (Figure 2).

Feature Extraction

For each segment of movement data, the following features were extracted:

- Zero crossing rate (ZCR): This feature was calculated as the rate of sound energy signal passing its mean value level, and then smoothed using a median filter with a window size of 2 minutes. ZCR is related to respiratory rate as it quantifies the speed of fluctuations normally caused by the alternation of respiratory phases in the sound energy signal.
- Baseline movements (BaseLine): To extract this feature, first, the absolute derivative of movement (dMabs) was calculated. dMabs includes abrupt spikes caused by body motions that are superimposed over the intensity of respiratory movements. Accordingly, by smoothing dMabs, BaseLine movement, which is related to the intensity of respiratory movements, was derived.
- Body motions: By subtracting the BaseLine from dMabs, motion spikes were extracted. Within a 2-minute window moving every second, the occurrence rate of the motion spikes with amplitude more than double the 90th percentile was calculated. The occurrence rate was smoothed using 30-second and 1-hour windows to calculate features Spike30s and Spike1h, respectively.

From each segment of the filtered sound, spectral auto-correlation was calculated. In the autocorrelation signal,
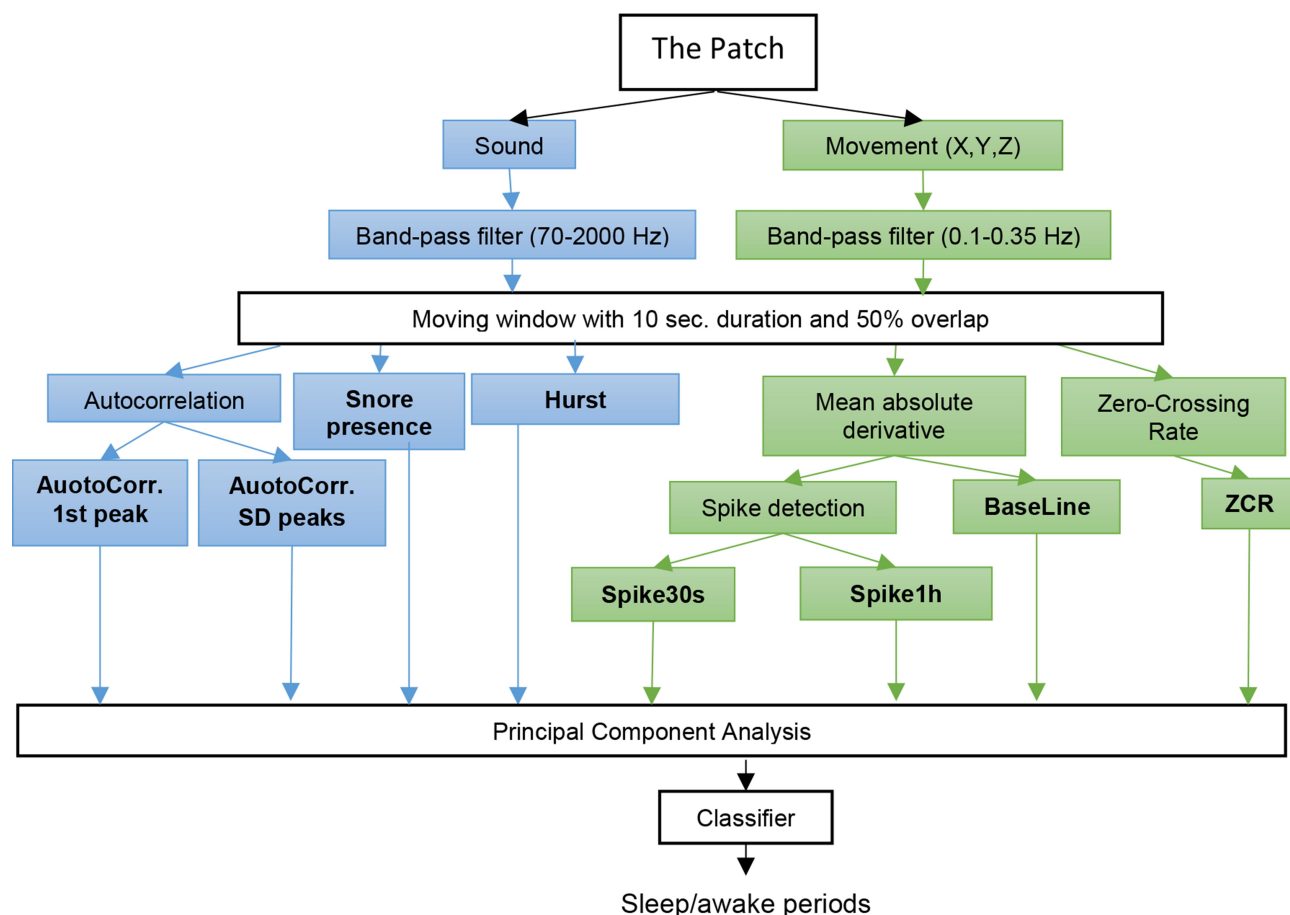
Figure 2 Sleep/wakefulness detection algorithm using tracheal sounds and movements.

all the local maximums were found as peaks. Then, the following features were extracted:

- AuotoCorr. 1st peak: the temporal location of the first peak after excluding the zero-lag autocorrelation. The feature quantifies the periodicity of the respiratory related sound.
- AuotoCorr. SD peaks: the standard deviation of the amplitude of the peaks, which accounts for the resemblance of the breaths occurring in the segment.
- Hurst exponent: this feature was calculated to quantify the speed of reduction in the resemblance of the respiratory phases.[20]

Finally, the presence of snoring in each epoch, which indicated sleep state was extracted. To detect snoring, tracheal sound was filtered using a 2nd order zero-phase Butterworth bandpass filter with 0.2 to 2000 Hz bandwidth. Then, four features including zero-crossing rate, power spectral density of the rising slope of zero-crossing, sound energy and variance were extracted from the

segments of sound chosen using a moving window with 60 ms and 50% overlap. In addition, recorded tracheal movements were filtered using 5th order zero-phase high-pass Butterworth with 5Hz cut-off frequency. Then, the variance of the summation of movements in three directions was calculated within a one-second window every 0.5 sec. Random Forest was used for snore detection for every second.

Figure 3 shows an example of the extracted features from tracheal sound and movements.

Machine Learning Algorithm

Based on the extracted features, a support vector machine (SVM) with Gaussian kernel was used to classify each epoch into sleep or awake, which is known to be less sensitive to outliers of the data.[21] To avoid overfitting and to deal with the imbalanced data in subjects with high sleep efficiency (SE) who have very few epochs of awake data, subjects were divided into those with SE<80% ($SE_{less80\%}$) and SE>80% ($SE_{more80\%}$). To define train and test datasets, 4-fold cross-validation was performed.
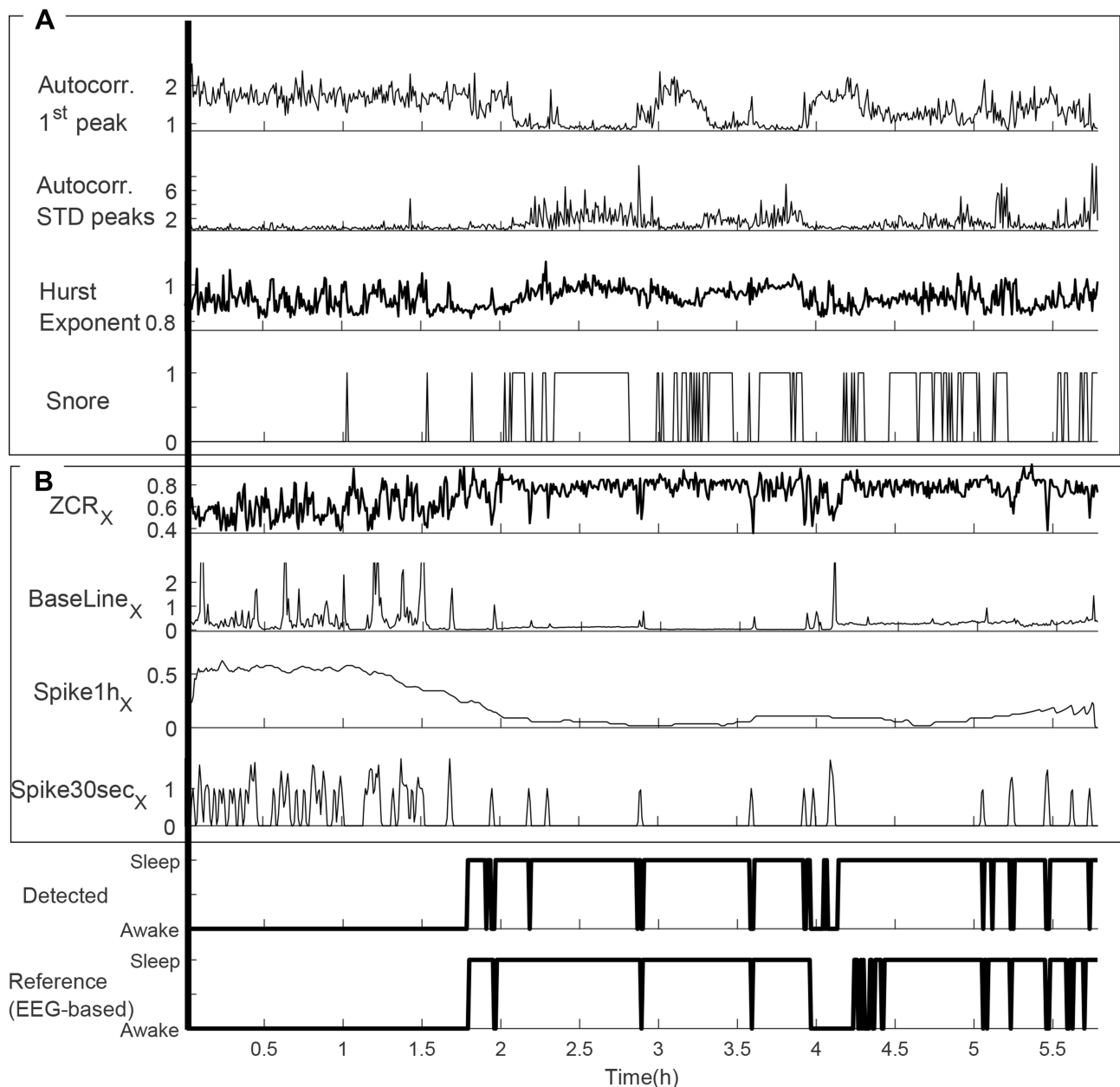
Figure 3 Sleep/wakefulness detection using the extracted features from (**A**) sound and (**B**) movement for a subject with sleep efficiency of 63.60% (estimated as 63.40%), sleep latency of 108 (estimated as 107 min) and total sleep time of 330 min (estimated as 312 min).

Subjects with ($SE_{less80\%)}$ were divided into 4 folds, 3 folds were interchangeably used as training data. The combination of the remaining fold of $SE_{less80\%}$ and data from subjects in $SE_{more80\%}$ group were used as the test data (Figure 4). For each fold of cross-validation, principal component analysis (PCA) was performed on the training data to reduce the number of features. Among the extracted principal components, those with eigenvalues less than 0.001 were removed. The remaining components were used to train the classifier. To validate the algorithm,

the classifier was evaluated using the same chosen components extracted from the test data. Finally, the results of the classifier to identify each epoch into sleep/wakefulness were used to assess the quality of sleep and to estimate sleep efficiency, total sleep time and sleep latency.

We compared the performance of the proposed algorithm with actigraphy. The accelerometer in The Patch can capture body movements. Therefore, extracted movement spikes (Spike30s and Spike1h features) were used separately to train the same algorithm to simulate the
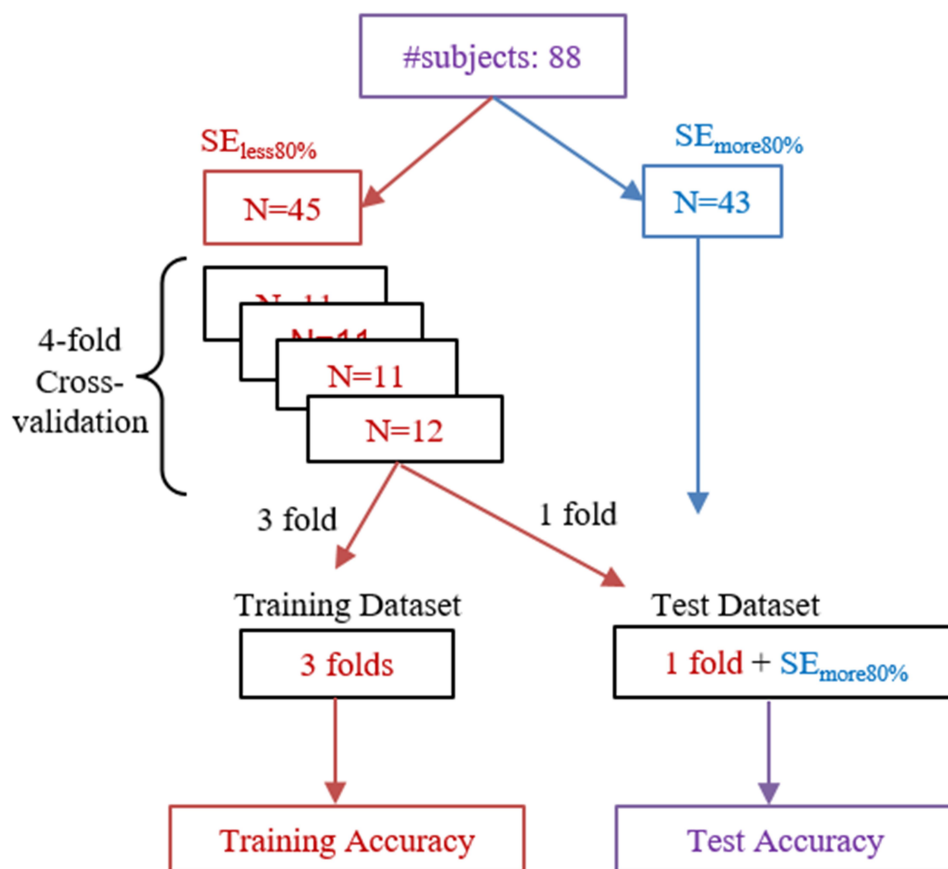
Figure 4 Four-fold cross-validation for training and testing the mathematical model. Training was performed using the data of those with sleep efficiency less than 80%.
Abbreviations: $SE_{more80\%}$, subset of subjects with sleep efficiency more than 80%; $SE_{less80\%}$, subset of subjects with sleep efficiency less than 80%.

actigraphy method. Similarly, two other models were trained using only sound features or movement features for more comparison.

Statistical Analysis

Statistical analyses were performed in R (i386 3.4.1) software. Shapiro–Wilk test was used for checking the normality of data. To compare the number of females between groups $SE_{less80\%}$ and $SE_{more80\%}$, we used the Chi-square test. Welch unpaired *t*-test or Wilcoxon rank sum test were used to compare other characteristics between the two groups. Also, *t*-test was used to compare the estimated sleep quality parameters with their PSG-based values. Performance of the sleep/wakefulness classification algorithm was quantified by sensitivity (sleep), specificity (wakefulness), accuracy, F1-score and Cohen's Kappa (κ). A sample size of at least 73 subjects was calculated to maintain a statistical power of 0.80 ($\alpha=0.05$) to achieve $\kappa=0.7$ with precision of 0.2 for two states classification for each subject, which was less than 90 subjects included in this study. One-way Analysis of Variance (ANOVA) was

used to compare the value of the extracted features across wakefulness, REM and non-REM stages. Performance of the classification algorithm trained with all the extracted features was assessed among healthy participants (AHI<5) and those with mild (5≤AHI<15), moderate (15≤AHI<30) and severe (30≤AHI) sleep apnea using one-way ANOVA. In case of significant differences, Tukey Post-hoc analysis was performed. Finally, to assess the amount of agreement between PSG-based and estimated sleep quality parameters, Pearson or Spearman correlation and Bland-Altman plot were performed.

Results

Out of 90 participants recruited for this study, data of two participants were excluded due to low quality of sound signals and misplacement of The Patch. Hence, a total of 88 subjects (age: 53±15 years, 42 females) with a body mass index (BMI) of 29.6±6.2 kg/m^2 with Epworth Sleepiness Scale of 8±4 were considered for this study. Demographics and sleep quality parameters of the

Table 1 Demographics and Sleep Structure of Participants

Characteristics	Total (N = 88)	SE more80% (N=43)	SE less80% (N=45)	p-value
Female (%)	47.7%	44.2%	51.1%	0.662
Body Mass Index (kg/m^2)	29.6 ± 6.2	28.9 ± 5.4	30.3 ± 6.8	0.634
Age (years)	53 ± 15	48 ± 13	57 ± 15	0.014*
AHI (events/hr)	12.6 (0.2–86.7)	11.2 (0.2–83.7)	13.9 (0.4–86.7)	0.406
Sleep efficiency (%)	76.5 ± 15.1	88.3 ± 5.5	65.1 ± 12.3	<0.001*
Sleep latency (min)	12 (0–147)	8 (0–42)	23 (0 −147)	<0.001*
Total sleep time (min)	334 (86–476)	380 (293–477)	240 (86–371)	<0.001*
Epworth sleepiness scale	8 ± 4	8 ± 4	8 ± 5	0.923

Notes: Values are reported as mean ± standard deviation in case of normality; otherwise median (minimum − maximum) are reported. *p value <0.05.
Abbreviations: SE more80%, subset of subjects with sleep efficiency more than 80%; SE less80%, subset of subjects with sleep efficiency less than 80%; AHI, apnea/hypopnea index.

participants are detailed in Table 1. Forty-three subjects had sleep efficiency of less than 80% and 45 subjects had sleep efficiency of more than 80%. Age was significantly higher in those with sleep efficiency of less than 80% (p=0.01). BMI, AHI, and number of females were similar between the groups (p>0.1 for all). Subjects with sleep efficiency less than 80% had shorter total sleep time (p<0.001) and longer sleep latency (p<0.001).

Figure 5 depicts the range of variations in the extracted features during wakefulness, NREM and REM sleep. Compared to wakefulness, AutoCorr. SD peaks, Hurst exponent and ZCR were significantly larger during NREM (p<0.001 for all except Hurst exponent with p=0.02). In contrast, the values of AutoCorr. 1st peak, BaseLine, Spike30s, and Spike1h in all dimensions were lower during NREM than during wakefulness (p<0.001 for all, Figure 5). The movement related features were significantly different during REM compared to wakefulness (p<0.001), and statistically the same compared to NREM. However, among sound features, AutoCorr. SD peaks and Hurst exponent were significantly higher during REM compared to wakefulness (p=0.03 and p=0.006, respectively). Significant lower values of AutoCorr. SD peaks were observed in REM to NREM (p<0.001), while AutoCorr. 1st peak was significantly higher (p=0.003). No significant changes were found between REM and NREM in the Hurst exponent.

Using all the extracted features, the accuracy of the sleep/wakefulness detection algorithm was 85.04±0.99% with κ=0.64±0.02 in the train dataset. In the SE$_{more80\%}$ test dataset, the accuracy was 86.20±8.98% with κ=0.50±0.15 while in the SE$_{less80\%}$ test dataset, accuracy was 82.04 ±8.03% with κ=0.58±0.19 (Table 2). Accuracy and

Cohen's Kappa score were lower for the other models trained using subsets of the features.

Table 3 shows the performance of the proposed detection algorithm among different AHI groups. While the accuracy was slightly less in severe sleep apnea patients compared to the healthy subjects (80.41±10.30 vs 86.38 ±6.39), the changes were not significant. However, the Kappa score (p=0.005) and specificity (p=0.01) were significantly lower in the severe sleep apnea group compared to the healthy group.

By comparing the estimated and PSG-based sleep quality parameters, no significant differences were found in sleep efficiency [75.74 (23.58–98.04) % vs 78.80 (19.17–97.75) %, p=0.3], sleep latency [15 (0–157) min vs 16 (0–159) min, p=0.9] and total sleep time [595 ± 133 min vs 609±143 min, p=0.5]. Furthermore (Figure 6), strong significant correlations were observed between the detected and PSG-based sleep efficiency (r=0.70, p<0.001), sleep latency (r=0.71, p<0.001) and total sleep time (r=0.78, p<0.001).

Discussion

In this study, we developed an algorithm to detect sleep/wakefulness periods using tracheal sounds and movements in patients with sleep apnea. The main findings of our study are that: 1) the extracted features from tracheal sounds and movements were significantly different between sleep and wakefulness; 2) the proposed algorithm can detect sleep/wakefulness overnight with an accuracy of 84.08% and κ of 0.54; and 3) the parameters to assess sleep quality based on our proposed method were similar to those derived from gold standard PSG.
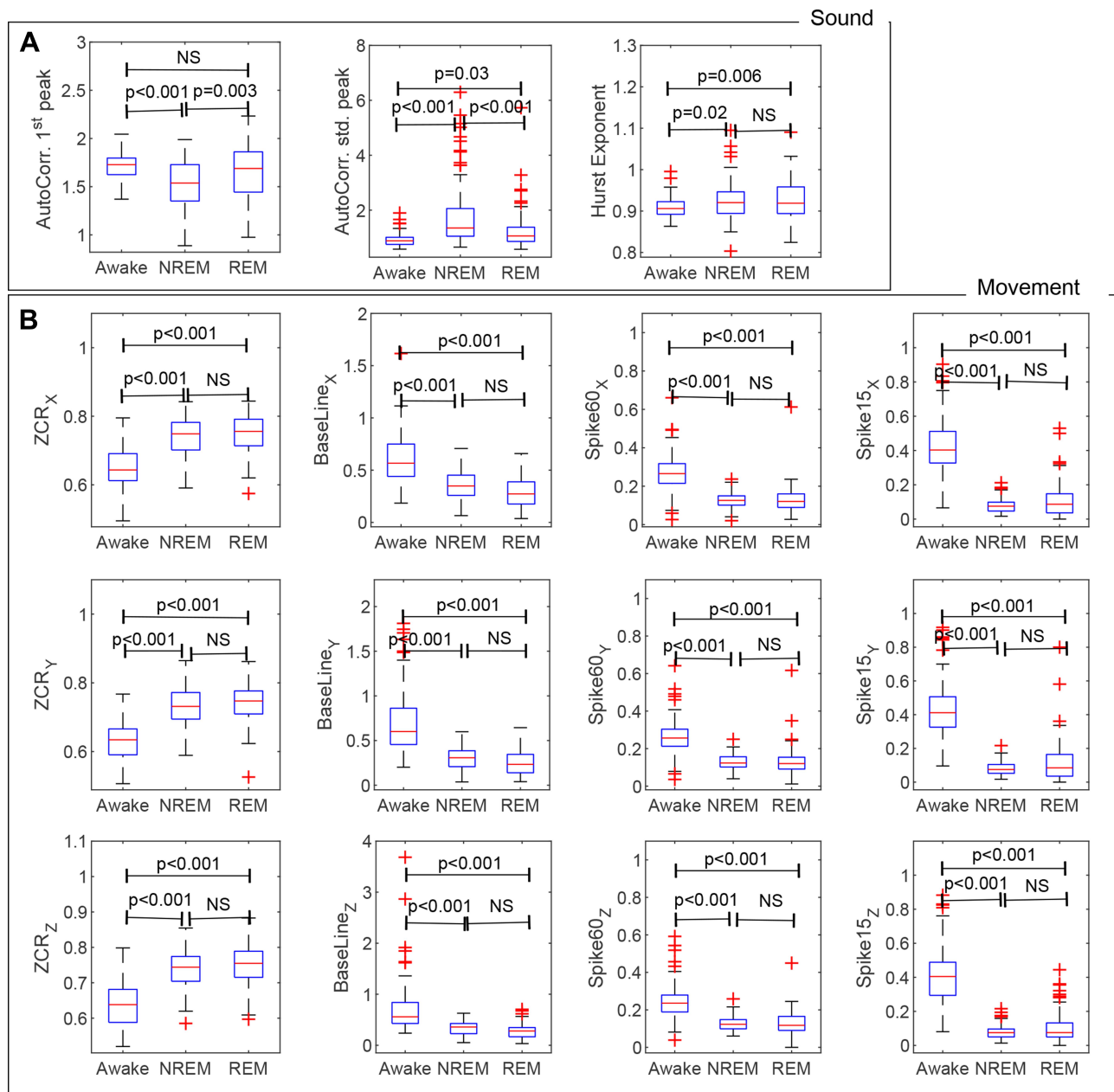
Figure 5 Comparison of various features extracted from (**A**) tracheal sound and (**B**) movements during wakefulness with NREM and REM. Box plots demonstrates median, first and last quartiles.
Abbreviations: AutoCorr. 1st peak, the average temporal location of the first peak in autocorrelations calculated over epochs; AutoCorr. SD peak, the average SD of the autocorrelations peaks; ZCR, zero-crossing rate; NS, not significant.

The proposed algorithm is the first to detect sleep/wakefulness states in full-night data including tracheal sounds and movements. It was successful to detect sleep and wakefulness epochs with high sensitivity (sleep detection) and specificity (wakefulness detection). This work was presented at the World Sleep congress 2019 as an abstract presentation with interim findings. The poster's abstract was published in *Sleep Medicine* [doi: 10.1016/j.sleep.2019.11.740]. Most of the previous studies on sleep/wakefulness detection have reported high sensitivity and accuracy, but with low value or no results for specificity. This was due to the imbalanced higher portion of sleeping epochs in the data recorded during the sleep test that challenged the learning algorithms. In this study, to overcome the imbalanced nature of the sleep/wakefulness data, subjects were categorized into two groups based on the sleep efficiency cut-off of 80%. Higher sleep efficiency is associated with fewer wakefulness periods. Thus, subjects

Table 2 Performance Comparison of the Proposed Sleep/Wakefulness Detection Algorithm and the Method Based on Spontaneous Body Movements

			SEN (Sleep)	SPC (Awake)	F1	ACC	κ
Movement features	Train data Test data	$SE_{less80\%}$ $SE_{less80\%}$ $SE_{more80\%}$ All	91.41 ± 1.18 88.78 ± 10.79 89.33 ± 7.23 89.05 ± 9.18	62.13 ± 3.46 60.09 ± 19.48 66.67 ± 20.15 63.31 ± 19.97	85.35 ± 0.57 83.28 ± 11.55 91.69 ± 4.42 87.39 ± 9.73	81.57 ± 0.35 78.82 ± 10.07 86.34 ± 6.50 82.49 ± 9.27	0.55 ± 0.02 0.49 ± 0.22 0.46 ± 0.16 0.48 ± 0.20
Sound features	Train data Test data	$SE_{less80\%}$ $SE_{less80\%}$ $SE_{more80\%}$ All	78.99 ± 2.61 78.73 ± 16.42 77.93 ± 18.06 78.34 ± 17.15	68.03 ± 1.39 78.73 ± 18.29 68.52 ± 19.87 68.47 ± 18	78.89 ± 1.80 78.93 ± 12.45 84.02 ± 14.02 81.42 ± 13.41	75.89 ± 1.58 75.85 ± 9.52 76.88 ± 18.06 76.35 ±12.58	0.46 ± 0.03 0.46 ± 0.19 0.33 ± 0.16 0.40 ± 0.19
Spike30s + Spike1h	Train data Test data	$SE_{less80\%}$ $SE_{less80\%}$ $SE_{more80\%}$ All	90.89 ± 1.18 90.36 ± 7.63 89.15 ±6.88 89.77± 7.2	54.15 ± 1.68 54.52 ± 19.48 65.15 ± 21.09 59.71 ± 20.86	82.97 ± 1.41 82.71 ± 11.10 91.51 ± 3.76 87.01 ± 9.41	77.60 ± 1.20 77.23 ± 10.31 85.87 ± 5.70 81.45 ± 9.40	0.47 ± 0.02 0.46 ± 0.20 0.43 ± 0.18 0.45 ± 0.19
All the features	Train data Test data	$SE_{less80\%}$ $SE_{less80\%}$ $SE_{more80\%}$ All	90.50 ± 1.60 87.72 ± 11.24 88.02 ± 10.56 87.86 ± 10.85	72.89 ± 1.75 69.87 ± 18.29 73.10 ± 18.81 71.44 ± 18.51	87.56 ± 1.36 85.10 ± 10.45 91.31 ± 6.76 88.13 ± 9.33	85.04 ± 0.99 82.04 ± 8.03 86.20 ± 8.98 84.08 ± 8.36	0.64 ± 0.02 0.58 ± 0.19 0.50 ± 0.15 0.54 ± 0.18

Note: The results are reported as mean ± standard deviation.
Abbreviations: SEN, sensitivity; SPC, specificity; F1, F1 score; ACC, accuracy; κ, Cohen's kappa coefficient.

with high sleep efficiency were excluded from the training dataset. Based on this approach, we were able to improve specificity by 20% (SPC=71.44%), with a sensitivity of 87.86% and robustness in different AHI groups.

Respiratory patterns can change from wakefulness to sleep due to the reduction in the respiratory drive and activity of pharyngeal dilator muscles, which results in shallower and more regular breathing patterns.[22,23] The studies conducted by Dafna et al[5,6] showed that the shapes of respiratory sounds recorded by a non-contact ambient microphone as quantified by autocorrelation are more identical (higher periodicity) during consecutive breaths in sleep than those during wakefulness. They obtained sensitivity of 92.2%, but low specificity of 56.6% for sleep/wakefulness detection with κ of 0.51.[5] In another study,[6] they were able to improve the learning algorithm and obtained average accuracy of 91.7% and κ of 0.68,

however they have not reported the performance of the algorithm in detecting sleep and wakefulness separately. This information is important, since sleep apnea patients often have fragmented and less efficient sleep. Furthermore, the quality of ambient microphone can be deteriorated by ambient noises in home settings.

In contrast, other studies have shown that the respiratory related sounds recorded over the trachea are more robust against interfering noises,[12,24] while their energy is highly correlated to the pattern of respiratory airflow.[25] Due to the special housing of the one-directional microphone used to record tracheal sounds, they are less sensitive to ambient noises[24] compared to non-contact methods to record breathing sounds. Also, close placement of the sensor to the source of vibrations in the airway caused by respiratory cycles increases the signal to noise ratio and makes tracheal sounds less sensitive to the quiet breathing.

Table 3 Performance Evaluation of the Sleep/Wakefulness Detection Algorithm in Different AHI Groups

AHI Group	N	SE (%)	SEN (Sleep)	SPC (Awake)	F1	ACC	κ
AHI < 5	27	78.3 ± 13.1	89.34 ± 6.78	78.23 ± 15.57	90.44 ± 5.47	86.38 ± 6.39	0.60 ± 0.14
5 ≤ AHI <15	20	75.5 ± 14.7	87.73 ± 9.22	69.58 ± 22.14	87.92 ± 5.96	82.94 ± 7.63	0.52 ± 0.19
15 ≤ AHI < 30	21	76.6 ± 14.5	87.62 ± 12.24	73.32 ± 17.78	88.34 ± 9.02	83.82 ± 9.02	0.56 ± 0.19
AHI ≥ 30	20	74.8 ± 18.9	84.09 ± 16.07	61.56 ± 17.42*	83.41 ± 15.85	80.41 ± 10.30	0.43 ± 0.18*

Notes: SE was statistically similar across AHI groups. *Adjusted *p*-value <0.05 compared to healthy group (AHI<5). The results are reported as mean ± standard deviation.
Abbreviations: AHI, apnea/hypopnea index; SE, sleep efficiency; SEN, sensitivity; SPC, specificity; F1, F1 score; ACC, accuracy; κ, Cohen's kappa coefficient.
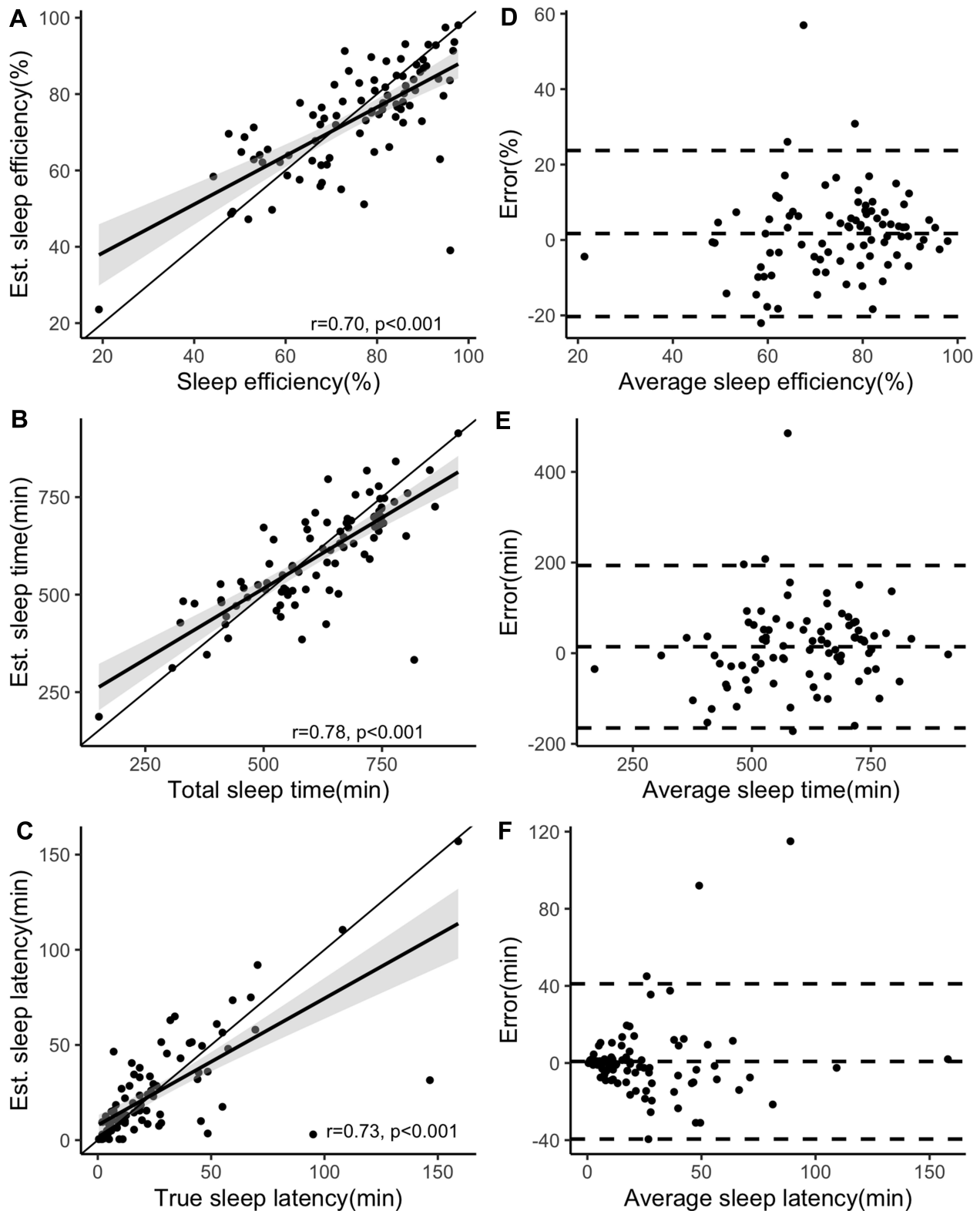
Figure 6 Agreement analyses between estimated and PSG-based measures of sleep quality. Line with gray shades represents least square line with confidence interval (CI) of 95%. (**A** and **D**) sleep efficiency assessed by Spearman's rank correlation with CI = (0.58–0.80). (**B** and **E**) sleep time assessed by Pearson's product-moment correlation with CI = (0.68–0.85). (**C** and **F**) sleep latency assessed by Spearman's rank correlation with CI = (0.61–0.81).

1018

Nature and Science of Sleep 2020:12

To analyze the respiratory patterns, autocorrelation of tracheal sound energy was extracted in this study. Moreover, from the calculated autocorrelation of tracheal sounds, the Hurst exponent was extracted. In a study by Soltanzadeh and Moussavi,[7] the Hurst exponent of the bispectrum of tracheal sound was found to be higher across sleep stages than during wakefulness and can be used for differentiating sleep from wakefulness with 100% accuracy. However, their study was applied on a limited number of segments and its performance was not assessed on overnight data or on sleep apnea population. One reason could be high computational cost of bispectrum calculation for long overnight data. Therefore, in this study, the Hurst component was calculated from the autocorrelation of sound energy. In-line with the findings in Soltanzadeh and Moussavi,[7] the Hurst exponent was higher during sleep indicating a more regular pattern of breathing compared to wakefulness. However, unlike the finding in Soltanzadeh and Moussavi,[7] no significant difference was found between REM and NREM of full-night data.

Sleep/wakefulness stages can affect the respiratory related movements recorded over the chest or the trachea. In this study, changes in respiratory related tracheal movements were quantified by ZCR and BaseLine. Compared to wakefulness, ZCR demonstrated higher values during sleep, presumably due to more regular breathing. On the other hand, increase in the BaseLine feature during sleep indicated higher intensity in respiratory related movements (Figure 3). Despite the increase in the baseline during sleep, the average value of the BaseLine feature was higher in wakefulness (Figure 5) than sleep, since there were abrupt spikes in the accelerometer signal during wakefulness due to body motions.

In healthy populations, body motions are at a minimum level during sleep. These motions have been analyzed for sleep detection in the studies based on actigraphy with sensitivity of more than 90%, but specificity less than 55%.[1,26,27] Low specificity in actigraphy is mostly due to intervals when the subject is awake with no motion. Also, in individuals with sleep apnea, actigraphy can find lots of movements during sleep caused by the presence of apneas/hypopneas that can mistakenly be scored as wakefulness. A combination of actigraphy with other features such as heart rate variability may improve the accuracy of sleep/wakefulness detection.[28–30] Our proposed method extracts features from body movement which are recorded by tracheal movements (Spike30s and Spike1h features) to simulate actigraphy features. The importance of these results is that a combination of body movement with respiratory related sounds that can be recorded over the trachea by a portable and convenient sleep screening device can have significant clinical applications beyond sleep detection, such as assessment of respiration and severity of sleep apnea. Moreover, the analysis based on only body motion features resulted in lower accuracy compared to the model trained by a combination of body motion features with those related to the respiratory sounds. In a study on sleep detection by Kalkbrenner et al,[31] tracheal sounds were used in combination with actigraphy over the chest. Despite comparable sleep detection accuracy and ability to assess respiration with our study, they used more complex hardware settings including a microphone over the neck, chest bands, and connecting wires. Embedding the accelerometer along with the microphone in one casing, The Patch presents a more convenient recording of tracheal signals.

The main limitation of this study is that misplacement of the wearable device over the neck reduces the signal quality, which affects the performance of the detection algorithm. This happened in the data collection of two (out of 90) participants in our study. Another limitation is related to the BaseLine feature. Based on our observations, BaseLine represents the effect of respiratory related movements superimposed by body motions. Further analyses are required to decompose these two effects and analyze them separately. This could be incorporated in addressing another limitation of this study on differentiating between different sleep stages. In previous studies, sleep stages were detected overnight by Dafna et al[5,6] using an ambient microphone and in a few short segments of sleep using tracheal sound by Soltanzadeh and Moussavi.[7] In this study, significant changes were found between REM and other states only in sound features. More in-depth analysis to differentiate various sleep stages should be addressed in future studies. For example, low frequency and high frequency components of heart rate are known to change from wakefulness to sleep and across sleep stages.[21,32,33] Heart sounds can be osculated over the trachea.[9] In future studies, we will analyze the tracheal heart sound for accurate extraction of heart rate and relevant features for sleep staging. Finally, this algorithm was evaluated on a population that was referred to the sleep laboratory for sleep apnea assessment. Further studies are needed to assess the proposed algorithm in a wider population.

Conclusion

This is the first algorithm developed for detecting sleep/wakefulness over full-night sleep data as the combination of tracheal sounds and movements in a population including individuals with sleep apnea. Sleep apnea is an underdiagnosed disorder with adverse societal and clinical outcomes such as higher rates of car/work accidents,[34,35] cardiovascular problems[36] and neurocognitive deficits.[32,37] Developing portable devices can facilitate the diagnosis of sleep apnea. In this regard, extracting total sleep time can significantly improve the performance of the portable devices to estimate AHI and sleep quality parameters, such as sleeping efficiency. Accurate estimation of sleeping time can increase the accuracy of estimated AHI especially in individuals with low sleep efficiency. The results of this study combined with our previous studies for estimating AHI based on tracheal sounds provide strong evidence that the proposed modality can be used to develop robust portable devices for monitoring sleep-related breathing disorders.

Acknowledgments

Disclosure

References

1. Marino M, Li Y, Rueschman MN, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*. 2013;36(11):1747–1755. doi:10.5665/sleep.3142

2. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. *Sleep Breath*. 2012;16(3):913–917. doi:10.1007/s11325-011-0585-y

3. Al-Angari H Evaluation of chin EMG activity at sleep onset and termination in obstructive sleep apnea syndrome. Paper presented at: 2008 Computers in Cardiology; 2008.

4. Hwang S, Chung G, Lee J, et al. Sleep/wake estimation using only anterior tibialis electromyography data. *Biomed Eng Online*. 2012;11 (1):26. doi:10.1186/1475-925X-11-26

5. Dafna E, Tarasiuk A, Zigel Y. Sleep-wake evaluation from whole-night non-contact audio recordings of breathing sounds. *PLoS One*. 2015;10(2):e0117382. doi:10.1371/journal.pone.0117382

6. Dafna E, Tarasiuk A, Zigel Y. Sleep staging using nocturnal sound analysis. *Sci Rep*. 2018;8(1):13474. doi:10.1038/s41598-018-31748-0

7. Soltanzadeh R, Moussavi Z. Sleep stage detection using tracheal breathing sounds: a pilot study. *Ann Biomed Eng*. 2015;43 (10):2530–2537. doi:10.1007/s10439-015-1290-y

8. Hafezi M, Montazeri N, Zhu K, Alshaer H, Yadollahi A, Taati B. Sleep apnea severity estimation from respiratory related movements using deep learning. Paper presented at: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2019.

9. Kalkbrenner C, Eichenlaub M, Rüdiger S, Kropf-Sanchen C, Rottbauer W, Brucher R. Apnea and heart rate detection from tracheal body sounds for the diagnosis of sleep-related breathing disorders. *Med Biol Eng Comput*. 2018;56(4):671–681. doi:10.1007/s11517-017-1706-y

10. Montazeri A, Giannouli E, Moussavi Z. Assessment of obstructive sleep apnea and its severity during wakefulness. *Ann Biomed Eng*. 2012;40(4):916–924. doi:10.1007/s10439-011-0456-5

11. Nakano H, Hirayama K, Sadamitsu Y, et al. Monitoring sound to quantify snoring and sleep apnea severity using a smartphone: proof of concept. *J Clin Sleep Med*. 2014;10(01):73–78. doi:10.5664/jcsm.3364

12. Penzel T, Sabil A. The use of tracheal sounds for the diagnosis of sleep apnoea. *Breathe*. 2017;13(2):e37–e45. doi:10.1183/20734735.008817

13. Saha S, Kabir M, Montazeri N, et al. Apnea-hypopnea index (AHI) estimation using breathing sounds, accelerometer and pulse oximeter. *ERJ Open Res*. 2019;5(suppl3):P63.

14. Solà-Soler J, Fiz JA, Torres A, Jané R Identification of obstructive sleep apnea patients from tracheal breath sound analysis during wakefulness in polysomnographic studies. Paper presented at: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2014.

15. Yu L, Ting C-K, Hill BE, et al. Using the entropy of tracheal sounds to detect apnea during sedation in healthy nonobese volunteers. *Anesthesiology*. 2013;118(6):1341–1349. doi:10.1097/ALN.0b013e318289bb30

16. Nakano H, Furukawa T, Tanigawa T. Tracheal sound analysis using a deep neural network to detect sleep apnea. *J Clin Sleep Med*. 2019;15 (8):1125–1133. doi:10.5664/jcsm.7804

17. Yadollahi A, Giannouli E, Moussavi Z. Sleep apnea monitoring and diagnosis based on pulse oximetry and tracheal sound signals. *Med Biol Eng Comput*. 2010;48(11):1087–1097. doi:10.1007/s11517-010-0674-2

18. Berry RB, Brooks R, Gamaldo CE, Harding SM, Marcus C, Vaughn B The AASM manual for the scoring of sleep and associated events. Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine; 2012.

19. Yuan G, Drost NA, McIvor RA. Respiratory rate and breathing pattern. *McMaster Univ Med J*. 2013;10(1):23–25.

20. Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*. 2002;316(1–4):87–114. doi:10.1016/S0378-4371(02)01383-3

21. Hoak J. *The Effects of Outliers on Support Vector Machines*. Portland State University; 2010.

22. Carberry JC, Jordan AS, White DP, Wellman A, Eckert DJ. Upper airway collapsibility (Pcrit) and pharyngeal dilator muscle activity are sleep stage dependent. *Sleep*. 2016;39(3):511–521. doi:10.5665/sleep.5516

23. Strohl KP, Butler JP, Malhotra A. Mechanical properties of the upper airway. *Compr Physiol*. 2012;2(3):1853–1872.

24. Yadollahi A, Moussavi Z. Automatic breath and snore sounds classification from tracheal and ambient sounds recordings. *Med Eng Phys*. 2010;32(9):985–990. doi:10.1016/j.medengphy.2010.06.013

25. Yadollahi A, Montazeri A, Azarbarzin A, Moussavi Z. Respiratory flow–sound relationship during both wakefulness and sleep and its variation in relation to sleep apnea. *Ann Biomed Eng*. 2013;41 (3):537–546. doi:10.1007/s10439-012-0692-3

26. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. *Sleep*. 2007;30(10):1362–1369. doi:10.1093/sleep/30.10.1362

27. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. *Sleep Med Rev*. 2011;15(4):259–267. doi:10.1016/j.smrv.2010.10.001

28. Aktaruzzaman M, Rivolta MW, Karmacharya R, et al. Performance comparison between wrist and chest actigraphy in combination with heart rate variability for sleep classification. *Comput Biol Med*. 2017;89:212–221. doi:10.1016/j.compbiomed.2017.08.006

29. Devot S, Dratwa R, Naujokat E Sleep/wake detection based on cardiorespiratory signals and actigraphy. Paper presented at: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology; 2010.

30. Lewicke A, Sazonov E, Schuckers S Sleep-wake identification in infants: heart rate variability compared to actigraphy. Paper presented at: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2004.

31. Kalkbrenner C, Brucher R, Kesztyüs T, Eichenlaub M, Rottbauer W, Scharnbeck D. Automated sleep stage classification based on tracheal body sound and actigraphy. *GMS German Med Sci*. 2019;17.

32. Buratti L, Viticchi G, Falsetti L, et al. Vascular impairment in Alzheimer's disease: the role of obstructive sleep apnea. *J Alzheimers Dis*. 2014;38(2):445–453. doi:10.3233/JAD-131046

33. Ebrahimi F, Setarehdan S-K, Ayala-Moyeda J, Nazeran H. Automatic sleep staging using empirical mode decomposition, discrete wavelet transform, time-domain, and nonlinear dynamics features of heart rate variability signals. *Comput Methods Programs Biomed*. 2013;112(1):47–57. doi:10.1016/j.cmpb.2013.06.007

34. Bhattacherjee A, Chau N, Sierra CO, et al. Relationships of job and some individual characteristics to occupational injuries in employed people: a community-based study. *J Occup Health*. 2003;45(6):382–391. doi:10.1539/joh.45.382

35. Haraldsson P-O, Carenfelt C, Diderichsen F, Nygren Å, Tingvall C. Clinical symptoms of sleep apnea syndrome and automobile accidents. *Orl*. 1990;52(1):57–62. doi:10.1159/000276104

36. Leung RS, Douglas Bradley T. Sleep apnea and cardiovascular disease. *Am J Respir Crit Care Med*. 2001;164(12):2147–2165. doi:10.1164/ajrccm.164.12.2107045

37. Buratti L, Luzzi S, Petrelli C, et al. Obstructive sleep apnea syndrome: an emerging risk factor for dementia. *CNS Neurol Disord Drug Targets*. 2016;15(6):678–682. doi:10.2174/1871527315666160518123930