# Apnea–Hypopnea Index Prediction Using Electrocardiogram Acquired During the Sleep-Onset Period

Da Woon Jung, Su Hwan Hwang, Yu Jin Lee, Do-Un Jeong, and Kwang Suk Park*, *Senior Member, IEEE*

***Abstract*—The most widely used methods for predicting obstructive sleep apnea are based on clinical or anatomico-functional features. To improve exactitude in obstructive sleep apnea screening, this study aimed to devise a new predictor of apnea–hypopnea index. We hypothesized that less irregular respiration cycles would be observed in the patients with more severe obstructive sleep apnea during the sleep-onset period. From each of the 156 and 70 single-lead electrocardiograms collected from the internal polysomnographic database and from the Physionet Apnea-ECG database, respectively, the 150-s sleep-onset period was determined and the respiration cycles during this period were detected. Using the coefficient of variation of the respiration cycles, obtained from the internal dataset, as a predictor, the apnea–hypopnea index predictive model was developed through regression analyses and k-fold cross-validations. The apnea–hypopnea index predictability of the regression model was tested with the Physionet Apnea-ECG database. The regression model trained and validated from the 143 and 13 data, respectively, produced an absolute error (mean $\pm$ SD) of $3.65 \pm 2.98$ events/h and a Pearson's correlation coefficient of 0.97 ($P < 0.01$) between the apnea–hypopnea index predictive values and the reference values for the 70 test data. The new predictor of apnea–hypopnea index has the potential to be utilized in making more reasoned clinical decisions on the need for formal diagnosis and treatment of obstructive sleep apnea. Our study is the first study that presented the strategy for providing a reliable apnea–hypopnea index without overnight recording.**

***Index Terms*—Autonomic dysfunction, obstructive sleep apnea (OSA), respiration cycle, sleep-onset.**

## I. Introduction

OBSTRUCTIVE sleep apnea (OSA) is the most prevalent type of sleep-related breathing disorder. OSA is characterized by recurrence of complete or partial upper airway obstruction during sleep, intermittently causing cessations of breathing (apneas), or reductions in airflow (hypopneas) notwithstanding the effort to inhale [2]. With increasing number of obesity and elderly population, OSA has been considered a serious clinical concern. OSA is an independent contributor to excessive daytime somnolence, neurocognitive deficits, tiredness, irritability, and depression [4]. Furthermore, undiagnosed and untreated OSA is a risk factor for life-threatening complications such as cardiovascular and neurovascular diseases, metabolic disorders, and altered immune function [8]. Within the postoperative environment, more frequent cardiorespiratory complications were reported in the unattended OSA patients as compared to the patients treated for OSA prior to their surgery [11].

The gold standard for diagnosing OSA is an overnight in-laboratory polysomnography (PSG). However, the use of PSG has been limited due to its cumbersome method, and cost and resource constraints. Thus, the issue of developing a simple and reliable method for diagnosing OSA without PSG has been raised. Moreover, it has been reported that a large proportion of subjects referred for overnight PSG actually did not have severe OSA [12]. It would be useful if a predictive tool is developed for separating low risk OSA patients from high risk ones who need PSG and active therapeutic intervention [16]. Prediction of OSA would be also helpful in enabling anesthesiologists and surgeons to reduce postoperative complications related to OSA [1], [3], [11].

One of the most widely used methods for predicting OSA is based on clinical features including demographic (e.g., age and gender) and anthropometric [e.g., body mass index (BMI) and neck circumference] variables, OSA symptoms, and comorbid conditions (e.g., snoring, hypertension, and daytime symptoms such as fatigue and excessive daytime sleepiness) [1]–[3], [5]–[7], [13], [15], [16], [18]–[27]. Information on these variables are generally obtained through questionnaires and/or doing an interview. Another conventional method is based on anatomical and functional characteristics of OSA patients' upper airway in terms of acoustic speech signal, breathing sound, and airway pressure signal [9], [10], [12], [14], [17], [28], [29]. Both of these two methods have revealed limitations. The "clinical features" model produces less validity because much of the data relies on the patient's subjective report. The "anatomico-functional characteristics" model proves to be less usable because of demand on the patients to perform specific tasks such as pronunciation of vowels, deep breathing, and negative expiratory pressure test. Furthermore, all the previous studies on predicting OSA presented incomplete results in that they just

Manuscript received November 25, 2015; revised April 10, 2016; accepted April 12, 2016. Date of publication April 14, 2016; date of current version January 18, 2017. This work was financially supported by Samsung Electronics, Inc. *Asterisk indicates in corresponding author.*

D. W. Jung and S. H. Hwang are with the Interdisciplinary Program for Biomedical Engineering, Seoul National University Graduate School.

Y. J. Lee and D.-U. Jeong are with the Department of Psychiatry and Behavioral Science, Seoul National University College of Medicine, and the Center for Sleep and ChronobiologySeoul National University Hospital.

*K. S. Park is with the Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul 110-799, South Korea (e-mail: pks@bmsil.snu.ac.kr).

Digital Object Identifier 10.1109/TBME.2016.2554138

provide a dichotomously classified outcome without respect to multiple values of apnea–hypopnea index (AHI), a quantitative evaluation measure for OSA.

This study aimed to devise a new effectual predictor of AHI using a neurophysiological approach. We hypothesized that the difference in AHI would be reflected in the irregularity of the respiration cycles during the sleep-onset period. The hypothesis was developed based on the following studies: during the sleep-onset period, repeated fluctuations between sleep and arousal states are observed and transition from arousal to sleep (or from sleep to arousal) is, respectively, associated with relative hypoventilation (or hyperventilation) causing an increase (or decrease) in arterial $pCO_2$ (carbon dioxide partial pressure) [30]. The arterial $pCO_2$ is known as a stimulator of the autonomic nervous system (ANS). In non-OSA subjects, active and sensitive responses of the ANS to the alternate stimulations during the sleep-onset period would result in irregular respiration cycles. On the other hand, OSA patients' unresponsive ANS to the alternate stimulations during the sleep-onset period due to autonomic dysfunction [31] would cause less irregular respiration cycles. It was also possible to hypothesize that more severe OSA would be related with more severe autonomic dysfunction producing almost regular respiration cycles during the sleep-onset period.

## II. Materials and Methods

### A. Datasets

Two different datasets were used in this study. The first one is the internal dataset that consists of the polysomnographic recordings collected from the 156 subjects who had undergone diagnostic overnight PSG because of suspected OSA at the Center for Sleep and Chronobiology of Seoul National University Hospital between January 2012 and August 2012. The polysomnographic recordings were scored by certified sleep technologists and verified by sleep physicians in accordance with the 2007 American Academy of Sleep Medicine (AASM) manual [32]. The number of apneas and hypopneas per hour of sleep were defined as apnea index (AI) and hypopnea index (HI), respectively. The AHI value was determined by summing AI and HI. According to the AHI value, the subjects were classified into non-OSA (AHI $<$ 5 events/h) subjects, mild OSA (5 $\leq$ AHI $<$ 15 events/h), moderate OSA (15 $\leq$ AHI $<$ 30 events/h), and severe OSA (AHI $\geq$ 30 events/h) patients [32]. Gender ratio-, age-, and BMI-matched the four groups consisted of 48 (non-OSA), 47 (mild OSA), 34 (moderate OSA), and 27 (severe OSA) individuals, respectively. All subjects were of Mongoloid origin. Exclusion criteria for all groups were: 1) the presence of other psychiatric or medical condition known to be associated with ANS and 2) the use of medications known to influence on sleep or ANS function. Demographic and anthropometric characteristics and sleep parameters of the subjects are summarized in Table I. Among the physiological data routinely acquired during PSG with a sampling frequency of 250 Hz, the single-lead electrocardiogram (ECG) (lead II) and the nasal pressure measured using a nasal cannula/pressure transducer (PTAF 2, Pro-Tech, Woodinville, WA, USA) were collected from each subject. The Institutional Review Board of Seoul National University Hospital approved this retrospective study (IRB No. 1405-075-581).

The second dataset includes the 70 single-lead ECG in the Physionet Apnea-ECG database [33]. The mean and SD of the ECG acquisition duration are 491.8 and 31.6 min, respectively (Min.–Max.: 401–578 min). The electrocardiographic recordings were acquired from 32 subjects (25 males) with an age (mean $\pm$ SD) of $43.8 \pm 11.1$ years (Min.–Max.: 27–63 years) and a BMI value (mean $\pm$ SD) of $28.1 \pm 7.1$ kg/m$^2$ (Min.–Max.: 19.2–45.3 kg/m$^2$). The dataset consists of 23, 5, 11, and 31 recordings with non-OSA, mild, moderate, and severe OSA severity, respectively. An AHI value (mean $\pm$ SD) of $28.0 \pm 27.5$ events/h (Min.–Max.: 0–93.5 events/h) was presented from the dataset. Each electrocardiographic recording, originally sampled at 100 Hz, was resampled at 250 Hz.

In both datasets, each AHI value was designated as $\mathrm{AHI}_{\mathrm{Refer}}$ denoting the AHI reference value. $\mathrm{AHI}_{\mathrm{Refer}}$ of zero was substituted with 0.0416 derived by the lower limit on the AHI for the maximum sleep duration. In other words, if there was at least an apneic or hypopneic event during a maximum sleeping time of 24 h, the AHI was obtained by dividing one (event) into 24 (h), yielding a value of 0.04166... . Thus, an AHI value of 0.0416 can connote that there was no apneic or hypopneic event during sleep. This substitution allowed natural log transformation of $\mathrm{AHI}_{\mathrm{Refer}}$.

### B. Data Analysis

***1) Sleep-Onset Period Detection:*** In the 2007 AASM manual, sleep-onset is defined as the time when reach first epoch (30-sec segment) scored as any sleep stage other than wakefulness. The gold standard method for classifying sleep stages requires electroencephalographic, electrooculargraphic, and electromyographic monitoring. In this study, sleep-onset was determined using ECG to improve usability. The algorithm for identifying sleep-onset was designed based on the effect of sleep involving significantly slower heart rate (HR) compared with the HR observed during wakefulness [34]. After detecting R peaks on ECG, HR values could be obtained from RR intervals. Fig. 1 represents a flowchart for determining sleep-onset epoch. Prior to applying the process in Fig. 1, the threshold (TH) was calculated as follows:

$$\mathrm{TH} = \mathrm{HR2}_{\mathrm{avg}} - 1.96 \times \mathrm{HR2}_{\mathrm{SD}}$$

where $\mathrm{HR2}_{\mathrm{avg}}$ and $\mathrm{HR2}_{\mathrm{SD}}$ are the mean and standard deviation of the HR samples collected after removing the outliers (values that fall outside of 1.96$*$SD from the mean) among the HR samples acquired for the first 2 min of recording (from first to fourth epoch). Starting from the fifth epoch, the HR samples in each consecutive epoch were examined one-by-one until a sleep-onset epoch was identified. When the number of HR samples, successively not exceeding the TH, was greater than half of the total number of HR samples in the epoch, the epoch was determined as a sleep-onset epoch. The sleep-onset period was defined as the period including a sleep-onset epoch, and the two previous and the two next epochs (if sleep-onset epoch is a sixth epoch, the sleep-onset period includes from fourth to eighth epoch ranging from 90 to 240 s).

***2) Respiration Cycles Detection:*** The ECG-derived respiration (EDR) signals were acquired by summing the amplitudes in an R-peak-centered window of 100 ms (from 50 ms

TABLE I
DEMOGRAPHIC AND ANTHROPOMETRIC CHARACTERISTICS AND SLEEP PARAMETERS OF INTERNAL DATASET

| | Non-OSA | Mild OSA | Moderate OSA | Severe OSA |
|---|---|---|---|---|
| Sample size (male/female) | 48 (36/12) | 47 (35/12) | 34 (25/9) | 27(20/7) |
| Age (years) | 40.5 ± 18.1 | 42.0 ± 11.7 | 40.9 ± 10.8 | 41.6 ± 12.6 |
| BMI (kg/m$^2$) | 24.5 ± 2.9 | 25.1 ± 3.2 | 24.8 ± 2.7 | 24.7 ± 3.0 |
| AHI (events/h) | 1.5 ± 1.3 | 10.7 ± 2.8 | 22.5 ± 4.9 | 43.1 ± 8.6 |
| TRT (min) | 499.5 ± 44.6 | 494.3 ± 35.8 | 496.2 ± 38.2 | 489.7 ± 39.0 |
| SL (min) | 12.6 ± 15.2 | 10.2 ± 9.8 | 10.5 ± 7.4 | 7.8 ± 6.9 |
| SE (%) | 88.5 ± 10.4 | 86.8 ± 10.2 | 87.7 ± 9.5 | 86.9 ± 9.7 |

OSA, obstructive sleep apnea; BMI, body mass index; AHI, apnea–hypopnea index; TRT, total recording time; SL, sleep-onset latency; SE, sleep efficiency.
Subjects were classified according to their AHI value into non-OSA (AHI $<$ 5 events/h), mild OSA (5 $\leq$ AHI $<$ 15 events/h), moderate OSA (15 $\leq$ AHI $<$ 30 events/h), and severe OSA (AHI $\geq$ 30 events/h) groups.
Data are presented as the mean ± SD.

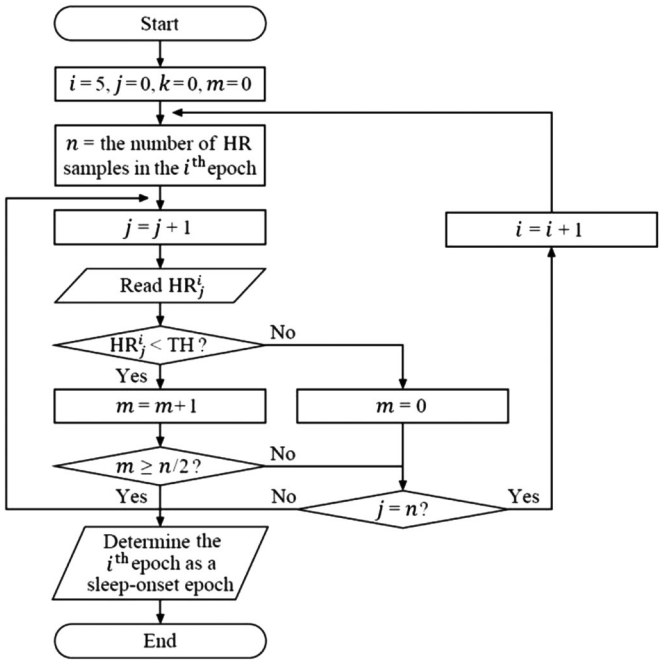


Fig. 1. Flowchart for determining sleep-onset epoch. HR and TH denote heart rate and threshold, respectively. $\mathrm{HR}_j^i$ indicates the value of $j$th HR sample in the $i$th epoch.

before R peak to 50 ms after R peak) [35]. In the sinusoidal waveforms of the EDR signals, a time span between two successive maximum peaks was defined as an EDR cycle. The coefficient of variation (CV), defined as the ratio of the standard deviation to the mean and expressed as a percentage, was used for quantifying irregularity. The CV of EDR cycles obtained during the sleep-onset period was designated as $\mathrm{EDR_{CV}}$.

***3) Regression Analysis:*** The regression analyses along with k-fold cross-validation tests were performed with the internal dataset using Curve Expert Professional software (v.2.0.4, www.curveexpert.net, USA). In the regression modeling, the dependent and explanatory variables were the natural log-transformed $\mathrm{AHI_{Refer}}$ and $\mathrm{EDR_{CV}}$, respectively, and a significant probability level of 95% was applied.

The AHI predictability of the regression model that provided the smallest k-fold cross-validation error was tested with the Physionet Apnea-ECG database.

## III. RESULTS

Because of the absence of physiological data required to identify reference sleep-onset period and respiration cycles in the Physionet Apnea-ECG database, the performance on detecting the sleep-onset period and respiration cycles was assessed only from the internal dataset.

All statistical tests were conducted using SPSS statistics software (v.21.0, SPSS Inc., Chicago, IL, USA).

### A. Sleep-Onset Period Detection Performance

From any PSG recording used in this study, rapid eye movement sleep was not observed during the sleep-onset period.

The quantitative evaluation for the difference between the sleep-onset period derived from the PSG result and that detected by our method was conducted using a correct detection rate (CDR). A CDR was defined as a percentage of the number of epochs correctly detected as the epochs composing the sleep-onset period. As a result, a CDR (mean ± SD) of $78.0 \pm 24.2\%$ was obtained, which denotes an epoch-shifted detection on average.

### B. Respiration Cycles Detection Performance

For the dataset consisted of the respiration cycles acquired from the nasal pressure signal and EDR cycles during sleep-onset period in each PSG recording, a mean comparison test and a correlation analysis were carried out. The significant differences were not observed in all datasets (Shapiro-Wilk normality test $P > 0.05$ and paired sample t-test $P > 0.05$, or Shapiro-Wilk normality test $P < 0.05$ and Wilcoxon test $P > 0.05$). For all datasets, a Pearson's correlation coefficient (mean ± SD) of 0.96 ± 0.02 was reported (all $P < 0.01$).

Between the CV computed from the nasal pressure signal-derived respiration cycles and $\mathrm{EDR_{CV}}$, an absolute error (mean ± SD) of 1.27 ± 1.01% was exhibited. Fig. 2 displays the respiration cycles obtained from two subjects belonging to different OSA severity groups. Both subjects exhibited high similarity

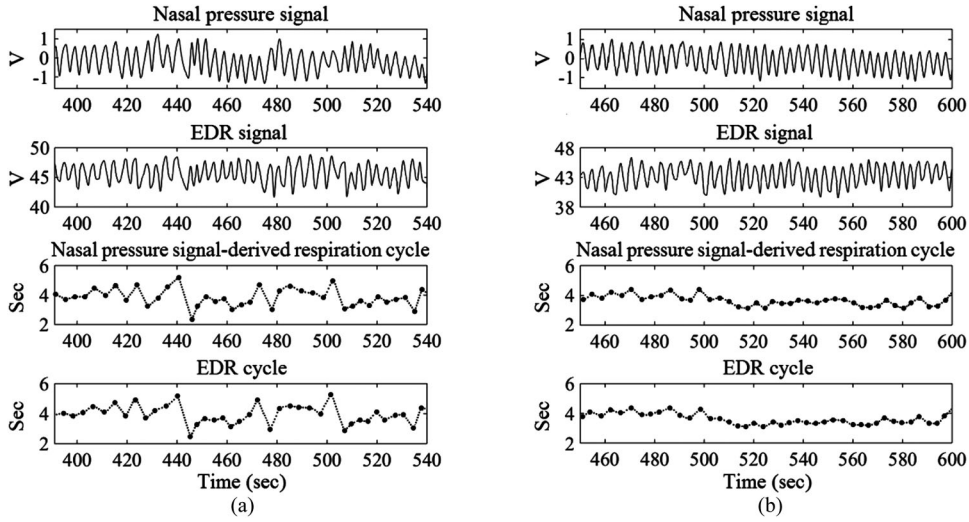


Fig. 2. Nasal pressure and EDR signals, and nasal pressure signal-derived respiration and EDR cycles obtained by computing the time spans between two successive maximum peaks in nasal pressure and EDR signals, respectively. The time of data acquisition (*x*-axis) corresponds to the 150-s sleep-onset period determined by our method. (a) An illustrative case of the non-OSA subject. (b) An illustrative case of the severe OSA patient.

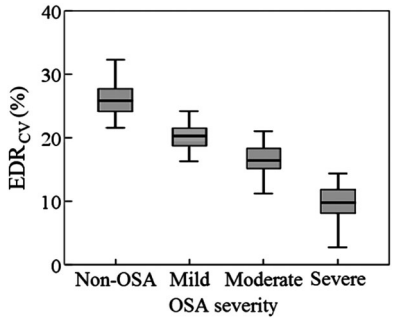


Fig. 3. Box-and-whisker plot showing the median (bold horizontal line in the box) with the twenty-fifth to seventy-fifth percentiles and the highest and lowest value within 1.5 times of the interquartile range (whiskers outside the box). There were significant differences in the CV of EDR cycles during the sleep-onset period, $EDR_{CV}$, according to OSA severity (one-way ANOVA, Levene's variance homogeneity test: $P < 0.001$, Dunnett T3 Post-hoc test: all $P < 0.001$).

between the nasal pressure signal-derived respiration and EDR cycles. During the 150-s sleep-onset period, the non-OSA subject showed relatively irregular respiration cycles [see Fig. 2(a); the CVs computed from the nasal pressure signal-derived respiration and EDR cycles were 25.5 and 24.7%, respectively]. On the other hand, less irregular respiration cycles were observed in the severe OSA patient [see Fig. 2(b); the CVs computed from the nasal pressure signal-derived respiration and EDR cycles were 7.6 and 8.0%, respectively].

### C. AHI Prediction Performance

The difference in $EDR_{CV}$ according to OSA severity is represented by the box-and-whisker plot in Fig. 3. A smaller mean of $EDR_{CV}$ was observed in the more severe OSA group. Concretely, the $EDR_{CV}$ values (mean $\pm$ SD) were computed at $25.9 \pm 2.6\%$, $20.1 \pm 1.6\%$, $16.5 \pm 2.2\%$, and $9.4 \pm 3.2\%$ in the non-, mild, moderate, and severe OSA groups, respectively.

TABLE II
CROSS-VALIDATION STATISTICS FOR AHI PREDICTIVE MODELS

| Cross-validation | MAE (events/h) | | Pearson's correlation coefficient | |
|---|---|---|---|---|
| | Training set | Validation set | Training set | Validation set |
| 2-fold | $3.48 \pm 0.50$ | $3.54 \pm 0.37$ | $0.97 \pm 0.01$ | $0.97 \pm 0.01$ |
| 4-fold | $3.52 \pm 0.20$ | $3.52 \pm 0.70$ | $0.97 \pm 0.01$ | $0.97 \pm 0.02$ |
| 6-fold | $3.51 \pm 0.11$ | $3.56 \pm 0.65$ | $0.97 \pm 0.00$ | $0.97 \pm 0.01$ |
| 12-fold | $3.51 \pm 0.06$ | $3.57 \pm 0.81$ | $0.97 \pm 0.00$ | $0.97 \pm 0.02$ |

MAE, mean absolute error.
Data are presented as the mean $\pm$ SD.
For all Pearson's correlation coefficients, $P < 0.01$.

One-way ANOVA followed by Dunnett's multiple comparison test revealed the significance of differences in $EDR_{CV}$ among the OSA severity groups (all $P < 0.001$).

The statistical results of the regression analyses along with the 2-, 4-, 6-, and 12-fold cross-validation tests are summarized in Table II. The statistics on mean absolute error (MAE) and Pearson's correlation coefficient in Table II were computed from the $AHI_{Estim}$, denoting AHI estimate, and the $AHI_{Refer}$. Prior to arranging $AHI_{Estim}$, the AHI estimate of the value less than 0.0416 was substituted with zero.

Fig. 4 represents the regression model that exhibited the smallest average of the MAEs for the training and validation sets. The best-fitting curve (see Fig. 4, solid line) to the 143 training data (see Fig. 4, hollow circles) with an AHI value (mean $\pm$ SD) of $18.01 \pm 19.32$ events/h produced an MAE of 2.79 events/h for the remaining 13 validation data (see Fig. 4, filled triangles) with an AHI value (mean $\pm$ SD) of $20.62 \pm 23.27$ events/h.

For the 70 test data, the regression model (see Fig. 4, solid line) provided an absolute error (mean $\pm$ SD) of $3.65 \pm 2.98$ events/h and a Pearson's correlation coefficient of 0.97 ($P < 0.01$) between the $AHI_{Estim}$ and the $AHI_{Refer}$. Fig. 5 displays

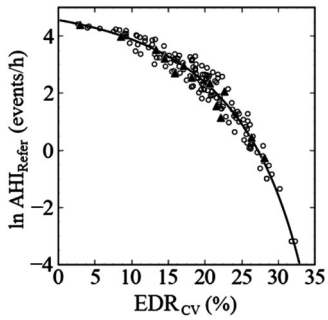


Fig. 4. Scatter plot showing the relationship between the natural log-transformed AHI reference values reported from PSG ($\mathrm{AHI_{Refer}}$) and the CV of EDR cycles during the sleep-onset period ($\mathrm{EDR_{CV}}$) (training data, hollow circles; validation data, filled triangles). The best-fitting curve (solid line) to the training data was derived from the Exponential Association function.

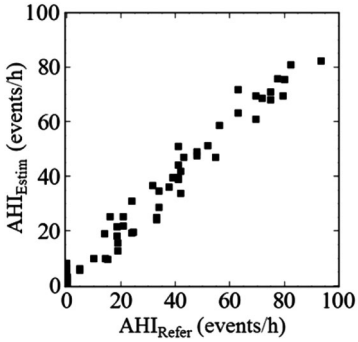


Fig. 5. Scatter plot showing the relationship between the AHI estimates provided by our method ($\mathrm{AHI_{Estim}}$) and the AHI reference values reported from PSG ($\mathrm{AHI_{Refer}}$) for the test data.

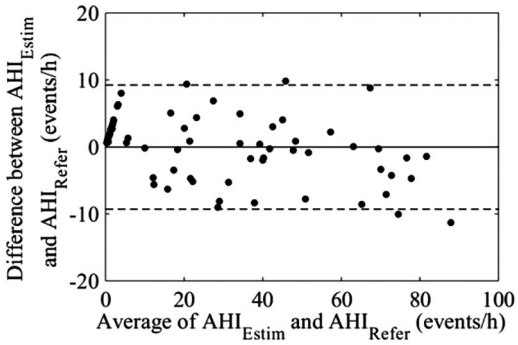


Fig. 6. Bland–Altman plot of the differences between the AHI estimates obtained by our method ($\mathrm{AHI_{Estim}}$) and the AHI reference values reported from PSG ($\mathrm{AHI_{Refer}}$) against the averages for the test data. The solid line represents the mean of differences between the $\mathrm{AHI_{Estim}}$ and $\mathrm{AHI_{Refer}}$, and the dashed lines denote the 95% limits of agreement ($\pm$ 2*SD of the differences).

the scatter plot of the $\mathrm{AHI_{Estim}}$ and the $\mathrm{AHI_{Refer}}$. Fig. 6 shows the Bland–Altman plot of the differences between the $\mathrm{AHI_{Estim}}$ and the $\mathrm{AHI_{Refer}}$ against the averages. The mean of differences (see Fig. 6, solid line) was computed at –0.03 events/h and the 95% limits of agreement (see Fig. 6, dashed lines) were observed from –9.30 to 9.24 events/h. Table III summarizes the OSA diagnostic performance at AHI cutoffs of ≥ 5, 10, 15, 20, 25, and 30 events/h. The averages for the five different AHI cutoffs were

TABLE III
OSA DIAGNOSTIC PERFORMANCE FOR TEST SET

| | AHI cutoff value (events/h) | | | | | |
|---|---|---|---|---|---|---|
| | ≥ 5 | ≥ 10 | ≥15 | ≥ 20 | ≥ 25 | ≥ 30 |
| Sensitivity (%) | 100.0 | 93.3 | 95.2 | 91.7 | 93.5 | 90.3 |
| Specificity (%) | 87.0 | 100.0 | 96.4 | 94.1 | 92.3 | 97.4 |
| PPV (%) | 94.0 | 100.0 | 97.6 | 94.3 | 90.6 | 96.6 |
| NPV (%) | 100.0 | 89.3 | 93.1 | 91.4 | 94.7 | 92.7 |
| Accuracy (%) | 95.7 | 95.7 | 95.7 | 92.9 | 92.9 | 94.3 |
| Cohen's kappa coefficient | 0.90 | 0.91 | 0.91 | 0.86 | 0.86 | 0.88 |
| ROC-AUC | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |

AHI, apnea–hypopnea index; PPV, positive predictive value; NPV, negative predictive value; ROC-AUC, area under the receiver operating characteristics curve.
Cohen's kappa coefficients ranging from 0.81 to 1.00 indicate almost perfect agreement.

TABLE IV
OSA SEVERITY CLASSIFICATION PERFORMANCE FOR TEST SET

| | | Estimated OSA severity | | | | |
|---|---|---|---|---|---|---|
| | | Non-OSA | Mild | Moderate | Severe | Total |
| OSA severity determined from PSG | Non-OSA | 20 | 3 | 0 | 0 | 23 |
| | Mild | 0 | 4 | 1 | 0 | 5 |
| | Moderate | 0 | 2 | 8 | 1 | 11 |
| | Severe | 0 | 0 | 3 | 28 | 31 |
| | Total | 20 | 9 | 12 | 29 | 70 |

OSA, obstructive sleep apnea; PSG, polysomnography.
OSA severity was categorized according to apnea-hypopnea index (AHI) into "Non-OSA" (AHI < 5 events/h), "Mild" (5 ≤ AHI < 15 events/h), "Moderate" (15 ≤ AHI < 30 events/h), and "Severe" (AHI ≥ 30 events/h).

sensitivity 94.0%, specificity 94.5%, positive predictive value (PPV) 95.5%, negative predictive value (NPV) 93.5%, accuracy 94.5%, and Cohen's kappa coefficient 0.89 corresponding to almost perfect agreement. The OSA severity classification performance was also assessed. Table IV, a 4-by-4 contingency table, includes the number of subjects classified according to their OSA severity determined by $\mathrm{AHI_{Refer}}$ (rows) and by $\mathrm{AHI_{Estim}}$ (columns). In Table IV, the diagonal elements correspond to the number of subjects correctly categorized into their OSA severity in each group, while the off-diagonal elements represent the number of subjects erroneously categorized. Among the 23 non-OSA subjects, three individuals were misclassified as mild OSA patients. In mild, moderate, and severe OSA patient groups, 1 of 5, 3 of 11, and 3 of 31 patients were misclassified, respectively. When a correct classification rate (CCR) was derived by a percentage of the number of subjects correctly categorized into their OSA severity, our method achieved a CCR of 85.7%. For the subjects who were overestimated their OSA severity, the differences between each $\mathrm{AHI_{Estim}}$ and the upper limit of each actual OSA severity criterion on AHI were explored and they showed an average value of 2.10 events/h. In the subjects who were underestimated their OSA severity, the average difference between each $\mathrm{AHI_{Estim}}$ and the lower limit of each actual OSA severity criterion on AHI was 3.58 events/h. No one was misclassified beyond his contiguous severity groups, for instance, any OSA patient with mild severity was not classified as a severe OSA patient by our method.

TABLE V
COMPARISON OF OSA PREDICTION PERFORMANCE WITH EXISTING METHODS

| AHI cutoff value (events/h) | Tool | Subject | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|
| ≥ 5 or > 5 | STOP questionnaire* [1] | 122 AHI > 5, 55 AHI ≤ 5 | 66 | 60 |
| | STOP-Bang questionnaire* [1] | 122 AHI > 5, 55 AHI ≤ 5 | 84 | 56 |
| | ASA questionnaire* [3] | 122 AHI > 5, 55 AHI ≤ 5 | 72 | 38 |
| | Wisconsin questionnaire* [5] | 38 AHI ≥ 5, 112 AHI < 5 | 95 | 64 |
| | Berlin questionnaire* [7] | 62 AHI > 5, 42 AHI < 5 | 86 | 95 |
| | Features from acoustic speech signal [10] | 67 AHI ≥ 5, 26 AHI < 5 | 79 / 84 | 83 / 86 |
| | Electrocardiogram* (this study) | 47 AHI ≥ 5, 23 AHI < 5 | 100 | 87 |
| ≥ 10 or > 10 | Self-developed clinical prediction model [13] | 189 AHI ≥ 10, 221 AHI < 10 | 94 | 28 |
| | Self-developed clinical prediction model* [15] | 104 AHI ≥ 10, 46 AHI < 10 | 99 | 80 |
| | Self-developed clinical prediction model [18] | 248 AHI ≥ 10, 122 AHI < 10 | 76–96 | 13–54 |
| | Features from acoustic speech signal [29] | 23 AHI > 10, 12 AHI ≤ 10 | 93 | 80 |
| | Electrocardiogram* (this study) | 45 AHI ≥ 10, 25 AHI < 10 | 93 | 100 |
| ≥ 15 or > 15 | Features from tracheal breath sound [12] | 22 AHI > 15, 30 AHI < 15 | 95 | 81 |
| | STOP questionnaire* [1] | 71 AHI > 15, 106 AHI ≤ 15 | 74 | 53 |
| | STOP-Bang questionnaire* [1] | 71 AHI > 15, 106 AHI ≤ 15 | 93 | 43 |
| | ASA questionnaire* [3] | 71 AHI > 15, 106 AHI ≤ 15 | 79 | 37 |
| | Wisconsin questionnaire* [21] | 54 AHI ≥ 15, 548 AHI < 15 | 87 | 40 |
| | Berlin questionnaire* [24] | 70 AHI ≥ 15, 30 AHI < 15 | 54 | 97 |
| | Clinical and biochemical predictors [6] | 36 AHI ≥ 15, 63 AHI < 15 | 89 | 81 |
| | Features from oronasal airway pressure [28] | 21 AHI ≥ 15, 20 AHI < 5 | 81 | 95 |
| | Electrocardiogram* (this study) | 42 AHI ≥ 15, 28 AHI < 15 | 95 | 96 |
| ≥ 30 or > 30 | STOP questionnaire* [1] | 39 AHI > 30, 138 AHI ≤ 30 | 79 | 49 |
| | STOP-Bang questionnaire* [1] | 39 AHI > 30, 138 AHI ≤ 30 | 100 | 37 |
| | ASA questionnaire* [2] | 39 AHI > 30, 138 AHI ≤ 30 | 87 | 36 |
| | Berlin questionnaire* [2] | 39 AHI > 30, 138 AHI ≤ 30 | 87 | 46 |
| | Clinical and biochemical predictors [6] | 23 AHI ≥ 30, 76 AHI < 30 | 96 | 71 |
| | Features from tracheal breath sound [9] | 13 AHI ≥ 30, 10 AHI < 30 | 89 | 85 |
| | Features from tracheal breath sound [12] | 15 AHI > 30, 17 AHI < 5 | 92 | 85 |
| | Negative expiratory pressure test [14] | 24 AHI ≥ 30, 24 AHI < 5 | 96 | 96 |
| | Features from tracheal breath sound* [17] | 58 AHI ≥ 30, 122 AHI ≤ 5 | 78 | 86 |
| | Electrocardiogram* (this study) | 31 AHI ≥ 30, 39 AHI < 30 | 90 | 97 |

AHI, apnea–hypopnea index; STOP, snoring, tiredness, observed apnea, and high blood pressure; STOP-Bang, STOP, BMI, age, neck circumference, and gender; ASA, American Society of Anesthesiologists.
The numbers shown before and after the forward slash correspond to the result for male and female, respectively.
The tool with * (an asterisk) was verified from the out-of-sample test and that without * was validated from the in-sample test.

## IV. DISCUSSION

In this study, we suggested a new method for predicting AHI by adopting a neurophysiological approach. During the sleep-onset period, less irregular respiration cycles were observed in the subjects with more severe OSA, which could be elucidated by autonomic dysfunction in OSA patients. The CV of EDR cycles during the 150-s sleep-onset period served as an effective predictor of the AHI.

In terms of OSA diagnostic performance, our method can be compared with previously suggested prediction methods. In Table V, sensitivities and specificities in OSA diagnosis reported by different prediction methods are summarized according to AHI cutoff values. Several studies were not reviewed because of insufficient data regarding the study subjects or reliability parameters. When there were two or more studies evaluating the OSA diagnostic performance of the identical prediction method, the results of the study presenting the highest average of sensitivity and specificity were included in Table V. In the screening of OSA using an AHI cutoff value of ≥ 5 or > 5 events/h, Wisconsin questionnaire (Berlin questionnaire) exhibited the highest sensitivity (specificity) of 95% (95%) with the specificity (sensitivity) of 64% (86%) [5], [7]. The highest sensitivity in identifying OSA using an AHI cutoff value of ≥ 10 events/h was reported as 99% with the highest specificity of 80% by an in-house-developed clinical prediction model [15]. The use of features extracted from acoustic speech signals also allowed 80% specificity with 93% sensitivity for an AHI cutoff value of > 10 events/h [29]. With regard to the prediction of moderate or severe OSA using an AHI cutoff value of ≥ 15 or > 15 events/h, the analysis on tracheal breath sound (Berlin questionnaire) exhibited the highest sensitivity (specificity) of 95% (97%) with the specificity (sensitivity) of 81% (54%) [12], [24]. The highest sensitivity (specificity) in severe OSA screening using an AHI cutoff value of ≥ 30 or > 30 events/h was reported as 100% (96%) by STOP-Bang questionnaire (negative expiratory pressure test) with the specificity (sensitivity) of 37% (96%) [1], [14]. For comparison, the OSA diagnostic performance provided by our method is also presented in the last row of each AHI cutoff value in Table V. In comparison with existing OSA prediction methods, our method exhibited generally high specificity along with sufficient sensitivity. The high specificity, denoting few false positives, ensures high applicability of our method as a diagnostic measure for OSA enabling efficient utilization of sleep clinic resources, thereby reducing healthcare costs. The sufficient sensitivity provided by our method would be potentially effective to reduce the risk of OSA-related postoperative complications by helping early detection of patients who need careful perioperative management.

Although many existing OSA prediction methods represented their promising performance on OSA screening, their usefulness

has a limitation in that the conjecture classified dichotomously is not sufficient as a corroborating evidence in clinical decision-making associated with the need and urgency of PSG and treatment. To the best of our knowledge, this is the only study that provided the AHI predictive values without overnight recording and verified its prediction ability for the multicategorized OSA severity. It should be noted that all the subjects misclassified their OSA severity by our method were found in "gray zone" where included the AHI values around a discrete boundary between contiguous severity groups.

One limitation of our study is associated with the methodology that depends on ECG. Satisfying the physiological and clinical conditions as well as the technological condition related with good-quality ECG is a prerequisite for reliable detection of sleep-onset period and effective acquisition of EDR. Another limitation of our study is associated with the usability. Although this study was based on the conventional electrocardiography, the technological advances in the electrocardiography devices with respect to their size and communication features facilitate their connection to mobile devices, thus simple systems can be required for ECG monitoring. An application of simplified ECG acquisition systems to our method can contribute to improving the potential for predicting night-to-night variability of AHI in out-of-sleep laboratory environment that allows a long-term follow-up study of OSA patients.

## V. Conclusion

OSA is associated with the abnormal autonomic nervous function. We hypothesized that the unresponsive ANS of OSA patients to the alternate stimulations during the sleep-onset period would cause regular respiration cycles. This hypothesis was validated from the significantly smaller mean of CV of EDR cycles, acquired during the sleep-onset period, in the more severe OSA patients. Using the CV of EDR cycles acquired during the sleep-onset period as a predictor, the model that can provide reliable AHI predictive values was developed. This study has the potential to address the growing need for an effectual OSA prediction measure.

## References


[1] F. Chung *et al.*, "STOP questionnaire: A tool to screen patients for obstructive sleep apnea," *Anesthesiology*, vol. 108, pp. 812–821, 2008.

[2] R. J. Farney *et al.*, "The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: Relation to polysomnographic measurements of the apnea/hypopnea index," *J. Clin. Sleep Med.*, vol. 7, pp. 459–465, 2011.

[3] F. Chung *et al.*, "Validation of the Berlin questionnaire and American Society of Anesthesiologists checklist as screening tools for obstructive sleep apnea in surgical patients," *Anesthesiology*, vol. 108, pp. 822–830, 2008.

[4] R. D. Chervin, "Sleepiness, fatigue, tiredness, and lack of energy in obstructive sleep apnea," *Chest*, vol. 118, pp. 372–379, 2000.

[5] S. K. Sharma *et al.*, "Prevalence and risk factors of obstructive sleep apnea syndrome in a population of Delhi, India," *Chest*, vol. 130, pp. 149–156, 2006.

[6] J. B. Dixon and L. M. Schachter, "Predicting sleep apnea and excessive day sleepiness in the severely obese: Indicators for polysomnography," *Chest*, vol. 123, pp. 1134–1141, 2003.

[7] S. K. Sharma *et al.*, "Validation of the modified Berlin questionnaire to identify patients at risk for the obstructive sleep apnoea syndrome," *Indian J. Med. Res.*, vol. 124, pp. 281–290, 2006.

[8] T. Kendzerska *et al.*, "Untreated obstructive sleep apnea and the risk for serious long-term adverse outcomes: A systematic review," *Sleep Med. Rev.*, vol. 18, pp. 49–59, 2014.

[9] J. Solà-Soler *et al.*, "Identification of obstructive sleep apnea patients from tracheal breath sound analysis during wakefulness in polysomnographic studies," in *Proc. IEEE 36th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2014, pp. 4232–4235.

[10] E. Goldshtein *et al.*, "Automatic detection of obstructive sleep apnea using speech signals," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 5, pp. 1373–1382, Dec. 2010.

[11] S. A. Chung *et al.*, "A systemic review of obstructive sleep apnea and its implications for anesthesiologists," *Anesth. Analg.*, vol. 107, pp. 1543–1563, 2008.

[12] A. Montazeri *et al.*, "Assessment of obstructive sleep apnea and its severity during wakefulness," *Ann. Biomed. Eng.*, vol. 40, pp. 916–924, 2012.

[13] S. Viner *et al.*, "Are history and physical examination a good screening test for sleep apnea?" *Ann. Intern Med.*, vol. 115, pp. 356–359, 1991.

[14] S. Romano *et al.*, "Upper airway collapsibility evaluated by a negative expiratory pressure test in severe obstructive sleep apnea," *Clinics*, vol. 66, pp. 567–572, 2011.

[15] S. D. Kirby *et al.*, "Neural network prediction of obstructive sleep apnea from clinical criteria," *Chest*, vol. 116, pp. 409–415, 1999.

[16] A. Abrishami *et al.*, "A systematic review of screening questionnaires for obstructive sleep apnea," *Can. J. Anaesth.*, vol. 57, pp. 423–438, 2010.

[17] A. M. Cameron and Z. Moussavi, "A novel expert classifier approach to pre-screening obstructive sleep apnea during wakefulness," in *Proc. IEEE 36th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2014, pp. 4236–4239.

[18] J. A. Rowley *et al.*, "The use of clinical prediction formulas in the evaluation of obstructive sleep apnea," *Sleep*, vol. 23, pp. 929–942, 2000.

[19] B. D. Crocker *et al.*, "Estimation of the probability of disturbed breathing during sleep before a sleep study," *Amer. Rev. Respir. Dis.*, vol. 142, pp. 14–18, 1990.

[20] V. Hoffstein and J. Szalai, "Predictive value of clinical features in diagnosing obstructive sleep apnea," *Sleep*, vol. 16, pp. 118–122, 1993.

[21] T. Young *et al.*, "The occurrence of sleep-disordered breathing among middle-aged adults," *N. Engl. J. Med.*, vol. 328, pp. 1230–1235, 1993.

[22] W. W. Flemons *et al.*, "Likelihood ratios for a sleep apnea clinical prediction rule," *Amer. J. Respir. Crit. Care Med.*, vol. 150, pp. 1279–1285, 1994.

[23] D. F. Dinges, "A survey screen for prediction of apnea," *Sleep*, vol. 18, pp. 158–166, 1995.

[24] N. C. Netzer *et al.*, "Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome," *Ann. Intern. Med.*, vol. 131, pp. 485–491, 1999.

[25] S. D. Ross *et al.*, "Systematic review and meta-analysis of the literature regarding the diagnosis of sleep apnea," *Sleep*, vol. 23, pp. 519–532, 2000.

[26] W. H. Tsai *et al.*, "A decision rule for diagnostic testing in obstructive sleep apnea," *Amer. J. Respir. Crit. Care Med.*, vol. 167, pp. 1427–1432, 2003.

[27] J. Rodsutti *et al.*, "A clinical decision rule to prioritize polysomnography in patients with suspected sleep apnea," *Sleep*, vol. 27, pp. 694–699, 2004.

[28] P. Caseiro *et al.*, "Screening of obstructive sleep apnea using Hilbert–Huang decomposition of oronasal airway pressure recordings," *Med. Eng. Phys.*, vol. 32, pp. 561–568, 2010.

[29] M. Kriboy *et al.*, "Detection of obstructive sleep apnea in awake subjects by exploiting body posture effects on the speech signal," in *Proc. IEEE 36th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2014, pp. 4224–4227.

[30] J. Trinder *et al.*, "Respiratory instability during sleep onset," *J. Appl. Physiol.*, vol. 73, pp. 2462–2469, 1992.

[31] U. Wiklund *et al.*, "Autonomic cardiovascular regulation in patients with obstructive sleep apnoea: A study based on spectral analysis of heart rate variability," *Clin. Physiol. Funct. Imag.*, *vol*. 20, pp. 234–241, 2000.

[32] C. Iber *et al.*, *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 1 ed. Westchester, IL, USA: Amer. Acad. Sleep Medicine, 2007.

[33] T. Penzel *et al.*, "The apnea-ECG database," *Comput. Cardiol.*, *vol*. 27, pp. 255–258, 2000.

[34] D. W. Jung *et al.*, "Nocturnal awakening and sleep efficiency estimation using unobtrusively measured ballistocardiogram," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 1, pp. 131–138, Aug. 2013.

[35] R. Atri and M. Mohebbi, "Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal," *Physiol. Meas.*, vol. 36, pp. 1963–1980, 2015.


**Authors'** photographs and biographies not available at the time of publication.