SCIENTIFIC INVESTIGATIONS

Electrocardiogram Recording as a Screening Tool for Sleep Disordered Breathing

Conor Heneghan, Ph.D.[1,2]; Philip de Chazal, Ph.D.[2]; Silke Ryan, M.D., Ph.D.[3]; Chern-Pin Chua, B.E.[1]; Liam Doherty, M.D.[3]; Patricia Boyle, B.S.[3]; Philip Nolan, M.B., Ph.D.[4]; Walter T. McNicholas, M.D.[4]

[1]School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Dublin, Ireland; [2]BiancaMed Ltd, NovaUCD, University College Dublin, Dublin, Ireland; [3]The Respiratory Sleep Disorders Unit, St. Vincent's University Hospital, Dublin, Dublin, Ireland; [4]The Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland

Study Objectives: Transient changes in heart rate associated with obstructive apneas have been suggested for screening of sleep disordered breathing (SDB). This study prospectively compares the outcomes of an automated ECG-based SDB screening tool with simultaneous polysomnography.

Methods: The previously-developed automated algorithm was applied to a single channel ECG obtained during standard overnight polysomnography (92 subjects) to obtain an apnea-hypopnea index (AHI) estimate. Using AHI thresholds of <5 and ≥15 to define absence and presence of SDB, respectively, we determined the likelihood ratios of the proposed technique.

Results: The automated algorithm achieved positive and negative likelihood ratios of 2.16 and 0.08. Estimated and reference AHI were highly correlated (r = 0.88). Pathologically insignificant arrhythmia in some subjects had no discernible impact on the algorithm.

Conclusions: ECG-based assessment provides a simple but limited means of recognizing subjects with obstructive sleep apnea.

Keywords: Obstructive sleep apnea, sleep apnea, Holter monitoring, Holter electrocardiography, screening

Citation: Heneghan C; de Chazal P; Ryan S; Chua CP; Doherty L; Boyle P; Nolan P; McNicholas WT. Electrocardiogram recording as a screening tool for sleep disordered breathing. *J Clin Sleep Med* 2008;4(3):223-228.

O bstructive sleep apnea syndrome (OSAS) is associated with significant cardiovascular morbidity,[1-3] independent of potential confounding factors such as diabetes, dyslipidemia, and visceral obesity, and is one of the leading identifiable causes of hypertension.[4] However, although OSAS is relatively prevalent in middle-aged adults[5] and preschool children,[6] only 10% to 15% have been diagnosed.[7] Diagnosis and treatment is important as effective therapy leads to significant reductions in cardiovascular mortality and nonfatal cardiovascular events.[2,8]

The standard for OSAS diagnosis is overnight attended polysomnography.[9] However, resource constraints in many countries mean polysomnography suffers from low availability.[10] Therefore there is considerable interest in the development of reliable low-cost techniques for identification of subjects with OSAS, and the professional societies in sleep medicine maintain an ongoing interest in the evaluation of potential technologies which can address this issue.[11]

We previously reported an algorithm[12] for identifying patients with OSAS by recognizing both autonomic and respiratory effort-related changes in the ECG that are associated with apneic events. This study prospectively validates the performance of our previously developed algorithm on a single channel of overnight ECG recording, without retraining the parameters in the pattern recognition component of the algorithm.

MATERIALS AND METHODS

Subjects

Subjects were recruited from patients referred to the Sleep Disorders Clinic at St. Vincent's University Hospital, Ireland, for evaluation of suspected OSAS. Consecutive males who were free from other medical disorders and not commenced on regular medication were recruited. We also recruited healthy male subjects from the general population, age- and BMI-matched to the sleep clinic group. Only males were studied as the referral pattern to our sleep clinic is predominantly male, and thus there were insufficient numbers of females available to provide reliable data. The study was approved by the Hospital's Ethics Committee, and all subjects provided written informed consent.

Diagnostic Polysomnography

Standard overnight attended polysomnography (including one channel of modified lead V2 ECG) was conducted in a

Table 1—Summary of Subject Characteristics and Polysomnography Results, Grouped by Apnea Severity

Characteristic	Non-OSAS	Mild OSAS	Moderate to Severe OSAS	F (p-value)
n	27	15	50	-
Age (yr)	41.6 ± 6.2	43.1 ± 7.6	41.9 ± 7.5	0.23 (p = 0.80)
Body mass index (kg/m^2)	30.9 ± 3.3	31.0 ± 3.8	36.4 ± 7.6	9.08 (p < 0.001)
Apnea-hypopnea index (/h)	1.3 ± 0.9	10.3 ± 3.0	56.3 ± 30.4	60.75 (p < 0.001)
Epworth Sleepiness Scale	8.7 ± 5.6	11.1 ± 4.7	13.8 ± 5.5	7.85 (p < 0.001)

Non-OSAS: AHI < 5; Mild OSAS: 5 ≤ AHI < 15; Moderate OSAS: 15 ≤ AHI < 30; Severe OSAS: AHI ≥ 30. F (p-value): Results of one-way ANOVA.

dedicated sleep laboratory at room temperature and humidity. An experienced sleep technologist supervised data acquisition, and subsequently annotated the respiratory events (including obstructive, mixed, and central apneas and hypopneas, and periodic breathing episodes). The sleep technologist was blind to the output of the automated ECG analysis system; likewise, the automated ECG analysis was performed blind to the results of polysomnography scoring.

Apneas were defined as complete cessation of airflow ≥ 10 sec and hypopneas as reduction of respiratory signals ≥ 10 sec associated with oxygen desaturation > 4% and/or arousal. Obstructive events were distinguished from central events by the presence or absence of paradoxical thoracic and abdominal movements during apneas or hypopneas.

ECG Analysis

The ECG analysis algorithm performs an automated classification of the single-channel ECG obtained from polysomnography. This algorithm divides a recording into 1-min epochs and then estimates the probability of each epoch being an "apneic" or "normal respiration" minute, using a technique called linear discriminant classification. The per-epoch classifications are then combined across each recording to obtain a "minutes of SDB per hour" measure, which is then mapped to an estimated AHI using a linear transformation. Further technical details are presented in Appendix A.

In addition, standard Holter reports (Biomedical Systems, MO, USA) were obtained on the ECG recordings to objectively assess the technical aspects of the recordings and to screen for pathologically significant arrhythmias.

AHI Estimation

As a variety of solutions have been proposed as diagnostic alternatives to polysomnography, the American Academy of Sleep Medicine (AASM) has published guidelines for assessing agreement between diagnostic devices.[13] In line with these guidelines, we use Bland and Altman plots and likelihood ratios. To facilitate comparison with previous work, we also report (Pearson moment) correlation coefficients.

Per-Subject Screening

For apnea screening, the AASM guidelines suggest focusing on likelihood ratios and allowing for a "grey zone," where the result of the screening test is accepted as "indeterminate." This offers an advantage over the commonly used single knife-edge

threshold approach, as it does not penalize small errors due to natural variability.

Accordingly, we adopted the commonly used thresholds for defining clinically significant SDB and defined the result of the Holter screening test (for OSAS) as negative when the estimated AHI was < 5, positive if estimated AHI was ≥ 15, and indeterminate if estimated AHI was between 5 and 15. For comparison with previous work, we also report our results using a single knife-edge threshold of 15.

RESULTS

Ninety-eight subjects were recruited, and all completed the protocol. Data from 6 subjects could not be analyzed due to poor quality ECG recordings as assessed by the Holter reports. Table 1 summarizes the demographic data and polysomnography results, grouped by apnea severity. No subject had clinical evidence of other medical disorders. Most of the subjects received testing for full blood count, liver and renal function, cardiac enzymes, resting ECG, and assessment of cardiovascular risk factors; all results were within normal limits. Lastly, none of the subjects had central sleep apnea syndrome (CSAS) (highest proportion of central events per subject was 50%).

AHI Estimation

There was a strong correlation between estimated AHI and AHI obtained from polysomnography across a wide range of

Table 2—Confusion Matrix of Identifying Subjects with SDB Based on PSG AHI and AHI Estimated from Holter Analysis

Holter Estimate	Polysomnography Classification			
	N	**I**	**A**	**Row Total**
n	**10**	1	**2**	13
i	9	9	2	20
a	**8**	5	**46**	59
Column Total	27	15	50	92

A/a: Apnea/ clinically significant SDB (AHI ≥15), N/n: No apnea/ clinically insignificant SDB (AHI <5), I/i: Indeterminate test results (5 ≤ AHI < 15). Each cell indicates the number of subjects who had been classified as N, I or A by polysomnography and as n, **i** or a by Holter analysis. For example cell (n,A) shows that 2 subjects were classified as having apnea by polysomnography but Holter analysis indicated they do not have apnea. For calculation of (conditional) likelihood ratios, i.e., likelihood ratios conditional on obtaining either a positive or negative result, only the numbers in boldface were considered.
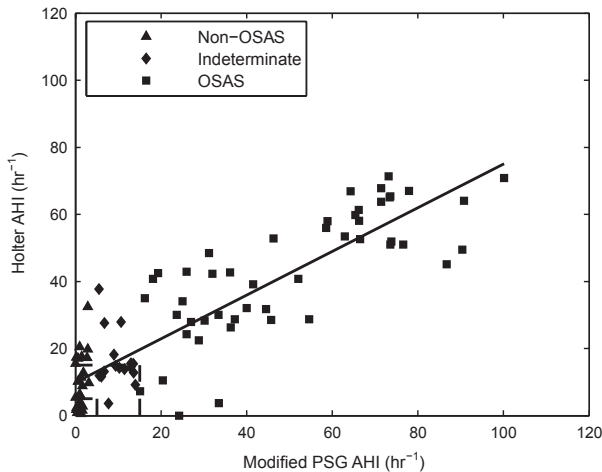
Figure 1—Scatter plot of estimated AHI and modified PSG AHI with best fit regression line. Subjects are grouped by apnea severity according to their PSG AHI score. Modified PSG AHI: Reference AHI obtained from polysomnography, excluding central apnea and hypopneas. Holter AHI: Estimated AHI obtained from ECG analysis. AHI thresholds of 5 and 15 are marked. One subject with a modified PSG AHI of 167 is not shown to improve the figure's clarity.
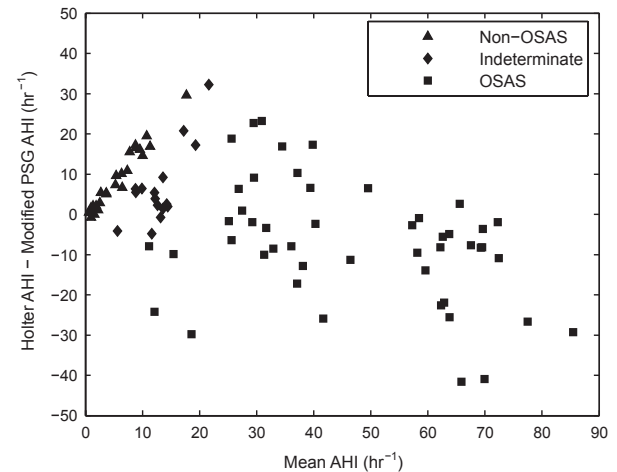

Figure 2—Bland-Altman plot of estimated AHI obtained from ECG analysis and modified PSG AHI obtained from PSG, *excluding* central apnea and hypopneas. One subject with a modified PSG AHI of 167 is not shown to improve the figure's clarity.

DISCUSSION

AHI Estimation

apnea severity (r = 0.88, p < 0.001). Corresponding Bland and Altman analysis showed a bias in estimated AHI (the Holter system underestimates the true AHI for high values of AHI). When this bias is removed, the standard deviation of the difference between the true and estimated AHI is 12.3/h.

In addition, we performed analysis using a modified PSG AHI which *excluded* central apnea and hypopnea events. The rationale for this analysis is that the original system was developed on datasets that had no central apneas or hypopneas. The correlation coefficient between the estimated AHI and modified PSG AHI is 0.88 (p < 0.001; Figure 1). With estimation bias removed, the standard deviation of the difference between the modified PSG AHI and the estimated AHI was 12.1/h (Figure 2).

Per-Subject Screening

Table 2 presents the screening results in the form of a confusion matrix obtained by classifying subjects according to the preselected AHI bands. As shown, positive and negative likelihood ratios were 2.16 (46/48 divided by 8/18) and 0.08 (2/48 divided by 10/18) respectively. According to AASM guidelines,[11] this represents "little change" and a "large reduction" in terms of increasing or decreasing pre-test probabilities respectively. 14% of subjects (11/77) obtained indeterminate test outcomes.

If a single AHI threshold of 15 was used, the sensitivity of our Holter screening test was 92%, specificity 69%, positive predictive value 78%, and negative predictive value was 88%.

Table 3 lists the types of arrhythmia encountered. None of the subjects had pathologically significant arrhythmia. Inspection of the results for subjects with more significant amounts of ectopy or bradycardia revealed that the presence of ectopy or bradycardia did not significantly affect the screening results. There was also no correlation between AHI and the percentage of ectopic beats (r = −0.05, p = 0.64).

The potential clinical significance of this study is the ability of the system to estimate a subject's AHI using just a single-channel surface ECG. The results also suggest the utility of the system as a screening tool for subjects with suspected OSAS, as overnight ECGs can be obtained relatively easily using a Holter monitor.

While there is ongoing debate in the sleep medicine community on the true value of AHI in estimating disease severity and significance, it remains the major objective variable by which treatment decisions are usually made. Accordingly, we believe that an automated apnea screening system should not only provide a broad yes/no/maybe output, but also some specific estimate of apnea severity as judged by AHI.

In that respect, our system provides an AHI estimate that correlates well with that obtained from polysomnography, although there is a considerable deviation between reference and estimated AHI. Also, the system in general overestimated true AHI at lower AHI and underestimates occurred at higher AHI. However, this may be primarily due to methodological differences in scoring of hypopneas between the St Vincent's Sleep

Table 3—List of Arrhythmias Encountered During Holter Analysis of the Single-Channel ECG Obtained from Polysomnography

Type	No of subjects	Range
Isolated/paired ventricular ectopy	40	0.1–1[a]
Isolated/paired supraventricular ectopy	64	0.1–3[a]
Bradycardia (<50 beats per minute)	61	0.1–83[a]
Bigeminy runs	3	1–2[b]
Supraventricular tachycardia runs	2	1[b]
Ventricular tachycardia runs	0	–
Pauses (>2 sec)	3	1–8[b]

[a] Percentage of total heart beats analyzed
[b] Number of episodes

Laboratory and the Marburg Sleep Laboratory, from which the data for developing the algorithm were obtained. For example, the Marburg Sleep Laboratory requires compensatory hyperventilation as a marker of hypopnea, whereas this is not a criterion in St. Vincent's. This may introduce a systematic bias as noted above; however, in this study we wished to prospectively test the existing algorithm on a new database rather than retraining it to perform optimally on the St Vincent's database. Underestimation at higher AHI could also be due to the time resolution of the system, which is unable to resolve multiple events occurring within a 1-min interval.

The results also demonstrated the robustness of the algorithm in the presence of limited occurrences of central and periodic breathing events, even though it was developed using a database without such events. Overall, in this study, central events accounted for an AHI of 3.0 ± 4.8/h. However, the performance of the system in subjects with predominantly central apnea is unclear.

Per-Subject Screening

The false positives were due to cyclic variations in heart rate (CVHR) episodes[14] being identified during periods that were not scored as apneic events by polysomnography. However, in many cases these periods were annotated as "possible events," and in fact, 4 of these subjects were assessed as probably having upper airway resistance syndrome (UARS) based on symptoms of loud snoring and significant daytime sleepiness. This indicates that the system cannot distinguish between UARS and OSAS, and would be a limitation of the system. For the false negatives, the RR tachograms revealed that the subjects did not exhibit sufficiently obvious bradycardia and tachycardia patterns during apneic events to be detected by the system.

For an AHI value of 5 to 15, conservative treatment is usually prescribed, which consists of weight loss, the use of a lateral sleep position, and avoidance of sedatives and alcohol.[15] Therefore it is possible that the sleep specialist may recommend conservative treatment for subjects with an indeterminate test outcome, or consider further evaluation with polysomnography.

The performance of our system if a single threshold was used is comparable with previous work. Roche et al have presented several reports on automated recognition of subjects with OSAS using analysis of heart rate variability,[16] inter-beat interval times[3] and wavelet analysis of the RR interval series.[17] They obtained sensitivity ranging from 83% to 92%, and specificity ranging from 52% to 96%. However, their techniques do not provide any temporal information about the occurrence of the apneic events, nor do they attempt to map their output variables to AHI.

Confounding Factors and Limitations

As a caveat, AHI derived from polysomnography is itself relatively inaccurate since there can be considerable inter-observer disagreement.[18] Therefore we should expect the margin of error on AHI scoring to be quite large, even among trained human experts. This would tend to emphasize that the AHI estimated from the Holter screening system will only be one component in forming an overall clinical judgment, and that factors such as BMI, excessive daytime sleepiness, and self-reporting of symptoms should also be elicited.

There has been strong interest within the sleep medicine community in tools that can approximate the outputs of polysomnography with a minimal number of signals. Within this context, we used only one ECG channel. Furthermore, because our focus is on RR intervals and respiratory modulation of the QRS complex, we do not expect performance to be significantly degraded by using only one ECG channel. In addition, the modified V2 lead was used, which generally provides good QRS complexes.

From an electrophysiology point of view, factors such as thoracic impedance, temperature and humidity can affect ECG recordings. However, as we did not record these parameters, we could not analyze the influence of these parameters on ECG and hence the proposed system's ability to identify apneic events. Also, we did not focus on balancing the number of controls (subjects with AHI < 5, 27 of 92 subjects) and number of apneic subjects (AHI ≥ 15, 50 of 92 subjects) as our focus was on the utility of the proposed system in a typical group of patients encountered in the sleep clinic. However, we did want to control for comorbidity and hence we only recruited consecutive patients without comorbidities.

The performance of the ECG analysis algorithm in subjects with low or minimal heart rate variability (such as in heart failure patients) is unclear. However, we would expect performance to be degraded as the cyclic variations in heart rate (CVHR) patterns in these subjects would be diminished. However, we would need to conduct a specific study in a target population with known reduced heart rate variability in order to definitively assess the impact of reduced HRV. Roche et al. did suggest a useful technique for compensating for reduced HRV by looking at differences between daytime and nighttime HRV parameters, rather than absolute values of nighttime HRV.[16]

There are two further limitations to our study in terms of assessing its general applicability. The study population was male only, and hence the performance of the algorithm in detecting OSAS in women is unknown. We particularly note this in the context of well-accepted differences in the presentation of OSAS in women (e.g., higher ratio of hypoponeas to apneas than in men),[19] so it would be not unexpected if the performance of the algorithm was different in women, and indeed might need to be specifically "tuned" for a female population. However, we have tested the algorithm in a small number of females (4 subjects; data not presented) and found no major difference in performance compared to males. Secondly, we cannot extrapolate the results of this study to its efficacy in a general population as a "screening" tool. Subjects presenting at the Sleep Laboratory in many cases have already undergone informal assessment by their primary care physician who had identified them as possible OSAS sufferers. Therefore the studied population has a much higher prevalence of OSAS than might be expected in a population selected solely for screening.

Despite these caveats, we believe that an important first step in assessment of this methodology is to obtain prospective evaluation of the automated algorithm in a well-controlled population and environment, rather than proceeding directly to evaluation in a more general population in their home environment.

To conclude, we have presented an automated ECG analysis algorithm for use as a means to recognize subjects with sleep apnea. The main advantages of this system are that it provides (a) a direct estimate of AHI, and (b) the temporal sequence of apnea events during the night, which can assist a clinician in forming a diagnosis. In addition, we demonstrated that the automated system is robust to a range of observed intermittent arrhythmias, as might be encountered in a typical population being screened for sleep apnea.

ACKNOWLEDGMENTS

ABBREVIATIONS

AASM	American Academy of Sleep Medicine
AHI	Apnea-hypopnea index
CSAS	Central sleep apnea syndrome
CVHR	Cyclical variations in heart rate
EDR	ECG-derived respiratory signal
ESS	Epworth Sleepiness Scale
OSAS	Obstructive sleep apnea syndrome
PSG	Polysomnography
SDB	Sleep disordered breathing

REFERENCES

1. McNicholas WT, Bonsignore MR, the Management Committee of EU COST ACTION B26. Sleep apnoea as an independent risk factor for cardiovascular disease: current evidence, basic mechanisms and research priorities. Eur Respir J 2007;29:156-78.
2. Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. Lancet 2005;365:1046-53.
3. Roche F, Duverney D, Court-Fortune I, et al. Cardiac interbeat interval increment for the identification of obstructive sleep apnea. Pacing Clin Electrophysiol 2002;25:1192-9.
4. Chobanian AV, Bakris GL, Black HR, et al. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Hypertension 2003;42:1206-52.
5. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993;328:1230-5.
6. Gislason T, Benediktsdottir B. Snoring, apneic episodes, and nocturnal hypoxemia among children 6 months to 6 years old. An epidemiologic study of lower limit of prevalence. Chest 1995;107:963-6.
7. Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. Sleep 1997;20:705-6.
8. Doherty LS, Kiely JL, Swan V, McNicholas WT. Long-term effects of nasal continuous positive airway pressure therapy on cardiovascular outcomes in sleep apnea syndrome. Chest 2005;127:2076-84.
9. Kushida CA, Littner MR, Morgenthaler T, et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. Sleep 2005;28:499-521.
10. Flemons WW, Douglas NJ, Kuna ST, Rodenstein DO, Wheatley J. Access to diagnosis and treatment of patients with suspected sleep apnea. Am J Respir Crit Care Med 2004;169:668-72.
11. American Thoracic Society Documents. Executive summary on the systematic review and practice parameter for portable monitoring in the investigation of suspected sleep apnea in adults. Am J Respir Crit Care Med 2004;169:1160-3.
12. de Chazal P, Heneghan C, Sheridan E, Reilly R, Nolan P, O'Malley M. Automated processing of the single lead electrocardiogram for the detection of obstructive sleep apnea. IEEE Trans Biomed Eng 2003;50:686-96.
13. Flemons WW, Littner MR. Measuring agreement between diagnostic devices. Chest 2003;124;1535-42.
14. Guilleminault C, Connolly S, Winkle R, Melvin K, Tilkian A. Cyclical variation of heart rate in sleep apnea syndrome. Lancet 1984;323:126-31.
15. Flemons WW. Clinical practice. Obstructive sleep apnea. N Engl J Med 2002;347:498-504.
16. Roche F, Gaspoz JM, Court-Fortune I, et al. Screening of obstructive sleep apnea syndrome by heart rate variability analysis. Circulation 1999;100:1411-15.
17. Roche F, Pichot V, Sforza E, et al. Predicting sleep apnea from heart period: a time-frequency wavelet analysis. Eur Respir J 2003;22:937-42.
18. Collop NA. Scoring variability between polysomnography technologists in different sleep laboratories. Sleep Med 2002;3:43-7.
19. O'Connor C, Thornley KS, Hanly PJ. Gender differences in the polysomnographic features of obstructive sleep apnea. Am J Respir Crit Care Med 2000;161:1465-72.

APPENDIX A: AUTOMATED ECG ANALYSIS ALGORITHM

The algorithm takes in a single channel of ECG recording. QRS complexes were first detected, leading to the set of RR intervals. The analysis does not attempt to distinguish between RR intervals of normal sinus rhythm origin and non-sinus rhythm depolarizations. Numerical simulations have also confirmed that infrequently occurring ectopic beats have a negligible effect on the analysis. An ECG-derived respiratory (EDR) signal was calculated by finding the area of the QRS complex. The ECG amplitude is modulated by respiration, and sampling and interpolation of the enclosed ECG area at QRS detection points will yield an estimate of respiratory effort at the ribcage level, as previously reported.[12]

For classification purposes, the ECG recording was then broken into epochs of 1-min duration. A set of numerical features was calculated for each 1-min epoch. These features, which are a subset of the features described by deChazal, et al.,[12] include the values of the RR interval-based spectrum at 32 evenly spaced intervals from 0 to 0.5 cycles/interval, the mean and standard deviation of the RR intervals, and the standard heart rate variability parameters NN50 and SDSD, calculated over

the 1-min epoch. Other features calculated include 32 evenly spaced values of the EDR spectrum. In this manner, each 1-min epoch is associated with a 74-member vector of features.

These vectors were then fed into a linear discriminant classifier whose output is the estimated probability (between 0 and 1) of that 1-min epoch containing episodes of OSA. To estimate the subject's apnea-hypopnea index (AHI), the number of apneic minutes in the whole recording was summed and divided by the estimated total sleep period to obtain a "minutes of sleep disordered breathing per hour" measure. Finally, this "minutes per hour" measure was then mapped to an estimated AHI using a scaling factor.