ERJ Express. Published on January 24, 2007 as doi: 10.1183/09031936.00091206

Automatic detection of sleep-disordered breathing from a single channel airflow record

Hiroshi Nakano, Takeshi Tanigawa*, Tomokazu Furukawa and Sankei Nishima

Dept of Pulmonology, Fukuoka National Hospital, Fukuoka, Japan

*Dept of Public Health Medicine, Graduate School of Comprehensive Human Sciences University of Tsukuba, Tsukuba, Japan

Running head

Automatic detection of apnea/hypopnea

Correspondence

H. Nakano, Dept of Pulmonology, Fukuoka National Hospital, 4-39-1 Yakatabaru, Minami-ku, Fukuoka, 811-1394, Japan.

E-mail nakano_h@palette.plala.or.jp

Tel +81-92-5655534

Fax +81-92-9669444

Copyright 2007 by the European Respiratory Society.

Abstract

Single channel airflow monitors developed for screening for sleep-disordered breathing (SDB) have conflicting results for their accuracy. We hypothesized that the analytic algorithm is crucial for the performance and tried to develop a novel computer algorithm. We employed 399 polysomnography records including a thermal sensor signal. The first 100 records were used in the development of the algorithm and the remainder for validation. In addition, 119 polysomnography records including a thermocouple signal and a nasal pressure signal were used for the validation. The algorithm was designed to obtain a time series (flow-power) using power spectral analysis that expresses fluctuation in the airflow signal amplitude. From the time series the algorithm detects transient falls of the flow-power and calculates flow-RDI defined as the number of the falls per hour.

In the validation group, the areas under receiver operating characteristic curves for diagnosis of SDB (apnea-hypopnea index ≥5) were 0.96, 0.95, and 0.95, for the records of the thermal sensor, thermocouple, and nasal pressure system, respectively. The diagnostic sensitivity / specificity of the flow-RDI were 96%/76%, 88%/80%, and 97%/77%, respectively.

These results suggest that a single channel airflow monitor can be used to detect SDB automatically if the analytic algorithm is optimized.

Key words: sleep apnea, screening, signal processing, power spectral analysis

Introduction

Sleep-disordered breathing (SDB) is known to be a relatively common condition which can affect 24% and 9% of middle-aged men and women, respectively, when it is defined by five or more episodes of apnea or hypopnea per hour of sleep regardless of symptoms (1). Many studies have indicated that SDB, even in mild stage, is strongly associated with cardiovascular morbidity (2–4) as well as traffic accidents (5–6). Thus undiagnosed SDB is considered as an important public health problem (7). In clinical setting, the exact diagnosis of SDB is made by polysomnography (PSG), which is a labor intensive and time consuming process, and is thus impractical to perform as a part of a general population health check-up. In Japan, several surveys using pulse oximetry have been performed and a high prevalence of SDB was found in working populations (8). However, this method has a limitation because oximetry cannot always detect apnea events especially for non-obese subjects (9). Recently, Ayappa et al (10) reported that manual analysis of the flow signal could provide a result similar to that obtained by full PSG analysis. Moreover, several single-channel airflow monitors have been developed for automated detection of SDB (11–13). Studies for the accuracy of these monitors have provided conflicting results, and it is unclear whether the performance depends on the signal acquisition system including the sensor or on the analytic algorithm. We hypothesized that automated analysis of a single-channel airflow record may have a good diagnostic performance for SDB when the analytic algorithm is fully optimized. To test the hypothesis, we developed a novel computer algorithm for flow signal analysis to detect SDB and validated it against a series of PSG records. We developed the algorithm for

a thermal flow sensor, but also validated it for another type of thermal sensor and a nasal pressure sensor.

Method

Subjects

We used a retrospective sample of diagnostic PSG records from January 2002 to September 2003 (group 1), when a thermal sensor was used to detect airflow. This series contained 418 PSG records from the patients who were referred for suspected SDB from general practitioners or consulted our clinic because of symptoms suggestive of SDB (habitual snoring: n = 355, daytime sleepiness: n = 351, witnessed apnea: n = 346, and nocturnal choking: n = 140). From the sample, the first 105 records (development group) were used for the development of the algorithm, and the latter 313 records (validation group) were used for the validation of the algorithm. From the development group, five records were excluded because of poor signal quality. From the validation group, 14 records were excluded because of failure in channels other than airflow. We did not exclude the records with poor airflow signal from the validation group, even if the record was difficult to interpret by visual inspection, because the aim of the development of the algorithm was for use in a fully automated analysis of the airflow record. The characteristics of the patients are listed in Table 1.

Another series of diagnostic PSG records from July 2005 to December 2005 (group 2), when both of a nasal prong pressure transducer and a thermocouple sensor were used to detect airflow, were used to evaluate the applicability of the algorithm to other types of flow sensor.

This series contained 120 PSG records, of which only one record was excluded because of failure in a respiratory movement channel. These PSG records included either or both nasal prong pressure record (n=117) and thermocouple record (n=116).

Polysomnography

The sensors for PSG were fixed by technicians in the hospital, but the sensors were not monitored after the start of the recording. PSG was recorded using a polygraph system (EEG7414, Nihon Kohden, Tokyo, Japan). Oronasal airflow was monitored with a thermal sensor (group 1: a polyvinylidene fluoride film; PVDF sensor, Dymedix Corporation, MN, USA; group 2: a thermocouple sensor, Pro-Tec, Mukilteo, WA, USA). In group 2 a nasal prong pressure transducer (PTAF, Pro-Tec, Mukilteo, WA, USA) was also used to monitor airflow. Thoracic and abdominal respiratory movements were monitored by respiratory inductive plethysmography (RIP; Respitrace, Ambulatory Monitoring Inc., Ardsley, NY, USA), which was calibrated by an isovolume maneuver before the test. Signals of airflow and respiratory movement were recorded at a sampling frequency of 10 Hz. Oxyhemoglobin saturation was monitored using a pulse oximeter (OLV-3100, Nihon Kohden, Japan) at the fastest response mode and recorded at a sampling frequency of 1 Hz. Sleep stages were scored manually according to the standard criteria (14). The oximetry data were analyzed automatically with a personal computer to detect any SaO_2-dip $\geq$ 3% lasting for 10 to 120 s. The oxygen desaturation index (ODI) was defined by the number of SaO_2-dips per hour of examination. Hypopnea (including apnea) was defined as follows (A) a $\geq$ 50% reduction of amplitude in the RIP sum signal lasting 10 s or longer, or (B) a discernible reduction ($\geq$30%

and <50%; duration ≥10 s) in the RIP sum signal associated with a ≥3% oxygen desaturation or an arousal (15, 16). The scoring of apnea-hypopnea events was performed by automated detection based on the RIP amplitude criteria followed by a thorough manual review and editing. The manual review was done on the computer screen and the amplitude of each signal was adjusted to be adequate for visual analysis. We used the RIP sum channel primarily for hypopnea identification; however in a few cases where airflow signal correlated with change in oxygen saturation better than RIP signal, we used the raw airflow channel. The scoring was done before the development of the algorithm for the current study. The apnea-hypopnea index (AHI) was calculated as the number of apnea and hypopnea events per hour of sleep. We did not score respiratory effort related arousals (RERAs). Because of the scoring method outlined above, in most cases, the thermal sensor signal did not contribute to the AHI value virtually.

Development of algorithm

For the automated analysis of the airflow signal, the following procedures (Fig. 1) were developed using the PSG data of the development group.

1. Generation of flow-power time series

The first step of the algorithm is generation of a time series which express the variation of the respiratory airflow. For this step we used a fast Fourier transformation (FFT) of the airflow signal to obtain the short-time power spectrum within the bandwidth of respiratory frequency. We designated this power spectrum as “flow-power.” The time window used for the FFT was the Hanning window of 128 points (12.8 s). The bandwidth of respiratory

frequency was set at 0.167 to 0.533 Hz (10 to 32 cycles per min) because respiratory frequencies between apnea events were observed within this range in the 100 subjects. FFT was performed every 2 s (overlapped windows). The flow-power was logarithmically transformed and expressed in dB (the maximal value was set as 0dB).

$$\text{Power in dB} = 10 \times \log_{10}(\text{Power/maximal Power}).$$

Thereafter, the flow-power time series was low-pass filtered to obtain a smooth contour.

2. Detection of the flow-power dip

The second step of the algorithm is detection of the SDB events. The typical change in flow-power at the SDB event was a transient decrease, which we decided to detect when there was a decrease greater than the threshold value within 30 s and when it started to resume within 90 s (Fig. 1B). We designated such a decrease in the flow-power as "flow-power dip."

3. Tuning of the algorithm

To determine the above-mentioned threshold value for the detection of flow-power dips, we compared the concordance between the number of apnea-hypopnea events detected by PSG and the numbers of flow-power dips at various threshold values. Thereafter, we detected flow-power dips at the determined threshold and obtained a respiratory disturbance index (flow-RDI), which we defined as the number of flow-power dips per hour of examination. To determine an appropriate flow-RDI cutoff value, we constructed a receiver-operating characteristic curve (ROC) to diagnose three levels of SDB (AHI≥5, 10, and 15).

Validation of the algorithm

We applied the developed algorithm to the PVDF thermal sensor record in group1 validation group as well as nasal pressure record and thermocouple record in group 2 to calculate flow-RDI. The square root of the pressure was used as the flow signal for the nasal prong pressure transducer system (17).

The relationship between the flow-RDI and the AHI was evaluated using Pearson's correlation coefficient. The agreement between these data was analyzed as described by Bland and Altman (18). The diagnostic ability of the flow-RDI for SDB was evaluated in terms of sensitivity and specificity.

To evaluate event-by-event agreement, we compared the events detected by three methods of automated detection by the algorithm, visual inspection of the thermal sensor record, and that of full PSG records, during mid one hour sleep in 100 consecutive PSG records. Visual detection of apnea/hypopnea events (50% or more reduction of amplitude lasting 10 s or longer) from the thermal sensor record was done on the computer screen, which visualized only the thermal sensor record. The event-by-event agreement was evaluated using the proportion of specific agreement (PSA) (19, 20). In addition, to know the effect of artifacts and irregular breathing while awake on the flow-RDI, we calculated the awake flow-RDI for 10 min or longer continuous awake segments in 50 patients in group 1 who had low sleep efficiency (<66%).

Comparison with other methods

1. Comparison with a conventional automatic analysis of flow signal

We compared the diagnostic performance of the developed algorithm with that of a conventional automated analysis using the flow channel data of the gorup1 validation group. We employed a computer program (Screening.exe: limited version of Seep Ukiha (21) for flow signal alone, NGK sparkPlug Co. Ltd, Nagoya, Japan) as the conventional analysis of the airflow signal. This program is designed to detect a flow signal amplitude decrease (>50%; duration ≥ 10 s) from baseline. We designated the number of events per hour of examination obtained by the program as "conventional RDI." The diagnostic performance of the conventional RDI was evaluated as described above. To compare the diagnostic sensitivity and specificity between the flow-RDI and the conventional RDI, the ROCs were constructed and the areas under the curves (AUCs) were calculated. The comparisons of the AUCs were performed by the Hanley and McNeil method (22).

2. Comparison with oximetry

The diagnostic ability of the flow-RDI and that of ODI were compared in the group 1 validation group. A separate comparison was also conducted for normal weight group (body mass index <25) and overweight-obese group (body mass index ≥ 25).

Results

Development of algorithm

1. Flow-power time series

A representative case is shown in Fig. 2. Transient decreases in the flow-power (flow-power dip) corresponding to SDB events were clearly identified in the filtered

flow-power time series. Even in the case of a noise contaminated airflow signal, the filtered flow-power time series depicted SDB events clearly (Fig. 3).

2. Determination of the threshold for the flow-power dip

In the development group, the intraclass correlation coefficients between the number of flow-power dips and the number of apnea/hypopnea events were 0.72, 0.93, 0.94, 0.91, and 0.87, for the flow-power dip thresholds of 3 dB, 4.5 dB, 6 dB, 7.5 dB and 9 dB, respectively. As a result, we determined the appropriate threshold to be 6 dB.

3. Determination of the cutoff value for flow-RDI

From the ROCs for the diagnosis of three levels of SDB in the development group, we determined the appropriate flow-RDI cutoff values to be 5, 7.5, and 10 for the SDBs of AHI≥5, 10, and 15, respectively. The sensitivity and specificity at these cutoff values were 92% and 80%, 93% and 84%, and 90% and 83%, respectively.

Validation of the algorithm

1. Agreement between the flow-RDI and AHI (Table 2)

The correlation coefficients between the flow-RDI and the AHI were high regardless of flow sensor type. Bland and Altman analysis showed that the difference between the flow-RDI and the AHI were narrower for the nasal pressure sensor compared with the other sensors. The bias in group 1 (PVDF sensor) is shown in Fig. 4. The bias was highly correlated with AHI (Fig. 4B).

2. Diagnostic ability of the flow-RDI

The diagnostic sensitivity and specificity of the flow-RDI in group 1 validation group are shown as ROCs (Fig. 5). The sensitivity and specificity at the preset flow-RDI cutoff value are shown in Table 2. The sensitivity was relatively high (0.95-0.97) and the specificity was moderate (0.73-0.82) for the PVDF sensor as well as the nasal pressure sensor. On the other hand, the sensitivity was lower and the specificity was higher for the thermocouple sensor as compared to those for the other sensors. The cutoff value of the flow-RDI for the thermocouple sensor to obtain sensitivity and specificity similar to the other sensors was 4.0, 5.0, and 7.5, for AHI $\geq$ 5, 10 and 15, respectively.

3. Event-by-event analysis

The analysis of event-by-event agreement showed that the ratio of number of events detected by the automated analysis to that of total PSG apnea-hypopnea events was 77%, 73% and 89% for the PVDF, thermocouple and nasal pressure sensors, respectively (Fig. 6). The ratio of number of false positive events to that of total events detected by the automated analysis was 9%, 11% and 14% for the PVDF, thermocouple and nasal pressure sensors, respectively (Fig. 6).

The PSA between the PSG analysis and the automated analysis was 0.833, 0.802 and 0.874 for the PVDF, thermocouple and nasal pressure sensors, respectively. The PSA between the PSG analysis and the flow record manual analysis was 0.835, 0.806 and 0.893 for the PVDF, thermocouple and nasal pressure sensors, respectively.

4. Analysis of waking time

The mean flow-RDI while awake was 13.2 ± 11.5. The awake flow-RDI had weak but significant correlation with the patient's AHI ($r = 0.28$, $p = 0.045$). The mean awake RDI for the patients with AHI < 15 was significantly lower than that for the patients with AHI ≥ 15 (6.9 ± 6.1 vs. 15.2 ± 12.2; $p = 0.0036$ by Welch's t-test).

Comparison with other methods

1. Comparison with a conventional automated analysis program

The correlation coefficient between the conventional RDI and the AHI was 0.888 (95%C.I. 0.861–0.910). Bland and Altman analysis showed that the mean ± 1.96 SD difference between the flow-RDI and the AHI was −3.2 ± 22.2. The AUCs were 0.903, 0.924, and 0.919 for AHI ≥ 5, 10, and 15, respectively, all of which were significantly lower than those of the flow-RDI ($p=0.000002$, 0.000020, and 0.000035, for AHI ≥ 5, 10, and 15, respectively).

2. Comparison with oximetry analysis

The diagnostic sensitivity and specificity of the ODI at the same cutoff value as flow-RDI were 94% and 84%, 93% and 87%, and 92% and 82% for the AHI cutoff values of 5, 10, and 15, respectively. The AUCs were 0.966, 0.977, and 0.961 for AHI ≥ 5, 10, and 15, respectively. All these parameters for diagnostic ability were similar to those of the flow-RDI. However, the diagnostic sensitivity of ODI was influenced by body habitus. The diagnostic sensitivity of ODI in the normal weight group was 86%, 87%, and 84% for AHI ≥

5, 10, and 15, respectively, while that in the overweight-obese group was 99%, 97%, and 97%, respectively. By contrast, the diagnostic sensitivity of the flow-RDI in the normal weight group was 90%, 95%, and 96% for AHI $\geq$ 5, 10 and 15, respectively, while in the obese-overweight group it was 99%, 96%, and 94%, respectively.

Discussion

We developed a new algorithm for fully automated detection of SDB from a single channel airflow signal. The algorithm utilizes power spectral analysis to detect the variation of airflow signal amplitude. In the validation phase of this study, it was shown that the algorithm had a relatively high diagnostic performance when the AHI by the manual full PSG analysis was used as the diagnostic gold standard. The diagnostic sensitivity was less affected by the subjects' body habitus compared with an oximetry analysis. The algorithm was applicable not only to the thermal sensor that was used for the development of the algorithm but also to another type of thermal sensor and a nasal pressure sensor.

Technical aspect

This algorithm is unique in several points. First, it utilizes power spectral analysis, by which it can lessen the effect of noise contamination because fluctuation of the airflow signal outside the possible frequency range of respiration can be eliminated. Second, it generates a time series of logarithmic power of airflow fluctuation (flow-power). The change in logarithmic power is proportional not to the difference but to the ratio between two points. When a 6 dB decrease is observed in this time series, it means that the mean amplitude during about 10 seconds decreased to one-half of that during the preceding part. This property

provides an advantage that the same scale can be used even when the level of flow signal changes due to change in the condition of the airflow sensor (e.g. a change in location) during a whole night's recording. Third, it processes the flow-power time series by a low-pass filter, which can eliminate changes irrelevant to apnea-hypopnea cycles. Fourth, it detects a transient decrease in flow-power (flow-power dip). The algorithm can detect the flow-power dip in a simple automated way. We think these characteristics of the algorithm make the single channel monitoring relatively accurate in the diagnosis of SDB.

Most of automated analyses of airflow to detect apnea/hypopnea events are considered to employ waveform analysis, which detects the events based on decrease in amplitude of the airflow waveform. Although the conventional waveform analysis program which we tested in this study follows the apnea/hypopnea definition literally, a considerable number of the events detected by the program seemed not to be real events at a visual inspection. A pattern recognition method on waveform is thought to be susceptible to various kinds of artifact. To our knowledge, only one study (25) other than ours has used power spectral analysis for airflow signal. Matrot (25) used short-time Fourier transform to calculate the spectral power profile of the respiratory signal, which is the same method with ours. However, the other part of analysis is different between the Matrot's method and ours. Martot used normalization of signal before power spectral analysis and detected apnea / hypopnea using a threshold for the spectral power. We used no normalization of signal, but we used logarithmic transformation for the spectral power. Such different strategy is considered to be relevant to the different setting. Martot developed the program to monitor activity and apneas of mice for relatively short time, while we developed the program to monitor breathing events of the human all

through the night. In our setting, the combination of the normalization and the threshold is thought to work incompletely. In another study (26) we have also used power spectral analysis to detect SDB from tracheal sound record. However, the purpose of power spectral analysis is different because in the tracheal sound analysis the power spectral analysis is used to calculate sound power which is related to airflow velocity but in the present algorithm the power spectral analysis is used to obtain the magnitude of airflow signal fluctuation corresponding to respiration.

Performance of the algorithm

The flow-RDI obtained by the algorithm has a relatively accurate diagnostic performance despite the fact that it uses only one channel and it uses completely automated analysis.

It is difficult to compare the performance of the algorithm with the other algorithms because almost all of them can be applicable only for the monitors originally equipped with them. We could compare the performance of the developed algorithm with that of a conventional automatic analysis program on the market, demonstrating the superiority of the algorithm. Several previous studies on the diagnostic accuracy of portable monitors have shown that automated scoring of respiratory signals is far less accurate compared with manual scoring (23, 24). By contrast, we demonstrated the performance of the algorithm is almost comparable with manual analysis of flow signal alone by the event-by-event analysis. From these facts, we believe the developed program is superior to many of conventional algorithms.

The Bland and Altman plot showed considerable systematic bias between the flow-RDI and the AHI. The bias, which was highly correlated with the AHI, means systematic

underestimation of the flow-RDI. The underestimation is considered to be due to two reasons. One is the failure of the algorithm to detect a certain proportion of hypopnea events, which was demonstrated by the event-by-event analysis and could cause underestimation of 23%. The other reason is that the flow-RDI uses examination time as the denominator while the AHI uses sleep time. Mean sleep efficiency in the validation group (78%) can cause underestimation of 22%. However, this underestimation is considered to be partially counterbalanced by detection of events during waking time. We think the counterbalance is not accidental but relevant to SDB, because the awake flow-RDI was correlated with the AHI. The resultant actual effect of underestimation was 26% of AHI. The event-by-event analysis also showed false detection of 200 events per 100 hours. This false detection was not correlated to the AHI, and could cause flow-RDI to be overestimated by 2.0/hour. We think that this overestimation explains the positive bias for subjects with low AHI in AHI vs. bias plot.

The AHI-dependent systematic bias between the AHI and the flow-RDI is considered to be an important advantageous feature of the algorithm in two points. One is that such bias can be partially compensated by multiplying the flow-RDI by a specific coefficient. The other is that subjects with relatively low AHI (AHI<20) are less affected by the bias, resulting in less adverse effect for the diagnostic ability. Most of other reports (11, 13, 23, 24) on validation of automated analysis do not show such an AHI-dependent systematic bias but rather a random and larger bias.

Performance of sensors

We originally developed the algorithm for the analysis of thermal flow sensor signal. It is well known that the change in thermal flow signal is not linear to the change in actual airflow (27) and hypopnea detection by thermal flow sensor may underestimate SDB.

The thermal sensor that we used in group 1 utilizes a polyvinylidene fluoride (PVDF) film, which is reported to have very fast response to temperature change and an excellent ability to detect hypopnea as compared with conventional thermal sensors (28). Although we could not compare the PVDF sensor and the thermocouple sensor directly, the event-by-event agreement was better for the PVDF sensor in group 1 than for the thermocouple sensor in group 2. Moreover, it was shown that the nasal pressure sensor was far better than the thermocouple sensor in view of the event-by-event agreement in group 2. These findings agree with the known difference in the sensitivity among these sensors.

However, diagnostic ability of the flow-RDI expressed as AUC of receiver operating characteristic curve was not so different among them. The high AUC means that the flow-RDI at an appropriate cutoff value can have high diagnostic ability. Actually, the appropriate cutoff values were lower for the thermocouple sensor and equivalent or higher for the nasal pressure sensor than those for the PVDF sensor, corresponding to the difference of the sensor characteristics. We think that the above-mentioned characteristics of the algorithm may make such adjustment capable to improve the diagnostic ability of the low-performance sensor. Another possible way to adjust to the sensor sensitivity is to choose an appropriate threshold value to detect the flow-power dip, which we have not tested in the current study.

Study limitations

This study has several limitations. First, the diagnostic ability of a monitor to diagnose SDB must be dependent on the population studied. In the present study, the subjects belonged to the population with relatively high probability of having SDB because they had symptoms suggestive of SDB. Therefore, the finding in the present study may not be applicable to general populations. However, the characteristic of the flow-RDI that the diagnostic sensitivity was less affected by subjects' body habitus as compared to oximetry analysis is one of advantageous features of it because mean BMI of general population tend to be lower than that of clinic samples. Second, this study used a flow signal channel from PSG record to validate the algorithm as a simulation of single-channel monitoring. A real validation of the algorithm needs to use the flow signal record in the home. Third, we did not include RERAs in the SDB events. If these were included, the estimated diagnostic ability may have become worse. Forth, although the algorithm is relatively immune to noise contamination, a record which is occupied exclusively by noise may be interpreted as having no events. For a practical use, it may not be so difficult to register noise segments as invalid segments based on the power spectral property. However, in this study we did not implement such a process, because we intended to show the usefulness of flow-power dip analysis. Finally, the single channel flow analysis has inevitable disadvantages; one is inability to calculate sleep time as the denominator of AHI and the other is inability to distinguish obstructive and central events. However, for the purpose of mass screening for SDB, the advantage of simplicity, and easy applicability might predominate over such disadvantages.

In conclusion, we have developed a new computer algorithm which can automatically detect SDB from the airflow record alone. The system needs no manual edit and is relatively

accurate. We think that not only the performance of sensors but also the analytic algorithm is crucial for development of simple portable monitor applicable for mass-screening for SDB.

References

1. 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.

2. Peppard PE, Young T, Palta M, Skatrud J. Prospective study of the association between sleep-disordered breathing and hypertension. N Engl J Med 2000;342:1378–84

3. Nieto FJ, Young TB, Lind BK, Shahar E, Samet JM, Redline S, D'Agostino RB, Newman AB, Lebowitz MD, Pickering TG. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. JAMA 2000;283:1829–1836.

4. Shahar E, Whitney CW, Redline S, Lee ET, Newman AB, Javier Nieto F, O'Connor GT, Boland LL, Schwartz JE, Samet JM. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. Am J Respir Crit Care Med 2001;163: 19–25.

5. Young T, Blustein J, Finn L, Palta M. Sleep-disordered breathing and motor vehicle accidents in a population-based sample of employed adults. Sleep 1997;20:608–613.

6. Terán-Santos J, Jimenez-Gomez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. N Engl J Med 1999;340:847–851.

7. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea. A population health perspective. Am J Respir Crit Care Med 2002;165:1217–39

8. Tanigawa T, Horie S, Sakurai S, Iso H. Screening for sleep-disordered breathing at workplaces. Industrial Health 2005;43:53-57.

9. Nakano H, Ikeda T, Hayashi M, Ohshima E, Itoh M, Nishikata N, Shinohara T. Effect of body mass index on overnight oximetry for the diagnosis of sleep apnea. Respir Med 2004;98:421–7

10. Ayappa I, Norman RG, Suryadevara M, Rapoport DM. Comparison of limited monitoring using a nasal-cannula flow signal to full polysomnography in sleep-disordered breathing. Sleep 2004;27:1171-9.

11. Shochat T, Hadas N, Kerkhofs M, Herchuelz A, Penzel T, Peter JH, Lavie P. The SleepStripTM: an apnoea screener for the early detection of sleep apnoea syndrome. Eur Respir J 2002; 19:121–126.

12. Wang Y, Teschler T, Weinreich G, Hess S, Wessendorf TE, Teschler H. Validation of microMESAM as screening device for sleep disordered breathing. Pneumologie 2003; 57:734-740.

13. Almeida FR, Ayas NT, Otsuka R, Ueda H, Hamilton P, Ryan FC, Lowe AA. Nasal pressure recordings to detect obstructive sleep apnea. Sleep and Breathing 2006;10:62-9.

14. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring systems for sleep stages of human subjects. National Institute of Health, Publ.No.204,1968

15. American Academy of Sleep Medicine. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. Sleep 1999;22:667-89.

16. Bonnet M, Carley D, Carskadon M, et al. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep 1992;15:173-84.

17. Farre R, Rigau J, Montserrat JM, Ballester E, Navajas D. Relevance of linearizing nasal prongs for assessing hypopneas and flow limitation during sleep. Am J Respir Crit Care Med 2001;163:494–497.

18. Bland JM, Altman DG. Statistical methods assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.

19. Ayappa I, Norman RG, Krieger AC, Rosen A, O'Malley RL, Rapoport DM. Non-invasive detection of respiratory effort-related arousals (RERAs) by a nasal cannula/pressure transducer system. Sleep 2000;23:763-771.

20. Fleiss JL. Statistical methods for rates and proportions. John Wiley&Sons, New York, 2003; pp.599-608.

21. Hashizume Y, Kuwahara H, Uchimura N, Mukai M, Shirakawa S, Satomura T, Takeuchi N, Tanaka J, Maeda H. Examination of accuracy of sleep stages by means of an automatic sleep analysis system 'Sleep Ukiha' Psychiatry Clin Neurosci 2001; 55:199-200.

22. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristics curves derived from the same cases. Radiology 1983;148:839-843.

23. Calleja JM, Esnaola S, Rubio R, Dura´n J. Comparison of a cardiorespiratory device versus polysomnography for diagnosis of sleep apnoea. Eur Respir J 2002;20: 1505–10.

24. Dingli K, Coleman EL, Vennelle M, Finch SP, Wraith PK, Mackay TW, Douglas NJ. Evaluation of a portable device for diagnosing the sleep apnoea/hypopnoea syndrome. Eur Respir J 2003; 21: 253-9.

25. Matrot B, Durand E, Dauger S, Vardon G, Gaultier C, Gallego J. Automatic classification of activity and apneas using whole body plethysmography in newborn mice. J Appl Physiol 2005; 98:365-370

26. Nakano H, Hayashi M, Ohshima E, Nishikata N, Shinohara T. Validation of a new system of tracheal sound analysis for the diagnosis of sleep apnea-hypopnea syndrome. Sleep 2004;27:951-7.

27. Farre R, Montserrat JM, Rotger M, Ballester E, Navajas D. Accuracy of thermistors and thermocouples as flow-measuring devices for detecting hypopnoeas. Eur Respir J 1998;11:179-82.

28. Berry RB, Koch GL, Trautz S, Wagner MH. Comparison of respiratory event detection by a polyvinylidene fluoride film airflow sensor and a pneumotachograph in sleep apnea patients. Chest 2005; 128:1331-1338

Fig. 1　Schematic display of signal processing in the algorithm.

A. The procedures to obtain flow-power time series from raw data of airflow signal are shown for a 60 seconds segment which includes an apnea event.

B. The filtered flow-power time series usually showed such a fall at an apnea / hypopnea event. We designate such a fall as flow-power dip, which is defined by a fall more than a threshold value within 30 seconds followed by a normalization part starting within 90 seconds. The computer algorithm detects any decrease as a possible start point of the flow-power dip and determines whether the following part satisfies the above conditions.

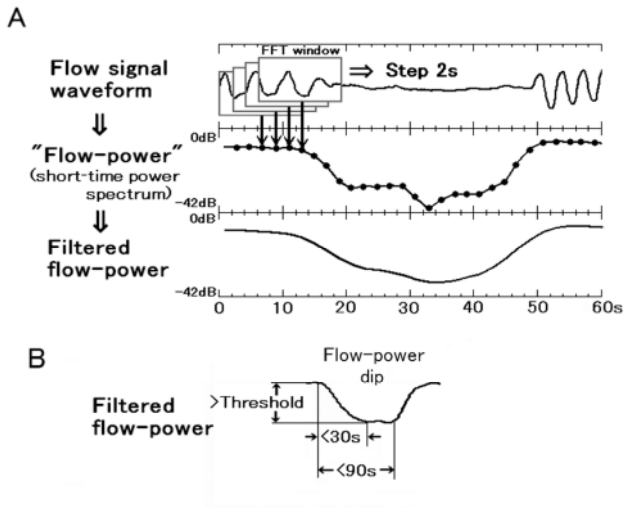

Fig. 2　In this 10 minutes flow trace, there are 14 events of apnea or hypopnea. Filtered flow power shows a clear transient decrease (flow-power dip; arrow) which corresponds to an apnea or hypopnea event. The change in SaO_2 (SaO_2 dip; arrow) follows that of flow-power with a time-lag of about 30 seconds.

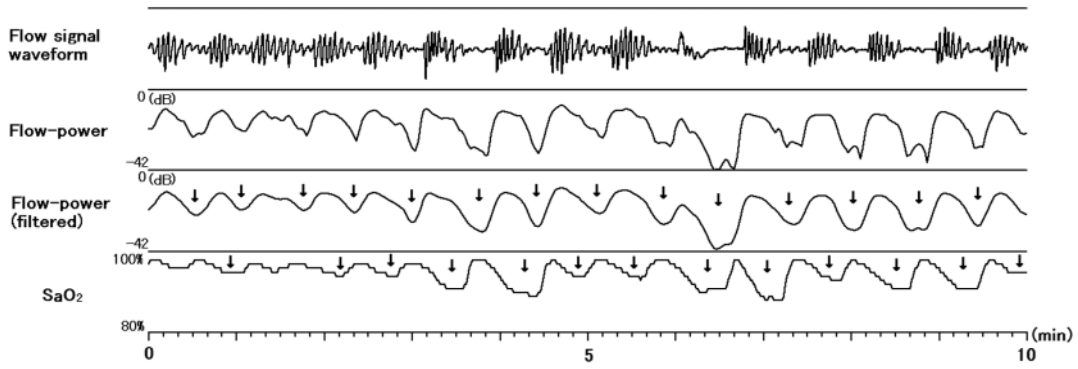

Fig. 3　Two cases with poor airflow signal quality.

The arrows in the flow-power trace indicate the positions of automatically detected flow-power dips. The boxes in the RIP-SUM traces indicate apnea / hypopnea events detected by visual inspection of full PSG records. A: The amplitude of the flow signal decreased in the latter part of the trace. All three events were correctly detected by the automatic flow analysis. B: The flow signal was very poor because of the contamination of noises. Although the flow signal was difficult to interpret visually, 7 out of 8 events were correctly detected by the automatic flow analysis.

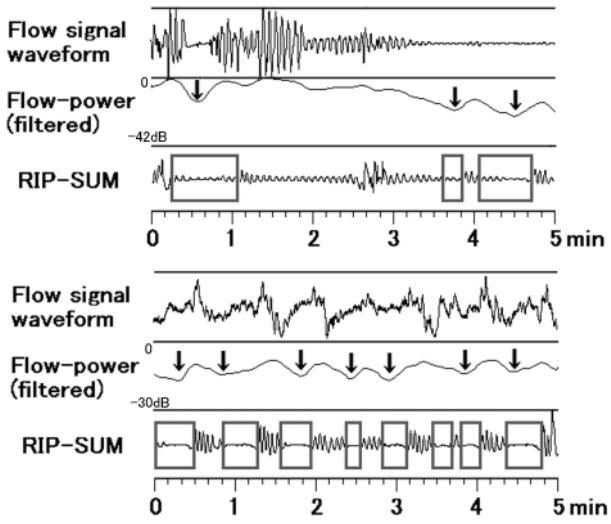

Fig. 4　A: Bland and Altman plots showing the variance between the flow-RDI and the AHI. The lines indicate the bias (mean difference) and the limit of agreement (mean ± 1.96SD). B: The relationship between the AHI and the flow-RDI vs. AHI difference. The line represents the regression line.

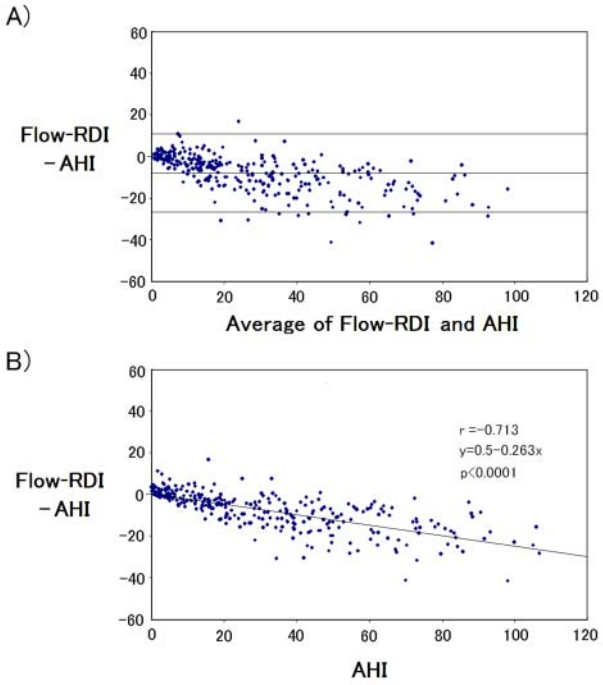

Fig. 5　The receiver operating characteristic (ROC) curve showing the relationship between diagnostic sensitivity and specificity of the flow-RDI in the validation group when AHI cutoff values of 5, 10 and 15 are used. The figures in the panels are the cutoff values of the flow-RDI. The figures in bold letters represent cutoff values of flow-RDI determined in the development group.

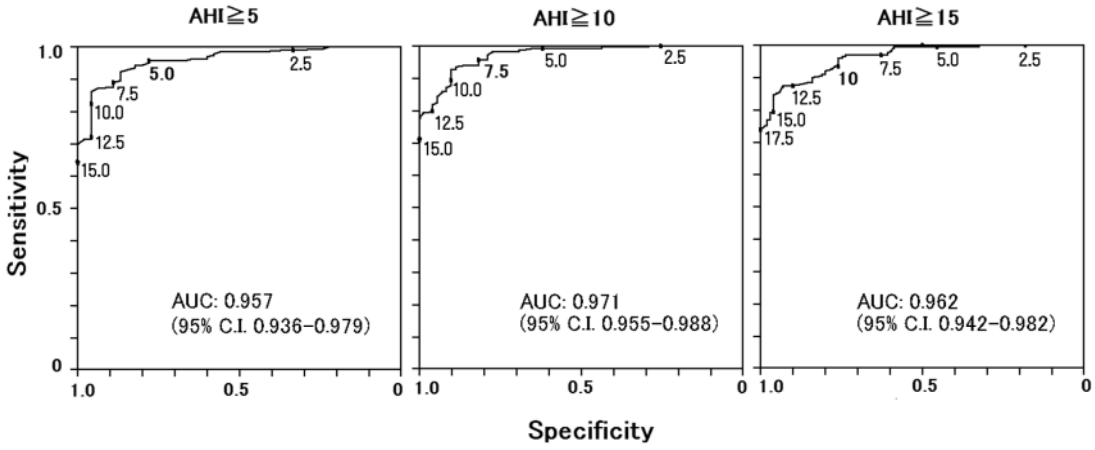

Fig. 6　Event-by-event agreement between the automatic flow analysis and full PSG manual analysis. The white rectangle represents the events detected by full PSG manual analysis, and the grey rectangle represents the flow-power dips detected by automatic flow analysis. The overlapped part represents the events detected by the both methods.

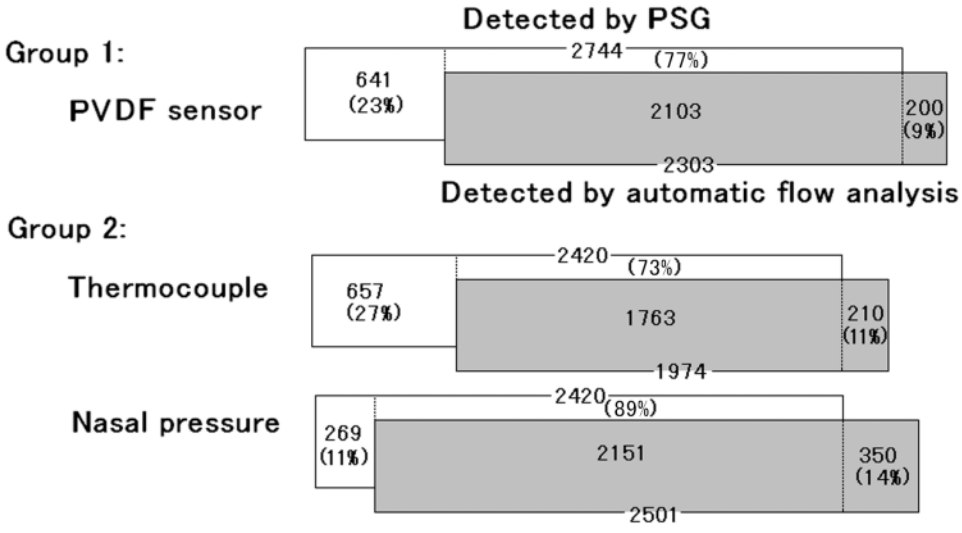

Table 1 Characteristics of subjects in group 1

	Development group		Validation group	
	Mean or [Median]	SD or [IQR]	Mean or [Median]	SD or [IQR]
n	100		299	
Men, no. (%)	79(79)		254(85)	
Age, y	51.1	13.7	48.6	13.9
BMI, kg/m2	25.8	4.1	26.0	4.3
Epworth Sleepiness Scale	10.4	5.1	11.3	5.4
Sleep efficiency, %	75	13.5	78.4	14
AI, no/h	[8.4]	[1.3-34.0]	[6.5]	[1.7-25.6]
AHI, no/h	[24.4]	[12.4-51.8]	[25.6]	[11.2-48.7]
AHI≧5, no. (%)	90(90)		254(85)	
AHI≧10, no. (%)	81(81)		228(76)	
AHI≧15, no. (%)	70(70)		200(67)	
ODI, no/h	[17.8]	[7.9-40.4]	[19.0]	[7.3-36.5]
%Time while SpO2<90%	[1.4]	[0.3-9.6]	[1.8]	[0.3-8.6]
Flow-RDI, no/h	[17.4]	[8.5-37.7]	[17.4]	[8.4-35.4]

IQR: interquartile range; BMI: body mass index; AI: apnea index; AHI: apnea-hypopnea index; ODI: oxygen desaturation index at the threshold of 3%;

Table 2 Agreement between the AHI and the flow-RDI

		Group 1	Group 2	
Flow sensor type used to obtain flow-RDI		PVDF sensor	Thermocouple	Nasal pressure
No. of subjects		299	116	117
Correlation between AHI and flow-RDI				
	r (95% C.I.)	0.94 (0.93, 0.96)	0.95 (0.93, 0.96)	0.97 (0.95, 0.98)
Difference between AHI and flow-RDI				
	Mean (95% C.I.)	-7.9 (-26.5, 10.7)	-7.1 (-23.2, 9.0)	-2.5 (-15.4, 10.4)
Disgnostic ability of flow-RDI*				
AHI>5	sensitivity / specificity	0.96 / 0.76	0.88 / 0.80	0.97 / 0.77
	AUC	0.96	0.95	0.95
AHI>10	sensitivity / specificity	0.96 / 0.82	0.92 / 0.90	0.97 / 0.76
	AUC	0.97	0.96	0.97
AHI>15	sensitivity / specificity	0.95 / 0.76	0.86 / 0.90	0.97 / 0.73
	AUC	0.96	0.95	0.98

95% C.I.: 95 % confidence interval; AHI: apnea-hypopnea index; AUC: area under the receiver-operating characteristic curve;
*The diagnostic sensitivity and specificity of the flow-RDI at three levels of AHI cutoff value are shown. The preset cutoff values of the flow-RDI are 5, 7.5, 10 for AHI cutoff values of 5, 10 and 15, respectively.