SLEEP-DISORDERED BREATHING

Predictive value of automated oxygen saturation analysis for the diagnosis and treatment of obstructive sleep apnoea in a home-based setting

V Jobin, P Mayer, F Bellemare

Thorax 2007;62:422–427. doi: 10.1136/thx.2006.061234

See end of article for authors' affiliations

Correspondence to: François Bellemare, PhD, Laboratoire du sommeil, CHUM–Hôtel-Dieu, 3840 St-Urbain, Montréal, Québec, Canada H2W 1T8; bellemare@videotron.qc.ca

Received 22 February 2006
Accepted 20 August 2006
Published Online First
24 January 2007

Obstructive sleep apnoea-hypopnoea (OSAH) is a highly prevalent disorder affecting 2–4% of adult men and women.¹ However, the standard diagnostic test for OSAH, that is, supervised in-laboratory overnight polysomnography (PSG), is time consuming and costly. Because of insufficient human, physical and financial resources, accessibility to this form of testing in some places is limited, and the diagnosis of OSAH is consequently delayed.² To overcome such difficulties, several portable monitors have been proposed that range in complexity from full PSG to oximetry alone.³ Although portable monitoring does not constitute a standard of care, in specialised centres such as ours it can help to reduce the burden for full PSG studies. Monitors that record respiratory variables together with oximetry but without electroencephalography (EEG) and electromyography (EMG) are particularly attractive because the same definition of OSAH as in standard PSG can be employed, and the test can be self-administered by patients in their home, thereby increasing accessibility. This type of monitoring offers high diagnostic sensitivity and specificity (>90%).⁴ However, given the number of signals being recorded, and because these studies are unattended, the risk of technical failures is higher. Furthermore, the scoring of such studies remains labour intensive as it requires expertise and manual editing.

Although a diagnosis of OSAH can also be made by oximetry alone,⁵ wide variability in sensitivity and specificity (range 31–100%) has been reported for this type of monitoring.^{3 4 6–11} Furthermore, the reliability of automatic scoring algorithms present on some of these devices has been questioned.^{3 4} However, in a large prospective study of consecutive patients, Vasquez *et al*¹² reported excellent agreement between the respiratory disturbance index (RDI) derived from the automated analysis of oxygen saturation (SpO₂) of the Remmers Sleep Recorder or RSR (RDI_{Rem}) and the apnoea-hypopnoea index (AHI) derived from the manual scoring of standard PSG

Background: A portable monitor for the automated analysis of episodic nocturnal oxygen saturation or SpO₂ (the Remmers Sleep Recorder, RSR) has been proposed for the diagnosis of obstructive sleep apnoea-hypopnoea (OSAH). A study was undertaken to compare the diagnostic performance of automated analysis with the manual scoring of polygraphic data by a more comprehensive respiratory monitor (the Suzanne recorder) used simultaneously in their intended home environment.

Methods: The respiratory disturbance indexes of the two monitors were compared in 94 consecutive adult patients suspected of having OSAH and who were deemed eligible for home-based investigation.

Results: The RSR overestimated the number of respiratory events associated with a $\geq 4\%$ fall in SpO₂ by 13% ($p < 0.005$) but underestimated the number of apnoeas and hypopnoeas defined on the basis of respiratory variables alone or their association with a $\geq 4\%$ fall in SpO₂ or autonomic arousals by 38–48% ($p < 0.0001$). In addition to these significant biases, the limits of agreement in all instances were wide, indicating a poor concurrence between the two monitors.

Conclusion: The automated analysis of SpO₂ with the RSR cannot be substituted for the manual scoring of polygraphic data with the more comprehensive respiratory monitor in the diagnosis of OSAH in an ambulatory home-based setting.

that was minimally altered when EEG-based arousals were incorporated in the definition of AHI. Compared with previous reports on oximetry, RDI_{Rem} offered excellent diagnostic sensitivity (>95%) and specificity (>85%). Because the analysis is automated and does not require a trained technician, it should be more widely applicable than more comprehensive respiratory monitors. The purpose of the present study was to assess the performance of this simpler monitor in comparison with the more comprehensive portable respiratory monitor currently employed in our ambulatory setting for the diagnosis of OSAH.

METHODS

Patients were recruited after referral to the sleep clinic of a university affiliated tertiary care centre. All patients who were prescribed an ambulatory trial by a sleep specialist were eligible for enrolment. In our centre, patients are generally considered eligible for an ambulatory investigation unless they meet one or more of the following exclusion criteria: inability to perform an ambulatory PSG, history of neuromuscular disease, severe lung disease (forced expiratory volume in 1 s (FEV₁) <50% predicted), unstable coronary artery disease or referral for parasomnia. Of the 513 consecutive patients eligible for ambulatory investigation between March 2003 and May 2004, 104 were enrolled in the study. The limit for enrolment was set by the number of available monitors (six Suzanne recorders and one RSR; see below). The RSR was offered to patients on a “first come first served” basis. The protocol was approved by the Institutional Human Ethics Committee of CHUM–Hôtel Dieu and all participants provided written informed consent.

Abbreviations: AHI, apnoea-hypopnoea index; ESS, Epworth Sleepiness Scale; FEV₁, forced expiratory volume in 1 s; OSAH, obstructive sleep apnoea-hypopnoea; PSG, polysomnography; RDI, respiratory disturbance index; RSR, Remmers Sleep Recorder; SpO₂, oxygen saturation

The number of patients recruited was sufficient to detect a difference in RDI of >5 events/h between the two monitors with a probability of 0.01 and a power of 90%.

Data acquisition devices

A computerised polysomnographic monitor (Suzanne Polysomnographic Recording System, Nellcor Puritan Bennett (Melville) Ltd, Ottawa, Ontario, Canada), hereafter referred to as the Suzanne recorder, and a standardised montage served as the reference for the study. Nasal pressure was recorded by means of a nasal cannula, breathing movements by means of piezoelectric respiratory effort belts (Ultima Respiratory Effort Sensor, Model 0522S, Braedon Medical Corp, Carp, Ontario, Canada) placed around the torso and abdomen, body position by a sensor placed over the thorax and held in place by a thoracic belt, SpO_2 by a finger probe (Oximax Durasensor, Model DS-100A, Tyco Healthcare, Pleasanton, California, USA) and snoring by means of a miniature microphone attached to the garments (Suzanne Microphone, Model Y-70615, Nellcor Puritan Bennett (Melville) Ltd). With the RSR (SnoreSat, SagaTech Electronics Inc, Calgary, Alberta, Canada), SpO_2 was recorded by means of a second finger probe (Oximax Durasensor, Model DS-100A, Tyco Healthcare) attached to the free hand as well as snoring and body position changes by means of a specially designed sensor fixed with adhesive tape over the sternal notch.

Study protocol

After medical assessment the patients were asked to complete a sleep quality questionnaire incorporating the Epworth Sleepiness Scale (ESS).¹³ They were then instructed how to use the two portable devices simultaneously at night. They were asked to indicate on a log the time they went to bed, the time lapse before sleep onset, the wake time and any problems that occurred during recording. The following morning the data from the two recorders were downloaded to a desktop computer for analysis (see below). At study completion a follow-up questionnaire was sent to all treated patients attending the sleep clinic. Specifically, patients were asked to report the number of months they had been on treatment, whether they were still using their device and, on average, how many hours per night. They were also asked to report whether they were satisfied with their treatment and to complete the ESS questionnaire.

Data analysis

The data from the two monitors were analysed separately and in a blind fashion by one of three trained sleep technicians. Data from the Suzanne recorder were scored manually using Sandman software (Nellcor Puritan Bennett (Melville) Ltd).¹⁴ A RDI was calculated based on the number of apnoeas and hypopnoeas per hour of recording time (RDI_{Suz}). Apnoea was scored when no breathing could be detected on the nasal pressure trace for at least 10 s. Hypopnoea was scored when the

amplitude of respiratory movements detected by the nasal pressure sensor or the thoracoabdominal effort belts decreased by more than 50% below baseline or when the amplitude of respiration was reduced by less than 50% but was followed by a fall in SpO_2 of $\geq 4\%$.¹⁴ An extended RDI ($\text{RDI}_{\text{Suz}+}$) was also calculated that incorporated, in addition to apnoeas and hypopnoeas, respiratory events associated with probable autonomic arousals,¹⁵ defined as increases in pulse rate of >5 beats/min immediately after a respiratory event characterised by the clear presence of inspiratory flow limitation, a decrease in flow of $<30\%$ of the baseline value and no change in SpO_2 . Finally, for direct comparison with RDI_{Rem} , a restricted RDI ($\text{RDI}_{\text{Suz}-}$) was defined as the number of respiratory events associated with a fall in SpO_2 of $\geq 4\%$. The Suzanne recorder samples the SpO_2 signal at 12 Hz and reports the average value over the preceding second every second.

For the RSR, no specific scoring was done since the recordings are automatically analysed by the software supplied with the instrument. With this software the raw data can be viewed but no editing is possible. RDI_{Rem} is calculated automatically as the number of episodes per hour of oximeter probe-on time with falls in $\text{SpO}_2 \geq 4\%$ below baseline. The RSR samples the SpO_2 signal at 1 Hz. Details concerning the detection algorithm of this monitor can be found elsewhere.¹² RDI_{Rem} appears automatically in the report sheet along with a graphic display of SpO_2 , snoring and body position as well as a series of related information including recording time, oxygen sensor probe-on time, minimum, maximum and mean SpO_2 . This report can be visualised and printed.

Statistical analysis

Differences between the two devices were compared using non-parametric statistics and the Wilcoxon signed rank test. The bias and limits of agreement between the two instruments were calculated¹⁶ and receiver operating characteristic (ROC) curves generated for selected diagnostic RDI cut-off values. Values are expressed as group mean (SEM) or median (interquartile range, IQR), depending on their distribution. The statistical level of significance was set at $p < 0.05$.

RESULTS

Patient characteristics

Of the 104 patients studied, 10 returned technically inadequate studies, leaving a total of 94 patients (62 men) for final analysis (table 1).

Technical considerations

Differences in recording time between the two monitors, although significant, were small (table 2). By contrast, wide variations were found in the time spent in different body positions (data not shown), presumably related to the different features or displacement of body position sensors, preventing meaningful comparisons from being made in different body

Table 1 Characteristics of study subjects

	Minimum	Maximum	Mean	SEM
Age	22	76	49.27	1.12
Body weight (kg)	51	178	89.51	2.20
Height (m)	1.47	2.16	1.72	0.01
BMI (kg/m^2)	16.1	55.9	30.47	0.78
Neck circumference (cm)	31	53	39.92	0.40
Adjusted neck circumference (cm)*	36	59	45.71	0.53
Epworth sleepiness score	0	21	11.72	0.53

BMI, body mass index calculated as the ratio of body weight (in kg) and height squared (in metres).

*Clinical prediction score which is equal to the value of neck circumference (in cm) adjusted for history of hypertension ($+4$ cm) and reporting of frequent snoring ($+3$ cm) and of witnessed apnoeas ($+3$ cm). For details see Flemmons.²¹

Table 2 Comparative results between the Suzanne and Remmers monitors

	Suzanne	Remmers	p Value*
RDI (n/h)	12.9 (5.4–28)	8.8 (2.9–18.2)	<0.001
Recording time (min)	431 (379–496)	429 (381–486)	<0.022
SpO ₂ mean (%)	95.9 (94.9–96.8)	95.1 (93.9–96.1)	<0.001
SpO ₂ min (%)	88.5 (84–91)	85 (80–87)	<0.001
ΔSpO ₂ (%)	7.7 (4.8–11.1)	10.4 (7.9–13.9)	<0.001
% time SpO ₂ <90% (%)	0.15 (0–1.4)	0.6 (0.1–1.6)	<0.001

*Wilcoxon signed rank test.

Values are median (interquartile range).

RDI, respiratory disturbance index measured as the number of events/h recording time; SpO₂ mean, mean oxygen saturation measured by pulsed oximetry over the recording time period; SpO₂ min, minimum oxygen saturation recorded by pulse oximetry during the recording time period; ΔSpO₂, arithmetic difference between SpO₂ mean and SpO₂ min.

Figure 1 Bland-Altman plots showing the distribution of the differences between the respiratory disturbance index (RDI) of the Remmers Sleep Recorder (RDI_{Rem}) and the RDI of the Suzanne recorder calculated according to several definitions of respiratory events (RDI_{Suz+}, extended RDI; RDI_{Suz-}, restricted RDI). For a definition of the Suzanne recorder RDIs, see footnote to table 3. In each panel the horizontal solid line represents the mean difference and the broken lines ± 2 SD of the observed differences.

Figure 2 Receiver operating characteristic (ROC) curves of the respiratory disturbance index (RDI) of the Remmers Sleep Recorder for two diagnostic cut-off values of the Suzanne recorder RDI calculated according to several definitions of respiratory events: Solid lines represent events defined on the basis of their association with a $\geq 4\%$ fall in SpO₂ (RDI_{Suz-}). Dotted lines represent events defined on the basis of a $\geq 50\%$ reduction in mechanical variables or $\geq 4\%$ fall in SpO₂ (RDI_{Suz+}). Dash-dotted lines represent events defined on the basis of their association with a $\geq 50\%$ reduction in mechanical variables or a $\geq 4\%$ fall in SpO₂ or autonomic arousals (RDI_{Suz+}).

Table 3 Predictive value of automated analysis of the Remmers sleep recorder

Diagnostic cut-off	ROC area	Sensitivity (%)	Specificity (%)	Patients correctly classified (%)	LR+	LR–
RDI_{Suz-}						
≥ 5	0.95	94.1	74.4	85.1	3.68	0.08
≥ 10	0.96	90.9	82.0	85.1	5.04	0.11
≥ 15	0.99	95.2	87.7	89.4	7.72	0.05
RDI_{Suz}						
≥ 5	0.85	75.3	81.0	76.6	3.95	0.30
≥ 10	0.90	67.9	87.8	76.6	5.57	0.37
≥ 15	0.91	62.8	96.1	80.9	16.0	0.39
RDI_{Suz+}						
≥ 5	0.84	73.7	83.3	75.5	4.42	0.31
≥ 10	0.82	60.6	96.4	71.3	17.0	0.41
≥ 15	0.94	51.9	97.5	71.3	20.7	0.49

RDI_{Suz-} , diagnostic cut-off value of the respiratory disturbance index (RDI) calculated by the Suzanne recorder based on respiratory events associated with a fall in SpO_2 of at least 4%; RDI_{Suz} , diagnostic cut-off value of the RDI calculated by the Suzanne recorder; RDI_{Suz+} , diagnostic cut-off value of the extended RDI calculated by the Suzanne recorder and incorporating respiratory events associated with autonomic arousals; ROC area, area under the receiver operating characteristic curve; LR+, likelihood ratio for a positive test result; LR–, likelihood ratio for a negative test result.

postures. Mean and minimum SpO_2 from these two monitors differed significantly, the RSR yielding consistently lower readings. The difference between mean and minimum SpO_2 was also significantly greater for the RSR, suggesting greater sensitivity of its oxygen sensor.

Measures of agreement

The mean (SE) of the differences between RDI_{Rem} and RDI_{Suz-} was small but statistically significant (1.58 (0.76); $p < 0.0001$), indicating a positive instrument bias. The distribution of these differences is shown in fig 1 in the form of a Bland-Altman plot. Visual inspection indicated that systematic differences occurred near the diagnostic threshold (ie, for RDI values between 5 and 20 events/h, where the impact of the discrepancy was most significant).

The mean (SE) values of the differences between RDI_{Rem} and RDI_{Suz} (–6.04 (0.94); $p < 0.0001$) and between RDI_{Rem} and RDI_{Suz+} (–10.16 (1.03); $p < 0.0001$) were larger, reflecting the addition of respiratory events with $< 4\%$ falls in SpO_2 and with autonomic arousals, respectively. These negative biases are substantial, representing approximately 38% and 48% of all respiratory events included in RDI_{Suz} and RDI_{Suz+} , respectively. The Bland-Altman plots in fig 2 illustrate the distribution of these differences. Again, visual inspection indicated that large differences between the two recorders occurred near the diagnostic threshold where these differences mattered. Furthermore, the limits of agreement were wide, suggesting poor agreement between the two monitors.

Predictive value of the automated analysis

The sensitivity and specificity of RDI_{Rem} varied with the diagnostic RDI cut-off value selected as well as with the

RDI_{Suz} definition (table 3, fig 2). Compared with RDI_{Suz-} , RDI_{Rem} had reasonable sensitivity at the same cut-off points but specificity (74%) and likelihood ratio for a positive test result (3.68) near the diagnostic threshold were low. When compared with RDI_{Suz} and RDI_{Suz+} , the sensitivity of RDI_{Rem} was markedly reduced at all cut-off points, but specificity was minimally altered. Because of this lower sensitivity, the likelihood ratio for a negative test result increased markedly at all cut-off points. At the diagnostic cut-off value of ≥ 5 events/h, only between 76% and 77% of the patients were correctly classified by RDI_{Rem} . At this cut-off point, a diagnosis of OSAH could be made with confidence (ie, 100% specificity) when RDI_{Rem} was ≥ 20 events/h ($RDI_{Suz} \geq 5$) or ≥ 10 events/h ($RDI_{Suz+} \geq 5$). With these cut-off points, however, only between 29% and 53% of OSAH patients were identified by RDI_{Rem} .

Outcome

Of the 94 patients, 49 met the minimum diagnostic criteria (≥ 5 events/h) using RDI_{Suz-} , of whom two had RDI_{Rem} of < 5 (Table 4). Adding respiratory events with $> 50\%$ reduction in respiratory movements but $< 4\%$ fall in SpO_2 in the definition of RDI_{Suz} led to 25 additional patients meeting this criterion, of whom 16 had $RDI_{Rem} < 5$. Adding respiratory events associated with autonomic arousals in the definition of RDI_{Suz+} led to 12 additional patients meeting the minimum diagnostic criteria, of whom 9 had $RDI_{Rem} < 5$.

Of these 94 patients, 13 were referrals from external centres. Of the 81 patients attending our clinic, 5 were lost to follow-up and 5 were referred for a second evaluation, leaving 71 new cases. Fifty-five patients with $RDI_{Suz+} \geq 5$ events/h were prescribed a treatment trial with continuous positive airway pressure or a mandibular advancement prosthesis. A follow-up

Table 4 Patient distribution according to minimum required diagnostic criteria

Criterion	n	RDI_{Rem}	RDI_{Suz-}	RDI_{Suz}	RDI_{Suz+}	ESS	n/Rx	ΔESS
$RDI_{Suz+} < 5$	8	1.9 (0.7)	1.0 (0.3)	2.0 (0.3)	3.7 (0.3)	12.9 (2.1)	0	–
$RDI_{Suz-} \geq 5$	49	23.2 (2.5)	22.4 (2.8)	32.4 (2.9)	36.7 (2.8)	11.3 (0.7)	27/35	–5.2 (1.0)
$RDI_{Suz} \geq 5$	25	4.4 (0.5)	1.7 (0.2)	9.9 (1.0)	14.3 (1.5)	12.0 (1.1)	7/12	–3.4 (1.7)
$RDI_{Suz+} \geq 5$	12	5.1 (1.7)	2.2 (0.5)	3.3 (0.4)	7.6 (0.8)	12.1 (1.1)	4/8	–6.8 (2.8)
Total	94	14.1 (1.7)	12.5 (1.8)	20.1 (2.0)	24.2 (2.0)	11.7 (0.5)	38/55	–5.0 (0.8)

Values are mean (SE).

RDI_{Rem} , respiratory disturbance index (RDI) of the Remmers Sleep Recorder; RDI_{Suz-} , restricted RDI of the Suzanne recorder; RDI_{Suz} , RDI of the Suzanne recorder; RDI_{Suz+} , extended RDI of the Suzanne recorder; ESS, Epworth sleepiness score at baseline; n/Rx, number of respondents at follow-up over number of treated patients; ΔESS , change in Epworth sleepiness score with treatment relative to baseline.

questionnaire sent to all these patients 6 months to 1 year after completion of the study was returned by 41. Compliance with treatment, defined as the number of patients who reported using their device regularly, was 95%, and all but 4 patients were satisfied with their treatment. Patients reported having used their device for 1.5–9 h/day over a period of 0.5–17 months. The ESS score decreased from an average of 13.0 (range 3–21) before treatment to 8.0 (range 1–14) while on treatment ($p < 0.0005$; $n = 38$). The response to treatment was independent of whether the diagnosis was based on RDI_{Suz-} , RDI_{Suz} or RDI_{Suz+} criteria (equality of medians test, $p = 0.45$; table 4).

DISCUSSION

This study shows that the limits of agreement of the two monitors are large, leading to low sensitivity and specificity of the RSR. As an oximeter, the automated analysis algorithm of the RSR overestimated by 13% the number of respiratory events with $\geq 4\%$ fall in SpO_2 . As an apnoea-hypopnoea monitor, the RSR underestimated the number of respiratory events by 38–48%. Only between 29% and 53% of OSAH patients, identified on the basis of RDI_{Suz} or RDI_{Suz+} , were identified with confidence by the RSR.

Comparison of RDI_{Rem} and RDI_{Suz}

In a previous study by Vasquez *et al*¹² no significant bias was found between RDI_{Rem} and PSG-derived AHI, even when EEG-based arousals were included in the definition of AHI. In the present work, a significant positive instrument bias was apparent when respiratory events were defined on the basis of their association with $\geq 4\%$ fall in SpO_2 , RDI_{Rem} exceeding RDI_{Suz-} by 13%. By contrast, when respiratory events associated with a $< 4\%$ fall in SpO_2 or with autonomic arousals were included in the definition of RDI_{Suz} , a negative instrument bias was evident, RDI_{Rem} underestimating RDI_{Suz} by 38% and RDI_{Suz+} by 48%.

RDI_{Rem} values in excess of RDI_{Suz-} may have been caused by inadvertent changes in the SpO_2 signal unrelated to respiratory events detected by automated analysis with RSR, but filtered out by manual scoring with the Suzanne polygraphic recorder. However, as the data in table 2 suggest, differences in the sensitivity of the oxygen sensors or analysis algorithms of the two recorders may also contribute to explain this finding.

On the other hand, RDI_{Rem} values smaller than RDI_{Suz} or RDI_{Suz+} are probably explained by different definitions of respiratory events. This should also contribute to an explanation of the difference between our results and those of Vasquez *et al*. In their study, hypopnoea would be scored if there was a fall in SpO_2 of $\geq 4\%$ whereas, in our work, a fall in SpO_2 of this magnitude was required only when the decrease in nasal flow or thoracoabdominal movements was $< 50\%$.¹⁴ Because apnoeas and hypopnoeas longer than 10 s duration can occur without significant changes in SpO_2 ,¹⁷ events associated with a fall in SpO_2 of $< 4\%$ would contribute to RDI_{Suz} and RDI_{Suz+} but not to RDI_{Rem} , thereby increasing the bias in our study compared with theirs.

In our investigation the inclusion of autonomic arousals in the definition of RDI_{Suz+} heightened instrument bias significantly whereas, in the study by Vasquez *et al*, no significant bias was noted when EEG-based arousals were included in the definition of hypopnoea.¹² In their trial, however, the inclusion of EEG-based arousals did not change the AHI significantly, implying a negligible contribution of these events to AHI. By contrast, in the present study, using autonomic arousals as a surrogate for EEG microarousals,¹⁵ RDI_{Suz+} significantly exceeded RDI_{Suz} , explaining the greater bias found when this criterion was included in the definition of RDI.

Although a negative bias could be anticipated in our study based on the different event definitions, its magnitude could not be predicted. In a previous investigation by Redline *et al*, adding a $\geq 4\%$ desaturation criterion as a prerequisite for scoring a respiratory event reduced the median RDI by 85% (from 29.3 to 4.4) and the prevalence estimate by about 55% (from 100% to 45%).¹⁸ In that study, however, flow was measured with a thermocouple, a method which may not reflect flow accurately.¹⁴ The closest comparison in our investigation would be between RDI_{Suz} and RDI_{Suz-} , showing a 66% reduction in median RDI (from 12.9 to 5.6) and a 26% decrease in the prevalence estimate (from 78% to 52%). The fact that, on average, 38–48% of respiratory events in the present study could be associated with a fall in SpO_2 of $< 4\%$ may be understood when considering that the duration of the events as well as the pre-event oxygen saturation level, end-expiratory lung volume and metabolic rate all contribute to determine the magnitude of the fall in SpO_2 during obstructive events.¹⁷ Because these factors can be expected to vary substantially in different patients, they may also contribute to explain the large inter-individual variability in the difference in RDI and, hence, the poor agreement between the two monitors.

One could question the clinical significance of including respiratory events associated with $< 4\%$ fall in SpO_2 or with autonomic arousals in the definition of RDI_{Suz} . As shown by Guilleminault *et al*,¹⁹ sleep fragmentation can occur in a number of patients with OSAH in association with periodic increments of respiratory efforts but without significant oxygen desaturation. These patients can be very symptomatic and benefit from treatment. Furthermore, as demonstrated experimentally by Martin *et al*,²⁰ even subcortical autonomic arousals can increase sleepiness. In the present study the inclusion of respiratory events with $< 4\%$ fall in SpO_2 or with autonomic arousals in the definition of RDI_{Suz} led to 37 additional patients being classified as having OSAH, of whom 50% were prescribed a treatment trial with benefits in most of them (table 4). These data suggest that a substantial number of patients with OSAH identified on the basis of these criteria can benefit from treatment, thereby justifying their inclusion in the definition of RDI.

Implications of the findings

Significant negative bias and poor agreement between the two monitors both contributed to reduce the diagnostic sensitivity and specificity of the RSR. Using a diagnostic cut-off value of ≥ 15 events/h, Vasquez *et al* reported 98% sensitivity,¹² a value comparable to that of the present study when a $\geq 4\%$ fall in SpO_2 was required to define a respiratory event with the Suzanne recorder. However, when respiratory events associated with a $< 4\%$ fall in SpO_2 or with autonomic arousals were included in the definition of RDI_{Suz} , the sensitivity of the RSR at the same cut-off point was markedly reduced to between 52% and 63%. These sensitivities are within the range reported previously for home oximetry (range 31–98%).⁴ The implication of this is that, for a substantial number of patients, additional investigations would have been required had the RSR been employed as the sole instrument. According to our results, for a positive diagnosis of OSAH (defined as ≥ 5 events/h) to be made with 100% confidence using the RSR would require that RDI_{Rem} be ≥ 20 or ≥ 10 events/h, depending on whether autonomic arousals are included or not in the definition of RDI_{Suz} . Because only between 33% and 49% of all cases met these criteria, a second study with a more comprehensive respiratory monitor would have been needed in 51% or more of patients. Although a cost-benefit analysis of the RSR under such circumstances has not been performed, the results are unlikely to show significant benefits. Furthermore, because a

second study would entail additional delay in the diagnosis of OSAH in a substantial number of patients, substitution of the more comprehensive respiratory monitor by the automated analysis of the RSR is not warranted.

With the RSR, no editing of RDI_{Rem} was possible. Spo₂ traces, however, could be reviewed. Other signals such as flow and respiratory effort belts may also be added and reviewed. As several studies have pointed out, visual inspection of Spo₂ tracings can improve the diagnostic sensitivity of portable monitors. In an investigation by Douglas *et al.*,⁸ 66% of patients with OSAH were correctly identified by visual inspection of Spo₂ traces alone with no false positives. Using a $\geq 4\%$ fall in Spo₂, only 41% of patients with OSAH were correctly identified with 97% specificity. In a study by Sériès *et al.*⁹ the visual appreciation of periodic fluctuations in Spo₂ correctly identified 98% of patients with OSAH with a specificity, however, of only 48%. Although this has not been examined in the present work, it is conceivable that visual inspection of Spo₂ traces and of other signals such as flow by a sleep specialist could improve the diagnostic sensitivity of the RSR.

Another limitation of our study is that the performance of the RSR in the home environment was not compared with diagnostic in-laboratory PSG. Differences between EEG-based sleep time and recording time, night to night variability and different environments are all factors that can affect RDI and, hence, the diagnostic performance of the RSR. By using the two recorders simultaneously in the home environment, we avoided these potential confounders. While this has advantages with respect to defining the performance of automated analysis, further validation comparing home-based monitoring with in-laboratory PSG is warranted. However, because the added variability associated with the home environment can be expected to reduce the diagnostic sensitivity and specificity of the RSR further, such a comparison is likely to reinforce our conclusions.

In summary, this study shows that, in spite of a significant association between the RDI of the two monitors, the low diagnostic sensitivity and specificity of RDI_{Rem} near the diagnostic threshold are such as to deter substitution of the more comprehensive respiratory monitor.

ACKNOWLEDGEMENTS

The Remmers Sleep Recorder used in this study was provided by Sagatech Electronics Inc, Calgary, Alberta, Canada. The authors acknowledge the editorial assistance of Ovid Da Silva, Research Support Office, Research Center, CHUM.

Authors' affiliations

V Jobin, P Mayer, F Bellemare, Laboratoire du sommeil, Centre hospitalier de l'Université de Montréal (CHUM)—Hôtel-Dieu, Montréal, Québec, Canada

FB was supported by Medigas, OSR Medical and the Laboratoires Biron.

Competing interests: None.

REFERENCES

- 1 Young T, Palta M, Dempsey J, *et al.* The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;**328**:1230–5.
- 2 Flemons WW, Douglas NJ, Kuna ST, *et al.* Access to diagnosis and treatment of patients with suspected sleep apnea. *Am J Respir Crit Care Med* 2004;**169**:668–72.
- 3 Ferber R, Millman R, Coppola M, *et al.* ASDA standards of practice: portable recording in the assessment of obstructive sleep apnea. *Sleep* 1994;**17**:378–92.
- 4 Flemons WW, Littner MR, Rowley JA, *et al.* Home diagnosis of sleep apnea: a systematic review of the literature. *Chest* 2003;**124**:1543–79.
- 5 ATS/ACCP/AASM Taskforce Steering Committee. Executive summary on the systematic review and practice parameters for portable monitoring in the investigation of suspected sleep apnea in adults. *Am J Respir Crit Care Med* 2004;**169**:1160–3.
- 6 Cooper BG, Veale D, Griffiths CJ, *et al.* Value of nocturnal oxygen saturation as a screening test for sleep apnea. *Thorax* 1991;**46**:586–8.
- 7 Stradling JR, Crosby JH. Predictors and prevalence of obstructive sleep apnea and snoring in 1001 middle aged men. *Thorax* 1991;**46**:85–90.
- 8 Douglas NJ, Thomas S, Jan MA. Clinical value of polysomnography. *Lancet* 1992;**339**:347–50.
- 9 Sériès F, Marc I, Cormier Y, *et al.* Utility of nocturnal home oximetry for case finding in patients with suspected sleep apnea hypopnea syndrome. *Ann Intern Med* 1993;**119**:449–53.
- 10 Gyulay S, Olson LG, Hensley MJ, *et al.* A comparison of clinical assessment and home oximetry in the diagnosis of obstructive sleep apnea. *Am Rev Respir Dis* 1993;**147**:50–3.
- 11 Olson LG, Ambrogetti A, Gyulay SG. Prediction of sleep-disordered breathing by unattended overnight oximetry. *J Sleep Res* 1999;**8**:51–5.
- 12 Vasquez J, Tsai WH, Flemons WW, *et al.* Automated analysis of digital oximetry in the diagnosis of obstructive sleep apnea. *Thorax* 2000;**55**:302–7.
- 13 Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. *Chest* 1993;**103**:30–6.
- 14 American Academy of Sleep Medicine Task Force. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* 1999;**22**:667–88.
- 15 Pitson DJ, Stradling JR. Autonomic markers of arousal during sleep in patients undergoing investigation for obstructive sleep apnea, their relationship to EEG arousals, respiratory events and subjective sleepiness. *J Sleep Res* 1998;**7**:53–9.
- 16 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.
- 17 Findley LJ, Ries AL, Tisi GM, *et al.* Hypoxemia during apnea in normal subjects: mechanisms and impact of lung volume. *J Appl Physiol* 1983;**55**:1777–83.
- 18 Redline S, Kapur VK, Sanders MH, *et al.* Effects of varying approaches for identifying respiratory disturbances on sleep apnea assessment. *Am J Respir Crit Care Med* 2000;**161**:369–74.
- 19 Guilleminault C, Stoohs R, Clerk A, *et al.* A cause of excessive daytime sleepiness. The upper airway resistance syndrome. *Chest* 1993;**104**:781–7.
- 20 Martin SE, Wraith PK, Deary IJ, *et al.* The effect of nonvisible sleep fragmentation on daytime function. *Am J Respir Crit Care Med* 1997;**155**:1596–601.
- 21 Flemons WW. Obstructive sleep apnea. *N Engl J Med* 2002;**347**:498–504.

Access the latest content chosen by our Editors

BMJ Journals editors select an article from each issue to be made free online immediately on publication. Other material is free after 12 months to non-subscribers. Access the Editor's Choice from the home page—or expand your horizons and see what the other BMJ Journals editors have chosen by following the links on any BMJ Journal home page.