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

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 ≥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 ≥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.

Obstructive sleep apnoea-hypopnoea (OSAH) is a highly prevalent disorder affecting 2–4% of adult men and women. 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. To overcome such difficulties, several portable monitors have been proposed that range in complexity from full PSG to oximetry alone. 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%). 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, wide variability in sensitivity and specificity (range 31–100%) has been reported for this type of monitoring. Furthermore, the reliability of automatic scoring algorithms present on some of these devices has been questioned. However, in a large prospective study of consecutive patients, Vasquez et al 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 (RDIRem) and the apnoea-hypopnoea index (AHI) derived from the manual scoring of standard PSG that was minimally altered when EEG-based arousals were incorporated in the definition of AHI. Compared with previous reports on oximetry, RDIRem 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 nasoanalanta, 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, SpO2 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), SpO2 was recorded by means of a second finger probe (Oximax Durasensor, Model DS-100A, Tyco Healthcare) attached to the free hand as well as recording 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). 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). A RDI was calculated based on the number of apnoeas and hypopnoeas per hour of recording time (RDI\textsubscript{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 SpO2 of >4%. An extended RDI (RDI\textsubscript{Suz}\textsubscript{E}) was also calculated that incorporated, in addition to apnoeas and hypopnoeas, respiratory events associated with probable autonomic arousals, 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 SpO2. Finally, for direct comparison with RDIRem, a restricted RDI (RDI\textsubscript{Suz\textsubscript{R}}) was defined as the number of respiratory events associated with a fall in SpO2 of >4%. The Suzanne recorder samples the SpO2 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. RDIRem is calculated automatically as the number of episodes per hour of oximeter probe-on time with falls in SpO2 >4% below baseline. The RSR samples the SpO2 signal at 1 Hz. Details concerning the detection algorithm of this monitor can be found elsewhere. RDIRem appears automatically in the report sheet along with a graphic display of SpO2, snoring and body position as well as a series of related information including recording time, oxygen sensor probe-on time, minimum, maximum and mean SpO2.

**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 calculated 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 positions.

---

**Table 1** Characteristics of study subjects

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>22</td>
<td>76</td>
<td>49.27</td>
<td>1.12</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>51</td>
<td>178</td>
<td>89.51</td>
<td>2.20</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.47</td>
<td>2.16</td>
<td>1.72</td>
<td>0.01</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>16.1</td>
<td>55.9</td>
<td>30.47</td>
<td>0.78</td>
</tr>
<tr>
<td>Neck circumference (cm)</td>
<td>31</td>
<td>53</td>
<td>39.92</td>
<td>0.40</td>
</tr>
<tr>
<td>Adjusted neck circumference (cm)*</td>
<td>36</td>
<td>59</td>
<td>43.71</td>
<td>0.53</td>
</tr>
<tr>
<td>Epworth sleepiness score</td>
<td>0</td>
<td>21</td>
<td>11.72</td>
<td>0.53</td>
</tr>
</tbody>
</table>

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.
Table 2  Comparative results between the Suzanne and Remmers monitors

<table>
<thead>
<tr>
<th></th>
<th>Suzanne</th>
<th>Remmers</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDI (n/h)</td>
<td>12.9 (5.4–28)</td>
<td>8.8 (2.9–18.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Recording time (min)</td>
<td>431 (379–496)</td>
<td>429 (381–486)</td>
<td>&lt;0.022</td>
</tr>
<tr>
<td>SpO2 mean (%)</td>
<td>95.9 (94.9–96.8)</td>
<td>95.1 (93.9–96.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SpO2 min (%)</td>
<td>88.5 (84–91)</td>
<td>85 (80–87)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ΔSpO2 (%)</td>
<td>7.7 (4.8–11.1)</td>
<td>10.4 (7.9–13.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>% time SpO2 &lt;90% (%)</td>
<td>0.15 (0–1.4)</td>
<td>0.6 (0.1–1.6)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*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 pulse 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 (RDIRem) and the RDI of the Suzanne recorder calculated according to several definitions of respiratory events (RDISuz+, extended RDI; RDISuz−, restricted RDI). For a definition of the Suzanne recorder RDI, see footnote to table 3. In each panel the horizontal solid line represents the mean difference and the broken lines ± 2SD 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 >4% fall in SpO2 (RDISuz+). Dotted lines represent events defined on the basis of a >50% reduction in mechanical variables or >4% fall in SpO2 (RDISuz). Dash-dotted lines represent events defined on the basis of their association with a >50% reduction in mechanical variables or a >4% fall in SpO2 or autonomic arousals (RDISuz−).
postures. Mean and minimum \text{SpO}_2 from these two monitors differed significantly, the RSR yielding consistently lower readings. The difference between mean and minimum \text{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 RDIRem and RDISuz 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 RDIRem and RDISuz (−6.04 (0.94); p<0.0001) and between RDIRem and RDISuz (−10.16 (1.03); p<0.0001) were larger, reflecting the addition of respiratory events with <4% falls in \text{SpO}_2 and with autonomic arousals, respectively. These negative biases are substantial, representing approximately 38% and 48% of all respiratory events included in RDIRem and RDISuz, 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 RDIRem varied with the diagnostic RDI cut-off value selected as well as with the RDISuz definition (table 3, fig 2). Compared with RDISuz, RDIRem 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 RDISuz and RDISuz, the sensitivity of RDIRem 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 RDIRem. At this cut-off point, a diagnosis of OSAH could be made with confidence, since 100% specificity when RDIRem was ≥20 events/h (RDISuz ≥5) or ≥10 events/h (RDISuz ≥5). With these cut-off points, however, only between 29% and 53% of OSAH patients were identified by RDIRem.

**Outcome**

Of the 94 patients, 49 met the minimum diagnostic criteria (≥5 events/h) using RDISuz, of whom 20 had RDIRem of <5 (Table 4). Adding respiratory events with >50% reduction in respiratory movements but <4% fall in \text{SpO}_2 in the definition of RDISuz led to 25 additional patients meeting this criterion, of whom 16 had RDIRem <5. Adding respiratory events associated with autonomic arousals in the definition of RDISuz led to 12 additional patients meeting the minimum diagnostic criteria, of whom 9 had RDIRem <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 RDISuz ≥5 events/h were prescribed a treatment trial with continuous positive airway pressure or a mandibular advancement prosthesis. A follow-up

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Predictive value of automated analysis of the Remmers sleep recorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic cut-off</td>
<td>ROC area</td>
</tr>
<tr>
<td>RDISuz &gt;5</td>
<td>0.95</td>
</tr>
<tr>
<td>≥10</td>
<td>0.96</td>
</tr>
<tr>
<td>≥15</td>
<td>0.99</td>
</tr>
<tr>
<td>RDSuz &gt;5</td>
<td>0.85</td>
</tr>
<tr>
<td>≥10</td>
<td>0.90</td>
</tr>
<tr>
<td>≥15</td>
<td>0.91</td>
</tr>
<tr>
<td>RDSuz &gt;5</td>
<td>0.84</td>
</tr>
<tr>
<td>≥10</td>
<td>0.82</td>
</tr>
<tr>
<td>≥15</td>
<td>0.94</td>
</tr>
</tbody>
</table>

RDSuz, diagnostic cut-off value of the respiratory disturbance index (RDI) calculated by the Suzanne recorder; RDIRem, diagnostic cut-off value of the Remmers sleep recorder; LR+ likelihood ratio for a positive test result; LR-, likelihood ratio for a negative test result.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Patient distribution according to minimum required diagnostic criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td>n</td>
</tr>
<tr>
<td>RDISuz &lt;5</td>
<td>8</td>
</tr>
<tr>
<td>RDISuz ≥5</td>
<td>49</td>
</tr>
<tr>
<td>RDSuz &gt;5</td>
<td>25</td>
</tr>
<tr>
<td>RDSuz ≥5</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
</tr>
</tbody>
</table>

Values are mean (SE).

RDIRem, respiratory disturbance index (RDI) of the Remmers Sleep Recorder; RDISuz, restricted RDI of the Suzanne recorder; RDISuz, RDI of the Suzanne recorder; RDIRem, extended RDI of the Suzanne recorder; ESS, Epworth sleepiness score at baseline; 5/Rx, number of responders at follow-up over number of treated patients; ΔESS, change in Epworth sleepiness score with treatment relative to baseline.
questionable changes in the SpO2 signal unrelated to respiratory
RSR underestimated the number of respiratory events by 38–
or thoracoabdominal movements was
magnitude was required only when the decrease in nasal flow
definition of hypopnoea.12 In their trial, however, the inclusion
sensitivity of the oxygen sensors or analysis algorithms of the
events detected by automated analysis with RSR, but filtered
RDISuz, RDISuz 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 >4% fall in SpO2. 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 RDIRem and RDISuz
In a previous study by Vasquez et al13 no significant bias was
found between RDIRem 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 >4% fall in SpO2, RDIRem exceeding
RDISuz by 13%. By contrast, when respiratory events associ-
ated with a <4% fall in SpO2 or with autonomic arousals were
included in the definition of RDISuz, a negative instrument bias
evident, RDIRem underestimating RDISuz by 38% and
RDISuz, by 48%.
RDIRem values in excess of RDISuz = may have been caused by
inadvertent changes in the SpO2 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, RDIRem values smaller than RDISuz or
RDISuz are probably explained by different definitions of
respiratory events. This should also contribute to an explana-
tion 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 SpO2 of >4% whereas, in our work, a fall in SpO2 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 SpO2,15 events associated with a fall in
SpO2 of <4% would contribute to RDISuz and RDI Suz, but not to
RDIRem, thereby increasing the bias in our study compared with
theirs.

In our investigation the inclusion of autonomic arousals in
the definition of RDISuz heightened instrument bias signifi-
cantly 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,17 RDISuz significantly
exceeded RDISuz, 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 ≥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%).13 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 RDISuz 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 SpO2 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 SpO2 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 SpO2 or with
autonomic arousals in the definition of RDISuz. As shown by
Guilleminault et al,12 sleep fragmentation can occur in a
number of patients with OSAH in association with periodic increments
of respiratory efforts but without significant oxygen desatura-
tion. These patients can be very symptomatic and benefit from
treatment. Furthermore, as demonstrated experimentally by
Martin et al,16 even subcortical autonomic arousals can increase
sleepiness. In the present study the inclusion of respiratory
events with <4% fall in SpO2 or with autonomic arousals in the
definition of RDISuz 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 ≥4% fall in
SpO2 was required to define a respiratory event with the
Suzanne recorder. However, when respiratory events associated
with a <4% fall in SpO2 or with autonomic arousals were
included in the definition of RDISuz, the sensitivity of the RSR
at the same cut-off point was markedly reduced to between
52% and 63%. These sensitivities are within the same range
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
RDIRem be ≥20 or ≥10 events/h, depending on whether
autonomic arousals are included or not in the definition of
RDISuz. 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 RDIRem 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., 66% of patients with OSAH were correctly identified by visual inspection of SpO2 traces alone with no false positives. Using a ≥4% fall in SpO2, only 41% of patients with OSAH were correctly identified with 97% specificity. In a study by Séries et al., 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 RDIRem near the diagnostic threshold are such as to deter substitution of the more comprehensive respiratory monitor.

ACKNOWLEDGEMENTS

The REMS 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.

References