Accuracy of Oximetry for Detection of Respiratory Disturbances in Sleep Apnea Syndrome*

Patrick Lévy, MD, PhD; Jean Louis Pépin, MD; C. Deschaux-Blanc, MSc; B. Paramelle, MD; and Christian Brambilla, MD

Study objective: The cost and inconvenience of polysomnography make simplified techniques of screening desirable in the strategy of diagnosis of sleep apnea syndrome (SAS). We have evaluated, in a prospective study of 301 consecutive patients referred for suspected sleep disorders, an index (Δ index) that detects apneic events by quantifying arterial oxygen saturation (SaO₂) variability.

Setting: Regional sleep laboratory taking referrals from general practitioners and specialists.

Design: Classic polysomnography was the gold standard, with 15 apneas plus hypopneas per hour (RDI) being used as a threshold for definition of obstructive sleep apnea (OSA). Oximetry was recorded over the same night. Signal variability was quantified as a function of time, using digital processing of oximetric data. Sensitivity, specificity, and positive and negative predictive values of oximetry testing were calculated.

A receiver operating characteristic (ROC) curve was constructed representing the comparative courses of sensitivity and 1-specificity at different thresholds of Δ index.

Results: Three hundred one patients were included (age, 56±12 years). Their RDI was 30±24. For a Δ threshold at 0.6, the sensitivity of oximetry for the diagnosis of OSA was 98% and the specificity was 46%. The positive and negative predictive values for diagnosing SAS were 77% and 94%, respectively. The three false-negative cases had a relatively high awake SaO₂ (97 vs 93.9±2.8%), a moderate RDI (23.3±1.6), and were less obese than the other patients (body mass index: 25±5 vs 33±5). The 58 false-positive cases had an RDI of 8±4, an awake SaO₂ of 93.1±3.6 vs 94.1±2.6 for the rest of the population (p=0.01). Finally, the false-positive cases had more airways obstruction (FEV₁/VC=72±13 vs 77±15%; p=0.026). Using a Δ value of 0.8 leads to a sensitivity of 90% with 19 false-negative cases but with a higher specificity of 75%.

Conclusions: A nocturnal oximetry test with a Δ index below 0.6 is helpful in ruling out the diagnosis of SAS in patients being screened for this condition, as this yielded only three negative test results in 301 screening procedures.


BMI=body mass index; FN=false-negative; FP=false-positive; RDI=respiratory disturbance index; apnea plus hypopnea index (events per hour of sleep); ROC curve=receiver operating characteristic curve; SaO₂=arterial oxygen saturation; SAS=sleep apnea syndrome; TN=true-negative; TP=true-positive

Key words: detection of sleep apnea; oximetry; sleep apnea

Sleep apnea syndrome (SAS) can be defined by the occurrence of more than 10 or 15 apneas plus hypopneas per hour of sleep.¹ Cardiovascular²-⁴ and neuropsychological⁵ morbidity has been demonstrated in untreated sleep apnea. This morbidity, plus the occurrence of an increased risk of auto crash⁶ and a relatively increased mortality⁷,⁸ make treatment imperative. The prevalence of sleep apnea is estimated at between 4 to 8% of the population.⁹,¹⁰ Polysomnography is needed to confirm the diagnosis.¹¹ The high prevalence of the disease and the cost and inconvenience of polysomnography make simplified techniques desirable for diagnosing SAS. Numerous techniques have been developed to detect SAS using questionnaires,¹² oximetry,¹³-¹⁵ and tracheal sounds,¹⁶ either alone or in combination.¹⁷-¹⁹

Using oximetry alone, the presence of repetitive fluctuations in the arterial oxygen saturation (SaO₂) signal, even without rigid criteria for the amplitude of the SaO₂ decline, should improve its accuracy as a screening test for SAS¹⁵,²⁰ compared with using predetermined levels of SaO₂ decline. Moreover, assessing this SaO₂ variability by computer analysis provides objective results independent of individual experience. We have previously reported an index (Δ index) that quantified SaO₂ variability and therefore allowed us to detect apneic events using pulse oximetry. This first study, however, involved only 26 patients.¹₁ Thus, we have evaluated this index in a prospective study

*From the Department of Respiratory Medicine (Drs. Lévy, Pépin, Paramelle, and Brambilla); and Sleep Laboratory, AGIR, Hôpital Care Regional Association for Respiratory Insufficiency (Drs. Pépin and Deschaux-Blanc); and the PRETA Laboratory, Department of Physiology, Faculty of Medicine, Joseph Fourier University, (Dr. Lévy), Grenoble, France.

ADRAR, COMARES provided grants for this study.

Manuscript received January 12, 1995; revision accepted September 19.

Reprint requests: Dr Levy, Département de Pneumologie, Unité Sommeil et Respiration, CHU de Grenoble, BP 217 X, 38043 Grenoble, France.
involving 301 consecutive patients referred for suspected sleep disorders.

**Materials and Methods**

**Patient Referral**

Patients were referred to our regional respiratory laboratory with suspected sleep-related breathing disorders or simple snoring by general practitioners and private and hospital specialists. Patients receiving oxygen therapy or home ventilation were excluded.

**Sleep Studies**

Standard polysomnography was performed in the sleep laboratory by measurement of EEG, electro-oculogram, chin electromyogram, and ECG. Nasal and oral thermistors and uncalibrated inductance plethysmography were used to measure air flow and respiration, respectively. Sleep was staged manually according to international criteria, and apneas and/or hypopneas were detected manually. Apnea was defined as a cessation of respiration of more than 10 s duration. Hypopnea was defined as a 50% reduction in the flow signal compared with the baseline value of immediately preceding breaths, persisting for more than 10 s. Oxygen desaturation was not a requisite criterion if a microarousal was present with reduction in airflow. Apneas were classified as central when there was no air flow and no respiratory movement. Obstructive apneas were scored when airflow was absent but respiratory efforts continued. SAS was defined as a combined apnea plus hypopnea index (RDI, events per hour of sleep) of 15 or greater.1

**Nocturnal Oximetry**

Nocturnal oximetry was performed simultaneously during sleep studies using a transcutaneous finger probe (Biox 3700 or 3740; Ohmeda; Boulder, Colo). The signal recorded was the 12-s sampling frequency for digital storing in the oximeter (Biox) (minimal value of six measurements over 12 s) with a 3-s response. Warley et al23 studied the ability of this algorithm to reproduce cycle lengths of SaO2 (similar to that observed in sleep apnea). They concluded that 12-s sampling frequency allowed reasonable resolution of SaO2 variability. There was some underestimation of the peak SaO2 in recovery postapnea but the signal shape and variability were preserved. We have quantified that signal variability was as a function of time, using a digital processing of oximetric data (original software: OXYMAT): 

\[
\Delta = \frac{1}{n} \sum \frac{d(SaO2)}{d(t)} \quad (12\text{-s intervals})
\]

The \( \Delta \) index corresponds to the sum of the absolute variations between two successive points, divided by the number of intervals; \( d(SaO2) \) = number of intervals; and \( d(t) \) = time. The analysis of an overnight recording takes less than 30 s and the software is available from us on a noncommercial basis.23

This index measures the variations between successive oxygen saturation data at constant time intervals. Therefore, if saturation is nearly constant during the night, the \( \Delta \) index will be very low, the SaO2 variation being minimal. The nocturnal SaO2 profile in SAS is due to repetitive sequences of apneas and resumptions of respi-ration. This leads to SaO2 oscillations and \( \delta SaO2 \) is then high for nearly all sampling intervals. This logically leads to high \( \Delta \) values.

The other usually derived indexes of desaturation, such as mean nocturnal SaO2, minimal nocturnal SaO2, and cumulative time spent below 90% SaO2 are also calculated.

**Data and Statistical Analysis**

The characteristics of the patients were reported using mean values and SDs (mean±SD). The accuracy and specificity of oximetry (\( \Delta \) Index) for screening in SAS: Nocturnal oximetry was used as the test and polysomnography as the gold standard for the correct classification of apneic vs nonapneic patients (RDI < 15/h). The \( \Delta \) value was rounded to the nearest decimal point. The number of true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) oximetry results were then determined. For the 301 patients, sensitivity (TP/[TP+FN]), specificity (TN/[TN+FP]) and positive (TP/[TP+FP]) and negative predictive values (TN/[TN+FN]) were calculated. A receiver operating characteristic (ROC) curve[28] was constructed, representing the comparative course of sensitivity and 1-specificity (1 minus the specificity expressed as a decimal) at different thresholds. On such a curve, if the threshold is displaced toward the right on the horizontal limb, a minimal gain in sensitivity results in a major loss of specificity. Conversely, a displacement toward the left on the vertical limb results in a minimal increase in specificity while sensitivity is dramatically decreased. For screening purposes, one chooses a high sensitivity in order not to falsely exclude from further investigation patients having the disease in

---

**Table 1—Patient Characteristics: Age, BMI, RDI, FEV1/VC**

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>56±12</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>32±8</td>
</tr>
<tr>
<td>RDI, nb/h</td>
<td>30±24</td>
</tr>
<tr>
<td>FEV1/VC, %</td>
<td>76±15</td>
</tr>
<tr>
<td>Wake SaO2, %</td>
<td>94±3</td>
</tr>
</tbody>
</table>

*Data are expressed as mean±SD; nb=number.

---

**Table 2—Sensitivity and Specificity at Various Thresholds of \( \Delta \) Index**

<table>
<thead>
<tr>
<th>( \Delta ) Index Threshold</th>
<th>Sensitivity, % (CI)*</th>
<th>Specificity, % (CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>100 (100-100)</td>
<td>10 (4-16)</td>
</tr>
<tr>
<td>0.6</td>
<td>98 (96-100)</td>
<td>46 (37-55)</td>
</tr>
<tr>
<td>0.8</td>
<td>90 (86-94)</td>
<td>75 (67-83)</td>
</tr>
<tr>
<td>1.2</td>
<td>69 (62-76)</td>
<td>93 (88-98)</td>
</tr>
<tr>
<td>1.5</td>
<td>56 (49-63)</td>
<td>96 (92-100)</td>
</tr>
<tr>
<td>1.8</td>
<td>46 (39-53)</td>
<td>100 (100-100)</td>
</tr>
<tr>
<td>2.1</td>
<td>37 (30-44)</td>
<td>100 (100-100)</td>
</tr>
<tr>
<td>3.0</td>
<td>19 (13-25)</td>
<td>100 (100-100)</td>
</tr>
</tbody>
</table>

*CI=95% confidence intervals.
question (FN). Conversely, for treatment decisions, one would choose a higher specificity in order not to inflict further investigation or treatment on patients without the disease (FP).

The Mann-Whitney test was used for comparison between subgroups.

A p value less than 0.05 was considered significant.

RESULTS

Patients

Three hundred one patients were prospectively studied. Their anthropometric data and respiratory disturbance indexes (RDI) by polysomnography are shown in Table 1; 193 patients (64%) were classified as apneic (RDI ≥15) and could be divided in 24% and 40% having, respectively, between 15 and 30, and more than 30 apneas plus hypopneas per hour of sleep. At an RDI threshold of 10/h, 221 (73.4%) of these patients would have a diagnosis of obstructive sleep apnea.

Accuracy of Oximetry

A significant correlation was found between Δ index and RDI (r=0.72, p=0.0001) (Fig 1). Other desaturation indexes frequently cited were less correlated with RDI than Δ index (mean nocturnal SaO2: r=−0.39; minimal nocturnal SaO2: r=−0.47; cumulative time spent below 90% SaO2: r=0.42, all p=0.0001).

Sensitivity and specificity using various thresholds for Δ index are indicated in Table 2. The ROC analysis relating sensitivity and 1-specificity is shown in Figure 2.

Using a Δ value of 0.6 gives a sensitivity of 95% with only three FN results but with a low specificity of 46%. A nocturnal oximetry test with a Δ index below 0.6 is helpful in ruling out the diagnosis of SAS in patients with a low likelihood of having this syndrome, because a negative test result (Δ index <0.6) yielded only three positive polysomnograms giving a negative predictive value of 94%. The positive predictive value at this cutoff point was 77%.

The three FN cases had a relatively high awake SaO2 (97 vs 93.9±2.8%), a moderate RDI (23.3±1.6), and were less obese than the other patients (body mass index [BMI], 25±3 vs 33±8). The mean Δ value of these three patients was 0.4±0.1. None eventually required continuous positive airway pressure treatment. The 38 FP cases had an RDI of 8±4, an awake SaO2 of 93±3.6 vs 94.1±2.6 for the rest of the population (p=0.01). Finally, the FP cases had more airway obstruction (FEV1/VC=72±13 vs 77±15%; p=0.026).

Establishing a diagnosis of obstructive sleep apnea at an RDI of 10 events per hour would give 11 FN results at Δ 0.6 with a sensitivity of 95% and specificity of 52.5% and a negative predictive value of 79%.

Using a threshold Δ score of 0.8, the sensitivity was 90% and the specificity was 75% for the diagnosis of sleep apnea. This leads to positive and negative predictive values of 87% and 81% respectively. The 19 FN cases identified at this cutoff level had a mean Δ value of 0.62±0.1, a high mean awake SaO2 (95±2%), a moderate RDI (22±9%), and were less obese than the other patients (BMI=29±6 vs 33±5; p<0.025).

There were 27 FP cases who had a mean Δ value at 1.03±0.3, an RDI of 8±4, a lower awake SaO2 of 91.5±4 vs 94±3 for the other subjects (p<0.0005), and a high BMI at 34±8. Finally, the FP cases had more airflow obstruction (FEV1/VC=68±15 vs 77±14%; p=0.005).

DISCUSSION

Nocturnal desaturation constitutes one of the consequences of sleep apnea and hypopnea. Our results indicate that an appropriate analysis of SaO2 variability can be valuable as an initial screening test in helping to decide whether to proceed to further investigation of patients suspected of sleep-disordered breathing. The rationale for identifying respiratory disturbances by SaO2 variability derives from the association of repetitive fluctuations in the SaO2 signal with the succession of respiratory events, whatever the amplitude of the SaO2 decline.

Farney et al25 first used visual analysis of oximetry as a screening and diagnostic test for SAS, but this involved detailed analysis of 3-min epochs of tracing played at slow speed, which hardly speeds up the diagnostic process compared with polysomnography. Visual inspection of the oximetry profile can be useful as shown by Cooper et al.13 but this relies on an experienced observer and in addition a number of studies have to be described as unreliable or even uninterpretable.
A number of other studies have evaluated the use of oximetry in simplifying the diagnostic strategy in SAS. However, there have been conflicting results. Some studies have shown a sensitivity as low as 50% and specificity of up to 95%, while conversely, other studies have found a high sensitivity for the diagnosis of SAS.

There are several reasons for such discrepancies. First, the populations studied can vary from one study to another. The proportion of mild SAS and/or patients with little desaturation associated with repeated hypopneas may be different in these studies. Age, pulmonary function, and degree of obesity are different factors influencing nocturnal desaturation which may lead to differences in terms of specificity and sensitivity. The main reason for these differences is, however, the method of analysis of the oximetry data. Studies reporting low sensitivity have used a method of quantification of nocturnal desaturation using software counting the number of times per hour that brief decreases of a given magnitude (generally a 4% decrease) in oxygen saturation occur during sleep.

Serriès et al., like us, used overall recognition of the pattern of abnormality in the oximetry profile, i.e., repetitive fluctuations of SaO₂ whatever the magnitude. These authors performed manual scoring, which is by definition dependent on subjective assessment. The diagnostic accuracy of such a method may depend on the level of experience of the reader of the trace. An objective assessment based on a software analysis provides a simpler and more accessible way of detecting SaO₂ variability. Serriès et al. found a sensitivity at 98% with a specificity of 48% for SAS at an RDI cutoff of 10. At an RDI of 20/h, they had no FN results but had a specificity of 38.8%. In the present study, using a cutoff point of 0.6 and RDI of 15/h, we have also found a high sensitivity (98%) with a similar degree of specificity (46%). There is only a small loss of sensitivity to 95% using an RDI of 10.

The characteristics of signal sampling may potentially modify the accuracy of such an analysis. Serriès et al. have used a 2-s frequency while we have used the minimal value found over a 12-s period which constitutes the standard way of storing digital data in several oximeters. If the sampling interval were shortened, the SaO₂ variation would consequently be lower between one point and another. This would lead to different Δ thresholds and related sensitivity and specificity and ROC curve would need to be calculated for each different sampling interval. However, a higher sampling frequency may not automatically lead to a more accurate analysis of the recording. In our study, a Δ value of 0.6 (Table 2 and Fig 2) leads to a sensitivity and specificity nearly identical to those of Serriès et al. The accuracy of Δ determination combined with higher sampling frequency remains to be studied, particularly to see if specificity can be increased without losing sensitivity. Our oximetry profile gives a value of the cumulative time spent desaturated to less than 90% SaO₂, thereby indicating indirectly the cumulative nadirs of desaturation that may be important in clinical decision making.

The question of sensitivity and specificity of a particular analysis calls into question the prior probability of the disorder for which a procedure is being performed and the reason for the investigation. If the test is to be used as a screening procedure, then one needs to emphasize the sensitivity in order to reduce the number of FNs to a minimum. In the situation of sleep-disordered breathing, an FP test will subsequently be “sorted out” by further investigation such as polysomnography. Thus, choosing a Δ cutoff point of 0.6 yielded only three FN results in 301 unselected patients. Thus, the next phase of use of our analysis should attempt to calculate the ruling-in and ruling-out gain of the information.

The characteristics of the FN cases at both thresholds of 0.6 and 0.8 are of interest. The combination of high awake SaO₂ together with low BMI and therefore probable normal lung volumes could together possibly lead to a reduced likelihood of desaturation during apneas because of being on a flatter point of the oxygen desaturation curve. A poor quality of sleep could possibly result in a low Δ index despite a high RDI (e.g., 50 apneas during a night of only 120 min of sleep leads to a Δ index of 0.7 but an RDI of 25/h of sleep). This was not the case for our three FN cases at Δ threshold of 0.6. None of these three patients have required nasal continuous positive airway pressure.

The FP cases had, conversely, a lower awake SaO₂ and were obese although their BMI was not statistically different when compared with the other patients. They also had a significantly reduced FEV₁/VC ratio reflecting a higher incidence of airflow obstruction. All these factors could have favored nonapneic desaturations thereby resulting in FP results.

The guidelines for the investigation of patients with possible SAS differ from one country to another. The British guidelines state that oximetry alone, or oximetry plus video recording, may be sufficient to diagnose the condition in some patients. Conversely, American recommendations state that polysomnography should be performed in all the patients. Undoubtedly, financial factors may favor the choice of a simplified strategy of diagnosis. However, the recent description by Guillemaintal et al. of an upper airway resistance syndrome, corresponding to sleep fragmentation in snorers, related to periodic increases in respiratory effort and repeated arousals, could require the performance of complete polysomnography in all snorers. Nevertheless, although this syndrome can be responsible for daytime sleepiness, any associated cardiovas-
cular morbidity remains to be established. As daytime hypersonsomnolence is such an important symptom in SAS, such patients would be deemed to have a high prior probability of SAS and thus would be further investigated if Δ was low.

We conclude that a computer-derived index of SaO₂ variability in nocturnal oximetry allows an easy and widely applicable screening test for the absence of SAS. It should be noted, however, that in the presence of a hypersonsomnolent patient, a negative oximetry result should always be followed by a complete polysomnography in order to exclude upper airway resistance syndrome or periodic leg movements.

ACKNOWLEDGMENT: We thank M. Colonna, PhD, for technical assistance.

REFERENCES