Cost-effectiveness Analysis of Nocturnal Oximetry as a Method of Screening for Sleep Apnea-Hypopnea Syndrome*

Laurence J. Epstein, MD, FCCP, and Gina R. Dolac, MD

Study objective: Determine the utility of nocturnal oximetry as a screening tool for sleep apnea-hypopnea syndrome (SAHS) compared with polysomnography (PSG).

Design: Cost-effectiveness analysis based on retrospective review of overnight sleep studies.

Setting: United States Air Force tertiary teaching hospital.

Patients: One hundred consecutive patients evaluated for SAHS by overnight sleep study.

Intervention: Participants underwent PSG and oximetry on the same night. Patients with obstructive sleep apnea had a continuous positive airway pressure trial.

Measurements: Oximetry was abnormal when ≥10 events per hour occurred. Two criteria were evaluated. A “deep” pattern of >4% change in oxyhemoglobin saturation to ≤90%, and a “fluctuating” pattern of repetitive short-duration fluctuations in saturation. The diagnostic accuracy of both methods was compared with PSG. Cost-effectiveness of screening oximetry was compared with PSG alone and use of split-night studies.

Results: The fluctuating pattern had a greater sensitivity and negative predictive value, while the deep pattern had a greater specificity and positive predictive value. Oximetry screening using the fluctuating pattern was not as sensitive as PSG for detecting patients with mild disease; 17 of 28 patients (61%) with normal oximetry results had treatable conditions detected by PSG. Cost analysis showed that screening oximetry would save $4,290/100 patients but with considerable loss of diagnostic accuracy.

Conclusion: Screening oximetry is not cost-effective because of poor diagnostic accuracy despite increased sensitivity using the fluctuating pattern. Greater savings, without loss of diagnostic accuracy, may be achieved through increased utilization of split-night PSGs.

(CHEST 1998; 113:97-103)

Key words: cost-effectiveness; obstructive sleep apnea; polysomnography; pulse oximetry

Abbreviations: AHI=apnea-hypopnea index; BMI=body mass index; CPAP=continuous positive airway pressure; OSA=obstructive sleep apnea; PSG=polysomnogram; SAHS=sleep apnea-hypopnea syndrome; SaO2=oxyhemoglobin saturation; SDB=sleep disordered breathing; UARS=upper airway resistance syndrome

Obstructive sleep apnea (OSA) is a common disorder with significant cardiovascular morbidity and effects on daytime performance. Recent studies have estimated the prevalence of the sleep apnea-hypopnea syndrome (SAHS), OSA plus clin-
they found that pulse oximetry was specific but not sensitive for detecting OSA. The addition of a clinical score improved sensitivity but did not eliminate false-negative results. Series et al reported improvement in detection by using different criteria for positive screening oximetry: the presence of repetitive, short-duration fluctuations in SaO2 without any absolute value decrease in saturation. They concluded that SAHS could be ruled out by normal pulse oximetry results using their criteria, and that patients with abnormal oximetry screens should go on to PSG for definitive diagnosis. The improvement in detection led to an elevation in the false-positive rate, which resulted in a large number of additional PSGs.

To be an effective screening tool for SAHS, pulse oximetry must be able to screen out patients without disease and detect patients with all levels of disease severity in a manner that is less expensive than current diagnostic procedures. We investigated the utility of pulse oximetry as a screening tool for SAHS by comparing both diagnostic criteria directly with PSG and evaluating the cost-effectiveness of a diagnostic algorithm employing screening oximetry.

**Materials and Methods**

**Patients**

We retrospectively reviewed the sleep studies of 100 consecutive patients referred for evaluation for possible sleep-disordered breathing (SDB) who underwent PSG in our sleep center. We studied 93 men and seven women, ages 19 to 72 years (mean, 39.3 years), with a mean body mass index (BMI) of 28.2 ± 4.4 kg/m2 (mean ± SD). Our referral base consists primarily of active duty military personnel, but also includes family members and retired military members. Patients were referred to the sleep center by their primary care physicians. After examination by a sleep medicine physician, a sleep study was ordered because of suspicion of the presence of SAHS.

**Protocol**

Sleep studies were performed in a monitored facility. All measurements were recorded on a 16-channel polygraph (Grass Instruments; Quincy, Mass) running at 10 mm/s and included the determination of sleep stages (with EEG, electro-oculogram, submental electromyogram), nasal/oral airflow with thermistor (Grass Instruments), EOG, thoracoabdominal movements by inductive plethysmography (Ambulatory Monitoring Inc; Ardsley, NY), limb electromyogram, and SaO2 with pulse oximeter (Ohmeda; Madison, Wis). The oximeter signal was also sent to a strip chart recorder that provided a separate real-time tracing of SaO2.

Sleep stages were defined in 30-s epochs according to standard criteria. An apneic event was defined as a cessation of nasal/oral airflow for at least 10 s, while a hypopnea was defined as a >50% decrement in nasal/oral airflow for at least 10 s associated with either an arousal or oxyhemoglobin desaturation ≥4%. An arousal was defined by an EEG frequency shift to alpha range for at least 3 s. An apnea-hypopnea index (AHI) was calculated using the following formula: number of apneas+number of hypopneas/total sleep time (hours).

Split-night PSGs were performed on patients with significant SDB. After the first episode of rapid eye movement sleep or 3 h in bed, whichever came first, the sleep technicians evaluated the records for OSA. Patients with an AHI ≥20 were started on a regimen of continuous positive airway pressure (CPAP) that was titrated to a pressure level adequate to eliminate OSA and snoring.

**Data Analysis**

PSG recordings were scored manually by a single registered PSG technician. The diagnosis of SAHS was made when the AHI on PSG was >10/h. The oximetry strip chart tracings were examined by a single physician interpreter blinded to the results of the PSG. Oximetry was scored by two methods: the standard “deep” pattern of >4% desaturation to below 90% (Fig 1), and the “fluctuating” pattern of either large desaturations or low-amplitude periodic fluctuations using neither a minimum decrease in SaO2 levels nor a threshold minimum saturation level (Fig 2). Oximetry results were classified as abnormal by the presence of repetitive desaturations (>10/h) followed by a rapid return to baseline SaO2 levels. The interpretations were verified by a second scoring physician, blinded to the PSG results and the results of the first scorer, on 10 randomly selected tracings, with a 100% concurrence rate. Scoring of both oximetry and PSG was performed only on the pre-CPAP diagnostic portions of the split-night studies. Other types of SDB diagnosed by PSG included mild OSA (5 ≤ AHI ≤ 10), and upper airway resistance syndrome (UARS), defined by the presence of a discrepancy between the arousal index and the AHI with evidence of a respiratory etiology for the arousals, such as crescendo snoring or paradoxical abdominal/thoracic movements.

**Cost Analysis**

Two diagnostic algorithms were compared using the data from the initial PSGs and oximetry (Fig 3). In the first algorithm, described by Series et al., oximetry was used as the initial screening test, and all patients with abnormal oximetry results went on to PSG for definitive diagnosis. For the second algorithm, all patients had an initial PSG. In both algorithms, all patients diagnosed as having SAHS underwent a CPAP titration trial for treatment. The CPAP titration trial is equivalent to PSG in performance and cost. Assumptions made for the cost analysis were as follows: the results of the most sensitive oximetry criteria (deep vs fluctuating) would be used in the cost analysis, and CPAP titration would require only a single additional night. The cost analysis was performed using price estimates obtained by telephone survey of local sleep centers. The mean charge reported for oximetry was $294 (range, $125 to $464); the mean charge for interpreted PSG and CPAP titration trial was $1,123, (range, $1,000 to $1,217). The costs were calculated at both the mean and the lowest prices.

Previous studies have shown that certain patients can undergo both diagnosis and CPAP titration in a single night of study (“split-night” study). In a split-night study, CPAP titration is initiated if a predetermined OSA severity threshold is crossed during the initial portion of the study. Sanders et al have shown a high correlation between indexes of OSA obtained in the first 2 h of a study compared with those calculated over the entire night. We performed an additional analysis to evaluate the effect on cost of utilizing split-night studies. The cost of the initial PSG algorithm was recalculated based on the number of patients with an AHI ≥20 who had split-night studies. The cost of a split-night study is the same as a diagnostic PSG or a CPAP titration. For each split-night study, only one PSG charge was added to the total instead of two.
Figure 1. Deep pattern. Example of oximetry tracing scored as abnormal by deep pattern; repetitive desaturations >4% to below 90%.

Figure 2. Fluctuating pattern. Example of oximetry tracing scored as abnormal by the fluctuating pattern; low-amplitude periodic fluctuations without threshold levels for change in SaO₂ levels or minimum saturation level. In particular, note variations on the right side of the tracing.

RESULTS

A diagnosis of SAHS was made by PSG in 53 of the 100 patients studied, with a mean AHI of 32.4±22.1. The ability of oximetry to detect SAHS varied according to the diagnostic criteria used, and is shown in Table 1. A comparison of the diagnostic accuracy of the two oximetry criteria for detecting an AHI
>10/h, and the effect of weight on diagnostic accuracy, is summarized in Table 2. Use of the fluctuating pattern for oximetry interpretation resulted in a higher sensitivity and negative predictive value, but decreased specificity and positive predictive value. Both methods were more sensitive in patients with an increased BMI. The diagnostic algorithms were applied using the fluctuating pattern because of the greater sensitivity of this method. The resulting values were utilized in the cost analysis.

The results of the cost analysis are summarized in Table 3. Use of screening oximetry prior to PSG would have saved $4,290 per 100 patients evaluated compared to initial PSG testing if all patients with SAHS had required a second night study for CPAP titration. The cost savings calculated with the lowest prices increased to $17,500 per 100 patients. However, despite the cost savings, 17 of 28 patients with normal oximetry results (61%) had a treatable disorder missed by screening oximetry alone. Two patients had an AHI >10 but normal oximetry results (false negatives) and 15 others had an elevated arousal index identified on PSG, indicating a disorder of sleep fragmentation (mean arousal index of 18.5±5.8/h). Diagnoses in this group included eight patients with mild OSA (5/h>AHI≤10/h) and additional respiratory related arousals due to increased upper airway resistance, four with UARS, one patient with periodic limb movements of sleep, and two with periodic limb movements of sleep plus mild OSA.

Split-night studies were successfully performed in 21 of 33 patients with an AHI≥20. The decrease in the number of PSGs performed due to the use of split-night studies resulted in the greatest cost savings (Table 3), with a reduction of $19,293/100 patients compared to the initial oximetry algorithm.

**Discussion**

This study compared two different diagnostic criteria for interpreting nocturnal pulse oximetry with standard PSG evaluation for the detection of SAHS. In addition, we evaluated the cost-effectiveness of pulse oximetry as a screening tool for the detection of SAHS. We found that the use of screening pulse oximetry resulted in only a small cost savings with a significant reduction in diagnostic accuracy.

Our study confirms the findings of Series et al4 that using a less rigid criterion for interpreting oximetry improved its ability to detect SAHS, increasing both the sensitivity and negative predictive value by almost 20% to >90% each. However, this improvement was achieved with a significant reduction in specificity, with 29% of all abnormal oximetry test results being false-positives. Despite the improvement in SAHS detection, 61% of patients with normal results of oximetry, 17% of all the patients, had a treatable sleep disorder that would have remained undetected.

Applying a diagnostic algorithm utilizing oximetry as a screening tool to guide decisions on which patients should undergo PSG did result in a small cost savings compared to using PSG alone. However, the consequence of using screening oximetry was that a significant number of patients with sleep disorders that cause excessive sleepiness would remain undiagnosed and untreated. The high cost of sleepiness-related motor-vehicle, work-related, and home-based accidents10 could negate these small cost savings.

**Table 1—Contingency Tables Comparing Oximetry and PSG Using Two Oximetry Diagnostic Criteria**

<table>
<thead>
<tr>
<th>Oximetry</th>
<th>Fluctuating Pattern</th>
<th>Deep Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AHI &gt;10</td>
<td>AHI ≤10</td>
</tr>
<tr>
<td>Abnormal</td>
<td>51</td>
<td>21</td>
</tr>
<tr>
<td>Normal</td>
<td>2</td>
<td>26</td>
</tr>
</tbody>
</table>

**Figure 3.** Diagnostic decision trees. A: algorithm utilizing oximetry as initial screening test. B: Algorithm using PSG as initial test.
Screening oximetry was most successful in detecting SAHS in patients with a high likelihood of having OSA (those with an elevated BMI) or those with more severe disease (frequent desaturations >4%). However, this is the subset of SAHS patients most likely to undergo successful split-night studies, which we found to give the greatest cost savings. In patients with milder disease, normal results of nocturnal oximetry did not mean the absence of SDB. As a result, additional testing would be necessary, negating the usefulness of the screening tool.

There are several reasons for the discrepancy between our results and those of Series et al.4 One is difference in methods. The Series et al study did not compare oximetry and PSG on the same night. There could have been discrepancies between the oximetry and the PSG due to first-night in laboratory effects or positional effects11 if studied on different nights. All of our studies were performed in the sleep laboratory, while the Series et al study utilized home oximetry. This should have improved the quality, and thus accuracy, of the oximetry tracings in our study, since technicians could correct problems giving poor signals. Our definition of hypopnea was slightly different in that a reduction in airflow was scored as a hypopnea if it was accompanied by a desaturation or an arousal, in contrast to only being scored if accompanied by a desaturation in the Series et al study. This may have led to a higher number of abnormal studies by PSG in our study. The number of studies affected by this difference is likely small; however, this highlights the problem of detecting patients with mild disease using oximetry. Guillemi-nault et al12 have shown that clinically significant sleep fragmentation, amenable to treatment, can occur in the absence of desaturation. Another possible source of bias was the use of split-night PSGs that reduced the amount of time available to detect desaturations by oximetry to ≤3 h. However, this did not influence the outcome of the study since oximetry results were abnormal in all of the patients who had split-night studies.

A major difference between the two studies is in the populations investigated. Although both populations were referred for suspected SDB, the characteristics of the two groups were different in one major factor, body weight. The BMI for the Series et al4 group was 31.7±0.8 vs 25.2±0.4 kg/m² (mean±SE) in ours. The population we studied, drawn from military personnel and their families, tended to be thinner and, as we have shown previously, have milder disease.13 As a result, the average AHI for the patients with SAHS was greater in the Series et al study, 38.1±2.5, than in ours, 32.4±2.2 (mean±SE). The effect of body mass on the accuracy of nocturnal oximetry can be seen in Table 2; the sensitivity and negative predictive values are less in the BMI <30 group compared to the >30 group. In both studies, the false-negative oximetry tests occurred in patients with lower BMIs, 23.3±1.0 in ours and 29.8±2.7 kg/m² in the Series et al study.

It is important to distinguish between the use of overnight oximetry as a diagnostic vs a screening tool. This study shows that the fluctuating criterion for scoring oximetry is accurate for detecting an AHI >10, as shown by the high sensitivity and positive predictive value. However, its usefulness as a diagnostic tool is limited by the high rate of false-positives, which requires that all patients with abnormal oximetry results undergo PSG anyway to confirm the diagnosis. Svansburg et al14 similarly found that OSA could be diagnosed by oximetry with a high sensitivity but a low specificity, requiring that patients with abnormal results of studies needed further evaluation with PSG. Douglas et al15 found that oximetry could diagnose up to two thirds of patients suspected of having SAHS, but could not detect disease in the rest, mostly patients with mild

---

**Table 2—Diagnostic Accuracy of Oximetry for Detecting an AHI >10**

<table>
<thead>
<tr>
<th>BMI, kg/m²</th>
<th>Deep</th>
<th>Fluctuating</th>
<th>Deep</th>
<th>Fluctuating</th>
<th>Deep</th>
<th>Fluctuating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>73.6</td>
<td>96.2</td>
<td>65.6</td>
<td>93.8</td>
<td>85.7</td>
<td>100</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>89.4</td>
<td>55.3</td>
<td>84.6</td>
<td>56.4</td>
<td>87.5</td>
<td>37.5</td>
</tr>
<tr>
<td>Positive PV, %</td>
<td>88.6</td>
<td>70.8</td>
<td>77.8</td>
<td>63.8</td>
<td>94.7</td>
<td>80.8</td>
</tr>
<tr>
<td>Negative PV, %</td>
<td>75.0</td>
<td>92.9</td>
<td>75.0</td>
<td>91.7</td>
<td>70</td>
<td>100</td>
</tr>
</tbody>
</table>

*Deep=* >4% desaturation to <90%; fluctuating=repetitive fluctuations in saturation; PV=predictive value.

---

**Table 3—Cost Analysis**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial oximetry</td>
<td>$167,529</td>
</tr>
<tr>
<td>Initial PSG</td>
<td>$171,819</td>
</tr>
<tr>
<td>Initial PSG/split-night</td>
<td>$148,236</td>
</tr>
</tbody>
</table>
Multiple studies have compared PSG and oximetry in an attempt to find a screening test for SAHS. Several studies have been able to obtain high sensitivity for detecting OSA by manipulating the AHI level detected and criteria for an abnormal test result.16-22 However, they all reported low specificity values, did not take into account milder forms of disease, and concluded that all abnormal oximetry results required PSG for confirmation. Three other studies evaluated oximetry vs PSG with respect to severity of disease and found that oximetry was a suitable screen for patients with moderate and severe OSA but was inadequate for patients with milder cases.15,21,24 Only one study looked specifically at patients with milder forms of disease, including UARS. Yamashiro and Kryger25 studied patients with all the varieties of SDB and compared results from same night oximetry and PSG. Oximetry detected all patients with moderate and severe OSA. However, 30% of the patients who were diagnosed normal by oximetry were found to have UARS on PSG. Similar to us, they concluded that oximetry was a poor screening tool since normal results did not rule out disease and all abnormal results required further PSG evaluation. To our knowledge, our study is the only one to apply a cost-effectiveness analysis to the issue of screening for SAHS with oximetry.

We conclude that the sensitivity of oximetry as a screening test for SAHS can be improved with the use of less rigid criteria for interpretation of an abnormal test result. However, the use of nocturnal oximetry as a screening tool does not appear to be justified on the basis of our cost-effectiveness analysis. Only minor cost savings are achieved by screening patients with oximetry before PSG with a significant loss of diagnostic accuracy. When compared with oximetry screening, greater savings, without loss of diagnostic accuracy, may be achieved through increased utilization of split-night studies. Methods of screening for OSA other than oximetry, such as multiple parameter monitors or portable PSG, should be evaluated for cost-effectiveness.

ACKNOWLEDGMENTS: The authors thank William Beninati, MD, and Isabel Diaz for assistance in data collection, and James H. Henderson, MD, and Dorothy Cunningham, MD, for their editorial assistance.

REFERENCES


5 Rechtschaffen A, Kales A. A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. Los Angeles: Brain Information Service, Brain Information Institute, University of California, 1968


10 Leger D. The cost of sleep-related accidents: a report for the national commission on sleep disorders research. Sleep 1994; 17:84-93


