Oximeter Performance*

The Influence of Acquisition Parameters

David G. Davila, MD, FCCP; Kathy C. Richards, PhD; Buddy L. Marshall, RPSGT; Patricia S. O'Sullivan, EdD; Ty G. Gregory, RPSGT; Valerie J. Hernandez, RPSGT; and Shirley I. Rice, RPSGT

Study objective: This study was designed to determine whether different desaturation indexes (DIs) would be obtained in patients with sleep-disordered breathing by systematically altering two acquisition parameters: the recording setting and the display mode.

Design: Prospective clinical study.

Setting: Community sleep-disorders center.

Patients: The study included 75 patients who were suspected of having sleep-disordered breathing.

Interventions: Each patient had simultaneous pulse oxyhemoglobin saturation (Sp O₂) traces at three recording settings (3 s, 6 s, and 12 s) during the diagnostic phase of split-night polysomnography. On-line and memory displays of those data at each recording setting were obtained. DIs for ≥ 3% desaturation events per hour were calculated for each of the six traces.

Results: The mean on-line DIs significantly differed from each other, with slower (longer) recording settings resulting in lower values than faster (shorter) settings. The memory DIs all significantly underestimated the on-line DIs. Pearson correlations ranged from 0.82 to 0.90 between the on-line/memory DI pairs, but Bland-Altman analysis detected disagreement at higher levels of disordered breathing.

Conclusions: These findings confirm that significantly different SpO₂ data are obtained at various acquisition options. The recording setting and display mode parameters should be disclosed in all reports employing oximetry with the fastest recording setting and on-line display mode preferable for case finding of sleep-disordered breathing.

(CHEST 2002; 122:1654–1660)

Key words: oximetry; sleep; sleep apnea syndromes

Abbreviations: BMI = body mass index; CPAP = continuous positive airway pressure; DI = desaturation index; MEM3 = memory at 3 s; MEM6 = memory at 6 s; MEM12 = memory at 12 s; OL3 = on-line at 3 s; OL6 = on-line at 6 s; OL12 = on-line at 12 s; RDI = respiratory disturbance index; SpO₂ = pulse oxyhemoglobin saturation

Pulse oximeters are widely used for measuring pulse oxyhemoglobin saturation (SpO₂) in sleep medicine. The degree of desaturation recorded per disordered-breathing event can be influenced by a number of clinical factors, including underlying cardiopulmonary function. Technical factors including sensor position and selectable options for data acquisition can also affect the timing and depth of desaturation.

Multiple recording settings are often available that average SpO₂ data over different time intervals to allow for flexible use in a variety of clinical settings. There are reports that these recording settings can influence the morphology of the SpO₂ signals and quantification of such data in patients with known sleep-disordered breathing. For example, in limited samples, Farre et al³ noted changes in the shape of the SpO₂ signals, and Kendrick et al⁴ found the countable desaturation rate to vary by a factor of three depending on the recording setting.

Studies of the performance of oximetry in populations suspected of having sleep-disordered breathing have often included data displayed either on-line in real-time on paper in attended study facilities,
or downloaded from the memory to paper from unattended settings of the oximeter. Although the results from both formats have been included together without distinction in past reviews, quantitative differences in data acquired from these two display modes (on-line vs memory) have been suspected. Wiltshire et al.17 and Kendrick et al.18 reported that desaturation data from the memory underestimate that obtained on-line by approximately 5 to 10 events per hour.

This study was designed to determine whether different desaturation indexes (DIs) would be obtained by systematically altering two acquisition parameters—the recording setting and display mode—and to quantify those differences. It was hypothesized that if differing data sets resulted from different acquisition settings, the performance of oximetry with respect to screening and case finding for sleep-disordered breathing might be influenced and that this might explain some of the variability in these measures previously reported in the literature.

**Materials and Methods**

**Design, Setting, and Sample**

This was a descriptive prospective study of consecutive patients referred by primary care providers, medical specialists, and surgeons to a sleep-disorders center at a large community hospital. After a detailed history and physical examination by a board-certified sleep-disorders specialist indicated clinical evidence of sleep-disordered breathing, patients were randomly assigned by clerical staff who had no knowledge of the study to one of five sleep rooms for overnight attended polysomnography. Patients assigned to a room that contained specialized oximetry equipment were invited to participate in this institutional review board-approved study.

**Techniques**

Standard techniques for recording EEG, electrooculogram, electromyogram, ECG, sonogram, airflow, and respiratory effort were employed, and signals were displayed on paper via polygraphs (Grass model 78; Astro-Med; West Warwick, RI). In addition, three oximeters (Ohmeda 35740; Datex-Ohmeda; Madison, WI) providing values within ± 2% variation awake were connected to three identical Ohmeda paper chart recorders with 50% and 100% calibration signals placed on graph paper (Graphic Controls; Buffalo, NY), allowing a 1-mm vertical deflection to correspond to a 1% change in SpO2. Each patient had three sensors (Ohmeda Flex II Sensor; Datex-Ohmeda) placed on the distal aspect of three fingers, which were connected to the three separate oximeters. Each oximeter was set to one of three different recording settings. The term recording setting used here is identical to averaging time 1–4 and response time 6 that have been previously employed in operating manuals and the literature. The three recording settings (3 s, 6 s, and 12 s) equate with “fast,” “normal,” and “slow” averaging or response times.3

Three hard-copy traces of the on-line SpO2 signals were obtained simultaneously with oximetry chart drive paper running at 20 cm/h with full-scale pen response < 0.4 s at all three recording settings during the diagnostic phase of polysomnography. The three oximetry traces per study were identical in length and averaged 3.94 h in duration. At the end of the diagnostic phase, the polysomnography and oximetry paper recordings were stopped, the trend output option was selected from each oximeter, and three additional SpO2 signals were displayed from the memory of each oximeter on paper running at 20 cm/min at the same pen response resulting in traces identical in length to those obtained on-line. Thus, there were six experimental conditions for the SpO2 data: on-line at 3 s (OL3), memory at 3 s (MEM3), on-line at 6 s (OL6), memory at 6 s (MEM6), on-line at 12 s (OL12), and memory at 12 s (MEM12).

The six hard-copy oximetry recordings per patient were examined, and the number of ≥ 3% desaturation events per tracing was manually counted. The number of desaturations divided by the time in bed was calculated to generate a DI for each tracing. All traces were reviewed by a scorer certified by the American Board of Sleep Medicine. To assess intrarater reliability, 12 traces were randomly selected and blindly rescored (r = 0.997). There was an average scoring difference of three desaturation events per trace.

**Statistical Analysis**

Age, respiratory disturbance index (RDI), and body mass index (BMI) were compared by t test, whereas gender and acceptance of continuous positive airway pressure (CPAP) prescription were compared by χ² for the final study group and the nonstudy group (those refusing consent plus those consenting but rejected later). Descriptive statistics were obtained for the calculated DIs at each of the three recording settings and in the two display modes. A two-factor, within-subjects, repeated-measures analysis of variance was performed to determine the following: (1) whether there were differences between on-line and memory DI values regardless of the recording setting; (2) whether there were differences in recording setting values regardless of display mode; and (3) whether there was interaction between display mode and recording setting. Significant effects at p < 0.05 were followed by post hoc tests. A Value on the post hoc tests was set at 0.001 to allow for multiple comparisons.

Additionally, to examine the relationship between on-line and memory displays, two approaches were taken. Pearson correlation coefficients (one-tailed) were calculated for the mean on-line and memory DI pairs. A second approach, the Bland-Altman method, was used to assess the agreement between the on-line and memory values at each setting.

**Results**

**Sample**

From January through August of 1998, 140 patients were invited into the study and 109 gave written informed consent. Mechanical equipment malfunctioning occurred in 6 patient studies, and technologist error (eg, failure to properly download memory data) was noted in 11 studies. In 17 other patients, more than four movement artifacts per hour (eg, the writer pen moved nonphysiologically) were found in one or more of the six SpO2 traces, which resulted in the exclusion of the data. Seventy-five patients had acceptable data for analysis and are reported here. There were 58 men and 17 women.
(mean age, 51 years; range, 21 to 74 years). Gender, age, RDI, BMI, and acceptance of CPAP prescription were compared for the study group (n = 75) and the nonstudy group (n = 65). No significant differences between the groups were found per t test for age, RDI, or BMI, nor by $\chi^2$ for gender or CPAP acceptance.

**Spo$_2$ Traces**

Visual comparison of the Spo$_2$ traces obtained at the different recording settings showed progressive changes in the detail of the analog signals. Figure 1 is a segment of an Spo$_2$ tracing from a single patient acquired simultaneously in on-line real-time fashion at three different recording settings. Although the overall morphology of the three on-line traces is quite similar, there is an apparent loss of resolution from the 3-s to 6-s to 12-s recording settings. Figure 1 also allows direct comparison of the Spo$_2$ traces acquired on-line with those obtained from the memory of the oximeter. The memory traces appear similar to the on-line version, but there is a consistent loss of detail in the signal at each recording setting.

**Desaturation Analysis**

The mean DI values and their SDs per recording setting and display mode are displayed in Table 1. Repeated-measures analysis of variance for the DIs indicated that there was no significant interaction between display mode and recording setting. However, there was a significant effect for display mode regardless of recording setting. On-line values were significantly higher than the memory DIs. The analysis also showed significant differences in the DIs depending on whether the recording setting was 3 s, 6 s, or 12 s. Post hoc analysis, followed by pair-wise comparisons, indicated that all pairs differed significantly from each other ($p < 0.001$).

Pearson correlation coefficients for the on-line and memory pairs showed a strong relationship between the DIs from the two display modes: OL3/MEM3 pair, $r = 0.902$ ($p < 0.001$); the OL6/MEM6 pair, $r = 0.861$ ($p < 0.001$); and the OL12/MEM12 pair, $r = 0.818$ ($p < 0.001$). The high correlation values between the on-line and memory display modes suggested that one display mode could be substituted for the other. However, since the analyses indicated statistically significant differences in the DIs, the relationship between on-line and memory was explored further. Bland-Altman analysis has been suggested as an alternative to Pearson correlation when values are likely to be correlated, but where the differences between the measures could be clinically meaningful. Figures 2–4 are Bland-Altman plots that show the level of agreement between the DIs obtained from the two display modes (on-line and memory) at each recording setting. Inspection of the distribution of data points on each plot revealed that the agreement between the pairs varied across the range of disordered breathing, as demonstrated on the x-axis. At lower levels of disordered breathing, the DI pairs showed higher agreement, demonstrating smaller differences and staying clustered well within the SD bars. However, at higher levels of disordered breathing, there was more variance in agreement, with some differences between the two display modes exceeding the SD limits.

![Figure 1. Simultaneous Spo$_2$ traces from a single patient.](image)

<table>
<thead>
<tr>
<th>Display Mode/Recording Setting</th>
<th>Mean DI</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>OL3</td>
<td>29.90</td>
<td>27.89</td>
</tr>
<tr>
<td>MEM3</td>
<td>14.77</td>
<td>15.32</td>
</tr>
<tr>
<td>OL6</td>
<td>22.53</td>
<td>22.72</td>
</tr>
<tr>
<td>MEM6</td>
<td>8.94</td>
<td>10.39</td>
</tr>
<tr>
<td>OL12</td>
<td>14.37</td>
<td>15.93</td>
</tr>
<tr>
<td>MEM12</td>
<td>5.29</td>
<td>6.61</td>
</tr>
</tbody>
</table>
Desaturation Thresholds

Given that six DIs were calculated for each of the 75 patients, a total of 450 total DIs were generated from the six experimental conditions. Using a DI threshold of five events per hour, 163 of the 450 traces (36%) would have been classified as negative (< 5/h). At a DI threshold of 10/h, 249 of the total traces (55%) would be negative (< 10/h). The order of the six settings from lowest to highest in terms of the percentage missed at each setting was as follows: OL3 (6%), OL6 (13%), MEM3 (14%), OL12 (17%), MEM6 (20%), and MEM12 (28%) at a threshold of < 5/h, and OL3 (8%), OL6 (12%), MEM3/OL12 (16%/16%), MEM6 (21%), and MEM12 (26%) at a threshold of < 10/h.

Discussion

Recording Settings

This study indicates that significantly different DIs were obtained from the three recording settings. These data confirm and extend the morphologic and numeric changes in SpO2 data reported by Farre et al3 and Kendrick et al4 from variable settings in this acquisition parameter. The oximeter manufacturer reports that a 30 sample per second moving time average mechanism is used in the Ohmeda model 3740, which results in 90 samples, 180 samples, and 360 samples for the 3-s, 6-s, and 12-s recording settings, respectively.1 The running weighted average value per time set is then plotted every 1/3 s at the 3-s setting, every 2/3 s for the 6-s setting, and every 4/3 s for the 12-s setting. The progressive smoothing influence in the on-line SpO2 traces as seen in Figure 1 likely reflects the differential averaging and plotting frequencies of the recording settings.

A pertinent concern for sleep facilities with regard to this particular acquisition parameter and oximeter (Ohmeda model 3740) is that two different generations of this same model default to two different settings (6 s and 12 s) if not manually set at each recording session. Neither automatically defaults to the 3-s setting. Thus, each time the oximeter is powered on, a particular setting would need to be

![Figure 2. Bland-Altman plots of on-line and memory DIs at the 3-s recording setting. The y-axis shows the differences between on-line (OL) and memory (MEM) DI pairs per patient. The x-axis shows the average DI (sum of pairs divided by 2) of those pairs, and is a scale of severity of disordered breathing. The hatched horizontal bars represent SDs above and below the average difference between the on-line and memory DIs, represented as a solid bar.](http://journal.publications.chestnet.org/pdfaccess.ashx?url=/data/journals/chest/21984/)
selected that could predispose to recording variability in large busy facilities where multiple generations of oximeters might be in use. The newer Ohmeda model 3900 also is said to default to the 12-s recording setting, although user-definable options for the 3-s and 6-s settings are available.

Display Modes

The differences in on-line and memory DIs reported here also support and extend the findings of Wiltshire et al\textsuperscript{17} and Kendrick et al\textsuperscript{18} with respect to the effects of the different display modes. Compar-

**Figure 3.** Bland-Altman plots of on-line and memory DIs at the 6-s recording setting. The y-axis shows the differences between on-line and memory DI pairs per patient. The x-axis shows the average DI (sum of pairs divided by 2) of those pairs, and is a scale of severity of disordered breathing. The hatched horizontal bars represent SDs above and below the average difference between the on-line and memory DIs, represented as a solid bar. See Figure 2 for expansion of abbreviations.

**Figure 4.** Bland-Altman plots of on-line and memory DIs at the 12-s recording setting. The y-axis shows the differences between on-line and memory DI pairs per patient. The x-axis shows the average DI (sum of pairs divided by 2) of those pairs, and is a scale of severity of disordered breathing. The hatched horizontal bars represent SDs above and below the average difference between the on-line and memory DIs, represented as a solid bar. See Figure 2 for expansion of abbreviations.
ison of the on-line and memory pairs showed high correlation by Pearson coefficients, implying a strong overall relationship between mean DI values from the two display modes at the three recording settings, although there was some deterioration in this relationship from the 3-s to 6-s to 12-s settings. Bland-Altman plots, however, showed that the level of agreement between the on-line and memory pairs was poor at times and varied with levels of severity. Thus, it would appear that although memory values track with the on-line data, they systematically underestimate it, and the variance is greatest at more severe levels of sleep-disordered breathing. The manufacturer (Datex-Ohmeda) reports that the memory or trend function of the oximeter selects the lowest value from contiguous 12-s windows of data at each recording setting and plots approximately five values per minute. This selective memory and limited plotting function likely explains the drop-off in resolution seen from on-line to memory display modes and the variance in agreement at higher levels of disordered breathing.

Of clinical importance with respect to both acquisition parameters is that the absolute differences in the DIs among the three recording settings and between on-line and memory displays could have an impact on case finding depending on the DI threshold employed. Based on these data, many of the same patients detected at faster recording settings with on-line displays using commonly cited DI cutoffs of 5 to 10/h would be missed at slower recording settings, especially if the memory display mode were employed, which is often the case in outpatient settings where even higher thresholds (10 to 20/h) have been used.

Another related clinical concern with respect to the downloaded display format is that Medicare is proposing that oximeters have this capability to assess oxyhemoglobin saturation in sleep when determining the need for home oxygen therapy for cardiopulmonary disorders. Dependence on data from this memory display mode, especially if slower recording settings were chosen, could significantly underestimate the burden of desaturation in sleep and possibly preclude some from qualifying for this therapy. This could also be a concern for clinicians who employ computerized scoring systems, such as Profox (Profox Associates; Escondido, CA), that can analyze stored oximetry data.

**Prior Publications**

The literature suggests that oximetry has variable reliability as a case finding tool for sleep-disordered breathing when using the polysomnography-derived RDI as the criterion standard. A review of these prior investigations reveals inconsistent oxyhemoglobin desaturation thresholds in the definitions for the desaturation events on oximetry (1 to 5%) and for the respiratory disturbance events on polysomnography (1 to 4%) within and across studies. The numeric cutoffs for DIs (5 to 15/h) and RDIs (5 to 25/h) also varied considerably within and among the reports. Hence, a wide range of sensitivity and specificity values have been cited for DIs in estimation of the RDIs, likely related to these methodologic issues.

In addition, a further inspection of these primary publications reveals that the acquisition parameters on the oximeters used for DI calculation and on polysomnography for the RDI were rarely explicitly mentioned. Of note, five studies employed relatively slower recording settings, 6 to 12 s, and four studies used memory data. In four other reports, relatively faster settings (3 to 6 s) were employed, with three of the four acquiring data on-line. The sensitivity with which disordered breathing was detected was lower (51 to 75%) and specificity was higher (86 to 96%) in studies that used relatively slower recording settings and the memory display mode. Higher sensitivity (94 to 100%) and lower specificity (45 to 64%) values were obtained with faster recording settings and more frequent use of the on-line display mode. These trends likely reflect the fact that more averaged values are plotted at slower recording settings, and selective sampling and plotting is employed by the memory function. Using the visual inspection method of Series et al, the smoother analog SpO₂ traces resulting from slower recording settings, and the memory display mode might make some cases of sleep-disordered breathing appear equivocal and thus lower sensitivity by increasing the false-negative rate. In contrast, the faster recording settings and on-line traces are likely to show more detail and appear more positive by the visual method, as well as by a counting method of desaturation events. Thus, while numerous definitional factors could influence the performance of oximetry in case finding of sleep-disordered breathing, the employment of multiple acquisition parameters may, in part, explain some of the variance in prior reports as well.

**Conclusion**

Significantly different DIs are obtained from an oximeter when different recording settings and display modes are selected, which appear to reflect the processing algorithms of the instrument. Clinicians should be aware of these operating characteristics and should preferentially employ the fastest-available recording settings in the on-line display mode to...
optimize case finding of sleep-disordered breathing. Nonsystematic use of both acquisition parameters may explain some of the variable performance reports in the literature of oximetry in case finding for sleep-disordered breathing. These two acquisition parameters should always be disclosed to allow for better comparisons of reported data. Efforts to standardize measurement techniques should include specifications regarding the acquisition parameters of oximeters for clinical and research use.26,27

ACKNOWLEDGMENT: The authors thank Peter C. Gay, MD, Susan M. Harding, MD, and Donald L. Bliwise, PhD, for review of this article, and Regina Huddleston, Jay Jordan, Leah Osbahr, Bridgett Tharp, and Brandon Phillips for technical assistance.

REFERENCES

27 Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research; the report of the American Academy of Sleep Medicine Task Force. Sleep 1999; 22:667–689