Comparison of Home Oximetry Monitoring With Laboratory Polysomnography in Children*

Valerie G. Kirk, MD, FCCP; Shelly G. Bohn, BSc; W. Ward Flemons, MD; and John E. Remmers, MD

Study objectives: To measure the accuracy and reliability of a portable home oximetry monitor with an automated analysis for the diagnosis of obstructive sleep apnea (OSA) in children.

Design: Prospective cohort study.

Setting: Alberta Lung Association Sleep Center, Alberta Children’s Hospital Sleep Clinic.

Study subjects: Consecutive, otherwise healthy children, aged 4 to 18 years, presenting to the Pediatric Sleep Service at the Alberta Children’s Hospital for assessment of possible OSA.

Interventions: All subjects underwent 2 nights of monitoring in the home with an oximetry-based portable monitor with an automatic internal scoring algorithm. A third night of monitoring was done simultaneously with computerized laboratory polysomnography according to American Thoracic Society guidelines.

Measurements and results: Both test-retest reliability of the portable monitor-based desaturation index (DI) between 2 nights at home and between laboratory and home were high using the Bland and Altman analysis (mean agreement, 0.32 and 0.64; limits of agreement, –8.00 to 8.64 and –0.75 to 6.50, respectively). The polysomnographic apnea-hypopnea index (AHI) agreed poorly with the portable monitor DI (mean difference, 1.27; limits of agreement, –12.02 to 15.02). The sensitivity and specificity of the monitor for the identification of moderate sleep apnea (polysomnography AHI > 0.5/h) were 67% and 60%, respectively.

Conclusion: Portable monitoring based only on oximetry alone is not adequate for the identification of OSA in otherwise healthy children. (CHEST 2003; 124:1702–1708)

Key words: abbreviated monitoring; ambulatory monitoring; children; diagnosis; obstructive sleep apnea; oximetry; sleep apnea syndromes

Abbreviations: AHI = apnea-hypopnea index; DI = desaturation index; OSA = obstructive sleep apnea; RDI = respiratory disturbance index
OSA in children has been studied. Oximetry interpretations were done manually using both oximetry and pulse waveform data. The positive predictive value was 97%, suggesting an abnormal study finding would eliminate the need for full polysomnography. However, the sensitivity of overnight oximetry was low (40%), confirming that a normal oximetry result does not rule out a diagnosis of OSA in children.4 Subsequent to this report, Urschitz et al5 published reference values for nocturnal home pulse oximetry in young children. We elected to evaluate a simple, oximetry-based portable monitor that has been validated in adults suspected of having sleep apnea6 to determine its diagnostic accuracy and reliability in a population of children referred to a pediatric sleep clinic with suspected OSA.

**MATERIALS AND METHODS**

**Patient Selection**

A consecutive sample of otherwise healthy children aged 4 to 18 years were studied. The minimum age of 4 years was selected to minimize data loss related to incomplete home studies and to ensure that all study subjects were in a low surgical risk group.7,8 All patients were referred for evaluation of suspected OSA to the Pediatric Sleep Service at the Alberta Children’s Hospital by community general practitioners, pediatricians, or otolaryngologists between October 1, 2000 and January 31, 2002. Children with complex medical conditions were excluded (Table 1) from enrollment. Ethics approval for the study was received from the Conjoint Health Research Ethics Board, Faculty of Medicine, University of Calgary. Consent to participate was obtained from all primary caregivers, and direct assent was obtained for all children aged ≥ 12 years.

**Laboratory Polysomnography**

Computerized laboratory polysomnography was performed according to American Thoracic Society guidelines1 using the Sandman NT (Nellcor Puritan Bennett; Ottawa, ON) and included ECG (C2-A1, C3-A2, O1-A2, O2-A1); electro-oculogram; submental electromyogram; ECG; oxygen saturation monitoring (N200; Nellcor Puritan Bennett); chest, abdominal wall, and arm channel movements using respiratory inductance plethysmography (Respitrace Plus; SensorMedics Corporation; Yorba Linda, CA); bilateral tibial electromyograms; nasal/oral airflow using a thermostor device (Edentec; Eden Prairie, MN); nasal pressure (Ultima Airflow Pressure Sensor Model 0590; Braebon Medical Corporation; Carp, ON), end-tidal carbon dioxide monitoring (Model 1265; Novametrix Medical Systems; Wallingford, CT), and transcutaneous carbon dioxide monitoring (Radiometer Compact Combined pCO2/pO2 Monitoring System; TCM3; Radiometer Medical; Copenhagen, Denmark). Sleep architecture was determined using standard criteria.9 A trained sleep technician, blind to portable monitor results, scored sleep architecture and respiratory events using specific scoring criteria (Table 2).1,4 OSA was defined as an apnea-hypopnea index (AHI) > 1/h of total sleep time in order to be consistent with current literature.1,4,10,11

**Protocol**

Three different recordings using the oximetry monitor were performed. During laboratory polysomnography, the oximetry unit received data from an additional oxygen saturation monitor. Two unattended studies using only the portable monitor were performed in the patients’ homes as well. These were done within 1 week of the overnight laboratory polysomnography. Patients were randomized to determine the order of laboratory and home studies. For unattended home study nights, the child’s caregiver received the portable monitor from a trained technician at the hospital and was instructed on its use. The caregiver applied the sensor to the child before going to sleep: the oxygen saturation finger/toe probe. The monitor was initiated and then turned off/unplugged in the morning. After a night of home study, the unit was either returned to the sleep laboratory or a technician would go to the home in order to download the data to a personal computer for automatic analysis.

**Scoring of Portable Monitor Studies**

Unlike polysomnography and standard oximetry testing that require manual interpretation, the portable monitor (SnoreSat; SagaTech Electronics; Calgary, AB, Canada) uses a nonproprietary automated analysis algorithm to interpret received oximetry data. Only the oxygen saturation signal (sampled at 1 Hz) is used to determine the monitor desaturation index (DI). The oximeter board uses an exponential filter to average the signal sent to the portable monitor. Specifically, the averaging calculation is based on a four-beat exponential average and an eight-beat sliding average for pulse rate. All averages are updated on a beat-by-beat basis. With the four-beat exponential average, the effect of each measurement gradually decreases beat by beat. Each measurement initially counts for one fourth of the average. This weight is decreased by multiplying three fourths on each succeeding beat: beat 1 × ¼ = 0.250; beat 2 × ¼ = 0.1875; beat 3 × ¼ = 0.1406; beat 4 × ¼ = 0.1055 (and so on... this is an infinite series). The following calculation is then made on each valid beat: new average = old average + (new oxygen saturation – old average)/4; in other words, the new average value is the prior average plus one fourth the difference of the new saturation value from the prior average value. For pulse rates > 112, the 4 is replaced by an 8; for rates > 225, the 8 becomes 16.

The SnoreSat algorithm sequentially scans each recorded oxygen saturation value, and whenever a drop in a sampled oxygen saturation value is detected, the program assigns an event marker to that reading. When an increase in oxygen saturation is detected, the program determines if at least three consecutive event markers (ie, three consecutive drops in recorded oxygen saturation readings) were present prior to this rise. If this criterion is met and if the lowest oxygen saturation value is > 3% lower than the baseline oxygen saturation, then a respiratory

---

**Table 1—Exclusion Criteria**

<table>
<thead>
<tr>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down syndrome</td>
</tr>
<tr>
<td>Pierre Robin syndrome</td>
</tr>
<tr>
<td>Cleft palate</td>
</tr>
<tr>
<td>Cerebral palsy</td>
</tr>
<tr>
<td>Craniofacial synostosis syndromes (Apert, Crouzon)</td>
</tr>
<tr>
<td>Neuromuscular disease</td>
</tr>
<tr>
<td>Congenital heart disease</td>
</tr>
<tr>
<td>Achondroplasia</td>
</tr>
<tr>
<td>Myelomeningocele</td>
</tr>
</tbody>
</table>

---

**CHEST / 124 / 5 / NOVEMBER, 2003**

© 2003 American College of Chest Physicians
disturbance is recorded. Baseline saturation was calculated as a moving time average, defined as the mean of the top fifth percentile of oxygen saturation values over the 5 min preceding the event. The DI is calculated by dividing the total number of respiratory disturbances by the total monitor “probe-on” time. Probe-on time is the total time the oximeter reports a valid oxygen signal. Although the automated analysis algorithm is not able to detect arousals from sleep, this did not affect the sensitivity and specificity of the monitor when compared with polysomnography in a large group of adults with OSA.

Data Analysis

Agreement between the polysomnography-derived AHI and the simultaneously recorded portable monitor-generated DI was examined using the technique described by Bland and Altman.12 The analysis compared results obtained using the new measurement technique (portable monitor) with the established one (polysomnography) by calculating the mean of the differences between the two measurements and then plotting the means against the average value. The mean and limits of agreement (95th confidence intervals) were calculated to identify how tightly and consistently the two techniques agreed within the full spectrum of disease. For example, a mean difference of $< 1$ with ninety-fifth confidence intervals of $-1.5$ and $1.5$ would describe very close agreement with little systematic bias. Agreement between the DI reported for nights 1 and 2 of the home SnoreSat study was examined as well as the agreement between the DI reported during the laboratory SnoreSat study and the mean of the two home studies using the Bland-Altman analysis (Stat Version 6.0; Stata Corporation; College Station, TX). This analysis technique was used to ensure accurate interpretation of the relationship between these two measurement tools, and is superior to simple correlation calculations when comparing a new tool to a standard testing technique, particularly when a range of test results is anticipated.

Sensitivity and specificity of the portable monitor for the identification of polysomnography-proven OSA (AHI > 1/h) were calculated. Receiver operator curves were generated by calculating the sensitivity and specificity of the portable monitor using specific DI case designations (0.5, 1.0, 2.0, 5.0, and 10.0 events per hour). This analysis allows for identification of the optimal case designation DI: the DI at which the monitor displays the highest sensitivity and specificity. This analysis was also performed for the identification of polysomnography-proven moderate OSA (AHI > 5/h) in order to determine if the portable monitor would be more sensitive and/or specific for identifying more severely affected children.

RESULTS

During the enrollment period, a total of 58 children (32 boys) met inclusion criteria and were enrolled into the study. Complete data were available for 57 children, indicating that is was feasible to obtain data without a technician present. One subject withdrew after having difficulty sleeping with the portable monitor at home during the first of 3 study nights. Seven children were $\geq 13$ years old (three boys). The remaining subjects were evenly distributed between 4 years and 7 years of age ($n = 25, 13$ boys) and 8 to 12 years of age ($n = 26, 16$ boys). The prevalence of OSA was 79%. Polysomnography results are summarized in Table 3. Forty-six percent ($n = 27$) had moderate OSA with a mean AHI of 14.4/h and a mean oxygen saturation nadir of 84.6%. End-tidal carbon dioxide maximal and mean levels were slightly elevated in all groups; however, the range of maximal values was higher in the children with OSA.

Measures of Agreement

The polysomnography-derived AHI and the oximetry-based monitor-derived DI showed poor agreement (Fig 1). The means and limits of agreement between AHI and DI were dependent on the AHI cutoff points and were much tighter for the subgroup of patients with an AHI $< 10$/h (mean, $-0.37$; limits of agreement, $-9.17$ to 8.43). Visual inspection of the Bland and Altman plot indicates that clinically significant differences between portable monitor DI and polysomnography-derived AHI occurred at all AHI values with a systematic bias noted at higher AHI values. The portable monitor consistently underestimated the frequency of respiratory events at higher AHI values (Fig 1). The Bland and Altman analysis of the portable monitor used in the laboratory compared with values obtained when used in the home demonstrated good agreement with no systematic bias regardless of what night of home study was used for the comparison (Fig 2). Similar results were obtained by analyzing night-to-night results of portable monitoring in the home, indicating good repeatability of the monitor (Fig 3).

The sensitivity and specificity of the SnoreSat device depended on the DI case designation criterion values chosen for diagnosing OSA. Receiver operator curves (Fig 4) demonstrate the effect of

---

**Table 2—Definitions for Respiratory Events Scored on Laboratory Polysomnography**

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstructive apnea</td>
<td>$\geq 80%$ decrease in amplitude on the respiratory inductance plethysmographic sum channel of any duration associated with continued and paradoxical chest and abdominal wall motion</td>
</tr>
<tr>
<td>Hypopnea</td>
<td>A 50 to 80% decrease in amplitude on the respiratory inductance plethysmographic sum channel associated with a $\geq 4%$ drop in oxygen saturation measurement and/or a cortical arousal</td>
</tr>
<tr>
<td>Arousal</td>
<td>A change in EEG pattern of $\geq 3$ s in duration. In rapid eye movement sleep, an increase in chin electromyogram amplitude must accompany the EEG changes.</td>
</tr>
</tbody>
</table>
changing the threshold for a positive portable monitor result (i.e., sensitivity and specificity for identifying polysomnography-proven OSA at various DI cutoff points). At a DI cutoff value of >5/h, the automated analysis algorithm had a sensitivity of 66.7% (95% confidence interval, 54 to 79%) and specificity of 60% (95% confidence interval, 47 to 73%) for identifying children with moderate OSA (polysomnography AHI >5/h). The portable monitor misclassified 9 children with moderate OSA (polysomnography AHI >5/h; range, 5.6 to 27.2/h).

### Table 3—Polysomnography Findings in the Study Population of 58 Children

<table>
<thead>
<tr>
<th>Parameters</th>
<th>AHI &lt;1 (n = 11)</th>
<th>AHI ≥1 (n = 47)</th>
<th>AHI ≥5 (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Median (Range)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Total sleep time, min</td>
<td>381.8 (50.7)</td>
<td>372.2 (350.0–407.1)</td>
<td>418.3 (33.2)</td>
</tr>
<tr>
<td>AHI, events/h of total sleep time</td>
<td>0.4 (0.3)</td>
<td>0.3 (0.0–0.9)</td>
<td>9.4 (9.9)</td>
</tr>
<tr>
<td>Sleep efficiency, %</td>
<td>58.2 (5.5)</td>
<td>85.9 (78.0–94.2)</td>
<td>90.5 (4.2)</td>
</tr>
<tr>
<td>Minimum oxygen saturation, %</td>
<td>91.4 (1.8)</td>
<td>91.7 (89.2–94.1)</td>
<td>86.3 (6.7)</td>
</tr>
<tr>
<td>Mean oxygen saturation, %</td>
<td>97.0 (1.0)</td>
<td>97.0 (95.4–98.6)</td>
<td>96.8 (1.2)</td>
</tr>
<tr>
<td>Maximum transcutaneous PCO₂ mm Hg</td>
<td>53.0 (5.7)</td>
<td>52.2 (43.1–62.3)</td>
<td>52.0 (4.7)</td>
</tr>
<tr>
<td>Mean transcutaneous PCO₂ mm Hg</td>
<td>44.7 (4.3)</td>
<td>44.4 (38.0–50.7)</td>
<td>44.4 (3.7)</td>
</tr>
<tr>
<td>Maximum end-tidal CO₂ mm Hg</td>
<td>55.1 (2.3)</td>
<td>55.2 (52.0–57.9)</td>
<td>54.3 (3.4)</td>
</tr>
<tr>
<td>Mean end-tidal CO₂ mm Hg</td>
<td>47.8 (2.7)</td>
<td>47.6 (41.9–51.8)</td>
<td>46.8 (2.9)</td>
</tr>
</tbody>
</table>

**Figure 1.** Bland and Altman comparison of laboratory polysomnography AHI and oximetry DI, illustrating the agreement between laboratory-derived AHIs and the DIs derived from simultaneous laboratory oximetry monitoring. Analysis of subgroups (i.e., average AHI/DI ≤10/h and >10/h) shows closer agreement among those with lower respiratory event indexes and a strong systematic measurement bias at higher values.

**Figure 2.** Bland and Altman comparison of the DIs using the portable monitor in the home and in the laboratory (Lab), illustrating the agreement between home and laboratory-based ambulatory monitoring. The average DI for the 2 nights of home study was compared with the laboratory-based DI. The chart illustrates the mean differences and limits of agreement when analysis is done for nights 1 and 2 separately.
as normal, and 12 normal children (polysomnography AHI < 1.0) as having moderate OSA (i.e., false-negative rate of 33%, false-positive rate of 40%).

The means of the differences between AHI and DI was large. Visual inspection of the Bland and Altman plot would suggest that systematic bias occurs at both lower and higher AHI values (overestimation by the portable monitor at low levels, and underestimation at high levels). The sensitivity and specificity of the oximeter depended on the DI case designation criterion values chosen for diagnosing OSA (Table 3). At a DI and AHI cutoff value of > 5/h, the automated analysis algorithm had a sensitivity of only 67% and a specificity of only 68%. The algorithm missed a diagnosis of moderate OSA (AHI > 5/h; range, 5.6 to 27.2/h) in 8 children (false-negative rate of 14%) and misdiagnosed moderate OSA in 13 children (false-positive rate of 22%).

In all but one case, complete data were obtained from the home studies, indicating excellent feasibility of home monitoring for children. The algorithm showed remarkable consistency between home and laboratory (Fig 2, r = 0.86). There was little night-to-night variability as well (Fig 3, r = 0.83).

**DISCUSSION**

Despite excellent feasibility and repeatability of the SnoreSat monitor used in the home, poor agreement with the reference standard diagnostic test for OSA in children was found. Although our data confirm that home monitoring without sleep technicians is feasible, the results obtained with the portable monitor did not agree sufficiently with standard testing to advocate for indiscriminate home oximetry testing as an alternative to laboratory polysomnography. The repeatability of the portable monitor from night to night and between home and laboratory settings confirms a systematic error with respect to identifying OSA in children.

Obstructive respiratory events are uncommon in children and are generally considered significant if present at a frequency of > 1/h. A cutoff point of five events per hour should clearly identify children with moderate abnormalities; that is, children most likely at increased risk of adverse outcomes. The portable monitor had poor sensitivity and specificity for identifying this group of children, even when analyzed with various DI cutoff points (Figs 3, 4). The discrepancy between our results and those reported by Vazquez et al highlights the dissimilarity between adults with OSA and children with OSA.

There are important differences in the physiology...
and pathophysiology of sleep and OSA in children that may impact our ability to obtain equivalent accuracy with identical technological devices used for adults. Firstly, the cutoffs for clinically significant OSA in adults are much higher, thus eliminating the need for discrimination between what may be considered low and very low AHI values. Children with OSA also frequently display prolonged episodes of partial upper airway obstruction without discreet abrupt desaturations. These episodes may be missed completely by oximetry-based testing due to the omission of ventilatory parameters such as end-tidal carbon dioxide or transcutaneous carbon dioxide monitoring and measurements of chest and abdominal wall motion.

Lastly, children are not only more restless during sleep, and those with OSA have been shown to have significantly more movement arousals than age-matched control subjects, increasing oximetry movement artifact signals. Scholle and Zwacka reported a mean of 20.4 movement arousals per hour in a group of untreated children with OSA. Movements also occur in response to airway obstruction or partial obstruction, further increasing the potential for movement artifact, which may be associated with low saturation recordings.

Inaccurate data can be identified and eliminated during the standard manual scoring procedure for laboratory polysomnography based on inspection of the data and inclusion of the pulse waveform variability into the determination of "bad data" or movement artifact. This likely explains some of the discrepancy between our findings and those published by Brouillette et al. Although using artifact-free recording time would be expected to reduce the overall DI in children, Urschitz et al. reported that drops in oxygen saturation were actually quite common (mean, 1.2/h; 95th percentile, 3.9/h) in their population of 90 unselected children. They speculate that the increased frequency of desaturation during sleep in their population compared with the previously published normative data on children may be due to the more restricted age range, stricter definition of a desaturation event (> 4% vs ≤ 4%), or the different monitoring setting. Unlike the study by Brouillette et al., artifact-free time was not identified manually but by data analysis software provided by the device manufacturer, raising the possibility of underestimation of artifact-free recordings.

Although ambulatory monitoring for the diagnosis of OSA in children has not been validated, many clinicians and researchers have adopted this method of diagnosis and are reporting outcomes and associations based on the results of ambulatory monitoring. We were able to identify only two prior studies in which home studies were compared with laboratory polysomnography results in children. Goodwin et al. performed 157 unattended polysomnographic home studies in a self-selected group of school children in order to identify the prevalence of sleep-disordered breathing in their population. Only 5 of the 157 children underwent laboratory studies, and all were performed “within 7 weeks” of the home study; no significant differences were seen. A similar design was employed by Jacob et al., who studied 21 young children in the home and in the laboratory. In both of these studies, the home studies were almost identical to the laboratory studies, and technicians went to the homes to set up equipment and make additional measurements and observations. Although good correlation was reported by both groups of investigators, these studies are not comparable to the ambulatory monitoring frequently done on a research and clinical basis employing unobserved, ambulatory monitoring with limited channels.

Several milestone publications relating epidemiologic data, behavioral and cognitive associations, and outcomes data of OSA in children are based on abbreviated, ambulatory monitoring techniques. Given the importance of these reported associations and the potential effect on changes in clinical care recommendations, it is imperative that these findings be confirmed using validated instruments before adopting changes in medical practice. Our findings suggest that, contrary to the recent American Academy of Pediatrics guidelines, even a positive abbreviated test result in a child with symptoms of OSA may need to be viewed with some skepticism, as the false-positive rate was unacceptably high in our study population.

Interpretation of data that are obtained using inaccurate measurement tools is obviously hazardous, particularly if changes in medical management or recommendations are affected. For example, evidence supports a correlation between preoperative polysomnography indexes and perioperative complications in children undergoing adenotonsillectomy for OSA. In a population of 349 children undergoing adenotonsillectomy, Wilson et al. reported an increased risk of postoperative respiratory complications (odds ratio, 7.2) for children with a polysomnography AHI ≥ 5/h. Clearly, if perioperative morbidity can be predicted based on polysomnography parameters, the agreement between ambulatory monitors and laboratory polysomnography in children must be confirmed before ambulatory monitoring can replace laboratory polysomnography as a preoperative assessment tool.

There are a few limitations to this study that may have impacted our results. The oximeter calculated
events based on probe-on time and not artifact-free time. This may explain some of the disagreement with laboratory polysomnography during which manual inspection of the pulse waveform signal is used to identify artifact-free time. Although unselected, otherwise healthy children were enrolled, they may not accurately represent the typical pediatric population undergoing evaluation for OSA. Our sleep service is located at a tertiary level institution and offers the only access to complete pediatric diagnostic sleep facilities in Western Canada.

Although laboratory polysomnography is the accepted “gold standard” test, there are currently no scientific data correlating polysomnography indexes to adverse clinical medical outcomes in children. Even among the abundant adult sleep apnea literature, there is no conclusive evidence that the risks of cardiovascular and cerebrovascular complications are related to specific polysomnographic measurements or that treatment of OSA results in improvement of these medical complications.22 Given this limitation of our “gold standard” measurement tool, comparison of polysomnography indexes with the portable oximetry determined respiratory disturbance indexes (RDIs) may not be the optimal approach to validation of a home monitor. Clinically, decisions regarding treatment of children with OSA are based on a combination of factors, including the daytime behavioral and learning problems experienced and reported by the child, parents, and teachers, as well as the results of diagnostic testing.

Clearly, simple oximetry measurements, with or without motion sensitivity, cannot replace laboratory polysomnography, given the poor sensitivity. Perhaps use of a motion-sensitive oximeter in association with some measure of airflow and respiratory effort will improve on both the sensitivity and specificity for identifying children with OSA, yet still allow for ambulatory monitoring to replace polysomnography in otherwise healthy children. Given the differences between adults and children with OSA, one cannot assume that diagnostic instruments previously validated in an adult population are equally as accurate when used in children.

ACKNOWLEDGMENT: The authors thank Dr. C. Bjornson for her assistance with data analysis and Eric Hadjuk for biomedical support.

REFERENCES