Agreement Between Oxygen Desaturation Index and Apnea–Hypopnea Index in Adults With Suspected Obstructive Sleep Apnea at an Altitude of 2240 m

Luis Torre-Bouscoulet, Armando Castorena-Maldonado, Rocío Baños-Flores, Juan Carlos Vázquez-García, María Sonia Meza-Vargas, and Rogelio Pérez-Padilla

Clínica de Sueño, Instituto Nacional de Enfermedades Respiratorias, México, DF, Mexico

OBJECTIVE: Altitude can affect the diagnostic accuracy of portable monitors used to diagnose suspected obstructive sleep apnea syndrome on the basis of oxygen desaturation measurements. The aim of this study was to determine agreement between the desaturation index measured by oximetry and the apnea–hypopnea index measured by polysomnography in Mexico City (2240 m above sea level). We also wished to determine agreement between the desaturation index and the respiratory disturbance index measured by monitoring airflow with a single-channel recording device.

METHODS: We used standard polysomnography and nocturnal oximetry (Remmers Sleep Recorder, Sagatech, Calgary, Alberta, Canada) to simultaneously measure the apnea–hypopnea index and the desaturation index, respectively, in a group of 38 patients aged over 18 years with suspected obstructive sleep apnea syndrome. In a second group of 30 patients, we compared the desaturation index to the respiratory disturbance index, which we measured using a single-channel device monitoring nasal airflow (ApneaLink, ResMed Corp., Poway, CA, USA).

RESULTS: The mean (SD) intraclass correlation coefficient between the apnea–hypopnea index and the desaturation index was 0.89 (0.03) (95% confidence interval, 0.83-0.96), and the mean of the differences was –0.9 (14.2). The mean intraclass correlation coefficient for the desaturation index and the respiratory disturbance index was 0.93 (0.02) (95% confidence interval, 0.89-0.97), and the mean of the differences was –6.6 (8.3).

CONCLUSIONS: Agreement was high between the desaturation index and both the apnea–hypopnea index and the respiratory disturbance index in adults with suspected obstructive sleep apnea syndrome in Mexico City.

Key words: Sleep apnea. Altitude. Portable monitor. Apnea–hypopnea index.


Introduction

Polysomnography is the gold standard for diagnosing obstructive sleep apnea syndrome (OSAS). There is, however, an almost universal tendency to simplify diagnostic procedures through the use of portable monitors,
which are classified as type-2, type-3, or type-4 monitors depending on their complexity and the number of signals they record. The diagnostic yield of each of these systems varies with respect to that of the gold standard. Type-4 portable monitors are not accepted as a definitive diagnostic tool.

The automatic analysis of overnight pulse oximetry is a type-4 study in which a range of algorithms are used to measure a patient’s desaturation index (DI), defined by the number of desaturations per hour of recording. Other type-4 monitors calculate a patient’s respiratory disturbance index (RDI) on the basis of a nasal airflow signal (nasal pressure).

Although type-4 portable monitors were developed as a means of screening patients with suspected OSAS, they are now also used for diagnostic purposes. They offer a number of advantages: they are accessible, they normally perform automatic analyses, they do not need to be managed by highly skilled personnel, and they can be used to perform sleep disturbance studies at the patient’s home.

It is essential to compare the diagnostic accuracy of portable monitoring systems with that of polysomnography in validation studies. The pulse oximetry algorithm used in the present study was previously validated by Vázquez et al.\(^1\) at an altitude of 1049 m above sea level. That study, which involved 241 patients, yielded a correlation coefficient of 0.97, a sensitivity of 98%, and a specificity of 88%. Other portable monitor validation studies have also been performed, but at lower altitudes than that of Mexico City.\(^2,4\) This represents a drawback for techniques based on pulse oximetry. The altitude of Mexico City (2240 m above sea level), for example, could affect the performance of the automated analysis algorithms used in pulse oximetry systems as the inhabitants of this city are very near the point where the hemoglobin dissociation curve drops away sharply. In such a situation, a change in PaO\(_2\) accentuates the change in oxygen saturation (SaO\(_2\)), which means that desaturation is more likely to occur following minor decreases in PaO\(_2\). This in turn is likely to result in an overestimation in which the DI would be higher than the apnea–hypopnea index (AHI).

The aim of the present study was to evaluate the agreement between the DI and both the AHI and the RDI in patients with suspected OSAS in Mexico City.

Patients and Methods

We conducted a cross-sectional study consisting of 2 parts: the first comparing the DI and the AHI and the second comparing the DI and the RDI. The study was approved by the institutional review board for science and bioethics and prior informed consent was obtained from all the patients.

Part 1 (Comparison of Desaturation Index and Apnea–Hypopnea Index)

We consecutively enrolled patients aged over 18 years who were referred with suspected OSAS to the sleep clinic at the Instituto Nacional de Enfermedades Respiratorias, Mexico City. All the patients simultaneously underwent standard overnight polysomnography and pulse oximetry at the clinic. Exclusion criteria included severe lung disease (forced expiratory volume in 1 second or forced vital capacity of <50% of predicted), neuromuscular disease, daytime hypoxemia, confirmed heart disease, insomnia, symptoms indicative of restless legs, psychiatric disease, or use of sleeping pills.

Standard baseline polysomnography was performed using a Grass recording device (Grass-Telefactor, Astro Med Inc, West Warwick, Richmond, USA) arranged in a monopolar configuration. The following signals were measured: 2 electroencephalogram channels (C3-A2, O1-A2), 2 electrooculogram channels (O1-A2, O2-A1), electromyography (of the masseter muscle and pelvic muscle), electrocardiogram, and microphone. The respiratory signals measured were inductance plethysmography (Respirtrace, Ambulatory Monitoring Inc, Ardsley, New York, USA) (2 channels), respiratory flow (nasal pressure and thermistor), expired PCO\(_2\), and pulse oximetry. The polysomnograms were scored according to the criteria of Rechtschaffen and Kales.\(^8\) Respiratory signals were scored manually in accordance with current, internationally accepted guidelines.\(^9\) Obstructive apnea was defined as the complete cessation of or a >50% reduction in airflow lasting longer than 10 seconds, accompanied by persistent respiratory effort, with or without oxygen desaturation or arousal. Hypopnea was defined as a reduction in airflow of <50% for longer than 10 seconds, detected by thermistor or nasal cannula, accompanied by oxygen desaturation of ≥4% or an arousal of 1.5 seconds or longer.\(^9\)

The portable monitor used to measure the DI (Remmers Sleep Recorder, Sagatech, Calgary, Alberta, Canada) recorded 5 simultaneous signals: respiratory flow, pulse oximetry, heart rate, body position, and snoring. The automated analysis algorithm was based purely on oximetry and a respiratory event was defined as an increase in SaO\(_2\) preceded by a reduction of more than 3% in baseline oxygen saturation. The baseline oxygen saturation value was defined as the mean of the fifth percentile of all the SaO\(_2\) values recorded during the 5 minutes preceding the event. The diagnostic yield of the monitor used had been determined previously in a study performed in Calgary, Canada, at an altitude of 1049 m above sea level.\(^1\)

Part 2 (Comparison of Desaturation Index and Respiratory Disturbance Index)

To conduct this part of the study, we enrolled patients aged over 18 years with suspected OSAS. Although the inclusion and exclusion criteria were the same as those applied in the first part of the study, the patients were not the same. Each patient was connected to 2 monitors simultaneously: one measured the DI and the other measured apnea and hypopnea via a single-channel nasal airflow monitoring system. The studies were performed at the sleep clinic.

The DI was measured using the same recording device as that used in part 1. The RDI was calculated using a single-channel airflow detection system that measured nasal pressure via nasal cannula (ApneaLink; ResMed Corp, Poway, California, USA). The corresponding automated algorithm considered that an apneic event had occurred when there was a ≥280% reduction in airflow and that a hypopneic event had occurred when there was a reduction in airflow of at least 50%. The signal was manually analyzed to discard periods containing artifacts. We excluded patients in whom 10% of the data recorded had artifacts or was of poor quality. The monitor had been previously validated against polysomnography in Germany at an altitude of 116 m above sea level.\(^6,7\)
Statistical Analysis

Data were expressed as means (SD). The degree of agreement between the DI and both the AHI and the RDI was calculated using the Pearson and Spearman correlation coefficients. Agreement between matching variables was measured using the intraclass correlation coefficient. Measurement errors were obtained by calculating the mean and standard deviation of the individual differences, and displayed using modified Bland-Altman plots. For part 1, we also calculated sensitivity, specificity, and the area under the curve (AUC). Several apnea–hypopnea cutoff values were used as reference values and compared with several DI cutoff values. Statistical significance was set at a value of $P<.05$. All analyses were performed using the Stata software package (Stat Version 9.2, StatCorp, College Station, Texas, USA).

Results

Part 1

The characteristics of the patients that participated in the 2 parts of the study are shown in Table 1. A total of 38 patients (28 men) with a mean (SD) age of 46 (10.4) years were included in part 1 of the study. Figure 1 shows the correlations observed between the DI and the AHI. The Pearson and Spearman correlation coefficients both had a value of 0.89. The intraclass correlation coefficient for the DI and the AHI was 0.89 and the mean difference (AHI–DI) was –0.9; this would imply that the AHI was slightly higher than the DI. Figure 2, however, shows how individual differences varied slightly as the AHI increased.

Because the differences between the AHI and the DI (AHI–DI) were uniformly negative for an AHI cutoff value of less than 35, we calculated the intraclass correlation coefficient according to subgroups. In patients with an AHI of less than 35 (n=13), the intraclass coefficient correlation was 0.59 (with a mean difference of AHI–ID of –7.6), while in patients with an AHI of greater than 35 (n=25), it was 0.72 (mean difference of AHI–ID of 2.5).

Table 2 shows the diagnostic yield based on a comparison between 4 reference AHI cutoff points (5, 10, 15, and 20) and 4 DI cutoff values (5, 10, 15, and 20). The largest AUC was obtained for a DI cutoff value of 10 and an AHI cutoff value of 5.
PART 2

Thirty patients (25 men) with a mean age of 49.8 (12.9) were enrolled. Figure 3 shows the correlation between the DI and the AHI. The Pearson and Spearman correlation coefficients were 0.96 and 0.94, respectively, and the intraclass correlation coefficient was 0.93. The mean difference (RDI–DI) was negative (–6.6), indicating that the number of desaturations exceeded the number of apneas and hypopneas measured by nasal airflow (Figure 4).

Discussion

OSAS is a serious public health problem. Its high prevalence (estimated to be 2% in women and 4% in men\textsuperscript{10,11}), combined with the increased recognition it has received in recent decades, has led to a spectacular rise in the demand for studies of OSAS.\textsuperscript{12} Pellicer-Cís car,\textsuperscript{12} for example, recently indicated that OSAS has now become the most common reason for referral to a respiratory medicine specialist in Spain. In view of the high prevalence of the disease, the increased public demand for studies in patients with suspected OSAS, and the limited access to diagnosis and treatment,\textsuperscript{13} more efficient diagnostic and care measures are required. Portable monitors are a useful diagnostic alternative to polysomnography studies, which are both costly and offered only by a reduced number of sleep clinics. Type-4 monitors, such as those based on single-channel airflow and pulse oximetry, have only been indicated for screening purposes to date.\textsuperscript{1}

The growing tendency to use simplified monitoring devices is not limited to developing countries, however, given that an increasing number of developed countries are also focusing their efforts on finding a way to meet the needs of a growing number of patients at a reduced cost.\textsuperscript{2,14} It is therefore necessary to investigate the diagnostic yield of each and every one of the portable monitors on the market and to analyze the variables that could possibly influence their performance.

High altitude, such as that of Mexico City, is one variable that, at least theoretically, could affect the diagnostic yield of pulse oximetry systems. We did not find this to be the case, however, when we compared oxygen desaturation values obtained by automatic measurement with AHI values obtained by polysomnography. In the circumstances described, the DI provided by the Remmers Sleep Recorder is a reliable indicator of AHI in Mexico City, a finding which coincides with those for other altitudes.\textsuperscript{5} Further studies involving a greater number of patients, however, are necessary to determine whether the high degree of agreement we observed is maintained in patients grouped by disease severity. Our findings revealed that the intraclass correlation coefficient decreased in patients with an AHI of less than 35; this implies that the diagnostic yield of pulse oximetry might vary in different groups of patients.

\begin{table}[ht]
\centering
\begin{tabular}{|c|cccc|}
\hline
\multicolumn{2}{|c|}{DI Cutoff Values} & \multicolumn{3}{c|}{AHI Cutoff Value (Polysomnography)} \\
\hline
\multirow{2}{*}{\text{DI Cutoff Values}} & \multicolumn{1}{c|}{Sensitivity} & \multicolumn{1}{c|}{Specificity} & \multicolumn{1}{c|}{AUC} & \multicolumn{1}{c|}{AUC} \\
\hline
5 & 1 & 0.5 & 0.75 & 0.1 \\
10 & 1 & 0.2 & 0.6 & 0.14 & 0.64 & 0.57 & 0.28 & 0.2 \\
15 & 1 & 0.1 & 0.7 & 0.64 & 0.1 & 0.8 & 0.28 & 0.2 \\
20 & 1 & 0.6 & 0.9 & 0.82 & 0.93 & 0.78 & 0.71 & 0.6 \\
\hline
\end{tabular}
\caption{Diagnostic Yield for Desaturation Index (DI) Cutoff Values Compared to Apnea–Hypopnea (AHI) Cutoff Values (Reference Values) Obtained by Polysomnography*}
\end{table}

*\text{AUC indicates area under the curve.}
The sleep recorder that we used to measure the DI had previously been validated by Vázquez et al. In their study, apnea was defined as the absence of oronasal airflow for at least 10 seconds and hypopnea as a reduction in chest and abdominal movement for at least 10 seconds accompanied by a ≥4% decrease in oxygen saturation, with or without arousal. We used a slightly different definition of both apnea and hypopnea in our study. Obstructive apnea was defined as the complete cessation of or a >50% reduction in airflow lasting longer than 10 seconds, accompanied by persistent respiratory effort, with or without oxygen desaturation or arousal. Hypopnea on the other hand was defined as a <50% reduction in airflow, detected by thermistor or nasal cannula, for longer than 10 seconds, accompanied by oxygen desaturation of ≥4% or an arousal of 1.5 seconds or longer. Despite these minimal differences in scoring criteria, when we compared the AHI obtained by polysomnography and the DI obtained by oximetry, our results were very similar to those of Vázquez et al. Although the present study was designed to evaluate agreement and not diagnostic yield, we calculated the AUCs by using the polysomnography AHI values as reference values and comparing these to a range of DI cutoff values. An AUC of 1 was obtained when a DI cutoff value of 10 was compared to a reference cutoff value of 5. This means that a DI cutoff value of 10 does not generate classification errors when compared to a reference cutoff value of 5 (which has been used to define sleep-disordered breathing). From a clinical perspective, the high rate of true positive and true negative results obtained for a DI cutoff value of over 10 compensates for the relatively low agreement observed between the DI and the AHI in patients with an AHI of less than 35. When we analyzed an AHI cutoff value of 10 (the strictest definition of OSAS), the best AUC was obtained for a DI cutoff point of 15 respiratory events per hour. Although automated analysis algorithms vary from one oximeter to another, generally speaking, the higher the oxygen desaturation level used to define a respiratory event, the greater the specificity (and the lower the sensitivity) of the oximeter. Our results cannot be extrapolated to other oximeters or higher altitudes; this would require a similar validation to the one used in our study.

There was also strong agreement between the DI and the RDI. Both indices were generated automatically by their corresponding algorithms, and the DI was uniformly higher than the RDI by 6.6 events per hour. This difference increased with disease severity (Figure 4). The hypopnea setting used by the Apnea Link airflow sensor to begin measurement of RDI is a decrease in airflow of greater than 50%. This might, at least partly, explain why the RDI was lower than the DI. That is, because the Apnea Link device does not have an oxygen sensor, it is not able to detect certain hypopneas (low sensitivity) with the above setting, a factor which could, theoretically, adversely affect the RDI. Indeed, the study that validated this index against polysomnography was performed at near-sea level, where hypopnea events associated with desaturation might be less frequent given where subjects at this altitude lie on the hemoglobin dissociation curve. It is important to highlight this point because, although airflow (measured ideally by pneumotachography) is the method of choice for detecting apneas and hypopneas, it has been reported that oxygen desaturation (caused by apnea and hypopnea) can also have a range of harmful health effects due to intermittent hypoxemia. In the original validation study, conducted at an altitude of 116 m, the mean difference between the RDI and the AHI obtained by polysomnography was 3.8 (range, 11.1 to –3.5). Sensitivity was 100% and specificity was 87.5% for a cutoff value of 10, which was associated with the largest area under the curve (0.95). In the present study, we did not compare the RDI to the AHI obtained by polysomnography. Nonetheless, it is unlikely that the difference observed between the DI and the RDI (6.6 more events per hour were recorded for DI) (Figure 4) would lead to classification errors in terms of the decision to prescribe continuous positive airway pressure as a difference of just 6.6 events per hour would not influence the physician’s decision in this respect.

In conclusion, we observed a strong correlation between the AHI obtained by polysomnography and the AHI and DI measured by a Remmers sleep recorder in adults with suspected OSAS. We also observed a strong correlation between the DI and the RDI measured using a single-channel airflow detector, although the RDI was lower than the DI by 6 events per hour of recording on average. The altitude of Mexico City did not significantly affect the diagnostic yield of the portable oximeter we analyzed. Further studies are required to evaluate the performance of other monitors in Mexico City and at higher altitudes.

REFERENCES