

SCIENTIFIC INVESTIGATIONS

Validation of a Self-Applied Unattended Monitor for Sleep Disordered Breathing

Indu Ayappa, Ph.D.; Robert G. Norman, Ph.D.; Vijay Seelall, M.D.; David M. Rapoport, M.D.

Division of Pulmonary and Critical Care Medicine, New York University School of Medicine, New York, NY

Study Objectives: To evaluate the validity of the Apnea Risk Evaluation System (ARES™) Unicorder, a self-applied, limited-channel portable monitoring device for the evaluation of sleep disordered breathing (SDB).

Design: Prospective study with blinded analysis.

Setting: Sleep disorder center, academic institution.

Participants: Eighty patients with suspected obstructive sleep apnea hypopnea syndrome (OSAHS) and 22 volunteers.

Interventions: N/A.

Measurements and Results: Subjects used the ARES™ Unicorder at home for 2 nights using only written instructions. Within 2 weeks, they returned to the laboratory for full nocturnal polysomnography (NPSG) with simultaneous monitoring with the Unicorder. NPSGs were scored manually to obtain an apnea-hypopnea index based on Medicare guidelines (AHI4%) and a respiratory disturbance index (RDI). ARES™ studies were autoscored and reviewed to obtain indices based on equivalent definitions i.e., AHI4%^{ARES} and apnea hypopnea (events with 1% desaturation) index (AHI1%^{ARES}). Indices from the NPSG were compared to the in-lab ARES™ and in-home ARES™ indices using mean differences and the intraclass correlations (ICC). For the in-lab comparison, there was high concordance between AHI4%^{NPSG} and AHI4%^{ARES} (ICC = 0.96, mean difference = 0.5/hour) and RDI^{NPSG} and AHI1%^{ARES} (ICC = 0.93, mean difference = 3.2/hour). For NPSG

versus In-Home ARES™ comparison, there was good concordance between AHI4%^{NPSG} and AHI4%^{ARES} (ICC = 0.8, mean difference = 4.1/hour) and RDI^{NPSG} and AHI1%^{ARES} (ICC = 0.8 mean difference = 8.6/hour). The diagnostic sensitivity of in-lab ARES™ for diagnosing SDB using an RDI cut-off of 15 per hour was 95% and specificity was 94%, with a positive likelihood ratio (LR+) = 17.04, and negative likelihood ratio (LR-) = 0.06. For in-home ARES™ data the sensitivity was 85% and specificity 91% (LR+ = 9.34, LR- = 0.17). There was good agreement between the manually scored NPSG SDB indices and the autoscoring ARES™ algorithm.

Conclusions: ARES™ Unicorder provides acceptably accurate estimates of SDB indices compared to conventional laboratory NPSG for both the simultaneous and in-home ARES™ data. The high sensitivity, specificity, and positive and negative likelihood ratios obtained in the group we studied supports the utility of an ambulatory limited-monitoring approach not only for diagnosing sleep disordered breathing but also to rule out SDB in suitably selected groups.

Keywords: Sleep disordered breathing, home monitoring, obstructive sleep apnea/hypopnea syndrome, diagnosis, portable monitoring

Citation: Ayappa I; Norman RG; Seelall V; Rapoport DM. Validation of a self-applied unattended monitor for sleep disordered breathing. *J Clin Sleep Med* 2008;4(1):26-37.

Disclosure Statement

This study was supported by grants from Advanced Brain Monitoring and National Institute of Health NCRR M01RR00096. Dr. Ayappa has received research support from Fisher & Paykel Healthcare and Korosensor. Dr. Norman has received research support from Fisher & Paykel Healthcare. Dr. Seelall has indicated no financial conflicts of interest. Dr. Rapoport has received research support from Fisher & Paykel Healthcare, Genzyme, Guidant, Korosensor, and St. Jude Medical; has performed speaking engagements for Genzyme, Guidant, Resironics, ResMed, and St. Jude Medical; and has consulted for Invacare, Sanofi-Aventis, Boehringer Ingelheim, and Restore Medical. Drs. Ayappa, Norman, and Rapoport all hold US patents and intellectual property rights covering techniques and analysis algorithms for the diagnosis of OSAHS and techniques for administering CPAP.

Submitted for publication May, 2007

Accepted for publication August, 2007

Address correspondence to: Indu Ayappa, PhD, NYU School of Medicine, 462 First Ave, NBV7W54, New York, NY 10016; Tel: (212) 263-6407; Fax: (212) 263-7445; E-mail: indu.ayappa@med.nyu.edu

The high prevalence,^{1,2} morbidity,³ and economic impact^{4,5} of sleep disordered breathing (SDB) has prompted increasing interest in expedited and economical strategies for diagnosis and exclusion of disease. The complexity of a standard nocturnal polysomnogram (NPSG) performed in the laboratory places a large patient burden and economic cost to society and may contribute to limitations of access to care. Prolonged waiting times of up to 10 months⁶ have been reported in the United States, particularly in Medicaid populations (personal observation), and up to several years in other countries⁶ and are likely to increase with recent recommendations for screening of vulnerable populations.^{6,7} This lack of access has been highlighted by the recently published Institute of Medicine report that has recommended “examination of new technologies... especially portable monitoring.”⁵

Recently published studies and a growing use in Europe and Asia have suggested that unattended limited monitoring focusing on respiratory measures may provide sufficient information for the diagnosis and exclusion of sleep SDB, especially if combined with appropriate clinical assessment.⁸⁻¹¹ The Apnea Risk Evaluation System (ARES™) is one such system designed

for unattended limited monitoring that has been shown to have high sensitivity and specificity to the diagnosis of SDB, along with a low failure rate.⁹ Recently, the ARES™ Unicorder was modified to include measurement of airflow using a nasal cannula connected to a pressure transducer, which we have shown to provide important information for the diagnosis of SDB.¹² The Unicorder has several features—including use of the nasal cannula, full disclosure of raw data, the possibility of recording multiple nights of data, and easy self-application—which are desirable in an unattended limited monitor.¹³

The present study reevaluates the ARES™ Unicorder as modified to include the nasal cannula, as this allows the airflow monitoring to meet criteria proposed by the American Academy of Sleep Medicine (AASM) taskforce (e.g., use of a validated measure of airflow and desaturation).^{14,15} The 2 purposes of this study were (1) to compare the respiratory disturbance index (RDI; all SDB events) and the apnea-hypopnea index (AHI; events with a 4% desaturation as required by the Centers for Medicare & Medicaid Services) obtained from conventional in-lab NPSG with similar indices obtained from the ARES™ Unicorder and (2) to evaluate the autoscoring algorithms in the device with and without technician editing.

METHODS

Subjects

All patients presenting to the NYU Sleep Disorders Center for evaluation of sleep complaints suggestive of SDB between April 2005 and August 2006 were eligible to participate in the study. The only 2 exclusion criteria were the inability to read English (the ARES™ Unicorder instructions were only available in English at the time of recruitment) and the inability to wear any device on the forehead. Eighty patients and 22 additional healthy subjects, recruited by word of mouth, were enrolled. The healthy subjects were recruited with the understanding that they had no sleep complaints. However, they were not excluded for snoring or when further questioning suggested mild SDB. The normal volunteers were included in order to amplify the range of SDB breathing, especially at the low end of the spectrum. Because the majority of patients coming to our laboratory have a high likelihood of having obstructive sleep apnea hypopnea syndrome (OSAHS), we used volunteers who had no specific sleep complaints in order to get a large enough number of true negatives (to calculate specificity). All subjects were paid \$20 for participation in the study. A full clinical evaluation was performed on all subjects by a sleep physician, which included a full medical history, detailed snoring history, physical examination, global evaluation for the presence of excessive daytime somnolence, and administration of the Epworth Sleepiness Scale. Demographic data are shown in Table 1.

Protocol

For all subjects, home-based limited monitoring with the ARES™ Unicorder was performed first. The Unicorder and its associated questionnaire was either given to the subjects at the time of their initial visit to the NYU sleep laboratory or mailed to the subject's home. Standard *written* instructions were pro-

Table 1—Demographic Data From 97 Subjects Who had Either Home or Lab Study with ARES™

	Patients (n = 77)			Volunteers (n = 20)		
	Mean	Min	Max	Mean	Min	Max
Gender, number						
Men	60			9		
Women	17			11		
Age, y	46	26	74	36	19	73
BMI, kg/m ²	30	21	70	24	19	32
Severity of SDB						
RDI range, no./h	0-5	6-10	11-20	21-40	40-104	
Subjects, number	13	14	20	23	27	
Mean ESS score	6 ± 4	8 ± 6	9 ± 4	9.5 ± 6	10 ± 5	

Data are expressed as mean, minimum and maximum, except for gender, which is number in each group. BMI refers to body mass index; ESS, Epworth Sleepiness Scale.

vided for application of the Unicorder, and no additional in-person training was provided beyond these. Randomization of the order of home unattended and laboratory use of the ARES™ Unicorder was *not* performed, as we wished to capture success rate of the home use of this device in naïve subjects. The written instructions were for subjects to use the device for 2 full nights at home and return it for download of the stored physiologic signals in full disclosure format.

Within 2 weeks of the home study, subjects returned to the sleep laboratory for a standard diagnostic NPSG, performed by a trained sleep technician according to the usual clinical protocol for the NYU Sleep Laboratory. A new Unicorder recording was initiated simultaneously with the laboratory NPSG by the sleep technician.

ARES™ UNICORDER

The ARES™ consists of the Unicorder device, a self-administered questionnaire, and off-line analysis software.⁹ The data from the questionnaire were not used in the present analyses, as the focus was on the relationship of the SDB indices obtained in an unattended setting to those obtained in the laboratory. Automated analysis on the downloaded physiologic data from the Unicorder is performed with custom software (ARES™ Insight) to obtain SDB indices. The software allows all signals and events scored by auto analysis to be visualized and edited by a technician before tabulation.

The ARES™ Unicorder is worn on the forehead and does not require additional wires to external devices such as an oximeter probe. It measures oxygen saturation (SpO₂) and pulse rate from reflectance oximetry, air flow from a nasal cannula and pressure transducer, snoring levels via a calibrated acoustic microphone, and head movement actigraphy and head position from accelerometers. The device also provides audible alerts during the study if poor-quality airflow or SpO₂ is detected so that the subject can reposition the device.

LABORATORY NPSG

Each NPSG consisted of a full night spent in the NYU Sleep Disorders Center. Recordings of frontal, central, and occipital

electroencephalogram, electrooculogram, and submental electromyogram were used to monitor sleep. Leg movements were monitored with an anterior tibialis electromyogram. A unipolar electrocardiogram was used for cardiac monitoring. Oxygen saturation was monitored with a pulse oximeter. A nasal cannula/pressure transducer system (Pro-Tech PTAF2, Mukilteo, WA) was used to measure airflow, along with an oral thermistor to detect mouth breathing. A single nasal cannula was placed on the patient, and the tubing was split to provide pressure inputs to both the Pro-Tech pressure transducer of the NPSG system and the ARES™ Unicorder. Preliminary testing showed that this setup did not affect the pressure tracing obtained in either device. Snoring was monitored using the Pro-Tech PTAF2 system. Chest wall and abdominal movement were monitored with piezoelectric strain gauges (EPMSystems, Inc, Midlothian, VA.) Subjects underwent either a full-night NPSG or a split-night NPSG (diagnostic study and continuous positive airway pressure titration) when severe SDB was observed, as per the usual clinical protocol.

Analysis of Signals

ARES™ UNICORDER

Automated event analysis was performed using custom (ARES™ Insight) software. SDB events were identified as follows. Apnea was cessation of airflow for 10 seconds. Hypopnea4% included events identified as a hypopnea (airflow < 50%) with a 3.5 or greater percentage desaturation and 1% resaturation. In addition Hypopnea1% events were determined as those hypopneas with a minimum 1% desaturation and resaturation and at least 1 surrogate arousal indicator (head movement, changes in snoring, or changes in pulse rate); this definition is similar to the AASM inclusive criteria for all SDB events. In order to calculate the valid recording time (denominator for SDB indices), the ARES™ automatically excludes (a) the first 15 minutes and last 3 minutes of the study, (b) periods in the record with poor or bad airflow or oxygen saturation (SpO₂) quality, (c) periods with excessive head movement, (d) any period the patient was upright, and/or (e) the first 30-seconds after each change of head position.

Each automatically scored ARES™ study was subjected to technical review by visual inspection of the signals. In this manual review, we allowed the technician only limited editing options: (a) change to an earlier start time based on indicators suggesting sleep onset (i.e., absence of head movements or appearance of apnea, hypopnea, flow limitation, or snoring), (b) override the exclusion of periods identified by the automated analysis as invalid (this typically occurred when airflow or SpO₂ quality was poor but obstructive events could be readily distinguished using all available signals), and (c) exclude additional periods of unacceptable SpO₂ quality not detected by the auto analysis. Finally, the technician was allowed to edit the number of events by excluding automatically detected events but the technician did *not* insert new events undetected by the automated algorithms. This was based on prior experience that the auto analysis rarely misses events but may overscore in some instances. The duration of the technical review averaged 5 minutes per record and was performed at Advanced Brain Monitoring without knowledge of other patient data or NPSG results.

NOCTURNAL POLYSOMNOGRAMS

All scoring of the NPSGs was performed by a single trained sleep technician at the NYU Sleep Disorders Center who was blinded to the results of the Unicorder studies and their scoring. Sleep was manually scored on 30-second epochs using the standard criteria of Rechtschaffen and Kales.¹⁶ Brief arousals from sleep were manually scored using the AASM criteria.¹⁷ For the full NPSG, the total sleep time scored by Rechtschaffen and Kales criteria was used as the denominator to calculate the SDB indices.

Respiratory signals were scored manually for the presence of SDB events as follows. *Respiratory events* longer than 10 seconds were scored while viewing all channels on the full NPSG. *Apneas* were identified when the airflow amplitude on the nasal cannula was less than 10% of baseline. *Hypopneas*4% (Medicare definition) were identified when airflow amplitude was reduced by 30% from baseline and the event was followed by a 4% or greater oxygen desaturation.¹⁷ An additional definition of hypopnea was used for the calculation of the RDI_{NPSG}, in which events were identified when airflow amplitude was less than 50% of baseline or, alternatively, whenever a discernable change occurred in the airflow amplitude (generally amplitude was between 50% and 80% of the baseline) and the event was followed by a 4% oxygen desaturation¹⁷ within 30 seconds. Respiratory effort-related arousals were identified whenever a discernable change was seen in the airflow amplitude and electroencephalographic arousal occurred within 5 seconds of its end.¹⁸ Thermistor and the chest wall/abdominal movement detectors were only used when the nasal cannula signal was uninterpretable, which occurred only rarely in these studies (< 1% of time) because technicians intervened immediately when the signal failed. A single rater was used to minimize interrater variability.

Tabulation

For the ARES™ studies (home and in-lab), SDB indices were calculated from the technician-edited data.

$$AHI4\%_{\text{ARES}} = \frac{\# \text{Apneas} + \text{Hypopneas}4\% \text{ on Unicorder}}{\text{Valid Time}}$$

$$AHI1\%_{\text{ARES}} = \frac{\# \text{Apneas} + \text{Hypopneas}1\% \text{ on Unicorder}}{\text{Valid Time}}$$

For ARES™ home studies, the number of events and valid time refer to cumulative data from 2 nights of recording. The AHI1% was selected as an analog to the RDI inclusive of respiratory effort-related arousals recommended by the AASM task force.^{14,15}

For the NPSG, the following indices were calculated.

$$AHI4\%_{\text{NPSG}} = \frac{\# \text{Apneas} + \text{Hypopneas}4\% \text{ on Full NPSG}}{\text{Total Sleep Time}}$$

$$RDI_{\text{NPSG}} = \frac{\# \text{Apneas} + \text{Hypopneas} + \text{RERAs on Full NPSG}}{\text{Total Sleep Time}}$$

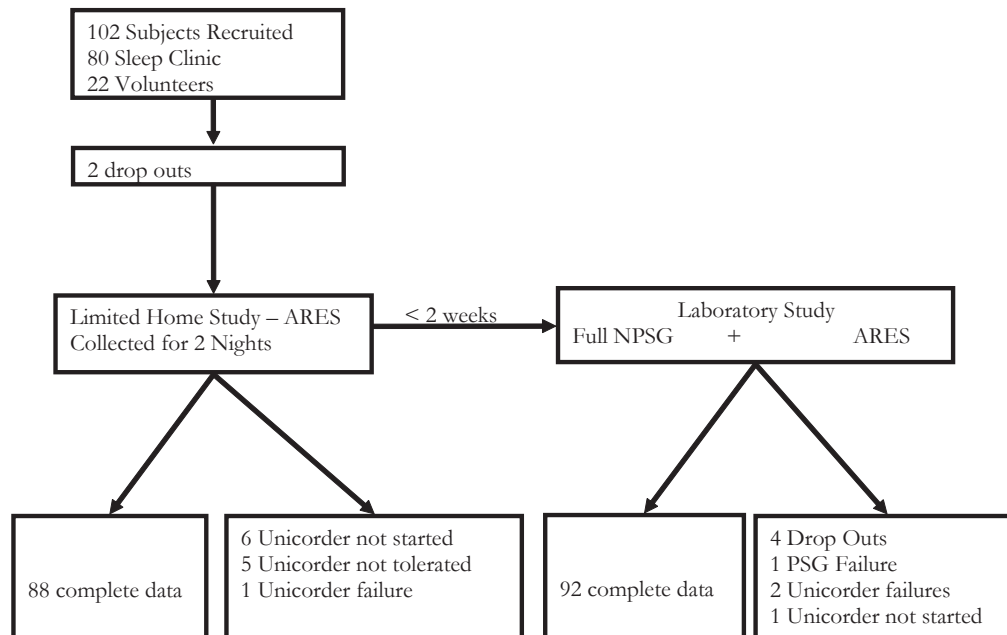


Figure 1—Flow chart of the study design.

The following comparisons were made for the indices described above: (1) simultaneous ARES™ (attended) versus lab NPSG and (2) home ARES™ (unattended) versus lab NPSG, and (3) ARES™ automated analysis versus technician-edited data for all ARES™ lab and home data.

Statistics

Variability in the ARES™ indices was evaluated against the gold standard of full NPSG by calculating the mean bias, 95% confidence intervals (CI) and limits of agreement (± 2 SD) and shown using Bland-Altman plots. Agreement was measured with a mixed-effects intraclass correlation coefficient (ICC). Typical clinical cut-offs of 5, 10, and 15 events per hour for the AHI4% and 10 and 15 events per hour for the RDI/AHI1% were used to calculate the diagnostic sensitivity, specificity, 95% confidence intervals, and positive and negative likelihood ratios.

The study was approved by the institutional review board of New York University School of Medicine, the NYU General Clinical Research Center, and the Health and Hospitals Corporation of New York City, and all subjects signed informed consent prior to entering the study. This study was partially funded by a grant from Advanced Brain Monitoring, and all automated and initial technical review of Unicorder studies was performed by Advanced Brain Monitoring staff, as per the clinical model advocated by Advanced Brain Monitoring for use of their device. However, Advanced Brain Monitoring played no role in patient selection, NPSG evaluation, data collation, or final data analysis and interpretation, which were performed entirely by NYU Investigators.

RESULTS

A total of 102 subjects (80 patients and 22 volunteers) were enrolled in the study. Two subjects dropped out prior to study

initiation due to scheduling issues. In addition, 6 subjects did not initiate home monitoring after being given the ARES™ Unicorder to take home, citing “lack of time” as the reason. Five subjects reported that they were unable to tolerate the Unicorder for more than 1 hour. Finally a technical failure prevented data collection in 1 subject. Thus, there were 88 successfully collected home recordings with the Unicorder.

Of the 102 subjects, 96 returned to the NYU Sleep laboratory (within 2 weeks of the home Unicorder study) for the in-laboratory NPSG. In this group, there was 1 NPSG failure (nasal cannula signal unusable), 1 study in which the Unicorder was not started, and 2 studies with Unicorder technical failures in the laboratory. Of note, 2 of the 5 subjects who could not tolerate the Unicorder in the home had no difficulty using it in the laboratory. Thus, only 3% (3/93) of subjects had trouble tolerating the Unicorder once recordings were initiated. Five subjects called the technical phone support line, which is provided by Advance Brain Monitoring and available throughout the night, prompted by the audible prompts. In 2 of these, the subject resolved the problem by repositioning the Unicorder, and, in 3, a new Unicorder was mailed to the patient. Figure 1 summarizes the results of data collection.

In general, signal loss was not a problem once a study was initiated. Only 4 patients had less than 3 hours of valid recording time in the home over 2 nights. Although subjects tended to use the device on average 45 minutes less the second night, only 6 subjects used it for only 1 night.

The above data collection resulted in 86 subjects who had complete data in *both* ARES™ home and laboratory NPSG, whereas 92 subjects had complete data in the simultaneous ARES™ and laboratory NPSG.

Simultaneous NPSG Versus ARES™ Unicorder

The total sleep time on the 92 NPSGs (14 were split-night studies) was 309 ± 115 minutes (mean $\pm$ SD), and the valid

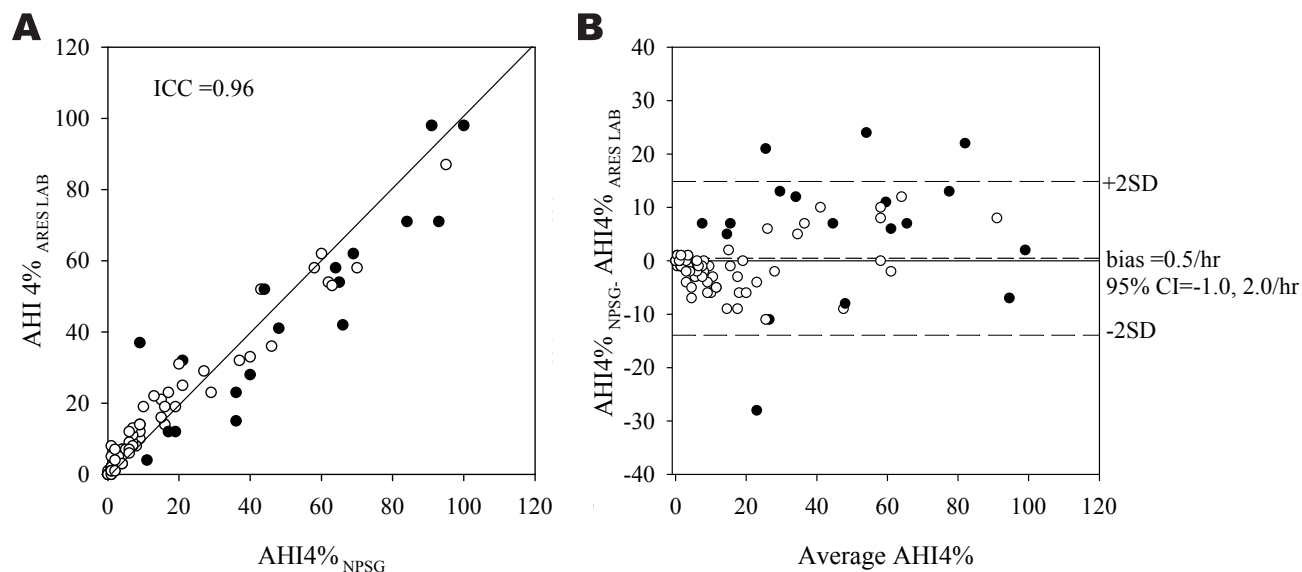


Figure 2-A—Relationship between apnea-hypopnea index $AHI4\%_{NPSG}$ and $AHI4\%_{ARES LAB}$. Each point represents 1 subject. The closed circles represent studies with less than 3 hours of recording time. Fourteen study pairs had less than 3 hours of recording time on both nocturnal polysomnography (NPSG) and Apnea Risk Evaluation System (ARESTTM) studies, and 4 had less than 3 hours in only the ARESTTM study. **B:** Bland Altman plot showing the difference between the AHI4% (NPSG-ARES lab) plotted against the average of the 2 AHI4%s. The greatest differences are in the subjects with short recording durations. 95%CI refers to 95% confidence interval.

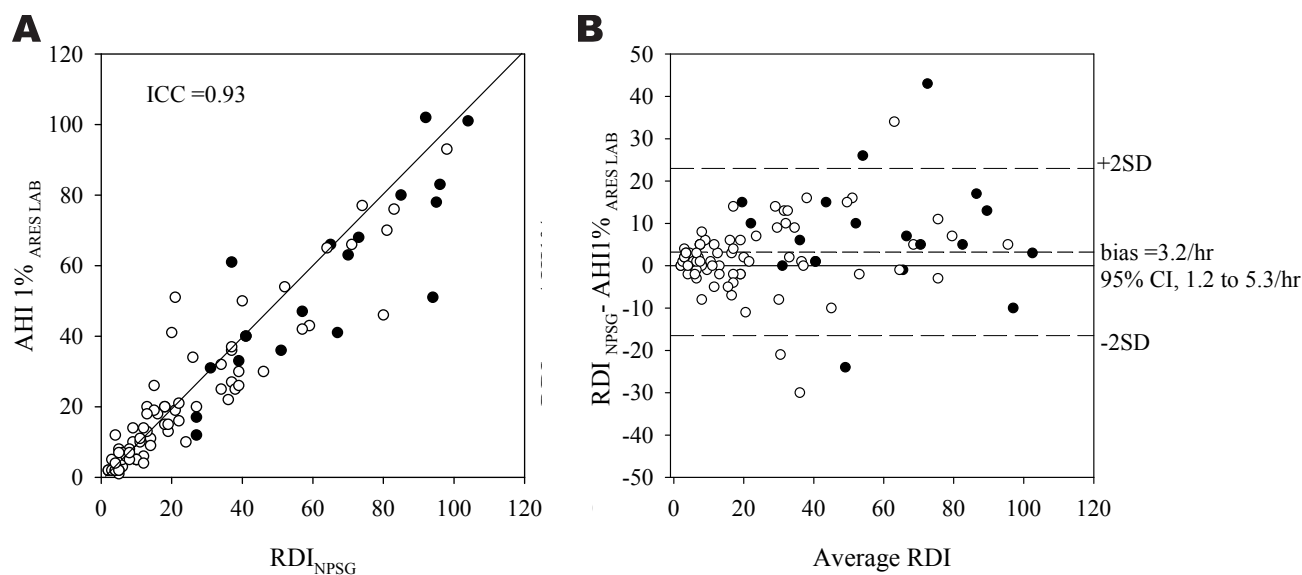


Figure 3-A—Relationship between the respiratory disturbance index derived from the nocturnal polysomnogram (RDI_{NPSG}) conducted in the lab and the apnea-hypopnea index based on the Apnea Risk Evaluation System (ARESTTM) studies conducted in the lab ($AHI1\%_{ARES LAB}$). Each point represents 1 subject. As in Figure 2, the closed circles represent studies with less than 3 hours of recording time. **B:** Bland Altman plot showing the difference between the RDI_{NPSG} and $AHI1\%_{ARES LAB}$ plotted against the average of the 2. 95%CI refers to 95% confidence interval.

time on ARESTTM was 309 ± 103 minutes. Figure 2 compares the AHI4% obtained from the NPSG with the AHI4% obtained from the technician-edited simultaneous ARESTTM study. There was no significant bias between the NPSG and ARESTTM study (mean ($AHI4\%_{NPSG} - AHI4\%_{ARES}$) = 0.5/h, 95% CI: -1.0 to 2.0/h). When the data from the volunteers were excluded from the analyses, the results for the 73 patients were similar to that of the entire group (ICC = 0.95, bias = 0.7/h, 95% CI: -1.2 to 2.6/h). Table 2 shows the sensitivity, specificity, 95% confidence intervals, and positive and negative likelihood ratios of the ARESTTM AHI4% for diagnosing SDB, as defined by multiple cut-offs for the in-lab NPSG.

The 4 subjects with the largest difference ($> 14.4/h = 2 SD$) between NPSG and ARESTTM AHI4% all had studies lasting less than 3 hours (3/4 due to split-night protocol). The mean bias did not change when all studies with less than 3 hours of recorded data were excluded and was -0.8 per hour (95% CI: -1.8 to 0.2/h, ICC = 0.98). However, the limits of agreement were much narrower ($\pm 2SD$: -9.2 to 7.6/h). The excluded studies consisted of 14 studies with split-night protocol in which both ARESTTM and NPSG were of short duration and 4 studies in which only the ARESTTM collection was short.

Comparisons of the RDI and AHI1% obtained from the simultaneous NPSG and ARESTTM were performed in a similar

Table 2a—Diagnostic Classification: Lab-Based ARES™ Compared with Lab-Based Nocturnal Polysomnography in All 92 Subjects

	Prevalence	Sensitivity	95% CI		Specificity	95% CI		LR+	LR-
			LL	UL		LL	UL		
AHI4%, no./h									
≥ 5	0.60	0.98	0.89	1.00	0.84	0.67	0.93	6.05	0.02
≥ 10	0.42	0.97	0.85	1.00	0.85	0.72	0.93	6.46	0.03
≥ 15	0.39	0.92	0.76	0.98	0.95	0.84	0.99	17.11	0.09
RDI, no./h									
> 10	0.75	0.91	0.81	0.96	0.87	0.65	0.97	7.00	0.10
≥ 15	0.61	0.95	0.84	0.99	0.94	0.80	0.99	17.04	0.06

Table 2b—Diagnostic Classification: Lab-Based ARES™ Compared with Lab-Based Nocturnal Polysomnography in 73 Subjects, Which Excludes Volunteers

	Prevalence	Sensitivity	95% CI		Specificity	95% CI		LR+	LR-
			LL	UL		LL	UL		
AHI4%, no./h									
≥ 5	0.71	0.98	0.88	1.00	0.76	0.52	0.91	4.12	0.03
≥ 10	0.51	0.97	0.84	1.00	0.78	0.60	0.89	4.38	0.03
≥ 15	0.47	0.91	0.75	0.98	0.92	0.78	0.98	11.85	0.10
RDI, no./h									
> 10	0.85	0.95	0.86	0.99	0.73	0.39	0.93	3.49	0.07
≥ 15	0.74	0.94	0.84	0.99	0.89	0.65	0.98	8.97	0.06

CI refers to confidence interval; LL, lower limit; UL, upper limit; LR+, positive likelihood ratio; LR-, negative likelihood ratio; AHI, apnea-hypopnea index; RDI, respiratory disturbance index.

manner to that used for the AHI4% and are presented in Figure 3 and Table 2. For the subset of patients alone (excluding volunteers), the results were similar to those of the entire group (ICC = 0.92, bias = 3.3/h, 95% CI: 0.8 to 5.9/h).

Lab NPSG Versus Home ARES™

The mean valid time from the home 86 ARES™ studies was 594 ± 171 minutes (total of 2 nights). Figure 4A-B compares the AHI4% obtained from the NPSG with the AHI4% obtained from the technician-edited home ARES™ study. The overall bias between the NPSG and ARES™ study was greater than the simultaneously acquired data (mean $[AHI4\%_{NPSG} - AHI4\%_{ARES}] = 4.1/h$, 95% CI: 0.8 to 7.3/h). Figure 4C-D compares the data after excluding the 14 split-night studies conducted in the laboratory. The mean bias was lower (1.8/h) and the 95% CI was -1.2 to 4.7 per hour. Tables 3a-b show the sensitivity, specificity, 95% confidence intervals, and positive and negative likelihood ratios of using the home ARES™ AHI4% for diagnosing SDB when SDB is defined by multiple cut-offs for the in-lab NPSG.

Comparisons of the RDI and AHI1% obtained from the laboratory NPSG and home ARES™ were performed in a manner similar to the comparison done for the AHI4% and are presented in Figure 5 and Tables 3a-b.

When the data from the volunteers were excluded from the analyses, the mean differences were slightly higher in the 67 patients, compared with the entire group. Comparing $AHI4\%_{NPSG} - AHI4\%_{ARESHome}$, the ICC was 0.76, the bias was 5.2 per hour, and the 95% CI was 1.0 to 9.4 per hour, and comparing $RDI_{NPSG} - AHI1\%_{ARESHome}$, the ICC was 0.76, the bias was 10.3 per hour, and the 95% CI was 5.9 per hour to 14.6 per hour. The sensitivity, specificity, and likelihood ratio data are presented in Table 3c.

ARES™ Lab Versus ARES™ Home

AHI4% obtained in the lab was higher than in the home (19 vs 16/h, $p = 0.05$). The AHI1% obtained in the lab was higher than in the home (27 versus 23/h, $p = 0.01$). In the laboratory study, patients spent a significantly greater percentage of the recording in the supine position (60% vs 46%, $p < 0.01$).

ARES™ Autoscoring with and without Technician Review

Overall, the effect of technician editing, as described in the Methods, was minimal. Figure 6 compares the autoscored AHI1% (RDI equivalent) before and after technician review, in the 181 ARES™ studies, including both those done in the home and in the lab. With respect to diagnostic accuracy (using an RDI of 15 or more per hour in the NPSG studies as the gold standard) the auto-scoring without technician review resulted in a sensitivity of 95% and a specificity of 86% for simultaneously acquired data. Technical review resulted in little change in the diagnostic accuracy (see Table 2). The largest change resulting from the type of technician review performed in this study was a change in sections marked as valid and considered for analysis (see Figure 7), but this did not have a net impact on the RDI (see Figure 6). Overall, 11 lab ARES™ studies and 21 home ARES™ had a longer than 30-minute change in valid recording time that resulted from technical review. In those studies in which review resulted in an increase in valid recording time, this was generally because auto-excluded periods with questionable signal quality were re-included by the manual review because the other signals provided sufficient evidence of valid events.

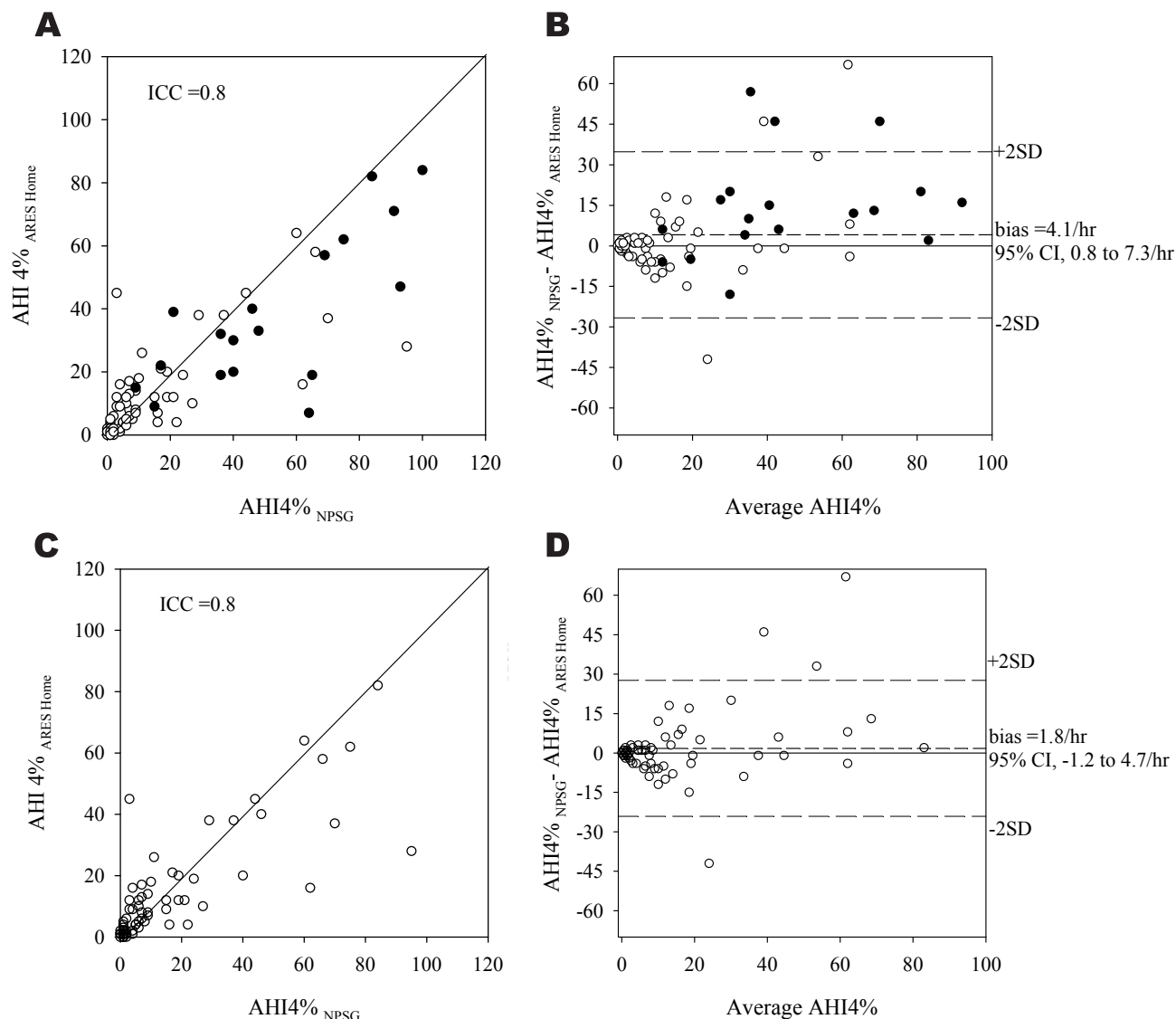


Figure 4—A—Relationship between apnea-hypopnea index (AHI) derived from the nocturnal polysomnogram ($AHI4\%_{NPSG}$) conducted in the lab and the AHI based on the Apnea Risk Evaluation System (ARES™) studies conducted in the home. Each point represents 1 subject. As in figure 2, the closed circles represent studies with less than 3 hours of recording time. **B**: Bland Altman plot showing the difference between the $AHI4\%$ (nocturnal polysomnogram-ARES Home) plotted against the average of the 2 $AHI4\%$ s. Panels **C** and **D**: Relationship between $AHI4\%_{NPSG}$ and $AHI4\%_{ARES\ Home}$ excluding data from the split-night studies. 95%CI refers to 95% confidence interval.

DISCUSSION

The present study shows that SDB indices similar to those recommended by the AASM Task Force and required by the Center for Medicare and Medicaid Services can be calculated from the signals acquired with the ARES™ Unicorder and that these indices of SDB correlate with those acquired during conventional laboratory NPSG. This was true both for the simultaneously acquired ARES™ data and when ARES™ studies acquired in the home were compared with the laboratory NPSG. Of note, patients used the Unicorder in the home with no technician input and written instructions only. Furthermore, there was good agreement between the SDB indices obtained from the manually scored NPSG and the autoscoring algorithm used in the ARES™ system.

Based on the likelihood ratios, the indices obtained using the ARES™ Unicorder yielded a “large increase” ($LR+ = 17.04$) in disease probability based on an RDI of greater than 15 per

hour and a “large reduction” ($LR- = 0.06$) in disease probability using an RDI of less than 15 per hour.¹⁹ For the home study on a separate night, the likelihood ratios are slightly lower (9.34 for $LR+$ and 0.17 for $LR-$), though this is likely due to night-to-night variability.

Normal volunteer subjects (recruited by word of mouth) were included in the present study in order to amplify the range of SDB, especially at the low end of the spectrum. The majority of patients coming to our laboratory have a high likelihood of being diagnosed with OSAHS. In order to obtain a large enough number of true negatives (to calculate a meaningful value for specificity) and also to have a wide enough range of AHI and RDI values at the low end of the spectrum, we used volunteers who had no specific sleep complaints. Using the dataset of patients alone, 19 of 73 (25%) had an RDI less than 15 per hour. With the inclusion of the volunteers, 36 of 92 (40%) had an RDI less than 15 per hour. There was no change in performance of the device when the data from the volunteers were excluded from the analysis.

Table 3a—Diagnostic Classification: Home-Based ARES™ Compared with Lab-Based Nocturnal Polysomnography in all 86 Subjects

	Prevalence	Sensitivity	95% CI		Specificity	95% CI		LR+	LR-
			LL	UL		LL	UL		
AHI4%, no./h									
≥ 5	0.60	0.90	0.78	0.96	0.79	0.62	0.91	4.39	0.12
≥ 10	0.43	0.86	0.70	0.95	0.82	0.67	0.91	4.71	0.17
≥ 15	0.41	0.74	0.56	0.87	0.88	0.75	0.95	6.31	0.29
RDI, no./h									
>10	0.76	0.86	0.75	0.93	0.90	0.68	0.98	9.05	0.15
≥15	0.62	0.85	0.72	0.93	0.91	0.75	0.98	9.34	0.17

Table 3b—Diagnostic Classification: Home-Based ARES™ Compared with Lab-Based Nocturnal Polysomnography in 72 Subjects, which Excludes Split-Night Studies

	Prevalence	Sensitivity	95% CI		Specificity	95% CI		LR+	LR-
			LL	UL		LL	UL		
AHI4%, no./h									
≥ 5	0.53	0.87	0.71	0.95	0.79	0.62	0.91	4.22	0.17
≥ 10	0.33	0.88	0.66	0.97	0.83	0.69	0.92	5.25	0.15
≥ 15	0.31	0.68	0.45	0.85	0.90	0.77	0.96	6.82	0.35
RDI, no./h									
> 10	0.71	0.82	0.69	0.91	0.90	0.68	0.98	8.65	0.20
≥ 15	0.54	0.82	0.66	0.92	0.91	0.75	0.98	9.03	0.20

Table 3c—Diagnostic Classification: Home-Based ARES™ Compared with Lab-Based Nocturnal Polysomnography 67 Patients, which Excludes Volunteers

	Prevalence	Sensitivity	95% CI		Specificity	95% CI		LR+	LR-
			LL	UL		LL	UL		
AHI4%, no./h									
≥ 5	0.73	0.92	0.80	0.97	0.67	0.41	0.86	2.76	0.12
≥ 10	0.52	0.89	0.72	0.96	0.72	0.53	0.86	3.15	0.16
≥ 15	0.49	0.76	0.57	0.88	0.82	0.65	0.93	4.29	0.29
RDI, no./h									
>10	0.87	0.90	0.78	0.96	0.78	0.40	0.96	4.03	0.13
≥15	0.75	0.84	0.71	0.93	0.81	0.54	0.95	4.50	0.19

CI: Confidence Interval, LL: Lower Limit, UL:Upper Limit, LR+: Positive Likelihood Ratio, LR-:Negative Likelihood Ratio; ; AHI, apnea-hypopnea index; RDI, respiratory disturbance index.

By definition, any limited monitoring provides a denominator of the SDB indices that is not the same as total sleep time but, rather, some fraction of the total time of recording. Nevertheless, our data show that, if the target is the Medicare AHI4%, the ARES™ Unicorder provides an acceptably accurate estimate of the NPSG index. In contrast, if the purpose of a study is screening, then an index of SDB that detects subtle disease (e.g., the RDI) may be more appropriate than the AHI4%. The present study shows that the ARES™ Unicorder also appropriately provided such an index, thus making it a useful tool to *rule out* significant SDB in the clinical situations in which this device is likely to be used.

In the present data set, the ambulatory (home) recording with the ARES™ monitor was always performed *first*, followed by the in-lab study within 2 weeks rather than in random order. This design was specifically chosen so that the home recording would not be influenced by any bias (experience) as a result of patients knowing their diagnosis or having acquired information or experience

as to application of sensors (especially from the simultaneous in-lab use of the ARES™ monitor) during the in-lab study. This allowed us to evaluate the ease of use and study failure rate of the self-applied ambulatory device in the setting in which it is intended to be used. We feel that it is unlikely that lack of randomization in the order of the home versus lab tests introduced a systematic bias in the diagnostic measures, as we know of no evidence that the performance of a sleep study has any effect on SDB measures obtained in subsequent studies. If anything, the second study tends to be easier for many subjects, as there is a tendency for a “first-night” effect to interfere with sleep.

The failure rate of the ARES Unicorder based on the number of successful studies in the home was 12% (12/100), as opposed to a failure rate of 1% for the NPSG. Six of the failed home studies were in patients who did not start the home study (citing lack of time) despite having agreed to the protocol. This may have been an unintended consequence of always following the home study with a laboratory study: because all patients were

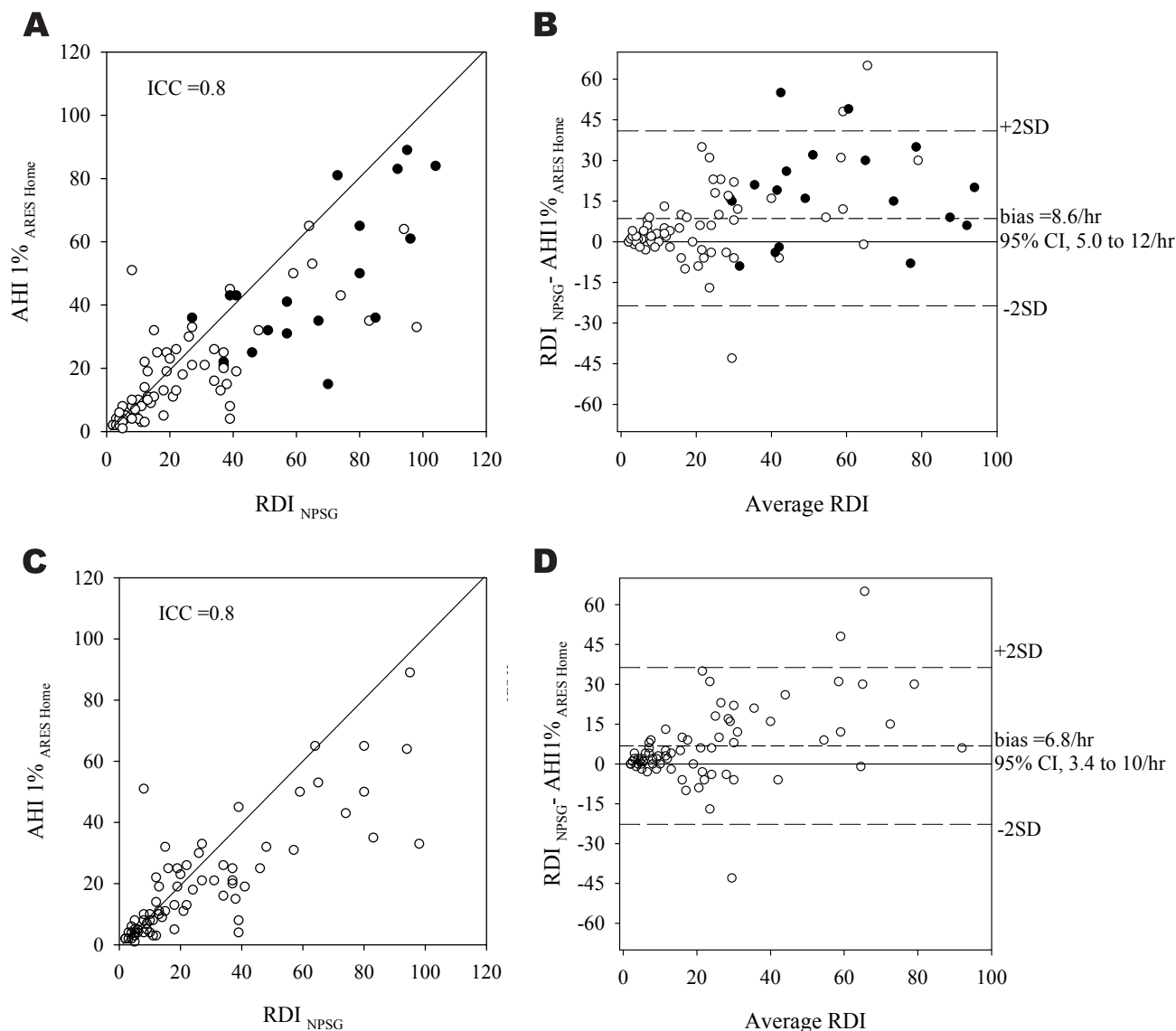


Figure 5—A—Relationship between the respiratory disturbance index derived from the nocturnal polysomnogram (RDI_{NPSG}) conducted in the lab and the apnea-hypopnea index (AHI) based on the Apnea Risk Evaluation System (ARES™) studies conducted in the home (AHI1%_{ARES Home}). Each point represents 1 subject. As in figure 2, the closed circles represent studies with less than 3 hours of recording time. **B**—Bland Altman plot showing the difference between the RDI_{NPSG} and AHI1%_{ARES Home} plotted against the average of the 2. **5-C** and **5-D**—Relationship between RDI_{NPSG} and AHI1%_{ARES Home} excluding data from the split-night studies. 95%CI refers to 95% confidence interval.

aware that an in-lab NPSG was scheduled, there may have been less incentive for them to perform the home monitoring than would have been the case if this were the only scheduled evaluation. From the above, the failure rate of the ARES Unicorder™ when utilized by a patient was 6%.

When ARES™ and NPSG data were acquired simultaneously, only 4 of 92 studies showed a difference in AHI4% greater than 14 per hour (2 SD). Factors to which differences in simultaneously acquired SDB indices can be attributed include the following:

(i) Valid signal time and duration of sleep: In the absence of electroencephalogram-defined sleep, the denominator of the AHI and RDI during limited studies is not the total sleep time, and this will always remain a limitation of not electroencephalographically monitoring sleep. In the analysis of the data collected using the Unicorder, the denominator is obtained from the total recording time by subtracting the time during which surrogates of arousal are detected, including prolonged large

head movements and upright position.⁹ In addition, periods of poor signal (airflow and saturation) are removed from both the numerator and denominator. The effect of changing the denominator used in the SDB will necessarily be greatest in short studies. In fact, our data suggest that the largest differences in AHI4% between the NPSG and the ARES™ were in subjects who had less than 3 hours of recording time and, in 1 subject, when the majority of the (short) recording time was spent in wake. When the overall duration of the study was short (which in 14 subjects was due the use of a split-night protocol in the laboratory), small changes in denominator and/or numbers of events have the greatest potential to have an impact on the index and appear to have contributed to differences between ARES™ and NPSG indices. From a practical point of view, this problem is likely to be minimized if short periods of unattended recordings are avoided (e.g., only collect data for full nights or even 2 nights). In fact, our data suggest that limiting analysis to ARES™ studies with durations

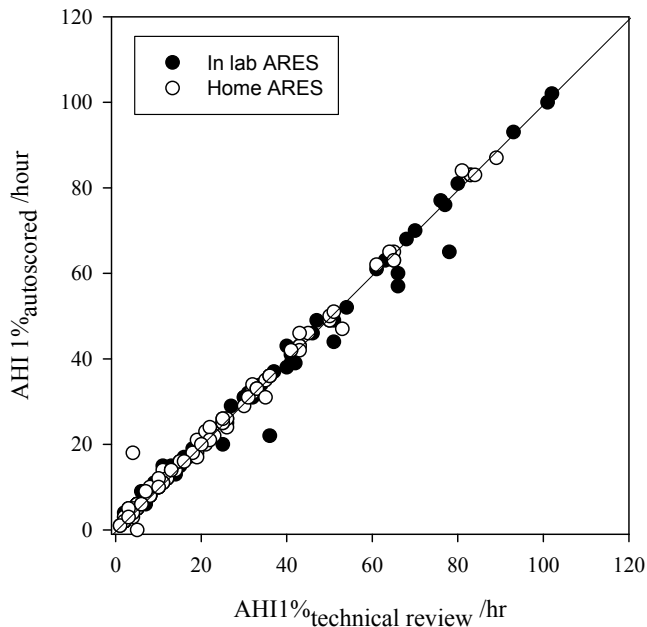


Figure 6—Relationship between the apnea-hypopnea index (AHI1%) autoscored with and without technical review for all Apnea Risk Evaluation System (ARES™) studies. The closed circles represent in-lab ARES™ studies, and the open circles represent in-home ARES™ studies.

of 3 hours or longer may provide the most accurate estimate of the NPSG AHI4%.

(ii) Difference in the way the oximeter signal is acquired and processed and used in defining events: We have previously shown that changing the oximeter used can significantly affect SDB indices in attended NPSG studies.²⁰ In the present study, some of the differences between the simultaneous ARES™ and

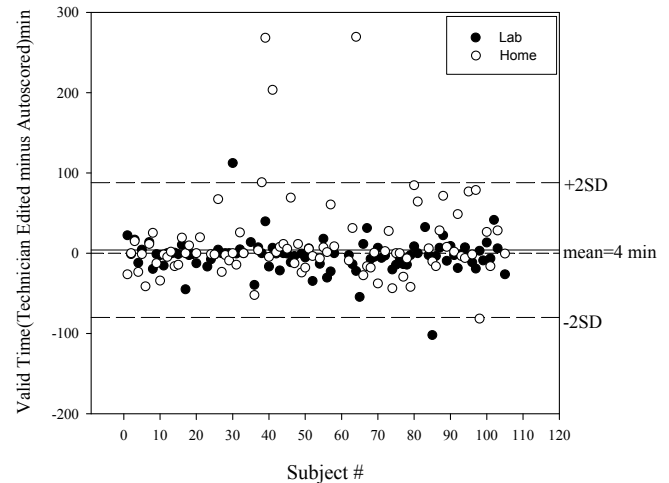


Figure 7—Change in valid time as a result of technical review plotted by subject number for all Apnea Risk Evaluation System (ARES™) studies. Closed circles represent in-lab ARES™ studies, and open circles represent in-home studies.

NPSG indices could be attributed to differences between the ARES™ reflectance oximeter and the Masimo transmission oximeter (Masimo Radical, Irvine, CA) used in our laboratory NPSG. There were also differences in the scoring algorithms utilized, and these may have contributed to differences in the SDB indices independent of the oximetry-signal acquisition. An example of this is that the ARES™ oximeter reports saturation with a resolution of 0.1%, in contrast with the lab NPSG oximeter, which reports saturation with a 1% resolution. Thus, the NPSG oximeter is subject to different rounding of values near 4%, and the ARES™ algorithm allows a cut-off of 3.5% for hypopnea.

(iii) Effects of the automated analysis used in the ARES™

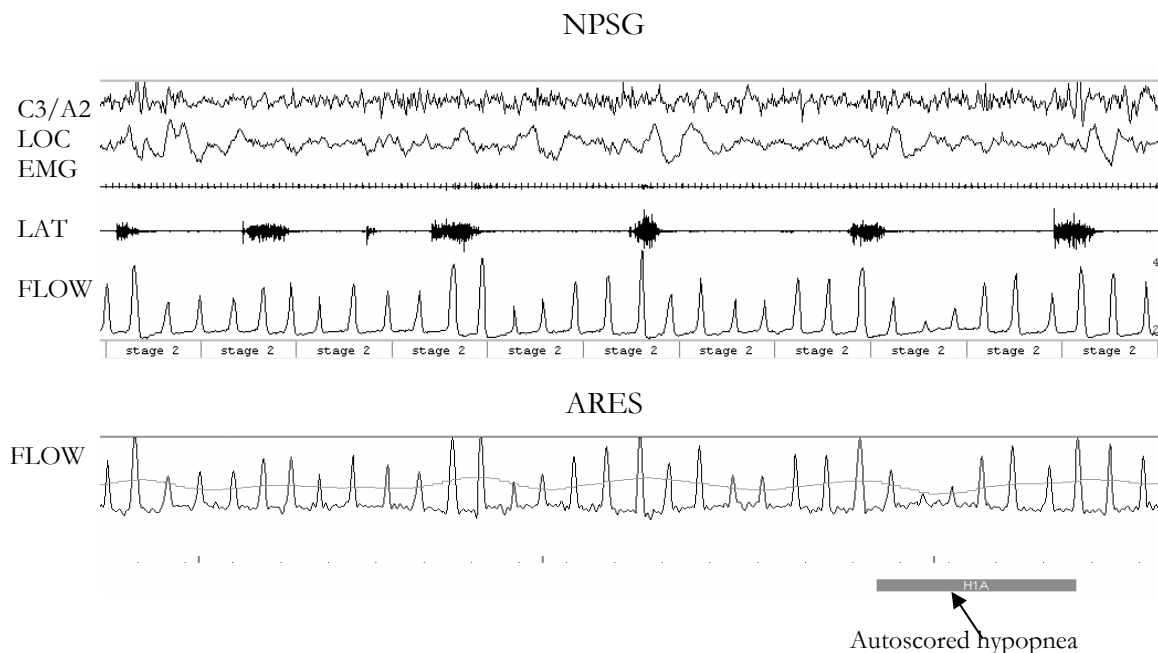


Figure 8—This figure illustrates the impact of periodic leg movements (PLMs) on the sleep-disordered breathing (SDB) index in a subject in whom the RDI_{NPSG} was lower than the $AHI1\%_{ARES\text{ LAB}}$ by 21/h. The top panel shows the signals from the nocturnal polysomnogram (NPSG), in which no SDB events are scored on the airflow channel by the technician due to the presence of clearly identifiable PLMs. The lower panel shows the Apnea Risk Evaluation System (ARES™) data collected simultaneously and the autoscored hypopnea.

versus manual scoring in the NPSG: Although autoscoring saves time and increases reproducibility of scoring, it is a source of deviation from manual scoring and needs to be validated. As is customary in clinical practice, all scored NPSGs were reviewed manually after initial scoring, including the autoscored ARES™ studies. As shown by our data, this did not produce large changes in the overall SDB indices produced by the automated ARES™ scoring algorithm. Although not directly assessed, this suggests that manual scoring of the ARES™ studies would have been similar to the results from the automated and manually reviewed scoring. The manual review of the ARES™ autoscored raw signals, in fact, provided a high level of confidence in the automated analysis, and our overall results comparing these automated SDB indices with the simultaneous lab NPSG-derived indices was sufficiently good that we did not perform an event-by-event analysis in the 92 studies in which this could have been done.

(iv) Ambiguity of respiratory signals if period leg movements are present: As previously reported,⁸ periodic arousals resulting from period leg movements may create apparent respiratory events due to the large breaths associated with the periodic leg movement-induced arousals. In 1 patient, this resulted in an elevated RDI (but not AHI4%) during the ARES™ analysis, compared with the RDI obtained during the NPSG (See Figure 8.) This was because the technician scoring the full NPSG was able to visualize the period leg movements and incorporate this information in her scoring.

When ARES™ and NPSG data were acquired on separate nights, there was, as expected, a greater deviation in the SDB indices. It is likely that, in addition to the factors listed above, there was an important contribution from night-to-night variability, including positional and sleep-stage effects. The differences we report are in the same range as the night-to-night variability reported for data collected on multiple nights of full polysomnography (as part of the Sleep Heart Health Study) in the home or 1 night in the lab and 1 night in the home.^{21,22} The ICCs reported in these studies were 0.81 (for 2 home studies)²² and 0.77 (lab versus home study),²¹ with 15% to 25% of the subjects crossing a threshold of 5 or 10 for RDI4% based on 2 separate nights of full polysomnographic recording. Consistent with subjects in other studies,²³ our subjects spent significantly more time in the supine position in the laboratory, compared with in the home, and this was reflected in the difference between ARES™ lab versus home SDB indices. This reinforces the advantage of doing studies in the more typical home environment because the data may be more representative of the typical disease burden.

In summary, the present data again confirm that it is possible to obtain SDB indices comparable to those obtained by full-lab NPSG from data acquired by an unattended, limited diagnostic device, at least in subjects suspected of having only SDB or of having no sleep disorder. Furthermore, it is noteworthy that these results were obtained with a device that was self applied by patients in the home with only written instructions and analyzed with an automated algorithm. Although the present study did not test patient outcomes nor did it explore a full diagnostic algorithm, as was recently reported by Mulgrew et al,²⁴ our intent was solely to evaluate the validity of using an unattended limited-data-acquisition device to obtain SDB indices in a population suspected of having SDB. The high sensitivity and specificity obtained in the group we studied (which included

asymptomatic volunteers) support the potential utility of an ambulatory limited-monitoring approach as part of an algorithm designed not only for diagnosing SDB (case finding), but also to rule out SDB (screening) in suitably selected groups.

Our data are in agreement with those of other studies comparing SDB indices obtained using limited monitoring^{8,10,11,25,26} or unattended NPSG²¹ compared with in-lab full NPSG. The demonstration that limited monitoring can provide a valid representation of SDB, as measured by in-lab NPSG along with a home continuous positive airway pressure titration, has the *potential* to increase access to underserved populations at a reduced cost and with minimal patient burden, but this needs to be tested directly. In addition, ambulatory, limited studies should allow acquisition of data from large numbers of subjects not easily studied in the lab to address the critical issue of the variability of SDB and its consequences in multiple populations.

ACKNOWLEDGMENTS

We thank Dan Levendowski and Tim Zavora for technical assistance with the ARES™ studies and Rakhil Kanevsakaya who performed the scoring of the laboratory nocturnal polysomnograms.

REFERENCES

1. Kripke DE, Ancoli-Israel S, Klauber MR, Wingard DL, Mason WJ, Mullaney DJ. Prevalence of sleep-disordered breathing in ages 40-64 years: a population-based survey. *Sleep* 1997;20:65-76.
2. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328:1230-5.
3. Peppard PE, Young T, Palta M, Skatrud J. Prospective study of the association between sleep-disordered breathing and hypertension. *N Engl J Med* 2000;342:1378-84.
4. Phillipson EA. Sleep apnea--a major public health problem. *N Engl J Med* 1993;328:1271-3.
5. Colton HR, Altevogt BM, eds. Committee on Sleep Medicine and Research and Board on Health Sciences. *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*. Washington: The National Academies Press; 2006.
6. Flemons WW, Douglas NJ, Kuna ST, Rodenstein DO, Wheatley J. Access to diagnosis and treatment of patients with suspected sleep apnea. *Am J Respir Crit Care Med* 2004;169:668-72.
7. Joint Commission on Accreditation of Healthcare Organization. Potential 2008 National Patient Safety Goals. 2007. Available at: <http://www.jointcommission.org/Standards/FieldReviews>. Accessed on August 28, 2007.
8. Ayappa I, Norman RG, Suryadevara M, Rapoport DM. Comparison of limited monitoring using a nasal-cannula flow signal to full polysomnography in sleep-disordered breathing. *Sleep* 2004;27:1171-9.
9. Westbrook PR, Levendowski DJ, Cvetinovic M, et al. Description and validation of the apnea risk evaluation system: a novel method to diagnose sleep apnea-hypopnea in the home. *Chest* 2005;128:2166-75.
10. Dingli K, Coleman EL, Vennelle M, et al. Evaluation of a portable device for diagnosing the sleep apnoea/hypopnoea syndrome. *Eur Respir J* 2003;21:253-9.
11. Yin M, Miyazaki S, Ishikawa K. Evaluation of type 3 portable monitoring in unattended home setting for suspected sleep apnea: factors that may affect its accuracy. *Otolaryngol Head Neck Surg* 2006;134:204-9.

12. Hosselet JJ, Ayappa I, Norman RG, Krieger AC, Rapoport DM. Classification of sleep-disordered breathing. *Am J Respir Crit Care Med* 2001;163:398-405.
13. Westbrook, P. R. Survey regarding limited diagnostic systems for sleep apnea. *J Clin Sleep Med* 2007;3:318-20.
14. Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* 1999;22:667-89.
15. Iber C, Ancoli-Israel S, Chesson A, Quan SF for the American Academy of Sleep Medicine. The AASM Manual for the Scoring of Sleep and Associated Events: Rules Terminology and Technical Specifications. 1st ed. Westchester, IL: American Academy of Sleep Medicine; 2007.
16. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques, and scoring system for sleep states of human subjects. Washington: US Government Printing Office; 1968.
17. Meoli AL, Casey KR, Clark RW, et al. Hypopnea in sleep-disordered breathing in adults. *Sleep* 2001;24:469-70.
18. Ayappa I, Norman RG, Krieger AC, Rosen A, O'Malley RL, Rapoport DM. Non-invasive detection of respiratory effort-related arousals (RERAs) by a nasal cannula/pressure transducer system. *Sleep* 2000;23:763-71.
19. Flemons WW, Littner MR. Measuring agreement between diagnostic devices. *Chest* 2003;124:1535-42.
20. Zafar S, Ayappa I, Norman RG, Krieger AC, Walsleben JA, Rapoport DM. Choice of oximeter affects apnea-hypopnea index. *Chest* 2005;127:80 -8.
21. Iber C, Redline S, Kaplan Gilpin AM, et al. Polysomnography performed in the unattended home versus the attended laboratory setting—Sleep Heart Health Study methodology. *Sleep* 2004;27:536-40.
22. Quan SF, Griswold ME, Iber C, et al. Short-term variability of respiration and sleep during unattended nonlaboratory polysomnography—the Sleep Heart Health Study. *Sleep* 2002;25:843-9.
23. Metersky ML, Castriotta RJ. The effect of polysomnography on sleep position: possible implications on the diagnosis of positional obstructive sleep apnea. *Respiration* 1996;63:283-7.
24. Mulgrew AT, Fox N, Ayas NT, Ryan CF. Diagnosis and initial management of obstructive sleep apnea without polysomnography: a randomized validation study. *Ann Intern Med* 2007;146:157-66.
25. Pittman SD, Ayas NT, MacDonald MM, Malhotra A, Fogel RB, White DP. Using a wrist-worn device based on peripheral arterial tonometry to diagnose obstructive sleep apnea: in-laboratory and ambulatory validation. *Sleep* 2004;27:923-33.
26. Reichert JA, Bloch DA, Cundiff E, Votteri BA. Comparison of the NovaSom QSG, a new sleep apnea home-diagnostic system, and polysomnography. *Sleep Med* 2003;4:213-8.