There is a general recognition that sleep staging and sleep architecture are important for the understanding of sleep pathology. In normal subjects, deeper and less-fragmented sleep have been associated with more nocturnal blood pressure dipping. Abnormal sleep architecture also appears to be an important feature in clinical conditions like depression. In patients with obstructive sleep apnea (OSA), decreased REM continuity and deep sleep quantity constitute hallmark features which resolve following treatment.

The conventional objective measure of sleep and associated breathing disorders is polysomnography (PSG). However, PSG is a cumbersome, complex, and expensive technique. Therefore, unattended portable monitors for the diagnosis of OSA were introduced into the AASM guidelines. Despite the fact that portable monitors detect OSA with reasonable accuracy, most of the devices do not record sleep itself. The ability to extract information on sleep stage and sleep architecture from the limited physiological measurements in a portable device therefore would be desirable.

The Watch-PAT100 (PAT recorder; Itamar Medical, Caesarea, Israel) is a wrist-worn ambulatory sleep recorder based on peripheral arterial tone (PAT) signal, pulse rate, actigraphy and pulse oximetry. It has been shown to accurately detect OSA (in laboratory and ambulatory settings), autonomic arousals, and sleep/wake status. The validity of using PAT and actigraphy signals derived from the PAT recorder to detect REM and light/deep sleep has also been demonstrated in two small studies. In the current study, we aimed to validate the sleep staging algorithm of the PAT recorder device in a large population including normal subjects and patients with suspected OSA. The original goal of this multi-center study was to test the accuracy of the PAT recorder in detecting sleep disordered breathing. Since that time a novel algorithm to assess sleep staging...
based on these signals was developed, and we aimed to test this algorithm in a “post hoc” assessment of the original data set. We hypothesized that the software for analysis of autonomic changes associated with different sleep stages could be used to stage sleep in this large clinical cohort.

**METHODS**

**Overview of Study Design**

Two hundred twenty-eight subjects were recruited from Haifa, Israel; Boston, US; and Skara, Sweden, but 1 subject was rejected due to noisy signal. Detailed information on this multi-center study cohort has been previously reported. In brief, the cohort consisted of 17 normal volunteers, 139 patients referred to the sleep laboratory due to suspected OSA, and 71 subjects randomly drawn from a population-based cohort of subjects undergoing ambulatory PSG studies to investigate the association between sleep disordered breathing and blood pressure. The original cohort included normotensive subjects (random population sample) and patients with hypertension. Signed IRB-approved consent forms were obtained prior to enrollment in each institution.

All participants underwent an overnight PSG study and simultaneous recording with a Watch-PAT100 (PAT recorder). PSG data were manually scored by experienced technicians blinded to the PAT recorder data. The PAT recorder signals were automatically scored using a novel software package which determines wake, light sleep, deep sleep, and REM sleep (zzzPAT, Itamar Medical, Caesarea, Israel). Data from the 2 monitoring systems were synchronized and compared on an epoch-by-epoch basis in the subsequent analysis.

**PSG Recording**

Overnight PSG was performed according to standard protocol/criteria using electroencephalogram, electrooculogram, submental and bilateral anterior tibialis electromyography, electrocardiogram (ECG), nasal-oral airflow (thermistors and nasal pressure), chest and abdominal wall motion (piezo or impedance belts), body position and arterial oxygen saturation. The Embla system (Embla, Broomfield, USA) was used at the Skara and Haifa centers, whereas the Alice system (Respironics, Pittsburgh, USA) was used in Boston. The respiratory disturbance index (RDI) was calculated as the number of respiratory events (apnea, hypopnea, and RERA – respiratory effort related arousal) divided by the actual sleep duration determined by the PSG. The 227 subjects were stratified into 4 subgroups based on PSG determinations of OSA severity: (1) RDI < 10/h, normal range; (2) 10 ≤ RDI < 20/h, mild OSA; (3) 20 ≤ RDI < 40/h, moderate OSA; and (4) RDI ≥ 40/h, severe OSA.

**WP100 Recording**

The PAT recorder device has previously been described in detail. Briefly, this battery-powered, wrist-mounted device records PAT signal (finger arterial pulse wave volume), pulse rate derived from the PAT signal, oxyhemoglobin saturation, and wrist activity (derived from actigraphy). A continuous synchronization signal generated by the PAT recorder was recorded on both the PAT recorder device and the PSG for epoch-by-epoch comparison.

**Automatic zzzPAT Algorithm**

The sleep stage detection algorithm used in the PAT recorder is based on the PAT and the actigraphy signals. After scoring wake epochs, sleep epochs were classified as REM/NREM sleep by the time series of the PAT amplitude and PAT-derived inter-pulse periods. Then, the NREM sleep epochs were further categorized as either light (mirror sleep stage 1 and 2) or deep (mirror sleep stage 3 and 4) sleep. The detailed algorithms of sleep/wake detection from actigraphy signal and sleep staging from the PAT signal have previously been described.

In brief, sleep/wake detection is based on assessment of movements and their occurrences (periodic or sporadic) while the sleep stage detections (REM, deep/light sleep) are based on the spectral components of the PAT signal. A set of 14 variables derived from the PAT signal amplitude time series as well as the pulse rate and their conditional probabilities were computed within a 5-min sliding window advanced by 30-sec epochs. A 2-step algorithm was performed to combine and weigh each of the features. The first step was to filter each of the features by defining a ± 5-min window around each epoch, allowing for smoothing around the epoch under consideration (Neighboring Filter). The second step was done by selected weighting to minimize the differences between the PSG staging and the PAT derived staging. It should be noted that this algorithm was developed on a prior, separate set of patients (training set), no further algorithm development was done in the current study population (validation set).

**Data Analysis**

Sleep and sleep stage data from PSG and the PAT recorder were compared on an epoch-by-epoch basis using several approaches. The sensitivity, specificity and agreement of the PAT recorder derived sleep staging and sleep parameters with the scored PSG data were examined in different OSA severity subgroups. Agreement was considered when both methods had the same score for a given epoch. Inter class correlation (ICC) and Bland-Altman plots were created for RDI. Sleep stages from PSG scoring and the PAT recorder automatic analyses were compared and a confusion matrix was computed on a total of 198,815 pooled epochs. Cohen κ coefficients for the entire set of sleep stages were calculated. All values are presented as mean ± SD with p < 0.05 being considered statistically significant.

**RESULTS**

A total of 227 subjects were analyzed in the study (age 49 ± 14 y, body mass index 29 ± 6 kg/m², and RDI 30 ± 23 events/h).

**REM Sleep Detection**

The sensitivity, specificity, and agreement of the algorithm to detect REM sleep in different OSA severity groups ranged between 59% and 94%. For normal subjects, these were 62.7 ± 28.3, 92.9 ± 4.7, and 88.5 ± 4.6, respectively; for mild OSA these were 68.9 ± 20.1, 91.9 ± 6.1, and 87.9 ± 5.6; for moderate OSA 66.9 ± 25.9, 92.2 ± 5.8, and 88.5 ± 5.9; and for severe OSA 59.2 ± 31.1, 94.2 ± 5.4, and 90.0 ± 5.3, respectively. The overall agreement of PSG scored REM sleep and the PAT recorder scored REM sleep was 88.7% ± 5.5%. The severity of
OSA (and the location of acquisition) did not have a substantial effect on the algorithm accuracy with regard to REM sleep detection. The 2 methods provided similar REM latency and REM percentage (237 ± 148 vs. 225 ± 159 epochs and 14.4% ± 6.5% vs. 19.3% ± 8.7%, respectively).

Light/Deep Sleep Detection
The sensitivity, specificity, and agreement in detecting deep versus light sleep in subgroups categorized by severity of OSA ranged between 64% and 96%. For normal individuals these were 68.8 ± 22.6, 91.1 ± 4.9, and 87.1 ± 5.1, respectively; for mild OSA these were 63.6 ± 23.9, 91.3 ± 5.1, and 86.4 ± 4.5; for moderate OSA 65.9 ± 25.1, 91.2 ± 5.2, and 87.5 ± 6.0; and for severe OSA 73.2 ± 27.1, 96.3 ± 3.8 and 93.5 ± 4.9, respectively. The total agreement of the 2 methods was 88.6% ± 5.9% for light/deep sleep detection. The overall agreement in terms of sleep stage (wake/light sleep/deep sleep/REM sleep) is presented in Table 1. The agreement between automatic scoring using the PAT recorder and the PSG was 66.0% ± 8.8%. Confusion matrices of epochs detected by the 2 methods and classified as wake/light sleep/deep sleep/REM sleep stages are presented in Table 2. The total Cohen κ coefficient for all the stages was 0.475 (95% CI = 0.472-0.479).

The agreements for other sleep related measures by PSG and the PAT recorder, respectively, were sleep latency (57 ± 31 vs. 43 ± 45 epochs, p < 0.05, ICC = 0.57, p < 0.01), sleep efficiency (78.4% ± 9.9% vs. 78.8% ± 13.4%, NS, ICC = 0.62, p < 0.01), and total sleep time (690 ± 152 vs. 690 ± 154 epochs, NS, ICC = 0.79, p < 0.01).

Sleep Apnea Detection
The agreement between methods to detect a respiratory disturbance event was 80% (event by event detection). The ICC for RDI determined by the 2 methods was 0.87, p < 0.05. A Bland-Altman plot for RDI by the 2 methods is presented in Figure 1. There was a high sensitivity, specificity, and agreement for the zzzPAT algorithm to detect OSA subjects based on a RDI threshold of 10/h, with an area under curve of 0.96.

Discuss
In this large multi-center cohort study, we demonstrated a moderate accuracy when using PAT and actigraphy signals to

<table>
<thead>
<tr>
<th>Level of OSA severity</th>
<th>No</th>
<th>Sleep Stages Agreement, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>38</td>
<td>65.4 ± 9.5</td>
</tr>
<tr>
<td>Mild</td>
<td>54</td>
<td>64.7 ± 6.6</td>
</tr>
<tr>
<td>Moderate</td>
<td>82</td>
<td>64.2 ± 7.8</td>
</tr>
<tr>
<td>Severe</td>
<td>53</td>
<td>70.6 ± 10.4</td>
</tr>
<tr>
<td>All</td>
<td>227</td>
<td>66.0 ± 8.8</td>
</tr>
</tbody>
</table>

Agreement is defined as the number of epochs where agreement exists in any specific state divided by the total number of epochs (shown as mean ± SD).

Table 1—Overall agreement of the PAT recorder device with PSG for detection of stage (wake/light sleep/deep sleep/REM sleep) in subgroups of subjects categorized by OSA severity (epoch-by-epoch comparison).

Table 2—A confusion matrix of overall number of epochs detected by the 2 methods and classified as wake/light sleep/deep sleep/REM sleep stages. The columns represent epochs scored from the PSG, while the rows represent epochs scored automatically by the zzzPAT algorithm.

<table>
<thead>
<tr>
<th>PAT recorder ↓</th>
<th>PSG →</th>
<th>Wake</th>
<th>REM sleep</th>
<th>Light sleep</th>
<th>Deep sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake</td>
<td>26423</td>
<td>2988</td>
<td>11654</td>
<td>441</td>
<td></td>
</tr>
<tr>
<td>REM sleep</td>
<td>1776</td>
<td>18467</td>
<td>9196</td>
<td>330</td>
<td></td>
</tr>
<tr>
<td>Light sleep</td>
<td>11665</td>
<td>8175</td>
<td>71266</td>
<td>12409</td>
<td></td>
</tr>
<tr>
<td>Deep sleep</td>
<td>949</td>
<td>470</td>
<td>8255</td>
<td>14351</td>
<td></td>
</tr>
</tbody>
</table>

The total Cohen κ coefficient for all the stages was 0.475 (95% CI = 0.472-0.479).

<table>
<thead>
<tr>
<th>PAT recorder ↓</th>
<th>PSG →</th>
<th>Wake</th>
<th>Sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake</td>
<td>26423</td>
<td>15083</td>
<td></td>
</tr>
<tr>
<td>Sleep</td>
<td>14390</td>
<td>142919</td>
<td></td>
</tr>
</tbody>
</table>

Cohen κ coefficient for sleep wake was 0.549 (95% CI = 0.544- 0.553).

<table>
<thead>
<tr>
<th>PAT recorder ↓</th>
<th>PSG →</th>
<th>REM sleep</th>
<th>NREM sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>REM sleep</td>
<td>18467</td>
<td>9526</td>
<td></td>
</tr>
<tr>
<td>NREM sleep</td>
<td>8645</td>
<td>106281</td>
<td></td>
</tr>
</tbody>
</table>

Cohen κ coefficient for detecting REM from NREM (among sleep stages detected, not including wake) was 0.592 (95% CI = 0.586- 0.597).

<table>
<thead>
<tr>
<th>PAT recorder ↓</th>
<th>PSG →</th>
<th>Deep sleep</th>
<th>Light sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep sleep</td>
<td>14351</td>
<td>8255</td>
<td></td>
</tr>
<tr>
<td>Light sleep</td>
<td>12409</td>
<td>71266</td>
<td></td>
</tr>
</tbody>
</table>

Cohen κ coefficient for detecting deep from light stage 0.456 (95% CI = 0.449-0.463).
Based on these assumptions, it may be hypothesized that an algorithm that includes PAT, pulse rate and the actigraphy may be able to detect sleep/wake status with reasonable accuracy. Regardless of the method used, essentially all studies have reported that OSA is associated with increased sympathetic activity. Furthermore, unlike the situation in normal subjects, blood pressure and sympathetic activity remain high during sleep in OSA patients. These findings highlight the challenges of using autonomic signals to detect sleep stages in OSA patients. However, peak sympathetic activity usually is seen at the termination of respiratory events. This enables a sophisticated algorithm that defines a local baseline and calculates changes over different time periods to detect both state specific and event specific changes. Indeed, despite the general increase of sympathetic tone and large respiratory event related sympathetic activations in OSA, the algorithm used in this study was found to detect specific autonomic changes associated with sleep and differentiate sleep stages with a reasonable accuracy.

In this study, we have demonstrated that the PAT signal and actigraphy from the PAT recorder device may be useful to differentiate wake from sleep, and to stratify sleep into light, deep and REM stages. In the era of an emerging need for simple monitors to diagnose OSA, these findings are meaningful. While most portable devices do not assess sleep or sleep stages, previous studies. Based on these assumptions, it may be hypothesized that an algorithm that includes PAT, pulse rate and the actigraphy may be able to detect sleep/wake status with reasonable accuracy.

Regardless of the method used, essentially all studies have reported that OSA is associated with increased sympathetic activity. Furthermore, unlike the situation in normal subjects, blood pressure and sympathetic activity remain high during sleep in OSA patients. These findings highlight the challenges of using autonomic signals to detect sleep stages in OSA patients. However, peak sympathetic activity usually is seen at the termination of respiratory events. This enables a sophisticated algorithm that defines a local baseline and calculates changes over different time periods to detect both state specific and event specific changes. Indeed, despite the general increase of sympathetic tone and large respiratory event related sympathetic activations in OSA, the algorithm used in this study was found to detect specific autonomic changes associated with sleep and differentiate sleep stages with a reasonable accuracy.

In this study, we have demonstrated that the PAT signal and actigraphy from the PAT recorder device may be useful to differentiate wake from sleep, and to stratify sleep into light, deep and REM stages. In the era of an emerging need for simple monitors to diagnose OSA, these findings are meaningful. While most portable devices do not assess sleep or sleep stages,
the index calculations are based on total recording time or self-reported sleep time which both may lead to an overestimation of actual sleep time and a dilution of the indices of the respiratory event indices. A portable device with the ability to detect wakefulness, REM sleep, light and deep sleep may therefore improve the standards of home diagnosis of OSA.

It may be argued that the overall 66% agreement for sleep state classification reported in this study is insufficient for clinical use. However, the Cohen κ coefficients for the various stages ranged from 0.46-0.59, which are considered as moderate agreement according to Cohen κ criteria. Moreover, these results are essentially within the variability range reported in some studies comparing registered PSG scorers and similar to the agreement reported in comparisons between automated PSG scoring and manual scoring. Multiple scoring sites, various recording techniques and a mixture between patients and normal subjects were used in our study, which may have increased biological and methodological variability compared to prior single center studies. However, the accuracy of the algorithm was similar across these various conditions. Therefore, we believe our results are reasonable and may justify further consideration of our algorithm. Moreover, recent study has demonstrated that specific reflection of autonomic signals during sleep could provide clinical and functional information that may not be obtained by traditional PSG recordings.

Some limitations of the study deserve mention. First, the study population did not include children or patients with specific movement or neurological disorders, limiting the generalizability of our study. Second, the current study was conducted before implementation of the new AASM scoring manual. However, we combined the sleep stages 3 and 4 into a single deep sleep stage which is identical to stage N3 classified in the AASM manual. Third, as stated above, we did not validate scoring accuracy between labs, which can introduce variability. However, we did not encounter differences between labs in the accuracy of the PAT recorder device when compared to PSG.

In conclusion, we have presented data on the accuracy of an automated sleep staging algorithm based on PAT and actigraphy from the PAT recorder device in normal subjects and patients. Multiple scoring sites, various recording techniques and a mixture between patients and normal subjects were used in our study, which may have increased biological and methodological variability compared to prior single center studies. However, the accuracy of the algorithm was similar across these various conditions. Therefore, we believe our results are reasonable and may justify further consideration of our algorithm. Moreover, recent study has demonstrated that specific reflection of autonomic signals during sleep could provide clinical and functional information that may not be obtained by traditional PSG recordings.

Some limitations of the study deserve mention. First, the study population did not include children or patients with specific movement or neurological disorders, limiting the generalizability of our study. Second, the current study was conducted before implementation of the new AASM scoring manual. However, we combined the sleep stages 3 and 4 into a single deep sleep stage which is identical to stage N3 classified in the AASM manual. Third, as stated above, we did not validate scoring accuracy between labs, which can introduce variability. However, we did not encounter differences between labs in the accuracy of the PAT recorder device when compared to PSG.

In conclusion, we have presented data on the accuracy of an automated sleep staging algorithm based on PAT and actigraphy from the PAT recorder device in normal subjects and patients with various severities of OSA. These results are of substantial interest in the era of a shift toward unattended ambulatory sleep recordings.

REFERENCES

3. McArdle N, Douglas NJ. Effect of continuous positive airway pressure on sleep architecture in the sleep apnea-hypopnea syndrome: a randomized controlled trial. Am J Respir Crit Care Med 2001;164:1459-63
7. Pittman SD, Ayas NT, MacDonald MM, et al. 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
28. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968;70:213-20
J Hedner, DP White, A Malhotra et al

Address correspondence to: Giora Pillar, M.D., Rambam Medical Center, Sleep Lab, 6 Efron St., Haifa, 31096, Israel; Tel: 972-4-8542646; Fax: 972-4-8542441; E-mail: gpillar@tx.technion.ac.il

DISCLOSURE STATEMENT

This study was partially supported by a non-restrictive grant from Itamar Medical Ltd. Dr. Hedner has participated in clinical trials sponsored by Schering Plough, Merck, Philips Respironics, Weinmann and Volvo. He has received honorariums for lectures from Weinmann and Nycomed. Dr. White is a consultant and has received research grants from Itamar Medical, and is currently the Chief Medical Officer of Philips Respironics. Atul Malhotra has received consulting and/or research income from Itamar, SHC, Apnex, SGS, Philips, Ethicon, Medtronic, Pfizer, SHC, Merck, Sepracor, and Cephalon. Dr. Herscovici is a former employee of Itamar Medical Ltd. Dr. Pittman is a former consultant to Itamar Medical and is an employee of Philips Respironics. Dr. Zou has indicated no financial conflict of interest. Dr. Grote has participated in clinical studies sponsored by Schering Plough, Merck, Philips Respironics, Weinmann, Mundipharma and Volvo. He has also received honorariums for speaking activities for Weinmann, Philips Respironics, Resmed, and Nycomed. Dr. Pillar is a consultant to Itamar Medical and Discover Medical Devices and has received research grants from Cephalon Ltd., Intec Pharma, and Actelion Ltd.