https://doi.org/10.5664/jcsm.9300

JCSM | Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Night-to-night variability in obstructive sleep apnea using peripheral arterial tonometry: a case for multiple night testing

Samuel Tschopp, MD[1]; Wilhelm Wimmer, PhD[1,2]; Marco Caversaccio, MD[1], Urs Borner, MD[1,*]; Kurt Tschopp, MD[3,*]

*[1]Department of Otorhinolaryngology, Head and Neck Surgery, Inselspital, University Hospital and University of Bern, Bern, Switzerland; [2]Hearing Research Laboratory, ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland; [3]Department of Otorhinolaryngology, Head and Neck Surgery, Kantonsspital Baselland, Liestal, Switzerland; *Contributed equally*

Study Objectives: Night-to-night variability of obstructive sleep apnea severity is considerable and may depend on the diagnostic modality used. We investigated the night-to-night variability using peripheral arterial tonometry (PAT).

Methods: Home sleep apnea testing was performed in 51 patients during 3 consecutive nights using PAT. Patients referred to our sleep clinic were screened and prospectively recruited for this study. All recordings were automatically and manually scored according to the PAT scoring guidelines.

Results: No systematic differences in PAT-derived apnea-hypopnea index (pAHI) were found between the nights. The night-to-night variability was comparable between manually and automatically scored data. pAHI varied in 35% of patients more than 10 events/h between the nights. The obstructive sleep apnea severity of 24% of patients was misclassified when using 1 night compared to the average of all nights. On average, pAHI varied by 57% from night to night. The variability of pAHI could partially be explained by the variability of time spent in the supine position with more time supine leading to a higher pAHI. On measuring a subsequent night, 12–14% of patients spontaneously fulfilled the commonly accepted criteria for treatment success without any intervention.

Conclusions: With repeated recordings of PAT, we found no first night effect. However, there is considerable night-to-night variability similar to values found for polysomnography, which can partially be explained by the variability of time spent in the supine position. Obstructive sleep apnea severity was frequently misclassified due to the night-to-night variability. Our findings make a strong case for multiple testing in the diagnostic work-up of obstructive sleep apnea patients.

Key words: sleep apnea, obstructive, obstructive sleep apnea, sleep study, peripheral arterial tonometry, home sleep apnea testing, misclassification, night-to-night variability

Citation: Tschopp S, Wimmer W, Caversaccio M, Borner U, Tschopp K. Night-to-night variability in obstructive sleep apnea using peripheral arterial tonometry: a case for multiple night testing. *J Clin Sleep Med*. 2021;17(9):1751–1758.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Night-to-night variability of obstructive sleep apnea severity is considerable and has large implications for diagnostics, clinical decision making, and outcome evaluation. The correct assessment of obstructive sleep apnea severity is especially important to initiate the appropriate treatment for patients.

Study Impact: This is the first study using peripheral arterial tonometry to demonstrate that there is a high night-to-night variability of obstructive sleep apnea severity but no first night effect. Our results make a strong case for multiple night recordings in the diagnostic work-up of patients with obstructive sleep apnea to improve the accuracy of sleep testing.

INTRODUCTION

The underlying pathophysiology of obstructive sleep apnea (OSA) is the repetitive narrowing and collapse of the upper airway during sleep.[1] The main parameter in measuring OSA severity is the apnea-hypopnea index (AHI). The classification of AHI has been standardized by the American Academy of Sleep Medicine has been last updated in 2020 with the American Academy of Sleep Medicine Scoring Manual, Version 2.6.[2] This standardization aims to facilitate and enable the comparison of sleep recordings. In research, AHI is the most widely used measure for classification of OSA severity, reporting the success of therapeutic interventions, and in epidemiological studies of OSA prevalence.

Although the American Academy of Sleep Medicine scoring guidelines allow for precise determination of AHI, this parameter is subject to considerable night-to-night variability (NNV) in repeated sleep recordings.[3–5] Several studies demonstrated this for polysomnography (PSG) and home sleep apnea testing (HSAT) using cardiorespiratory polygraphy.[6–11] NNV was sizable but similar with both methods, suggesting that it might not be specific to the type of sleep examination.

Several factors leading to a higher NNV have been identified, such as obesity, smoking, and financial strain.[12] Sforza et al[13] and Störbel et al[14] found a higher variability in patients with less severe OSA, while Bittencourt et al[6] and Aarab et al[4] found no association or the opposite effect in their collective. The NNV can be partially explained by the variability of time spent in the

supine position.[15] White et al[16] found that rostral fluid shift is a contributing factor to NNV, which might be especially important in patients with heart failure, hypertension, or renal disease. However, clear predictors for NNV remain allusive.

An additional obstacle in sleep apnea testing is the first night effect (FNE), which describes changes to the sleep pattern in sleep studies compared to natural sleep. The FNE is caused by the unfamiliar environment of a sleep laboratory and possibly also by the recording equipment itself.[17,18] Studies investigating FNE in OSA diagnosis have shown conflicting results. While several studies found no or a minimal FNE,[9,19–21] others show a considerable difference between the first and subsequent nights.[12,13,15,18,22,23] Furthermore, the disruption of normal sleep pattern might not only affect the first night but extend to several nights.[18] OSA severity is generally underestimated on the first night.[8,13,15,23] An FNE has been well demonstrated for in-lab PSG and to a lesser extent also for PSG at home, affecting mainly rapid eye movement (REM) sleep, indicating that the recording device may contribute to the phenomenon.[18,19] However, FNE has not been investigated yet for peripheral arterial tonometry (PAT) in HSAT.

PAT is a relatively novel technique for HSAT. PAT devices measure several physiological functions such as pulse waveform, heart rate and heart rate variability, oximetry, actimetry, body position, snoring, and newer devices, and also chest motion as a respiratory effort signal. Sleep-related breathing disturbances and autonomic arousals are detected using a combination of PAT signal, oximetry, and actigraphy. From this information, a PAT-derived AHI (pAHI) is calculated. Validation studies have found a strong correlation between PAT and PSG for AHI and sleep stages.[24,25] PAT devices are categorized as Type 3 devices according to the American Academy of Sleep Medicine, together with respiratory polygraphy for HSAT.[2] The investigation of the NNV using peripheral-arterial tonometry was suggested by Zhang et al 2020.[26]

To our knowledge, there are no studies that evaluate the NNV for PAT, which is an increasingly used method for HSAT. The goal of this study is to describe the variability for PAT measurements based on 3 consecutive recordings with respect to 1) NNV of AHI, 2) FNE, 3) misclassification rate of OSA severity, 4) standard error of measurement, and 5) minimal detectable difference. The knowledge of these parameters is indispensable for clinicians to decide how many measurements are appropriate for a given clinical question.

METHODS

Between 2017 and 2018, patients with suspected OSA at our clinic were screened for their suitability for multiple nights of sleep apnea testing using PAT. All patients consented to the use of their data, and the study was approved by the local Ethics Committee. The inclusion criteria were heavy snoring or suspected OSA, absence of severe cardiac or neurological comorbidity (American Society of Anthologists class I or II), willingness to undergo recording of sleep parameters on 3 consecutive nights, and a signed informed consent.

In all patients, a detailed sleep history was taken including the Epworth Sleepiness Scale and self-reported snoring intensity on a visual analog scale between 1 and 10 (1 meaning no snoring and 10 meaning unsupportable snoring). All patients were measured during 3 consecutive nights using the WatchPAT 200 system (Itamar Medical, Caesarea, Israel). All patients were personally instructed on correct device installation. They were advised to follow their regular sleep habits and refrain from alcohol and sleep medication to obtain representative sleep studies. The recordings were checked for good signal quality in all channels. Scoring was performed automatically using the proprietary, validated computed-based algorithm (Itamar Medical, Caesarea, Israel) and manually by an experienced sleep technician according to the guidelines for WatchPAT scoring.[26] When not otherwise stated, the manually scored data were used. Sleep studies with less than 4 hours of sleep time or incomplete recording of the signal channels were excluded. For all recordings, the duration and proportion of sleep, wake, and sleep stages (light sleep, deep sleep, and REM sleep) were calculated. Obstructive breathing was characterized by pAHI, PAT-derived respiratory disturbance index, and oxygen desaturation index as a total of the night and depending on sleeping position and sleep stage. Statistics of pAHI and PAT-derived respiratory disturbance index were calculated both with automatically and manually scored data for comparison. The OSA severity was categorized using pAHI as no OSA: < 5 events/h, mild: $5 \leq 15$ events/h, moderate $15 \leq 30$ events/h, and severe: $>$ 30 events/h. In interventional sleep medicine, a reduction of baseline AHI by more than 50% and posttherapeutic AHI equal or less than 20 events/h is generally regarded as a responder criterion according to Sher et al.[27] Although no intervention took place between the recordings in our study, we evaluated if the NNV would lead to false-positive responders according to the Sher et al criteria.

For the classification of positional OSA, we used a ratio of supine AHI to nonsupine AHI of ≥ 2.[28,29] Further subgroups of positional OSA were determined according to the criteria proposed by Joosten et al[30] and Mador et al.[31] Supine-predominant OSA was defined as the ratio of supine to nonsupine AHI ≥ 2 with a nonsupine AHI > 5 events/h and supine-isolated OSA as the ratio of supine AHI to nonsupine AHI ≥ 2 with a nonsupine AHI < 5 events/h. The same was done for REM-associated OSA using the generally accepted definition of a ratio of REM to non-REM (NREM) AHI of ≥ 2.[32]

Baseline characteristics of the patient and differences between the 3 nights were analyzed using an ANOVA. The reliability of pAHI was assessed using the intraclass correlation coefficient, which describes the concordance of repeated measurements of the same patient. To quantify the range of the true measurement value the standard error of measurement (SEM) was calculated according to Fleiss and Kingman.[33,34] The SEM indicates the distribution of repeated measurements around the "true" value. As such it is a measure for the reliability of a test and 67% of results from repeated measurements will lie within the range of ± SEM around the true value. The minimal detectable difference (MDD) is the difference with which 2 measurements can be regarded as significantly different. The MDD for a 95% confidence interval (CI) is the SEM multiplied by 1.96

Table 1—Comparison between the 3 nights.

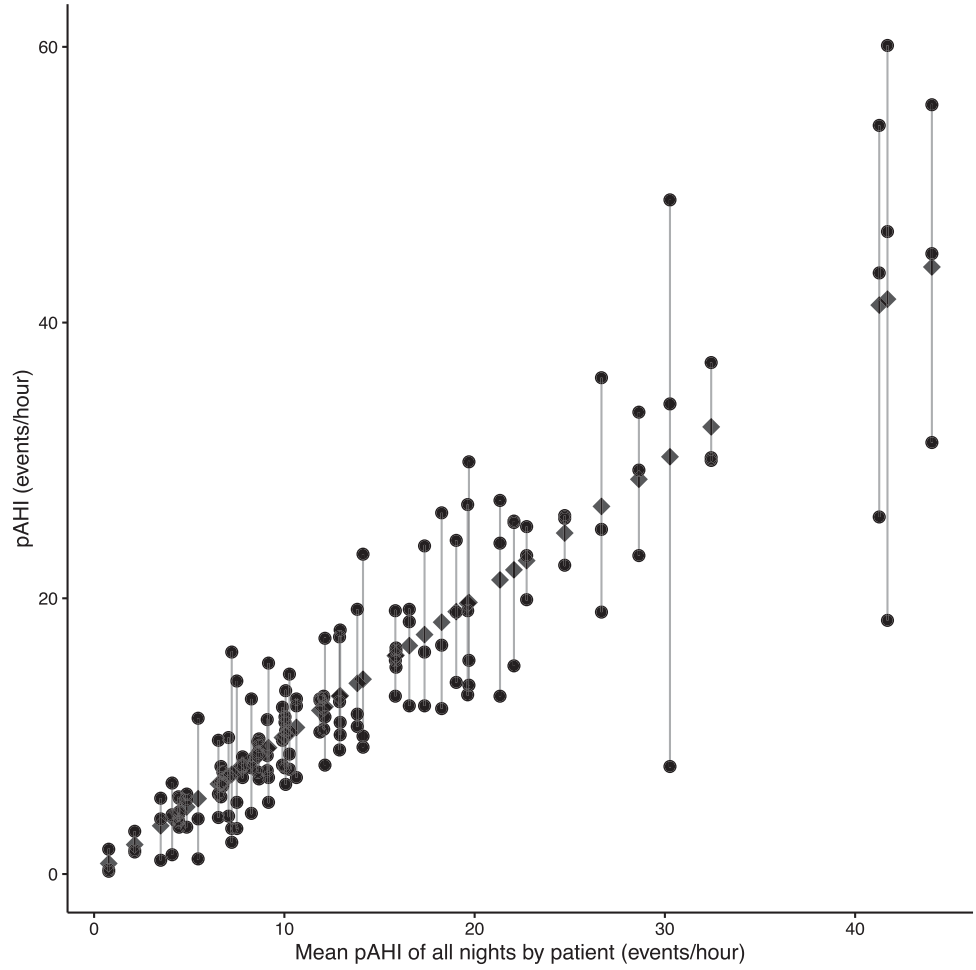
	Night			*P*
	1	2	3	
Recording time (hours)	7.9 ± 1.3	7.5 ± 1.2	7.1 ± 1.2	< .01
Total sleep time (hours)	6.8 ± 1.2	6.5 ± 1.1	6.0 ± 1.1	< .01
Manual scoring				
pAHI total (events/h)	15.0 ± 12.0	14.8 ± 11.6	13.9 ± 11.0	.87
pAHI in REM sleep (events/h)	18.6 ± 15.3	18.5 ± 14.8	18.7 ± 14.5	.99
pAHI in NREM (events/h)	13.8 ± 12.3	13.4 ± 12.0	11.1 ± 9.4	.44
pAHI in supine position (events/h)	27.7 ± 23.9	27.2 ± 21.3	24.2 ± 19.7	.69
pAHI in nonsupine position (events/h)	8.7 ± 9.6	9.2 ± 9.3	7.5 ± 7.8	.61
pAHI difference supine and nonsupine (events/h)	19.06 ± 22.13	18.04 ± 18.33	17.16 ± 17.25	.89
pAHI difference REM and NREM (events/h)	4.76 ± 13.65	5.15 ± 13.12	7.64 ± 12.19	.49
Automatic scoring				
pAHI total (events/h)	23.1 ± 13.3	24.0 ± 12.8	21.4 ± 14.3	.62
pAHI in REM sleep (events/h)	28.9 ± 16.5	30.7 ± 15.5	29.7 ± 16.0	.85
pAHI in NREM (events/h)	21.1 ± 14.1	21.6 ± 13.5	17.6 ± 13.0	.29
pAHI in supine position (events/h)	36.1 ± 24.8	35.9 ± 22.1	32.2 ± 22.1	.64
pAHI in nonsupine position (events/h)	16.9 ± 13.4	19.3 ± 11.9	15.1 ± 12.2	.25
Cartwright Index	11.5 ± 39.8	6.60 ± 13.8	6.7 ± 13.7	.57
REM Association	2.1 ± 2.4	2.1 ± 2.1	2.8 ± 4.1	.42
pRDI total (events/h)	18.9 ± 11.6	18.9 ± 11.5	18.1 ± 11.2	.93
pRDI in REM sleep (events/h)	22.0 ± 14.7	21.9 ± 14.5	22.28 ± 14.2	.99
pRDI in NREM (events/h)	17.9 ± 12.1	17.7 ± 12.0	15.7 ± 10.2	.58
ODI total (events/h)	10.1 ± 9.6	10.1 ± 7.8	9.2 ± 8.3	.82
ODI in REM sleep (events/h)	13.8 ± 13.3	14.6 ± 12.8	14.4 ± 12.3	.95
ODI in NREM (events/h)	8.93 ± 9.42	8.5 ± 7.6	6.8 ± 6.8	.37
Mean oxygen saturation (%)	94.5 ± 1.2	94.4 ± 1.2	94.3 ± 1.90	.79
Mean oxygen desaturation (%)	91.8 ± 1.8	91.5 ± 1.9	91.7 ± 1.8	.71
Time below 90% oxygen saturation (min)	3.8 ± 8.5	3.8 ± 8.9	3.7 ± 7.6	.99
Time below 90% oxygen saturation (%)	1.0 ± 2.3	0.9 ± 1.9	1.0 ± 2.1	.95
Mean heart rate (beats/min)	63.8 ± 8.9	64.2 ± 7.6	62.6 ± 8.0	.64
Supine time (% of total sleep time)	42.5 ± 23.5	40.4 ± 24.3	43.7 ± 24.4	.80
REM sleep (% of total sleep time)	23.6 ± 6.9	23.4 ± 6.2	22.6 ± 6.6	.72
Light sleep (% of total sleep time)	60.3 ± 10.1	60.1 ± 10.1	60.9 ± 9.7	.92
Deep sleep (% of total sleep time)	16.1 ± 5.8	16.4 ± 6.0	16.5 ± 5.8	.95
Sleep time (% of recording time	87.2 ± 6.3	86.5 ± 4.8	85.4 ± 4.8	.21
Wake time (% of recording time	12.8 ± 6.3	13.5 ± 4.8	14.6 ± 4.8	.21

Values are mean ± SD. *P* values are calculated using one-way analysis of variance. Cartwright Index = pAHI in supine position/pAHI in nonsupine position, REM Association = pAHI in REM sleep/pAHI in NREM sleep. NREM sleep = non–rapid eye movement sleep, ODI = oxygen desaturation index, pAHI = peripheral arterial tonometry–derived apnea-hypopnea index, pRDI = peripheral arterial tonometry–derived respiratory disturbance index, REM sleep = rapid eye movement sleep.

and the square root of 2. The MDD is important to clinicians when assessing treatment effects.[34]

We used a linear mixed-effects model to assess the NNV and differences in pAHI, with a fixed effect for the night (first night, second night, and third night) and a subject-level random effect to account for paired measurements. We included time supine (in minutes) as a fixed effect interacting with night to evaluate the influence of the sleeping pose on the pAHI outcome. All other covariates (ie, total pAHI, supine and nonsupine pAHI, percent in REM sleep, total sleep time, recording time, weekday, age, smoking habits [yes/no], presence of tonsils [yes/no], and body mass index) showed no statistical significance, and were excluded from the model. Medication and alcohol consumption were not included, since these data were not available.

Figure 1—An overview of all pAHI measurements grouped by the mean of a patient.



The circles indicate individual measurements. The gray squares give the mean of a patient over all measurements. pAHI = peripheral arterial tonometry–derived apnea-hypopnea

The misclassification rate was calculated as the rate of different OSA categories in a given night compared to the average of all 3 nights. Variability in positional and REM-associated OSA was assessed using the raw agreement and Cohen's kappa.[30] The raw agreement was calculated as the proportion of patients who tested in the same category on all three nights. Cohen's kappa coefficient (κ) measures the reliability of categorical items and controls for agreement occurring by chance.

The statistical analysis was performed using R Studio (Boston, MA). $P < .05$ was considered statistically significant.

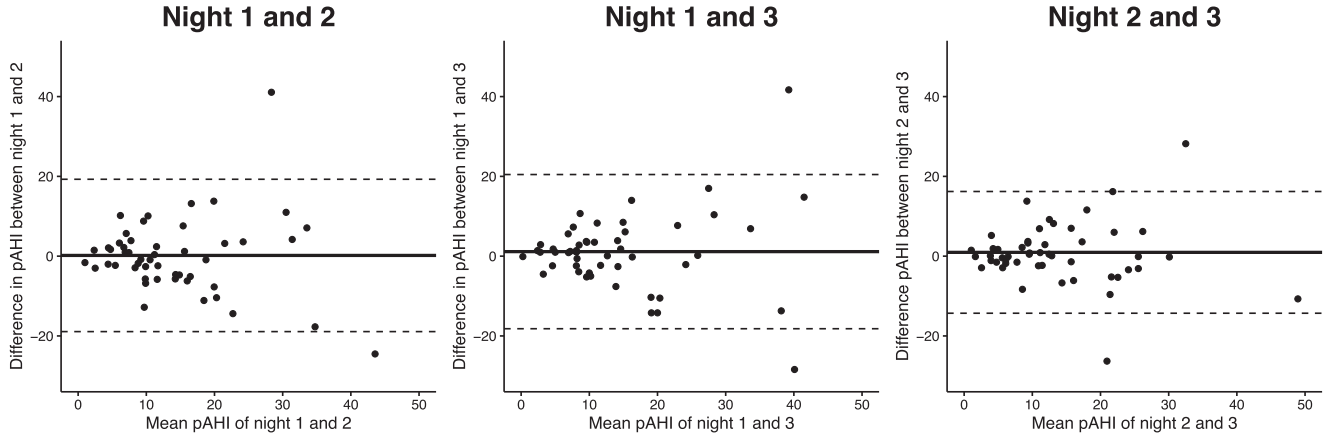
RESULTS

PAT recordings were obtained in 52 patients on 3 consecutive nights. One patient was excluded due to a technical failure in 2 of 3 recordings, leaving 51 data sets for final analysis. The patient collective was predominantly male (49 men, 2 women), middle-aged (51 ± 9 years), and with a mean body mass index of 26.0 ± 2.5. kg/m^2. No patient was performing shift work. The mean pAHI for the entire patient collective and all recordings was 14.5 ± 11.5 events/h. The mean Epworth Sleepiness Scale was 7.7 ± 4.9 and snoring intensity 8.2 ± 1.7. An overview of clinical and sleep parameters is given in **Table 1**.

Figure 1 gives a graphical overview of all measurements sorted by a patient's mean pAHI overall measurements. A considerable NNV of pAHI was found and illustrated using Bland–Altman plots in **Figure 2**. A similar NNV was found for pAHI in supine, nonsupine, REM, and NREM as well as time spent in supine position between the three nights.

In 35% (18 of 51 patients) of patients, the pAHI varied ≥ 10 events/h, and in 65% (33/51) more than ≥ 5 events/h between the 3 nights. On average, pAHI varied by 56.7% from the previous night. Compared to the average of all 3 nights, OSA severity was misclassified in 25% (13/51), 24% (12/51), and 22% (11/51) on the first, second, and third night, respectively. The categorization of OSA severity by night is given in **Table 2**. The Sher criteria for responders to surgery, defined as pAHI < 20 events/h and pAHI reduction > 50% compared to baseline, were fulfilled in 14% (7/51) and 12% (6/51) from the first to

Figure 2—Bland–Altman plots comparing pAHI of the 3 measurements.



pAHI = peripheral arterial tonometry–derived apnea-hypopnea.

Table 2—Agreement of OSA severity.

	Night 1	Night 2	Night 3
Severe OSA	6 (12%)	3 (6%)	4 (8%)
Moderate OSA	12 (24%)	16 (31%)	13 (25%)
Mild OSA	27 (53%)	24 (47%)	25 (49%)
No OSA	6 (12%)	8 (16%)	9 (18%)
Category changes from previous night		24 (47%)	18 (35%)
Average change of pAHI from previous night (%)		69.6	43.7

The number and percent of patients are given. The raw agreement for OSA classification was 0.37, with a *K* value of 0.36 and a *P* value < .001. Note that the *P* value reports the statistical significance of *K*, not the agreement between the nights. OSA categories are defined as no OSA: pAHI < 5 events/h, mild: pAHI 5 ≤ 15 events/h, moderate: pAHI 15 ≤ 30 events/h, severe ≥ 30 events/h. OSA = obstructive sleep apnea, pAHI = peripheral arterial tonometry–derived apnea-hypopnea index.

Table 3—Agreement of positional and REM-associated OSA between nights.

	Night 1	Night 2	Night 3	Raw Agreement	*K*	*P*
Positional OSA				0.52	0.24	< .01
Yes	34 (69%)	30 (63%)	32 (71%)			
No	15 (31%)	18 (38%)	13 (29%)			
Category changes from previous night		15 (31%)	13 (29%)			
Positional OSA subcategories				0.33	0.29	< .01
Supine isolated OSA	17 (36%)	12 (25%)	18 (40%)			
Supine predominant OSA	17 (36%)	18 (38%)	14 (31%)			
No positional OSA	15 (31%)	18 (38%)	13 (29%)			
Category changes from previous night		22 (46%)	18 (40%)			
REM-associated OSA				0.45	0.22	< .01
Yes	17 (33%)	18 (37%)	22 (45%)			
No	34 (67%)	31 (63%)	27 (55%)			
Category changes from previous night		19 (39%)	15 (31%)			

Number and percent of patients are given. Positional OSA is defined as a ratio of pAHI in supine to pAHI in nonsupine ≥ 2. The subcategories are defined as supine isolated OSA if the ratio of pAHI in supine to pAHI in nonsupine is ≥ 2 and pAHI nonsupine < 5 events/h. Supine predominant OSA as a ratio of pAHI in supine to pAHI in nonsupine ≥ 2 and pAHI nonsupine ≥ 5 events/h. OSA with REM-associated OSA is defined as pAHI REM/pAHI NREM ≥ 2. The criteria for REM-associated and positional OSA were not available in all patients. Therefore, the number of patients in the subgroup analyses do not always add up to 51. Note that the *P* value reports the statistical significance of *K*, not the agreement between the nights. NREM = non–rapid eye movement, OSA = obstructive sleep apnea, pAHI = peripheral arterial tonometry–derived apnea-hypopnea index, REM = rapid eye movement.

second and second to third night, respectively. This was caused solely by NNV without any intervention having taken place between the nights.

pAHI in supine and nonsupine positions, as well as pAHI in REM and NREM, showed a sizable variability (**Figure S1** and **Figure S2** in the supplemental material). The categorization and the agreement of positional and REM-associated OSA are given in **Table 3**. Over all nights, mean supine pAHI was 26.4 ± 21.6 events/h and mean nonsupine pAHI was 8.4 ± 8.9 events/h. The average pAHI over 3 nights for REM sleep and NREM sleep was 18.6 ± 14.7 events/h and 12.8 ± 11.4 events/h, respectively. The difference between supine and nonsupine pAHI was 16 events/h and significantly higher compared to the difference between pAHI of REM and NREM sleep of 5.8 events/h. REM proportion of total sleep time was not statistically different over the 3 nights and, on average, 23.2 ± 6.5%.

The intraclass correlation coefficient of average measures for pAHI showed a good agreement between the nights with of 0.87 (95% CI 0.79–0.92). The SEM of pAHI calculated according to Fleiss and Kingman[33] was 6.9 event/h (95% CI, 5.5–8.3) when analyzing 2 nights and 6.5 events/h (95% CI, 5.2–7.7) when all 3 measurements were included. We calculated the minimal detectable difference of pAHI for 2 and 3 measurements as 19.1 events/h and 18.0 events/h, respectively. The night-to-night variability was not influenced by the scoring method used. Automatically scored data yielded comparable results to manually scored data. Detailed results for automatically scored data are given in **Table S1**. Oxygen desaturation index showed a similar NNV compared to pAHI (**Figure S3**) with an intraclass correlation coefficient of 0.83 (95% CI, 0.73–0.90) and an SEM of 5.3 (95% CI, 4.3–6.4).

pAHI increased significantly with time in supine decubitus. For every minute spent in the supine position, pAHI increased by 0.04 events/h. The sequence of nights had no influence on time spent in the supine position. Further analysis using the linear mixed-effects model showed no significant influence of any other sleep parameter on NNV.

No sleep parameter showed a significant difference between the 3 nights (**Table 1**). The only difference between the nights was decreasing recording time and total sleep time. This finding can be explained by the fact that most of the recordings were performed during the weekend with patients sleeping longer on Friday and Saturday nights, but slept shorter on Sunday since they had to rise earlier on Monday. The mean end time of the recordings on the third night was considerably earlier at 6:06 AM compared to 7:05 AM and 6:55 AM for the first and second night, respectively. The mean pAHI of the nights was 15.0 ± 12.0 events/h, 14.8 ± 11.6 events/h, 13.9 ± 11.0 events/h, sequentially.

A linear mixed-effects model showed that the sequence of nights had no significant influence on pAHI or any other sleep parameter, finding no evidence for an FNE in PAT.

DISCUSSION

To our knowledge, this is the first study investigating the NNV for peripheral-arterial tonometry (WatchPAT). We found a similar NNV using PAT compared to previous studies using PSG and respiratory polygraphy.[4,6,10,11,20] This indicates that NNV is a general phenomenon occurring during sleep regardless of the applied recording method. We found that the NNV was not influenced by the scoring method. In our collective, the intraclass correlation coefficient of pAHI was 0.87, which is considered a good agreement and compares well to values found in the literature for PSG and respiratory polygraphy.[3,20] In 24% of measurements, the OSA severity was misclassified compared to the average of all nights. This misclassification rate is corroborated by similar findings of other studies.[6–8,21]

The Sher et al[27] criteria are commonly accepted for the assessment of responders of a given treatment. In our study, 12–14% of patients fulfilled these criteria on a subsequent night. Since no intervention has been performed a substantial portion of patients would be falsely considered responders simply due to variability in OSA severity. This remarkable finding has to be considered when evaluating responder rates which were obtained from single recordings.

Our finding of the SEM and minimal detectable difference for PAT are comparable to values found for respiratory polygraphy and PSG.[4,10] In our study, the SEM of all pAHI measurements was 6.4 events/h and MDD 18 events/h. Anitua et al[10] found an SEM of 4.6 events/h for repeated measures using respiratory polygraphy. Aarab et al[4] found for PSG an SEM of 4.6 events/h and MDD of 12.8 events/h. This suggests that assessing OSA severity is accompanied by a major uncertainty due to NNV and large differences are required to reach statistical significance. Multiple testing could slightly reduce this uncertainty. These findings have a major impact on the assessment of therapeutic effects and should be implemented in clinical guidelines, which is currently not the case.

NNV could be partially explained by time spent in supine decubitus. Our model suggests an increase of pAHI by 0.04 events/h for every minute spent in the supine position using PAT. Other studies have found similar positional effects.[15] Interestingly, the positional effect seems to be more pronounced in mild to moderate OSA, whereas severe OSA is less influenced by the sleeping position.[31,35] Pevernagie and Shepard[29] found that patients with strong positional OSA tend to avoid supine sleeping, possibly to reduce the time spent in the position with the most sleep disruptions. In our study, patients changed their categorization of positional and REM-associated OSA over 3 nights in 48% and 55%, respectively. Therefore, positional and REM-associated effects contributing to NNV are highly variable from night-to-night and possibly neutralizing each other. Therefore, caution may be necessary when diagnosing positional OSA from a single night recording.

In the literature, several factors have been shown to contribute to NNV such as time spent in the supine position, OSA severity, and FNE.[14,15,36] However, when using a linear mixed-effects model, we could not identify other factors than the supine time that significantly contribute to the NNV of OSA severity. However, NNV is not only influenced by sleep parameters, but also by patient characteristics such as smoking, emotional stress due to financial problems, obesity, and rostral fluid shift.[12,16] In conclusion, NNV is multifactorial and might not

be easily described by a model relying only on sleep parameters derived from sleep recordings.

The FNE is a phenomenon that has been documented for PSG in various studies.[6,13,22] FNE is characterized by a reduced total sleep time, decreased sleep efficiency, an increased proportion of light sleep, less REM sleep, and more sleep stage transitions, indicating a more fragmented sleep.[22] FNE may be attributed to the unfamiliar surroundings of a sleep lab, but possibly also to the measuring method itself. The latter assumption is corroborated by a study using PSG at home over 4 consecutive nights.[18] It revealed an FNE, albeit smaller than that described for in-lab PSG affecting mainly REM sleep with an adaptation process extending up to the fourth night. Moreover, a study with mostly older people with insomnia showed that anxiety may also play a role in FNE.[37] No FNE has been found for respiratory polygraphy.[38] In our study, we found no evidence of an FNE with PAT. PAT measurements are performed at home. The PAT device does not need a nasal cannula and does not hinder natural nightly movements. The familiar sleep environment and minimal disturbance could explain the absence of an FNE in PAT. PAT allows the measurement of REM sleep and we found no evidence, that PAT affected REM sleep in any way. Our findings are of interest especially in comparison to those by Le Bon et al[18], who found an FNE with PSG at home, particularly for REM sleep adaptation.

This study has several limitations. Our patient collective was rather homogenous, predominantly middle-aged not obese men with mild to moderate OSA and few comorbidities. This leads to fewer confounding factors but limits the generalizability of this study. In the patient collective were only 2 women, limiting the application of the results to female patients. Although PAT has good accordance with PSG, the latter remains the gold standard in OSA diagnosis. Manual scoring of PAT recordings can significantly improve the concordance with PSG.[26] In our study, manual scoring was performed by an experienced sleep technician to minimize interrater variability and improve comparability with PSG. Scoring was performed only once, and therefore rescoring reliability cannot be reported. The analysis of automatically scored data showed no significant difference regarding NNV. Most measurements were conducted over the weekend. Alcohol consumption and sleep medication were not recorded or controlled. Furthermore, the WatchPAT 200 was used for this study, which cannot differentiate between obstructive and central apneas.

Several studies analyzing the NNV in OSA diagnostics conclude that multiple measurements might be necessary.[8,11,13,15,23] This is especially recommended when using a method with a FNE because OSA severity is mostly underestimated on the first night of testing.[36] Repeated measurements are costly and resource intensive. Sleep laboratories often have waiting times, leading to delayed diagnosis. Repeated measurements would compound this problem.

The authors see multiple HSAT using PAT as a viable alternative in selected patients with a high pretest probability of OSA and no neurological or cardiovascular comorbidities influencing sleep patterns. OSA needs long-term treatment, which makes the correct assessment of OSA severity especially important to avoid costly overtreatment or undertreatment. Based on only 1 night of recording, the misclassification rate of OSA severity is unacceptably high. The number of recordings recommended should depend on the clinical situation. In a patient with severe OSA and corresponding symptoms, additional measurements might be unnecessary because of lacking clinical consequences. In patients with diverging sleep study results and clinical symptoms, multiple testing allows for a better estimation of the individual patient's OSA characteristics. In our opinion, the diagnosis of positional OSA should be substantiated by the recording of multiple nights before recommending positional treatment. Repeated measurements using PAT offer a cost-effective way in the diagnostic work-up of OSA patients.

In summary, we found similar and sizable night-to-night variability comparable to previous studies using PSG or respiratory polygraphy for sleep apnea testing. For PAT, there is no evidence of a FNE. The considerable variability has to be born in mind when counseling patients for possible treatment options, reporting the success of therapeutic interventions and designing scientific research. This variability has to be considered when reporting treatment success due to the sizable proportion of false-positive responders to treatment. This aspect is also important when interpreting study results. We believe that repeated, nondisturbing measurements in the familiar sleep surroundings of a patient's home are a promising way to quantify sleep better and diagnose obstructive sleep apnea.

ABBREVIATIONS

AHI, apnea-hypopnea index
CI, confidence interval
FNE, first night effect
HSAT, home sleep apnea testing
MDD, minimal detectable difference
NNV, night-to-night variability
NREM, non–rapid eye movement
OSA, obstructive sleep apnea
pAHI, peripheral arterial tonometry–derived apnea-hypopnea index
PAT, peripheral arterial tonometry
PSG, polysomnography
REM, rapid eye movement
SEM, standard error of measurement

REFERENCES

1. Remmers JE, deGroot WJ, Sauerland EK, Anch AM. Pathogenesis of upper airway occlusion during sleep. *J Appl Physiol*. 1978;44(6):931–938.
2. Berry RB, Quan SF, Abreu AR, et al; for the American Academy of Sleep Medicine. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Version 2.6. Darien, IL: American Academy of Sleep Medicine; 2020.
3. Prasad B, Usmani S, Steffen AD, et al. Short-term variability in apnea-hypopnea index during extended home portable monitoring. *J Clin Sleep Med*. 2016;12(6):855–863.
4. Aarab G, Lobbezoo F, Hamburger HL, Naeije M. Variability in the apnea-hypopnea index and its consequences for diagnosis and therapy evaluation. *Respiration*. 2009;77(1):32–37.

5. Fietze I, Glos M, Zimmermann S, Penzel T. Long-term variability of the apnea-hypopnea index in a patient with mild to moderate obstructive sleep apnea. *J Clin Sleep Med*. 2020;16(2):319–323.
6. Bittencourt LRA, Suchecki D, Tufik S, et al. The variability of the apnoea-hypopnoea index. *J Sleep Res*. 2001;10(3):245–251.
7. Levendowski DJ, Zack N, Rao S, et al. Assessment of the test-retest reliability of laboratory polysomnography. *Sleep Breath*. 2009;13(2):163–167.
8. Levendowski D, Steward D, Woodson BT, Olmstead R, Popovic D, Westbrook P. The impact of obstructive sleep apnea variability measured in-lab versus in-home on sample size calculations. *Int Arch Med*. 2009;2(1):2.
9. Gouveris H, Selivanova O, Bausmer U, Goepel B, Mann W. First-night-effect on polysomnographic respiratory sleep parameters in patients with sleep-disordered breathing and upper airway pathology. *Eur Arch Otorhinolaryngol*. 2010;267(9):1449–1453.
10. Anitua E, Duran-Cantolla J, Almeida GZ, Alkhraisat MH. Predicting the night-to-night variability in the severity of obstructive sleep apnea: the case of the standard error of measurement. *Sleep Sci*. 2019;12(2):72–78.
11. Stepnowsky CJ Jr, Orr WC, Davidson TM. Nightly variability of sleep-disordered breathing measured over 3 nights. *Otolaryngol Head Neck Surg*. 2004;131(6):837–843.
12. Zheng H, Sowers M, Buysse DJ, et al. Sources of variability in epidemiological studies of sleep using repeated nights of in-home polysomnography: SWAN Sleep Study. *J Clin Sleep Med*. 2012;8(1):87–96.
13. Sforza E, Roche F, Chapelle C, Pichot V. Internight variability of apnea-hypopnea index in obstructive sleep apnea using ambulatory polysomnography. *Front Physiol*. 2019;10:849.
14. Stöberl AS, Schwarz EI, Haile SR, et al. Night-to-night variability of obstructive sleep apnea. *J Sleep Res*. 2017;26(6):782–788.
15. Yalciner G, Bademez MA, Gul F. Association of sleep time in supine position with apnea-hypopnea index as evidenced by successive polysomnography. *Sleep Breath*. 2017;21(2):289–294.
16. White LH, Lyons OD, Yadollahi A, Ryan CM, Bradley TD. Night-to-night variability in obstructive sleep apnea severity: relationship to overnight rostral fluid shift. *J Clin Sleep Med*. 2015;11(2):149–156.
17. McCall C, McCall WV. Objective vs. subjective measurements of sleep in depressed insomniacs: first night effect or reverse first night effect? *J Clin Sleep Med*. 2012;8(1):59–65.
18. Le Bon O, Staner L, Hoffmann G, et al. The first-night effect may last more than one night. *J Psychiatr Res*. 2001;35(3):165–172.
19. Sharpley AL, Solomon RA, Cowen PJ. Evaluation of first night effect using ambulatory monitoring and automatic sleep stage analysis. *Sleep*. 1988;11(3):273–276.
20. Quan SF, Griswold ME, Iber C, et al.Sleep Heart Health Study (SHHS) Research Group. Short-term variability of respiration and sleep during unattended nonlaboratory polysomnography–the Sleep Heart Health Study. [corrected] *Sleep*. 2002;25(8):843–849.
21. Ahmadi N, Shapiro GK, Chung SA, Shapiro CM. Clinical diagnosis of sleep apnea based on single night of polysomnography vs. two nights of polysomnography. *Sleep Breath*. 2009;13(3):221–226.
22. Newell J, Mairesse O, Verbanck P, Neu D. Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples. *Psychiatry Res*. 2012;200(2-3):795–801.
23. Meyer TJ, Eveloff SE, Kline LR, Millman RP. One negative polysomnogram does not exclude obstructive sleep apnea. *Chest*. 1993;103(3):756–760.
24. Yalamanchali S, Farajian V, Hamilton C, Pott TR, Samuelson CG, Friedman M. Diagnosis of obstructive sleep apnea by peripheral arterial tonometry: meta-analysis. *JAMA Otolaryngol Head Neck Surg*. 2013;139(12):1343–1350.
25. Hedner J, White DP, Malhotra A, et al. Sleep staging based on autonomic signals: a multi-center validation study. *J Clin Sleep Med*. 2011;7(3):301–306.
26. Zhang Z, Sowho M, Otvos T, et al. A comparison of automated and manual sleep staging and respiratory event recognition in a portable sleep diagnostic device with in-lab sleep study. *J Clin Sleep Med*. 2020;16(4):563–573.
27. Sher AE, Schechtman KB, Piccirillo JF. The efficacy of surgical modifications of the upper airway in adults with obstructive sleep apnea syndrome. *Sleep*. 1996;19(2):156–177.
28. Cartwright RD. Effect of sleep position on sleep apnea severity. *Sleep*. 1984;7(2):110–114.
29. Pevernagie DA, Shepard JW Jr. Relations between sleep stage, posture and effective nasal CPAP levels in OSA. *Sleep*. 1992;15(2):162–167.
30. Joosten SA, O'Donoghue FJ, Rochford PD, et al. Night-to-night repeatability of supine-related obstructive sleep apnea. *Ann Am Thorac Soc*. 2014;11(5):761–769.
31. Mador MJ, Kufel TJ, Magalang UJ, Rajesh SK, Watwe V, Grant BJB. Prevalence of positional sleep apnea in patients undergoing polysomnography. *Chest*. 2005;128(4):2130–2137.
32. Oksenberg A, Arons E, Nasser K, Vander T, Radwan H. REM-related obstructive sleep apnea: the effect of body position. *J Clin Sleep Med*. 2010;6(4):343–348.
33. Fleiss JL, Kingman A. Statistical management of data in clinical research. *Crit Rev Oral Biol Med*. 1990;1(1):55–66.
34. Smith S. Reporting the apnoea hypopnea index (AHI). *J Clin Sleep Med*. 2007;3(3):321.
35. Mador MJ, Choi Y, Bhat A, et al. Are the adverse effects of body position in patients with obstructive sleep apnea dependent on sleep stage? *Sleep Breath*. 2010;14(1):13–17.
36. Le Bon O, Hoffmann G, Tecco J, et al. Mild to moderate sleep respiratory events: one negative night may not be enough. *Chest*. 2000;118(2):353–359.
37. Riedel BW, Winfield CF, Lichstein KL. First night effect and reverse first night effect in older adults with primary insomnia: does anxiety play a role? *Sleep Med*. 2001;2(2):125–133.
38. Teschler H, Hoheisel G, Schumann H, Wagner B, Konietzko N. [Validation of the Sleep-Doc-Porti system for ambulatory sleep apnea diagnosis]. *Pneumologie*. 1995;49(9):496–501.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication December 2, 2020
Submitted in final revised form March 18, 2021
Accepted for publication March 18, 2021
Address correspondence to: Samuel Tschopp, Department of Otorhinolaryngology, Head and Neck Surgery, Inselspital, University Hospital and University of Bern; Email: samuel.tschopp@insel.ch

DISCLOSURE STATEMENT

All authors have seen and approved this manuscript. The authors report no conflicts of interest.