SCIENTIFIC INVESTIGATIONS

Residual Events during Use of CPAP: Prevalence, Predictors, and Detection Accuracy

Joel Reiter, MD[1,2]; Bashar Zleik, MD[1,3]; Mihaela Bazalakova, MD, PhD[1,4]; Pankaj Mehta, MD[1,5]; Robert Joseph Thomas, MD, MMSc[1]

[1]Sleep Disorders Clinic, Departments of Medicine & Neurology, Beth Israel Deaconess Medical Center, Boston, MA; [2]Pediaric Pulmonary Unit, Department of Pediatrics, Hadassah-Hebrew University Medical Center, Jerusalem, Israel; [3]Department of Neurology, Staten Island University Hospital, Hofstra Northwell School of Medicine, Staten Island, NY; [4]Department of Neurology, Center for Sleep Medicine and Sleep Research, University of Wisconsin-Madison, Madison, WI; [5]Pulmonary Critical Care and Sleep, Covenant Clinic, Waterloo, IA

Study Objectives: To assess the frequency, severity, and determinants of residual respiratory events during continuous positive airway therapy (CPAP) for obstructive sleep apnea (OSA) as determined by device output.

Methods: Subjects were consecutive OSA patients at an American Academy of Sleep Medicine accredited multidisciplinary sleep center. Inclusion criteria included CPAP use for a minimum of 3 months, and a minimum nightly use of 4 hours. Compliance metrics and waveform data from 217 subjects were analyzed retrospectively. Events were scored manually when there was a clear reduction of amplitude ($\geq 30\%$) or flow-limitation with 2–3 larger recovery breaths. Automatically detected versus manually scored events were subjected to statistical analyses included Bland-Altman plots, correlation coefficients, and logistic regression exploring predictors of residual events.

Results: The mean patient age was 54.7 ± 14.2 years; 63% were males. All patients had a primary diagnosis of obstructive sleep apnea, 26% defined as complex sleep apnea. Residual flow measurement based apnea-hypopnea index (AHI_{FLOW}) > 5, 10, and 15/h was seen in 32.3%, 9.7%, and 1.8% vs. 60.8%, 23%, and 7.8% of subjects based on automated vs. manual scoring of waveform data. Automatically detected versus manually scored average AHI_{FLOW} was 4.4 ± 3.8 vs. 7.3 ± 5.1 per hour. In a logistic regression analysis, the only predictors for a manual AHI_{FLOW} > 5/h were the absolute central apnea index (CAI), (odds ratio [OR]: 1.5, p: 0.01, CI: 1.1–2.0), or using a CAI threshold of 5/h of sleep (OR: 5.0, p: < 0.001, CI: 2.2–13.8). For AHI_{FLOW} > 10/h, the OR was 1.14, p: 0.03 (CI: 1.1–1.3) per every CAI unit of 1/hour.

Conclusions: Residual respiratory events are common during CPAP treatment, may be missed by automated device detection and predicted by a high central apnea index on the baseline diagnostic study. Direct visualization of flow data is generally available and improves detection.

Keywords: sleep, auto-CPAP, residual apnea

Citation: Reiter J, Zleik B, Bazalakova M, Mehta P, Thomas RJ. Residual events during use of CPAP: prevalence, predictors, and detection accuracy. *J Clin Sleep Med* 2016;12(8):1153–1158.

INTRODUCTION

Obstructive sleep apnea is a common disorder affecting both men and women and leading to considerable morbidity and mortality.[1] Continuous positive airway pressure (CPAP) is the treatment of choice.[2] Modern positive airway pressure (PAP) devices measure and store airflow and pressure data, displaying the latter as residual event and compliance indices.[3] This allows for tracking of both efficacy and adherence.[2] However, vendor algorithms vary, and there are no specific guidelines or standards for capturing, measuring, or scoring the data.[2] High-resolution flow data can also be reviewed directly, enabling visual/manual assessment of events.

Several studies have examined the relationship between device detected events or devices reported apnea-hypopnea index (AHI) based on flow measurements (AHI_{FLOW}) and findings on polysomnography.[3–9] The findings mostly demonstrate good correlation,[5–8] one study showing device overestimation.[4] Event-by-event analysis in one study showed that the automatic detection had high specificity but only modest sensitivity (for a cutoff of AHI > 10/h, sensitivity was 0.58, and specificity 0.94), with good agreement for apneas but less so for detecting

BRIEF SUMMARY

Current Knowledge/Study Rationale: Initiation and adjustment of PAP devices in automatic mode is generally accepted as the standard of care for treatment of OSA without significant comorbidities. Device settings are adjusted remotely, based on event data from these devices; however, waveform data available is seldom reviewed. We sought to compare automatically detected versus manually scored events to analyze the accuracy of devices and identify predictors of residual AHI_{FLOW}.

Study Impact: It is demonstrated that direct review of device generated flow data identifies inadequate control, otherwise missed with sole review of automatic device detection indices. The baseline study central sleep apnea index is the main predictor of high AHI_{FLOW}.

hypopneas,[3] an automatic detection limitation also seen in other studies.[6,9] There is thus a concern that efficacy can be overestimated by the auto-titrating device.

We know of no studies comparing manually scored vs automated device-detected residual apnea-hypopnea indices, based on high resolution device-derived flow/waveform compliance data. Using the device output data, we manually scored respiratory events, comparing these with the number of

device-detected events. Based on our prior clinical experience it was our hypothesis that (1) the device algorithm captures far fewer events than those apparent on manual analysis of flow data, and (2) predictors of residual events may be identified.

METHODS

Database

The study was conducted at the Beth Israel Deaconess Medical Center (BIDMC), Boston, Massachusetts, USA. Institutional board review approval was part of a general protocol enabling review of archived sleep data. From mid-2012, the center's policy changed to indefinite PAP data tracking: the data modem was left active/reactivated continuously. This resulted in a weekly database acquisition of high-resolution flow waveform data, in addition to traditional compliance, automated AHI_{FLOW}, and leak information. Patients who were already on stable therapy were re-supplied with modems for entry into this long-term tracking paradigm. This EncoreAnywhere database was queried for data between January 1, 2013, and June 30, 2013. At the onset of this study the only devices capable of generating and storing waveform data in an online database were Phillips Respironics devices. These included the REMstar Auto, BiPAP Pro, BiPAP Auto, and BiPAP AutoSV Advanced.

Subject Selection and Supportive Data

Inclusion criteria included PAP compliance for ≥ 3 continuous months, with average nightly use ≥ 4 h (all nights). The most recently available single-night, high-resolution flow data samples were divided equally, and randomly, for manual event scoring among 5 sleep physicians at the BIDMC sleep clinic (RT, BZ, JR, MB, and PM). Datasets with high leak rates were excluded. Additional baseline data was collected from the patients' electronic medical records, including: age, gender, race, body mass index (BMI), comorbidities, and medications; baseline polysomnogram (PSG) data and CPAP titration data when available; treatment data including device, mode, vented vs. non-vented mask (for complex apnea), and oxygen supplementation. Complex sleep apnea patients were recognized by the clinical note and PSG report in accordance with the International Classification of Sleep Disorders, Third Edition.[10]

Polysomnography

All polysomnograms were performed at an AASM accredited sleep center, and used standard scoring criteria for sleep stages and respiratory events. For our analysis, we used the AHI where hypopneas were associated with 3% oxygen desaturation and/or an arousal.

Manual Scoring of Respiratory Events

The EncoreAnywhere database generates a Portable Document Format (pdf) of the detailed high-resolution flow data ("waveforms"). Each horizontal line contains 6 minutes of data. Events are tagged automatically when detected, including closed and open airway apnea, hypopnea, vibratory snoring, and periodic breathing. The data is "as is" and not amenable to compression or other forms of manipulation—the system is essentially a viewer. The investigators counted the number of automatically scored events and scored additional events manually using modifications of standard criteria (AHI_{FLOW}). Events were scored when there was a clear reduction of signal amplitude (≥ 30%) or clear flow-limitation with 2–3 larger recovery breaths. We did not try to differentiate apnea from hypopnea, and did not score periodic breathing periods, but tagged sample nights as periodic breathing present/absent. **Figure 1** shows a representative waveform sample, demonstrating automatically device-detected events and missed, unscored, events. **Figure 2** shows a period of stable breathing form the same patient. Interscorer variability was assessed by 5 randomly selected flow data samples that were given to all scorers.

Statistical Methods

Statistical analysis included summary measures (mean, standard deviation), Bland-Altman plots, and logistic regression for prediction of high residual AHI_{FLOW}. Analysis was performed using STATA 12.

RESULTS

Subjects and Demographics

A total of 217 consecutive subjects with laboratory polysomnogram and high resolution compliance flow data were included. One hundred one patients who were in the database during this period were excluded from analysis as home sleep studies preceded the use of auto-CPAP. The mean patient age was 54.7 ± 14.2 years; there were 63% males, 84% Caucasians, and mean BMI was 33.1 ± 7.7. All patients had a primary diagnosis of obstructive sleep apnea, 26% defined as complex sleep apnea. Comorbid conditions included hypertension in 45.6%, diabetes in 13.8%, heart failure in 3.7%, and ischemic heart disease in 8.2%. The average use of CPAP was 218 ± 32 days.

Polysomnography

Diagnostic polysomnography showed a total sleep time (TST) of 278.9 ± 144.1 min, sleep efficiency (SE): 76.7% ± 16.3%, N1: 19.3% ± 15.9%, N3: 12.3% ± 11.7%, REM sleep: 11% ± 9.1%. The diagnostic AHI_{ALT} using arousal and/or 3% oxygen desaturation hypopnea scoring criteria, was 41.7 ± 31.6 per hour of sleep. The respiratory disturbance index (RDI, including apnea hypopnea and respiratory effort related arousals) was 60.2 ± 31.8 per hour of sleep. The average central apnea index (CAI) on the diagnostic assessment was 3.1 ± 8 per hour of sleep. No patient fulfilled conventional criteria for central sleep apnea (CAI ≥ 5/h of sleep, and ≥ 50% of all events). Thus, these patients would be commonly characterized as obstructive sleep apnea patients. The polysomnographic features of diagnostic, titration, or split-night studies are summarized in **Table 1**. The only difference on the baseline polysomnogram between subjects with versus without high AHI_{FLOW} on treatment, were seen in compliance AHI_{FLOW} categories of ≥ 5 and 10/h, an elevated CAI: 0.6 ± 1.6/h vs. 5.3 ± 10.5/h (p = 0.004) and 1.3 ± 3.7/h vs. 8.5 ± 13.5/h (p < 0.001).

Figure 1—Snapshot from the EncoreAnywhere database, generated from high resolution flow ("waveform") data.
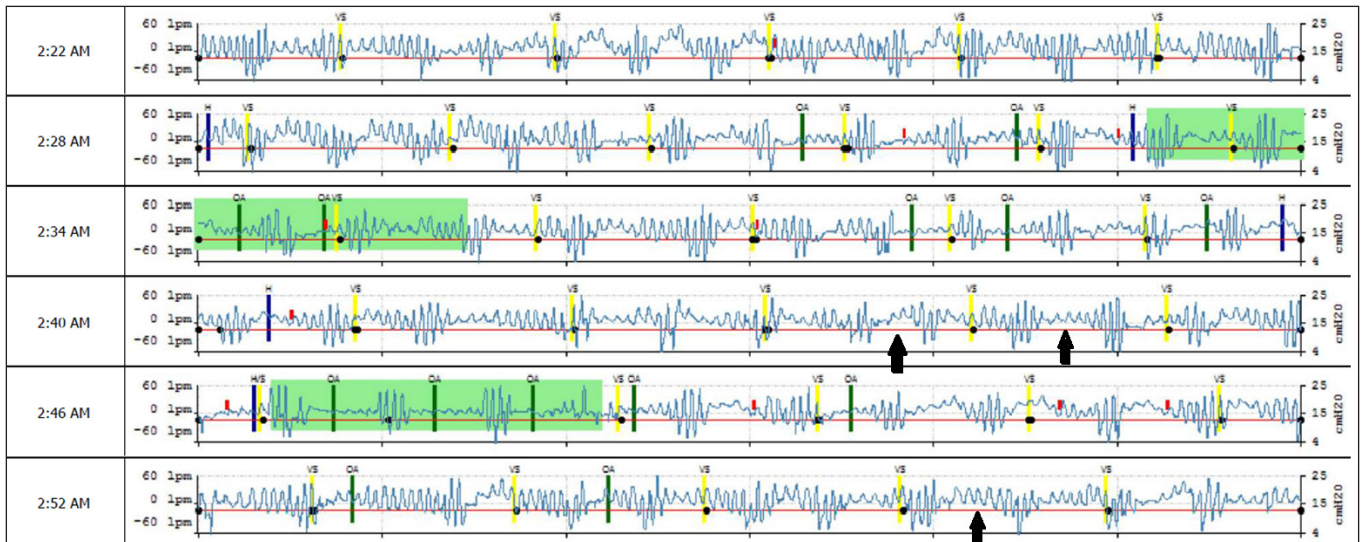
Each horizontal line is 6 minutes long. The red line is the pressure output, with a treatment auto-CPAP range set at 10–15 cm H_2O in this 46-year-old male, without heart failure, but severe periodic breathing and obstructions on the baseline assessment. Increasing or narrowing the pressure range of the auto-CPAP had minimal effect on the residual device-detected respiratory event indices. Highlighted in green is machine detected periodic breathing; green bars represent machine-detected obstructive apneas, yellow bars designate "vibratory snoring." Note large numbers of missed overt severe hypopneas/apneas (arrows) and periodic breathing with relatively short (approximately 20 seconds) cycle length.

Figure 2—Snapshot from the EncoreAnywhere database, generated from high-resolution flow ("waveform") data.
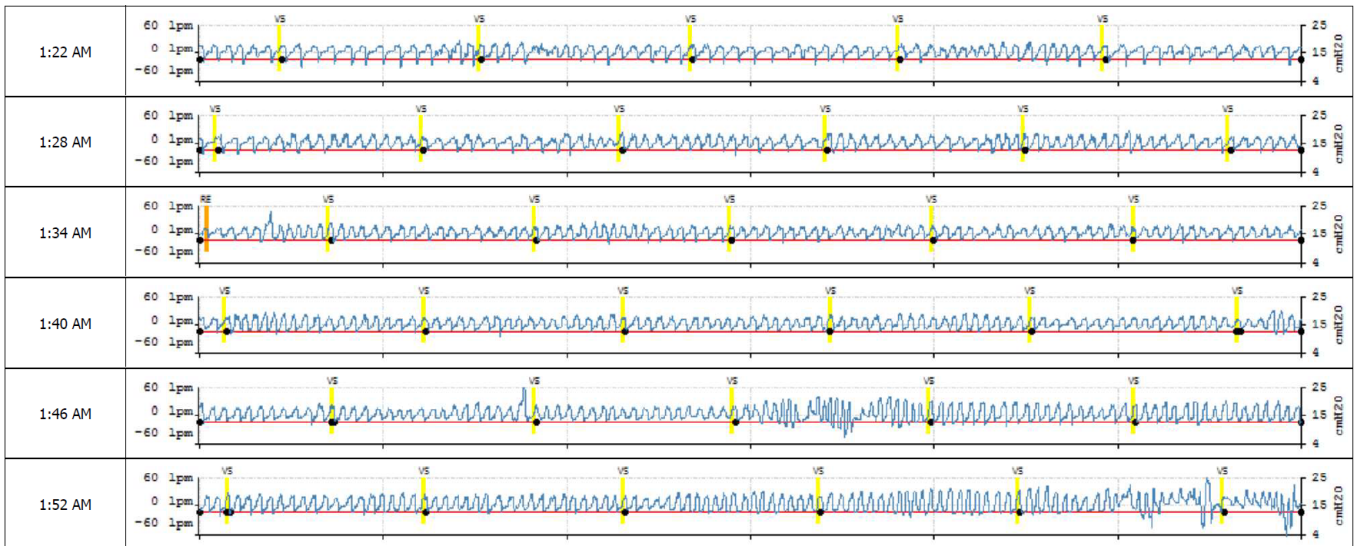
Same patient as in Figure 1, a period of stable breathing, showing the contrast and relative ease of visual recognition of stable breathing periods (with flow limitation in this instance) and periods with respiratory events. VS is machine estimated vibratory snoring.

Database Analysis

The average nightly CPAP use based on waveform compliance data was 6.3 ± 1.5 h, averaged over the previous 4 weeks. CPAP was used in the fixed mode in 76 (35%) subjects, although all subjects were treated with an auto-capable device. There were no significant differences in residual apnea-hypopnea indices in auto vs. fixed mode: 4.4 ± 4.1, p = 0.94 and 7.4 ± 5.4, p = 0.76, events/h of use. Interscorer variability of manual scoring of events revealed an agreement of 0.96 among the 5 scorers.

Using the AHI ≥ 5, 10, or 15 events/h thresholds, 70 (32.3%), 21 (9.7%), and 4 (1.8%) subjects showed residual apnea-hypopnea as estimated by the device. Using the same thresholds, 132 (60.8%), 50 (23%), and 17 (7.8%) had residual apnea-hypopnea by manual scoring. Periodic breathing was detected in 14 patients by the device algorithm and 42 patients by visual scanning of the plots. The mean device-detected AHI_{FLOW} was 4.4 ± 3.8/h and manually scored AHI_{FLOW} was 7.3 ± 5.1/h of use.

Table 1—Polysomnogram (PSG) variables: diagnostic and titration studies.

Measure	Diagnostic PSG	Titration PSG
Total sleep time (min)	278.9 ± 144.1	289.2 ± 91.8
Sleep efficiency (%)	76.7 ± 16.3	77 ± 14.7
N1 (%)	19.4 ± 15.9	11.1 ± 7.6
N3 (%)	12.4 ± 11.8	13.4 ± 11
REM (%)	11 ± 9.1	20.5 ± 10.6
Arousal index (events/h)	32.1 ± 31.2	19.3 ± 14.9
RDI (apnea, hypopnea, and RERA/h)	60.2 ± 31.8	25.9 ± 18.4
AHI_{ALT} (events/h)	41.7 ± 31.6	15.9 ± 15.5
$AHI_{4\%}$ (events/h)	27 ± 28.2	4.3 ± 7.1
Central apnea index (events/h)	3.5 ± 8	3.5 ± 6.5
Oxygen 3% desaturation index (events/h)	24.5 ± 25.3	8.7 ± 8.8
Minimum saturation (%)	81.3 ± 8.3	87.3 ± 5.7
Time under 90% saturation (min)	5 ± 11.4	5 ± 19.7
Periodic limb movement index (events/h)	17.8 ± 27	15.1 ± 24.1
PLM arousal index (events/h)	8.2 ± 15.5	4.1 ± 9

Predictors of Residual Apnea-Hypopnea

Logistic regression was used to explore the database for predictors of high residual AHI_{FLOW}. The central apnea index (CAI) on the baseline study was the only predictor of manual $AHI_{FLOW} \geq 5$/h of use, with an odds ratio (OR) of 1.5, p = 0.01 (CI: 1.1–2.0), per hour. That is, for every increase in CAI of 1/h, the risk of residual $AHI_{FLOW} \geq 5$/h increased by 50%. Using a baseline CAI ≥ 5/h, the OR was 5.0, p < 0.001 (CI: 2.2–13.8) if CAI ≥ 5/h threshold. The CAI was the only predictor for a manually-scored $AHI_{FLOW} > 5$/h. For manual $AHI_{FLOW} > 10$/h, OR was 1.14, p = 0.03 (CI: 1.1–1.3) per 1 CAI/h of sleep. This result did not change when adjusted for age, gender, race, baseline N1, SE, and diagnostic AHI.

Auto vs. Manual Scoring Agreements

The Bland-Altman comparison of auto and manual scoring showed the limits of agreement (Reference Range for difference): −11.4 to 5.5, a mean difference: −2.897 (CI: −3.465 to −2.329), range: 0.48 to 21.36. The Pitman test of difference in variance was: r = −0.357, n = 217, p < 0.001.

DISCUSSION

In our assessment of flow data from a CPAP device several months after the initiation of therapy, we found that: (1) residual respiratory abnormality is common; (2) device efficiency is overestimated leading to missed inadequate control in some; (3) the central apnea index on the baseline polysomnogram is the only predictor of high residual AHI_{FLOW}.

This is the first study analyzing the high resolution data available from modern PAP devices comparing visual/manual scoring to automated device scoring. Our results show that residual respiratory events are common with stable compliant CPAP use. In this study automated scoring detected approximately half of the patients who would have qualified as having mild residual OSA (32.3% vs. 60.8% by manual scoring), and less than a quarter of patients with moderate to severe residual OSA ($AHI_{FLOW} \geq 15$/h; 1.8% vs. 7.8%). Thus, despite the narrow difference found between automated and manual AHI_{FLOW} (4.4 ± 3.8 and 7.3 ± 5.1, respectively) a significant number of undertreated patients are missed when automated compliance metrics are examined alone.

Of all parameters investigated, the only significant predictor of a high residual apnea-hypopnea burden was the central apnea index on baseline diagnostic polysomnography. The physiological cause of this finding is not clear and beyond the scope of this study but may suggest causes other than anatomical for the high AHI_{FLOW}. Excessive air leak can cause central apnea during positive pressure titration,[11] and by extension may occur for the same reason during chronic PAP treatment, but was not the cause in our patients as excessive leak was part of our exclusion criteria. However, this finding, along with high AHI on the titration study (**Table 1**) suggest a possible role for treatment emergent sleep apnea and ventilatory instability leading to high residual AHI_{FLOW}, primarily on manual scoring.

Similar to our findings, prior studies that included polysomnographic analysis of device detection, the gold standard, showed overall good correlation but raised concern of overestimation of device efficacy. Huang et al.[9] performed a study of simultaneous first-night auto-CPAP titration along with full polysomnography in the sleep lab. They showed that the diagnostic study arousal index, central apnea index, and history of cardiac disease predicted a high AHI on the treatment PSG in the auto-titration mode. They recommended that patients with even mild central sleep apnea be excluded from auto-CPAP titration.

The ATS 2013 statement on CPAP adherence tracking systems noted that current clinical care systems are not optimally configured for examining this data.[2] As such, those patients in whom this data is likely to be of added benefit, to assess control, and determine the need for treatment parameter adjustments, need to be identified. One such group, as shown in this study, is that with high central apnea index on the baseline study. In addition, it appears intuitive that this data will be

helpful in those with inadequate OSA control, persistent symptoms and low compliance.

The strength of this study lies in its real-life perspective, with patients using their own devices at home, manual scoring, and the selection of patients several months after the start of therapy. We collected data that is readily obtainable from the patients' devices but is mostly overlooked in busy clinical practices, and even in the research setting. While polysomnography offers the most detail, and remains the obvious gold standard, it is often not readily available.

The study however has several limitations. The use of a single PAP provider may be such a limitation, however, in the interim waveform data has also become available from other providers or are viewable using freeware, and appears to show very similar trends (unpublished observation). Flow data in itself provides only limited data with no information on results of decreased flow such as arousals or desaturations, and possible effect of leak (the latter minimized in the study by excluding high leak datasets). While this is a limitation of the data and our results, and the airway can open without an arousal, repeated clusters of recovery breaths preceded by signal reduction or flow limitation on the flow-only signal are highly reminiscent of respiratory events that are routinely scored on polysomnography. Some of the events we scored included what auto-CPAP devices may tag as respiratory effort related arousals. Residual events, manual or automated, occur during periods of the recording usually flanked by periods of stable breathing, where there is minimal change of signal amplitude and morphology over several minutes. Thus, the higher the event number identified, regardless of type, the more likely stable breathing is proportionately reduced. While we did not evaluate clinical outcomes, it is reasonable to speculate that those with less stable breathing are at higher risk of residual symptoms. Complex sleep apnea, which likely plays a role in our findings, has been shown to occur in 2% to 20% of patients,[10] whereas in our population, a tertiary referral center emphasizing the treatment of this disorder, this was found in 26% of patients. This may have led to an overestimation and impact of the generalizability of our findings. Our assessment focused on data rather than clinical outcomes, and we cannot comment on the impact of residual respiratory events on our patients. We selected patients with high usage to maximize raw data collection for scoring, thus any contribution of residual events to poor compliance cannot be addressed.

Compliance with CPAP remains low despite advances in devices and interfaces. While some reject treatment immediately others, up to 25% in some studies, do so after an initial trial.[12] It is possible that inadequate control contributes to poor clinical response and thereby to non-compliance. Our findings suggest that device auto-detection may lead to underestimation of residual apnea-hypopnea and that this may be overcome by the review of high resolution waveform data.

ABBREVIATIONS

AHI, apnea-hypopnea index

AHI_{ALT}, apnea-hypopnea index using arousal and/or 3% oxygen desaturation hypopnea scoring criteria

AHI_{FLOW}, flow measurement based apnea-hypopnea index

BIDMC, Beth Israel Deaconess Medical Center

BMI, body mass index

CAI, central apnea index

CPAP, continuous positive airway pressure

OR, odds ratio

OSA, obstructive sleep apnea

PAP, positive airway pressure

PSG, polysomnogram

RDI, respiratory disturbance index

SE, sleep efficiency

TST, total sleep time

REFERENCES

1. 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.

2. Schwab RJ, Badr SM, Epstein LJ, et al. An official American Thoracic Society statement: continuous positive airway pressure adherence tracking systems. The optimal monitoring strategies and outcome measures in adults. Am J Respir Crit Care Med 2013;188:613–20.

3. Berry RB, Kushida CA, Kryger MH, Soto-Calderon H, Staley B, Kuna ST. Respiratory event detection by a positive airway pressure device. Sleep 2012;35:361–7.

4. Ueno K, Kasai T, Brewer G, et al. Evaluation of the apnea-hypopnea index determined by the S8 auto-CPAP, a continuous positive airway pressure device, in patients with obstructive sleep apnea-hypopnea syndrome. J Clin Sleep Med 2010;6:146–51.

5. Ikeda Y, Kasai T, Kawana F, et al. Comparison between the apnea-hypopnea indices determined by the REMstar Auto M series and those determined by standard in-laboratory polysomnography in patients with obstructive sleep apnea. Intern Med 2012;51:2877–85.

6. Prasad B, Carley DW, Herdegen JJ. Continuous positive airway pressure device-based automated detection of obstructive sleep apnea compared to standard laboratory polysomnography. Sleep Breath 2010;14:101–7.

7. Desai H, Patel A, Patel P, Grant BJ, Mador MJ. Accuracy of autotitrating CPAP to estimate the residual Apnea-Hypopnea Index in patients with obstructive sleep apnea on treatment with autotitrating CPAP. Sleep Breath 2009;13:383–90.

8. Cilli A, Uzun R, Bilge U. The accuracy of autotitrating CPAP-determined residual apnea-hypopnea index. Sleep Breath 2013;17:189–93.

9. Huang HC, Hillman DR, McArdle N. Control of OSA during automatic positive airway pressure titration in a clinical case series: predictors and accuracy of device download data. Sleep 2012;35:1277–83.

10. American Academy of Sleep Medicine. International classification of sleep disorders, 3rd ed. Darien, IL: American Academy of Sleep Medicine, 2014.

11. Montesi SB, Bakker JP, Macdonald M, et al. Air leak during CPAP titration as a risk factor for central apnea. J Clin Sleep Med 2013;9:1187–91.

12. Engleman HM, Wild MR. Improving CPAP use by patients with the sleep apnoea/hypopnoea syndrome (SAHS). Sleep Med Rev 2003;7:81–99.