HHS Public Access

Author manuscript

Sleep Breath. Author manuscript; available in PMC 2020 September 08.

Published in final edited form as:

Sleep Breath. 2018 December ; 22(4): 1179–1188. doi:10.1007/s11325-018-1735-2.

Addition of frontal EEG to adult home sleep apnea testing: does a more accurate determination of sleep time make a difference?

Matthew P. Light[1,2], **Thalia N. Casimire**[1], **Catherine Chua**[1,3], **Viachaslau Koushyk**[1], **Omar E. Burschtin**[1,3], **Indu Ayappa**[1,3], **David M. Rapoport**[1,3]

[1]Division of Pulmonary, Critical Care, and Sleep Medicine, New York University School of Medicine, New York, NY, USA

[2]Division of Pulmonary, Critical Care & Sleep Medicine, University of California, San Diego, 9300 Campus Point Drive #7381, La Jolla, CA 92037-7381, USA

[3]Division of Pulmonary, Critical Care and Sleep Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Abstract

Rationale—Home sleep apnea testing (HSAT) typically does not include electroencephalogram (EEG) monitoring for sleep assessment. In patients with insomnia and low sleep efficiency, overestimation of the sleep period can result from absence of EEG, which will reduce sleep disordered breathing (SDB) indices and may lead to a false-negative result.

Objective—To validate a single channel frontal EEG for scoring sleep versus wake against full EEG during polysomnography, and then to examine the utility of adding this single channel EEG to standard HSAT to prevent false-negative results.

Methods—Epoch-by-epoch validation for sleep scoring of single channel EEG versus full PSG was first performed in 21 subjects. This was followed by a separate retrospective analysis of 207 consecutive HSATs in adults performed in a university-affiliated sleep center using the Somte (Compumedics) HSAT with one frontal EEG as well as chin EMG, nasal airflow, oxyhemoglobin saturation, respiratory effort, pulse rate, and body position. Each study was scored twice, with ($HSAT_{EEG}$) and without the EEG signal visible ($HSAT_{Polygraphy}$), to calculate AHI4 and RDI and

Matthew P. Light, mlight@ucsd.edu.

Conflict of interest

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or nonfinancial interest (such as personal or professional relationships, affiliations, knowledge, or beliefs) in the subject matter or materials discussed in this manuscript.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

For this type of study formal consent is not required.

This work was previously presented in the form of a thematic poster at the American Thoracic Society annual meeting in Washington, D.C., May 2017.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s1132501817352) contains supplementary material, which is available to authorized users.

the effect on OSA diagnosis and severity. Analyses were repeated in 69 patients with poor sleep suggesting insomnia plus Epworth Sleepiness Scale < 7 as well as in 38 patients ultimately shown to have sleep efficiency < 70% on HSAT with EEG.

Measurements and main results—Single channel and full EEG during polysomnography agreed on sleep versus wake in 92–95% of all epochs. HSAT without EEG overestimated the sleep period by 20% (VST = 440 ± 76 min vs TST = 356 ± 82 min), had a false-negative rate of 8% by AHI4 criteria, and underestimated disease severity in 11% of all patients. Subgroup analysis of patients with subjective poor sleep suggesting insomnia did not change the results. Patients later shown to have low sleep efficiency had lower SDB indices and a 20.8% false negative rate of sleep apnea diagnosis.

Conclusions—Although overall false negative rates using $HSAT_{Polygraphy}$ were moderate, suggesting utility for ruling out OSA, there was a specific subgroup in whom there were significant missed diagnoses. However, we were unable to identify this subgroup a priori.

Keywords

Obstructive sleep apnea; Home sleep apnea testing; Electroencephalogram; OSA diagnosis; Portable monitoring

Introduction

The prevalence of obstructive sleep apnea (OSA), the most common form of sleep disordered breathing, is high and may be rising due to the increase in obesity in the general population [1]. Polysomnography (PSG) remains the diagnostic gold standard and includes an electroencephalogram-defined period of total sleep time (TST) which is used to calculate the AHI4 and RDI. PSG, however, is expensive, burdensome for patients, and can be subject to delay depending on laboratory availability [2]. Home polygraphy (i.e., respiratory monitoring alone) is becoming increasingly available and is considered validated for patients with a high pretest probability of OSA [3, 4]. Benefits include convenience, lower cost, and evaluation of patients in their normal sleep environment [5, 6]. Accordingly, the United States Centers for Medicare and Medicaid Services (CMS) allows prescriptions for continuous positive airway pressure (CPAP) based on results of home sleep apnea testing (HSAT), and the American Academy of Sleep Medicine (AASM) updated its clinical practice guideline to include portable monitoring in patients with high pretest probability of moderate to severe OSA [7–10]. What has been less well-studied is whether polygraphy, i.e., a recording without EEG, provides sufficient diagnostic accuracy to rule out obstructive sleep apnea.

As clinical practice and insurance reimbursement policies have evolved, our experience has been that more patients with significant comorbid medical (and sleep) disorders are funneled towards homebased testing and, according to recommendations from the AASM, may be at high-risk for a false negative result [11]. Included in this population are patients with low or moderate pretest probability of OSA and, of specific interest to this study, patients with insomnia and/or those who report poor or fragmented sleep. In the latter group, distinguishing sleep versus wake may play a critical role in defining the sleep-related

breathing disorder (SRBD) index (AHI4 and RDI, see definitions in “Methods” section) because it affects the denominator, “total sleep time” (TST). Several approaches have been proposed to reduce overestimating the sleep period, including removing areas of invalid oximetry signal and identifying probable sleep/wake from the appearance of the airflow signal or other surrogates for the EEG [12, 13]. While these approaches may reduce the denominator of the SRBD indices, the surrogates may be insufficiently sensitive to detect all periods of wake. New technology now allows HSAT devices to collect portable, albeit sometimes limited, EEG data and measure total sleep time. While there are potential downsides to implementation, such as additional training of sleep technicians and increased scoring time, the calculation of a correct SRBD index has important clinical implications, given that SRBD indices are ultimately used to guide therapy and may be required by insurance to approve CPAP. The present study attempts to quantitate the effect that addition of EEG to HSAT has on diagnosis and severity of OSA in the group currently being tested by this modality. In particular, we examined the effect that addition of EEG to HSAT had on meeting the AHI4 and RDI criteria to rule out OSA.

The specific aims of this study were to determine (1) whether a single frontal lead is a valid surrogate for the conventional EEG montage (used in full PSG) to determine sleep versus wake; (2) whether HSAT performed with addition of a single EEG lead changes the final diagnosis and/or severity of OSA reported; and (3) whether the impact of adding EEG to HSAT is greatest in patients with a pretest report of poor sleep.

Methods

Comparison of a single frontal EEG lead to usual polysomnography to determine sleep versus wake

To determine if a single lead frontal EEG can accurately determine sleep versus wake, nocturnal in-laboratory PSGs were performed with the Fz electrode repositioned to the forehead, simulating a self-applied lead at home. Sleep versus wake (and sleep stage) was then scored from the Fz lead and chin EMG tracings only (limited montage) and results were compared to a simultaneous conventional PSG EEG montage. Two independent sleep technicians scored the limited montage to evaluate reproducibility. Full PSG was scored once by AASM rules [14] using F4, C4, EOG, and EMG tracings. Percent agreement of epochs scored as sleep or wake and between scoring methods was tabulated.

Impact of adding a single frontal EEG lead to home sleep apnea testing

Sleep studies of consecutive adults who underwent home testing via a university-affiliated sleep center (Sleep Medicine Associates of NYC, New York, NY) between September 2013 and August 2014, and using the *SomtePSG* device (Compumedics, Abbotsford, Australia), were analyzed retrospectively by rescoring with different methods. The HSAT device was configured to record one channel of frontal EEG and chin EMG along with standard physiologic variables (nasal cannula airflow, oxyhemoglobin saturation, respiratory effort, pulse rate, and body position). During an initial outpatient evaluation an AASM sleep-boarded physician determined the appropriateness of prescribing HSAT for each patient according to AASM practice parameters [7]. HSATs were also performed in instances where

insurance approval for PSG was denied, but the HSAT was considered appropriate by the sleep physician due to a moderate suspicion of OSA with no other sleep diagnosis in the differential. Because we require a minimum valid recording time for HSATs, only studies with > 2 h if valid recording time were analyzed. Demographic and questionnaire data were collected, and patients were trained to apply the sensors and to use the device during a preliminary daytime session, supplemented by additional video and written material. Each patient then underwent unattended, self administered testing at home for one night.

The HSAT studies were downloaded and manually scored twice. The first scoring was always performed with the frontal EEG signal visible, which we refer to as $HSAT_{EEG}$. In this first scoring, EEG was used to determine sleep versus wake and sleep stages (N1, N2, N3, REM), and the duration of all epochs scored as sleep equaled TST. Next, to simulate standard HSAT without EEG, the EEG tracing was hidden to create a limited sleep montage $HSAT_{Polygraphy}$. The technician then reviewed the visible tracings for events epoch by epoch, adding and/or subtracting based on usual flow and oximetry findings. As the $HSAT_{EEG}$ contained periods that had previously been identified by EEG as wakefulness (respiratory events were therefore not tabulated), the second scoring (without EEG) usually resulted in a higher number of events, as well as a larger denominator of recording time. We did not randomize the order of (with and without EEG) scoring. We deemed this approach preferable to a complete rescore of the respiratory events in order to minimize the variability of respiratory scoring within sleep attributable to inter/intra scorer reliability (see "Discussion" section). Studies were reviewed by one of the study investigators when flagged by the technician for significant periods of invalid airflow or oximetry signal (32/207).

Since no standard exists for determining sleep time in the absence of EEG, valid signal time (VST) was used as a surrogate for TST (denominator of the AHI) and to define the area of valid respiratory event scoring. VST-was defined as the total recording time minus any periods with invalid flow or pulse-oximetry signals. The total recording time is defined as the period between the first appearance of both oximetry and nasal airflow signals, and end of the recording as identified by the disappearance of these two signals.

In both $HSAT_{EEG}$ and $HSAT_{Polygraphy}$, respiratory event scoring was predominantly done using the nasal cannula airflow signal. If the airflow signal quality was poor, chest and abdominal belts were used. AASM rules as revised in 2018 were used with the "ACCEPTABLE" definition for hypopnea and RERA [15]: Apnea was defined as absence of airflow (< 10% baseline) for > 10 s; hypopnea was defined as 30% reduction in airflow for > 10 s resulting in at least 4% desaturation. AHI4 was defined as the number of apneas plus hypopneas divided by the TST for $HSAT_{EEG}$ and apneas plus hypopneas divided by VST for $HSAT_{Polygraphy}$. RDI was calculated using the number of apneas plus hypopneas plus Respiratory Event Related Arousals (RERAs) divided by TST for $HSAT_{EEG}$, and number of apneas, hypopneas, and surrogates for RERAs divided by VST for $HSAT_{Polygraphy}$. "RERA" on $HSAT_{EEG}$ was defined as any visibly reduced airflow associated with inspiratory flow limitation that ended in arousal, but did not meet desaturation criteria for hypopnea. We have previously shown that RERAs defined by the AASM gold standard of esophageal manometry can be detected from the flow signal alone by using the appearance of inspiratory flow limitation and visibly reduced airflow followed by normalization of airflow

after the event [16]. The AASM scoring guidelines for HSAT do not address the rules for calculating the RDI, which requires scoring of respiratory events when there is no associated oxygen desaturation in the absence of EEG. In the present paper, airflow normalization at the end of a 10s event was used was used in $HSAT_{Polygraphy}$ as a surrogate for arousal [16–18]. We have previously demonstrated the validity of this approach by showing good agreement for RDI between in lab PSG and HSAT using the above definitions [17]. A presumed diagnosis of OSA was assigned separately by two criteria: for AHI4, OSA was assigned when AHI4 ≥ 5 events/h; for RDI, OSA was assigned when RDI ≥ 15 events/h [19]. Of patients diagnosed with OSA, severity of disease (also assigned separately for each respiratory event index) was defined as follows: Mild OSA was defined as AHI4 ≥ 5 to 15 events/h or RDI ≥ 15 and < 30 events/h. Severe OSA was defined as AHI4 ≥ 30 or RDI ≥ 45 events/h. Intermediate values of AHI4 and RDI were defined as Moderate OSA [19].

As pulse-oximetry is needed to classify hypopneas, studies with absent or invalid saturation signal for more than 50% of recording time were excluded. For each home study, sleep times (TST and VST), SRBD indices (AHI4 and RDI), and OSA diagnosis/severity were calculated for $HSAT_{EEG}$ and $HSAT_{Polygraphy}$ and compared. The study protocol was approved by the New York University institutional review board.

Impact of pretest patient report of poor sleep suggesting insomnia, and low sleep efficiency

Subjective poor sleep suggesting insomnia was defined as present in patients who answered “yes” to either of the following questionnaire items: “Do you have trouble falling asleep when you first go to bed?” or “ Do you awaken too early in the morning?” and who were without excessive daytime sleepiness (EDS) defined by ESS < 7. Differences in respiratory event indices and OSA diagnosis/severity between HSAT methods were then analyzed for those with and without poor sleep suggesting insomnia as so defined. These analyses were performed to see if poor sleep suggesting insomnia without sleepiness was predictive of a greater benefit from frontal EEG in scoring sleep-disordered breathing, and repeated in a group with subjective poor sleep suggesting insomnia without regard to ESS. A final subgroup was defined as those with poor sleep by EEG as ultimately recorded on the $HSAT_{EEG}$ (TST/VST ≤ 70%).

Statistical analysis

For validation of the limited EEG (Fz) montage for staging sleep, inter-scorer reliability with kappa statistic was performed on the PSG studies. Sleep versus wake agreement between limited EEG montage and full PSG EEG was also analyzed this way. When comparing $HSAT_{EEG}$ to $HSAT_{Polygraphy}$ sleep variables were tested for normality using the Shapiro-Wilk method. TST and valid signal time were compared using paired-sample *t* tests. Differences from $HSAT_{EEG}$ to $HSAT_{Polygraphy}$ in AHI4 and RDI were analyzed with Wilcoxon signed-rank tests. Cross-tabulations with Pearson chi-square tests were performed to analyze agreement for OSA diagnosis and severity, and reported as percent agreement between methods, and sensitivity of testing without EEG ($HSAT_{EEG}$ was used as the reference test). Missing demographic or questionnaire data occurred infrequently and was excluded from analyses for which it was required.

Results

Single-channel frontal versus full EEG during polysomnogram for distinguishing sleep from wake

Data in 21 subjects (ages 32–81 years) from 16 diagnostic PSGs and six CPAP titration studies were analyzed. The average number of 30s epochs/subject was 845 (range 669–990). Using only the limited EEG signal, the two scorers agreed on sleep versus wake in 91% (range 76%–97%) of epochs. When sleep stage as determined by limited EEG was compared to PSG, agreement for epoch by epoch sleep/wake was 92% (scorer 1) and 95% (scorer 2). Epoch by epoch sleep stage agreement between frontal EEG and PSG was 37%/49% for N1, 82%/81% for N2, 97%/99% for N3, and 80%/85% for REM sleep. The associated kappa coefficients were 0.72 (limited montage 1 vs PSG) and 0.73 (limited montage 2 vs PSG). These results are summarized in the supplemental data section (Tables S1, S2).

Comparison of HSAT with and without frontal EEG available to the scorer (*n* = 207)

14/221 HSAT studies were excluded for invalid pulse-oximetry signal during > 50% of recording time, leaving 207 HST studies (94%) for analysis. Baseline patient characteristics are shown in Table 1.

Table 2 shows the data on sleep times, respiratory events, and SDB indices for both HSAT methods. In accord with previously published results, sleep duration estimated by $HSAT_{Polygraphy}$ (VST = 440 ± 76 min) was longer than TST by $HSAT_{EEG}$ (TST = 356 ± 82 min) by about 20%. Thus, if the numerator (i.e., number of events) had remained constant across HSAT methods, this increase of the denominator alone predicts that not having EEG should reduce the AHI4 and RDI a similar amount, approximately 20%. Results show that by dividing actual events scored for each HSAT method by the corresponding defined sleep period, AHI4 by $HSAT_{Polygraphy}$ was 19% lower than $HSAT_{EEG}$. However, the RDI by polygraphy was only 8% lower than the RDI by $HSAT_{EEG}$, i.e., there was less impact on RDI by removal of the EEG. As can be seen in Table 2, the total number of events scored for the AHI4 was similar for $HSAT_{Polygraphy}$ and $HSAT_{EEG}$, whereas there were more RDI events scored by $HSAT_{Polygraphy}$ (183 ± 121 events scored by $HSAT_{Polygraphy}$ vs 158 ± 111 events by $HSAT_{EEG}$). This is because more RERAs were being scored in $HSAT_{Polygraphy}$ (compared with $HSAT_{EEG}$), which mitigated the effect of changing from TST to VST.

Figure 1a and b shows that there is a good correlation between $HSAT_{Polygraphy}$ and $HSAT_{EEG}$ for both SRBD indices (AHI4 and RDI). The line of identity and the associated Bland-Altman plots show that agreement is best at low SRBD indices. Tables 3 and 4 show the agreement statistics for OSA diagnosis, which establishes the clinical significance. For AHI4, there was agreement for OSA diagnosis in 197/207 (95%) subjects and on disease severity in 183/207 (88%). $HSAT_{Polygraphy}$ missed the diagnosis of OSA in 10/207 patients compared to HSATEEG, resulting in a false negative rate of 8.0%. It should be noted that all patients with a false-negative $HSAT_{Polygraphy}$ result had only mild disease (AHI 5–15 events/h). For RDI, there was agreement for OSA diagnosis in 189/207 (91%) of studies and

on disease severity in 164/207 (79%) studies. $HSAT_{Polygraphy}$ missed the diagnosis of OSA in 13/207 patients compared to $HSAT_{EEG}$, a false negative rate of 8.4%.

Subgroup analysis #1: subjects with pretest complaints of poor sleep suggesting insomnia (*n* = 69)

Compared to all patients, pretest poor sleepers with a low Epworth score had a similar Valid Signal Time (440 ± 76 vs 447 ± 62 min), and surprisingly, no significant difference in total sleep time (356 ± 82 vs 349 ± 69 min (Table 2 and supplement Table 3a). Accounting for changes in sleep time alone (holding respiratory events constant), sleep disordered breathing indices for $HSAT_{Polygraphy}$ were predicted to be ~ 22% lower than for $HSAT_{EEG}$. As seen in supplement Table 3a, the observed subgroup AHI4 by $HSAT_{Polygraphy}$ was reduced by 14%. However, RDI was only 10% lower by $HSAT_{Polygraphy}$. Similar to the results seen when all patients were included, in this subgroup of “poor sleep suggesting insomnia” more RDI events were scored by $HSAT_{Polygraphy}$ than $HSAT_{EEG}$ (172 ± 116 events scored by $HSAT_{Polygraphy}$ vs 143 ± 94 events by $HSAT_{EEG}$).

Using AHI4, methods of scoring agreed on diagnosis of OSA in 66/69 (96%) sleep studies, and agreed on OSA severity in 60/69 (87%) studies (Table 4). There were three missed OSA diagnoses in this subset of pretest poor sleepers, with a false negative rate of 7.9%, which was not different than in all patients (Table 4). Using RDI criteria, agreement for OSA diagnosis occurred in 60/69 studies (87%), and agreement for disease severity occurred in 53/69 (77%) studies. There were seven missed OSA diagnoses by RDI with a false negative rate of 14.3%. Sleep efficiency of pretest poor sleepers with low Epworth score was 79 ± 14%, but this was not significantly different from a sleep efficiency of 83 ± 13% observed in those who were not pretest poor sleepers.

Subgroup analysis #2: subjects found to have low sleep efficiency with EE G on the eventual HSAT(*n* = 38)

HSATs of 38 (18%) study patients had sleep efficiency < 70%. In this group, sleep duration by EEG was 266 ± 74 min (TST) versus VST without EEG of 456 ± 95 min (Supplement Table 3b), and sleep efficiency was 58 ± 10% of recording time, compared to 81 ± 14% for all patients. The difference between TST and VST alone predicts this group should have SRBD indices by $HSAT_{Polygraphy}$ ~ 42% lower than with addition of EEG to home testing. For the AHI4, the observed decrease was 39%. Once again the effect on RDI was less (24%). There were 173 ± 119 events scored by $HSAT_{Polygraphy}$ compared to 139 ± 99 events scored by $HSAT_{EEG}$. Using AHI4, methods with and without EEG agreed on diagnosis of OSA in 33/38 (87%) studies, and agreed on OSA severity in 27/38 (71%) studies (Table 4). $HSAT_{Polygraphy}$ missed 5 OSA diagnoses (false negative rate of 20.8%), more than that seen in all patients, and no patients were given a diagnosis based on results of $HSAT_{Polygraphy}$ alone. By RDI, agreement for diagnosis of OSA was 28/38 (74%) and agreement for disease severity was 19/38 (50%). There were ten missed OSA diagnoses by $HSAT_{Polygraphy}$, the false negative rate was 31.3%, and sensitivity to detect OSA by this method compared to $HSAT_{EEG}$ was only 69%.

Discussion

Our in-laboratory data show a high concordance between sleep/wake scoring by a reduced montage of one frontal EEG lead and full PSG (92–95%), suggesting that this may provide a useful way to obtain the actual TST in an ambulatory HSAT. Agreement was similar to a previous study of a limited EEG montage used in a self-applied HSAT that was derived from left and right electrooculogram and compared to full PSG scoring of sleep [20]. In this earlier study, there were substantial differences from our EEG methodology as we used a single frontal EEG lead and did not perform automated scoring [20]. Together these results indicate that a single frontal channel of EEG signal can provide a reasonably accurate determination of the sleep period (total sleep time), the area relevant to scoring of respiratory events during sleep (AHI4 and RDI). Once the scoring validity of a single EEG was confirmed, we then proceeded with a subsequent formal comparison of standard portable monitoring with and without frontal EEG. The agreement for stage-specific scoring (which is not relevant to the present study as all we used was TST) was 80%; this is comparable to what has been published for intra- and inter-scorer agreement [21, 22].

Many limitations of type 3 devices are known, one of which is an accepted degree of inaccuracy introduced by not identifying sleep/wake, or by using only physiologic surrogates for arousal such as movements, heart rate, and changes in sympathetic tone. Studies have shown that abbreviated sleep montages consistently overestimate the sleep period in the laboratory compared to simultaneous PSG [23]. However, this discrepancy has not been previously quantified during home studies. Our data show that portable monitoring with EEG reduces measured sleep duration (TST) compared with standard non-EEG polygraphy-based estimates (VST) by roughly 20% on average. For patients with low sleep efficiency however, even when studied in their usual sleep environment, we found a greater discrepancy between total sleep time and valid signal time, 40–50%. Unfortunately, we were not able to identify subjects with low sleep efficiency before their sleep study by history alone. Our results quantify the upper limits of expected error in AHI4 and RDI by HSAT without EEG, which are relatively high, but may be dependent on the population being tested.

A possible limitation of the analyses in our study is our choice of 4% rather than 3% oxygen desaturation for definition of hypopnea. However, almost all major epidemiologic studies examining sleep apnea and its outcomes have used the hypopnea definition with 4% oxygen desaturation, making this a definition that has relevance to the field [24–26]. While there have been multiple definitions for hypopnea proposed as standards by the AASM—including requiring 4%, 3% desaturation, and EEG arousal in various combinations, the current AASM recommendation is to use 3% oxygen desaturation and/or EEG arousal for hypopnea [15, 27, 28]. The AASM, however, also states it is “ acceptable ” to use 4% desaturation for hypopnea, and arousal for RDI, in part because CMS requires that hypopneas be limited to these events. For simplicity, most sleep laboratories report an AHI4 rather than an AHI3 when disregarding arousal from the EEG. As required by AASM rules, one must then also calculate and report an RDI composed of apneas, hypopneas with 4% oxygen desaturation and other events with arousal, and these are the definitions used in the present analysis. Using 3% rather than 4% oxygen desaturation criteria would raise both the

AHI_{EEG} and $AHI_{Polygraphy}$ (median difference reported in clinical and epidemiologic populations is 2.2–4.7/h [29, 30]), with indices being equally affected as the data are from the same oximeter. Thus, in our study we would expect analysis of the $AHI3_{EEG}$–$AHI3_{polygraphy}$ would differ minimally from $AHI4_{EEG}$–$AHI4_{polygraphy}$.

We also looked for changes to the numerator (number of sleep disordered breathing events) in the AHI4 and RDI across the two testing methods. A major concern was that considerably fewer respiratory events might be scored when EEG signal was part of the scoring montage, because areas of wake can be identified and are not examined for what would otherwise be valid respiratory events; these awake periods may contain irregularities mimicking apnea/hypopnea that are discarded if the EEG is known to show wake. Changes in the respiratory scoring produced by the presence of EEG would thus dilute any effect that differences in sleep duration alone would have on sleep apnea indices. In our study the scoring of respiratory events was performed once with EEG, then again in a second round with EEG hidden from view and editing of previous respiratory scoring. This strategy, we think, allowed us to best compare performance of the two testing modalities, albeit retrospectively, to simulate dual home sleep studies with and without EEG. By not doing two independent, randomly ordered respiratory scoring passes and adding events we may have increased agreement. However, this approach minimizes the variability of respiratory scoring within sleep that is attributable to inter/intra scorer reliability. We have previously shown high scoring agreement both event-by-event and of the overall RDI in full lab PSGs scored with and without visualization of the EEG [17]. In fact, the agreements were similar to inter- and intra-scorer agreement reported in other studies [31]. In the present study, we found that removal of EEG increased the number of respiratory events scored by ~ 13% in a patient population representative of current HSAT prescribing practices. Reasons observed in our data why EEG decreased respiratory event counts were (1) some apneas/hypopneas scored during periods of Valid Signal Time were discarded when EEG showed wake, and, more often (2) periods of inspiratory flow limitation (scored as a surrogate for RERA by $HSAT_{Polygraphy}$) often did not show an associated arousal when EEG was added to the sleep montage and so were discarded in that setting (not scored in $HSAT_{EEG}$) [14].

Thus, the modest effect of adding the frontal EEG on respiratory event indices results from the combined changes in sleep duration and number of events scored, with AHI4 more affected than RDI (see Table 2). This effect results in 5% misdiagnoses by AHI4 and 9% by RDI (predominantly false negatives), which is similar to published rates of misdiagnoses by HSAT compared to PSG [11]. Since our misdiagnoses occurred in subjects with low AHI, these false negative rates may increase in epidemiologic populations.

Our observation that home sleep testing ± EEG performed similarly in the subgroup with pretest complaints of poor sleep (± low Epworth score) as in all patients was not completely unexpected, as subjective insomnia complaints relate poorly to measured sleep efficiency. In contrast, our study showed that subjects with low sleep efficiency experienced the greatest benefit from adding EEG to home sleep testing. In this group the sensitivity to detect OSA by AHI4 using standard HSAT was only 79%, compared to 92% for all patients, and the false negative rate was 20.8% compared to 8.0%. For RDI, the effect of adding frontal EEG in the subjects with low sleep efficiency was further magnified (false negative rate was 31.3

vs 8.4% for the entire group). Since we were not able to identify the subjects most affected from pretest subjective complaints, identifying low efficiency sleepers before formal sleep evaluation remains an area of interest for future study. One might consider use of validated clinical tools like the Insomnia Severity Index or the shorter Brief Insomnia Questionnaire, as these may better predict sleep efficiency than our insomnia screening questions [32–34].

In order to evaluate the utility of a prediction equation on the agreement statistics, we calculated $AHI4_{Predicted}$ and $RDI_{Predicted}$ using the regression of AHI_{EEG} on $AHI4_{Polygraphy}$ ($AHI_{Predicted} = AHI_{Polygraphy}*1.14 + 0.29$; $RDI_{Predicted} = 0.53 + 1.05*RDI_{Polygraphy}$) and using the predicted values for diagnosis of OSA. The statistics using the predicted values for AHI4 showed better agreement, but for the RDI the agreement was the same or lower than found with the raw RDIs (Table 5). This finding is not unexpected as the main difference in AHI4 was in the denominator (TST vs VST) as opposed to in the RDI where the differences between polygraphy and EEG scoring were in both the numerator and denominator. Of note, increasing the sensitivity $HSAT_{Polygraphy}$ by lowering the RDI threshold for OSA to 10 events/h resulted in only two missed diagnoses when comparing to an RDI with EEG cutoff of 15.

There is a known impact on diagnosis of OSA of physiologic night-tonight variability of SDB events, even with full PSG. This has been reported to be in the range of 14–54% [35–37]. It is difficult to compare the impact of this variability to our current data as there are few data in the literature on night-night variability in specific populations such as those with known insomnia. At least one study has suggested the benefit of multiple nights of recording [38]. Presumably using the highest SDB index would minimize any dilution of the denominator due to a single night of poor sleep.

A strength of this study is that we evaluated a clinic population with a somewhat lower prevalence of OSA (53%) than other reports comparing HSAT to PSG [6]. This makes our results particularly relevant to the populations in which screening for OSA is being contemplated, in contrast to the case finding advocated by the AASM.

Conclusions

Total sleep time obtained by recording one channel of frontal EEG is similar to TST obtained from full PSG. Home sleep apnea testing without EEG overestimates the sleep period compared to testing with EEG by approximately 20% overall, and by more in those with low sleep efficiency. In our dataset of 207 patients typical of those seen in a US sleep center, a false negative rate for diagnosis of OSA by AHI4 ≥ 5 occurred in 8.0% of tested subjects, all of whom had mild sleep apnea when EEG was recorded. Severity of OSA was underestimated in 11% of all patients. In patients with low sleep efficiency, the false negative rate for diagnosis of OSA by AHI4 ≥ 5 rose to 20.8%. Whether these rates of misclassification are acceptable will depend on how and where an HSAT is used (i.e., epidemiologic vs clinical, screening vs case finding). While from history alone (prior to home study) we were unable to identify those patients who benefited most from adding an EEG to their home sleep study, it remains to be shown that in some populations this becomes predictable (i.e., CHF, untreated pain).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This research was funded in part by grants from the NIH K24HL109156 (Principal Investigator: Ayappa, Indu), and the Foundation for Research in Sleep Disorders.

Abbreviations

AHI4	Apnea-Hypopnea Index based on hypopneas with 4% oxygen desaturation
AASM	American Academy of Sleep Medicine
EEG	Electroencephalogram
ESS	Epworth Sleepiness Scale
HSAT	Home sleep apnea testing
OSA	Obstructive sleep apnea
PSG	Polysomnography
RDI	Respiratory Disturbance Index based on hypopneas with either 4% desaturation or arousal/arousal surrogate
SDB	Sleep disordered breathing
TST	Total sleep time
VST	Valid signal time

References

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM (2013) Increased prevalence of sleep-disordered breathing in adults. Am J Epidemiol 177(9):1006–1014 [PubMed: 23589584]
2. Flemons WW, Douglas NJ, Kuna ST, Rodenstein DO, Wheatley J (2004) Access to diagnosis and treatment of patients with suspected sleep apnea. Am J Respir Crit Care Med 169(6):668–672 [PubMed: 15003950]
3. Rosen CL, Auckley D, Benca R, FoldvarySchaefer N, Iber C, Kapur V, Rueschman M, Zee P, Redline S (2012) A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the Home PAP study. Sleep 35(6):757–767 [PubMed: 22654195]
4. Collop NA et al. (2011) Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. J Clin Sleep Med 7(5):531–548 [PubMed: 22003351]
5. Bruyneel M, Ninane V (2014) Unattended homebased polysomnography for sleep disordered breathing: current concepts and perspectives. Sleep Med Rev 18(4):341–347 [PubMed: 24388970]
6. Garg N et al. (2014) Homebased diagnosis of obstructive sleep apnea in an urban population. J Clin Sleep Med 10(8):879–885 [PubMed: 25126034]
7. Epstein LJ et al. (2009) Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. J Clin Sleep Med 5(3):263–276 [PubMed: 19960649]

8. Decision memo for continuous positive airway pressure (CPAP) therapy for obstructive sleep apnea (OSA) (2008)
9. Collop NA et al. (2007) Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med 3(7):737–747 [PubMed: 18198809]
10. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, Harrod CG (2017) Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of Sleep Medicine clinical practice guideline. J Clin Sleep Med 13(3):479–504 [PubMed: 28162150]
11. Zeidler MR, Santiago V, Dzierzewski JM, Mitchell MN, Santiago S, Martin JL (2015) Predictors of obstructive sleep apnea on polysomnography after a technically inadequate or normal home sleep test. J Clin Sleep Med 11(11):1313–1318 [PubMed: 26156951]
12. Ayappa I et al. (2009) Irregular respiration as a marker of wakefulness during titration of CPAP. Sleep 32(1):99–104 [PubMed: 19189784]
13. Bresler M et al. (2008) Differentiating between light and deep sleep stages using an ambulatory device based on peripheral arterial tonometry. Physiol Meas 29(5):571–584 [PubMed: 18460762]
14. Berry RB et al. (2012) Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. Deliberations of the sleep apnea definitions Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med 8(5):597–619 [PubMed: 23066376]
15. Malhotra RK, Kirsch DB, Kristo DA, Olson EJ, Aurora RN, Carden KA, Chervin RD, Martin JL, Ramar K, Rosen CL, Rowley JA, Rosen IM (2018) Polysomnography for obstructive sleep apnea should include arousal based scoring: an American Academy of sleep medicine position statement. J Clin Sleep Med 14(7):1245–1247 [PubMed: 29991439]
16. Ayappa I, Norman RG, Krieger AC, Rosen A, O'Malley RL, Rapoport DM (2000) Noninvasive detection of respiratory effort related arousals (REras) by a nasal cannula/pressure transducer system. Sleep 23(6):763–771 [PubMed: 11007443]
17. Ayappa I, Norman RG, Suryadevara M, Rapoport DM (2004) Comparison of limited monitoring using a nasal-cannula flow signal to full polysomnography in sleep disordered breathing. Sleep 27(6):1171–1179 [PubMed: 15532212]
18. Masa JF, Corral J, Gomez de Terreros J, Duran-Cantolla J, Cabello M, Hernández-Blasco L, Monasterio C, Alonso A, Chiner E, Aizpuru F, Zamorano J, Cano R, Montserrat JM, Collaborating group, Garcia-Ledesma E, Pereira R, Cancelo L, Martinez A, Sacristan L, Salord N, Carrera M, Sancho-Chust JN, Embid C (2013) Significance of including a surrogate arousal for sleep apnea-hypopnea syndrome diagnosis by respiratory polygraphy. Sleep 36(2):249–257 [PubMed: 23372273]
19. Ho V, Crainiceanu CM, Punjabi NM, Redline S, Gottlieb DJ (2015) Calibration model for apnea-hypopnea indices: impact of alternative criteria for hypopneas. Sleep 38(12):1887–1892 [PubMed: 26564122]
20. Stepnowsky C, Levendowski D, Popovic D, Ayappa I, Rapoport DM (2013) Scoring accuracy of automated sleep staging from a bipolar electroocular recording compared to manual scoring by multiple raters. Sleep Med 14(11):1199–1207 [PubMed: 24047533]
21. Norman RG et al. (2000) Inter observer agreement among sleep scorers from different centers in a large dataset. Sleep 23(7):901–908 [PubMed: 11083599]
22. Ruehland WR, Churchward TJ, Schachter LM, Lakey T, Tarquinio N, O'Donoghue FJ, Barnes M, Rochford PD (2015) Polysomnography using abbreviated signal montages: impact on sleep and cortical arousal scoring. Sleep Med 16(1):173–180 [PubMed: 25547033]
23. Pang KP, Gourin CG, Terris DJ (2007) A comparison of polysomnography and the Watch PAT in the diagnosis of obstructive sleep apnea. Otolaryngol Head Neck Surg 137(4):665–668 [PubMed: 17903588]
24. Marin JM et al. (2005) Longterm cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. Lancet 365(9464):1046–1053 [PubMed: 15781100]

25. Quan SF et al. (1997) The sleep heart health study: design, rationale, and methods. Sleep 20(12):1077–1085 [PubMed: 9493915]
26. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S (1993) The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 328(17):1230–1235 [PubMed: 8464434]
27. (1999) Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. Sleep 22(5):667–689 [PubMed: 10450601]
28. Iber C, AncoliIsrael S, Chesson AL Jr, Quan SF, for the American Academy of Sleep Medicine (2007) The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. Westchester
29. Redline S et al. (2000) Effects of varying approaches for identifying respiratory disturbances on sleep apnea assessment. Am J Respir Crit Care Med 161(2 Pt 1):369–374 [PubMed: 10673173]
30. Ruehland WR, Rochford PD, O'Donoghue FJ, Pierce RJ, Singh P, Thornton AT (2009) The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index. Sleep 32(2):150–157 [PubMed: 19238801]
31. Heitman SJ, Atkar RS, Hajduk EA, Wanner RA, Flemons WW (2002) Validation of nasal pressure for the identification of apneas/hypopneas during sleep. Am J Respir Crit Care Med 166(3):386–391 [PubMed: 12153976]
32. Kessler RC, Coulouvrat C, Hajak G, Lakoma MD, Roth T, Sampson N, Shahly V, Shillington A, Stephenson JJ, Walsh JK, Zammit GK (2010) Reliability and validity of the brief insomnia questionnaire in the America insomnia survey. Sleep 33(11):1539–1549 [PubMed: 21102996]
33. Bastien CH, Vallieres A, Morin CM (2001) Validation of the insomnia severity index as an outcome measure for insomnia research. Sleep Med 2(4):297–307 [PubMed: 11438246]
34. Gagnon C, Belanger L, Ivers H,Morin CM(2013) Validation of the insomnia severity index in primary care. J Am Board Fam Med 26(6):701–710 [PubMed: 24204066]
35. Dean RJ, Chaudhary BA (1992) Negative polysomnogram in patients with obstructive sleep apnea syndrome. Chest 101(1):105–108 [PubMed: 1729054]
36. Lord S et al. (1991) Night-to-night variability of disturbed breathing during sleep in an elderly community sample. Sleep 14(3):252–258 [PubMed: 1896727]
37. Bittencourt LR et al. (2001) The variability of the apnoea-hypopnoea index. J Sleep Res 10(3):245–251 [PubMed: 11696078]
38. Guerrero A, Embid C, Isetta V, Farre R, Duran-Cantolla J, Parra O, Barbé F, Montserrat JM, Masa JF (2014) Management of sleep apnea without high pretest probability or with comorbidities by three nights of portable sleep monitoring. Sleep 37(8):1363–1373 [PubMed: 25083017]

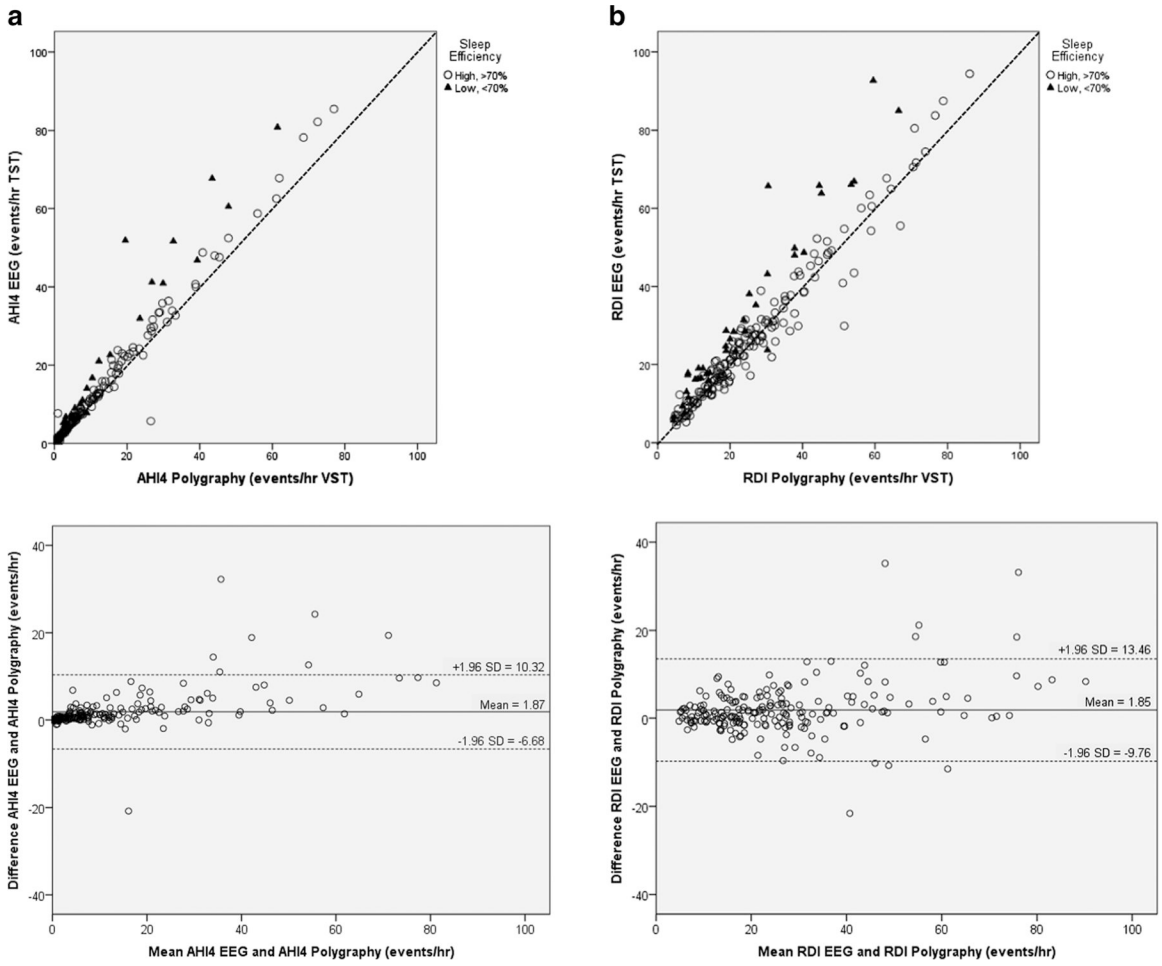

Fig. 1.

a AHI4 and **b** RDI scatter plots and associated Bland-Altman plots for all subjects. Scatter plots are grouped by high vs low sleep efficiency

Table 1

Study population baseline characteristics

Patients (*n* = 207)		
Demographics	**Mean**	**Range**
Gender	147 M/58 F	
Age (years)	47	16–86
BMI (kg/m^2)	28.5	17–51
Epworth score	7	0–22
Sleep history	**Number of patients**	**Percent of total patients**
Prior OSA diagnosis	37	18%
ESS ≥ 7	97	47%
Difficulty initiating sleep[a]	72	35%
Difficulty maintaining sleep[a]	106	51%

[a]Difficulty initiating sleep refers to patients who had a positive response to the question “do you have trouble falling asleep when you first go to bed?” Difficulty maintaining sleep refers to those patients who answered positively to the question “do you awaken too early in the morning?”

Table 2

Summary of events, sleep times, and respiratory event indices with and without EEG available to the scorer for all subjects (n = 207). Delta = $HSAT_{EEG}$ − $HSAT_{Polygraphy}$

Results of scoring HSAT with EEG vs without EEG (polygraphy)			
AHI4	# Events	Sleep time (min)	AHI4 (events/h)
	Mean ± SD	Mean ± SD	Median (IQR)
HSAT-EEG	75± 101	356 ±82	6.4 (1.6–17.8)
HSAT-polygraphy	80 ±106	440 ±76	5.4 (1.3–15.7)
Delta	− 4.7 ±1.0	−83.6±4.5	1.0
Delta (%)	− 6%	− 19%	19%
RDI	# Events	Sleep time (min)	RDI (events/h)
	Mean ± SD	Mean ± SD	Median (IQR)
HSAT-EEG	158±111	356 ±82	22.7 (14.4–33.3)
HSAT-polygraphy	183 ±121	440 ±76	21.0 (13.9–32.4)
Delta	−24.6±1.7	−83.6±4.5	1.7
Delta (%)	−13%	− 19%	8%

Table 3

Cross tabulations of OSA severity defined by AHI4 scoring with and without EEG for all subjects

a) All subjects			OSA severity AHI4 polygraphy			
			No OSA	**Mild**	**Moderate**	**Severe**
OSA severity AHI4 EEG	No OSA	Count	87	0	0	0
		% of row	100.0%	0.0%	0.0%	0.0%
	Mild	Count	10	52	2	0
		% of row	15.6%	81.3%	3.1%	0.0%
	Moderate	Count	0	4	23	0
		% of row	0.0%	14.8%	85.2%	0.0%
	Severe	Count	0	0	8	21
		% of row	0.0%	0.0%	27.6%	72.4%

Table 4

Utility of HSAT-polygraphy compared to HSATEEG by patient subgroup

	All subjects (*n* = 207)	**Poor sleep-low ESS pre-test (*n* = 69)**	**Low sleep efficiency post-test (*n* =38)**
Agreement for OSA diagnosis			
AHI4	95%	96%	87%
RDI	91%	87%	74%
Sensitivity			
AHI4	92%	92%	79%
RDI	92%	86%	69%
Specificity			
AHI4	100%	100%	100%
RDI	91%	90%	100%
False negative rate			
AHI4	8.0%	7.9%	20.8%
RDI	8.4%	14.3%	31.3%
Agreement for OSA severity			
AHI4	88%	87%	71%
RDI	79%	77%	50%

Table 5

Utility of HSAT-predicted from polygraphy compared to HSATEEG by patient subgroup

	All subjects (*n* = 207)	Poor sleep-low ESS pre-test (*n* = 69)	Low sleep efficiency post-test (*n* =38)
Agreement for OSA diagnosis			
AHI4	97%	96%	92%
RDI	90%	87%	79%
Sensitivity			
AHI4	96%	95%	88%
RDI	94%	90%	75%
Specificity			
AHI4	99%	97%	100%
RDI	77%	80%	100%
False negative rate			
AHI4	4.2%	5.3%	12.5%
RDI	5.8%	10.2%	25.0%

Calculations based on $AHI_{predicted} = AHI_{polygraphy}*1.14 + 0.29$; $RDI_{predicted} = 0.53 + 1.05*RDI_{polygraphy}$ using cutoffs of > 5/h for AHI4 and > 15/h for RDI for OSA diagnosis