Sleep Medicine 96 (2022) 20–27

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Polysomnographic validation of an under-mattress monitoring device in estimating sleep architecture and obstructive sleep apnea in adults

Feihong Ding [a, 1], Andrew Cotton-Clay [b, 1], Laura Fava [b], Venkat Easwar [b], Arthur Kinsolving [b], Philippe Kahn [b], Anil Rama [c], Clete Kushida [a, *]

[a] Division of Sleep Medicine, Stanford University, Stanford, CA, USA
[b] Fullpower Technologies Inc, Santa Cruz, CA, USA
[c] The Permanente Medical Group, San Jose, CA, USA

ARTICLE INFO

Article history:
Received 18 October 2021
Received in revised form
2 March 2022
Accepted 18 April 2022
Available online 22 April 2022

Keywords:
Sleeptracker-AI
Polysomnography
Validation
Sleep health
Obstructive sleep apnea

ABSTRACT

Objective: The objective of this study is to evaluate the validity of an under-mattress monitoring device (Fullpower Technologies) in estimating sleep continuity and architecture, as well as estimating obstructive sleep apnea in an adult population.

Methods: Adult volunteers (n=102, 55% male and 45% female, aged 40.6 ± 13.7 years with a mean body mass index of 26.8 ± 5.8 kg/m^2) each participated in a one-night unattended in-lab study conducted by Fullpower Technologies. Each participant slept on a queen-sized bed with Sleeptracker-AI Monitor sensors placed underneath the mattress. Standard polysomnography (PSG) was simultaneously recorded on the same night. Researchers (FD and CK) were provided de-identified sleep studies and datasets by Fullpower Technologies for analysis. Sleep continuity measures, 30-s epoch-by-epoch sleep stages, and apnea and hypopnea events estimated by an automated algorithm from the Sleeptracker-AI Monitor were compared with the PSG recordings, with the PSG recordings serving as the reference.

Results: Overall, the Sleeptracker-AI Monitor estimated similar sleep continuity measures compared with PSG. The Sleeptracker-AI Monitor overestimated total sleep time (TST) by an average of 6.3 min and underestimated wake after sleep onset (WASO) by 10.2 min. Sleep efficiency (SE) was similar between the Sleeptracker-AI Monitor and PSG (87.6% and 86.3%, respectively). The epoch-by-epoch accuracy of Sleeptracker-AI Monitor to distinguish 4-stage sleep (wake, light, deep, and REM sleep) was 79.0% (95% CI: 77.8%, 80.2%) with a Cohen's kappa of 0.676 (95% CI: 0.656, 0.697). Thirty-five participants (34.3%) were diagnosed with obstructive sleep apnea (OSA) with an apnea-hypopnea index (AHI) $\geq$ 5 based on PSG. Accuracy, sensitivity, and specificity for the Sleeptracker-AI Monitor to estimate OSA (an AHI $\geq$5) were 87.3% (95% CI: 80.8%, 93.7%), 85.7% (95% CI: 74.1%, 97.3%), and 88.1% (95% CI: 80.3%, 95.8%) respectively. The positive likelihood ratio (LR+) for AHI $\geq$5 was 7.18 (95% CI: 3.69, 14.0), and the negative likelihood ratio (LR-) for AHI $\geq$5 was 0.16 (95% CI: 0.072, 0.368).

Conclusion: The Sleeptracker-AI Monitor had high accuracy, sensitivity, and specificity in estimating sleep continuity measures and sleep architecture, as well as in estimating apnea and hypopnea events. These findings indicate that Sleeptracker-AI Monitor is a valid device to monitor sleep quantity and quality among adults. Sleeptracker-AI Monitor may also be a reliable complementary tool to PSG for OSA screening in clinical practice.

1. Introduction

Rapid advances in technology have led to unprecedented changes in every aspect of our life. Various user-friendly devices and apps have been developed for self-tracking and assessing an individual's fitness, sleep, and health. Consumer home devices have gained rising popularity among the general population for sleep

* Corresponding author. Division of Sleep Medicine, Stanford University, 450 Broadway St, Redwood City, CA, 94063, USA.
E-mail address: clete@stanford.edu (C. Kushida).
[1] Contributed equally to this work.

https://doi.org/10.1016/j.sleep.2022.04.010

monitoring [1]. These devices are accessible, easy-to-use, non-obtrusive, and available for longitudinal monitoring. Yet most currently available home sleep monitoring devices lack validation. The American Academy of Sleep Medicine (AASM), in their position statement about consumer sleep technology, highlights the need for validation of these devices against PSG, and for FDA approval before they can be used in clinical applications [2].

The Sleeptracker-AI Monitor (Fullpower Technologies, California, USA) is a commercially available, unobtrusive and non-wearable home sleep monitoring device. It is either pre-installed under a mattress or can be purchased separately by consumers. The device uses piezo-electric sensors that register the forces exerted through the mattress by features such as the subject's motion, respiration, heartbeats, and snoring vibrations. The goal of this study is to evaluate and validate the performance of the Sleeptracker-AI Monitor in assessing sleep continuity and architecture, as well as estimating obstructive sleep apnea (OSA).

2. Materials and methods

2.1. Participants

One hundred and two adult volunteers were recruited through online advertisement, followed by a phone call screening to assess their willingness and ability to complete an unattended in-lab study (sole inclusion criterion) with the only exclusion criterion of weight above 400 lbs. All potential participants gave their written informed consent to Fullpower Technologies to participate in the study. The study was conducted in agreement with the Declaration of Helsinki requirements [3]. All participants received a $50 gift card for their participation. Subject recruitment and data collection took place between Jul. 2019 and Mar. 2020. Researchers (FD and CK) participated in the study following data collection, and were provided de-identified sleep studies and datasets by Fullpower Technologies for analysis. The protocol for review and analysis of these de-identified sleep studies and datasets was submitted to the Stanford University Institutional Review Board (IRB) Research Compliance Office; the IRB determined that this protocol did not need IRB approval since it did not involve human subjects as defined in 45 CFR 46.102(f) or 21 CFR 50.3(g).

2.2. Measurements

2.2.1. PSG recordings

Data from in-lab nocturnal PSG (Philips Respironics ALICE 6 LDXs System) and Sleeptracker-AI Monitors were simultaneously collected from each participant in a Fullpower Technologies sleep research laboratory. A standard PSG montage was used, which included a 5-channel electroencephalography (EEG; F3/M2, F4/M1, C3/M2, O1/M2, O2/M1), bilateral electro-oculography (EOG), and chin electromyography (EMG). Respiratory parameters were obtained from an oronasal thermal airflow sensor, dual thoracoabdominal belts, and pulse oximetry. All electrodes were placed by Registered Polysomnographic Technologists (RPSGTs). Video recordings were conducted to assess nocturnal events.

PSG recordings were manually scored in 30-sec epochs according to the AASM scoring manual Version 2.5 [4] by an experienced RPSGT. The RPSGT was blinded to the Sleeptracker-AI Monitor data when scoring PSG data. PSG scoring was further independently verified by one of the study authors (CK), who is a board-certified sleep specialist; the PSGs provided to him for review were deidentified.

2.2.2. Sleeptracker-AI Monitor recordings

One Sleeptracker-AI Monitor setup consists of two sensors, one for each side of the bed, placed between the mattress and foundation and connected to one processing unit. For the study, five Sleeptracker-AI Monitors, with independent predictions, were set up in parallel to evaluate the sensitivity to sensor location, with sensors placed in parallel on both sides of a queen-size bed underneath the mattress (Fig. 1A). The sensor is 5.77 inches long, 3.04 inches wide, and 0.6 inches thick (Fig. 1B). Each sensor captured physical forces exerted through the mattress by the sleeper. These include: (1) body movements; (2) respiratory efforts through the forces of the chest and abdomen on the mattress; (3) heartbeats as a ballistocardiographic signal riding on top of the preceding signals; and (4) snoring by vibration transferred through the mattress. These signals were then processed via an automated algorithm that consists of signal processing, machine and deep learning models, and statistical inference techniques to separate these effects and produce the following estimates: (1) whether the bed is occupied; (2) when the bed is occupied, estimation of sleep vs wake; (3) when sleeping, estimation of light (N1+N2) vs deep (N3) vs REM sleep; and (4) estimation of apnea or hypopnea events. The algorithm has access, through these forces, to time series of motion, breathing, heartbeats, and snoring, and thus to amplitudes and frequencies of and variation in each parameter.

The parallel sensors (one on each side of the bed) were connected to one of the 5 Sleeptracker-AI processors, namely TB-5, PA-1, PA-2, PB-3, and PB-4 (Fig. 1A). Each Sleeptracker-AI processor produced independent sleep recordings with estimations. These sleep recordings were automatically generated and did not depend on lights-out or lights-on indicators.

2.3. Data processing

PSG lights-out and lights-on time were treated as recording start and end time respectively for both the Sleeptracker-AI Monitor and PSG during analysis. Participants turned out their lights at their habitual bedtime. Times when the participant was out of the bed and/or the PSG unit was unplugged were excluded from the analysis for both the Sleeptracker-AI Monitor and PSG (totaling 1.0% of epochs). Both the Sleeptracker-AI Monitor and PSG produced estimates of wakefulness or sleep stage of each 30-s epoch. Time stamps were used to synchronize epoch-to-epoch data between the Sleeptracker-AI Monitor and PSG, as much as possible within 15 s.

2.4. Statistical analyses

Data from each processor, each corresponding to two parallel sensors were independently analyzed. Results from sensors at the PB-3 location were presented for the primary results, as PB-3 is the preferred location for sensor placement per setup instructions. Performance of sensors at other locations were also analyzed, and the results are included in the Appendix.

Means and standard deviations were calculated for all sleep continuity measures, including total sleep time (TST), sleep onset latency (SOL), and wake after sleep onset (WASO) that were provided by the Sleeptracker-AI Monitor and PSG. Pearson correlations were calculated to assess the strength of associations of continuous measures between the Sleeptracker-AI Monitor and PSG. Agreement between the Sleeptracker-AI Monitor and PSG was estimated according to Bland-Altman plots. Similar to others [5–7], we defined as satisfactory an *a priori* difference between the Sleeptracker-AI Monitor and PSG $\leq$30 min for total sleep time (TST) and wake after sleep onset (WASO), and <5% for sleep efficiency (SE).

For sleep stage measures, epoch-by-epoch comparison

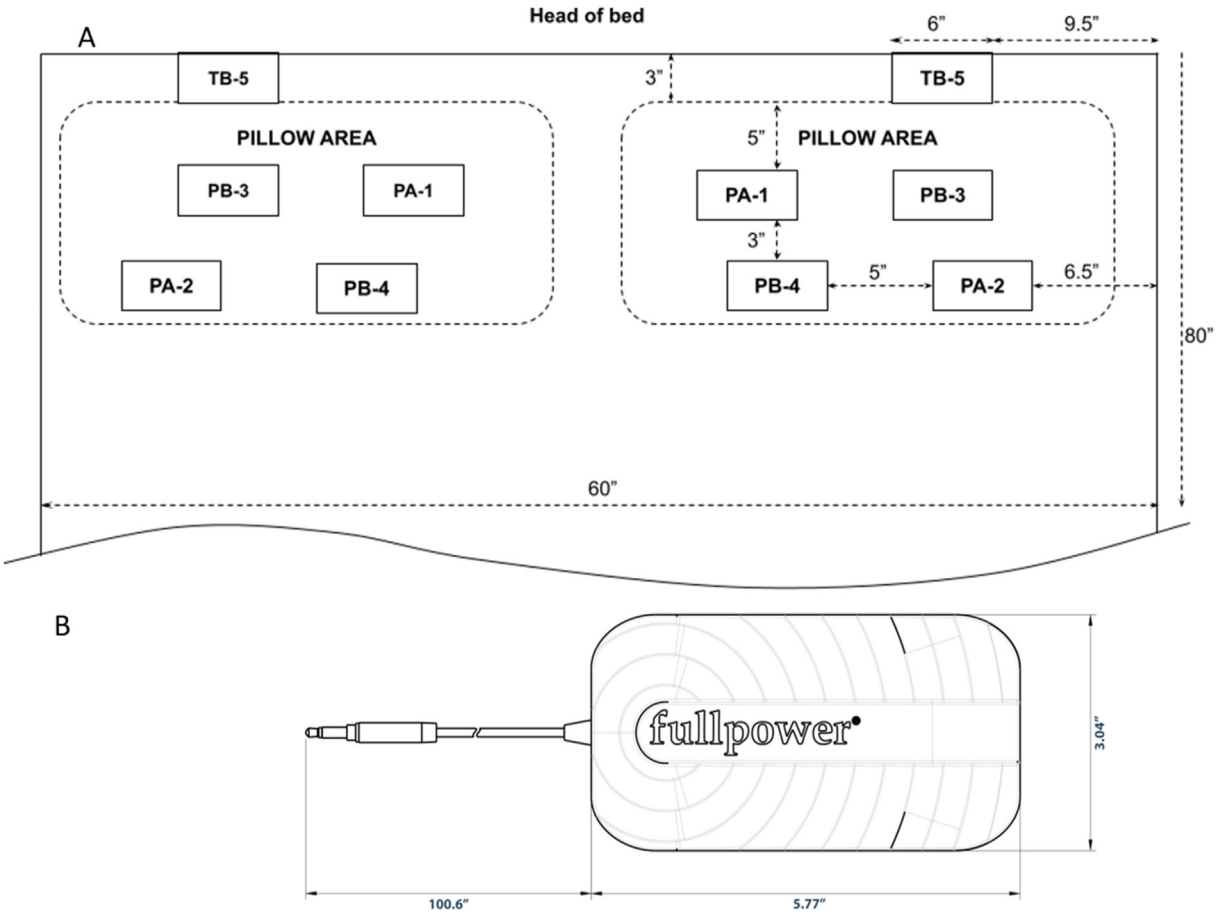

Fig. 1. One Sleeptracker-AI Monitor setup consists of two Sleeptracker-AI Monitor sensors placed between the mattress and foundation, each connected to one Sleeptracker-AI Monitor processing unit.
A) Sleeptracker-AI Monitor sensor positions tested. Each position corresponds to an independent Sleeptracker-AI Monitor, with independent predictions. TB-5: top, sensor B, processor 5; PA1-1: under-pillow, sensor A, processor 1; PA-2: under-pillow, sensor A, processor 2; PB-3: under-pillow sensor B, processor 3 (default); PB-4: under-pillow sensor B, processor 4.
B) Dimensions of Sleeptracker-AI Monitor.

between the Sleeptracker-AI Monitor and PSG was performed to calculate sensitivity, specificity and accuracy. The bootstrap resampling method, treating each subject's data as a single unit for resampling purposes, was used to estimate the 95% confidence interval of sensitivity and specificity [8]. Sensitivity is defined as the probability of the Sleeptracker-AI Monitor to positively identify a sleep stage when it is present on PSG. Specificity is defined as the probability of the Sleeptracker-AI Monitor to negatively identify a sleep stage when a sleep stage is not present on PSG. The accuracy represents the percentage of epochs with the same judgement between the Sleeptracker-AI Monitor and PSG. Cohen's kappa was calculated, which measures the agreement between two methods beyond what would be expected from chance alone. A kappa value of 0–0.2 is considered essentially no agreement, 0.2–0.4 low agreement, 0.4–0.6 moderate agreement, 0.6–0.8 high agreement, and 0.8–1.0 nearly perfect agreement [9].

For the apnea-hypopnea index (AHI), means and standard deviations were calculated. Paired t-tests were used for comparison. Sensitivity, specificity, and accuracy were calculated for the estimation of obstructive sleep apnea, corresponding to an AHI $\geq$5. P values less than 0.05 were considered statistically significant. Statistical analyses were performed using Python (Python Software Foundation, version 3.8.3).

3. Results

Demographic characteristics of all participants are presented in Table 1.

Table 2 summarizes sleep continuity measures and mean time spent in each sleep stage. Overall, the Sleeptracker-AI Monitor estimated similar estimates for sleep continuity compared with PSG. The Sleeptracker-AI Monitor overestimated TST by an average

Table 1
Demographic characteristics of 102 participants.

Characteristic	Mean ± SD (range), or n (%)
Age, years	40.6 ± 13.7 (range 18–72)
Male/Female	55 (53.9%)/47 (46.1%)
BMI, kg/m^2	26.8 ± 5.8 (range 17.7–44.6)
BMI <25	50 (49.0%)
25 $\leq$ BMI <30	28 (27.5%)
BMI $\geq$30	24 (23.5%)
AHI	6.7 ± 12.3 (range 0.0–75.6)
Any OSA (AHI $\geq$5)	35 (34.3%)
Moderate-Severe OSA (AHI $\geq$15)	11 (10.8%)

BMI: Body Mass Index.
AHI: Apnea-Hypopnea Index.
OSA: Obstructive Sleep Apnea.

Table 2
Means and correlation of sleep parameters between Sleeptracker-AI Monitor and polysomnography in 102 participants.

Sleep parameters	Sleeptracker-AI		PSG		Correlation coefficient (Rho)
	Mean (SD)	Range	Mean (SD)	Range	
TST, min	423.9 (83.3)	172.0, 679.0	417.6 (85.2)	150.0, 690.5	0.96**
SOL, min	15.6 (27.0)	0, 205.5	11.6 (11.6)	0, 55.0	0.34*
WASO, min	46.2 (37.0)	5.5, 211.0	56.4 (50.3)	7.0, 294.5	0.88**
SE, %	87.6 (8.7)	55.4, 98.4	86.3 (9.7)	52.1, 97.2	0.87**
Light, min	265.5 (62.2)	124.0, 465.5	241.6 (63.7)	64.5, 445.5	0.84**
Deep, min	62.8 (23.0)	4.5, 119.5	75.2 (32.5)	5.0, 173.0	0.61**
REM, min	96.3 (30.6)	30.0, 170.5	100.8 (39.6)	26.0, 223.0	0.76**

TST: total sleep time; SOL: sleep onset latency; WASO: wake after sleep onset; SE: sleep efficiency; REM sleep: rapid eye movement sleep.
**p value < 0.0001.
*p value < 0.05.

of 6.3 min and underestimated WASO by 10.2 min. The Sleeptracker-AI Monitor overestimated light sleep by 23.9 min and underestimated deep sleep by 12.4 min. However, all the mean differences between the Sleeptracker-AI Monitor and PSG statistics fell within an *a priori* established satisfactory range described above. Agreement for TST, SE, SOL and WASO are shown in Bland-Altman plots (Fig. 2). Correlations are high across all sleep parameters except SOL, which may be due to some large amplitude errors in this parameter which has a smaller scale (lower mean in minutes) than the other parameters considered.

Table 3 provides epoch-by-epoch classification performance of the Sleeptracker-AI Monitor against PSG within each sleep stage. For the 4-class classification to distinguish wake, light, deep and REM sleep, the Sleeptracker-AI Monitor achieved an accuracy of 79.0% with a Cohen's kappa of 0.676. For 3-class classification to distinguish wake, non-REM (NREM), and REM sleep, the accuracy was 86.6% with a Cohen's kappa of 0.733. For 2-class classification, the agreement was high across all sleep stages, with each stage-specific Cohen's kappa ranging between 0.601 and 0.772. Sensitivity and specificity were high across all stages. The sensitivity for detecting deep sleep was slightly lower compared to other sleep stages. See Fig. 3 for example hypnograms comparing the estimated staging from Sleeptracker-AI Monitor to PSG.

Table 4 shows the stage-by-stage cross tabulation of

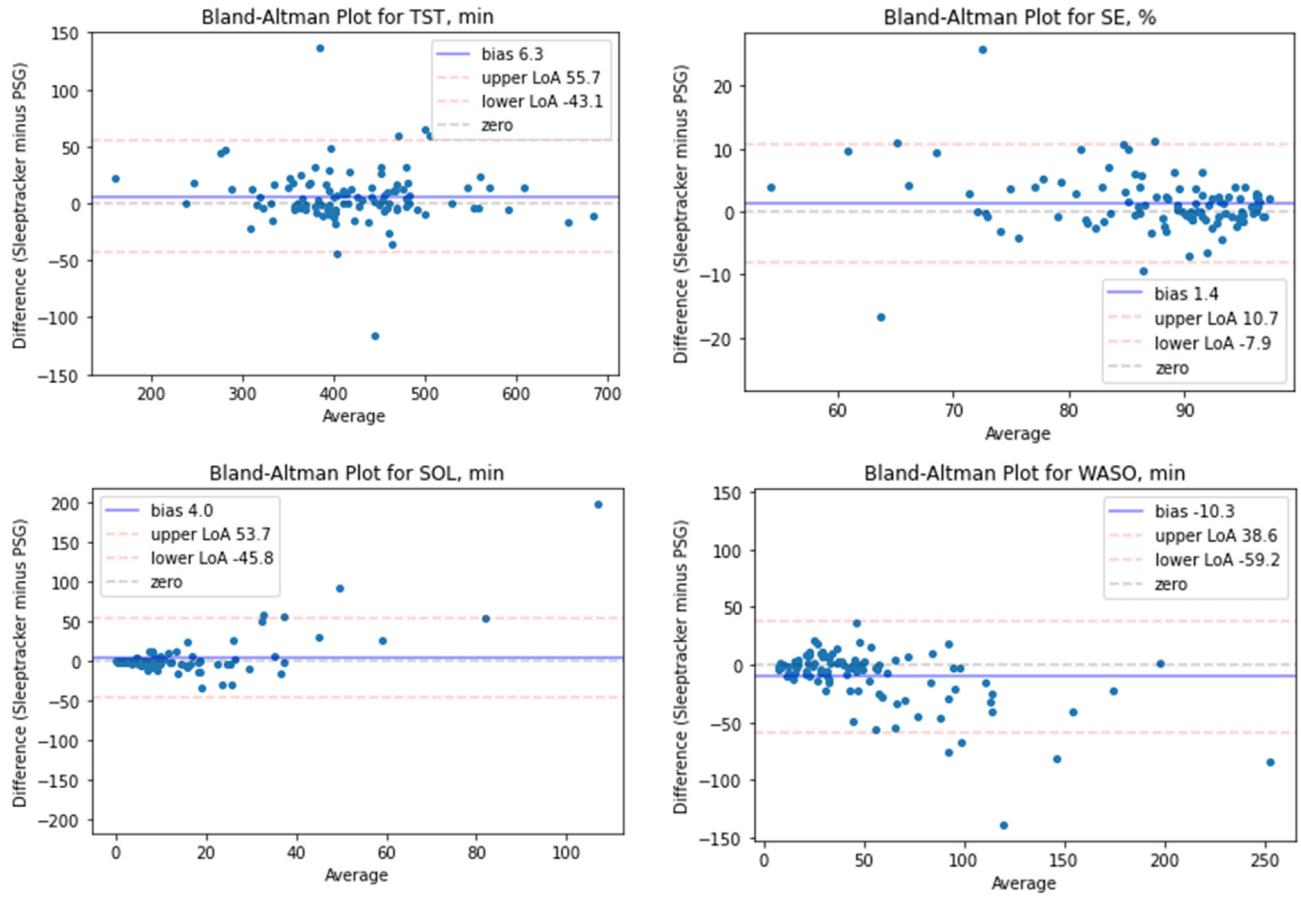

Fig. 2. Bland-Altman plots for total sleep time (TST), sleep efficiency (SE), sleep onset latency (SOL), and wake after sleep onset (WASO) recorded by Sleeptracker-AI Monitor and polysomnography (PSG). Mean (bias) of the differences between the Sleeptracker-AI Monitor and PSG outcome, and lower and upper agreement limits (mean difference ± 1.96 standard deviations) are displayed for each Bland-Altman plot.

Table 3
Epoch-by-epoch agreement, sensitivity and specificity between Sleeptracker-AI Monitor and polysomnography in 102 participants.

Sleep stage	Accuracy (95% CI), %	κ (95% CI)	Sensitivity (95% CI), %	Specificity (95% CI), %
Multiple classes				
Wake/N1+N2/N3/REM	79.0 (77.8, 80.2)	0.676 (0.656, 0.697)	n/a	n/a
Wake/NREM/REM	86.6 (85.4, 87.7)	0.733 (0.706, 0.756)	n/a	n/a
Single class detection				
Wake	93.3 (92.4, 94.1)	0.709 (0.672, 0.743)	71.3 (66.8, 75.7)	96.8 (96.1, 97.4)
Light	80.1(78.8, 81.2)	0.601 (0.576, 0.625)	84.8 (83.6, 86.2)	75.3 (72.9, 77.2)
Deep	92.0 (91.3, 92.7)	0.673 (0.644, 0.702)	65.6 (62.1, 68.9)	96.9 (96.3, 97.4)
REM	92.6 (91.6, 93.5)	0.772 (0.743, 0.800)	80.0 (76.5, 83.3)	95.9 (95.2, 96.5)

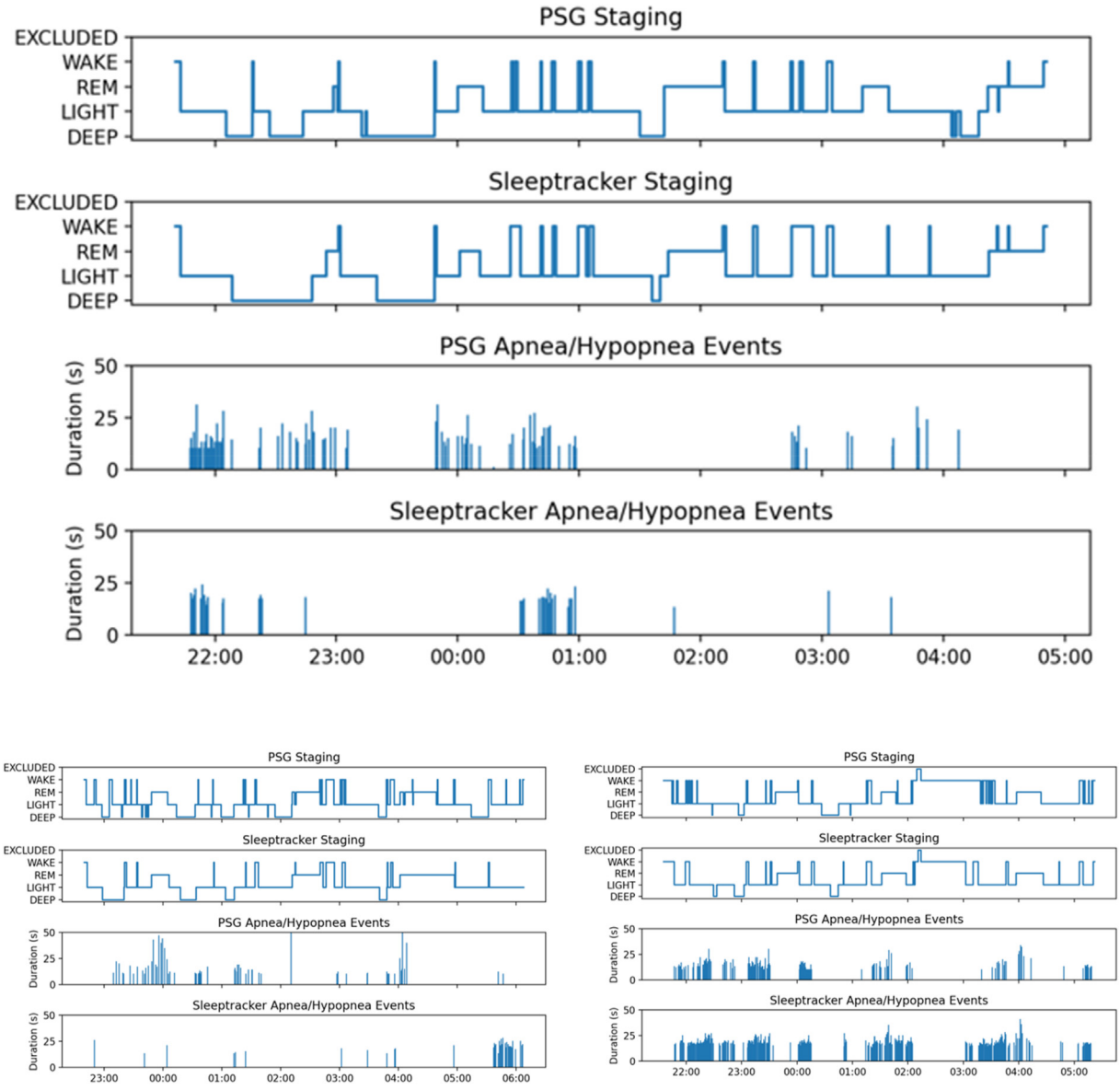

Fig. 3. Hypnograms and apnea/hypopnea events from PSG and Sleeptracker for three recordings, selected based on four-class per-epoch sleep staging accuracy: median (top), lower quartile (bottom left), upper quartile (bottom right). (Note that EXCLUDED includes unplugged status for PSG and off-bed status for Sleeptracker, as discussed.)

Sleeptracker-AI Monitor compared with PSG. The per-stage sensitivities appear along the diagonal. The Sleeptracker-AI Monitor tends to misidentify PSG N3 as light sleep (33.6% of PSG N3 epochs) and also misidentified 23.4% of wake epochs as light sleep. In contrast, there were very few errors misclassifying deep sleep as REM or vice-versa. This tendency for most of the errors to be misclassified between light sleep and another stage may be due to the fact that light sleep is the most prevalent stage and also that most transitions are between light sleep and another sleep stage, so errors in locating transitions cause errors of this type.

There are 35 participants (34.3%) found to have an AHI $\geq$5 based on PSG recordings. Mean AHI based on the Sleeptracker-AI Monitor

Table 4
Stage by stage cross tabulation of Sleeptracker-AI Monitor compared with polysomnography.

Sleeptracker-AI		PSG			
		Wake	Light	Deep	REM
Dichotomized accuracy[a], %		93.3	80.1	92.0	92.6
Cross Tabulation[b], %	Wake	71.32	4.87	0.71	0.92
	Light	23.42	84.92	33.59	19.04
	Deep	1.48	4.84	65.56	0.05
	REM	3.79	5.37	0.14	79.99

[a] Dichotomized % accuracy is the percentage of epochs where both the Sleeptracker-AI Monitor and PSG rate the stage as present or where both the Sleeptracker-AI Monitor and PSG rate the stage as not present.
[b] Cross Tabulation provides the percentage of epochs of each stage rated by the Sleeptracker-AI Monitor when PSG rated a given stage.

and PSG were 7.9 ± 14.1 and 6.7 ± 12.3 respectively, with a paired *t*-test p-value of 0.057. Central apnea events were minimal based on PSG. Accuracy, sensitivity, and specificity for the Sleeptracker-AI Monitor to estimate OSA (an AHI ≥5) were 87.3% (95% CI: 80.8%, 93.7%), 85.7% (95% CI: 74.1%, 97.3%), and 88.1% (95% CI: 80.3%, 95.8%) respectively (Table 5). The positive likelihood ratio (LR+) for AHI ≥5 was 7.18 (95% CI: 3.69, 14.0). There were 11 participants (10.8%) with AHI ≥15 based on PSG recordings. Statistics on Sleeptracker-AI Monitor's estimation of moderate-to-severe OSA (AHI ≥15) are shown in Table 5. See also Fig. 4 for a comparison of Sleeptracker-AI Monitor estimated AHI values to AHI values from PSG, and Fig. 3 for examples of estimated apnea-hypopnea events compared to PSG.

Subgroup analysis demonstrated excellent performance of the Sleeptracker-AI Monitor in estimating sleep continuity measures and classifying sleep stages among participants with OSA as well as those without OSA, and participants with normal BMI as well as those who are overweight or obese (Appendix Supplementary Table A1-8).

The performance of the Sleeptracker-AI Monitor sensors at other locations was similar to that of the PB-3 location (Appendix Supplementary Table B1-8). For two of the locations (PB-3 and PB-4), all N = 102 Sleeptracker-AI Monitor recordings were available. For three of the locations (TB-5, PA-1, and PA-2), only N = 100 Sleeptracker-AI Monitor recordings were available. This was due to missing sensor data for the remaining 2 study nights, during which the corresponding Sleeptracker-AI Monitor device either lost power or connectivity or otherwise failed to collect and upload sensor data for processing. Therefore, only N = 100 study nights were used for the analysis for those locations.

4. Discussion

This study directly compared the Sleeptracker-AI Monitor against PSG in estimating sleep continuity measures, sleep stages, and apnea and hypopnea events among adult volunteers.

The Sleeptracker-AI Monitor provided similar estimates of sleep continuity measures. The epoch-by-epoch agreement in estimating sleep stages was high, with a 4-class accuracy of 0.790 and Cohen's kappa of 0.676. The performance of the Sleeptracker-AI Monitor in estimating OSA (AHI ≥5) was excellent, with a sensitivity of 85.7%, and a positive likelihood ratio (LR+) of 7.18 when compared to PSG. It meets both the sensitivity and LR+ criteria (a sensitivity ≥82.5% and a LR+ ≥ 5) set for out-of-center testing devices for use in confirming OSA (AHI ≥5) among patients with a high pretest probability [10]. We conclude that the Sleeptracker-AI Monitor is a valid device in assessing sleep continuity measures and sleep architecture. It may also serve as a reliable tool for OSA screening.

Table 5
Accuracy, sensitivity, and specificity of Sleeptracker-AI Monitor in estimating obstructive sleep apnea among 102 participants compared with polysomnography.

	AHI ≥5	AHI ≥15
Accuracy (95% CI), %	87.3 (80.8, 93.7)	92.2 (86.9, 97.4)
Cohen's Kappa (95% CI)	0.723 (0.582, 0.864)	0.649 (0.414, 0.883)
Sensitivity (95% CI), %	85.7 (74.1, 97.3)	81.8 (59.0, 100.0)
Specificity (95% CI), %	88.1 (80.3, 95.8)	93.4 (88.3, 98.5)
Positive likelihood ratio (LR+) (95% CI)	7.18 (3.69, 14.0)	12.4 (5.44, 28.3)
Negative likelihood ratio (LR-) (95% CI)	0.162 (0.072, 0.368)	0.195 (0.055, 0.685)

Sleep is essential to health [11]. Buysse defines sleep health as a multidimensional pattern of sleep-wakefulness, adapted to individual, social, and environmental demands, that promotes physical and mental well-being [12]. To date, conventional in-lab PSG study remains the gold standard for evaluating sleep quality and diagnosing sleep disordered breathing. However, the high cost and level of expertise needed in performing and interpreting PSG prevent it from routine implementation in the general population for sleep health evaluation. In addition, good sleep health goes beyond the absence of sleep disorders, or a single sleep measure such as sleep duration. Buysse's perspective on good sleep health highlights the following dimensions of sleep: subjective satisfaction, appropriate timing, adequate duration, high efficiency, and sustained alertness during waking hours [12]. Thanks to rapid advances in artificial intelligence (AI), particularly machine learning algorithms, AI-based consumer home sleep monitoring devices have become easily accessible to the general population, allowing comprehensive and longitudinal sleep monitoring with automated output of sleep continuity and architecture measures [1,2]. Validation studies assessing the performance of such consumer home sleep monitoring devices against PSG are critical for interpreting automated output of sleep measures.

The Sleeptracker-AI Monitor is one of the few under-mattress sleep monitoring devices. Previously, limited validation studies have been conducted for under-mattress sleep devices among a small number of participants [13,14]. The conclusions from those studies were inconsistent. Tuominen et al. [13] compared an under-mattress Ballistocardiograph Beddit Sleep Tracker with PSG among 10 adults between age 18 to 30 with a BMI less than 30. The study showed poor correlation between the two devices in estimating sleep continuity measures and extremely poor agreement in sleep stage classification with a Cohen's kappa less than 0.1. Thus, the authors concluded that the Beddit Sleep Tracker is not a valid device to monitor sleep. Nagatomo et al. [14] compared another under-mattress sensor, the Nemuri SCAN against PSG among 11 critically ill patients in the intensive care unit. The accuracy, sensitivity, and specificity of patients' sleep between the Nemuri SCAN and PSG were 68%, 90%, and 39%. The poor specificity of the Nemuri SCAN was attributed by the authors to its inability to identify immobile wakefulness, which is often observed in critically ill patients.

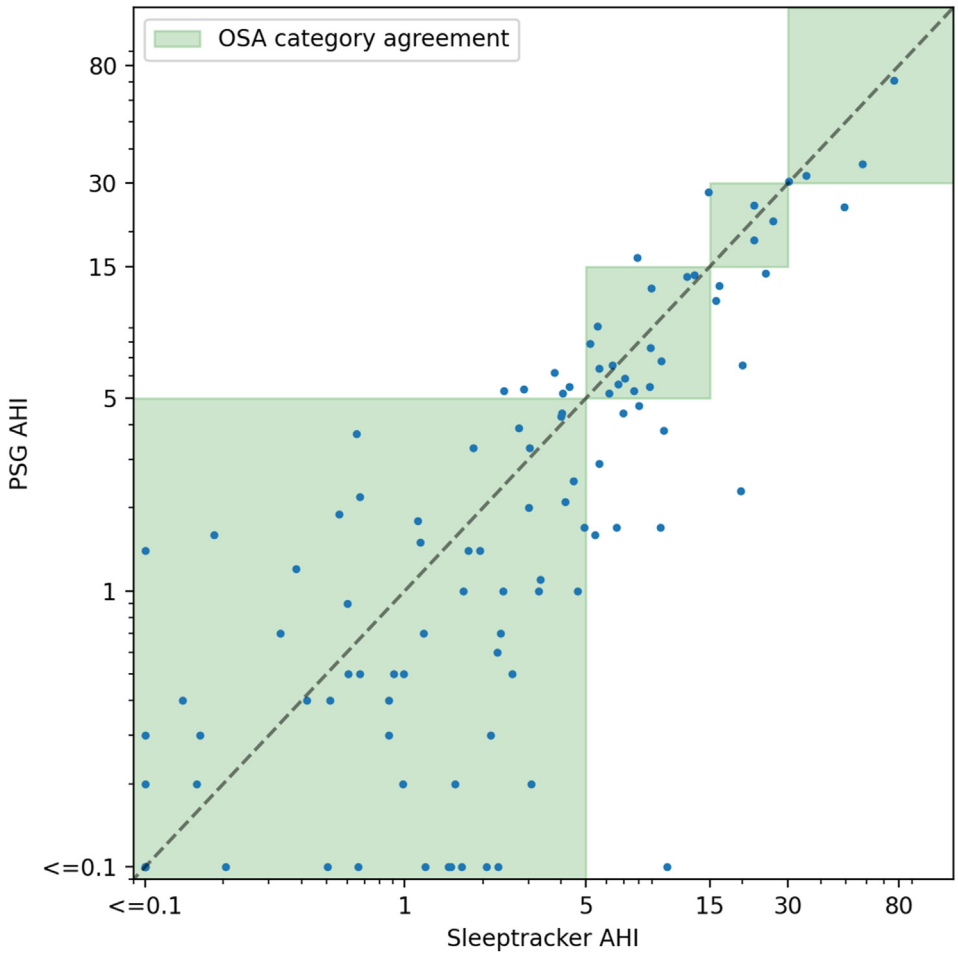

Fig. 4. Apnea-Hypopnea Index values (log scale) estimated by the Sleeptracker-AI Monitor compared to those from PSG, with regions where the resulting OSA categories (none, mild, moderate, severe) agree shaded.

Our study showed excellent performance of the Sleeptracker-AI Monitor in estimating sleep continuity measures and demonstrated sleep staging accuracy comparable to interrater reliability among PSG scorers. In 2013, the AASM published a study reporting that interscorer agreement for sleep stages averaged 82.6% [15]. The excellent performance of the Sleeptracker-AI Monitor, with accuracy (percent agreement with PSG) of 79.0% could be due to advances in technology. The Sleeptracker-AI Monitor uses highly sensitive piezo-electric sensors that register forces through the mattress, including those from body movement, respiratory effort, heartbeats, and snoring vibration. The piezo-electric signals are processed using digital signal processing and machine learning algorithms to automatically produce sleep estimates.

In patients with a high pretest probability of having OSA, the criteria for out-of-center testing devices to be used in diagnosing OSA are a LR+ $\geq$ 5 and a minimum sensitivity of 82.5% coinciding with a PSG-generated AHI $\geq$5 [10,16]. The performance of Sleeptracker-AI Monitor in estimating OSA meets the above criteria, although participants in this study were volunteers from the general population of adults. The performance of the Sleeptracker-AI Monitor in estimating OSA (an AHI $\geq$5) was also comparable to the Withings Sleep Analyzer [17]. In a recent validation study assessing the performance of the Withings Sleep Analyzer against PSG in detecting OSA among 118 patients with suspected OSA, the accuracy, sensitivity, and specificity for moderate-to-severe OSA (an AHI $\geq$15) were 92.6%, 88.0%, 88.6% respectively compared to PSG [17]. Note that in the Sleeptracker-AI study only 11 of the 102 subjects had AHI $\geq$15 and in the Withings study only 12 of the 118 subjects had AHI <5. Another mattress-based solution used capacity-coupled ECG and bioimpedance sensors [18] to predict OSA, though in this case on top of the mattress. In that study, the accuracy, sensitivity, and specificity for moderate-to-severe OSA (AHI $\geq$15) on a dataset of 36 test recordings of suspected OSA patients were 61.1%, 53.9%, and 80.0%, respectively compared to PSG [18].

There are several strengths to this study. First, this is one of the first studies investigating under-mattress home monitoring devices in estimating OSA, and it demonstrates excellent performance. Secondly, adults from the general public (weight <400 lbs and ability to undergo an unattended sleep study) were included, which improves the generalizability of the study findings. Thirdly, as an advantage of an under-mattress sleep monitoring device, data were collected passively with no subject interaction needed and objective sleep estimates were produced automatically.

Limitations of this study should be acknowledged. First, the current study was conducted in a sleep laboratory, which preplaced the Sleeptracker-AI sensor under the mattress. This may not reflect a real-world setting with possible misplacement of the sensor. But note that several sensor placements were tested in the general areas under the pillow region with consistent performance, which should mitigate this situation. Second, current validity results may not hold if the algorithm for sleep parameter and

architecture estimation is updated. However, upgrades of algorithms typically improve the performance of devices. Thirdly, this study is based on a single overnight recording, which does not allow the examination of the within-subject reliability. Collection of longitudinal data of subjects may allow investigation of the association of sleep and individuals' long-term health. Lastly, each participant slept on a queen-sized bed for this study. The influence of having a bed partner and the performance of simultaneously tracking two users sleeping on the same bed are currently being studied in the home; however, sleep laboratories do not typically study two patients at the same time in a single bed.

5. Conclusion

To conclude, the Sleeptracker-AI Monitor is a valid consumer home sleep monitoring device for assessing sleep continuity measures and sleep architecture. It may also serve as a reliable tool for sleep health evaluation and OSA screening in clinical practice.

Funding

None.

CRediT authorship contribution statement

Feihong Ding: Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Andrew Cotton-Clay:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Laura Fava:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – review & editing. **Venkat Easwar:** Project administration, Conceptualization, Methodology, Software, Investigation, Resources, Data curation, Writing – review & editing, Supervision. **Arthur Kinsolving:** Project administration, Conceptualization, Investigation, Resources, Supervision, Writing – review & editing. **Philippe Kahn:** Project administration, Conceptualization, Investigation, Resources, Supervision, Funding acquisition. **Anil Rama:** Writing – review & editing, Supervision. **Clete Kushida:** Project administration, Validation, Writing – original draft, Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

Andrew Cotton-Clay, Laura Fava, Venkat Easwar, Arthur Kinsolving, and Philippe Kahn declare that they have conflicts of interests. They are employees of Fullpower Technologies, Inc.

Acknowledgements

The authors appreciate Dr. Jennifer Zitser's constructive comments on the manuscript.

Abbreviations

AASM	American Academy of Sleep Medicine
AHI	Apnea-hypopnea index
CI	Confidence intervals
OSA	Obstructive sleep apnea
PSG	Polysomnography
SE	Sleep efficiency
SOL	Sleep onset latency
TST	Total sleep time
WASO	Wake after sleep onset

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sleep.2022.04.010.

References

[1] Kelly JM, Strecker RE, Bianchi MT. Recent developments in home sleep-monitoring devices. ISRN Neurol. 2012:768794. https://doi.org/10.5402/2012/768794.
[2] Khosla S, Deak MC, Gault D, Goldstein CA, Hwang D, Kwon Y, et al. Consumer sleep technology: an American Academy of sleep medicine position statement. J Clin Sleep Med 2018;14:877–80. https://doi.org/10.5664/jcsm.7128.
[3] World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310:2191–4. https://doi.org/10.1001/jama.2013.281053.
[4] Berry RB, Albertario CL, Harding S. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. American Academy of Sleep Medicine. Version 2.5. IL: Darien; 2018.
[5] Werner H, Molinari L, Guyer C, Jenni OG. Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. Arch Pediatr Adolesc Med 2008;162:350–8. https://doi.org/10.1001/archpedi.162.4.350.
[6] Meltzer LJ, Walsh CM, Traylor J, Westin AML. Direct comparison of two new actigraphs and polysomnography in children and adolescents. Sleep 2012;35:159–66. https://doi.org/10.5665/sleep.1608.
[7] de Zambotti M, Baker FC, Colrain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015;38:1461–8. https://doi.org/10.5665/sleep.4990.
[8] Efron B. Bootstrap methods: another look at the jackknife. Ann Stat 1979;7:1–26. https://doi.org/10.1214/aos/1176344552.
[9] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
[10] Collop NA, Tracy SL, Kapur V, Mehra R, Kuhlmann D, Fleishman SA, et al. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. J Clin Sleep Med 2011;7:531–48. https://doi.org/10.5664/JCSM.1328.
[11] Ramar K, Malhotra RK, Carden KA, Martin JL, Abbasi-Feinberg F, Aurora RN, et al. Sleep is essential to health: an American Academy of Sleep Medicine position statement. J Clin Sleep Med 2021. https://doi.org/10.5664/jcsm.9476.
[12] Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014;37:9–17. https://doi.org/10.5665/sleep.3298.
[13] Tuominen J, Peltola K, Saaresranta T, Valli K. Sleep parameter assessment accuracy of a consumer home sleep monitoring Ballistocardiograph Beddit sleep tracker: a validation study. J Clin Sleep Med 2019;15:483–7. https://doi.org/10.5664/jcsm.7682.
[14] Nagatomo K, Masuyama T, Iizuka Y, Makino J, Shiotsuka J, Sanui M. Validity of an under-mattress sensor for objective sleep measurement in critically ill patients: a prospective observational study. J Intensive Care 2020;8:16. https://doi.org/10.1186/s40560-020-0433-x.
[15] Rosenberg RS, Van Hout S. The American Academy of Sleep Medicine inter-scorer reliability program: sleep stage scoring. J Clin Sleep Med 2013;9:81–7. https://doi.org/10.5664/jcsm.2350.
[16] Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, et al. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of sleep medicine clinical practice guideline. J Clin Sleep Med 2017;13:479–504. https://doi.org/10.5664/jcsm.6506.
[17] Edouard P, Campo D, Bartet P, Yang R-Y, Bruyneel M, Roisman G, et al. Validation of the Withings Sleep Analyzer, an under-the-mattress device for the detection of moderate-severe sleep apnea syndrome. J Clin Sleep Med 2021;17:1217–27. https://doi.org/10.5664/jcsm.9168.
[18] Huysmans D, Castro I, Borzée P, Patel A, Torfs T, Buyse B, Testelmans D, Van Huffel S, Varon C. Capacitively-coupled ECG and respiration for sleep-wake prediction and risk detection in sleep apnea patients. Sensors 2021 Sep 25;21(19):6409. https://doi.org/10.3390/s21196409. PMID: 34640728; PMCID: PMC8512805.