scientific reports

OPEN

Validation of sleep-staging accuracy for an in-home sleep electroencephalography device compared with simultaneous polysomnography in patients with obstructive sleep apnea

Jaehoon Seol[1,2,3✉], Shigeru Chiba[2], Fusae Kawana[4], Saki Tsumoto[2,5], Minori Masaki[2,5], Morie Tominaga[6], Takashi Amemiya[6], Akihiro Tani[6], Tetsuro Hiei[6], Hiroyuki Yoshimine[7], Hideaki Kondo[8] & Masashi Yanagisawa[2,6,9,10,11✉]

Efforts to simplify standard polysomnography (PSG) in laboratories, especially for obstructive sleep apnea (OSA), and assess its agreement with portable electroencephalogram (EEG) devices are limited. We aimed to evaluate the agreement between a portable EEG device and type I PSG in patients with OSA and examine the EEG-based arousal index's ability to estimate apnea severity. We enrolled 77 Japanese patients with OSA who underwent simultaneous type I PSG and portable EEG monitoring. Combining pulse rate, oxygen saturation (SpO$_2$), and EEG improved sleep staging accuracy. Bland–Altman plots, paired t-tests, and receiver operating characteristics curves were used to assess agreement and screening accuracy. Significant small biases were observed for total sleep time, sleep latency, awakening after falling asleep, sleep efficiency, N1, N2, and N3 rates, arousal index, and apnea indexes. All variables showed > 95% agreement in the Bland–Altman analysis, with interclass correlation coefficients of 0.761–0.982, indicating high inter-instrument validity. The EEG-based arousal index demonstrated sufficient power for screening AHI ≥ 15 and ≥ 30 and yielded promising results in predicting apnea severity. Portable EEG device showed strong agreement with type I PSG in patients with OSA. These suggest that patients with OSA may assess their condition at home.

Unoptimized sleep, whether qualitative or quantitative, leads to dementia, depression, cardiometabolic health issues, and increased mortality[1,2]. Approximately one billion middle-aged adults worldwide are estimated to be affected by obstructive sleep apnea (OSA) with or without symptoms[3]. An epidemiological study revealed that OSA is the most rapidly increasing form of sleep disturbance[4].

OSA can be screened for in the laboratory and/or medical facilities using standard polysomnography (PSG) (type I PSG). This disease can also be diagnosed by monitoring oxygen saturation during sleep (e.g., 3% oxygen

[1]Faculty of Health and Sports Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8574, Japan. [2]International Institute for Integrative Sleep Medicine (WPI-IIIS), University of Tsukuba, 1-2 Kasuga, Tsukuba, Ibaraki 305-8550, Japan. [3]Department of Frailty Research, Center for Gerontology and Social Science, National Center for Geriatrics and Gerontology, Obu, Aichi 474-8511, Japan. [4]Cardiovascular Respiratory Sleep Medicine, Juntendo University Graduate School of Medicine, Tokyo 113-8421, Japan. [5]Ph.D. Program in Humanics, University of Tsukuba, Tsukuba, Ibaraki 305-8575, Japan. [6]S'UIMIN, Inc., Tokyo 151-0061, Japan. [7]Department of Respiratory Medicine, Inoue Hospital, Nagasaki, Nagasaki 850-0045, Japan. [8]Department of General Medicine, Institute of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8102, Japan. [9]Life Science Center for Survival Dynamics (TARA), University of Tsukuba, Ibaraki 305-8577, Japan. [10]R&D Center for Frontiers of Mirai in Policy and Technology (F-MIRAI), University of Tsukuba, Ibaraki 305-8575, Japan. [11]Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. ✉email: seol.jaehoon.ge@u.tsukuba.ac.jp; yanagisawa.masa.fu@u.tsukuba.ac.jp

1

desaturation index [ODI], apnea–hypopnea index [AHI], and respiratory events index [REI]). However, owing to the emotional burden of an unfamiliar sleep environment, accurate sleep measurement methods in home environments to ensure precise sleep assessments need to be established[5,6]. Furthermore, while the number of patients with OSA in Japan is estimated to be over 9 million[3], the number of PSG tests carried out annually is approximately 80,000[7], indicating a low test-to-number of patients.

To address these limitations, several comparative validations were conducted in both laboratory and home environments using equipment such as unattended PSG (e.g., based on the American Academy of Sleep Medicine [AASM] criteria, including type II, III, and IV PSG) and wearable devices such as accelerometers and smartphone applications[5,8–12]. Type II, III, and IV involve unattended full PSG with ≥ 7, 4–7, and 1–2 channels, respectively, measuring oxygen saturation, heart rate, respiratory bands, and airflow. These measurements conducted outside the sleep laboratory using portable monitoring devices have shown substantial agreement with type I PSG[5,12,13]. However, the use of electrodes and other components in existing type II-IV PSGs remains complex, prompting the need for the development of portable electroencephalography (EEG) devices that patients can easily wear to screen for conditions such as OSA[5]. A portable EEG device that integrates electrodes, which is simpler than PSG, has been used in studies involving healthy middle-aged and/or older adults[14,15]. The advantage of this device is that although it has a smaller number of electrodes than the existing PSGs, a variety of channels can be generated by combining montages, and its accuracy improves with a smaller number of electrodes[14] (Supplementary Fig. S1).

Accelerometers, such as actigraphs, which are easier to measure than type II–IV PSGs, exhibit high agreement with type I PSG in healthy participants for determining sleep/wake and sleep latency[8,11,16]. However, this remains an area for further investigation in patients with sleep disturbances[8,9,11]. Furthermore, the actigraph does not determine the sleep stage, and it consistently underestimates sleep parameters[5,16,17]. This inconsistency may be attributed to the high arousal response noise due to dyspnea, which leads to low accuracy[5].

As mentioned earlier, despite the high concordance between type II-IV PSG measurements and type I PSG measurements using various instruments[12], there are still barriers preventing patients with OSA from undergoing measurements on their own at home[12]. Therefore, we aimed to assess the level of agreement between our portable EEG device and type I PSG in patients with OSA. Additionally, it has been reported that the arousal index measured by PSG (i.e., EEG-based) can serve as a screening tool for identifying OSA[18,19]. Consequently, the process of OSA screening could be more accessible to a larger number of patients with OSA if a portable EEG device alone could effectively serve as a screening method without the need for supplementary devices. In this study, we aimed to investigate the screening capability of a portable EEG device in patients with OSA, determine the percentage of agreement with type I PSG, and evaluate the screening effectiveness using AHI thresholds of 15 and 30.

Results

The participant characteristics are presented in Supplementary Table S1, and the cumulative display of sleep architecture for all 77 participants is shown in Fig. 1. The average abdominal and neck circumferences were 90.4 ± 13.7 cm and 37.0 ± 4.1 cm, respectively. The total sleep time, sleep efficiency, and arousal index averaged 425.8 ± 69.5 min, 85.4 ± 9.4%, and 37.0 ± 15.1, respectively (Fig. 1 and Table 1). In this study, apnea was categorized using the AHI as mild (5–15) in 20 patients (25.9%), moderate (15–30) in 17 patients (22.1%), and severe (> 30) in 34 patients (44.1%). Six patients (7.9%) were suspected of having apnea during screening; however, they had an AHI of < 5. Table 1 presents the sleep variables measured using both the type I PSG and portable EEG devices. The results of the Bland–Altman analysis, paired t-test, and intraclass correlation coefficient (ICC) are presented in Figs. 2, 3 and 4. Significant small biases were observed in total sleep time, sleep latency, wake after sleep onset, sleep efficiency, N1, N2, and N3 rates, arousal index, and between AHI and REI ($P < 0.05$; Figs. 2, 4a). Notably, 95% of the values fell within the limits of agreement (LOA) for all variables, indicating a reasonable bias range of − 5.5–10.2 min and − 8.3–9.8% (Figs. 2, 4a). Furthermore, the ICC exceeded 0.75, indicating strong validity between the devices (Figs. 3, 4b).

The confusion matrix for the five sleep stages (Wake, N1, N2, N3, REM) out of a total of 75,677 epochs (Table 2). The agreement between the two devices was 10,032 (91.7%), 11,380 (55.8%), 24,921 (90.0%), 4,230 (66.6%), and 8,772 (85.2%) epochs for Wake, N1, N2, N3, and REM, respectively (Table 2). The overall kappa coefficient was 0.708.

The epoch-by-epoch analysis showed good accuracy in identifying sleep stages between the two devices: Wake (96.4), N1 (83.9), N2 (85.0), N3 (95.8), and REM (95.6) (Table 3). Furthermore, the specificity of the sleep stages was significant: Wake (97.1), N1 (93.8), N2 (82.2), N3 (98.6), and REM (97.3). The sensitivities for Wake (88.2), N2 (88.6), and REM (84.1) were higher than those of N1 (54.3) and N3 (53.9) (Table 3). The trend of high concordance for arousal, N2, and REM, while relatively low for N1 and N3, remained consistent even when the AHI levels were different. (Supplementary Table S2, Supplementary Figs. S2 and S3). Notably, N1 and N2 PABAKs showed reductions of 18.4% and 8.2%, respectively, in cases where AHI was ≥ 30 when compared with AHI < 15 (Supplementary Fig. S3). The arousal index obtained from the portable EEG device had an area under the curve (AUC) of 0.897, with sensitivity, specificity, and accuracy for AHI ≥ 15 being 0.784, 0.923, and 0.835, respectively. Additionally, the AUC, sensitivity, specificity, and accuracy for AHI ≥ 30 were 0.968, 0.971, 0.971, and 0.924, respectively (Fig. 5).

Discussion

This study is one of the few to confirm the concordance of portable EEG devices with type I PSG in patients with OSA[20]. N1 and N3 exhibited lower sensitivities than Wake, N2, and REM (Table 3). To examine the inter-rater variability of sleep staging, 10 skilled EEG analysts analyzed EEG data from 70 patients in the same sleep period,
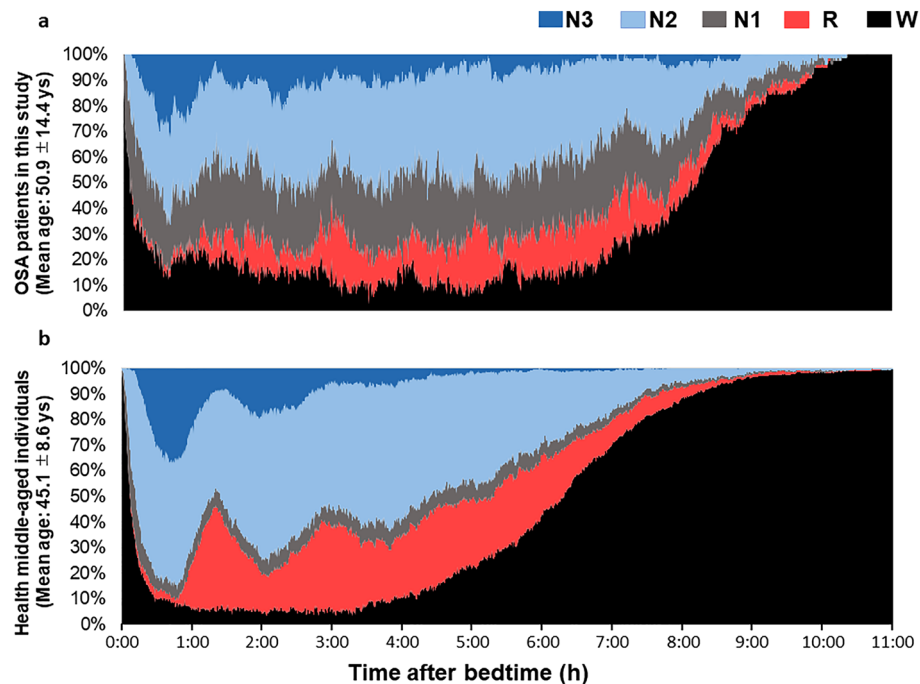
Figure 1. (**a**) Cumulative display of sleep architecture in all 77 patients with OSA in this study. (**b**) Cumulative display of sleep architecture in participants from a previous study who were healthy middle-aged individuals[14]. The percentage of patients in each sleep stage is shown for stage W (black), N1 (gray), N2 (light blue), N3 (blue), and R (red). W, wake; R, rapid-eye-movement sleep; N, non-REM sleep.

	Type I PSG	Portable EEG device
	Mean ± SD	Mean ± SD
Total sleep time, min	425.8 ± 69.5	420.4 ± 69.4
Sleep onset latency, min	10.7 ± 18.4	13.5 ± 24.1
WASO, min	62 ± 46	65.6 ± 48.1
Sleep efficiency, %	85.4 ± 9.4	84.2 ± 10.1
REM latency, min	110.5 ± 67.5	118.7 ± 77.6
N1, %	31.7 ± 13.4	23.3 ± 12.1
N2, %	42.7 ± 10.5	52.5 ± 10.9
N3, %	10.0 ± 7.8	8.0 ± 7.4
REM, %	15.6 ± 4.6	16.2 ± 5.7
Arousal index, index	37.0 ± 15.1	32.0 ± 15.6
AHI and REI, index	30.7 ± 22.7	22.5 ± 22.4

Table 1. Sleep parameters measured in both devices. *PSG* polysomnography, *WASO* wake after sleep onset, *REM* rapid eye movement sleep, *N* non-REM sleep, *AHI* apnea–hypopnea index, REI *REI* depending on SpO_2, AHI measure by type I PSG, and REI measured by portable EEG device with SpO_2.

with results consistent with those of previous studies. Specifically, there was poor agreement for N1 and N3[21]. Considering that the pattern of N1 decreased with the increasing severity of apneic symptoms (Supplementary Fig. S2), it might be necessary for future studies to further refine the EEG criteria for patients with apnea. A previous study[22] investigating the agreement between portable single-channel EEG devices and type I PSG in relatively healthy middle-aged and elderly persons (without apneic symptoms) revealed similar findings, although the overall accuracy, specificity, and sensitivity were lower than those observed in the present study[22]. Notably, the N1 value of 20.6 is low[22]. Referring to the confusion matrix results (Table 2), there was a reduction in sensitivity for determining N1 and N3 compared to the other stages (Table 3). The portable EEG device frequently misjudged type I PSG's N1 as N2 (28.2%) and N3 as N2 (32.7%).

The reasons for this misjudgment could not be determined in this study; however, there are three possibilities. First, as with the portable single-channel EEG device[21], the use of fewer channels than in type I PSG may have contributed to a decrease in sensitivity for N1, which requires more detailed assessments[23]. Indeed, concerning
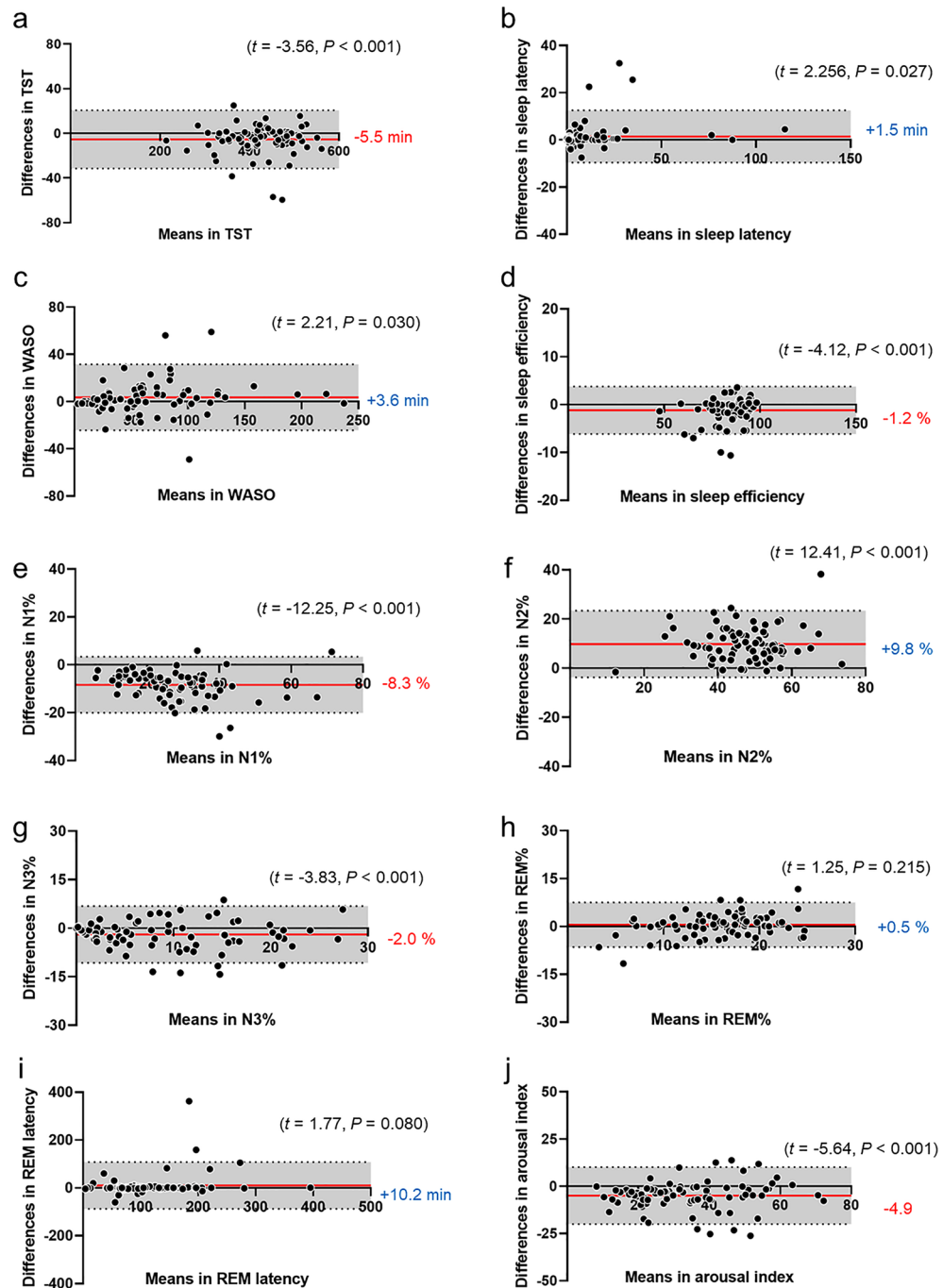
3

Figure 2. Bland–Altman plots comparing type I PSG and portable EEG device values. The Bland–Altman plots depict the mean bias (red solid line) and upper and lower limits of agreement (1.96 standard deviations from bias; black dashed lines) for each sleep parameter for the portable EEG device compared with type I PSG. The parentheses contain information about the t-test, and if the P-value is significant, it indicates that the bias, depicted to the right of the solid red line, is statistically significant. A positive solid line signifies an overestimation of the portable EEG device, while a negative line signifies an underestimation.

forehead EEG devices, as emphasized in previous studies, sweat artifacts, which manifest as high-amplitude and low-frequency waves, are associated with the challenge of observing the N2/N3 transition without an epoch delay, causing a decrease in sensitivity[24]. Second, patients with OSA experience more noise than healthy subjects[25], which could lead to the misclassification of N1 as Wake or REM or N2 as N3, even when analyzed by a sleep EEG analyst. This shows that the interprofessional agreement in EEG analysis was 82.0%, which is consistent with that of a previous study[26]. Although the sensitivity of N1 and N3 in determining sleep stage was low, the agreement was generally high; hence, there is a possibility of home measurements in the future with a portable EEG device.
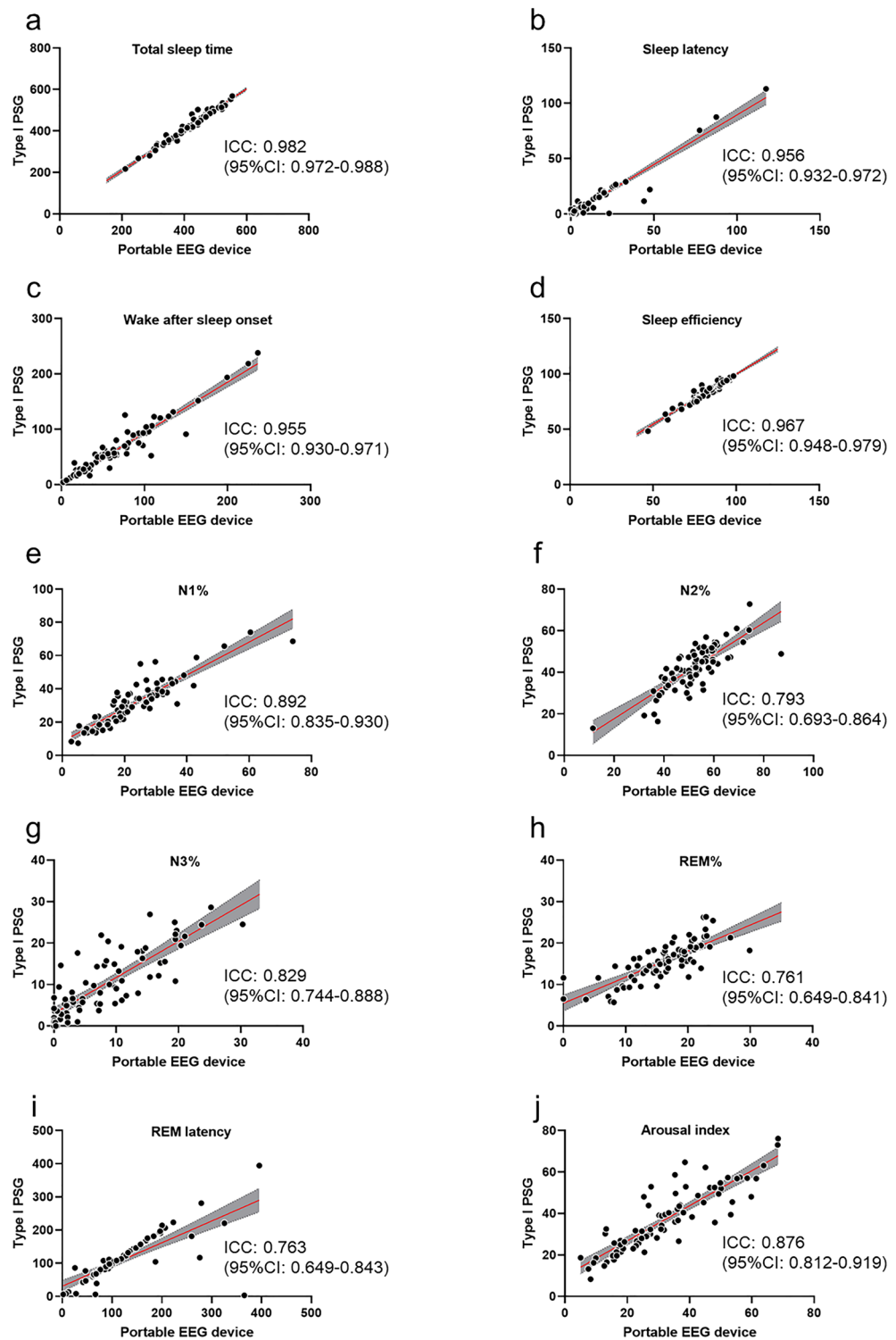
Figure 3. Relationship between the type I PSG and portable EEG device, showing the intraclass correlation coefficient (red solid line with 95% CI).

Third, drawing on our earlier studies[14,15], patients with OSA exhibited reduced N2 sleep and increased N1 sleep (there might have been a misinterpretation of N1 and N2) (Fig. 1a and b), even when their AHI was not greater than 100[27]. Since inter-rater reliability is lowest for N1 determination[28], it is reasonable that N1 determination performance would decrease with the severity of OSA in patients with increased N1.

To reduce the burden of OSA, various devices have been developed, including wearable devices (e.g., portable EEG devices and/or actigraphy) and smartphone applications[5,17]. Although numerous studies on healthy
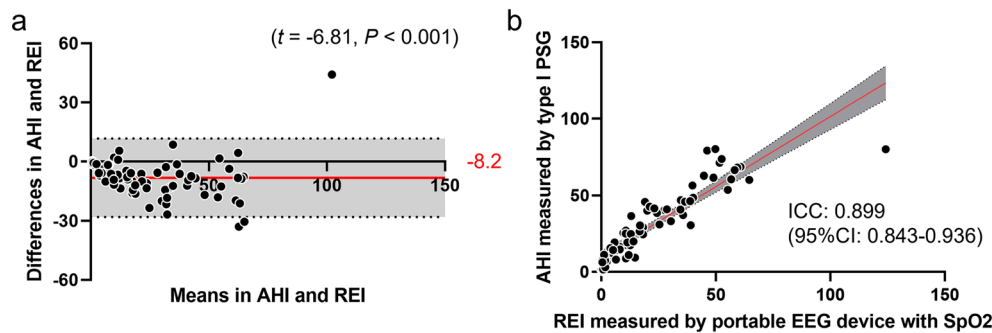
Figure 4. (**a**) Bland–Altman plot and (**b**) intraclass correlation coefficient was calculated between the AHI of type I PSG and the REI from portable EEG device with SpO$_2$. (**a**) The Bland–Altman plots depict the mean bias (red solid line) and upper and lower limits of agreement (1.96 standard deviations from bias; black dashed lines) for each sleep parameter for the portable EEG device compared with type I PSG. The parentheses contain information about the t-test, and if the *P*-value is significant, it indicates that the bias, depicted to the right of the solid red line, is statistically significant. A positive solid line signifies an overestimation of the portable EEG device, while a negative line signifies an underestimation. (**b**) Intraclass correlation coefficient with a red solid line and 95% confidence interval.

		Portable EEG device				
		Wake	N1	N2	N3	REM	Total
	Wake	10,032 (91.7)	668 (6.1)	155 (1.4)	1 (0.0)	83 (0.8)	10,939
	N1	1,623 (8.0)	11,380 (55.8)	5,749 (28.2)	121 (0.6)	1,511 (7.4)	20,384
Type I PSG	N2	114 (0.4)	1,590 (5.7)	24,821 (90.0)	849 (3.1)	229 (0.8)	27,703
	N3	4 (0.1)	46 (0.7)	2,075 (32.7)	4,230 (66.6)	0 (0.0)	6,355
	REM	91 (0.9)	906 (8.8)	527 (5.1)	0 (0.0)	8,772 (85.2)	10,296
	Total	11,864	14,590	33,427	5,201	10,595	75,677

Table 2. Confusion matrix for sleep stage classification in all epochs of all recordings (n = 75,677). Parentheses indicate percentage agreement between two devices. *PSG* polysomnography, *REM* rapid eye movement sleep, *N* non-REM sleep.

	Accuracy	Sensitivity	Specificity	PABAK
Wake	96.4 (2.8)	88.2 (13.7)	97.1 (2.9)	0.93 (0.1)
N1	83.9 (6.2)	54.3 (13.1)	93.8 (4.2)	0.68 (0.1)
N2	85.0 (5.5)	88.6 (8.5)	82.2 (7.3)	0.70 (0.1)
N3	95.8 (3.0)	53.9 (33.7)	98.6 (1.8)	0.92 (0.1)
REM	95.6 (2.1)	84.1 (15.5)	97.3 (2.1)	0.91 (0.1)

Table 3. Overall epoch-by-epoch analysis for type I PSG and portable EEG device. Standard errors are in parentheses. *PSG* polysomnography, *REM* rapid eye movement sleep, *N* non-REM sleep, *PABAK* prevalence and bias-adjusted kappa.

subjects have demonstrated results comparable to those of Type I PSG, only a limited number of studies have been conducted on patients with OSA[16]. In healthy subjects, there is high agreement between wearable devices and PSG[5]. However, it has been reported that assessing sleep onset latency is challenging in patients with OSA, although some studies have incorporated various algorithms to enhance accuracy[29–31].

Furthermore, this study demonstrated that the arousal index, calculated from a portable EEG device, can effectively screen for moderate (AHI ≥ 15) and severe (AHI ≥ 30) apnea. Notably, screening efficacy was higher when using AHI ≥ 30 as the criterion than AHI ≥ 15, as shown in Fig. 4. Our results confirmed that the AHI
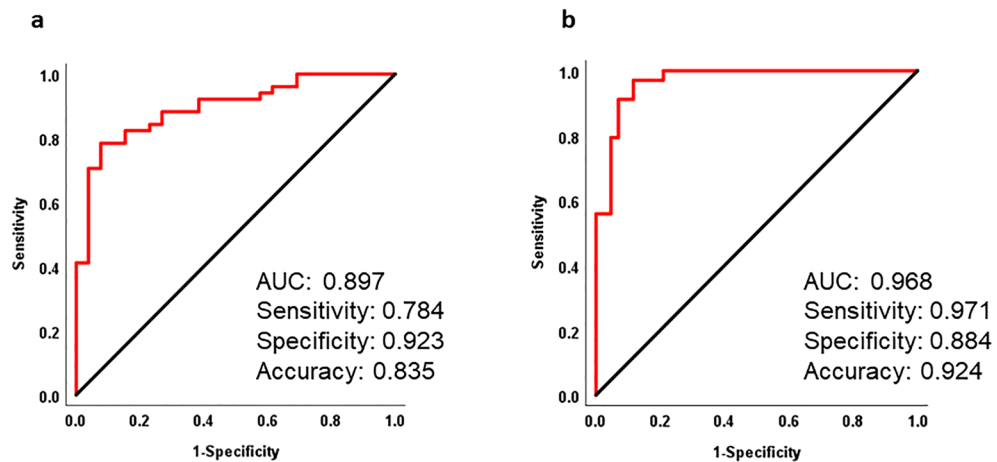
Figure 5. ROC curves for OSA diagnosis using the arousal index measured using the portable EEG device. (**a**) ROC curve for OSA diagnosis using the arousal index at an AHI threshold of ≥ 15 events/h. (**b**) ROC curve for OSA diagnosis using the arousal index at an AHI threshold of ≥ 30 events/h. OSA, obstructive sleep apnea. Both (**a**) and (**b**) ROC curves adjusted for sex, age, and BMI.

cut-off values of 15 and 30 or higher are an EEG-based arousal index ≥ 25 and ≥ 32, respectively. While previous studies have reported that the arousal index of type I PSG can be used to screen for moderate (AHI ≥ 15) and severe (AHI ≥ 30) apnea[18,19], results from the portable EEG device align in a manner that enables screening with comparable efficacy.

The use of portable EEG device systems has become increasingly institutionalized. Meta-analyses have reported respiratory event measurements (e.g., AHI, ODI, respiratory disturbance index, oxygen saturation, and lowest oxyhemoglobin saturation) from both type I PSG and portable EEG devices conducted at home to be within the range of 0.832–0.942[32]. This study further suggests that the sleep EEG of patients with apnea can be accurately assessed using a portable EEG device. Recent advancements in portable EEG devices, which are easier to use than type II–IV PSG, are expected to lead to further developments in the future. The portable EEG device used in this study uploads the EDF data to the cloud as soon as the overnight measurement is complete. Users can promptly review the previous night's sleep results upon waking up, utilizing machine learning and AI to determine their sleep stages[33,34]. In this study, a professional EEG analyst performed all assessments to ensure reliability. However, portable PSGs equipped with AI analysis capabilities are expected to advance further.

Our study had a few limitations. Previous studies have proposed integrating various portable devices that can be easily measured at home to improve measurement accuracy[21]. We also performed sleep staging by combining SpO_2 and pulse rate obtained with type I PSG on a portable EEG device and found a high concordance rate. However, further studies are required to determine a simple method for recording respiratory functions and heart rate variability on a portable EEG device. Second, to maintain natural sleep patterns, no restrictions were placed on sleep duration, resulting in varying total sleep times among the participants. Sleep stages were quantified as percentages relative to total sleep time. It is important to mention that participants were required to sleep at a medical institution's sleep laboratory to enable a direct comparison with type I PSG. Therefore, future studies are required to validate these findings in the home environment. Third, while this study demonstrated results comparable to those of type I PSG in screening for an AHI of 15 or 30 using the arousal index measured by a portable EEG device, it remains unclear whether it can be used to screen healthy individuals. This is because the study population included patients with apnea, suggesting that further research is required in this area. Finally, out of a total of 90 measurements taken with the portable EEG device, there were 11 cases (12%) of instrumental problems, including electrode disconnection and measurement time deviation. Failure rates ranging from 3–18% for measurements obtained solely with a portable EEG device have been reported[35], which is considered a reasonable failure rate given the study's simultaneous measurement of type I PSG and the portable EEG device. Additionally, future studies should examine the inter-rater validity of test–retest reliability.

Conclusions

In patients with OSA, our portable EEG device demonstrated good agreement with type I PSG, although more work is required to improve the agreement for N1 and N3 screening. Measurements could be conducted in a home setting, significantly reducing the burden on patients with OSA. Furthermore, the EEG-based arousal index, which predicts the degree of apnea without the need for a separate device to measure respiratory events, showed good results even with the portable EEG device in this study. This suggests that patients with OSA may be able to comfortably use this device in their homes to determine their degree of apnea.

Methods

Study design and participants

This cross-sectional study was conducted between August 2021 and March 2022. A total of 90 patients with OSA were recruited from the Sleep Center of Shunkaikai Inoue Hospital, Nagasaki, Japan. The inclusion criteria were age > 20 years and suspected of having an OSA with an AHI measurement ≥ 5 or a 3% ODI of ≥ 5 on an at-home test with a wearable device or a simple pulse oximeter. Instrumental problems, including measurement time deviations, were experienced with Type I PSGs (n = 2) and portable EEG devices (n = 5). Additionally, portable EEG devices with electrode-dislodgement measurement failures (n = 6) were excluded from the analysis. Finally, 77 patients with OSA were included in the analysis. Participants with suspected sleep apnea wore both type I PSG and portable EEG devices simultaneously and slept and woke up at their usual time. As soon as the patient went to bed, the lights were turned off, and the Type I PSG and portable EEG device were started. Similarly, as soon as the patient woke up, the lights were turned on, and the recording from both devices was stopped. In addition, sleep staging calculations were performed after matching the time of light on/off on both devices.

This study was performed in accordance with the principles of the Declaration of Helsinki, and all experimental protocols were approved by the Ethics Committee of the Japan Medical Association (Ref. 2014-4). Additionally, all participants provided informed consent.

Measurements

Standards polysomnography (type I PSG)
Type I PSG was recorded using Alice 6 LDE (Philips Respironics, PA, USA). The recording system of this device consisted of six EEG electrode sites (F3–M2, F4–M1, C3–M2, C4–M1, O1–M2, and O2–M1). Two electrooculograms and one submental electromyogram with pulse rate and SpO_2 were adopted and recorded during sleep. The records were scored every 30 s to classify the sleep stages as awake (stage W), non-rapid eye movement (REM) (stage N), N1, N2, N3, and REM. We calculated the percentage of total sleep time for each sleep stage. Measurements during sleep onset latency were classified as stage W. REM latency was measured until the first R stage, which appeared after sleep onset. Sleep efficiency was calculated as the sum of N1, N2, N3, and REM sleep (i.e., total sleep time) divided by the total time spent in bed, multiplied by 100. Wake after sleep onset (WASO) was defined according to standard AASM criteria[36]. Additionally, we calculated the arousal index during the sleep period as follows: the number of arousals on EEG divided by the total sleep time (h). The AHI was calculated by assessing the frequency of apneas and hypopneas based on oxygen saturation, nasal respiratory flow sensor readings, and respiratory capacity determined by thoracoabdominal movement per total sleep time measured using type I PSG. To assess whether the arousal index measured from each EEG could function as a screening tool for AHI, we categorized the AHI into moderate apnea (AHI ≥ 15) and severe apnea (AHI ≥ 30)[10].

Portable EEG device
A portable EEG was recorded using the Insomnograf K2 (S'UIMIN Inc., Tokyo, Japan), which is lightweight (162 g) and easily attached and detached because of the soft-sticking integrated electrodes (Supplementary Fig. S1). In healthy subjects, it demonstrated 86.9% agreement with type I PSG and a kappa coefficient of 0.80[14,15]. The recording system of this device consisted of four EEG electrodes (Fp1, Fp2, M1, and M2) and one reference electrode (Fpz), according to the 10–20 system. The montage was combined with four electroencephalogram derivations (Fp1–M2, Fp2–M1, Fp1–average M, and Fp2–average M), using Fp1–Fp2 and Fp2–Fp1 for left and right electrooculography, respectively, and M1–M2 for chin electromyography to analyze sleep staging. To enhance the accuracy of sleep staging based on respiratory events, we integrated the pulse rate and SpO_2 data from a type I PSG with EEG data from a portable EEG device for sleep-staging determination[37]. However, the portable EEG device, pulse, and SpO_2 data sheets were blinded as a separate set from the type I PSG. We calculated the REI depending on SpO_2 (REI) (i.e., it shares the same concept as AHI) by counting the total number of respiratory events based on oxygen saturation, pulse rate, and EEG-detected respiratory events and dividing by the total sleep time recorded by the portable EEG device. To classify the sleep stages, recordings were made every 30 s, and criteria similar to the type I PSG method described above were used[36]. Sleep staging for type I PSG and portable EEG devices was performed by the same sleep EEG analyst with 40 years of working experience. The portable EEG and type I PSG EDF were used to determine sleep staging by a sleep EEG analyst who had no visibility of the respective results. A third-party researcher linked both datasheets.

Potential confounders

To identify potential confounders, we included continuous variables such as age, body mass index (BMI), and abdominal and neck circumference (continuous variable). Categorical variables ("yes" or "no"), including sex ("male" or "female"), tobacco-smoking status ("current" or "previous/never"), alcohol consumption ("drinker" or "non-drinker"), and medical history of hypertension, diabetes, dyslipidemia, and hyperuricemia, were evaluated. Furthermore, depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale[38]. Finally, subjective sleep parameters were measured using the Epworth Sleepiness Scale[39] and the Athens Insomnia Scale[40].

Statistical analysis

Bland–Altman analysis was used to compare each device, illustrating individual night differences compared to type I PSG and portable EEG devices. This analysis also presented the overall levels of bias and the LOA. A statistically significant paired t-test was used to assess bias. Furthermore, the ICC and 95% confidence interval (CI) were calculated for both devices. ICC values > 0.75, 0.40–0.75, and < 0.40 were very good, fair to good, and poor, respectively[41]. Additionally, an epoch-by-epoch analysis was conducted across all 30 s epochs for each

device. This analysis involved calculating the accuracy, sensitivity, specificity, prevalence, and bias-adjusted kappa (PABAK) by referring to prior studies[42].

Receiver Operating Characteristic (ROC) curve analysis was performed to investigate the potential of the arousal index in screening for apnea. This analysis utilized moderate (AHI ≥ 15) and severe (AHI ≥ 30) apnea as dependent variables while adjusting for independent variables such as sex, age, BMI, and arousal index measured with a portable EEG device. The AUC, accuracy, sensitivity, and specificity were calculated.

Bland–Altman and epoch-by-epoch analyses were conducted using R software (R Foundation for Statistical Computing, Vienna, Austria), and the ICC and ROC curve analyses were conducted using SPSS version 26.0 (IBM Corporation, Armonk, NY, USA). A two-tailed statistical significance was set at $P < 0.05$ (two-tailed).

Data availability

The data used in this study was licensed to S'UIMIN, Inc. The data have not been made publicly available and can be used in the future for the development of medical devices and diagnostic technologies. Proposals and requests for data access should be emailed to the corresponding authors.

References

1. Di, H. *et al.* Evaluation of sleep habits and disturbances among US adults, 2017–2020. *JAMA Netw. Open* **5**, e2240788. https://doi.org/10.1001/jamanetworkopen.2022.40788 (2022).
2. Khan, H., Kella, D., Kunutsor, S. K., Savonen, K. & Laukkanen, J. A. Sleep duration and risk of fatal coronary heart disease, sudden cardiac death, cancer death, and all-cause mortality. *Am. J. Med.* **131**, 1499–1505. https://doi.org/10.1016/j.amjmed.2018.07.010 (2018).
3. Benjafield, A. V. *et al.* Estimation of the global prevalence and burden of obstructive sleep apnoea: A literature-based analysis. *Lancet Respir. Med.* **7**, 687–698. https://doi.org/10.1016/S2213-2600(19)30198-5 (2019).
4. Peppard, P. E. *et al.* Increased prevalence of sleep-disordered breathing in adults. *Am. J. Epidemiol.* **177**, 1006–1014. https://doi.org/10.1093/aje/kws342 (2013).
5. Scott, H., Lack, L. & Lovato, N. A systematic review of the accuracy of sleep wearable devices for estimating sleep onset. *Sleep Med. Rev.* **49**, 101227. https://doi.org/10.1016/j.smrv.2019.101227 (2020).
6. Johnson, D. A., Billings, M. E. & Hale, L. Environmental determinants of insufficient sleep and sleep disorders: Implications for population Hhealth. *Curr. Epidemiol. Rep.* **5**, 61–69. https://doi.org/10.1007/s40471-018-0139-y (2018).
7. The Ministry of Health, Labor and Welfare. The 8th NDB Open Data. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000177221_00012.html (2023).
8. Chinoy, E. D. *et al.* Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep* **44**, zsaa291. https://doi.org/10.1093/sleep/zsaa291 (2021).
9. Kim, K., Park, D. Y., Song, Y. J., Han, S. & Kim, H. J. Consumer-grade sleep trackers are still not up to par compared to polysomnography. *Sleep Breath* **26**, 1573–1582. https://doi.org/10.1007/s11325-021-02493-y (2022).
10. Xu, Y., Ou, Q., Cheng, Y., Lao, M. & Pei, G. Comparative study of a wearable intelligent sleep monitor and polysomnography monitor for the diagnosis of obstructive sleep apnea. *Sleep Breath* **27**, 205–212. https://doi.org/10.1007/s11325-022-02599-x (2023).
11. Miller, D. J., Sargent, C. & Roach, G. D. A validation of six wearable devices for estimating sleep, heart rate and heart rate variability in healthy adults. *Sensors (Basel)* **22**, 6317. https://doi.org/10.3390/s22166317 (2022).
12. Douglas, J. A. *et al.* Guidelines for sleep studies in adults—A position statement of the Australasian Sleep Association. *Sleep Med.* **36**, S2–S22. https://doi.org/10.1016/j.sleep.2017.03.019 (2017).
13. Bruyneel, M., Van den Broecke, S., Libert, W. & Ninane, V. Real-time attended home-polysomnography with telematic data transmission. *Int. J. Med. Inform.* **82**, 696–701. https://doi.org/10.1016/j.ijmedinf.2013.02.008 (2013).
14. Iwagami, M. *et al.* Association between electroencephalogram-based head characteristics and physical health in the general adult population. *Sci. Rep.* **13**, 21545. https://doi.org/10.1038/s41598-023-47979-9 (2023).
15. Seol, J. *et al.* Bidirectional associations between physical activity and sleep in older adults: A multilevel analysis using polysomnography. *Sci. Rep.* **12**, 15399 (2022).
16. Guillodo, E. *et al.* Clinical applications of mobile health wearable-based sleep monitoring: Systematic review. *JMIR Mhealth Uhealth* **8**, e10733. https://doi.org/10.2196/10733 (2020).
17. Mehrabadi, M. A. *et al.* Sleep tracking of a commercially available smart ring and smartwatch against medical-grade actigraphy in everyday settings: Instrument validation study. *JMIR Mhealth Uhealth* **8**, e20465. https://doi.org/10.2196/20465 (2020).
18. Ren, R. *et al.* Association between arousals during sleep and hypertension among patients with obstructive sleep apnea. *J. Am. Heart Assoc.* **11**, e022141. https://doi.org/10.1161/JAHA.121.022141 (2022).
19. Taguchi, Y. *et al.* Successful screening of sleep-disordered breathing using a pacemaker-based algorithm in Japan. *J. Cardiol.* **73**, 394–400. https://doi.org/10.1016/j.jjcc.2018.10.007 (2019).
20. Imtiaz, S. A. A Systematic review of sensing technologies for wearable sleep staging. *Sensors (Basel)* **21**, 1562. https://doi.org/10.3390/s21051562 (2021).
21. Younes, M. *et al.* Reliability of the American Academy of Sleep Medicine rules for assessing sleep depth in clinical practice. *J. Clin. Sleep Med.* **14**, 205–213. https://doi.org/10.5664/jcsm.6934 (2018).
22. Lucey, B. P. *et al.* Comparison of a single-channel EEG sleep study to polysomnography. *J. Sleep Res.* **25**, 625–635. https://doi.org/10.1111/jsr.12417 (2016).
23. Matsuo, M. *et al.* Comparisons of portable sleep monitors of different modalities: Potential as naturalistic sleep recorders. *Front. Neurol.* **7**, 110. https://doi.org/10.3389/fneur.2016.00110 (2016).
24. Myllymaa, S. *et al.* Assessment of the suitability of using a forehead EEG electrode set and chin EMG electrodes for sleep staging in polysomnography. *J. Sleep Res.* **25**, 636–645. https://doi.org/10.1111/jsr.12425 (2016).
25. Kapur, V. 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.* **13**, 479–504. https://doi.org/10.5664/jcsm.6506 (2017).
26. Danker-Hopfe, H. *et al.* Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J. Sleep Res.* **13**, 63–69. https://doi.org/10.1046/j.1365-2869.2003.00375.x (2004).
27. Sata, N. *et al.* Clinical, polysomnographic, and cephalometric features of obstructive sleep apnea with AHI over 100. *Sleep Breath* **25**, 1379–1387. https://doi.org/10.1007/s11325-020-02241-8 (2021).
28. Lee, Y. J., Lee, J. Y., Cho, J. H. & Choi, J. H. Interrater reliability of sleep stage scoring: A meta-analysis. *J. Clin. Sleep Med.* **18**, 193–202. https://doi.org/10.5664/jcsm.9538 (2022).

29. Hedner, J. *et al.* A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients. *Sleep* **27**, 1560–1566. https://doi.org/10.1093/sleep/27.8.1560 (2004).
30. Chae, K. Y. *et al.* Evaluation of immobility time for sleep latency in actigraphy. *Sleep Med.* **10**, 621–625. https://doi.org/10.1016/j.sleep.2008.07.009 (2009).
31. Dick, R. *et al.* AASM standards of practice compliant validation of actigraphic sleep analysis from SOMNOwatch™ versus polysomnographic sleep diagnostics shows high conformity also among subjects with sleep-disordered breathing. *Physiol. Meas.* **31**, 1623–1633. https://doi.org/10.1088/0967-3334/31/12/005 (2010).
32. Cagle, J. L. *et al.* Portable sleep study device versus polysomnography: A meta-analysis. *Otolaryngol. Head Neck Surg.* **168**, 944–955. https://doi.org/10.1002/ohn.179 (2023).
33. Horie, K. *et al.* Automated sleep stage scoring employing a reasoning mechanism and evaluation of its explainability. *Sci. Rep.* **12**, 12799. https://doi.org/10.1038/s41598-022-16334-9 (2022).
34. S'UIMIN inc. home page. https://www.suimin.co.jp/. Accessed 12 October 2023.
35. Collop, N. A. *et al.* 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**, 737–747 (2007).
36. Berry, R. B. *et al.* *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications* (American Academy of Sleep Medicine, 2017).
37. Kwon, S., Kim, H. & Yeo, W. H. Recent advances in wearable sensors and portable electronics for sleep monitoring. *iScience* **24**, 102461. https://doi.org/10.1016/j.isci.2021.102461 (2021).
38. Manson, S. M. *et al.* Depressive symptoms among American Indian adolescents: Psychometric characteristics of the Center for Epidemiologic Studies Depression Scale (CES-D). *J. Consult. Clin. Psychol.* **2**, 231–237. https://doi.org/10.1037/1040-3590.2.3.231 (1990).
39. Johns, M. W. A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep* **14**, 540–545 (1991).
40. Soldatos, C. R., Dikeos, D. G. & Paparrigopoulos, T. J. Athens Insomnia Scale: Validation of an instrument based on ICD-10 criteria. *J. Psychosom. Res.* **48**, 555–560. https://doi.org/10.1016/s0022-3999(00)00095-7 (2000).
41. Rosner, B. *Fundamentals of Biostatistics* 6th edn. (Thomson-Brooks/Cole, 2006).
42. De Zambotti, M., Cellini, N., Goldstone, A., Colrain, I. M. & Baker, F. C. Wearable sleep technology in clinical and research settings. *Med. Sci. Sports Exerc.* **51**, 1538–1557. https://doi.org/10.1249/MSS.0000000000001947 (2019).

Acknowledgements

Author contributions

S.C., M.T., T.A., A.T., T.H., H.K., and M.Y. designed the study, and H.Y. contributed to data collection. J.S., S.T., and M.M. performed the formal analysis. F.K. analyzed the EEG signals and sleep staging. J.S. wrote the first draft of the manuscript. S.C., F.K., S.T., M.M., M.T., T.A., A.T., T.H., H.Y., H.K., and M.Y. substantially contributed to the interpretation of the results and provided feedback on the main manuscript. All the authors reviewed the manuscript.

Funding

Competing interests

This study was funded by S'UIMIN Inc.; J.S. and F.K. received an honorarium from S'UIMIN Inc.; M.T., A.T., and T.A. are employees of S'UIMIN Inc. M.Y. and T.H. are the board members and hold stock in the company.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-53827-1.

Correspondence and requests for materials should be addressed to J.S. or M.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.