Article

Not peer-reviewed version

Evaluating the Feasibility of Low-Cost, Contactless Consumer Sleep-Tracking Devices as Measurement Tools for Preliminary Sleep Research

Huifang Zhai , Yonghong Yan * , Litao Gao , Siqi He , Xiaowan Dong , Xiang Cheng , Tao Hu

Posted Date: 26 February 2026

doi: 10.20944/preprints202602.1395.v1

Keywords:

sleep; actigraphy; contactless; validation; sensors; polysomnography

Article

Evaluating the Feasibility of Low-Cost, Contactless Consumer Sleep-Tracking Devices as Measurement Tools for Preliminary Sleep Research

Huifang Zhai [1,2], Yonghong Yan [1,2,*], Litao Gao [3], Siqi He [4], Xiaowan Dong [5], Xiang Cheng [1,2] and Tao Hu [1,2]

[1] School of Architecture and Urban Planning, Chongqing University, Chongqing 400044, China
[2] Key Laboratory of the Ministry of Education of Mountainous City and Towns Construction and New Technology, Chongqing University, Chongqing 400044, China
[3] School of Electrical and Electronic Engineering, Chongqing Vocational and Technical University of Mechatronics, Chongqing 402760, China
[4] College of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, China
[5] School of Urban Planning and Architecture, Henan University of Urban Construction, Henan 467036, China
* Correspondence: yanyonghong@cqu.edu.cn

Highlights

What are the main findings?

- For all sleep metrics except sleep stages, the iSleep S200G and Sleep Dot B501 showed no significant differences compared to PSG.
- The iSleep S200G and Sleep Dot B501 demonstrate comparable or even superior reliability levels to Actiwatch Spectrum in sleep/wake measurements and EBE agreement.

What are the implications of the main findings?

- This study offers a potential solution for conducting long-term, large-scale preliminary sleep research with low precision requirements under budget constraints.
- The performance validation results provide data references for researchers adopting the iterative upgrade of CCSTDs.

Abstract

Compared to polysomnography (PSG) and actigraphy, contactless consumer sleep-tracking devices (CCSTDs) are low-cost, user-friendly, and non-disruptive of sleep. This study evaluates the performance of two inexpensive and representative first-generation Chinese-made CCSTDs—the iSleep S200G and Sleep Dot B501—compared with PSG and actigraphy using standardized validation protocols. The objective is to assess their feasibility as alternatives for large-scale, long-term preliminary research that do not rely on single-day high-precision sleep data. Eleven healthy young adults (mean age = 26.5 ± 4.8 years) participated in a two-night sleep laboratory study using four devices in parallel. The iSleep S200G exhibited no significant differences compared to PSG in TST and SE, while the Sleep Dot B501 showed no significant differences in TST, SE, SOL, and WASO. The intraclass correlation coefficient values and epoch-by-epoch agreement of iSleep S200G and Sleep Dot B501 were comparable to or superior to actigraphy. Notably, the Epoch-by-epoch agreement of both devices are non-inferior to consumer-grade sleep devices already used for long-term, large-scale sleep monitoring. Therefore, even within budget constraints, first-generation CCSTDs can effectively meet the requirements for long-term, large-scale sleep monitoring without sleep stages detection. The results also provide data references for researchers adopting the iterative upgrade of CCSTDs.

Keywords: sleep; actigraphy; contactless; validation; sensors; polysomnography

 doi:10.20944/preprints202602.1395.v1

1. Introduction

Long-term, large-scale data collection for preliminary sleep assessments—such as sleep structure, habits, and rhythms—is constrained by the high costs and specialized operational requirements of gold-standard polysomnography (PSG) and silver-standard actigraphy, particularly in economically underdeveloped regions with low levels of education and poor living conditions. As a result, their widespread use in such settings is often impractical. To meet the needs of these unique application scenarios, there is an urgent need to explore a more suitable alternative monitoring solution. Such a device should be low-cost, easy to operate, and simple to process data, with strong environmental adaptability, enabling stable operation under rudimentary conditions and resistance to damage. Consumer-grade sleep monitoring devices offer a viable approach—recent advancements in technology have led to a reduction in the disparity between consumer-grade and research-grade sleep monitoring devices [1,2]. Numerous validation studies, using recognized validation standards [3–5], have also confirmed that many consumer-grade sleep monitors have achieved or surpassed the accuracy of actigraphy [6–10]. Several validated commercial sleep devices, including Huawei FIT 2 [11], Withings sleep analyzer [12], Readiband [8], HSL-102M [13], and SleepScore Max [8], have been extensively utilized in sleep monitoring experiments in various fields such as health lighting [14,15], healthcare [16,17], and sports [18–20]. These application results indicate that consumer sleep devices have unparalleled advantages in large-scale and long-term sleep monitoring and tracking.

However, the wide variety of consumer-grade sleep monitoring devices, coupled with significant price differences and a product update cycle far outpacing the pace of validation studies [21], complicates the selection process for researchers. At the same time, considering that real-world usage scenarios may involve certain vulnerable and sensitive individuals, such as burn victims, infants, and psychiatric patients, wearable consumer sleep devices may still cause interference [22–25]. Based on this, this study focuses on contactless consumer sleep-tracking devices (CCSTDs), selecting two representative low-cost first-generation CCSTDs in the Chinese market—the iSleep S200G and Sleep Dot B501—as validation subjects. Based on the validation results, this study evaluates the feasibility of using low-cost CCSTDs based on first-generation technology for long-term, large-scale preliminary sleep research. Simultaneously, it provides performance data references for researchers utilizing the iteratively upgraded CCSTDs. The performance and reliability of two CCSTDs were validated in a sleep laboratory. The study utilized a sample of healthy young adults and compared the results with PSG and actigraphy. The validation experiment followed the research reports and guidelines for sleep device performance proposed by the Sleep Research Society and Menghini et al. [4,26].

2. Materials and Methods

2.1. Participants

In this study, a total of 11 healthy young adults, including 6 men and 5 women, with an average age of 26.5 ± 4.8 years (mean ± SD) and a body mass index calculated based on height and weight of 21.9 ± 2.3 kg/m² (mean ± SD), were included as participants. These individuals had no neurological or psychiatric diagnoses, sleep disorders, or other medical conditions. Alcohol, caffeine, and any drugs were strictly prohibited during the research. Participants were recruited at Chongqing University between June and July 2020. This study was conducted in strict accordance with the ethical guidelines of the Declaration of Helsinki. It has been approved by the Ethics Committee of South China Normal University (No. SCNU-AOE-2025-018). All participants voluntarily signed written informed consent forms after receiving complete disclosure.

2.2. Sleep Monitors

2.2.1. Contactless Consumer Sleep-Tracking Devices

 doi:10.20944/preprints202602.1395.v1

For this study, we selected two representative CCSTDs. The iSleep S200G, recognized as the first CCSTD launched in China [27], was developed by XIKANG ALPS Technology Co., Ltd. (Shenyang, Liaoning Province, China) and released in 2014 at a price of $180[28]. The Sleep Dot B501, the affordable CCSTD (priced below $14), was manufactured by Sleepace Co., Ltd. (Shenzhen, China) and introduced in 2015[29].

The iSleep S200G uses a 2.45 GHz radio-frequency micro-power radar, which effectively penetrates non-metallic substances such as bedding and clothing, to monitor body movement and respiration [30]. The sleep monitoring device is placed approximately 0.5 to 0.8 meters from the body, oriented toward the chest. Sleep data can be accessed and monitored via the display screen and the associated application. The monitored sleep parameters include total time in bed, total sleep time (TST), total wake time (TWT), sleep efficiency (SE), light sleep (N1), moderate sleep (N2), and deep sleep (N3).

The Sleep Dot B501 is equipped with a magnetic clasp designed for attachment to the edge of a pillow. The device uses a 3-axis accelerometer to monitor sleep patterns by tracking body movements. It connects to a mobile application through Bluetooth, enabling users to access and analyze their sleep data. The sleep parameters that can be directly measured include total time in bed, TST, TWT, SE, and the various stages of sleep, such as N1, N2, and N3. Sleep onset latency (SOL) and wake after sleep onset (WASO) can be calculated using the recorded bedtime, sleep onset time, and TWT provided by the device. We additionally reached out to Sleepace to acquire the original data from the 60-second epochs. Light sleep, moderate sleep, and deep sleep in both the iSleep S200G and Sleep Dot B501 correspond to the N1, N2, and N3 stages in PSG, respectively.

2.2.2. PSG and Actigraphy

We used the SOMNOtouch™ RESP (SOMNOmedics GmbH & Co. KG, Germany), which is currently the most compact full-color touchscreen PSG device available on the market. This device was utilized to collect sleep data from various sources, including electroencephalography recorded at the F3, C3, O1, Fpz, and Cz sites with reference to the contralateral mastoids; bilateral electrooculography; mentalis electromyography; tibialis anterior electromyography; pulse oximetry; electrocardiography; and a respiratory monitoring sensor. At the beginning of the recordings, impedances were equal to or less than 10 kΩ. PSG electrode sites were measured and applied according to standard criteria, specifically following the International 10-20 System of Electrode Placement.

PSG sleep stages, including N1, N2, N3, and rapid eye movement (REM), as well as arousals, were manually assessed with the assistance of an experienced sleep technologist. The sleep summary and epoch-by-epoch (EBE) analysis employed the conventional 30-second PSG scoring epoch.

Participants utilized the Actiwatch Spectrum, a research-grade actigraphy device manufactured by Philips Respironics, Inc., in Murrysville, PA, USA. Participants were instructed to wear the actigraphy device on their non-dominant wrist throughout the entire night. We configured the devices to operate at a medium-threshold sensitivity of 40 activity counts per epoch and collected sleep data in 30-second epochs.

2.3. Procedures

Sleep data was collected from each participant over two nights in a sleep laboratory that had been converted from an apartment at Chongqing University. This data collection was conducted from August to September 2020. The laboratory replicates the sleeping environment typically found in a home. Only one participant is scheduled for the experiment each night, and all participants use the same equipment. Before the formal assessment, every participant underwent a period of acclimatization with the PSG device, lasting one to two nights. Participants arrived at the laboratory two hours before their usual bedtime. The researcher undertook the task of installing the PSG electrodes, configuring the Actiwatch Spectrum, and setting up the iSleep S200G and Sleep Dot B501 devices. The Actiwatch Spectrum activated automatically when worn, while the other devices

 doi:10.20944/preprints202602.1395.v1

activated simultaneously as participants began to fall asleep. Participants were instructed to follow their usual sleep routines, going to bed and waking up at their regular times, and to turn off their sleep devices immediately upon waking. The experimental study was conducted under strict supervision without any intervention. The outcome measures included TST, SOL, WASO, SE, light sleep (N1+N2), and deep sleep (N3).

2.4. Statistical Analyses

Sleep data collected from the iSleep S200G, Sleep Dot B501, and Actiwatch Spectrum devices were compared to PSG recordings. Statistically significant differences between each device and PSG were assessed using paired t-tests or Wilcoxon signed-rank tests, with corresponding p-values reported. The normality of the data was evaluated by conducting the Shapiro-Wilk test. For all continuous variables that adhered to a normal distribution, the mean and standard deviation (SD) were provided, and paired t-tests were used. For variables exhibiting a non-normal distribution, the median and interquartile range (IQR) were reported, and the Wilcoxon matched-pairs test was used. P-values that fell below the threshold of $p < 0.05$ were considered statistically significant. Hedges' g was reported as the effect size. Agreement and reliability between the PSG and each of the selected testing devices were evaluated using a two-way random effects model to calculate the intraclass correlation coefficient (ICC). ICC values indicate poor agreement for less than 0.50, moderate agreement for between 0.50 and 0.75, good agreement for between 0.75 and 0.90, and excellent agreement for greater than 0.90 [31].

A Bland-Altman plot was employed to visually represent the level of agreement between PSG and each of the chosen devices. The mean of two measurements was plotted on the x-axis, and the discrepancy between them was plotted along the y-axis. The calculation of the mean of the differences provides an estimation for the average deviation between PSG and other devices. The lower and upper limits of agreement (LOA) were determined by calculating the mean difference ±1.96 SD. LOA is a statistical method used to estimate the range within which a specific proportion of differences between measurements is expected to fall.

The present study employed EBE agreement statistics to assess the agreement between sleep and wake epochs recorded by each device in relation to the corresponding epochs scored by PSG. These statistics include sensitivity (true positive rate), specificity (true negative rate), positive predictive value (PPV), negative predictive value (NPV), and accuracy. Sensitivity refers to the proportion of sleep epochs correctly detected as sleep by the PSG device, while specificity represents the proportion of wake epochs correctly detected as wake by the PSG device. PPV indicates the proportion of sleep epochs scored by the device that were accurately identified as true PSG sleep, while NPV represents the proportion of wake epochs scored by the device that were accurately identified as true PSG wake. Accuracy, on the other hand, refers to the proportion of all PSG epochs correctly detected by the device [4].

All data were analyzed using IBM SPSS 20. GraphPad Prism 6 software was utilized to conduct Bland-Altman analyses to assess the level of agreement between two devices for each sleep parameter.

2.5. Missing Data Procedures

At each lab visit, the researchers carefully synchronized all devices and charged the Actiwatch Spectrum to collect sleep data from PSG, actigraphy, and CCSTDs during scheduled sleep periods. However, on one night, the polysomnography electrodes detached while the subject was asleep, which led to the exclusion of all monitoring device data for that specific night. Additionally, due to device malfunctions, data from five nights of iSleep S200G reports, two nights of Sleepace reports, and five nights of Actiwatch Spectrum reports were excluded from the analysis.

3. Results

 doi:10.20944/preprints202602.1395.v1

3.1. Sleep Summary Outcomes

Table 1 displays the summary measures of the iSleep S200G, Sleep Dot B501, and Actiwatch Spectrum compared to PSG. Figures 1–3 display the Bland-Altman plots for different devices corresponding to the data.

 doi:10.20944/preprints202602.1395.v1

Table 1. A comparative analysis of the outcomes between PSG and each of the selected devices.

Outcomes	n	PSG Mean±SD/ Med (IQR)	Device Mean±SD/ Med (IQR)	Bias[a]	LOA[b]	t/w(P)[c]	Effect size	ICC[d] [95%CI] *P*
TST								
iSleep[e]	16	438.66±65.90	420.94±60.93	-17.72	-96.20,60.76	1.77(.10)	0.28	0.78[0.48-0.92] <.001
Sleep Dot[f]	19	400.42±98.20	391.45±98.16	-8.97	-98.93,80.99	0.85(.41)	0.09	0.89[0.75-0.96] <.001
Actiwatch[g]	16	399.25±101.30	401.47±70.60	2.22	-115.54,119.98	-0.15(.88)	-0.03	0.77[0.46-0.92] <.001
SE								
iSleep	16	87.37% (75.65%,92.78%)	85.14%±4.26%	1.36%	-15.36%,17.96%	-.16(.88)	-0.17	0.48[-0.02-0.78] .03
Sleep Dot	19	84.40% (62.88%,89.90%)	82.48% (71.15%,89.23%)	1.92%	-18.80%,22.64%	-0.44(.66)	-0.14	0.71[0.40-0.88] <.001
Actiwatch	16	76.31%±15.24%	82.11%±8.66%	5.80%	-20.58%,32.18%	-1.72(.11)	-0.47	0.38[-0.07-0.72] .051
SOL								
Sleep Dot	19	30.20(20.10,66.70)	49.00(25.00,66.00)	5.46	-57.91,68.83	-0.78(.44)	-0.11	0.81[0.57-0.92] <.001
Actiwatch	16	37.75(19.80,109.65)	60.66±41.65	-4.71	-83.07,73.65	-0.36(.72)	0.09	0.72[0.36-0.89] .001
WASO								
Sleep Dot	19	32.00(18.10,76.90)	28.52(16.00,53.64)	-3.96	-74.15,66.23	0(*P*>.99)	0.10	0.64[0.27-0.84] .002
Actiwatch	16	33.70(24.54,70.13)	25.00(14.25,28.50)	-18.75	-78.92,41.42	-2.07(.04)	0.53	0.56[0.11-0.82] .004
Light Sleep								
iSleep	16	255.44±59.94	236.00±25.57	-19.44	-129.57,90.69	1.38(.19)	0.42	0.25[-0.23-0.64] .16
Sleep Dot	19	224.18±66.69	277.96±75.38	53.78	-71.35,178.91	-3.67(.002)	-0.76	0.47[-0.01-0.77] .003
Deep Sleep								
iSleep	16	41.09±29.28	187.50±50.75	146.41	44.65,248.17	-11.28(< .001)	-3.53	0.03[-0.04 - 0.18] 0.20
Sleep Dot	19	39.32±27.27	112.49±44.31	73.17	-15.83,162.17	-7.02(< .001)	-1.99	0.08[-0.09-0.35] 0.156

[a] Bias: the mean differences between the test device and PSG. [b] LOA: limits of agreement (MD±1.96 SD). [c] t/w(P): paired t-tests or Wilcoxon matched-pairs test. [d] ICC: intraclass correlation coefficient. [e] iSleep: iSleep S200G. [f] Sleep Dot: Sleep Dot B501. [g] Actiwatch: Actiwatch Spectrum.

 doi:10.20944/preprints202602.1395.v1

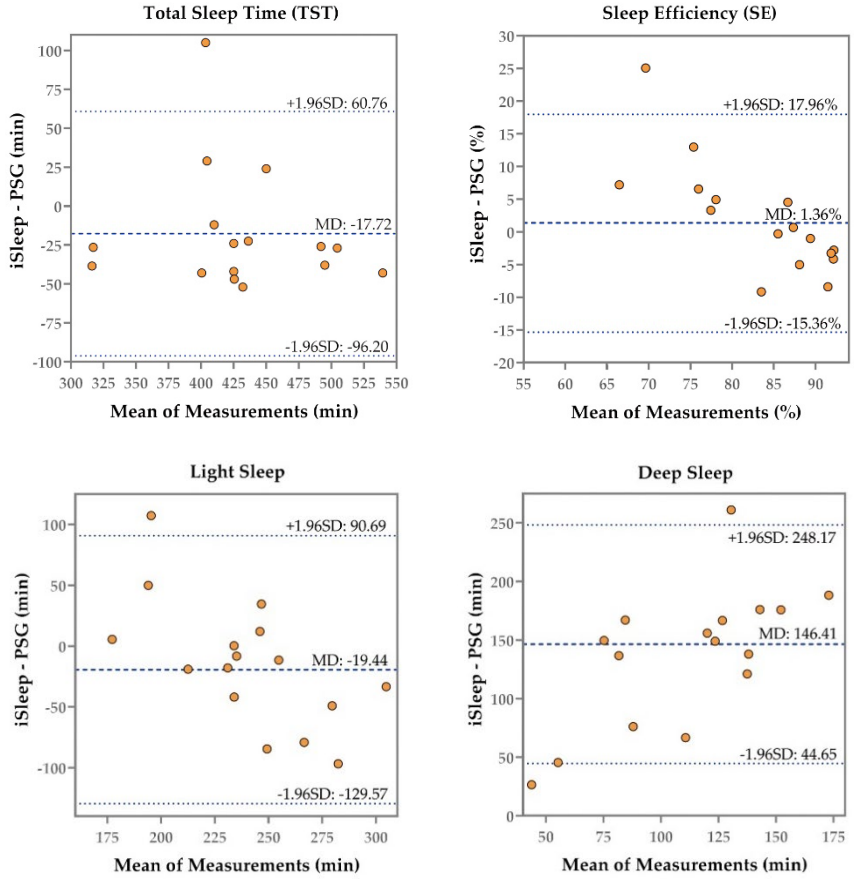
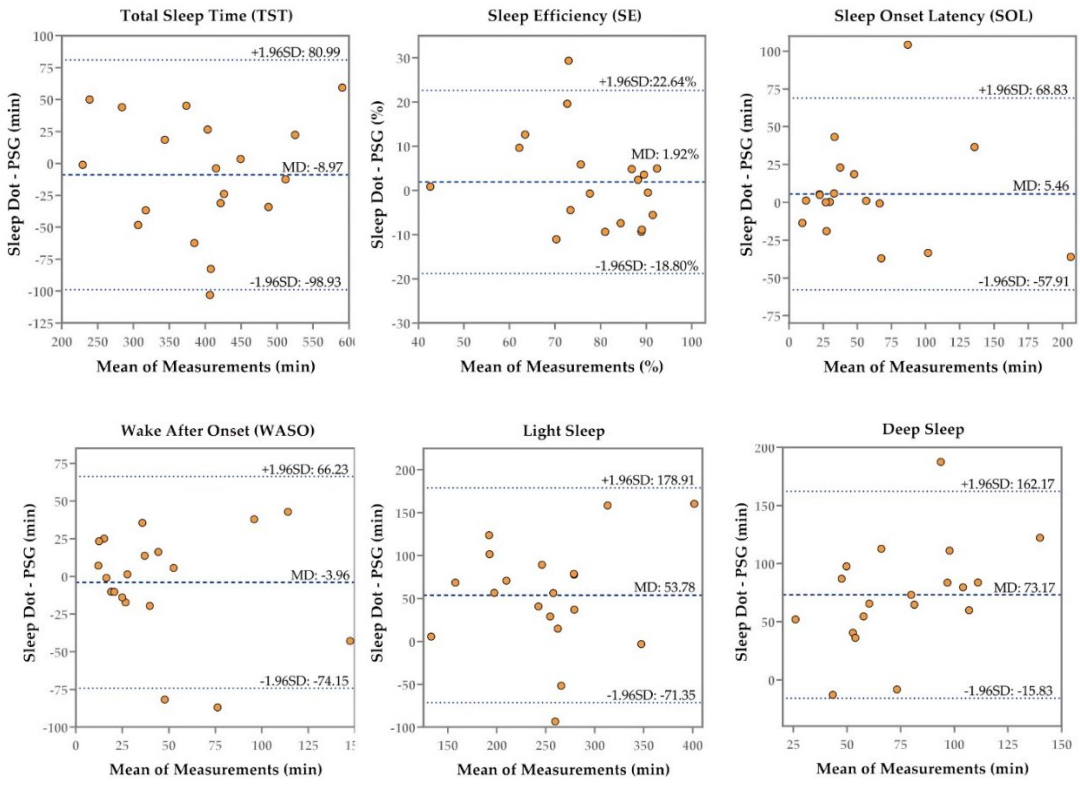


Figure 1. Bland-Altman plot illustrating the comparison of four outcomes (TST, SE, Light Sleep, and Deep Sleep) derived from the iSleep S200G and PSG recordings. The middle line represents the mean difference, while the upper and lower dashed lines indicate the upper and lower limits of agreement (mean difference ± 1.96 SD).



 doi:10.20944/preprints202602.1395.v1

Figure 2. Bland-Altman plot of the six outcomes (TST, SE, SOL, WASO, Light Sleep, and Deep Sleep) recorded by the Sleep Dot B501 and PSG. The middle line represents the mean difference, while the upper and lower dashed lines indicate the upper and lower limits of agreement (mean difference ± 1.96 SD).

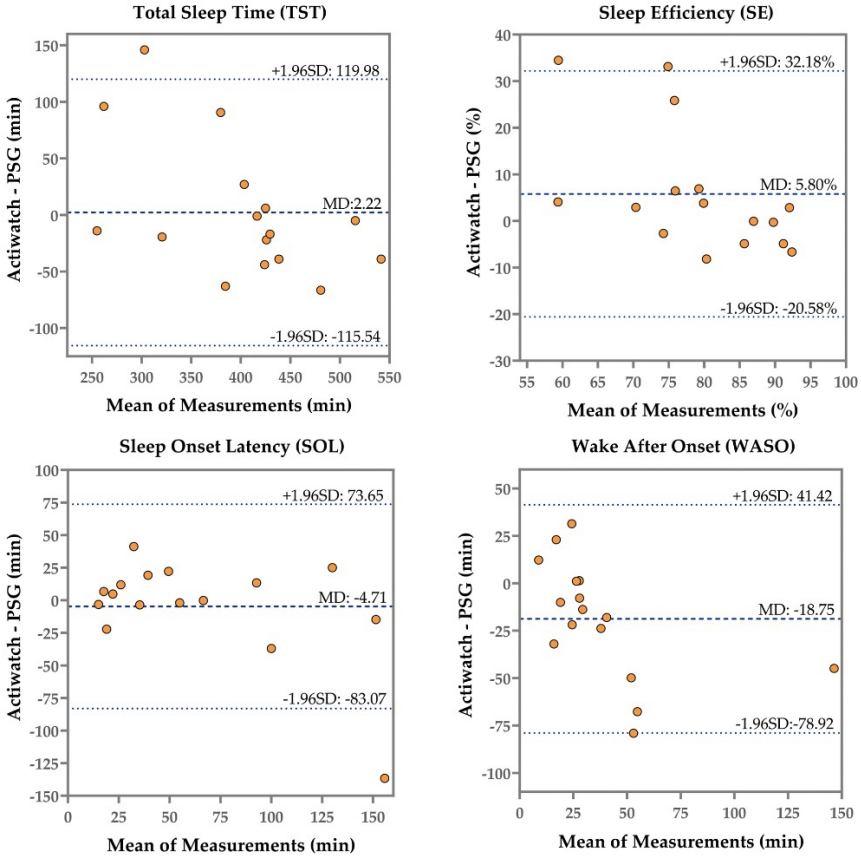


Figure 3. Bland-Altman plot of the four outcomes (TST, SE, SOL, WASO) recorded by the Actiwatch Spectrum and PSG. The middle line represents the mean difference, while the upper and lower dashed lines indicate the upper and lower limits of agreement (mean difference ± 1.96 SD).

Although the iSleep S200G, Sleep Dot B501 and Actiwatch Spectrum exhibited overestimation or underestimation in measuring TST, SE, SOL, and WASO, no significant differences were observed between these devices and PSG results for all metrics except WASO for the Actiwatch Spectrum ($P = 0.04$). The iSleep S200G and the Sleep Dot B501 devices were compared with PSG in the sleep stage. The results indicate that the iSleep S200G significantly overestimated deep sleep ($P < 0.001$), while the Sleep Dot B501 overestimated both light sleep ($P = 0.002$) and deep sleep ($P < 0.001$). However, the iSleep S200G underestimated light sleep by an average of 19.44 minutes, showing no significant difference compared to PSG. As shown in Figure 1, the Bland-Altman plot indicates that as the mean duration of light sleep measurements increases, the iSleep S200G shifts from overestimation to underestimation.

Compared with PSG, the iSleep S200G, Sleep Dot B501, and Actiwatch Spectrum showed good agreement in measuring TST, with ICC values of 0.78, 0.89, and 0.77, respectively. The Sleep Dot B501 showed the moderate agreement for SE (ICC=0.71), while the iSleep S200G (ICC=0.48) and Actiwatch Spectrum (ICC=0.38) exhibited poor agreement. The Actiwatch Spectrum showed lower agreement in both TST and SE compared to iSleep S200G and Sleep Dot B501. The Bland-Altman plots for SE from the iSleep S200G (Figure 1), Sleep Dot B501 (Figure 2), and Actiwatch Spectrum (Figure 3) indicate that when subjects exhibit higher mean SE values, the biases of each device relative to PSG decrease. Sleep Dot B501 demonstrated superior agreement to Actiwatch Spectrum in SOL and

 doi:10.20944/preprints202602.1395.v1

WASO. Figures 2 and 3 illustrate that the Bland-Altman plots for SOL and WASO between the Sleep Dot B501 and Actiwatch Spectrum show that when participants have lower SOL and WASO values, the biases from PSG exhibit the smallest magnitude and least variability. The iSleep S200G and Sleep Dot B501 both exhibit poor agreement in both light and deep sleep.

3.2. EBE Analysis Outcomes

The level of agreement between EBE in sleep and wake states, as compared to PSG for all nights, is displayed in Table 2.

Table 2. EBE agreement: sleep versus wake.

Device	Accuracy	Sensitivity	Specificity	PPV	NPV
Actiwatch Spectrum	0.83	0.92	0.58	0.88	0.57
iSleep S200G	0.85	0.91	0.65	0.92	0.49
Sleep Dot B501	0.88	0.94	0.68	0.91	0.74

The sensitivity of the iSleep S200G (0.91) and Sleep Dot B501 (0.94) compared to PSG was found to be excellent. The aforementioned observation suggests the significant proficiency of devices in accurately identifying PSG sleep epochs. Sleep Dot B501 exhibits higher sensitivity compared to Actiwatch Spectrum (0.92), whereas iSleep S200G demonstrates slightly lower sensitivity. However, the specificity of the two CCSTDs (iSleep S200G=0.65, Sleep Dot B501=0.68) was not as good as their sensitivity. Although both the iSleep S200G and Sleep Dot B501 showed reduced accuracy in detecting wake epochs compared to sleep epochs, these two devices achieved higher wake detection accuracy than the Actiwatch Spectrum (0.58).

Other indicators of agreement between EBE and PSG, as presented in Table 2, primarily align with the results of sensitivity and specificity. The PPV for both the iSleep S200G (0.92) and the Actiwatch Spectrum (0.91) consistently demonstrated high values. This finding implies that the sleep epochs recorded by these devices provide an accurate representation of sleep, as assessed through PSG. The NPV for both the iSleep S200G (0.49) and the Actiwatch Spectrum (0.57) devices was relatively lower, indicating that the device-scored wake epochs had a poorer ability to accurately reflect the PSG scoring compared to the device-scored sleep epochs. Sleep Dot B501 demonstrates a PPV of 0.91 while also demonstrating the highest NPV of 0.74. All devices demonstrate a similar level of accuracy.

4. Discussion

4.1. Principal Results

Overall, the two representative devices, iSleep S200G and Sleep Dot B501, selected for this experiment, despite their low cost and having been on the market for over eight years, still demonstrate respectable performance compared to PSG and actigraphy. Compared to PSG, the iSleep S200G showed statistically significant differences in deep sleep, the Sleep Dot B501 in both light and deep sleep stages, and the Actiwatch Spectrum in WASO. The EBE analysis results indicate that the iSleep S200G and Sleep Dot B501 demonstrated superior performance in sleep detection. They exhibited higher sensitivity in detecting sleep compared to wake, but relatively low specificity when compared to PSG. However, the accuracy of both devices was found to be lower during the sleep stages. The iSleep S200G demonstrated comparable levels of accuracy, sensitivity, and specificity to actigraphy. The Sleep Dot B501 exhibited superior accuracy and sensitivity compared to actigraphy, although it had a slightly lower level of specificity. The findings also demonstrated a high level of agreement for TST in iSleep S200G, as well as TST and SOL in Sleep Dot B501. The agreement was moderate for SE in both the iSleep S200G and the Sleep Dot B501, as well as for WASO in the Sleep Dot B501. The agreement levels of Actiwatch Spectrum were inferior to those of both iSleep S200G and Sleep Dot B501 regarding TST, SE, SOL, and WASO. The iSleep S200G and Sleep Dot B501

 doi:10.20944/preprints202602.1395.v1

demonstrated performance comparable to, and even superior to, the Actiwatch Spectrum in terms of EBE agreement in sleep and wake.

4.2. Comparison with Prior Work

In comparison to other CCSTDs with radar sensors, such as the ResMed S+ (accuracy: 0.88, sensitivity: 0.93, specificity: 0.51) [8], SleepScore Max (accuracy: 0.88, sensitivity: 0.94, specificity: 0.50) [8], Somnofy (accuracy: 0.76, sensitivity: 0.97, specificity: 0.72) [32], SleepMinder (accuracy: 0.77, sensitivity: 0.86, specificity: 0.52) [33], iSleep S200G consistently shows strong performance in terms of agreement with EBE. In comparison, when evaluating the Sleep Dot B501 against other consumer sleep devices that utilize accelerometers—such as the non-sleep staging Fitbit (which has an accuracy range of 0.81 to 0.91, a sensitivity range of 0.87 to 0.99, and a specificity range of 0.10 to 0.52), the sleep staging Fitbit model (accuracy: 0.90, sensitivity: 0.95 to 0.96, specificity: 0.58 to 0.69) [2], and the ActiGraph GT3X+ (accuracy: 0.89, sensitivity: 0.97, specificity: 0.24 to 0.25) [34] —the Sleep Dot B501 exhibits outstanding capabilities in sleep epoch identification.

The comparative analysis above demonstrates that both the iSleep S200G and Sleep Dot B501 exhibit high reliability in sleep monitoring. However, their suitability as experimental instruments across different sleep research scenarios can be evaluated by comparing them with other consumer-grade sleep monitoring devices already used in experiments.

Research Scenario 1: Long-term Sleep Monitoring. Among existing research, wristband-style sleep monitoring devices are the most widely used. We selected three high-performance devices for comparative analysis. POWER C J et al. [20] observed the sleep patterns of seven semi-professional female basketball players over a 13-week duration utilizing wrist-worn sleep monitors, Readiband™ (accuracy: 0.88, sensitivity: 0.94, specificity: 0.45) [8]. The sleep metrics included in the quantitative study were time in bed, TST, WASO, SE, sleep onset, and sleep offset (the time of day when last waking before getting up). Readiband was utilized in an additional sleep intervention trial focused on enhancing sleep, mood, and cognitive function in athletes. The research gathered sleep monitoring data from 56 e-sports participants across Korea, the United States, and Australia during a period of four weeks preceding and following the intervention. The evaluated sleep parameters comprised sleep onset time, TST, SOL, WASO, SE, time in bed, and wake-up time [35]. Compared with Readiband, Sleep Dot B501 exhibits superior performance across all sleep metrics necessary for the experiment and shows greater agreement in EBE. But the iSleep S200G has slightly lower accuracy and sensitivity, and it cannot offer data for SOL or WASO. Fritz H et al. [36] performed a study examining the influence of indoor air quality on sleep quality with the Fitbit Inspire HR. The Fitbit gadget exhibited an accuracy of 0.89, a sensitivity of 0.93, and a specificity of 0.45 [9]. The research gathered sleep data from 20 people over 77 days, encompassing assessments of TST, SE, REM, and non-rapid eye movement (NREM). Both the iSleep S200G and Sleep Dot B501 fulfill the experiment's requirements; however, the Sleep Dot B501 surpasses the Fitbit Inspire HR. A separate study investigating the effects of home isolation during the COVID-19 epidemic on sleep data employed the Garmin Fenix 6 Pro [37]. The research gathered sleep data, encompassing TST, light sleep, deep sleep, SOL, and REM, from 16 male professional fitness trainers over a duration of four consecutive months. The Garmin Fenix 6 Pro selected for this investigation is predicated on the validation outcomes of the Garmin Fenix 5S (accuracy: 0.87, sensitivity: 0.56, specificity: 0.92) [8]. Apparently, the Garmin Fenix 5S ostensibly excels in wakefulness monitoring but exhibits inferior sleep-tracking skills relative to the iSleep S200G and Sleep Dot B501. Neither the iSleep S200G nor the Sleep Dot B501 possesses the capability to track REM sleep.

Experimental Scenario 2: Sleep Monitoring in Large or Global Samples. Hirata T et al. [17] utilized the HSL-102M device (HSL-101: accuracy: 0.588, sensitivity: 0.963, specificity: 0.376) [13] to evaluate the population attributable fraction of home hypertension. A large-scale sleep monitoring study was conducted over 10 days, involving 1,474 participants. The iSleep S200G and Sleep Dot B501 are superior to the HSL-101 in terms of accuracy and specificity, and they only have slightly lower sensitivity. Another large-scale study examined the impact of smartphone usage on adolescents' sleep

 doi:10.20944/preprints202602.1395.v1

duration, involving 614 participants who provided TST data from 17 validated consumer-grade wearable sleep monitoring devices [38]. The Xiaomi Mi-Band (accuracy: 0.84, sensitivity: 0.99, specificity: 0.05) [39], the WHOOP Strap 3.0 (accuracy: 0.89, sensitivity: 0.95, specificity: 0.51) [40], and the Samsung Galaxy Watch (accuracy: 0.80-0.97) [41] are among the devices that demonstrate relatively satisfactory accuracy. The convenience of consumer sleep devices also simplifies the collection of data from large-scale sleep samples on a global level. SCOTT H et al. [16] used the Withings Sleep Analyzer, a device placed under the mattress, to examine the relationship between sleep duration, irregular sleep patterns, and hypertension in a substantial global sample encompassing over 2 million nights. The researchers gathered data on TST, sleep onset, sleep midpoint, and sleep offset from 15,526 participants aged 18 to 90. Research has proven that the equipment overestimates the TST of subjects by 25.8 minutes [19]. In contrast, the iSleep S200G and Sleep Dot B501 underestimated TST by 19.6 and 7.4 minutes, respectively. In another global study, researchers also employed the Withings Sleep Analyzer's consumer user database to evaluate the prevalence, variability, and diagnostic misclassification of obstructive sleep apnea across multiple nights. This database included nightly sleep data from 87,610 participants collected from July 1, 2020, to March 30, 2021 [42].

In conclusion, the iSleep S200G and Sleep Dot B501 exhibit superior performance in accuracy evaluations relative to other consumer sleep devices, while also being cost-effective and user-friendly. They can address sleep monitoring requirements in diverse experimental contexts, including long-term, large-scale, or global sample populations. As with other validated consumer sleep devices, the light and deep sleep data from the two validated consumer sleep devices cannot presently be utilized in experimental analyses due to technical limitations. The limitations lead to reduced accuracy in sleep stage detection. Simultaneously, the iSleep S200G cannot offer data on SOL or WASO. As a result, the device must be chosen based on the experimental requirements. CCSTDs are an optional and easy experimental instrument for small-scale and short-duration studies. The measurement bias of these devices requires critical evaluation. However, when performing large-scale, long-term sleep research on people who are unable or unwilling to wear sleep monitoring devices, relying on expensive PSG equipment and activity monitors for continuous monitoring is both impractical and challenging. In this situation, CCSTDs provide an innovative solution.

Currently, the rapid development of sleep-tracking technology has resulted in new devices and algorithms entering the market quicker than researchers can assess their performance [43,44]. We noted that the validation of some consumer sleep devices utilized in experimental investigations did not meet the established standards for validation. This issue encompasses the lack of critical analyses of the ICC and EBE agreements. Lack of validation studies necessitates relying on validation data from older generations when employing later-generation devices. Researchers assert that the new generation outperforms the previous generation.

The experimental results may be subject to unpredictable errors due to the lack of clarity regarding the measurement bias between the new and the ancient devices. Consequently, it is imperative to implement standardized validation experiments prior to the implementation of the new device.

4.3. Limitations

The study has limitations. First, the experiment's participants were limited to healthy young people. A more in-depth verification study is needed to ensure the accuracy of sleep monitoring in youngsters, older people who are more active during sleep, and individuals with sleep disorders. Second, while the sample size for this investigation is small, we rigorously vetted the subjects and sleep data to verify that they are representative and correctly reflect actual sleep patterns. Third, all of the individuals in this study slept alone, with no verification research conducted in settings including numerous beds in the same room or multiple people sleeping in the same bed.

5. Conclusions

 doi:10.20944/preprints202602.1395.v1

As representatives of the first generation of CCSTDs, the iSleep S200G and Sleep Dot B501 demonstrate excellent performance in monitoring sleep and wakefulness, achieving accuracy comparable to or even superior to actigraphy. Compared to the PSG, the iSleep S200G showed no significant differences in TST and SE, whereas the Sleep Dot B501 demonstrated no significant differences in TST, SE, SOL, or WASO. Both devices showed greater agreement than actigraphy across these metrics. However, like other validated consumer-grade sleep monitoring devices, the iSleep S200G and Sleep Dot B501 exhibit relatively poor accuracy in sleep stage detection. Based on the validation results of this study, even CCSTDs released eight years ago can still meet the needs for long-term, large-scale preliminary or exploratory sleep research. However, the choice of device should depend on the required sleep monitoring indicators. The Sleep Dot B501 in this study can monitor more sleep metrics than the iSleep S200G. Following the pattern of technological iteration, subsequent products generally outperform older versions. Therefore, their accuracy should not be worse than that of the two devices evaluated in this study. In addition to performance, cost is also an important consideration. Long-term preliminary sleep monitoring does not depend on data from a single day. Higher precision is not necessarily better, as it typically entails higher equipment costs. Therefore, a balance must be struck between performance and cost, and appropriate equipment should be selected based on the application's requirements to ensure the smooth progress of the study.

Author Contributions: Conceptualization, Y.Y.，S.H. and H.Z.; Validation, H.Z. and S.H.; methodology, H.Z. and S.H.; software, H.Z. and L.G.; formal analysis, H.Z., L.G.; investigation, H.Z., X.D., X.C., T.H. and S.H.; resources, H.Z. and S.H.; data curation, H.Z. and S.H.; writing—original draft preparation, H.Z., L.G.; writing—review and editing, H.Z., Y.Y., S.H., X.C., L.G., X.D. and T.H.; visualization, H.Z.; supervision, Y.Y.; project administration, H.Z.; funding acquisition, Y.Y.. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 51978097 and 51778081).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of South China Normal University (No. SCNU-AOE-2025-018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to thank Ning Yan from the Department of Neurology. **at** University-Town Hospital of Chongqing Medical University for his support during the experi-ment.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PSG	Polysomnography
CCSTDs	Contactless consumer sleep-tracking devices
TST	Total sleep time
TWT	Total wake time
SE	Sleep efficiency
SOL	Sleep onset latency
WASO	Wake after sleep onset
N1	Light sleep
N2	Moderate sleep
N3	Deep sleep
REM	Rapid eye movement

 doi:10.20944/preprints202602.1395.v1

EBE	Epoch-by-epoch
ICC	Intraclass correlation coefficient
LOA	Lower and upper limits of agreement
PPV	Positive predictive value
NPV	Negative predictive value

References

1. Zhai, H.; Yan, Y.; He, S.; Zhao, P.; Zhang, B. Evaluation of the Accuracy of Contactless Consumer Sleep-Tracking Devices Application in Human Experiment: A Systematic Review and Meta-Analysis. *Sensors (Basel, Switzerland)* **2023**, *23*, doi:10.3390/s23104842.
2. Haghayegh, S.; Khoshnevis, S.; Smolensky, M.H.; Diller, K.R.; Castriotta, R.J. Accuracy of Wristband Fitbit Models in Assessing Sleep: Systematic Review and Meta-Analysis. *Journal of medical Internet research* **2019**, *21*, e16273, doi:10.2196/16273.
3. Scott, H. Sleep and circadian wearable technologies: considerations toward device validation and application. *Sleep* **2020**, *43*, doi:10.1093/sleep/zsaa163.
4. Menghini, L.; Cellini, N.; Goldstone, A.; Baker, F.C.; de Zambotti, M. A standardized framework for testing the performance of sleep-tracking technology: step-by-step guidelines and open-source code. *Sleep* **2021**, *44*, doi:10.1093/sleep/zsaa170.
5. de Zambotti, M.; Menghini, L.; Grandner, M.A.; Redline, S.; Zhang, Y.; Wallace, M.L.; Buxton, O.M. Rigorous performance evaluation (previously, ?validation?) for informed use of new technologies for sleep health measurement. *Sleep health* **2022**, *8*, 263-269, doi:10.1016/j.sleh.2022.02.006.
6. Lee, X.K.; Chee, N.; Ong, J.L.; Teo, T.B.; van Rijn, E.; Lo, J.C.; Chee, M.W.L. Validation of a Consumer Sleep Wearable Device With Actigraphy and Polysomnography in Adolescents Across Sleep Opportunity Manipulations. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* **2019**, *15*, 1337-1346, doi:10.5664/jcsm.7932.
7. Kahawage, P.; Jumabhoy, R.; Hamill, K.; de Zambotti, M.; Drummond, S.P.A. Validity, potential clinical utility, and comparison of consumer and research-grade activity trackers in Insomnia Disorder I: In-lab validation against polysomnography. *Journal of sleep research* **2020**, *29*, e12931, doi:10.1111/jsr.12931.
8. Chinoy, E.D.; Cuellar, J.A.; Huwa, K.E.; Jameson, J.T.; Watson, C.H.; Bessman, S.C.; Hirsch, D.A.; Cooper, A.D.; Drummond, S.P.A.; Markwald, R.R. Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep* **2021**, *44*, doi:10.1093/sleep/zsaa291.
9. Chinoy, E.D.; Cuellar, J.A.; Jameson, J.T.; Markwald, R.R. Performance of Four Commercial Wearable Sleep-Tracking Devices Tested Under Unrestricted Conditions at Home in Healthy Young Adults. *Nat Sci Sleep* **2022**, *14*, 493-516, doi:10.2147/NSS.S348795.
10. Grandner, M.A.; Bromberg, Z.; Hadley, A.; Morrell, Z.; Graf, A.; Hutchison, S.; Freckleton, D. Performance of a multisensor smart ring to evaluate sleep: in-lab and home-based evaluation of generalized and personalized algorithms. *Sleep* **2023**, *46*, doi:10.1093/sleep/zsac152.
11. Zhang, Z.; Henzmann, S.; Hugli, G.; Qi, M.; Chen, W.; Lu, C.; Khatami, R. Validation of wearable sleep monitoring device based on cardiopulmonary coupling and accelerometer with comparison to polysomnography in adults. In Proceedings of the European journal of neurology, 2018; pp. 385-385.
12. Edouard, P.; Campo, D.; Bartet, P.; Yang, R.Y.; Bruyneel, M.; Roisman, G.; Escourrou, P. Validation of the Withings Sleep Analyzer, an under-the-mattress device for the detection of moderate-severe sleep apnea syndrome. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* **2021**, *17*, 1217-1227, doi:10.5664/jcsm.9168.
13. O'Hare, E.; Flanagan, D.; Penzel, T.; Garcia, C.; Frohberg, D.; Heneghan, C. A comparison of radio-frequency biomotion sensors and actigraphy versus polysomnography for the assessment of sleep in normal subjects. *Sleep and Breathing* **2014**, 1-8, doi:10.1007/s11325-014-0967-z.
14. Hou, D.; Lin, C.; Lin, Y. Diurnal Circadian Lighting Accumulation Model: A Predictor of the Human Circadian Phase Shift Phenotype. *Phenomics* **2022**, *2*, 50-63, doi:10.1007/s43657-021-00039-6.
15. Aries, M.B.C.; Fischl, G.; Lowden, A.; Beute, F. The relationship of light exposure to sleep outcomes among office workers. Part 1: Working in the office versus at home before and during the COVID-pandemic. *Lighting Research & Technology* **2022**, doi:10.1177/14771535221136096.

16. Scott, H.; Lechat, B.; Guyett, A.; Reynolds, A.C.; Lovato, N.; Naik, G.; Appleton, S.; Adams, R.; Escourrou, P.; Catcheside, P.; et al. Sleep Irregularity Is Associated With Hypertension: Findings From Over 2 Million Nights With a Large Global Population Sample. *Hypertension (Dallas, Tex. : 1979)* **2023**, *80*, 1117-1126, doi:10.1161/Hypertensionaha.122.20513.
17. Hirata, T.; Kogure, M.; Tsuchiya, N.; Miyagawa, K.; Narita, A.; Nochioka, K.; Uruno, A.; Obara, T.; Nakamura, T.; Nakaya, N.; et al. Impacts of the urinary sodium-to-potassium ratio, sleep efficiency, and conventional risk factors on home hypertension in a general Japanese population. *Hypertension research : official journal of the Japanese Society of Hypertension* **2021**, *44*, 858-865, doi:10.1038/s41440-021-00628-y.
18. Driller, M.W.; Dunican, I.C.; Omond, S.E.T.; Boukhris, O.; Stevenson, S.; Lambing, K.; Bender, A.M. Pyjamas, Polysomnography and Professional Athletes: The Role of Sleep Tracking Technology in Sport. *Sports* **2023**, *11*, doi:10.3390/sports11010014.
19. Edinborough, L.; Hill, J.; Jarvis, M.; Bruce-Low, S.; Pedlar, C.R. A bespoke sleep monitoring and sleep hygiene intervention improves sleep in an U18 professional football player: A case study. *Journal of sports sciences* **2023**, *41*, 350-357, doi:10.1080/02640414.2023.2213032.
20. Power, C.J.; Fox, J.L.; Teramoto, M.; Scanlan, A.T. Sleep Patterns Fluctuate Following Training and Games across the Season in a Semi-Professional, Female Basketball Team. *Brain Sci* **2023**, *13*, doi:10.3390/brainsci13020238.
21. Stone, J.D.; Rentz, L.E.; Forsey, J.; Ramadan, J.; Markwald, R.R.; Finomore, V.S., Jr.; Galster, S.M.; Rezai, A.; Hagen, J.A. Evaluations of Commercial Sleep Technologies for Objective Monitoring During Routine Sleeping Conditions. *Nature and Science of Sleep* **2020**, *12*, 821-842, doi:10.2147/nss.S270705.
22. Heglum, H.S.A.; Kallestad, H.; Vethe, D.; Langsrud, K.; Sand, T.; Engstrøm, M. Distinguishing sleep from wake with a radar sensor: a contact-free real-time sleep monitor. *Sleep* **2021**, *44*, doi:10.1093/sleep/zsab060.
23. Hendriks, M.M.S.; van Lotringen, J.H.; Vos-van der Hulst, M.; Keijsers, N.L.W. Bed Sensor Technology for Objective Sleep Monitoring Within the Clinical Rehabilitation Setting: Observational Feasibility Study. *JMIR mHealth and uHealth* **2021**, *9*, e24339, doi:10.2196/24339.
24. 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. *Journal of intensive care* **2020**, *8*, doi:10.1186/s40560-020-0433-x.
25. Zhang; Zenglian; Liu; Xin. Effect of the sleep nursing intervention on sleep status and gross motor development of infants. *Chin J Behav Med & Brain Sci* **2019**, *28*, 881-886.
26. Depner, C.M.; Cheng, P.C.; Devine, J.K.; Khosla, S.; de Zambotti, M.; Robillard, R.; Vakulin, A.; Drummond, S.P.A. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep* **2020**, *43*, doi:10.1093/sleep/zsz254.
27. Liang, S. The contactless sleep monitor iSleep achieved its crowdfunding goal in just one hour. Available online: https://tech.huanqiu.com/article/9CaKrnJFMWd (accessed on 2025-10-22).
28. Isleep Non Contact Sleep Monitor. Available online: https://www.made-in-china.com/showroom/manshu0608/product-detailWvwxiducrPkD/China-Isleep-Non-Contact-Sleep-Monitor.html (accessed on 2025-10-20).
29. Sleepace. Available online: http://webapi.test.sleepace.net/cn/mall.html (accessed on 2025-10-20).
30. DOU Yuan zhu, X.X.j.; BO Hong rui, X.L.s. Application of Radio Frequency Micro-Power Radar in Intelligent Sleep Management. *China Medical Devices* **2016**, *31*, 16-20.
31. Tkk, A.; Mylb, B.J.J.o.C.M. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. **2016**.
32. Toften, S.; Pallesen, S.; Hrozanova, M.; Moen, F.; Grønli, J. Validation of sleep stage classification using non-contact radar technology and machine learning (Somnofy®). *Sleep medicine* **2020**, *75*, 54-61, doi:10.1016/j.sleep.2020.02.022.
33. Pallin, M.; O'Hare, E.; Zaffaroni, A.; Boyle, P.; Fagan, C.; Kent, B.; Heneghan, C.; de Chazal, P.; McNicholas, W.T. Comparison of a novel non-contact biomotion sensor with wrist actigraphy in estimating sleep quality in patients with obstructive sleep apnoea. *Journal of sleep research* **2014**, *23*, 475-484, doi:10.1111/jsr.12126.

 doi:10.20944/preprints202602.1395.v1

34. Chase, J.D.; Busa, M.A.; Staudenmayer, J.W.; Sirard, J.R. Sleep Measurement Using Wrist-Worn Accelerometer Data Compared with Polysomnography. *Sensors (Basel, Switzerland)* **2022**, *22*, doi:10.3390/s22135041.
35. Bonnar, D.; Lee, S.; Roane, B.M.; Blum, D.J.; Kahn, M.; Jang, E.; Dunican, I.C.; Gradisar, M.; Suh, S. Evaluation of a Brief Sleep Intervention Designed to Improve the Sleep, Mood, and Cognitive Performance of Esports Athletes. *International journal of environmental research and public health* **2022**, *19*, doi:10.3390/ijerph19074146.
36. Fritz, H.; Kinney, K.A.; Wu, C.Y.; Schnyer, D.M.; Nagy, Z. Data fusion of mobile and environmental sensing devices to understand the effect of the indoor environment on measured and self-reported sleep quality. *Building and Environment* **2022**, *214*, doi:10.1016/j.buildenv.2022.108835.
37. Dergaa, I.; Ammar, A.; Souissi, A.; Fessi, M.S.; Trabelsi, K.; Glenn, J.M.; Ghram, A.; Taheri, M.; Irandoust, K.; Chtourou, H.; et al. COVID-19 lockdown: Impairments of objective measurements of selected physical activity, cardiorespiratory and sleep parameters in trained fitness coaches. *EXCLI J* **2022**, *21*, 1084-1098, doi:10.17179/excli2022-4986.
38. Mac Carthaigh, S.; Perry, J.; Griffin, C. Do objective data support the claim that problematic smartphone use has a clinically meaningful impact upon adolescent sleep duration? *Behav Inform Technol* **2022**, doi:10.1080/0144929x.2022.2136533.
39. Ameen, M.S.; Cheung, L.M.; Hauser, T.; Hahn, M.A.; Schabus, M. About the Accuracy and Problems of Consumer Devices in the Assessment of Sleep. *Sensors* **2019**, *19*, doi:10.3390/s19194160.
40. Miller, D.J.; Lastella, M.; Scanlan, A.T.; Bellenger, C.; Halson, S.L.; Roach, G.D.; Sargent, C. A validation study of the WHOOP strap against polysomnography to assess sleep. *Journal of sports sciences* **2020**, *38*, 2631-2636, doi:10.1080/02640414.2020.1797448.
41. El-Amrawy, F.; Nounou, M.I. Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial? *Healthc Inform Res* **2015**, *21*, 315-320, doi:10.4258/hir.2015.21.4.315.
42. Lechat, B.; Naik, G.; Reynolds, A.; Aishah, A.; Scott, H.; Loffler, K.A.; Vakulin, A.; Escourrou, P.; McEvoy, R.D.; Adams, R.J.; et al. Multinight Prevalence, Variability, and Diagnostic Misclassification of Obstructive Sleep Apnea. *American journal of respiratory and critical care medicine* **2022**, *205*, 563-569, doi:10.1164/rccm.202107-1761OC.
43. Goldstein, C. Current and Future Roles of Consumer Sleep Technologies in Sleep Medicine. *Sleep medicine clinics* **2020**, *15*, 391-408, doi:10.1016/j.jsmc.2020.05.001.
44. Baumert, M.; Cowie, M.R.; Redline, S.; Mehra, R.; Arzt, M.; Pepin, J.L.; Linz, D. Sleep characterization with smart wearable devices: a call for standardization and consensus recommendations. *Sleep* **2022**, *45*, doi:10.1093/sleep/zsac183.