https://doi.org/10.32895/MPR.25.00033

MedPharmRes (MPR) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	**Validation of a consumer-grade wearable activity tracker for sleep measurement in Vietnamese patients**
Running Title (within 10 words)	Validation of wearable activity tracker
Author	Toan Thanh Le[1], Son Trung Huynh[1], Xuan-Hieu Thi Nguyen[1], Vinh Nhu Nguyen[1], An Le Pham[1]
Affiliation	[1] Faculty of Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam
ORCID (for more information, please visit https://orcid.org)	Toan Thanh Le (https://orcid.org/0009-0003-5186-6103) Son Trung Huynh (https://orcid.org/0000-0003-3049-0541) Xuan-Hieu Thi Nguyen (https://orcid.org/0009-0005-9116-5066) Vinh Nhu Nguyen (https://orcid.org/0000-0002-8358-902X) An Le Pham (https://orcid.org/0000-0003-1186-0543)
Competing interests	The authors have no conflicts of interest to declare.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	None
Acknowledgements	We sincerely thank all the patients who participated in this study. We also extend our gratitude to Dr. Myunggu Jung and Dr. In-hong Choi from YONSEI University, for their invaluable assistance in refining the English language of this manuscript, as part of the collaborative MEF project with UMP. This article references Fitbit® products for non-commercial research purposes, in accordance with the fair use principles of Vietnam's Law on Science and Technology (No. 29/2013/QH13) and Law on Intellectual Property (No. 50/2005/QH11). This usage is undertaken in good faith, respects academic integrity, and is not intended to infringe upon any intellectual property rights. The mention of these products does not imply endorsement by the rights holders. Any concerns regarding potential rights violations should be directed to the corresponding author or MedPharmRes for resolution.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form. Authors can't change and add items, but you can delete items that are not applicable.	Conceptualization: TTL, STH, ALP. Data curation: STH, XHTN, VNN. Formal analysis: TTL, STH, XHTN, ALP. Methodology: TTL, XHTN, ALP. Software: STH, XHTN, VNN. Validation: TTL, VNN, ALP. Investigation: XHTN, VNN, STH.

	Writing - original draft: STH, TTL. Writing - review & editing: TTL, STH, XHTN, VNN, ALP
Ethics approval and consent to participate	All personal information was anonymized and kept confidential. Ethical approval was granted by the Institutional Review Board of the University of Medicine and Pharmacy at Ho Chi Minh City (approval number 998/HĐĐĐ - ĐHYD, dated October 20, 2023). Written informed consent was obtained from all participants before the commencement of the study in accordance with the Declaration of Helsinki.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Son Trung Huynh
Email address – this is where your proofs will be sent	huynhtrungson@ump.edu.vn
Secondary Email address	strhuynh@gmail.com
Address	University of Medicine and Pharmacy at Ho Chi Minh City, 217 Hong Bang street, Ward 11, District 5, Ho Chi Minh City, Vietnam
Cell phone number	+84 908 056 167
Office phone number	(+84-28) 3855 5780
Fax number	(+84-28) 3855 2304

Abstract

Introduction: The increasing use of wearable activity trackers (WAT) for home-based sleep assessment has raised necessity to clarify their accuracy, particularly in resource-limited settings. This study aimed to validate WAT by polysomnography (PSG) for measuring key sleep parameters—including total sleep time (TST), sleep efficiency (SE), wake after sleep onset (WASO), sleep onset latency (SOL), and sleep stage distribution—in a Vietnamese clinical population, with implications for primary care applications.

Methods: This cross-sectional study was conducted at the University Medical Center-Ho Chi Minh City, Vietnam, from December 2023 to July 2024. Sleep data were collected from 47 patients undergoing overnight PSG while simultaneously wearing a WAT. Sensitivity, specificity, and accuracy in detecting sleep versus wakefulness were assessed using epoch-by-epoch comparisons. Bland-Altman analysis was used to evaluate the agreement between WAT and PSG measurements, with mean differences and limits of agreement calculated for each sleep parameter.

Results: The WAT demonstrated high sensitivity (93%) but low specificity (44%) and an accuracy of 79% in identifying sleep versus wakefulness when compared to PSG. No significant differences were found between the two devices in measuring total sleep time (TST), sleep efficiency (SE), sleep onset latency (SOL), and sleep stages. However, the WAT significantly underestimated wake-after-sleep onset (WASO) compared to PSG ($p=0.011$).

Conclusion: The results are promising, but further confirmation in larger studies is required to confirm the utility of WAT in primary care settings in Vietnam.

Keywords: Wearable activity trackers, sleep measurement, diagnostic accuracy, primary care, Vietnam

1. Introduction

Sleep is a fundamental physiological process, vital for maintaining optimal health and well-being. Disruptions in sleep, particularly insomnia, can lead to a wide range of adverse effects, including impaired cognitive function, excessive daytime sleepiness, and declines in both physical and mental health, which ultimately diminish an individual's quality of life (1). Epidemiological studies have shown that insomnia affects a significant portion of the adult population, with prevalence rates ranging from 30% to 55% (2, 3). In the United States, the 2020

National Health Interview Survey reported that 14.5% of adults had difficulty initiating sleep, and 17.8% struggled to maintain sleep quality over a 30-day period.(4)

Accurate assessment of sleep is essential for effective diagnosis and management of insomnia. Polysomnography (PSG) is considered the gold standard for sleep evaluation, as it provides detailed physiological data on sleep patterns and abnormalities (5). However, PSG is costly, requires overnight monitoring in specialized centers, and may cause patient discomfort due to the unfamiliar environment and multiple sensors. Recent advancements have led to the development of wearable activity trackers (WAT), such as smartwatches and fitness trackers, which offer a more convenient and accessible alternative for home-based sleep monitoring. These devices leverage advanced technology to track and record sleep metrics, providing benefits such as affordability, ease of use, convenience, and the ability for individuals to monitor their sleep in familiar home environments (6).

In Vietnam, primary care physicians frequently encounter patients with sleep complaints, highlighting insomnia as a growing public health concern. However, limited research on the clinical utility of WAT hinders effective diagnostic solution in Vietnamese populations, particularly among individuals with suspected sleep disorders. This study aimed to validate WAT by comparing its sleep measurement capabilities with PSG in patients referred for overnight sleep assessment at the University Medical Center-Ho Chi Minh City (UMC HCMC), Vietnam. The findings can offer valuable insights into the feasibility and potential benefits of incorporating WAT into primary care practice.

2. Methods

2.1. Study design & participants:

2.1.1. Study design

A cross-sectional study was conducted at the Sleep Disorders Center, University Medical Center-Ho Chi Minh City (UMC HCMC). Polysomnography was conducted under standard conditions using the SOMNOlab 2 system, and participants wore consumer-grade WAT overnight. The WAT data were collected the following morning, and PSG data served as the reference standard.

2.1.2. Inclusion and exclusion criteria

The eligible participants were adults (≥18 years) referred for overnight sleep measurement by PSG and provided written informed consent.

Exclusion criteria encompassed:

- Known diagnosis of other primary sleep disorders (such as obstructive sleep apnea, restless legs syndrome, or circadian rhythm disorders);
- Severe psychiatric conditions (including major depression, anxiety disorders, schizophrenia, or bipolar disorder);
- Uncontrolled medical illnesses (such as heart failure or renal failure);
- Pregnancy or breastfeeding;
- Current use of sedatives, antidepressants, antipsychotics, antihistamines, or any substances known to affect sleep architecture.

2.1.3. Recruitment strategy

Participants were recruited using the convenience sampling method from the Sleep Disorders Center and then screened for eligibility based on the inclusion and exclusion criteria. Due to resource constraints at our center, only one PSG observation could be scheduled per night, which limited the total number of enrolled participants during the study period.

2.2. Sleep measurements and procedures

2.2.1. Polysomnography (PSG) as reference procedure

PSG was performed using the SOMNOlab 2 system, with sleep stages scored according to the American Academy of Sleep Medicine criteria (2017 updated version) (7). Collected metrics included total sleep time (TST), sleep efficiency (SE), sleep onset latency (SOL), wake after sleep onset (WASO), light sleep, deep sleep, and rapid eye movement (REM) sleep. Prior to each recording session, the SOMNOlab 2 polysomnography system was calibrated according to the manufacturer's instructions. This included impedance checks for all electrodes and verification of signal quality across all channels (EEG, EOG, EMG, ECG, airflow, and oximetry) before lights-off.

2.2.2. Fitbit Charge 5 as the validation instrument for WAT

Fitbit Charge 5 was chosen as the primary research instrument to validate the sleep measurement of WATs for their affordability and ease-of-use functionality in the context of limited resources. The device was worn from bedtime until the following morning in the laboratory setting. Each device was reset, updated to the latest firmware, and fully charged prior to use. WAT was placed on the participant's non-dominant wrist with skin contact visually confirmed by staff. Device placement was re-checked before lights-off. After each session, the WAT was immediately synchronized with the corresponding application for data transfer and accurate alignment with PSG records. Sleep parameters recorded by the WAT (including TST, SE, SOL, WASO, and sleep stage distributions) were matched to corresponding PSG data, based on the definitions provided on the mobile application and corresponding website.

2.3 Sample size calculation:

The sample size was estimated using the formula for sensitivity estimation provided by Buderer (8):

$$\text{n} = \frac{1.96 \text{ x sens x } (1 - \text{sens})}{d^2 \ x \ P}$$

Where:

- sens = Estimated sensitivity (0.95, based on previous study (9))

- d = Margin of error (0.2)

- P = Prevalence of insomnia in the target population (15%, based on epidemiological data from China (10))

Therefore, the minimum required sample size was calculated to be 34 participants.

While we acknowledged that a smaller margin of error (e.g., d = 0.1) would be more ideal for a validation study, the available population for PSG at our institution was limited. Notably, at the UMC HCMC, only one PSG test could be scheduled per night due to resource constraints and high clinical demand. Therefore, the number of participants enrolled was determined by the maximum feasible number of eligible patients during the study period, resulting in a final sample size of 47 observations after data cleaning.

2.4 Data collection methods

Data were collected using a structured form comprising three sections. Section 1 documented general information, including birth year, gender, height, weight, and BMI. Section 2 recorded metrics PSG metrics (lights-off/on times, TST, SOL, WASO, light sleep, deep sleep, REM sleep, and SE). Section 3 captured comparable metrics from WAT (TST, SOL, WASO, light sleep, deep sleep, REM sleep, and SE).

Demographic variables, including age, gender, height, weight, and BMI, were collected for all participants for descriptive analysis. No subgroup analyses or statistical adjustments for these variables were performed, as the study was primarily designed to validate the agreement between devices in the overall cohort.

Eligible participants referred for overnight PSG at the UMC HCMC were scheduled between 8:00 PM and 9:00 PM. Prior to the study, both PSG and WAT devices were calibrated. The WAT was set to "night-time mode" to minimize disturbances during sleep. Participants were instructed on the appropriate use of WAT and on PSG procedures. Data were retrieved the following morning as mentioned above.

2.5 Data analysis

PSG data categorized sleep stages into wake, Stage 1 (S1), Stage 2 (S2), Stage 3 (S3), Stage 4 (S4), and REM sleep. For comparison with results from WAT, PSG stages were grouped as follows: light sleep (S1 + S2), deep sleep (S3 + S4), and REM sleep (REM). Wake epochs were labeled as "0" and sleep epochs as "1," with corresponding time intervals recorded for analysis. For epoch-by-epoch comparison, both PSG and WAT data were divided into 30-second epochs. Epoch alignment was based on the lights-off and lights-on times recorded by the PSG system, ensuring that only epochs within the same time interval were included for analysis. Time synchronization was confirmed by matching the start and end times on both devices. Any epochs that were missing, incomplete, or identified as artifacts on either device were excluded from the analysis.

Sleep stage scoring for PSG data was performed by a single trained technician in accordance with the American Academy of Sleep Medicine (AASM) criteria, 2017 update, as mentioned above. Inter-rater reliability was not assessed. Sleep-wake classification by WAT was determined by the device's proprietary algorithm, without independent validation.

Descriptive statistics were used to summarize sleep parameters obtained from both devices.

Epoch-by-epoch comparisons (30-second epochs) were conducted to assess the WAT's ability to detect sleep and wake states relative to PSG. Sensitivity was defined as the proportion of correctly identified sleep epochs, specificity as the proportion of correctly identified wake epochs, and accuracy as the overall agreement between methods.

Agreement between PSG and WAT measurements for TST, SOL, WASO, light sleep, deep sleep, REM sleep, and SE were further analyzed using Bland-Altman plots. Mean differences and the 95% limits of agreement (LOA) were calculated to evaluate the magnitude and direction of potential biases. A positive mean difference indicated that the WAT underestimated a specific sleep variable relative to PSG, whereas a negative value reflected overestimation.

The normality of paired differences for each sleep parameter was assessed using the Shapiro–Wilk test, with visual inspection via Q–Q plots and histograms. Statistical analysis was performed using R (version 4.0.1), with significance set at a p-value < 0.05.

2.6 Ethics

Ethical approval was granted by the Institutional Review Board of the University of Medicine and Pharmacy at Ho Chi Minh City (approval number 998/HĐĐĐ - ĐHYD, dated October 20, 2023). Written informed consent was obtained from all participants before the commencement of the study in accordance with the Declaration of Helsinki.

3. Results

3.1 General characteristics of the participants

The study enrolled a total of 50 patients who met the inclusion criteria. After data cleaning and the exclusion of incomplete records, 47 participants (94% of the enrolled cohort) were eligible for analysis. The mean (SD) age was 48.42 (12.45) years (range, 28–73), and 66% were male. According to the Asia-Pacific classification of body mass index, 61.7% of participants were categorized as obese.

The participants had a mean total sleep time of approximately 336 minutes, corresponding to an average sleep efficiency of approximately 76% (Table 1). Most of their rest was spent in light sleep, followed by REM sleep, with deep sleep representing the smallest proportion. Additional details regarding median values and ranges for these parameters are provided in Table 1.

[Table 1]

3.2 Comparison between wearable activity trackers and polysomnography

As shown in Table 2, the WAT demonstrated a high sensitivity of 0.93 (SD 0.06; 95% CI, 0.91–0.95), indicating a high probability of correctly identifying sleep epochs. In contrast, the specificity for detecting wakefulness was lower at 0.44 (SD 0.19; 95% CI, 0.38–0.49), suggesting a tendency to misclassify wake epochs as sleep. The device's overall accuracy, representing the proportion of correctly classified epochs (sleep or wake), was 0.79 (SD 0.09; 95% CI, 0.76–0.81).

[Table 2]

3.3 Bland-Altman mean difference analysis of sleep parameters:

The Shapiro–Wilk test indicated that the paired differences for TST, SE, SoL, WASO, and LS were not normally distributed (all p < 0.05), while DS and REM differences did not show significant deviation from normality ($p > 0.05$). Figure 1 illustrates the Bland–Altman plots comparing sleep parameters measured by the WAT with those obtained from polysomnography (PSG). The WAT tended to overestimate TST by a mean (SD) of 30.07 minutes ($P = 0.835$) and SE by 6.15% ($P = 0.078$), though neither difference was statistically significant. Sleep onset latency (SoL) was slightly underestimated by 3.36 minutes ($P = 0.691$), which was also not significant. In contrast, WAT significantly underestimated wake after sleep onset (WASO) by a mean of 23.11 minutes ($P = 0.011$). No significant differences were detected in the measurement of light sleep ($P = 0.63$), deep sleep ($P = 0.475$), or REM sleep ($P = 0.995$) between WAT and PSG (Table 3).

[Figure 1]

[Table 3]

4. Discussion

4.1. Validation of wearable activity trackers in sleep measurement

The findings contribute to the growing body of work validating WAT, particularly among patients with sleep disorders. The unique context of our research lies in the focal point on a Vietnamese clinical population and the exploration of the potential application of a consumer-grade WAT - the Fitbit Charge 5 - in resource-limited primary care settings in Vietnam. We evaluated the WAT against the gold standard of polysomnography (PSG) in patients referred for overnight sleep assessment. We found that the WAT demonstrated a high sensitivity of 93% (95% CI: 91-95%) for detecting sleep, but its specificity was relatively low at 44% (95% CI: 38–49%), resulting in an overall accuracy of 79% (95% CI: 76–81%). In addition, there were no statistically significant differences between WAT and PSG for total sleep time (TST), sleep efficiency (SE), sleep onset latency (SoL), or sleep stage classification, suggesting that the device can reliably measure these parameters; however, the WAT significantly underestimated wake after sleep onset (WASO) by an average of 23.11 minutes ($p = 0.011$).

The high sensitivity indicates that the WAT is effective at identifying sleep epochs, which can aid in ruling out cases of clinically significant insomnia. Our sensitivity findings are comparable to those reported in previous studies on similar devices (11-14). In contrast, the relatively low specificity suggested that the device is less capable of accurately distinguishing wakefulness, which may lead to an overestimation of sleep duration. This lowered specificity, compared to studies on the Fitbit Charge 4 (12) could partly be due to the characteristics of the study population, including a higher obesity rate, which may impact the performance of wrist-based heart rate sensors (11, 13). Overall, the sleep tracker demonstrated moderate accuracy of 79% (95% CI: 76-81%) in identifying sleep and wake stages. The accuracy in this study (79%) is not significantly different from the accuracy reported for the Fitbit Charge 4 (86.5%) (12). Overall, our results are consistent with the systematic review by Haghayegh et al. (15), which reported sensitivity values ranging from 0.87 to 0.99 and specificity values from 0.10 to 0.52.

4.2. Comparison of sleep parameters

Our analysis revealed no significant differences between WAT and PSG regarding the measurement of total sleep time (TST), sleep efficiency (SE), sleep onset latency (SoL), and overall sleep stage distribution, suggesting that the device might reliably measure these parameters. Although the WAT tended to overestimate TST by approximately 30 minutes, this difference was not statistically significant ($p = 0.835$). Similarly, there was no significant difference in SE between the two methods ($p = 0.691$). These results suggest that the WAT can measure sleep duration with reliable accuracy, including total sleep time and the proportion of time spent sleeping relative to time in bed. Our findings are consistent with those reported by Dong et al. (12), who found no significant differences in TST using the Fitbit Charge 3. In contrast, previous studies reporting overestimations of TST and SE by other WAT devices relative to PSG (11, 14) were not corroborated by our results.

Furthermore, there was no significant difference in measuring SoL between the WAT and PSG ($p = 0.691$), indicating that the WAT accurately measures the time required for sleep initiation. Additionally, no significant differences were observed in the classification of sleep stages ($p > 0.05$), which suggests that the WAT can reliably categorize different sleep stages—including light sleep (LS), deep sleep (DS), and REM sleep). Our results regarding REM sleep align with previous findings on the Fitbit Charge 4 and Fitbit Sense, although Dong et al. (12) noted that the Fitbit Charge 4 tended to overestimate LS while significantly underestimating DS. These discrepancies may be attributed to differences in study populations, device versions, and other potential confounding factors.

A notable finding was the significant underestimation of wake-after-sleep onset (WASO) by the WAT (mean difference: 23.11 minutes, $p = 0.011$). Given that WASO is a key indicator of sleep quality and is linked to clinical outcomes such as daytime fatigue, impaired cognitive performance, and increased risks of chronic conditions (e.g., cardiovascular disease, diabetes)

(16), this discrepancy is clinically important. The underestimation may be attributed to the WAT's reliance on movement and heart rate signals for sleep/wake detection, which could lead to misclassification of brief awakenings as sleep (14, 17, 18). Additionally, factors such as age, obesity, and comorbid conditions might further impair the device’s wake detection capabilities (2, 19). The study population, consisting of patients referred for PSG due to suspected sleep disorders, might exhibit different sleep patterns and characteristics compared to healthy individuals, potentially contributing to the observed discrepancy in WASO measurement. Although our WASO findings are consistent with those from studies on the Fitbit Charge 2 and Fitbit Sense (11, 14), one study on the Fitbit Charge 4 reported no significant difference in WASO ($p = 0.6426$). Therefore, while the WAT shows promise in providing useful sleep metrics, clinicians should interpret WASO data with caution and consider complementary assessment tools when necessary.

4.3. Study strengths and limitations

This study has several strengths, including its pioneering validation of a commercially available sleep tracker in a Vietnamese clinical setting and its use of PSG as the reference standard, which enhances the reliability and validity of the findings. However, certain limitations should be noted. First, our sample size was relatively small due to resource constraints, specifically the limited availability of PSG at our institution. Second, we did not systematically collect detailed data on chronic comorbidities, although patients with significant psychiatric or medical conditions and medication use affecting sleep were excluded. Third, analyses were not adjusted for potential confounding factors (e.g., age, gender, BMI), given the sample size constraints and preliminary validation focus. Fourth, PSG scoring by a single technician precluded assessment of inter-rater reliability. Additionally, the proprietary algorithm of the WAT was not independently validated, and our epoch-by-epoch analysis excluded incomplete or artifact-containing epochs, potentially affecting data completeness. Finally, the assumption of normal distribution of paired differences required for Bland–Altman analysis was not fully met in our sample, although there were no strong or extreme outliers. As such, the agreement results should be interpreted with caution, especially in light of the modest sample size. At this stage, the evidence is not yet sufficient to recommend the WAT as a full substitute for polysomnography in clinical practice. Future research with larger samples, subgroup analyses, independent device validation, as well as investigations into cost-effectiveness and practical implementation, are recommended.

5. Conclusion

In summary, while the WAT demonstrated promising agreement with PSG for several key sleep parameters in a Vietnamese clinical population, these findings should be considered preliminary. Further studies are needed to confirm its utility and determine the appropriate role of WAT in routine sleep assessment, particularly in resource-constrained environments.

References

1. 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;17(10):2115-9.
2. Roth T. Insomnia: definition, prevalence, etiology, and consequences. J Clin Sleep Med. 2007;3(5 Suppl):S7-10.
3. Schutte-Rodin S, Broch L, Buysse D, Dorsey C, Sateia M. Clinical guideline for the evaluation and management of chronic insomnia in adults. J Clin Sleep Med. 2008;4(5):487-504.
4. Adjaye-Gbewonyo D, Ng AE, Black LI. Sleep Difficulties in Adults: United States, 2020. NCHS Data Brief. 2022(436):1-8.
5. 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(3):479-504.
6. 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(5):877-80.
7. Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF, et al. AASM Scoring Manual Updates for 2017 (Version 2.4). J Clin Sleep Med. 2017;13(5):665-6.
8. Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med. 1996;3(9):895-900.
9. Evenson KR, Spade CL. Review of Validity and Reliability of Garmin Activity Trackers. J Meas Phys Behav. 2020;3(2):170-85.
10. Aktaruzzaman M, Rivolta MW, Karmacharya R, Scarabottolo N, Pugnetti L, Garegnani M, et al. Performance comparison between wrist and chest actigraphy in combination with heart rate variability for sleep classification. Comput Biol Med. 2017;89:212-21.
11. Moreno-Pino F, Porras-Segovia A, Lopez-Esteban P, Artes A, Baca-Garcia E. Validation of Fitbit Charge 2 and Fitbit Alta HR Against Polysomnography for Assessing Sleep in Adults With Obstructive Sleep Apnea. J Clin Sleep Med. 2019;15(11):1645-53.
12. Dong X, Yang S, Guo Y, Lv P, Wang M, Li Y. Validation of Fitbit Charge 4 for assessing sleep in Chinese patients with chronic insomnia: A comparison against polysomnography and actigraphy. PLoS One. 2022;17(10):e0275287.

13. Menghini L, Yuksel D, Goldstone A, Baker FC, de Zambotti M. Performance of Fitbit Charge 3 against polysomnography in measuring sleep in adolescent boys and girls. Chronobiol Int. 2021;38(7):1010-22.
14. Ogasawara M, Takeshima M, Kosaka S, Imanishi A, Itoh Y, Fujiwara D, et al. Exploratory Validation of Sleep-Tracking Devices in Patients with Psychiatric Disorders. Nat Sci Sleep. 2023;15:301-12.
15. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of Wristband Fitbit Models in Assessing Sleep: Systematic Review and Meta-Analysis. J Med Internet Res. 2019;21(11):e16273.
16. Ramar K, Olson EJ. Management of common sleep disorders. Am Fam Physician. 2013;88(4):231-8.
17. de Zambotti M, Cellini N, Goldstone A, Colrain IM, Baker FC. Wearable Sleep Technology in Clinical and Research Settings. Med Sci Sports Exerc. 2019;51(7):1538-57.
18. Fitbit. Fitbit Charge 5™ 2024 [Available from: https://www.fitbit.com/global/us/products/trackers/charge5.
19. Taylor DJ, Mallory LJ, Lichstein KL, Durrence HH, Riedel BW, Bush AJ. Comorbidity of chronic insomnia with medical problems. Sleep. 2007;30(2):213-8.

Table 1. Sleep characteristics according to polysomnography

Variable	Mean ± SD	Median (IQR)	Min–Max
TIB, min	443.41 (84.85)	467 (444.5–489.5)	141–523.5
TST, min	336.20 (86.55)	353.5 (290–394)	62–491
SE, %	76.13 (12.62)	78 (70.7–84.6)	42.4–94.3
SoL, min	33.12 (25.32)	29.5 (12.5–49.5)	4–139.5
WASO, min	72.75 (47.42)	61 (33–105)	16–204
LS, min	258.55 (68.49)	267.5 (225–315.5)	61–342.5
DS, min	19.34 (31.13)	1 (0–25)	0–102
REM, min	61.19 (38.49)	55.5 (38–76)	0–149

Data are presented as mean (standard deviation), median (interquartile range), and range (minimum-maximum).

TIB indicates time in bed; TST, total sleep time; SE, sleep efficiency; SoL, sleep-onset latency; WASO, wake after sleep onset; LS, light sleep; DS, deep sleep; REM, rapid eye movement.

Table 2. Sensitivity, specificity, and accuracy of the wearable activity tracker in determining sleep-wake states

Value	Mean (SD)	95% CI
Sensitivity (actual sleep)	0.93 (0.06)	0.91–0.95
Specificity (actual wakefulness)	0.44 (0.19)	0.38–0.49
Accuracy (actual sleep + actual wakefulness)	0.79 (0.09)	0.76–0.81

Data are presented as mean (standard deviation), median (interquartile range), and range (minimum-maximum).
TIB indicates time in bed; TST, total sleep time; SE, sleep efficiency; SoL, sleep-onset latency; WASO, wake after sleep onset; LS, light sleep; DS, deep sleep; REM, rapid eye movement.

Table 3. Bland–Altman analysis of PSG vs WAT sleep variables

Variable	PSG, Mean (SD)	FBC, Mean (SD)	Bias (95% CI)	LLOA (95% CI)	ULOA (95% CI)	P Value
TST, min	336.2 (86.6)	366.3 (85.2)	–30.1 (–43.8 to –16.4)	–121.7 (–149.6 to –102.0)	61.5 (41.9 to 89.5)	0.835
SE, %	76.1 (12.6)	82.3 (10.5)	–6.2 (–8.8 to –3.5)	–23.9 (–29.3 to –20.1)	11.6 (7.8 to 17.0)	0.078
SoL, min	33.1 (25.3)	29.8 (24.6)	3.4 (–0.3 to 7.0)	–21.2 (–28.7 to –15.9)	27.9 (22.6 to 35.4)	0.691
WASO, min	72.8 (47.4)	49.7 (34.5)	23.1 (11.5 to 34.7)	–54.4 (–78.1 to –37.8)	84.0 (83.0 to 124.3)	**0.011**
LS, min	258.6 (68.5)	253.8 (67.2)	4.8 (–9.6 to 19.1)	–91.2 (–120.4 to –70.6)	100.7 (80.1 to 129.9)	0.854
DS, min	19.3 (31.1)	42.2 (28.7)	–22.9 (–30.2 to –15.4)	–72.3 (–87.3 to –61.7)	26.6 (16.0 to 41.6)	0.475
REM, min	61.2 (38.5)	70.3 (38.5)	–9.1 (–17.2 to –0.9)	–63.6 (–80.2 to –51.9)	45.4 (33.7 to 62.0)	0.995

Results of Bland–Altman analysis comparing polysomnography (PSG) and the wearable activity tracker (WAT) device. Data are shown as mean (standard deviation) for both PSG and FBC. Bias is the mean difference (PSG minus WAT), while the lower (LLOA) and upper (ULOA) limits of agreement are presented with 95% confidence intervals (CIs).

Abbreviations: *TST, total sleep time; SE, sleep efficiency; SoL, sleep-onset latency; WASO, wake after sleep onset; LS, light sleep; DS, deep sleep; REM, rapid eye movement; SD, standard deviation; CI, confidence interval; LLOA, lower limit of agreement; ULOA, upper limit of agreement.*

Note: Rounding was performed to one decimal place. All p values were calculated by Bland–Altman analysis.