OXFORD Sleep Research Society®

SLEEP Advances, 2026, **7**, zpag001

doi:10.1093/sleepadvances/zpag001

Advance Access publication 21 January 2026

Original Article

Original Article

Usability and stability of longitudinal at-home sleep evaluation using the Waveband electroencephalogram headband in an insomnia population

Silvia Frati Savietto*, Antoine Guillot, Jay Pathmanathan, Mason Harris, Kristy Nordstrom, M. Brandon Westover, Delphine Lemoine and Jacob Donoghue

Beacon Biosignals, Inc., Boston, MA, United States

*Corresponding author. Silvia Frati Savietto, Beacon Biosignals, Inc., 22 Boston Wharf Road, Boston, MA 02114, United States. Email: silvia.fratisavietto@beacon.bio.

Abstract

Study Objectives: Sleep monitoring outside of clinics could enhance care for insomnia and other sleep disorders but requires home systems that are easily operable and provide consistent data quality over multiple nights. We assessed the Waveband for usability by participants and feasibility of obtaining multi-night sleep data in the home setting.

Methods: 15 subjects with insomnia wore the Waveband electroencephalogram headband and an FDA-cleared wearable home sleep testing device, WatchPAT ONE ("WP1") for three nights. Usability was assessed via the System Usability Scale (SUS). Feasibility of participants to collect data was evaluated by examining stability of measured total sleep time in relation to measurements from the reference device (WP1) and data quality as evaluated by three human experts.

Results: Average SUS score was 69.7, meeting the 68-point threshold for good usability. Total sleep time recorded by the Waveband and WP1 devices showed a correlation of 87.3 per cent. All the recordings had an average of over 7 scorable hours of data per night.

Conclusions: Waveband demonstrated good usability by patients, was operable by patients, and generated interpretable data that provided stable sleep estimates across nights, comparable to an established home sleep testing device. The device has potential to advance patient care, sleep research, and clinical trials by enabling longitudinal ambulatory sleep assessment.

Key words: ambulatory sleep staging; wearable sensor; home sleep monitoring

Statement of Significance

Home sleep testing is a valuable resource to provide accurate assessment of sleep in a person's native environment. The ability to incorporate electroencephalogram (EEG) data into home sleep testing would be valuable for detailed assessment of sleep in a variety of neurological and psychiatric conditions with significant impacts on sleep. This study provides usability and feasibility data for the Waveband EEG headband, designed for patient use in the home environment to provide accurate EEG-based sleep staging. Results demonstrate that the Waveband device can be used by people without medical oversight to collect high quality data.

Introduction

Sleep is a critical component of health [1]. The primary method of evaluating sleep quality in a clinical setting traditionally hinges on polysomnography (PSG) [2] with a primary focus on electroencephalogram (EEG) measured brain activity to define clinical sleep stages [3]. PSG is typically performed as a single-night analysis conducted by a trained operator (technician or technologist) in a dedicated sleep lab facility. This process is expensive, resource-intensive, disruptive to natural sleep (both due to the alien recording environment of the sleep lab and the use of multiple wired contacts over the body), and not conducive to consecutive nights of monitoring, limiting its utility for longitudinal sleep assessment and reducing diagnostic yield. Sleep monitoring is thus generally restricted to diagnosing specific disorders amenable to limited sampling, rather than ongoing monitoring of conditions or symptoms [4]. Furthermore, in-lab PSG results may not reflect natural sleep, given the artificial environment and lack of reproducible measures. Consequently, there is a significant medical need for the ability to monitor sleep over multiple nights in the patient's natural home environment, without the inconvenience and expense of overnight clinic stays or sleep technologist visits. This is particularly relevant for conditions with variable symptoms, such as insomnia and narcolepsy, where long-term assessment could enhance understanding and management

Submitted: 23 June, 2025; **Revised:** 8 December, 2025; **Accepted:** 11 December, 2025

© The Author(s) 2026. Published by Oxford University Press on behalf of Sleep Research Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercialNoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

[5–8]. The Waveband EEG headband (formerly Dreem 3) represents an innovative solution, offering ambulatory sleep assessment via dry-electrode EEG that is relatively inexpensive, user-operable, and capable of providing consistent, high-quality data over many nights. This technology has the potential to advance patient care, clinical trials, and sleep research.

While several studies have used the Dreem 2 and 3 headbands in prior studies for sleep monitoring [9–15], formal studies specifically evaluating the usability and feasibility for longitudinal at-home use have not been previously published. In order to assess the usability and performance of the Waveband device in the home setting by untrained subjects, the Livie-1 study was conducted (NCT05611099, "Assessment of the stability, robustness and usability of the Waveband for EEG sleep monitoring over multiple nights in the home setting, in an insomnia population"). This study had three primary objectives (points 2 and 3 demonstrating feasibility of participants to capture data longitudinally at home):

(1) Usability: Assess usability of Waveband by patients with known or suspected insomnia over multiple nights, in a home environment.
(2) Stability: Assess the stability and robustness of Waveband data acquired by patients in their home environment. This was assessed by evaluating consistency of total sleep time (TST) output of the Waveband as compared to sleep determined by a device with 510(k) clearance for home based sleep assessment (the WatchPAT ONE, WP1, home sleep study device was chosen as the reference device based on its regulatory clearance for sleep–wake evaluation and proven home usability), over a three-night measurement period. It should be noted that detailed assessment of sleep staging accuracy was NOT a goal of this study, as a companion study (Octave-3) was performed to compare sleep staging outputs from Waveband to clinical gold-standard PSG recording.
(3) Quality: Assess quality of signals recorded with Waveband.

The study also had three exploratory objectives:

(1) Quantitative user feedback: Collect quantitative feedback on critical tasks related to using Waveband over multiple nights in a home environment.
(2) Qualitative user feedback: Collect qualitative information about perceptions, opinions, beliefs, and attitudes of individuals on the overall subjective experience of using Waveband over multiple nights in a home environment.
(3) Night-to-night variability: Compare intra-subject inter-night variability measured with Waveband and WP1.

An add-on study, Livie-2, examined the data quality obtained with the Waveband device. The hypothesis of these studies was that patients could operate the Waveband EEG Headband at home and without oversight from trained medical staff to collect data adequate for sleep staging, much like wristband sensors used for actigraphy-based sleep measurement. The results of these clinical trials plus a separate study comparing Waveband to in-lab PSG (Octave-3) were ultimately used to support FDA clearance of the Waveband device.

Materials and Methods

Study design

The present study utilizes data from the Livie-1 study, a prospective observational study, and Livie-2, a retrospective analysis of Livie-1 data. Livie-1 was designed to examine the ability of subjects to operate the Waveband device (formerly Dreem 3) at home over multiple nights, as compared to a wrist worn peripheral arterial tone signal and actigraphy device (WatchPAT One, WP1, Itamar Medical Ltd, a subsidiary of ZOLL Medical Corporation). The studies were conducted under a protocol approved by the Advarra IRB (protocol # Pro00065468, approval date August 16, 2022); all patients provided written informed consent. The study enrolled 15 subjects each instructed to record three nights of sleep, with the goal of obtaining 37 recordings.

Sample size calculations were based on a one-sided non-inferiority test, a standard approach in validation studies where the goal is to demonstrate that a new device performs no worse than a benchmark by more than a pre-specified margin. Specifically, we estimated that 37 records would provide 80 per cent power to detect non-inferiority on the "Wake" classification of a 5×5 confusion matrix, assuming a non-inferiority threshold of 0.667, an expected agreement of 0.735, a standard deviation of 0.15, and a one-sided alpha of 0.05. Accounting for a 10 per cent dropout rate, the enrollment target was 41 records. Each subject contributed up to three nights of data, making 15 subjects sufficient to achieve the required number of analyzable records. While the sample size was powered for non-inferiority, the main results (e.g. correlation, intraclass correlation coefficient [ICC]) are descriptive and do not involve formal hypothesis testing. All other statistical tests are two-tailed unless otherwise noted.

Study endpoints

Outcome measures for the primary objectives described above were:

(1) Usability: System Usability Scale (SUS): means, standard deviations, frequencies, and scores of the SUS questionnaire are presented using descriptive statistics. SUS is described below.
(2) Stability: Correlation between TST measured by Waveband versus WP1, on each of the three nights of recordings.
(3) Quality: Percentage of acceptable recordings as assessed by the formal grading system developed by Leach et al. for identification of sleep physiological signals [16]. The grading system is described in detail below.

Outcome measures for the exploratory objectives described above were:

(1) Quantitative user feedback: means, standard deviations, frequencies, and percentages of detailed questionnaire results used as descriptive statistics.
(2) Qualitative user feedback: end-users' perceptions regarding Waveband usability obtained after content analysis of the transcription of the individual debriefing interviews.
(3) Night-to-night variability: correlation of the intra-subject inter-nights ΔTST measured with the Waveband and WP1.

Eligibility criteria

Subjects were eligible to participate if they met all of the following criteria: (1) between the ages of 22 and 70 years (inclusive), (2) under the care of a sleep physician for diagnosed or suspected insomnia, (3) had a Wi-Fi connection at home, (4) had a smartphone to install an app required for the study (the Waveband phone app), (5) self-reported no drug or alcohol consumption 24 h before the start of recording, (6) agreed to abstain from drugs and alcohol during the 3 days of measurement, and (7) able to read, understand, and sign an informed consent form.

Subjects were excluded for any of the following conditions: (1) BMI ≥ 40; (2) sleep apnea treated with Continuous Positive Airway Pressure (CPAP); (3) currently enrolled in a clinical trial of any investigational drug; (4) head circumference < 53 cm or device fitting issues as determined during training; and (5) shift worker or subject working unusual hours. All participants maintained overnight sleep schedules.

The target sample size was 15 subjects, with 45 total nights to reach a threshold of at least 37 good quality recordings. Detailed study protocols are included as Appendices 1 and 2. A study flow diagram is shown in Figure S1.

Study duration

The total enrollment period was 20 days, including the 10 days for shipping the devices. Individual subjects were recorded at their home for three nights wearing Waveband and a simultaneously worn WP1.

Study procedures

Subjects were identified via three routes: by the Principal Investigator and Dreem Sleep Clinic of California clinical team; by referrals from the Principal Investigator and DSC clinical team colleagues in other private practices in California; and by advertising on social media for people who had experienced difficulty sleeping and insomnia symptoms. Subjects were identified and contacted by Dreem Sleep Clinic of California clinicians before being screened by the Principal Investigator (PI) or their designee for confirming recruitment eligibility. To enroll, subjects had to give informed written consent.

Enrollment in the study required subjects to self-operate the Waveband and WP1 devices. Both devices were shipped to subjects, and subjects were asked to download device apps onto their phones. Training was provided via video conference to ensure subjects understood how to set up both devices in the context of the study. The training session also allowed the trainer to verify that Waveband fit the subjects' head.

Subjects underwent three consecutive nights of Waveband and WP1 sleep assessment. During these three nights, subjects were asked to set up and use the devices autonomously. Subjects were allowed to do quiet activities (reading, film watching, chatting) but to instructed to start WP1 recording, followed by Waveband immediately before going to sleep. Subjects were permitted to sleep freely, and there was no prescribed sleep onset time, although all participants. Subjects were instructed to start recording with both devices just prior to sleep, which was estimated to take approximately 10 min. Subjects were instructed to stop recordings upon waking in the morning, without requirement for minimum time in bed or minimum recording duration. This procedure was repeated for the three consecutive nights of assessments. Daytime recording was not evaluated, including naps (if any, which were not explicitly forbidden or regulated). While alcohol and recreational drugs were to be avoided during the study, caffeine use was not restricted. At the end of each of these three nights, subjects were asked to ship back the Waveband and a telehealth interview was organized to collect user feedback and administer System Usability Scale questionnaires. WP1 devices did not have to be returned.

WP1 data quality was assessed using built-in device quality flags and recording completeness, as provided in the device export summary. Only nights in which WP1 recorded a full session without interruption and produced a valid sleep report were considered analyzable. Nights with device failure (e.g. incomplete recordings, missing outputs, or failure to generate a sleep report) were excluded from analysis. WP1 data were not manually edited or scored.

Waveband raw data were de-identified and saved on a secure server. The Waveband and WP1 devices were synchronized by aligning the start and end times of each recording based on timestamps recorded by the respective device apps. Time alignment was confirmed during preprocessing by visually inspecting the signal traces and sleep metrics (e.g. movement onset, sleep onset) to verify parallel trends. A difference of more than 2 min between the recorded durations or clear signal desynchrony was used as a threshold for exclusion. No recordings included in the analysis exceeded this threshold.

Study devices

The Waveband device is shown in Figure 1 with corresponding 10–20 EEG electrode positions as noted. The device is a wireless dry-electrode system which acquires, transmits, displays, and stores EEG records to produce automatic sleep staging. The Waveband headband records two types of physiological signals: (1) brain activity via dry electrode electroencephalography sensors (EEG), and (2) movements via a 3D accelerometer. An audio system delivering sounds via a bone conduction speaker is integrated in the frontal band but was not used in this study. The Waveband is composed of foam and fabric with an elastic band behind the head making it adjustable such that it is tight enough to be secure, but loose enough to minimize discomfort. The EEG signal is measured by two electrodes in the frontal band (frontotemporal positions correlating to F7 and F8) and two at the back of the head (occipital positions O1 and O2). The 3D accelerometer is embedded in the top part of the headband.

The WatchPat ONE (FDA cleared K183559), WP1: The WP1 device is a non-invasive home sleep assessment device for use with patients suspected of having sleep-related breathing disorders. WP1 is a diagnostic aid for the detection of sleep-related breathing disorders, sleep staging (Rapid Eye Movement (REM) Sleep, Light Sleep, Deep Sleep, and Wake), snoring level, and body position. It records several parameters, including PAT signal, pulse oximetry, and actigraphy. It uses actigraphy to provide a measurement of TST.

EEG data

EEG data were stored as European Data Format (EDF) files. The EEG data were recorded using the Waveband dry-EEG electrodes, which outputs data from the following bipolar channels: F7-O1, F8-O2, F7-F8, F7-O2, F8-O1. All recordings included in the Livie-1 study were analyzed (originally excluded data from Livie-1 were from subjects who removed the device during the night, subjects who were found to have an exclusionary medical condition, and a subject whose data was deemed uninterpretable for automated staging). Two recordings were included in this study that were not included in the original Livie-1 analysis because the WP1 failed. These two recordings were included here because WP1 data were not required for all analyses.

Signal quality was assessed using an established grading system for identification of sleep physiological signals, including differentiating valid signals from signals with artifact which may impede identification of sleep stages [16]. The protocol was performed by manual review and grading of epochs from EEG recordings taken using the Waveband during the Livie-1 study.

Sleep recordings (EDF files) were uploaded to Noxturnal Sleep Scoring software (6.0.2) to enable viewing and labeling of data for each 30 s epoch, aligning with segments that would have been scored for sleep. Each expert was instructed on how to load the

5 dry electrode EEG sensors:

A. 2 frontal sensors to measure frontal brain activity (F7 and F8)
B. 1 ground electrode on the frontal band (Fp1)
C. 2 sensors in the back of the headband to monitor occipital brain activity (O1 and O2)

Miscellaneous:

D. Bone conduction speaker for audio output
E. Power button
F. 3D accelerometer to measure head position and movement during sleep
G. Magnetic port for charging

Figure 1. Waveband and its sensors.

data, adjust the filters and display, score the data, export the labels, and upload and email the scoring file. EDF files were stored in an access-controlled folder. The first and last 10 min of each night were discarded from each recording, based on the prescribed analysis plan and estimates of the amount of time required for a subject to apply and activate or stop both devices.

Three qualified expert reviewers independently assessed the data. Expert reviewers had to have one or more of the following three qualifications: (1) board-certified neurologists trained in EEG assessment, (2) EEG technologists qualified to score EEG data, and (3) sleep technologists qualified to score EEG data.

Reviewers labeled each 30 s EEG epoch as unscorable (1) or scorable (2), defined as:

- Unscorable (1)—Artifact was so extensive as to make that 30-s epoch clinically unusable (data could not be used to score the sleep stage).
- Scorable (2)—Artifacts may or may not have been present but did not limit the ability to sleep stage the epoch. This could include brief artifacts, artifacts affecting only a subset of channels, or a combination of brief and spatially limited artifact. If, based on assessment by the expert, the epoch could be sleep scored despite presence of artifact, it was considered scorable even if artifact was severe on some channels or intermittent on all channels.

Starting at Epoch #1, grading data quality was performed as follows. The high pass filter was set to 0.3 Hz. The low pass filter was then set to 35 Hz. Number pad key "1" was entered for unscorable data and number pad key "2" for scorable EEG. Labels were exported to an excel spreadsheet.

Scorability assessments were limited to determining whether EEG epochs were of sufficient quality to allow scoring, but manual sleep staging was not performed in this study (this was performed with PSG as the gold standard in the Octave-3 study, reported separately). Automated sleep staging was not otherwise evaluated in this analysis.

Patient reported outcomes

The usability of Waveband was assessed by the System Usability Scale (SUS) questionnaire [17, 18] (Table S1) and presented to each subject during an end-of-the study interview. SUS is based on 10 standard questions that can be adapted to suit the system evaluated. SUS was historically developed to determine how users perceive the ease of use of software. It is now commonly used to assess usability of hardware and software products and product interfaces. It consists of 10 statements based on a Likert scale, each with five possible answers ranging from complete rejection to complete agreement: from 1 = strongly disagree to 5 = strongly agree. Responses to each question are summed using a weighted scale to provide scores ranging from 0 to 100 (higher scores indicating greater usability). Generally, systems in early development may expect to have a rating of 30, while more mature systems should rate 60–80. A SUS score above 68 shows that the device's usability is above average. SUS has demonstrated high reliability and validity for assessing medical software [17, 18].

A debriefing interview aimed to collect users' perspectives and perceptions on the device, complementing task fulfillment observations. Participants were asked to complete a secondary usability questionnaire aiming at collecting detailed feedback on specific tasks of Waveband intended use, including unboxing, setting up, launching, and stopping a recording (Table S2).

Adverse events

Adverse events (AEs) (defined in the Livie-1 protocol, Appendix 1) were recorded, and characterized by type, severity, seriousness, and relationship to the device. AEs and serious adverse events (SAEs) were defined in the detailed study protocol prior to study initiation. The list of AEs is presented in Table S3. No SAEs were reported. No further treatment of AEs was required.

Statistical analysis

Statistical analysis sets. The Evaluable Population was defined as all study participants who signed the informed consent, met all screening criteria, enrolled, and had data collected from both Waveband and WP1 that passed a data upload quality assessment on the same night and completed the final interview. The Evaluable Population was used for all summaries and analyses.

Overview of statistical analyses. Descriptive statistics for categorical variables include the number and percent of study participants with each characteristic. Correlation between Waveband and WP1 TST was assessed using Pearson's correlation coefficient. Agreement was further evaluated using the two-way mixed-effects ICC for absolute agreement. Confidence intervals

for both metrics were calculated using a nonparametric bootstrap procedure.

Demographic and baseline characteristic data summarization are provided, including study site, gender, race, age, BMI, and baseline characteristics related to medical history. Percentages are based on the number of study participants with non-missing values. A tabulation of the disposition of study participants is provided, including the number enrolled, the number assessed, and the reasons for study discontinuation. TST was summarized using descriptive statistics for each of Waveband and WP1 records. Incidence of AEs leading to early discontinuation of study participation is also presented. Means, standard deviations, frequencies, and scores of the SUS questionnaire are presented as descriptive statistics. Confidence intervals for performance statistics were calculated where relevant by bootstrapping. Sample size calculations were generated to achieve 80 per cent power to detect non-inferiority for agreement on awake segments (adjusted to account for three nights of recording and a 10 per cent dropout rate).

Scorability analysis. The statistical analysis was conducted on manually scored data. Epochs were considered of scorable quality if at least 2 out of 3 reviewers deemed the epoch scorable. Consensus scores were tabulated as total scorable epochs and epochs determined to be of not scorable quality were determined for each EDF representing a night of sleep for each subject.

An acceptance criterion of 90 per cent of all recordings having 4 h total time of scorable epochs was then applied to the data. This aligns with American Academy of Sleep Medicine PSG guidelines for a minimum of 4 h of data. Given that Waveband is intended for multiple night recording, at a 90 per cent success rate, we would expect independently distributed artifact and recording errors would result in a <1 per cent failure rate if used for two or more nights.

Note that all EEG epochs from each available recording were graded as scorable or unscorable by expert reviewers. The primary signal quality analysis computed the percentage of scorable epochs per night. This analysis included all recordings with usable Waveband data, regardless of overall quality. Separately, we reported the proportion of nights meeting a minimum threshold of 4 h of scorable data, based on AASM PSG guidelines, to assess clinical utility.

Results

Fifteen subjects with confirmed insomnia symptoms were enrolled. Fourteen completed the whole study (three consecutive nights of measurement, end of study questionnaire) while 1 subject completed two nights of measurement and the end of study questionnaire. Of the 14 subjects completing three nights, 1 was excluded due to a finding of severe OSA, 3 failed to activate the Waveband headband on one night. This resulted in 10 subjects with three nights of recording and 3 subjects with three nights of recording.

Demographics are shown in Table 1. Subjects averaged 52.6 years in age (range 28 to 65) and were predominantly female (5 [33.3 per cent] were males and 10 [66.6 per cent] were females, consistent with the demographics of insomnia).

Of those 15 subjects there were 45 possible nights; of which 34 recordings were included in the analysis set. Reasons for excluding records included incidental findings of severe obstructive sleep apnea (subject LVS014, excluded due to based on protocol, three nights excluded), subject withdrawal (LVS016, who withdrew, two nights with low quality data excluded and last night was not recorded), missing WP1 data (LVS008 night 1 and LVS010 night 3), and missing Waveband data (LVS007 night 3, LVS010 night 1, and LVS011 night 1). These are detailed in Table S4, and recordings included in the dataset are shown in Table 3. Therefore, missed or lost Waveband nights accounted for 3/45 nights, and missed or lost WP1 nights accounted for 2/45 nights.

Table 1. Demographic data on the evaluable population

Parameter (unit)	Statistics/ category	All evaluable subjects (*N* = 15)
Age (years)	Mean ± *SD*	52.6 ± 10.9
	Median	54.0
	Min; Max	28.0; 65.0
Sex	Female *n* (%)	10 (66.6)
	Male *n* (%)	5 (33.3)
Ethnicity	White *n* (%)	9 (60.0)
	Asian *n* (%)	3 (20.0)
	Hispanic *n* (%)	1 (6.6)
	Black *n* (%)	2 (13.3)
BMI	Mean ± *SD*	26.0 ± 4.4
	Median	25.6
	Min; Max	19.9; 34.2

Stability primary analysis: Correlation between TST between WP1 and Waveband

We first analyzed stability and robustness of Waveband data collection, by calculating the correlation between TST measured by the Waveband and WP1 on each of the three-night recordings. Over the 34 records, the average TST recorded by the Waveband was 419.2 min (std = 73.6), and the average TST recorded by WP1 was 407.4 min (std = 64.3).

A Bland–Altman analysis was performed to quantify agreement in TST between Waveband and WP1 across all recorded nights (34 nights from 13 subjects), as shown in Figure 2. Overall, both systems tracked total sleep similarly, with small and mostly consistent differences. As already noted, Waveband tended to report slightly longer TST. The distribution of paired differences was approximately normal (Shapiro–Wilk $W = 0.96$, $p = .27$). A linear mixed-effects model with random intercepts for subjects showed no evidence of proportional bias ($\beta = 0.02$ min per min TST, $t = 0.27$), indicating that the mean difference between devices did not vary systematically with TST. The estimated fixed-effects bias was small (intercept = 4.9 min), although substantial between-subject variability in bias was present (*SD* = 33 min). To characterize the variability of the differences, absolute residuals from the bias model were regressed on average TST in a second mixed-effects model. This model revealed no statistically significant heteroscedasticity ($\beta = -0.07$, $t = -1.27$), but predicted absolute residuals were converted to TST-dependent standard deviations and used to compute regression-based limits of agreement (LoA) following Menghini et al. [19]. The resulting LoA showed a slight narrowing across the TST range, with the upper LoA decreasing modestly (slope = −0.15) and the lower LoA increasing (slope = 0.19). Empirical 2.5th and 97.5th percentiles (−65.6 to +72.1 min) were included as non-parametric reference limits. Overall, the analysis indicates that Waveband exhibits a largely constant bias relative to the reference, with wide but stable night-to-night variability and only mild TST-dependent changes in the LoA. The correlation between the Waveband and the WP1 measures was 87 per cent [95% CI: 80%, 93%], demonstrating high correlation. We also calculated the ICC between TST recorded by the two devices. ICC

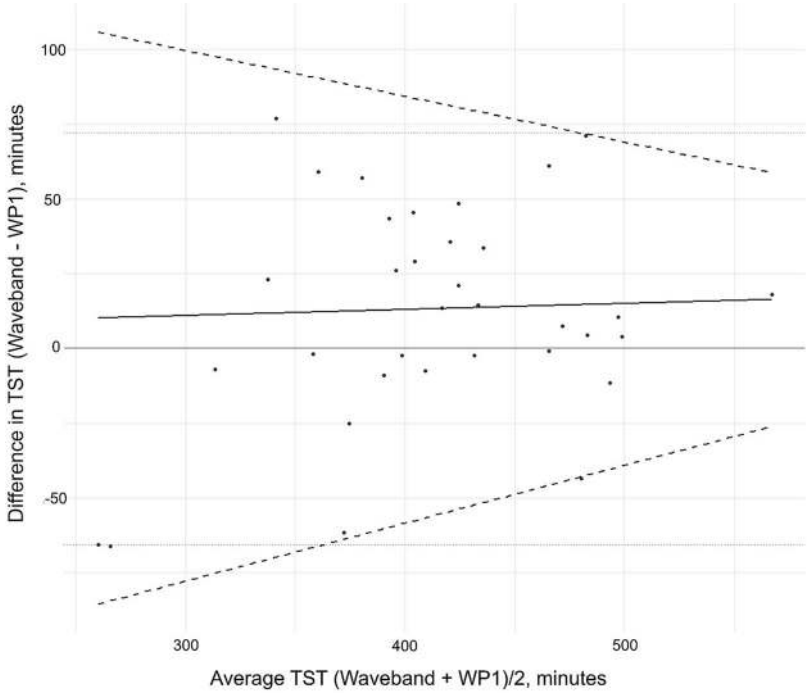

Figure 2. Bland–Altman plot comparing total sleep time (TST) between waveband and WP1. Differences (waveband – WP1) are plotted against average TST. The solid line shows the estimated mean bias, and dashed lines show regression-based limits of agreement derived from mixed-effects modeling. Dotted lines show the empirical 2.5th and 97.5th percentiles.

Table 2. Primary analysis—SUS score

Parameter	Statistics	All evaluable subjects ($N = 15$)
SUS score	*n*	15
	Mean ± *SD*	69.67 ± 18.02
	Median	72.5
	Min; Max	35.0; 92.5

reached 85 per cent with a 5 per cent lower bound confidence interval of 69 per cent indicating moderate agreement.

Stability secondary analysis: ΔTotal sleep time between WP1 and Waveband

The secondary stability analysis assessed the within-subject night-to-night variability in TST, calculated separately for Waveband and WP1. For each subject, we computed the difference in TST between consecutive nights (e.g. Night 1–Night 2 and Night 2–Night 3) for each device. Across 29 such pairs of nights, the average ΔTST was −4.03 min (std = 67.8) for Waveband and −4.40 min (std = 69.2) for WP1, indicating similar night-to-night variability. The correlation between the ΔTST values derived from each device was 91.6 per cent, and the ICC was 91.9 per cent, suggesting strong agreement in capturing inter-night variability.

Usability

The average SUS score of the Waveband in the evaluable population was 69.7 ± 18.0, with a minimum SUS score value of 35.0 and a maximum score of 92.5 (Table 2). Thus, Waveband met the primary usability outcome of the Livie-1 study (SUS ≥ 68).

Exploratory usability analysis: Quantitative and qualitative user feedback. We next investigated additional user feedback, both qualitative and quantitative. We focused especially on responses from subjects who did not consider the device to have good usability, to understand what could be improved to increase the user experience, and to understand if concerns raised impact the overall safety and effectiveness of the device. To differentiate between usability issues related to hardware, software, or to the entire system, detailed questionnaires were presented to the subjects, as described in Materials and Methods.

Usability feedback. No issues were reported related to unboxing the device and identification of the detachable components (battery and size adjusters). However, 33 per cent of subjects reported the instructions provided were not sufficient, and 13 per cent reported issues with understanding how to charge the device and transfer data.

As the associated phone app provided complementary instructions, including a step-by-step tutorial showing how to set up the device and position it, we next analyzed whether these complementary instructions allowed effective use of the device. Twenty six per cent of subjects reported they didn't understand the information provided by the mobile app. However, the first steps of the Headband's configuration and tutorial (downloading the app, login, pairing, and Wi-Fi set-up) were correctly understood by most participants, with only 26 per cent of participants having issues starting and stopping a recording from the app.

No issue was raised regarding how to wear the Waveband headband.

Eighty six per cent of participants were able to set up the system in <15 min. Most participants reported no issue in remembering how to use the system. Eighty seven per cent of participants reported no objections to using Waveband longer than three nights. One participant (6 per cent) raised concerns about comfort. This participant decided to stop the measurement period after two nights of recording.

Investigation of AEs

There were two reports of AEs related to participant discomfort using the study devices. One participant reported she experienced pain in her finger in the middle of the night and had to take off the WP1 device to sleep. The pain resolved by morning and the event has therefore classified as mild. The participant pursued the study the next night using the WP1 on another finger, without further incident. A second participant reported a headache the morning after wearing the headband; this event was described in the preceding section. The severity of the event was reported as mild since the headache disappeared a few hours after waking up.

Signal quality: Primary analysis

We recruited three expert reviewers to perform signal quality assessments: (1) Jay Pathmanathan, MD, PhD (board certified in neurology, clinical neurophysiology [EEG]), and actively licensed in Massachusetts and Pennsylvania, with 18 years in neurology practice; (2) Kristy Nordstrom, C.LTM, R.EEG.T is a registered EEG Technologist with 18 years of experience at Massachusetts General Hospital; and (3) Mason Harris, RPSGT, a registered PSG technologist with 10 years of experience.

The shortest total recording was 381.5 min (about 6 h), and the longest was 629 min (~10.5 h). Hundred per cent of recordings met the acceptance criterion of 4 h total time of scorable epochs with an average of 7.6 h scorable data per night recorded. Results are summarized in Table 3. Excluded recordings are summarized in Tables S4 (for Livie-1) and S5 (for Livie-2).

Table 3. Percentage of scorable epochs. Each recording is shown with the percent scorable epochs, total minutes deemed scorable by at least two reviewers, and total recording duration

Subject	% Epochs scorable	Scorable (min)	Total (min)
LVS001-1	91.43%	485.5	531.0
LVS001-2	89.27%	478.5	536.0
LVS001-3	99.53%	532.5	535.0
LVS002-1	99.43%	523	526.0
LVS002-2	99.16%	528.5	533.0
LVS002-3	99.26%	540	544.0
LVS003-1	100.00%	457.5	457.5
LVS003-2	96.74%	460	475.5
LVS003-3	98.50%	559	567.5
LVS004-1	100.00%	456.5	456.5
LVS004-2	97.87%	437.5	447.0
LVS004-3	99.53%	426.5	428.5
LVS006-1	98.29%	461	469.0
LVS006-2	96.14%	423	440.0
LVS006-3	95.41%	364	381.5
LVS007-1	95.74%	382.5	399.5
LVS007-2	97.74%	410	419.5
LVS008-2	98.89%	489.5	495.0
LVS008-3	98.64%	578.5	586.5
LVS009-1	92.98%	477	513.0
LVS009-2	88.14%	412.5	468.0
LVS009-3	93.26%	450	482.5
LVS010-2	86.09%	541.5	629.0
LVS011-1	86.07%	358.5	416.5
LVS011-2	92.84%	408.5	440.0
LVS012-1	98.03%	523.5	534.0
LVS012-2	99.32%	436.5	439.5
LVS012-3	99.55%	446.5	448.5
LVS013-1	99.76%	419.5	420.5
LVS013-2	97.45%	401	411.5
LVS013-3	99.88%	414.5	415.0
LVS015-1	98.04%	424.5	433.0
LVS015-2	98.99%	442.5	447.0
LVS015-3	98.97%	432	436.5
Average	96.50%	458.3	475.4
Median	98.17%	448.25	457
Standard deviation	3.98%	55.9	59.0
Minimum	86.07%	358.5	381.5
Maximum	100%	578.5	629.0

Exploratory analyses: Signal quality

Additional supportive analyses were conducted to assess consistency of the scoring guidelines when applied by each reviewer and whether data quality was stable over the course of the night.

Pairwise agreement between each pair of scorers (proportion of epochs on which two scorers agreed) was: scorer 1 versus 2, agreement (95% CI): 96.4% (94.9%, 97.4%); 1 versus 3: 88.1% (83.1%, 91.7%), 2 versus 3: 87.3% (82.3%, 90.9%), yielding an average pairwise agreement of 90.6.

To assess whether Waveband signal quality could systemically vary over the course of the night, we compared the percentage of scorable epochs during the first half of each night to the percentage of scorable epochs during the second half of each night. The average difference between the first and second half of the night was −0.28 per cent [95% CI: −2.21%, 1.52%]. Therefore, these data do not support a difference in scorability based on either half of the night.

Discussion

This study provides evidence supporting the usability and feasibility of operating the Waveband EEG headband system in the home environment (without technologist oversight). The findings demonstrate that the device has good usability, shows stability across nights comparable to an established home sleep testing device, and provides EEG signals of sufficient quality for both manual and automated sleep scoring.

A key goal was to test stability of the Waveband data collection for home use over multiple nights. High correlation was observed between the devices for TST across the three study nights. Both devices also showed comparable night-to-night variability in TST. This analysis showed high agreement between Waveband and WP1 in TST across three nights. However, this correlation reflects both consistency in participant sleep patterns and the devices' ability to detect those patterns. Thus, while these results support the utility of Waveband for multi-night home monitoring, they do not isolate test–retest reliability of the device itself in a population with highly variable sleep.

We also found that patients found Waveband usable. The high average usability (SUS) score indicates overall good usability of Waveband. This met the pre-specified threshold for adequate usability. The score is also comparable to published SUS scores for other consumer sleep tracking devices [20, 21]. Further analysis of user feedback identified specific areas that could be improved in future iterations, including enhancing the LED light indicators, refining app instructions for starting/stopping recordings, and providing more guided setup instructions for first-time users. Notably, participants in this study received no hands-on training with either device, instead receiving video teleconference and app-based training. Critically however, no issues were reported regarding device safety, and most participants were able to effectively use the system after one night of accommodation.

It must be noted that this study chose the WP1 system as comparator to assess TST acquired in the home setting. WP1 is not intended to evaluate insomnia, but the device has FDA clearance to evaluate wake, light sleep, deep sleep, and REM sleep stages for the detection of sleep-related breathing disorders. Because it can therefore evaluate TST, and is easy to use in the home setting, it was chosen as the reference for data quality and sleep evaluation. Importantly, many other sleep metrics are of critical importance. However, to evaluate the performance of the Waveband EEG headband in evaluation of sleep stage derived metrics we have conducted a separate study (Octave-3) comparing performance to the clinical gold standard of in-lab PSG. Those results are reported separately (Octave-3 reference is in submission, will be cited here, or this will be removed if submission incomplete). Additionally, the Dreem headband (a precursor of the Waveband device) was previously compared to PSG and found to have excellent sleep staging concordance [11]. This study was focused solely on usability of the Waveband device in the home setting and feasibility of participants to collect high quality data longitudinally without technologist support—deferring detailed sleep staging performance to the Octave-3 study. However, note that Beau et al. [15] compared Dreem 3 to in-lab PSG in insomnia patients and examined TST, sleep latency, wake after sleep onset, and sleep stages. They found good concordance, although also noted a tendency to overestimate TST. We also found a consistently longer TST when compared to WP1, though the difference was small (slightly <12 min). As previously noted, a separate validation study (Octave-3) was performed to examine Waveband versus PSG sleep staging (Octave 3 reference, currently in submission).

Limitations of this study include an insomnia population with varying severity, lack of daytime data to assess for daytime naps (or deviation from typical nap frequency), lack of data on weekday versus weekend use (which may alter sleep patterns), or caffeine use (which was not limited or examined in this study)—changes to any of which may have altered typical nocturnal sleep patterns. Therefore, the extent to which these factors could impact usability is unknown, as this study did not collect detailed information to characterize patient's unique insomnia presentations. However, in a general insomnia population Waveband demonstrated usability and feasibility to assess sleep over multiple nights. Another limitation is that this study used a per-protocol analysis including only nights with valid data from both devices. While appropriate for validation, this may underestimate real-world data loss. In clinical and research settings, such loss could require re-collection or result in missing data. Future studies should evaluate performance under intention-to-treat conditions and insomnia subpopulations. Nevertheless, all 15 participants who enrolled completed the SUS survey and are included. Thus, the impact of any selection bias is likely minimal for usability metrics.

The detailed signal quality analysis provides validation that the quality of Waveband EEG data is sufficient for accurate sleep scoring. Nearly all epochs met quality standards for scorability when rated by expert technologists and neurologists. The high inter-rater agreement further supports the reliability of this scoring methodology. These results confirm Waveband data quality is adequate for clinical and research use.

Conclusion

In conclusion, this study provides evidence for the usability, stability, and signal quality of the Waveband system for at-home sleep monitoring. The device demonstrated good usability based on high average SUS scores. User feedback also indicated that the headband was comfortable for multi-night use and participants successfully recorded at least two nights, although feedback highlighted points where patient instructions could use clarification. The device showed stability in performance across multiple nights, with minimal night-to-night variability in key sleep metrics compared to the reference device. Finally, the EEG signal quality was sufficient for manual and automated scoring, meeting pre-specified criteria for percentage of scorable epochs. Overall, these results support use of the Waveband system for longitudinal EEG sleep monitoring in the home environment. The device has the potential to advance patient care, sleep research, and clinical trials by enabling longitudinal ambulatory sleep assessment.

Supplementary material

Supplementary material is available at *SLEEP Advances* online.

Author contributions

Silvia Frati Savietto (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Methodology [equal], Project administration [lead]), Antoine Guillot (Conceptualization [equal], Data curation [equal], Formal analysis [lead], Investigation [equal], Methodology [lead], Writing—original draft [lead], Writing—review & editing [equal]), Mason Harris (Formal analysis [equal], Methodology [equal], Validation [equal]), Kristy Nordstrom (Formal analysis [equal]), Michael Brandon Westover (Formal analysis [equal], Writing—review & editing [equal]), Delphine Lemoine (Conceptualization [equal], Investigation [equal], Project administration [equal]), Jay Pathmanathan (Formal analysis [supporting], Validation [equal], Writing—original draft [supporting], Writing—review & editing [lead]), and Jacob Donoghue (Data curation [equal], Funding acquisition [lead], Project administration [equal], Resources [lead], Writing—original draft [supporting], Writing—review & editing [supporting]).

Disclosure statement

Financial disclosure: The authors of this paper are employees of Beacon Biosignals, which owns the Waveband Device. The authors receive salary and stock in Beacon Biosignals. The Livie-1 study was funded by Dreem SAS, which was subsequently acquired by Beacon Biosignals. S.F.S. and A.G. were employees of Dreem. The Livie-2 study was funded by Beacon Biosignals.

Non-financial disclosure: None declared.

Data availability

Data acquired for this study, including PSG and Waveband data are proprietary and cannot be shared, except for summary statistics provided in this article.

References

1. Institute of Medicine (US) Committee on Sleep Medicine and Research. *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*. Washington (DC): National Academies Press (US); 2006.
2. Berry RB, Brooks R, Gamaldo CE, *et al. The AASM Manual for the Scoring of Sleep and Associated Events. Rules, Terminology, Technical Specifications*. Darien, IL: Am Acad Sleep Med; 2012.
3. Berry RB, Brooks R, Gamaldo C, *et al.* AASM scoring manual updates for 2017 (version 2.4). *J Clin Sleep Med*. 2017;**13**(5): 665–666. https://doi.org/10.5664/jcsm.6576

4. Kushida CA, Littner MR, Morgenthaler T, *et al*. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep*. 2005;**28**(4):499–521. https://doi.org/10.1093/sleep/28.4.499
5. Vgontzas AN, Liao D, Bixler EO, Chrousos GP, Vela-Bueno A. Insomnia with objective short sleep duration is associated with a high risk for hypertension. *Sleep*. 2009;**32**(4):491–497. https://doi.org/10.1093/sleep/32.4.491
6. 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 JCSM Off Publ Am Acad Sleep Med*. 2008;**4**:487–504. https://doi.org/10.5664/jcsm.27286
7. McCall C, McCall WV. Comparison of actigraphy with polysomnography and sleep logs in depressed insomniacs. *J Sleep Res*. 2012;**21**(1):122–127. https://doi.org/10.1111/j.1365-2869.2011.00917.x
8. Marino M, Li Y, Rueschman MN, *et al*. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*. 2013;**36**(11):1747–1755. https://doi.org/10.5665/sleep.3142
9. González DA, Wang D, Pollet E, *et al*. Performance of the Dreem 2 EEG headband, relative to polysomnography, for assessing sleep in Parkinson's disease. *Sleep Health*. 2024;**10**(1):24–30. https://doi.org/10.1016/j.sleh.2023.11.012
10. Ong JL, Golkashani HA, Ghorbani S, *et al*. Selecting a sleep tracker from EEG-based, iteratively improved, low-cost multisensor, and actigraphy-only devices. *Sleep Health*. 2024;**10**(1):9–23. https://doi.org/10.1016/j.sleh.2023.11.005
11. Arnal PJ, Thorey V, Debellemaniere E, *et al*. The Dreem headband compared to polysomnography for electroencephalographic signal acquisition and sleep staging. *Sleep*. 2020;**43**(11). https://doi.org/10.1093/sleep/zsaa097
12. Blaskovich B, Bullón-Tarrasó E, Pöhlchen D, *et al*. The utility of wearable headband electroencephalography and pulse photoplethysmography to assess cortical and physiological arousal in individuals with stress-related mental disorders. *J Sleep Res*. 2024;**33**:e14123. https://doi.org/10.1111/jsr.14123
13. Chinoy ED, Cuellar JA, Jameson JT, Markwald RR. 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. https://doi.org/10.2147/NSS.S348795
14. G Ravindran KK, Della Monica C, Atzori G, *et al*. Evaluation of Dreem headband for sleep staging and EEG spectral analysis in people living with Alzheimer's disease and older adults. *Sleep*. 2025;**48**(8). https://doi.org/10.1093/sleep/zsaf122
15. Beau QAL, Boardman JM, Mellor A, Ogeil RP, Drummond SPA. Dreem headband performance in insomnia disorder: an in-lab comparison to polysomnography. *Sleep Med*. 2025;**134**:106698. https://doi.org/10.1016/j.sleep.2025.106698
16. Leach S, Chung K-Y, Tüshaus L, Huber R, Karlen W. A protocol for comparing dry and wet EEG electrodes during sleep. *Front Neurosci*. 2020;**14**:586. https://doi.org/10.3389/fnins.2020.00586
17. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Intl J Human–Computer Interact*. 2008;**24**:574–594. https://doi.org/10.1080/10447310802205776
18. Lewis JR. The system usability scale: past, present, and future. *Int J Human–Computer Interact*. 2018;**34**:577–590. https://doi.org/10.1080/10447318.2018.1455307
19. Menghini L, Cellini N, Goldstone A, Baker FC, 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**(2). https://doi.org/10.1093/sleep/zsaa170
20. Liang J, Xian D, Liu X, *et al*. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. *JMIR Mhealth Uhealth*. 2018;**6**(11):e11066. https://doi.org/10.2196/11066
21. Marquard JL, Howard J, LeBlanc R. Using novel data visualization methods to understand mobile health usability: exemplar from a technology-enabled sleep self-monitoring intervention. *CIN Comput Inform Nurs*. 2023;**41**(1):1–5. https://doi.org/10.1097/CIN.0000000000000970

© The Author(s) 2026. Published by Oxford University Press on behalf of Sleep Research Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercialNoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.
SLEEP Advances, 2026, **7**, 1–9
https://doi.org/10.1093/sleepadvances/zpag001
Original Article