Digital Biomarkers

Tools and Devices

Digit Biomark 2026;10:113–124
DOI: 10.1159/000551257

Received: August 15, 2025
Accepted: February 22, 2026
Published online: March 24, 2026

Methods for Wearable-Derived Pulse Rate Measures and Their Application to Modeling the Relationship between Pulse Rate and Motor Activity in Narcolepsy Type 1

Eena Kosik-Rose[a, b] Raul Torres[b] Brian Tracey[b] Tina Olsson[b] Dmitri Volfson[b] Francesco Onorati[b] Marta Karas[b]

[a]Department of Cognitive Sciences, University of California, San Diego, CA, USA; [b]Takeda Development Center Americas, Inc., Cambridge, MA, USA

Keywords
Narcolepsy · Wearable device · Pulse rate · Heart rate · Motor activity

Abstract
Introduction: Narcolepsy type 1 (NT1) is a chronic, rare neurological disorder caused by loss of orexin neurons that disrupts the sleep-wake cycle. Current research suggests a link between orexin neuropeptide deficiency and heart rate regulation. Wrist-worn devices enable continuous monitoring of physiological data, including heart rate. We aim to effectively utilize wearables in clinical research by developing robust methodologies to address challenges including data artifacts and missing data. ***Methods:*** Empatica wrist-worn smartwatch data were collected for 2 weeks at baseline in the TAK-861-2001 randomized, phase 2 clinical trial (NCT05687903) from participants with NT1. Epoch-level (minute-level) pulse rate data were rigorously screened for artifacts. Daily pulse rate measures (24-h, daytime, and nighttime averages) were derived. Numerical experiments were conducted to evaluate the stability of daily measures as a function of the proportion of valid minute-level data within a given day. Mean pulse rate and motor activity daily measures and the association between the two were quantified using linear mixed-effects models. ***Results:*** Participants (n = 110) had a mean age of 34.0 years and 51.8% were female. Approximately 23.5% of the epoch-level pulse rate observations were invalid due to non-wear, missing data, or artifacts. With a ≥70% threshold for valid daily wearable data, participants contributed an average of 10.3 days of data. Mean pulse rates were 72.8, 83.3, and 66.2 beats per minute and mean motor activities were 40.6, 61.5, and 10.3 epoch-level activity counts for 24-h, daytime, and nighttime averages, respectively. Pulse rate was significantly positively associated with motor activity daily measures, with up to 80% of pulse rate variance explained by motor activity levels and individual differences in baseline pulse rate and pulse rate-motor activity association. ***Conclusion:*** This work proposes a robust framework for preprocessing smartwatch pulse rate data for analyses in clinical research. The findings from our smartwatch data demonstrate a significant positive association between pulse rate and motor activity in participants with NT1 during a 2-week baseline period.

Francesco Onorati and Marta Karas shared senior co-authorship.

karger@karger.com
www.karger.com/dib

Correspondence to:
Marta Karas, marta.karas@takeda.com

Introduction

Narcolepsy type 1 (NT1) is a chronic, rare neurological disorder that disrupts the natural sleep-wake cycle, leading to symptoms like excessive daytime sleepiness, cataplexy, cognitive difficulties, and sleep-related symptoms (sleep paralysis, hallucinations, and disrupted nighttime sleep) [1–4]. NT1 impacts individuals' quality of life and daily functioning [5]. The orexin neuropeptide is involved in autonomic regulation, especially sympathetic tone and thermoregulation, where low orexin levels are considered the pathophysiological feature of NT1 [6]. Most evidence suggests that an autoimmune mechanism targets the orexin neurons, leading to their loss, ultimately driving NT1 symptoms [2, 7].

Current research suggests a link between orexin neuropeptide deficiency and heart rate regulation. A study in orexin knockout mice reported higher heart rates during rest time versus controls [8]. In humans, individuals with NT1 showed higher sleeping heart rates than matched controls [9, 10]. Therefore, monitoring daytime and nighttime heart rate may support the development of physiological markers of autonomic function to better characterize the NT1 phenotype.

Digital tools, such as wearable sensors, present promising avenues for investigating phenotypes in various patient populations, including those with NT1 [11–13]. These devices allow for continuous and high-resolution data collection in natural settings, capturing physiological and behavioral changes that might be missed during infrequent clinic visits [14]. Modern wrist-worn smartwatches are commonly equipped with accelerometer and photoplethysmography (PPG) sensors, among others, to track motor activity and pulse rate at different time resolutions, including minute and day levels [11]. With sleek designs and recent battery improvements, smartwatches are well positioned for weeks-long unobtrusive monitoring in large samples [15]. Throughout this paper, "motor activity" refers to the wearable-derived quantification of bodily acceleration measured by the device's accelerometer and is considered a proxy for physical activity volume, where physical activity is considered as bodily movement that results in energy expenditure above resting levels, encompassing activities of daily living and exercise. Additionally, "pulse rate" is used to describe the estimate derived from device's PPG data, as it more accurately describes the mechanical pulses detected at the wrist, rather than "heart rate," which refers to the electrical activity of the heart measured through methods such as electrocardiography [16].

However, wrist-worn device data in clinical research present analytical challenges, including missing data at the sub-day level (an interval during the day) and uncertainty about whether data points represent true measurements or artifacts [17]. Indeed, missing data may occur due to device malfunction, data transfer error, or a participant failing to wear the device [18]. In large epidemiological studies with 24-h continuous wear protocols for wrist-worn devices, varying thresholds for the minimum amount of valid accelerometer data per day have been used, including 95% [19], 90% [20], or 67% [21] of wear time per day. Guidance on handling missing wearable pulse rate data remains limited; however, some works applied a threshold of 1,000 min (69%) [22] or 1,008 min (70%) [23] of wear time per day.

The association between physical activity and pulse rate is well established [24, 25], with an episode of higher activity generally leading to a pulse rate increase within a short timeframe. At the day level, individuals may therefore show higher average daily pulse rates on days with higher overall physical activity. This highlights the importance of accounting for physical activity when analyzing long-term pulse rate data in clinical trials [26], especially when increased activity may be a part of the intended therapeutic effect. Without such considerations, true increases in average physical activity could lead to corresponding increases in pulse rate that might be misattributed to other factors, such as a direct effect of the investigational treatment, creating misleading conclusions.

This study aims to investigate the relationship between wearable-derived pulse rate and motor activity daily measures in participants with NT1 during a 2-week baseline period. To achieve this, we address two key challenges in wearable-derived pulse rate data analyses: (1) identifying physiologically implausible minute-level pulse rate values and (2) defining the minimum proportion of valid minute-level pulse rate data required for a day's inclusion in downstream analyses. These steps are critical in wearable data-based research, where artifacts, data loss, and inconsistent wear patterns can substantially bias physiological estimates.

Methods

Study Design and Population

Data used in this analysis were collected during the baseline period in the randomized, placebo-controlled, phase 2 clinical trial evaluating the efficacy and safety of TAK-861, an orexin receptor 2 agonist, in participants

with NT1 (NCT05687903) [27]. Further details on inclusion criteria and study design are provided in the online supplementary methods (for all online suppl. material, see https://doi.org/10.1159/000551257). Trial results are reported by Dauvilliers et al. [28]. The CONSORT reporting checklist is in online supplementary materials. This work uses only baseline data; no treatment period data or information on specific treatment arms or doses were included in the analysis.

Digital Technology and Data

Participants used Empatica's EmbracePlus wearable device (Empatica, Boston, MA, USA) to continuously record physiological and behavioral signals during the 14-day baseline period until 1 day before treatment initiation, and throughout the 8-week treatment period. Only wearable device data collected during baseline were used. Participants were instructed to wear the device continuously on their nondominant wrist, with ≥1 h of charging every other day. The vendor provided the sponsor with measures aggregated per nonoverlapping 1-min window ("epochs"). Epoch-level measures included activity counts (unitless measure of frequency and intensity of movements), device wear percentage (estimated percentage of time the device was worn in the epoch), and pulse rate (estimated number of heart beats in the epoch, expressed in beats per minute [bpm]).

Further device details are provided in the online supplementary methods. We refer to Empatica's EmbracePlus wearable device as "wearable device."

Digital Endpoints Derivation

Data Preprocessing

Our preprocessing aimed to identify and exclude artifactual or missing pulse rate estimates that do not reflect the participant's true physiological state, which we refer to as "invalid" data. In the first step, we excluded epochs with a pulse rate of 0, missing pulse rate values, or a device wear percentage <50% in the epoch. According to the vendor's documentation, possible reasons for missing pulse rate measures include device not worn or not worn correctly (off-wrist), device out of battery or charging, or low-quality PPG signal due to other reasons. This step excluded ~22% of the epoch-level observations across all participant-days at baseline.

Next, we excluded epochs with pulse rate values deemed physiologically implausible. These values were identified by examining the joint distribution of pulse rate and activity counts values. We constructed a two-dimensional grid with a bin size of 10 bpm for pulse rate and 10 units for activity counts. The grid served as a descriptive visualization tool to identify sparsely populated regions representing unlikely physiological combinations (e.g., very high pulse rate concurrent with minimal activity). We compared grids using bin sizes of 5, 10, and 20 units and selected the 10-unit resolution as the most interpretable balance between resolution and visual clarity. Regions showing abrupt declines in data density (approximately 20–50% relative to neighboring bins), consistent with nonphysiological patterns, informed the exclusion criteria definition for outlier removal. Thresholds were chosen to remove anomalies that were unlikely to represent valid biological phenomena. Concurrently, the selection process aimed to avoid introducing bias by selectively focusing on values close to an average or expected behaviors, thereby preserving the representativeness of the dataset. Specifically, values were excluded if the (1) pulse rate was ≥220 bpm, irrespective of activity counts, (2) pulse rate was 190–219 bpm and activity counts were <100, or (3) pulse rate was <40 bpm and activity counts were >30. The two-dimensional table with observation counts is visualized in online supplementary Figure S1, which excluded ~0.15% of the remaining epoch-level pulse rate observations. High pulse rates with low activity counts were typically found within long periods where the device recorded the maximum recordable pulse rate values, indicating a high-frequency oscillation artifact on the PPG signal during events when the device was not worn, which summed to ~0.03% of the data. In contrast, low pulse rates with high activity counts were rare and typically consisted of isolated epochs, indicating occasional errors of the pulse rate algorithm, which summed to ~0.04% of the data. The final 0.08% was excluded due to implausibly high pulse rate values, ≥220 bpm. Epoch-level activity count data were preprocessed to keep only observations from the epochs where pulse rate data were retained.

Daily Measures of Pulse Rate and Motor Activity

For each participant, three daily pulse rate measures from wearable data were calculated: 24-h (00:00 to 00:00), daytime (09:00 to 21:00), and nighttime (01:00 to 06:00) averages. Each measure was obtained by averaging the epoch-level pulse rates within the specified time windows. We also derived three motor activity measures using epoch-level activity counts: 24-h, daytime, and nighttime averages, calculated within the same time ranges as for pulse rate. Daytime and nighttime windows were defined following

the suggestion in O'Brien et al. [29] to exclude "retiring" and "rising" periods, chosen here as 21:01 to 00:59 and 06:01 to 08:59, respectively, thereby reducing the impact of age and lifestyle-related differences in sleep and wake times. Epoch-level observations excluded during preprocessing did not contribute to daily measure calculations. This dataset included ≤14 observations of each measure per participant, corresponding to the 14-day baseline period.

Statistical Data Analysis

The analytic sample consisted of all enrolled participants who contributed any wearable device data at baseline; only baseline data were analyzed. Numerical experiments were conducted to evaluate the stability of daily pulse rate measures as a function of the proportion of valid minute-level pulse rate data within a day. The analysis used all participant-days with at least 95% valid pulse rate data from the analytic sample. For each analysis window (24-h, daytime, nighttime) and each simulated level of invalid data (5%–95%, in 5% increments), we performed a two-stage bootstrap with 500 repetitions to capture both between- and within-participant variability. We defined the simulation grid in terms of the percentage of valid daily data, rather than fixed epoch counts, to improve generalizability across studies with potentially differing epoch lengths. Both the use of percentage-based thresholds and their coarse granularity are consistent with prior actigraphy literature, in which valid daily data time criteria are commonly specified using broad cutoffs, such as >95% (approximately <1.2 h invalid [19, 30]) or >90% (<2.4 h invalid [20]). In each bootstrap iteration, participants were sampled with replacement, and for each participant, 1 day was resampled with replacement. Within every resampled day, a contiguous block of minutes corresponding to the desired proportion of invalid data was randomly removed. The average pulse rate was then recalculated, and its deviation from the complete-data average (i.e., an error) was recorded. For each bootstrap repetition, we then computed metrics comparing pulse rate averages derived from complete data with those obtained from simulated incomplete data: error-based metrics, including the root mean square error (RMSE), mean absolute error, mean error (ME), standard deviation (SD) of the absolute error, and SD of the error, as well as agreement-based metrics, including Pearson's correlation coefficient (r), the concordance correlation coefficient, and the intraclass correlation coefficient (ICC [1, 3]; two-way mixed-effects model for absolute agreement of single measurements). These metrics were then aggregated across bootstrap repetitions using the median and interquartile range (IQR).

Based on the simulation results and prior pulse rate algorithm validation work [31], which used acceptance criteria of accuracy root-mean-square (calculated as square root of the mean of the squared deviations) ≤3 bpm under no-motion and ≤5 bpm under motion conditions for FDA-cleared wrist-worn devices, we selected a threshold for the minimum proportion of valid minute-level pulse rate data within a day required for inclusion of the day in downstream analyses. The threshold was chosen to cap the median RMSE at the lower end of the 3–5 bpm range for 24-h and daytime average pulse rate measures, while maximizing the number of participant-days retained. Ultimately, in our downstream analyses, a single threshold was applied across all three analysis windows (24-h, daytime, and nighttime) to maintain consistency and simplify interpretation. The threshold was applied at the day level, meaning that each participant-day pulse rate data was evaluated independently.

The following statistical analyses were conducted using only participant-days that met the established threshold. We estimated population means of daily pulse rate and motor activity measures at baseline, and the association between the two, using linear mixed-effects models, conducted in R (version 4.3.1). All models included a participant-level random intercept, and the model evaluating the association between pulse rate and motor activity additionally included a random slope. For nighttime average pulse rate and nighttime average motor activity, as part of a sensitivity analysis, we additionally fitted models to datasets in which values of nighttime average motor activity above the 95th and 80th percentiles, respectively, were removed. All other analyses were conducted in Python (version 3.12.4). Further details are provided in the online supplementary methods.

Results

Participant Characteristics

The study enrolled 112 participants with NT1. Our analytic sample consisted of 110 participants with wearable device data recorded at baseline. Among these, 101 had at least 1 day with 95% or more valid minute-level

Table 1. Participant demographic and clinical characteristics at baseline

Characteristic	Analytic sample (n = 110)
Age, mean (SD), years	34.0 (11.5)
BMI (SD)	27.0 (4.6)
Sex, *n* (%)	
Female	57 (51.8)
Male	53 (48.2)
Ethnicity, *n* (%)	
Not Hispanic or Latino	103 (93.6)
Hispanic or Latino	6 (5.6)
Unknown or not reported	1 (0.9)
Race, *n* (%)	
White	94 (85.5)
Asian	8 (7.3)
Black	6 (5.5)
Multiple or not reported	2 (1.8)
Average sleep latency on MWT, mean (SD), min[a]	4.5 (6.2)
Epworth sleepiness scale score, mean (SD)	18.5 (3.0)
Weekly cataplexy rate, mean (SD)	22.2 (24.1)

BMI, body mass index; MWT, Maintenance of Wakefulness Test. [a]Data not available for one participant.

pulse rate data. Baseline demographic and clinical characteristics of the analytic sample are shown in Table 1. These clinical characteristics indicate substantial difficulty maintaining wakefulness as measured through the Maintenance of Wakefulness Test, self-reported excessive daytime sleepiness, and high frequency of cataplexy episodes, consistent with the NT1 phenotype [28, 32].

Assessing Stability of Daily Pulse Rate Measures as a Function of Invalid Data

Bootstrapped wearable pulse rate measures showed consistent patterns across the 24-h, daytime, and nighttime windows (Fig. 1). As the proportion of invalid data increased, error-based metrics – the RMSE, mean absolute error, and SD of the absolute error between the reference pulse rate (from the complete window) and the pulse rate estimated from windows with simulated data loss – grew progressively larger (Fig. 1a, b, c), while agreement-based metrics, such as Pearson's correlation coefficient, decreased (Fig. 1d). A ≥70% threshold for valid minute-level pulse rate data within a day was selected to maximize the number of participant-days retained while keeping the median RMSE at the lower end of the 3–5 bpm range for 24-h and daytime average pulse rate measures; the choice of 70% also reflects the 5% grid used in the simulations, enabling simple thresholding. At 70% valid data (30% invalid data), we observed median RMSE (IQR) of 3.51 (0.18, 0.55), 2.61 (0.21, 0.64), and 1.56 (0.14, 0.43), for 24-h, daytime, and nighttime average pulse rates, respectively. This threshold also yielded median correlation (IQR) of r = 0.92 (0.01, 0.04), r = 0.96 (0.01, 0.02), and r = 0.99 (0.0, 0.01), for 24-h, daytime, and nighttime average pulse rates, respectively. Additional results including ME, SD of ME, ICC, and concordance correlation coefficient are reported in online supplementary Figure S2, depicting similar trends.

Wearable-Derived Pulse Rate and Motor Activity Baseline Measures

With the 70% threshold for valid daily wearable data, 1,155 (68.8%) of 1,680 participant-day observations were retained from 104 (94.5%) participants. Each participant contributed an average of 10.3 observations (range, 1–14) from the baseline period. At baseline, the population-level mean (95% CI) average pulse rates were 72.8 (71.5–74.0), 83.3 (81.9–84.6), and 66.2 (64.6–67.9) bpm for 24-h, daytime, and nighttime, respectively (Table 2). Population-level mean (95% CI) average motor activities were 40.6 (38.7–42.4), 61.5 (58.6–64.5), and 10.3 (9.1–11.5) epoch-level activity counts for 24-h,

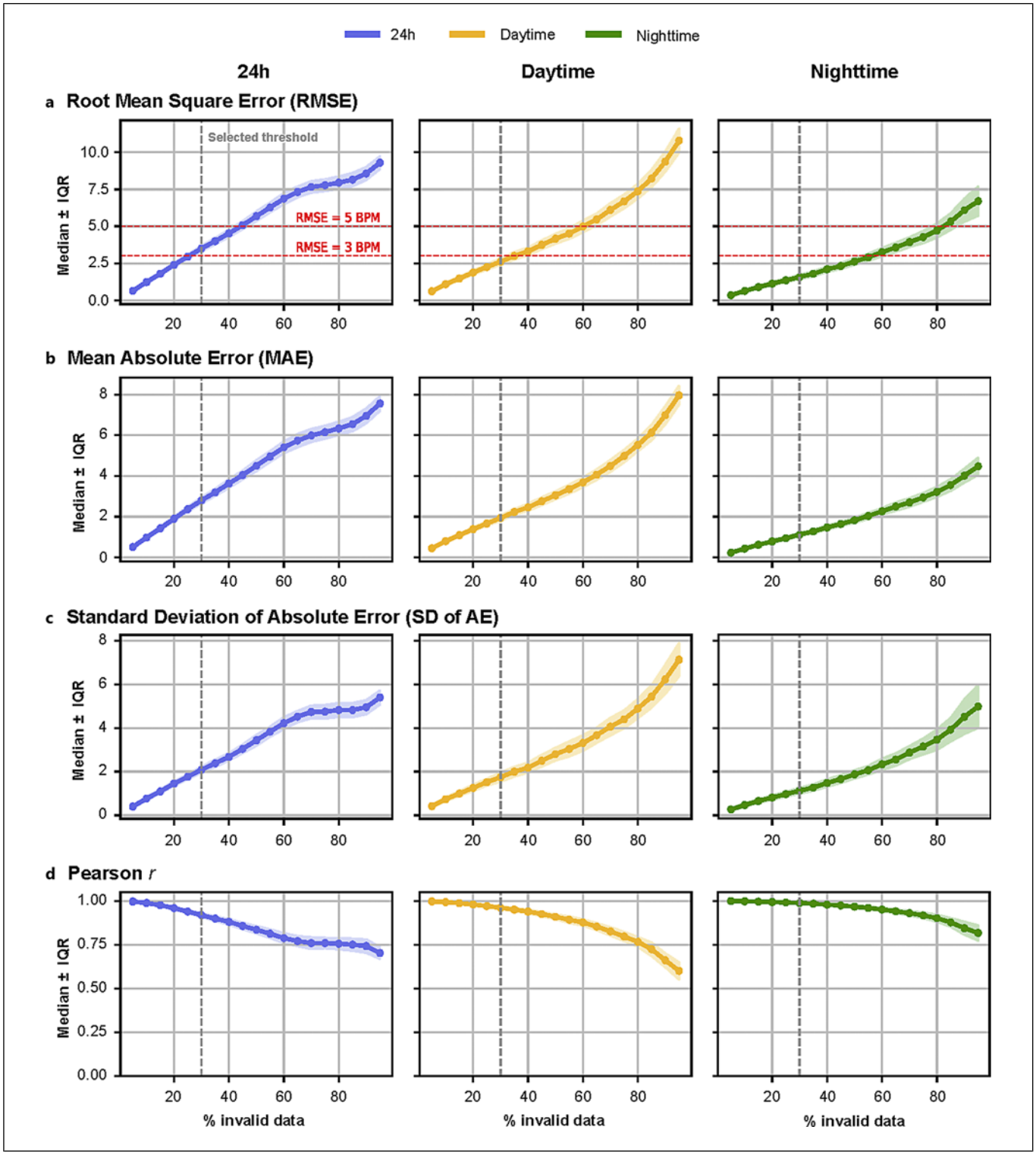

Fig. 1. Error-based and agreement-based metrics comparing the reference pulse rate (from the complete window) with the pulse rate estimated from windows with simulated data loss across analysis windows: 24-h (blue), daytime (yellow), and nighttime (green), shown as a function of the percentage of invalid data (x-axis). **a** Root mean square error (RMSE). **b** Mean absolute error (MAE). **c** SD of the absolute error. **d** Pearson's correlation coefficient (*r*). Points represent median values, and shaded regions indicate interquartile ranges (IQRs) of metric estimates across bootstrap repetitions.

Table 2. Population means of daily pulse rate and motor activity measures derived from wearable device data at baseline

No.	Model outcome, daily measure	Intercept coefficient estimate, population mean (95% CI)	Random intercept SD	Residual error SD
1	24-h average pulse rate, bpm	72.8 (71.5–74.0)	6.3	6.2
2	Daytime average pulse rate, bpm	83.3 (81.9–84.6)	6.4	7.1
3	Nighttime average pulse rate, bpm	66.2 (64.6–67.9)	8.3	5.9
4	24-h average motor activity, AC	40.6 (38.7–42.4)	8.6	10.2
5	Daytime average motor activity, AC	61.5 (58.6–64.5)	14.3	16.3
6	Nighttime average motor activity, AC	10.3 (9.1–11.5)	5.3	6.2

AC, epoch-level activity counts; bpm, beats per minute in the epoch. Daytime was 09:00 to 21:00; nighttime was 01:00 to 06:00. Means were estimated using linear mixed-effects models. SDs of the random intercept and residual error are reported.

daytime, and nighttime, respectively (Table 2). Overall, we observed similarity between SDs of the random intercepts and residual errors across measures, suggesting that both between-participant differences and within-participant day-to-day fluctuations contribute to the overall variability in baseline daily measures.

Association between Wearable-Derived Pulse Rate and Motor Activity Measures at Baseline

Table 3 presents the results of three linear mixed-effects models, each examining the association between daily measures of pulse rate and motor activity at different aggregation windows (24-h, daytime, and nighttime). All models show significant positive relationships between pulse rate and motor activity measures, with the largest magnitude of association observed during the nighttime period (effect estimate = 0.41; $p < 0.001$). Participant-level data and population-level trends are visualized in Figure 2. These demonstrate substantial variability in both participant-level daily pulse rate measures and the direction of the slope of the association between daily pulse rate and motor activity measures. A wide range of motor activity was observed even within nighttime, with higher values potentially reflecting the participant's wakefulness during this period on a given day. This is consistent with prior evidence that heart rate varies by sleep stage, generally higher during wakefulness and rapid eye movement (REM) sleep than during non-REM sleep [33]. Online supplementary Figure S3 illustrates epoch-level pulse rate and heart rate data from one participant over 2 days, comparing periods of relatively low and high nighttime motor activity. R^2c, which represents the proportion of variance in the outcome, namely, pulse rate, explained by both fixed and random effects, ranged from 0.76 for 24-h average to 0.80 for nighttime average measures (Table 3). ICCs computed from the random-effects variance components indicated good reliability of wearable-derived pulse rate measures across days (Table 3). The direction and magnitude of the association between nighttime pulse rate and motor activity remained consistent across sensitivity analyses in which values of nighttime average motor activity above the 95th and 80th percentiles, respectively, were removed (online suppl. Table S1).

Discussion

This study investigated the relationship between daily pulse rate and motor activity using a wrist-worn device in a large sample of participants with NT1 during a 2-week clinical trial baseline period. A significant positive association was found, with 76–80% of the variance in 24-h, daytime, and nighttime average pulse rates being explained by motor activity together with individual differences in both typical baseline pulse rate and the magnitude of each participant's pulse rate association with motor activity. These relatively high R^2c values suggest the models fit the data well and that motor activity is a strong predictor of pulse rate at the participant level. Furthermore, this emphasizes the need to account for

Table 3. Association between daily measures of pulse rate and motor activity derived from wearable device data at baseline

No.	Model outcome	Model covariate	Intercept coefficient estimate (95% CI)	Slope coefficient estimate (95% CI); *p* value	Random intercept SD	Random slope SD	Residual error SD	R^2c	ICC
1	24-h average pulse rate, bpm	24-h average motor activity, AC	57.1 (55.3–59.0)	0.39 (0.35–0.42); $p < 0.001$	6.7	0.10	4.6	0.76	0.68
2	Daytime average pulse rate, bpm	Daytime average motor activity, AC	65.1 (63.3–67.0)	0.30 (0.27–0.32); $p < 0.001$	6.8	0.09	4.9	0.77	0.66
3	Nighttime average pulse rate, bpm	Nighttime average motor activity, AC	62.2 (60.5–63.9)	0.41 (0.35–0.47); $p < 0.001$	8.4	0.17	4.6	0.80	0.77

AC, epoch-level activity counts; bpm, beats per minute in the epoch; ICC, intraclass correlation coefficient; R^2c, linear mixed-effects model conditional coefficient of determination. Daytime was 09:00 to 21:00; nighttime was 01:00 to 06:00. Effects were estimated using linear mixed-effects models. *p* values for fixed effects were obtained using *t* tests with Satterthwaite's approximation for degrees of freedom, as implemented in the lmerTest R package. SDs of the random intercept, random slope, and residual error are reported. Model's R^2c is reported.

motor activity as a potential confounder when interpreting changes in average daily pulse rates in NT1 clinical trials, particularly when the desired therapeutic effect may induce changes in physical activity (e.g., movements following the nighttime awakenings). Without such consideration, motor activity could confound longitudinal analyses, making it difficult to determine whether pulse rate changes are potentially due to treatment or simply variations in physical activity.

We propose a rigorous framework for preprocessing smartwatch pulse rate data in clinical research. Data-driven approaches were used to identify and exclude epoch-level data artifacts. Pulse rate data from a full analysis window compared to windows with varying levels of invalid data were used to determine the threshold for minimum valid daily wearable data required for statistical analysis. From that, we selected ≥70% valid data (≤30% invalid data) threshold. With this threshold, the median RMSE remained ≤3.51 bpm and Pearson's correlation ≥0.92 across all three analysis windows, in line with prior pulse rate algorithm validation work [31] that used acceptance criteria of accuracy root-mean-square ≤3 BPM under no-motion and ≤5 bpm under motion conditions for FDA-cleared wrist-worn devices (Empatica validation study [31]). Although separate thresholds could be derived for daytime, nighttime, and 24-h analyses, we selected a single relatively conservative threshold to maintain consistency across endpoints and simplify interpretation. The threshold value we selected (70%) aligns with the 1,000 min of wear time per day (69%) threshold previously applied by Radin et al. [22]. Our analysis focused on identifying a threshold for the proportion of valid data within a single day. Alternative or complementary approaches could be considered to address invalid wearable pulse rate data, such as participant- and minute-specific imputation strategies similar to those implemented in published minute-level actigraphy preprocessing frameworks (e.g., the arctools R package) [34].

Overall, our framework employed integrated data summaries, analysis of trends among outliers, subject matter expert consultation, and detailed reporting of data filtering, aligning with recommendations for quantitative reporting on adherence data [35]. These steps aim to enhance the reliability of analysis results, ensure transparency in reporting, and foster scientific trust.

The data and methods presented here contribute to the growing body of work paving the way for the use of wearable-derived measures in drug development. For a wearable-derived endpoint to be included in a marketing authorization package or label extension, regulatory acceptance would typically require analytical and clinical validation, clearly defined context of use, and proactive engagement with regulatory agencies to seek feedback

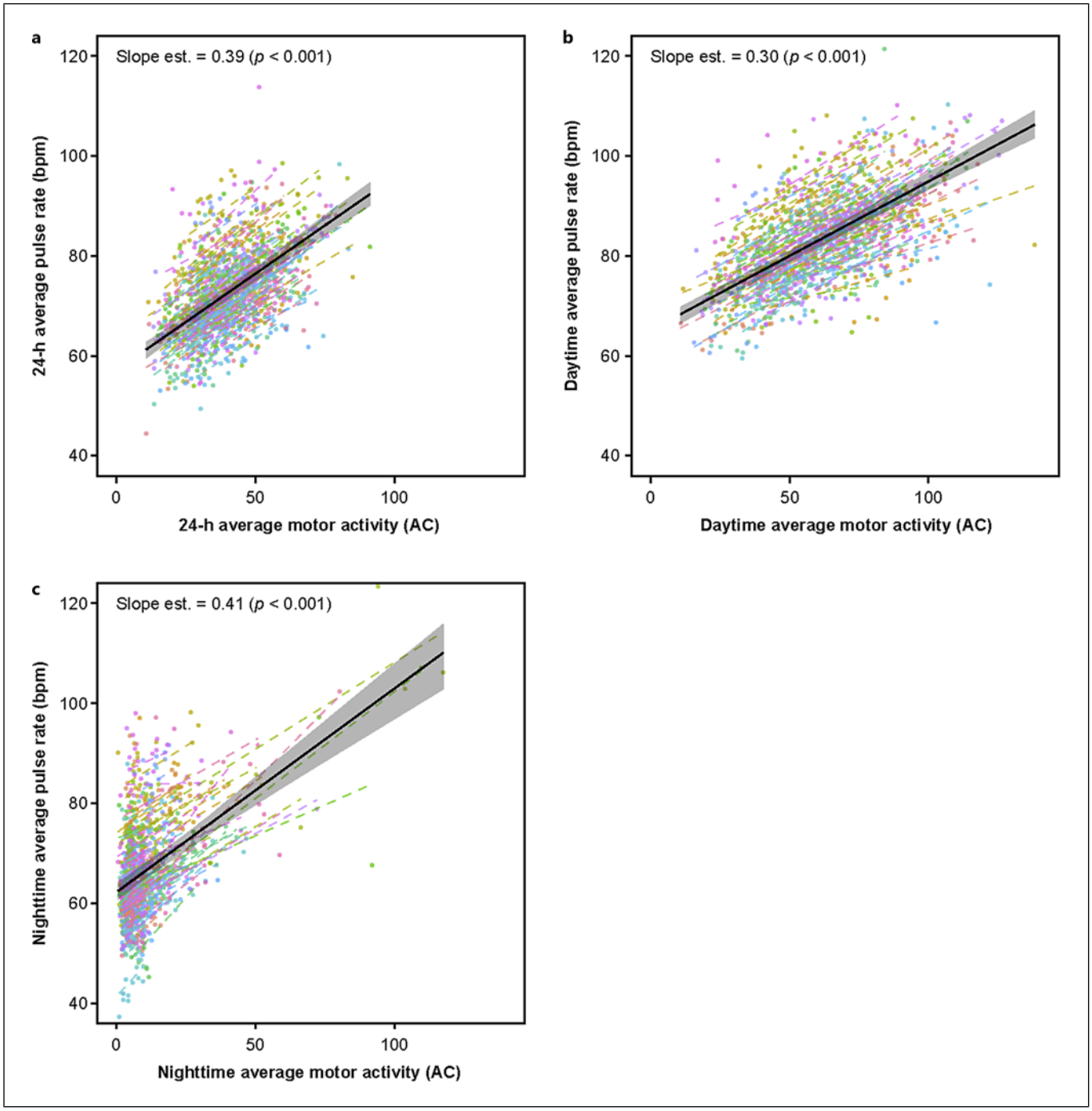

Fig. 2. Association between daily motor activity (x-axes) and pulse rate (y-axes) measures at baseline for the three aggregation windows. 24-h (**a**), daytime (**b**), and nighttime (**c**) associations were estimated using linear mixed-effects models. Data points represent observed values, and lines indicate participants' conditional means, both color-coded by participant consistently across panels. Black lines show the population-level regression fit, with gray bands representing 95% CIs. AC, epoch-level activity counts; bpm, beats per minute in the epoch. *p* values for fixed effects were obtained using *t* tests with Satterthwaite's approximation for degrees of freedom, as implemented in the lmerTest R package.

before inclusion in a marketing application [14]. While these requirements are beyond the scope of the present study, our work may inform future validation efforts aimed at qualification of wearable-derived pulse rate measures as potential primary or secondary endpoints.

A limitation of our preprocessing framework is that day-level measures were only included if each pulse rate measure (24-h, daytime, and nighttime average) had less than 30% invalid wearable data. Depending on the specific research objectives and characteristics of the study population, researchers might opt for having separate thresholds for 24-h, daytime, and nighttime averages to accommodate potential variations in wear patterns across the day. Such variations may arise due to factors such as comfort, aesthetics, or lifestyle preferences in the specific population. There is also a limitation of our minute-level motor activity measure (activity counts), which provides a high-level, wearable-derived proxy of physical activity volume based on body acceleration, but does not distinguish between specific activity modalities (e.g., ambulation, exercise, fidgeting) or behavioral context (e.g., purposeful activity vs. restlessness). Therefore, the present findings cannot characterize which specific motor activities contribute to the observed variance in pulse rate. Another limitation is that this work focuses solely on data from individuals with NT1. Although the relationship between pulse rate and motor activity is physiologically fundamental and expected to be present in other populations, these specific findings, including the strength of the association, may not be directly generalizable to other patient groups or healthy individuals. Finally, we acknowledge the potential impact of skin pigmentation on the accuracy of PPG-derived measures such as pulse rate, as melanin absorption of green light can weaken pulsatile signals used by the device algorithms [31]. Although skin pigmentation was not recorded in our study, 14.5% of participants self-reported a race other than White. Importantly, the EmbracePlus device and its pulse rate estimation algorithm have been validated for consistent performance across diverse demographic and physiological variables, including sex, age, skin tone, and body mass index, supporting the robustness of pulse rate estimates in our cohort [31].

In conclusion, wearable-derived pulse rate was significantly positively associated with motor activity in participants with NT1 during a 2-week baseline period, with up to 80% of pulse rate variance explained by motor activity levels and individual differences in baseline pulse rate and pulse rate-motor activity association. This work proposes a framework for utilizing smartwatch pulse rate data for analyses in clinical research to support evaluating the effects of investigative treatments.

Acknowledgments

Editorial assistance in formatting, proofreading, and copy editing was provided by Envision Catalyst, an Envision Medical Communications agency, a part of Envision Pharma Group.

Statement of Ethics

The TAK-861-2001 clinical trial protocol was reviewed and approved by the institutional review boards or ethics committees at each participating site. The full list of participating site and ethics committees can be found in Dauvilliers et al. [28]. The trial adhered to International Council for Harmonisation guidelines, the ethical principles specified in the Declaration of Helsinki, and all relevant local or regional regulatory requirements. Informed written consent was obtained from all participants.

Conflict of Interest Statement

All authors are or were employees of Takeda Development Center Americas, Inc., and stockholders in Takeda Pharmaceutical Company Limited at the time of the study.

Funding Sources

This work and clinical trial were funded by Takeda Development Center Americas, Inc. Takeda Development Center Americas, Inc., provided funding to Envision Catalyst for support in editing this paper. The funder was involved in the study design, analysis, and interpretation of results, and participated in the preparation, review and approval of the manuscript, and the decision to submit for publication.

Author Contributions

E.K-R., F.O., and M.K. contributed to the conceptualization, methodology, formal analysis, and writing the original draft. R.T. contributed to the methodology. B.T. and D.V. contributed to the conceptualization and methodology. E.K.-R., R.T., B.T., T.O., D.V., F.O., and M.K. reviewed and edited the manuscript and approved the final version for publication.

Data Availability Statement

The data that support the findings of this study are not publicly available due to sensitive participant information but are available from the corresponding author [M.K.] upon reasonable request.

References

1 Bassetti CLA, Adamantidis A, Burdakov D, Han F, Gay S, Kallweit U, et al. Narcolepsy - clinical spectrum, aetiopathophysiology, diagnosis and treatment. Nat Rev Neurol. 2019;15(9):519–39. https://doi.org/10.1038/s41582-019-0226-9

2 Dauvilliers Y, Arnulf I, Mignot E. Narcolepsy with cataplexy. Lancet. 2007;369(9560):499–511. https://doi.org/10.1016/S0140-6736(07)60237-2

3 Scammell TE. Narcolepsy. N Engl J Med. 2015;373(27):2654–62. https://doi.org/10.1056/NEJMra1500587

4 Cano CA, Harel BT, Scammell TE. Impaired cognition in narcolepsy: clinical and neurobiological perspectives. Sleep. 2024;47(9):zsae150. https://doi.org/10.1093/sleep/zsae150

5 Lividini A, Pizza F, Filardi M, Vandi S, Ingravallo F, Antelmi E, et al. Narcolepsy type 1 features across the life span: age impact on clinical and polysomnographic phenotype. J Clin Sleep Med. 2021;17(7):1363–70. https://doi.org/10.5664/jcsm.9198

6 Kukkonen JP, Holmqvist T, Ammoun S, Akerman KE. Functions of the orexinergic/hypocretinergic system. Am J Physiol Cell Physiol. 2002;283(6):C1567–91. https://doi.org/10.1152/ajpcell.00055.2002

7 Mignot E. Narcolepsy: genetics, immunology, and pathophysiology. In: Kyger M, Roth T, Dement W, editors. Principles and practice in sleep medicine. 6th ed. Elsevier; 2017. p. 855–72.

8 Bastianini S, Silvani A, Berteotti C, Elghozi JL, Franzini C, Lenzi P, et al. Sleep related changes in blood pressure in hypocretin-deficient narcoleptic mice. Sleep. 2011;34(2):213–8. https://doi.org/10.1093/sleep/34.2.213

9 Bosco A, Lopez R, Barateau L, Chenini S, Pesenti C, Pepin JL, et al. Effect of psychostimulants on blood pressure profile and endothelial function in narcolepsy. Neurology. 2018;90(6):e479–91. https://doi.org/10.1212/WNL.0000000000004911

10 Grimaldi D, Calandra-Buonaura G, Provini F, Agati P, Pierangeli G, Franceschini C, et al. Abnormal sleep-cardiovascular system interaction in narcolepsy with cataplexy: effects of hypocretin deficiency in humans. Sleep. 2012;35(4):519–28. https://doi.org/10.5665/sleep.1738

11 Gnarra O, van der Meer J, Warncke JD, Fregolente LG, Wenz E, Zub K, et al. The Swiss Primary Hypersomnolence and Narcolepsy Cohort Study: feasibility of long-term monitoring with Fitbit smartwatches in central disorders of hypersomnolence and extraction of digital biomarkers in narcolepsy. Sleep. 2024;47(9):zsae083. https://doi.org/10.1093/sleep/zsae083

12 Tracey B, Culp M, Fabregas S, Mignot E, Buhl DL, Volfson D. Novel biomarkers derived from the Maintenance of Wakefulness Test as predictors of sleepiness and response to treatment. Sleep. 2024;47(12):zsae148. https://doi.org/10.1093/sleep/zsae148

13 Vilela M, Tracey B, Volfson D, Barateau L, Cai A, Buhl DL, et al. Identifying time-resolved features of nocturnal sleep characteristics of narcolepsy using machine learning. J Sleep Res. 2024;33(6):e14216. https://doi.org/10.1111/jsr.14216

14 Di J, Karas M, Vlanjnic V. Novel digital wearable sensors for drug development in pharmaceutical industry. Current developments in biosensor applications and smart strategies. London, UK: IntechOpen; 2024.

15 Karas M, Bai J, Strączkiewicz M, Harezlak J, Glynn NW, Harris T, et al. Accelerometry data in health research: challenges and opportunities. Stat Biosci. 2019;11(2):210–37. https://doi.org/10.1007/s12561-018-9227-2

16 Schafer A, Vagedes J. How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. Int J Cardiol. 2013;166(1):15–29. https://doi.org/10.1016/j.ijcard.2012.03.119

17 Tackney MS, Cook DG, Stahl D, Ismail K, Williamson E, Carpenter J. A framework for handling missing accelerometer outcome data in trials. Trials. 2021;22(1):379. https://doi.org/10.1186/s13063-021-05284-8

18 Di J, Demanuele C, Kettermann A, Karahanoglu FI, Cappelleri JC, Potter A, et al. Considerations to address missing data when deriving clinical trial endpoints from digital health technologies. Contemp Clin Trials. 2022;113:106661. https://doi.org/10.1016/j.cct.2021.106661

19 Meng Q, Cui E, Leroux A, Mowry EM, Lindquist MA, Crainiceanu CM. Quantifying the association between objectively measured physical activity and multiple sclerosis in the UK Biobank. Med Sci Sports Exerc. 2023;55(12):2194–202. https://doi.org/10.1249/MSS.0000000000003260

20 Karas M, Muschelli J, Leroux A, Urbanek JK, Wanigatunga AA, Bai J, et al. Comparison of accelerometry-based measures of physical activity: retrospective observational data analysis study. JMIR Mhealth Uhealth. 2022;10(7):e38077. https://doi.org/10.2196/38077

21 Rezende LFM, Ahmadi M, Ferrari G, Del Pozo Cruz B, Lee IM, Ekelund U, et al. Device-measured sedentary time and intensity-specific physical activity in relation to all-cause and cardiovascular disease mortality: the UK Biobank cohort study. Int J Behav Nutr Phys Act. 2024;21(1):68. https://doi.org/10.1186/s12966-024-01615-5

22 Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. Lancet Digit Health. 2020;2(2):e85–93. https://doi.org/10.1016/S2589-7500(19)30222-5

23 Braem CIR, Yavuz US, Hermens HJ, Veltink PH. Missing data statistics provide causal insights into data loss in diabetes health monitoring by wearable sensors. Sensors. 2024;24(5):1526. https://doi.org/10.3390/s24051526

24 Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. JAMA. 2018;320(19):2020–8. https://doi.org/10.1001/jama.2018.14854

25 Nystoriak MA, Bhatnagar A. Cardiovascular effects and benefits of exercise. Front Cardiovasc Med. 2018;5:135. https://doi.org/10.3389/fcvm.2018.00135

26 Karimi F, Amoozgar Z, Reiazi R, Hosseinzadeh M, Rawassizadeh R. Longitudinal analysis of heart rate and physical activity collected from smartwatches. CF Trans Pervasive Comp Interact. 2024;6(1):18–35. https://doi.org/10.1007/s42486-024-00147-y

27 Takeda. A study of TAK-861 in participants with narcolepsy type 1. 2025 [Internet] [cited 2025 July 9]. Available from: https://clinicaltrials.gov/study/NCT05687903

28 Dauvilliers Y, Plazzi G, Mignot E, Lammers GJ, Del Rio Villegas R, Khatami R, et al. Oveporexton, an oral orexin receptor 2-selective agonist, in narcolepsy type 1. N Engl J Med. 2025;392(19):1905–16. https://doi.org/10.1056/NEJMoa2405847

29 O'Brien E, Parati G, Stergiou G, Asmar R, Beilin L, Bilo G, et al. European Society of Hypertension position paper on ambulatory blood pressure monitoring. J Hypertens. 2013;31(9):1731–68. https://doi.org/10.1097/HJH.0b013e328363e964

30 Leroux A, Xu S, Kundu P, Muschelli J, Smirnova E, Chatterjee N, et al. Quantifying the predictive performance of objectively measured physical activity on mortality in the UK biobank. J Gerontol A Biol Sci Med Sci. 2021;76(8):1486–94. https://doi.org/10.1093/gerona/glaa250

31 Chen W, Cordero R, Lever Taylor J, Pangallo DR, Picard RW, Cruz M, et al. Multicenter evaluation of machine-learning continuous pulse rate Algorithm on wrist-worn device. Digit Biomark. 2024;8(1):218–28. https://doi.org/10.1159/000542615

32 Dauvilliers Y, Mignot E, Del Rio Villegas R, Du Y, Hanson E, Inoue Y, et al. Oral orexin receptor 2 agonist in narcolepsy type 1. N Engl J Med. 2023;389(4):309–21. https://doi.org/10.1056/NEJMoa2301940
33 Cajochen C, Pischke J, Aeschbach D, Borbely AA. Heart rate dynamics during human sleep. Physiol Behav. 1994;55(4):769–74. https://doi.org/10.1016/0031-9384(94)90058-2
34 Karas MSJ, Urbanek J. Arctools: processing and physical activity summaries of minute level activity data. 2022. [Internet] Available from: https://CRAN.R-project.org/package=arctools
35 Olaye IM, Belovsky MP, Bataille L, Cheng R, Ciger A, Fortuna KL, et al. Recommendations for defining and reporting adherence measured by biometric monitoring technologies: systematic review. J Med Internet Res. 2022;24(4):e33537. https://doi.org/10.2196/33537