

Original Paper

Importance of Getting Enough Sleep and Daily Activity Data to Assess Variability: Longitudinal Observational Study

María Óskarsdóttir^{1,2*}, PhD; Anna Sigridur Islind^{1,2*}, PhD; Elias August^{2,3}, PhD; Erna Sif Arnardóttir^{1,2,3,4}, PhD; François Patou^{5,6}, PhD; Anja M Maier^{5,7}, PhD

¹Department of Computer Science, Reykjavík University, Reykjavík, Iceland

²Reykjavík University Sleep Institute, School of Technology, Reykjavík University, Reykjavík, Iceland

³Department of Engineering, Reykjavík University, Reykjavík, Iceland

⁴Internal Medicine Services, Landspítali University Hospital, Reykjavík, Iceland

⁵Department of Technology, Management and Economics, DTU-Technical University of Denmark, Copenhagen, Denmark

⁶Oticon Medical, Copenhagen, Denmark

⁷Department of Design, Manufacturing and Engineering Management, Faculty of Engineering, University of Strathclyde, Glasgow, United Kingdom

*these authors contributed equally

Corresponding Author:

María Óskarsdóttir, PhD

Department of Computer Science

Reykjavík University

Menntavegur 1

Reykjavík, 102

Iceland

Phone: 354 5996326

Email: mariaoskars@ru.is

Abstract

Background: The gold standard measurement for recording sleep is polysomnography performed in a hospital environment for 1 night. This requires individuals to sleep with a device and several sensors attached to their face, scalp, and body, which is both cumbersome and expensive. Self-trackers, such as wearable sensors (eg, smartwatch) and nearable sensors (eg, sleep mattress), can measure a broad range of physiological parameters related to free-living sleep conditions; however, the optimal duration of such a self-tracker measurement is not known. For such free-living sleep studies with actigraphy, 3 to 14 days of data collection are typically used.

Objective: The primary goal of this study is to investigate if 3 to 14 days of sleep data collection is sufficient while using self-trackers. The secondary goal is to investigate whether there is a relationship among sleep quality, physical activity, and heart rate. Specifically, we study whether individuals who exhibit similar activity can be clustered together and to what extent the sleep patterns of individuals in relation to seasonality vary.

Methods: Data on sleep, physical activity, and heart rate were collected over 6 months from 54 individuals aged 52 to 86 years. The Withings Aura sleep mattress (nearable; Withings Inc) and Withings Steel HR smartwatch (wearable; Withings Inc) were used. At the individual level, we investigated the consistency of various physical activities and sleep metrics over different time spans to illustrate how sensor data from self-trackers can be used to illuminate trends. We used exploratory data analysis and unsupervised machine learning at both the cohort and individual levels.

Results: Significant variability in standard metrics of sleep quality was found between different periods throughout the study. We showed specifically that to obtain more robust individual assessments of sleep and physical activity patterns through self-trackers, an evaluation period of >3 to 14 days is necessary. In addition, we found seasonal patterns in sleep data related to the changing of the clock for daylight saving time.

Conclusions: We demonstrate that >2 months' worth of self-tracking data are needed to provide a representative summary of daily activity and sleep patterns. By doing so, we challenge the current standard of 3 to 14 days for sleep quality assessment and call for the rethinking of standards when collecting data for research purposes. Seasonal patterns and daylight saving time clock change are also important aspects that need to be taken into consideration when choosing a period for collecting data and designing

studies on sleep. Furthermore, we suggest using self-trackers (wearable and nearable ones) to support longer-term evaluations of sleep and physical activity for research purposes and, possibly, clinical purposes in the future.

(*JMIR Form Res* 2022;6(2):e31807) doi: [10.2196/31807](https://doi.org/10.2196/31807)

KEYWORDS

wearable technology; nearable technology; internet of health care things; sleep; Withings; study duration; establishing standards; seasonality; mHealth; digital health

Introduction

Background

Sleep disorders and short sleep durations are some of the main health challenges in current times. Obstructive sleep apnea is one such disorder and is estimated to affect 1 billion adults worldwide [1]. Insomnia, defined as difficulties in initiating or maintaining sleep, outlines another common sleep disorder [2,3]. Short sleep duration, although not a sleep disorder, is also a major risk factor for adverse health effects and death [3-5]. The gold standard measurement setting for clinical assessment of sleep quality and sleep disturbances is the use of polysomnography for 1 night, typically performed in a hospital environment [6]. Polysomnography is expensive and requires individuals to sleep with several sensors attached to their face, scalp, and body, which is cumbersome [7,8]. Furthermore, data from such a study gives no indication on important routine aspects of sleep quality such as the average total sleep time (TST) of individuals, when they normally go to bed and wake up, whether they are affected by seasonal changes, or whether they have insomnia [9]. Instead, to assess free-living sleep conditions, multiple night recordings in the home environment need to be performed [9]. In the medical field, this is typically accomplished by using wrist actigraphy, which involves a small watch-like device with an embedded accelerometer that often also records ambient light conditions and skin temperature [9]. The use of actigraphy is accompanied by a subjective sleep log or a sleep diary. Clinical guidelines recommend that the individual wears the actigraphy for 7 to 14 days; however, 72 hours of recording is generally sufficient to bill for testing in the United States [9]. For research purposes, 5 to 7 days of actigraphy measurements are often used to assess sleep behavior [10]. These data are used to assess, for example, average sleep duration, chronotype (morningness vs eveningness, commonly referred to as *A-type* vs *B-type*), and similar sleep parameters of interest. This type of data can also be used to facilitate the analysis of individual sleep patterns and for clustering purposes to show trends at the group level [11].

Consumer-Grade Self-Tracking Technologies

More recently, consumer-grade self-tracking technologies that facilitate sleep data collection over longer periods have emerged [12]. Wearable technology (wearables) is an umbrella term for body-worn connected sensors [8]. Smartwatches are an example of such wearables and can capture information similar to actigraphy. Often, they collect even a wider range of physiological signals, such as heart rate, skin temperature, and oxygen saturation [13-15]. Other self-tracking technologies are nearable technologies (nearables), which can also be used to monitor physiological signals by close approximation to the

body. These are increasingly used in conjunction with wearables in health-related research studies [14-16]. For instance, and relevant to our study, they include connected mattresses to monitor sleep patterns in more detail [17]. In most cases, consumer-grade self-trackers are designed for the general purpose of activity tracking. However, their ability to monitor a broad range of physiological parameters means that they are now seriously being considered as alternatives to medical-grade technology for the monitoring of various clinical conditions [18,19]. In addition, the portability and affordability of these trackers open up opportunities for pursuing clinical research on larger cohorts of participants and for rethinking the implementation of remote monitoring care models in specific patient populations [20].

Recent years have seen a surge of research on sleep with consumer-grade self-trackers. Most of these studies focus on relating measurement from the wearable device to either mental or physical health and sometimes both [21,22]. In a few cases, the duration of data collection varies from days and weeks [23-26] to months and years [27,28]. In addition, large sample sizes obtained from the vast number of people who wear self-trackers in the general population have been leveraged to study and compare sleep patterns by age, gender, and BMI worldwide, as in the work presented by Jonasdóttir et al [12]. In terms of duration of data collection, a similar study associated shorter sleep duration and greater variability of sleep duration with increased BMI [28]. Furthermore, the large amount of data collected with self-trackers has encouraged the use of advanced machine learning techniques and deep learning to predict clinical outcomes more robustly [29,31]. Although some studies have taken on the task of observing participants over a longer time span than the gold standard for clinical assessment of sleep quality and sleep disturbances, to our knowledge, only a single study has compared data collected over 1 week with data collected over 2 weeks, concluding that the shorter period is sufficient [24]. Although the sleep research community acknowledges the need for longer periods of data collection with wearable and nearable (nonwearable that is placed near the body) self-trackers, the question of whether the participants should wear self-trackers for a longer time than the gold standard to generate a more insightful portrait of their sleep patterns remains unanswered [11,31].

Aims and Overview

The primary goal of this paper is to investigate whether the time span of 3 to 14 days is sufficient for data collection when performing sleep measurements at home using wearable and nearable sensors. We address the primary goal through the following research question: is 3 to 14 days of data collection sufficient to capture the sleep habits and fluctuations in sleep

patterns of an individual in a reliable way for research purposes? Our secondary goals are to investigate whether there is a relationship between sleep quality, physical activity, and heart rate and whether individuals who exhibit similar activity and sleep patterns in general and in relation to seasonality can be clustered together. We address the secondary goals through the following three research questions:

1. Is there a relationship between sleep quality, physical activity, and heart rate?
2. Can individuals who exhibit similar activity be clustered together in an insightful manner?
3. Are there significant differences between sleep patterns of individuals that are affected by seasonality and daylight saving time (DST) clock changes?

Our a priori hypothesis is that 3 to 14 days' worth of data are neither sufficient to capture a person's sleep habits nor sufficient to observe fluctuations in sleep patterns that might be important for research purposes.

Methods

Data Collection

This study was proposed in the context of the Stanford Medicine X–Digital Health Challenge [32]. It was executed under an ethical waiver from the central Danish National Committee on Health Research Ethics. The participants were recruited through advertisements in 2 local newspapers (*Søndagsavisen Vestegnen* and *Villabyerne*) distributed within Greater Copenhagen in Denmark. A total of 82 adults aged >50 years were screened. The first screening was conducted over the phone. Candidates were then scheduled for a home visit, during which the Montreal

Cognitive Assessment test was administered by a trained neuropsychologist. The Montreal Cognitive Assessment scores were collected but are not reported in this paper as it was outside the scope of this study (see [33] for details). Of the 82 individuals, 54 (66%) (aged 52–86 years; male: 35/54, 65%; female: 19/54, 35%) fulfilled the inclusion criteria of the study. All participants signed informed consent to join the study and agreed to share their data. At a second home visit, the participants were equipped with the wearable Withings Steel HR smartwatch (Withings Inc), tracking the number of steps and heart rate on a per-minute basis. Participants were also equipped with the nearable Withings Aura sleep mattress (Withings Inc), tracking the various phases of sleep (sleep onset latency, wake, light sleep, deep sleep, rapid eye movement [REM] sleep, and waking up times) on a per-minute basis [34]. The first day of data logging for the participants spanned from June 7, 2017, to September 25, 2017. Data logging stopped for all participants on December 28, 2017. Figure 1 shows an overview of the days for which data were acquired for all participants in the study. In addition, the participants' age, height, weight, and gender were noted upon entry into the study.

The study is based on data from the abovementioned devices—smartwatch (wearable) and sleep mattress (nearable)—and specifically the variables listed in Table 1. Most of the variables in Table 1 are either measured directly or calculated by the smartwatch during the day or sleep mattress during the night. In addition to those, we derived 2 commonly used variables in sleep research, namely, TST, which is the time in hours from falling asleep until final wake up, and sleep midtime, which is the midtime between falling asleep and final wake up.

Figure 1. Overview of data collection for the participants in the study. The dots indicate dates with measurements. Blue represents periods with complete data from the smartwatch during the day, and red represents data from the sleep mattress during the night.



Table 1. Overview and definitions of the variables used in our analysis and the self-trackers used to collect them.

Name	Description	Device
Daily step count	Number of steps during the day	Smartwatch
Diurnal heart rate–average	Mean heart rate during the day	Smartwatch
Nocturnal heart rate–average	Mean heart rate during the night	Sleep mattress
Total duration in bed	Time in hours from going to bed until getting out of bed	Sleep mattress
Total sleep time	Time in hours from falling asleep until final wake up	Sleep mattress
Sleep onset latency	Time in minutes from going to bed until falling asleep	Sleep mattress
Number of times awake	Count of how often the individual woke up during the night	Sleep mattress
Deep sleep duration	Time in hours spent in deep sleep	Sleep mattress
REM ^a sleep duration	Time in hours spent in REM sleep	Sleep mattress
Light sleep duration	Time in hours spent in light sleep	Sleep mattress
Sleep midtime	Midtime between falling asleep and final wake up	Sleep mattress

^aREM: rapid eye movement.

Group-Level Analyses

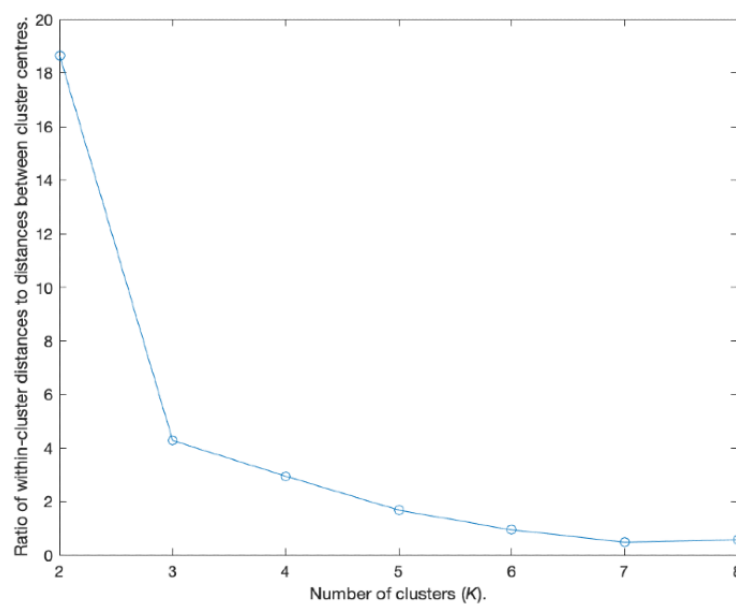
To inspect long-term changes and variability at the cohort level, we considered the participants for whom data were collected the longest, starting in June 2017 until December 2017. For this part of the cohort (25/54, 46%), we calculated the daily means and SDs of the measurements.

To inspect whether there are discernable patterns in the day-to-day activities of the participants, we used K -means clustering that aims to group together similar numerical data, where similarity is defined through the Euclidean distance, particularly, to partition N observations into K clusters [35]. First, we applied the method using only data from the last week

of the trial and then from the last 2 weeks and so on. We decided to do so because of the different starting points and to avoid seasonality effects.

To determine the number of clusters, we used the elbow method [36]. The number of clusters was chosen such that adding an additional one does not increase (much) the information gained. Specifically, we recorded the ratio of within-cluster distances of all clusters to distances between cluster centers and used Figure 2 to determine when it ceases to change (much). This created an *elbow* in the graph at $K=3$, after which not much change occurs. Note that to investigate which variables differed with statistical significance between clusters, we used the 2-sampled Student 2-tailed t test with $P < .05$ significance level.

Figure 2. Elbow test.



Individual-Level Analyses

To demonstrate the variability in sleep and daily activity at the individual level, 7% (4/54) of the participants were selected at random and studied in depth. Following the selection, their values were compared with those of other participants. Figure 3 shows the variable distributions of these four participants, which fell within the same range as that of the entire population. These participants were not meant to be a representative sample of the cohort, and the rationale behind our choice to show only data from 7% (4/54) of participants was to clearly demonstrate the variety in measurement patterns among participants without compromising the readability of the figures.

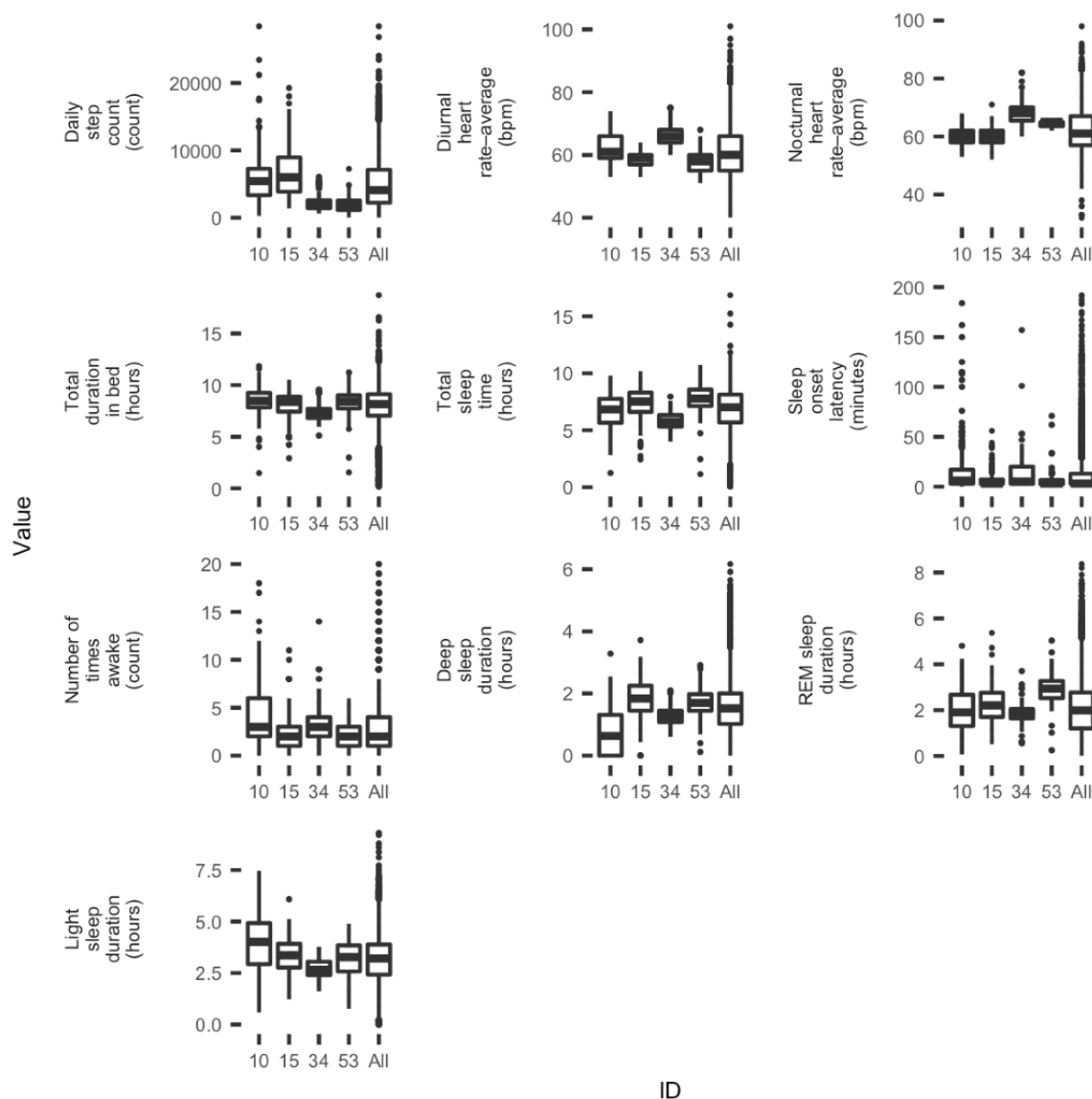
We considered 3 perspectives in the individual-level analyses. First, we assessed the day-to-day values for variables associated with sleep quality for a span of 1 week. The week was chosen at random. Subsequently, we calculated the weekly mean and SD for the different variables for a span of 10 weeks for the same 4 participants, where we normalized variable values for each participant by dividing by the largest value measured in

the collection period. These values showed how sleep and daily activity changed from day to day and week to week.

Second, we looked at the evolution of the SD of sleep and activity measurements. We calculated a rolling SD over 7 days with a 1-day moving window from the first week of October 2017 until the end of December 2017. Moreover, starting with the first 3 days of October 2017, we calculated the SD of each participant's measurements. Then, we added the next day and performed the calculation again. We repeated the procedure until 80 days had been added to the original 3 days. Thus, we obtained a sequence of SD values that described the variability in each participant's measurement.

Finally, to investigate seasonal effects and, in particular, the impact of the DST clock change on October 29, 2017, we used a 2-tailed t test to evaluate whether differences in the values of each of the 11 variables before and after the DST clock change were significant using a .05 significance level. For this, we considered 3 periods: (1) short-term: 15 days before and 15 days after October 29, 2017; (2) midterm: 30 days before and 30 days after October 29, 2017; and (3) long-term: 60 days before and 60 days after October 29, 2017.

Figure 3. Boxplots showing the distributions for the variables considered (daily step count, diurnal heart rate–average, nocturnal heart rate–average, total duration in bed, total sleep time, sleep onset latency, number of times awake, deep sleep duration, rapid eye movement sleep duration, and light sleep duration) for the four participants and the whole population. As sleep midtime is a circular variable, it is not considered in this figure. bpm: beats per minute; REM: rapid eye movement.



Results

Group-Level Analyses

Some of the participants in the study wore self-trackers for 6 months. This allowed us to look at trends over a longer period and assess seasonal patterns. Figure 4 shows daily means and SDs for TST, total duration in bed, sleep onset latency, sleep midtime, and daily step count. These variables showed the most evidence of seasonal effects. Major trends in the data indicate that total duration in bed increased, albeit the TST remained similar. The sleep onset latency leaped at the end of October 2017, when DST stops in Europe and the clock is set back by 1 hour. Clearly, the participants in this study were affected by this change, as shown by the increased time they took to fall asleep in the weeks after the change of the clock. We also saw

a downward trend in the number of steps throughout the 6-month period and fluctuations in the sleep midtime.

Clustering analysis resulted in the suggestion of 2 distinct cohorts of approximately the same size (25 participants each) and a third one that we neglected for its small size when >4 weeks of data were used (Figure 5); the third cluster with 6 participants was omitted. Table 2 shows the mean and SD of all the variables in the two cohorts. The 2-tailed *t* test results showed that only differences in *number of times awake* were statistically significant. On average, the difference between the cohorts in the number of times awake was 1 time. On average, most participants woke up <4 times per night, whereas 6% (3/54) of study participants woke up >5 times per night. We did not find statistically significant differences in gender, age, or BMI between the two clusters.

Figure 4. Seasonal differences in (A) daily step count, (B) total duration in bed, (C) total sleep time, (D) sleep onset latency, and (E) sleep midtime. The blue line indicates the local polynomial regression fit, and the red dashed line indicates the start of daylight saving time on October 29, 2017.

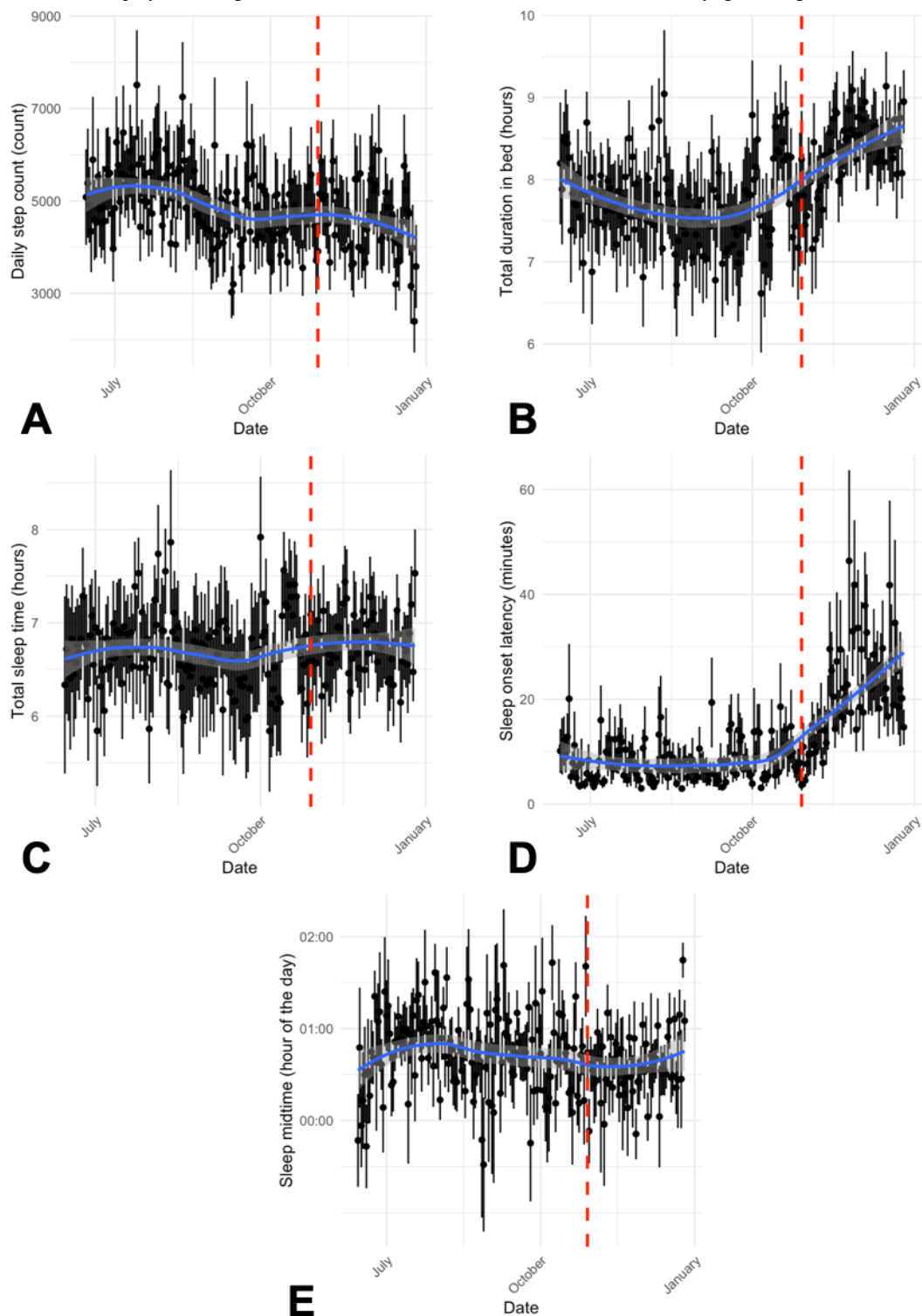


Figure 5. Cluster analysis revealed 2 cohorts (red and blue in the figure). The figure shows a difference in the variable number of times awake between the 2 cohorts. Other variables were less distinctive. REM: rapid eye movement.

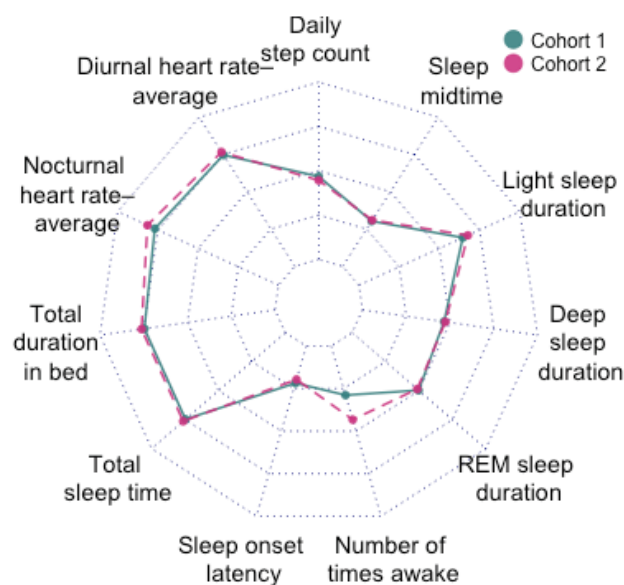


Table 2. Mean and SD of the 11 variables in the 2 cohorts.

Variable	Cohort 1, mean (SD)	Cohort 2, mean (SD)
Daily step count (count)	4895.97 (2772.25)	4996.31 (2230.61)
Diurnal heart rate-average (bpm)	58.89 (5.91)	61.85 (6.67)
Nocturnal heart rate-average (bpm)	61.10 (6.02)	62.66 (6.85)
Total duration in bed (h)	7.81 (0.82)	8.06 (1.24)
Total sleep time (h)	6.86 (0.99)	7.08 (1.29)
Sleep onset latency (min)	10.22 (5.25)	8.70 (4.05)
Number of times awake (count)	2.03 (0.93) ^a	3.06 (1.27) ^a
Duration of REM ^b sleep (h)	1.54 (0.34)	1.53 (0.63)
Duration of deep sleep (h)	2.18 (0.68)	2.23 (0.77)
Duration of light sleep (h)	3.13 (0.56)	3.32 (0.68)
Sleep midtime (time)	11,249.91 (2645.61)	11,493.10 (3465.17)

^aIndicates that the difference was statistically significant at the .05 confidence level.

^bREM: rapid eye movement.

Individual-Level Analyses

Figure 6 shows time-series data for 6 variables (total duration in bed, TST, sleep onset latency, sleep midtime, deep sleep duration, and REM sleep duration) collected for 1 week for 7% (4/54) of participants on a night-to-night basis. The remaining 5 variables (daily step count, diurnal heart rate-average, nocturnal heart rate-average, number of times awake, and light sleep duration) can be seen in Figure S1 in [Multimedia Appendix 1](#). The figures show a clear difference for each

participant on a day-to-day basis and among the four of them. The relationship between the duration of REM and deep sleep differed for the participants considered here. For participant 34, they were in sync, but not for the remaining participants. Finally, sleep onset latency appeared regular for all 4 participants, and of them, 2 (50%) had days where it peaked.

Figure 7 shows weekly averages over a period of 10 weeks for the same 7% (4/54) of participants and 6 of the variables (total duration in bed, TST, sleep onset latency, sleep midtime, deep

sleep duration, and REM sleep duration). The remaining 5 variables (daily step count, diurnal heart rate–average, nocturnal heart rate–average, number of times awake, and light sleep duration) can be seen in Figure S2 in Multimedia Appendix 1. Here, we see that the variation in measurements was even greater. For example, the sleep onset latency of participant 10 was gradually increasing, a pattern that can also be discerned in the total duration in bed and sleep midtime plots. In some weeks, the SD was large, which indicated that the values in those weeks spanned a wide range. The measurements of participant 53 showed stark fluctuation during the 10-week period.

Figure 8 shows the correlation between the weekly averages in Figure 7 for the 7% (4/54) of participants, which varied greatly. Participants 10, 15, and 34 had some positive and negative correlations between their variables. For example, for participant 10, there was a positive correlation between light sleep duration and TST and a negative correlation between light sleep duration and sleep onset latency. In contrast, for participant 53, most of the correlations were strongly positive or negative, showing great synergy. Only sleep onset latency showed little correlation with the other variables.

Figure 6. Daily parameters over a period of 1 week for the 4 participants. Each column and color represent one of the participants. REM: rapid eye movement.

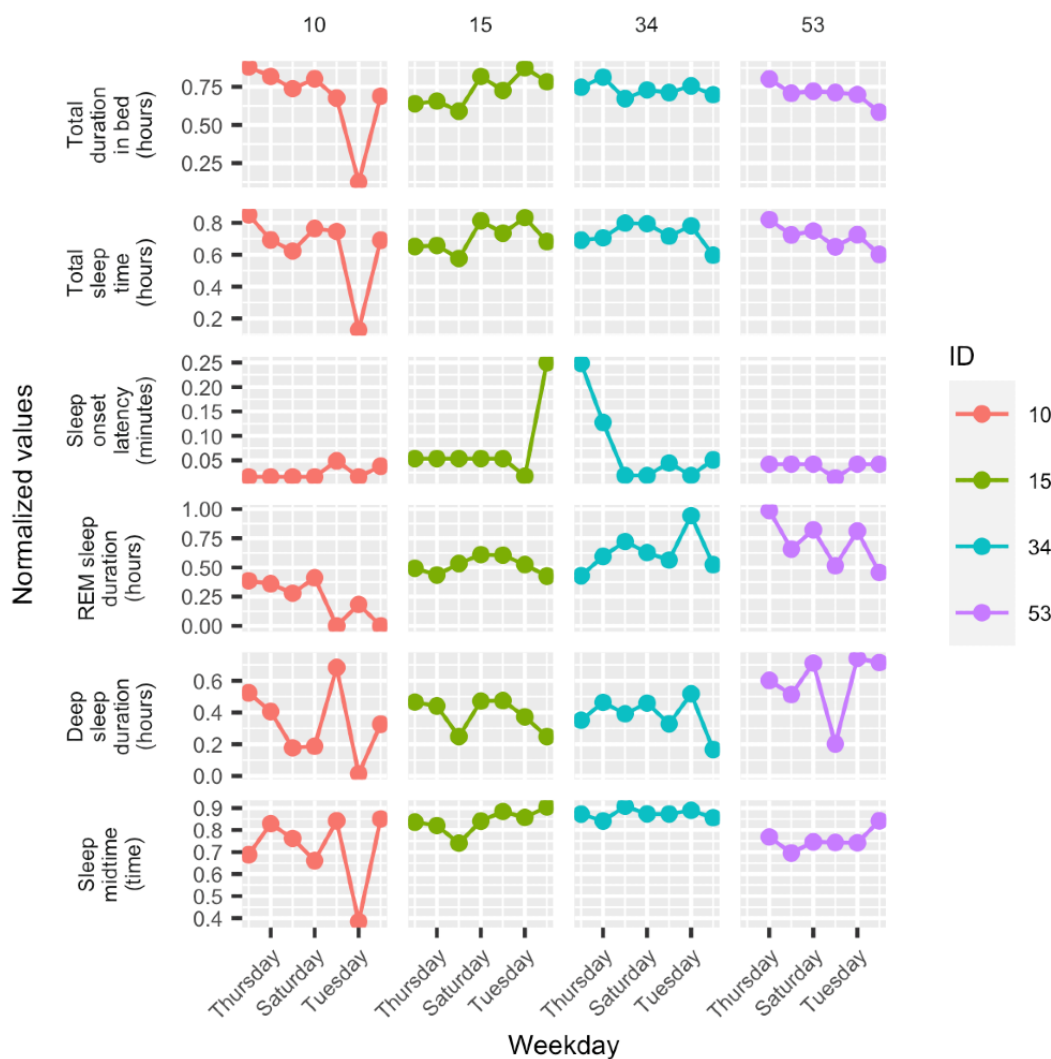


Figure 7. Average activity by week over a 10-week period for the 4 participants. The bars denote the SD within each week. REM: rapid eye movement.

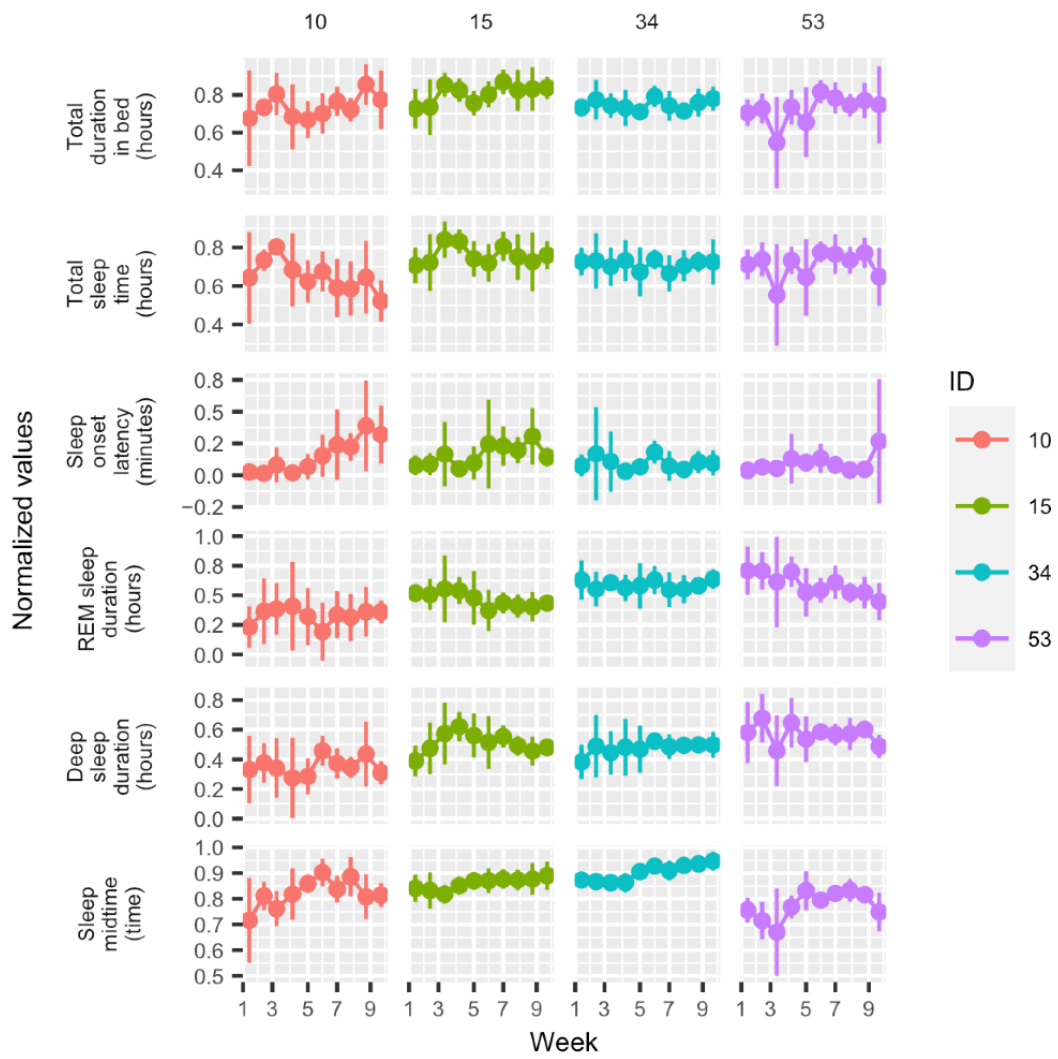
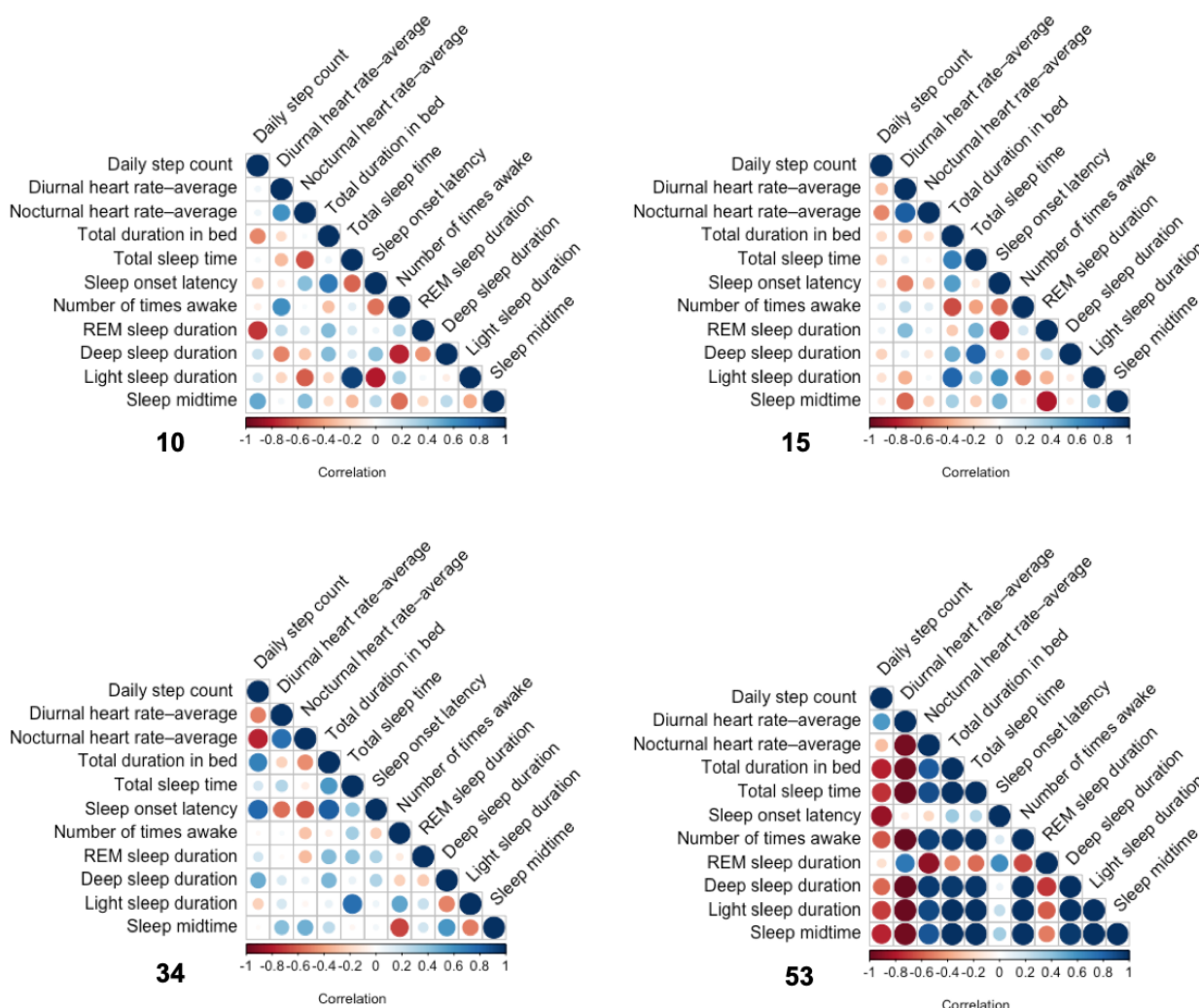


Figure 8. Correlation between the 11 variables for the 4 participants. REM: rapid eye movement.



Next, we investigated the variability in different variables by focusing on the changes in SD over time. Firstly, Figure 9 shows the rolling SD computed over 1 week with a rolling window of 1 day for 6 of the variables (TST, total duration in bed, sleep midtime, deep sleep duration, REM sleep duration, and sleep onset latency). The remaining 5 variables (daily step count, diurnal heart rate-average, nocturnal heart rate-average, number of times awake, and light sleep duration) can be seen in Figure S3 in Multimedia Appendix 1. From the figure, we see that the SD changed greatly throughout time for all participants and for all measurements. Participant 34 had little variation in measurements. The SD of TST and total duration in bed remains within 1 hour. However, in the first weeks, the SD of the sleep onset latency went up to 60 minutes. The other participants had greater fluctuations throughout the period, with SD of TST reaching 3 hours for participants 10 and 53. In addition, the variability in deep and REM sleep duration decreased over time. The data also shows that the variability in sleep onset latency had an increasing trend in the 10-week period.

Finally, Figure 10 shows the cumulative SD of the 7% (4/54) of participants for the same 6 variables. The remaining 5

variables can be seen in Figure S4 in Multimedia Appendix 1. These plots give a sense of the participants' overall variability over time and how it stabilized as more days were added to the data collection. The plots show that 1 week is not representative of someone's sleep behavior as it can change drastically from week to week.

We now assessed the effects of seasonality on sleep and sleep quality. More precisely, we investigated which, if any, of the variables were significantly different before and after the DST clock change when looking at short-term (15 days before and after October 29, 2017), midterm (30 days before and after October 29, 2017), and long-term (60 days before and after October 29, 2017) periods for the 7% (4/54) of participants. Table 3 and Figure 11 show which variables had a significant difference before and after October 29, 2017. The difference in sleep midtime was almost always significant. Also, long-term changes were the most statistically significant, and before the changing of the clock, the participants spent more time in REM sleep, the midtime of their sleep was earlier, and they fell asleep faster.

Figure 9. Rolling SDs over a 10-week period. These were calculated over 7 days with a 1-day rolling window. REM: rapid eye movement.

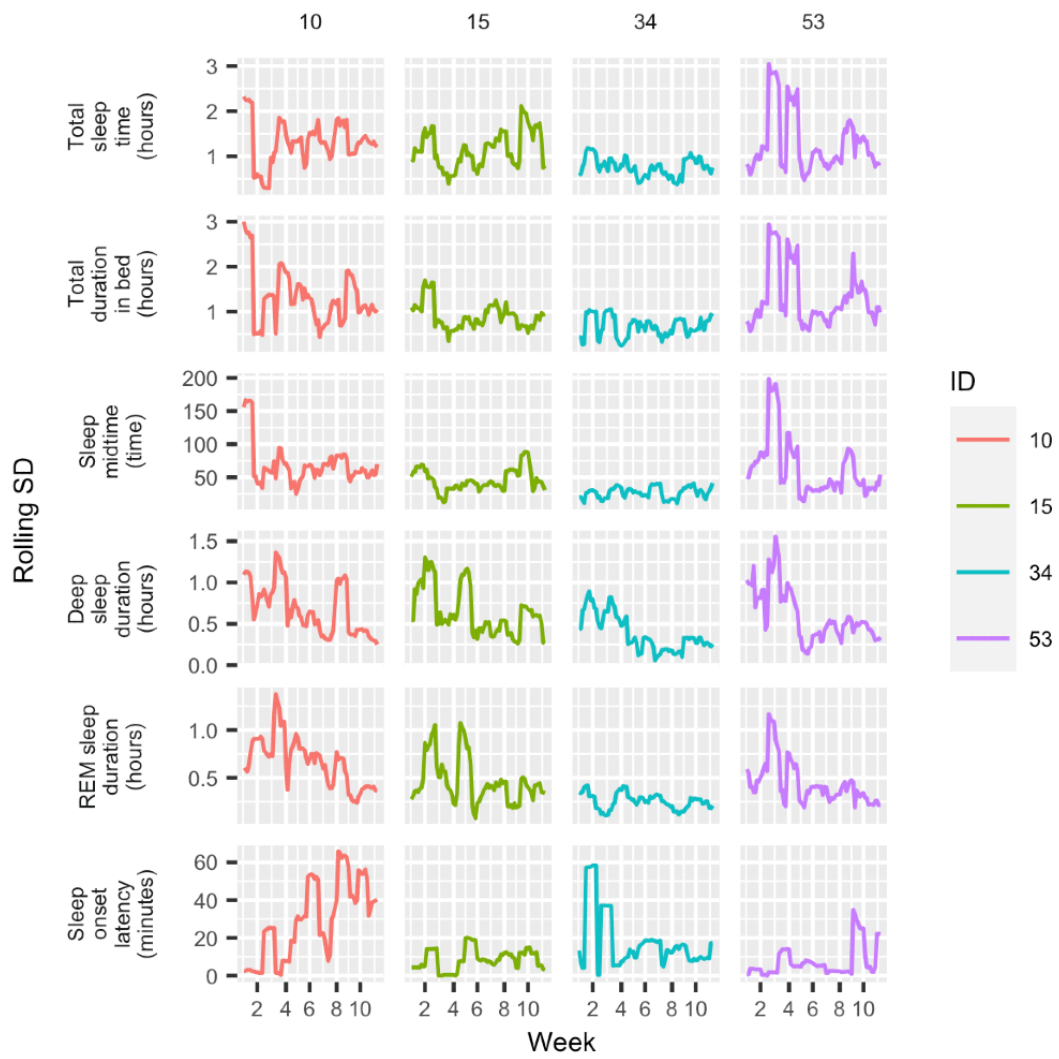


Figure 10. Cumulative SD over a 10-week period, adding 1 day at a time. REM: rapid eye movement.

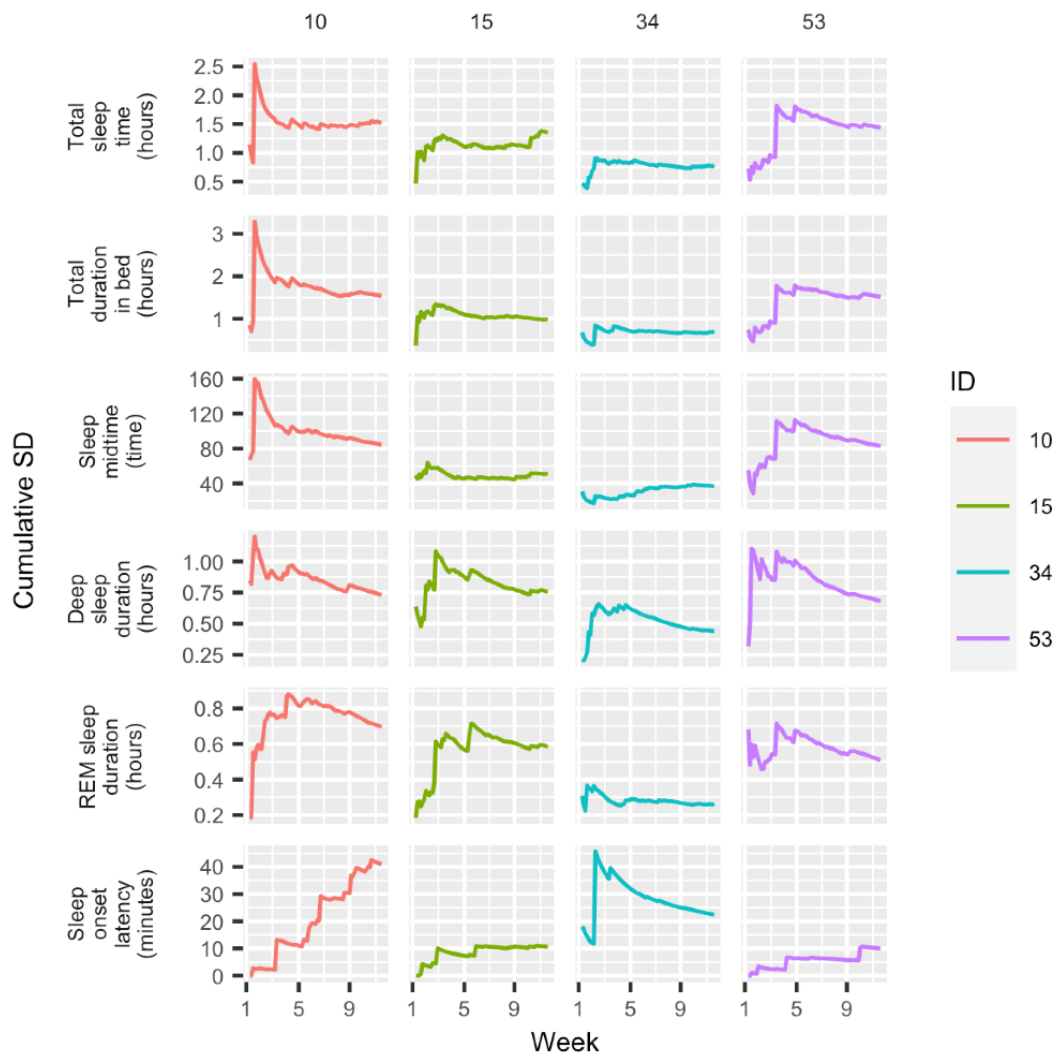


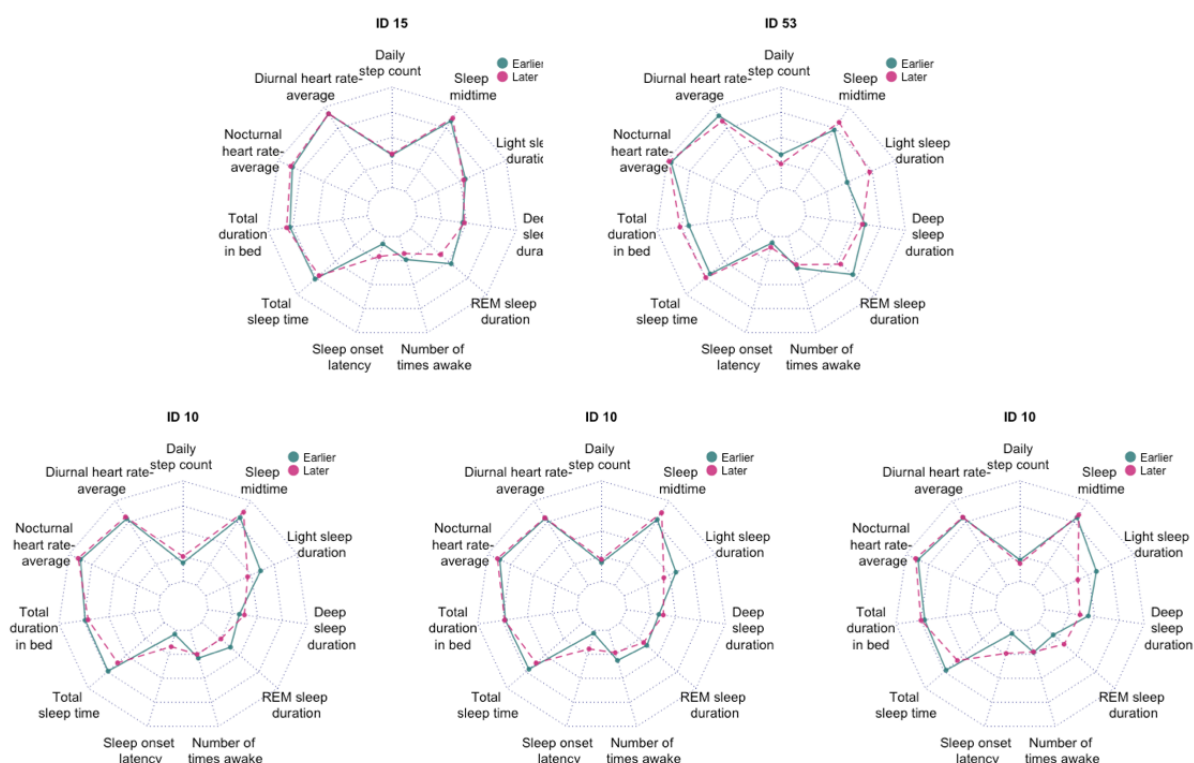
Table 3. Statistical significance of the difference in variables before and after October 29, 2017, for each of the four participants during the three periods.

Variables	Short term				Midterm				Long term			
	ID 10	ID 15	ID 34	ID 53	ID 10	ID 15	ID 34	ID 53	ID 10	ID 15	ID 34	ID 53
Daily step count												√ ^a
Diurnal heart rate–average						√		√				√
Nocturnal heart rate–average					√				√	√		
Total duration in bed												√
Total sleep time	√	√			√				√	√		
Sleep onset latency	√				√	√			√	√		
Number of times awake					√		√			√	√	
REM ^b sleep duration						√		√	√	√		√
Deep sleep duration										√		
Light sleep duration	√			√	√		√	√	√			√
Sleep midtime	√	√	√	√	√	√	√	√		√	√	√

^aIndicates a statistical significance at the .05 confidence level.

^bREM: rapid eye movement.

Figure 11. Mean values for the variables before and after October 29, 2017. Top row: long-term patterns for participants 15 and 53. Bottom row: short-term, midterm, and long-term patterns, which represent 15, 30, and 60 days before and after October 29, 2017, respectively, for participant 10. Both participants 15 ($P<.0001$) and 53 ($P<.001$) had significantly longer REM sleep before the change. Participant 10 fell asleep sooner, slept longer, and spent more time in light sleep before the change. REM sleep duration also changed from short term to long term. REM: rapid eye movement.



Discussion

Principal Findings

Although the sleep research community welcomes the advancement of consumer-grade self-trackers, including wearables and nearables, they are also widely aware of the numerous challenges that remain, especially regarding the need to validate the devices to ensure their accuracy and reliability [8]. Although many of these barriers are for wearable and nearable technology companies to solve, it is the responsibility of the sleep research community and the medical informatics community to make a collective effort and decide upon necessary and sufficient requirements for validating the devices [31]. As we show in this paper, the duration of the validation period is an essential but grossly overlooked factor. Guillodo et al [11] acknowledged the need for long-term sleep studies, which could help identify connections between sleep quality and health outcomes. However, few attempts have been made with data from wearables and nearables. Although data collected over a longer period is essential, it is also important to make a clear distinction between group-level and individual-level approaches when it comes to research goals, clinical value, and data analysis. Although data collected over an extended period in a large cohort can reveal interesting insights about sleep patterns of the general population [10], there is much potential in using wearables and nearable devices for individualized medicine approaches as well. The approach used in this paper, where we studied individual patterns, has been fruitful for understanding sleep patterns over time.

Another shortcoming in the sleep literature is that it views and analyzes individual nights instead of analyzing time series, where trends, seasonality, and other long terms patterns can be discovered. On that basis, we showed that sleep patterns vary highly from person to person, and, because of that, an individualized approach may be more appropriate than pooling the data per night for several individuals, as is common in the literature. Moreover, we can see that the type of wearable or nearable is not the main value; instead, the main value is in comparing data from the same device, for the same individual, over an extended period. It has been acknowledged that clinical practices should embrace the unique characteristics of individual patients and their patterns and seek to individualize patient care; clearly, the same should hold for sleep [37].

In this paper, our primary research goal was to investigate whether the gold standard, the traditional time span of up to 2 weeks, is sufficient for obtaining reliable data to assess sleep duration and sleep quality of an individual when performing sleep measurements at home using wearable and nearable sensors. Our answer to this question is no. Specifically, we showed that there is much variability in the self-tracker measurements for individual participants across time.

Furthermore, in our cohort analysis, we observed a clear distinction in the empirical data only when using sufficient data (>30 days) and could show the emergence of clusters that are robust to changes in the amount of data and the specific dates chosen for the analysis. However, when following individual behavior, an even longer period is needed, and we recommend >2 months.

The secondary research goal of this paper was to investigate whether there was a relationship among sleep quality, physical activity, and heart rate and whether within-group patterns in clusters of individuals exhibit similar activity and sleep patterns, both in general and in relation to seasonality. Our results show a seasonal effect on sleep patterns is related to the changing of the clock. This could both be because of overall seasonal changes and affected by the DST change, which has a significant effect on sleep patterns. This has been acknowledged previously, for instance, by international sleep and biological rhythm societies [38]. We show that there is much variability in the self-tracker measurements and apparent correlation between variables among participants [38].

Conclusions

In conclusion, analysis and exploration of time-series data have given new insights about collecting and analyzing data from self-trackers. The findings in this paper show that it is important to get enough sleep data when attempting to understand sleep patterns from self-trackers in depth. First, the gold standard is less useful as there is much variation in the measures, both on a day-to-day basis and a week-to-week basis. This means that when collecting data on individuals, we recommend a longer period to capture as much of this variability as possible. Second, the variation in the patterns in the data is high from person to person. Although cluster analysis indicates that some patterns seem common among groups of people, our individual observations indicate that the analysis should be conducted on a person-by-person basis by training algorithms to learn individual patterns. Thus, further analysis is needed to investigate the number of days suitable for data collection with self-trackers and whether these patterns and correlations observed are common among groups of people, particularly as our analyses were only based on data from self-trackers and additional information such as illnesses, exercise plans, medication, or medical history were not included. Further limitations of the study include recruitment bias as participants were not randomly selected but were included from a homogeneous sample, and the sample size of 54 individuals affects robust conclusions. The novel finding and call to action of this paper is to reconsider the gold standard in sleep research from 14 days to >3 months. The proposition of this paper is that wearables and nearables make this possible and appear promising for clinical research under free-living sleep conditions, such as at home.

Acknowledgments

The authors thank Withings Inc for providing the health devices—the smartwatch and the sleep mattress. The authors thank all the study participants and collaborators of the Stanford Medicine X–Digital Health Challenge.

Conflicts of Interest

ESA discloses lecture fees from Nox Medical, Philips, and ResMed, outside the scope of the current manuscript.

Multimedia Appendix 1

Additional figures showing trends in daily and weekly averages and weekly trends in cumulative and rolling SD for the five variables not included in the manuscript.

[\[DOCX File , 978 KB-Multimedia Appendix 1\]](#)

References

1. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MS, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med* 2019 Aug;7(8):687-698 [FREE Full text] [doi: [10.1016/S2213-2600\(19\)30198-5](https://doi.org/10.1016/S2213-2600(19)30198-5)] [Medline: [31300334](https://pubmed.ncbi.nlm.nih.gov/31300334/)]
2. Zeng L, Zong Q, Yang Y, Zhang L, Xiang Y, Ng CH, et al. Gender difference in the prevalence of insomnia: a meta-analysis of observational studies. *Front Psychiatry* 2020;11:577429 [FREE Full text] [doi: [10.3389/fpsy.2020.577429](https://doi.org/10.3389/fpsy.2020.577429)] [Medline: [33329116](https://pubmed.ncbi.nlm.nih.gov/33329116/)]
3. Kocevská D, Lysen TS, Dotinga A, Koopman-Verhoeff ME, Luijk MP, Antypa N, et al. Sleep characteristics across the lifespan in 1.1 million people from the Netherlands, United Kingdom and United States: a systematic review and meta-analysis. *Nat Hum Behav* 2021 Jan;5(1):113-122. [doi: [10.1038/s41562-020-00965-x](https://doi.org/10.1038/s41562-020-00965-x)] [Medline: [33199855](https://pubmed.ncbi.nlm.nih.gov/33199855/)]
4. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *Eur Heart J* 2011 Jun;32(12):1484-1492. [doi: [10.1093/eurheartj/ehr007](https://doi.org/10.1093/eurheartj/ehr007)] [Medline: [21300732](https://pubmed.ncbi.nlm.nih.gov/21300732/)]
5. Pienaar PR, Kolbe-Alexander TL, van Mechelen W, Boot CR, Roden LC, Lambert EV, et al. Associations between self-reported sleep duration and mortality in employed individuals: systematic review and meta-analysis. *Am J Health Promot* 2021 Jul;35(6):853-865. [doi: [10.1177/0890117121992288](https://doi.org/10.1177/0890117121992288)] [Medline: [33567861](https://pubmed.ncbi.nlm.nih.gov/33567861/)]
6. The AASM manual for the scoring of sleep and associated events. American Academy of Sleep Medicine. URL: <https://aasm.org/clinical-resources/scoring-manual/> [accessed 2021-12-31]
7. Fischer J, Dogas Z, Bassetti CL, Berg S, Grote L, Jennum P, Executive Committee (EC) of the Assembly of the National Sleep Societies (ANSS), Board of the European Sleep Research Society (ESRS), Regensburg, Germany. Standard procedures for adults in accredited sleep medicine centres in Europe. *J Sleep Res* 2012 Aug;21(4):357-368 [FREE Full text] [doi: [10.1111/j.1365-2869.2011.00987.x](https://doi.org/10.1111/j.1365-2869.2011.00987.x)] [Medline: [22133127](https://pubmed.ncbi.nlm.nih.gov/22133127/)]
8. Arnardóttir ES, Islind AS, Óskarsdóttir M. The future of sleep measurements: a review and perspective. *Sleep Med Clin* 2021 Sep;16(3):447-464. [doi: [10.1016/j.jsmc.2021.05.004](https://doi.org/10.1016/j.jsmc.2021.05.004)] [Medline: [34325822](https://pubmed.ncbi.nlm.nih.gov/34325822/)]
9. Ancoli-Israel S, Martin JL, Blackwell T, Buenaver L, Liu L, Meltzer LJ, et al. The SBSM guide to actigraphy monitoring: clinical and research applications. *Behav Sleep Med* 2015;13 Suppl 1:S4-38. [doi: [10.1080/15402002.2015.1046356](https://doi.org/10.1080/15402002.2015.1046356)] [Medline: [26273913](https://pubmed.ncbi.nlm.nih.gov/26273913/)]
10. Rognvaldsdóttir V, Guðmundsdóttir SL, Brychta RJ, Hrafnkelsdóttir SM, Gestsdóttir S, Arngrímsson SA, et al. Sleep deficiency on school days in Icelandic youth, as assessed by wrist accelerometry. *Sleep Med* 2017 May;33:103-108 [FREE Full text] [doi: [10.1016/j.sleep.2016.12.028](https://doi.org/10.1016/j.sleep.2016.12.028)] [Medline: [28449887](https://pubmed.ncbi.nlm.nih.gov/28449887/)]
11. Guillodo E, Lemey C, Simonnet M, Walter M, Baca-García E, Masetti V, HUGOPSY Network, et al. Clinical applications of mobile health wearable-based sleep monitoring: systematic review. *JMIR Mhealth Uhealth* 2020 Apr 01;8(4):e10733 [FREE Full text] [doi: [10.2196/10733](https://doi.org/10.2196/10733)] [Medline: [32234707](https://pubmed.ncbi.nlm.nih.gov/32234707/)]
12. Jonasdóttir SS, Minor K, Lehmann S. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. *Sleep* 2021 Feb 12;44(2):zsa169. [doi: [10.1093/sleep/zsa169](https://doi.org/10.1093/sleep/zsa169)] [Medline: [32886772](https://pubmed.ncbi.nlm.nih.gov/32886772/)]
13. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. *Info Systems J* 2018 Jan 29;28(6):995-1019. [doi: [10.1111/isj.12179](https://doi.org/10.1111/isj.12179)]
14. Mettler T, Wulf J. Physiolytics at the workplace: affordances and constraints of wearables use from an employee's perspective. *Info Systems J* 2018 Jul 06;29(1):245-273. [doi: [10.1111/isj.12205](https://doi.org/10.1111/isj.12205)]
15. Swan M. The quantified self: fundamental disruption in big data science and biological discovery. *Big Data* 2013 Jun;1(2):85-99. [doi: [10.1089/big.2012.0002](https://doi.org/10.1089/big.2012.0002)] [Medline: [27442063](https://pubmed.ncbi.nlm.nih.gov/27442063/)]
16. Lindroth T, Islind AS, Steineck G, Lundin J. From narratives to numbers: data work and patient-generated health data in consultations. *Stud Health Technol Inform* 2018;247:491-495. [Medline: [29678009](https://pubmed.ncbi.nlm.nih.gov/29678009/)]
17. Depner CM, Cheng PC, Devine JK, Khosla S, de Zambotti M, Robillard R, et al. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep* 2020 Feb 13;43(2):zsz254 [FREE Full text] [doi: [10.1093/sleep/zsz254](https://doi.org/10.1093/sleep/zsz254)] [Medline: [31641776](https://pubmed.ncbi.nlm.nih.gov/31641776/)]
18. Hwang Y, Kim MG, Rho J. Understanding Internet of Things (IoT) diffusion: focusing on value configuration of RFID and sensors in business cases (2008–2012). *Inf Develop* 2016 Jul 09;32(4):969-985. [doi: [10.1177/0266666915578201](https://doi.org/10.1177/0266666915578201)]
19. Islind AS. Platformization: co-designing digital platforms in practice. *EfterCancern*. 2018. URL: https://www.researchgate.net/publication/327175455_Platformization_Co-Designing_Digital_Platforms_in_Practice [accessed 2021-12-31]

20. Saganowski S, Kazienko P, Dziezyc M, Jakimów P, Komoszyńska J, Michalska W, et al. Review of consumer wearables in emotion, stress, meditation, sleep, and activity detection and analysis. *Emotion Recognition Using Wearables*. 2020. URL: https://www.researchgate.net/publication/341118509_Review_of_Consumer_Wearables_in_Emotion_Stress_Meditation_Sleep_and_Activity_Detection_and_Analysis [accessed 2021-12-31]
21. Thorpe JR, Forchhammer BH, Maier AM. Development of a sensor-based behavioral monitoring solution to support dementia care. *JMIR Mhealth Uhealth* 2019 May 30;7(6):e12013 [FREE Full text] [doi: [10.2196/12013](https://doi.org/10.2196/12013)] [Medline: [31199304](https://pubmed.ncbi.nlm.nih.gov/31199304/)]
22. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. *JMIR Form Res* 2019 Oct 17;3(4):e12346 [FREE Full text] [doi: [10.2196/12346](https://doi.org/10.2196/12346)] [Medline: [31625951](https://pubmed.ncbi.nlm.nih.gov/31625951/)]
23. Berryhill S, Morton CJ, Dean A, Berryhill A, Provencio-Dean N, Patel SI, et al. Effect of wearables on sleep in healthy individuals: a randomized crossover trial and validation study. *J Clin Sleep Med* 2020 May 15;16(5):775-783 [FREE Full text] [doi: [10.5664/jcsm.8356](https://doi.org/10.5664/jcsm.8356)] [Medline: [32043961](https://pubmed.ncbi.nlm.nih.gov/32043961/)]
24. Briscoe S, Hardy E, Pengo MF, Kosky C, Williams AJ, Hart N, et al. Comparison of 7 versus 14 days wrist actigraphy monitoring in a sleep disorders clinic population. *Chronobiol Int* 2014 Apr;31(3):356-362. [doi: [10.3109/07420528.2013.858163](https://doi.org/10.3109/07420528.2013.858163)] [Medline: [24304408](https://pubmed.ncbi.nlm.nih.gov/24304408/)]
25. Bradshaw DA, Yanagi MA, Pak ES, Peery TS, Ruff GA. Nightly sleep duration in the 2-week period preceding multiple sleep latency testing. *J Clin Sleep Med* 2007 Oct 15;3(6):613-619 [FREE Full text] [Medline: [17993043](https://pubmed.ncbi.nlm.nih.gov/17993043/)]
26. Melton BF, Buman MP, Vogel RL, Harris BS, Bigham LE. Wearable devices to improve physical activity and sleep: a randomized controlled trial of college-aged African American women. *J Black Stud* 2016 Jul 27;47(6):610-625. [doi: [10.1177/0021934716653349](https://doi.org/10.1177/0021934716653349)]
27. Kalmbach DA, Fang Y, Arnedt JT, Cochran AL, Deldin PJ, Kaplin AI, et al. Effects of sleep, physical activity, and shift work on daily mood: a prospective mobile monitoring study of medical interns. *J Gen Intern Med* 2018 Jun;33(6):914-920 [FREE Full text] [doi: [10.1007/s11606-018-4373-2](https://doi.org/10.1007/s11606-018-4373-2)] [Medline: [29542006](https://pubmed.ncbi.nlm.nih.gov/29542006/)]
28. Jaiswal SJ, Quer G, Galarnyk M, Steinhubl SR, Topol EJ, Owens RL. Association of sleep duration and variability with body mass index: sleep measurements in a large US population of wearable sensor users. *JAMA Intern Med* 2020 Dec 01;180(12):1694-1696 [FREE Full text] [doi: [10.1001/jamainternmed.2020.2834](https://doi.org/10.1001/jamainternmed.2020.2834)] [Medline: [32926073](https://pubmed.ncbi.nlm.nih.gov/32926073/)]
29. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, et al. Sleep quality prediction from wearable data using deep learning. *JMIR Mhealth Uhealth* 2016 Nov 04;4(4):e125. [doi: [10.2196/mhealth.6562](https://doi.org/10.2196/mhealth.6562)]
30. Zhai B, Perez-Pozuelo I, Clifton EA, Palotti J, Guan Y. Making sense of sleep: multimodal sleep stage classification in a large, diverse population using movement and cardiac sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020 Jun;4(2):1-33. [doi: [10.1145/3397325](https://doi.org/10.1145/3397325)]
31. de Zambotti M, Godino JG, Baker FC, Cheung JC, Patrick K, Colrain IM. The boom in wearable technology: cause for alarm or just what is needed to better understand sleep? *Sleep* 2016 Sep 01;39(9):1761-1762 [FREE Full text] [doi: [10.5665/sleep.6108](https://doi.org/10.5665/sleep.6108)] [Medline: [27397564](https://pubmed.ncbi.nlm.nih.gov/27397564/)]
32. Withings precision research challenge. *Stanford Medicine X*. URL: <https://medicinex.stanford.edu/precision-research-challenge/> [accessed 2021-12-31]
33. Maier AM, Bang MM, Özkil A, Forchhammer BH. Remember to remember: a feasibility study adapting wearable technology to the needs of people aged 65 and over with Mild Cognitive Impairment (MCI) and Alzheimer's dementia. In: *Proceedings of the International Conference on Engineering Design*. 2015 Presented at: International Conference on Engineering Design; Jul 27–30, 2015; Milan, Italy URL: <https://tinyurl.com/2p84k54s>
34. Edouard P, Campo D, Bartet P, Yang R, Bruyneel M, Roisman G, et al. Validation of the Withings Sleep Analyzer, an under-the-mattress device for the detection of moderate-severe sleep apnea syndrome. *J Clin Sleep Med* 2021 Jun 01;17(6):1217-1227. [doi: [10.5664/jcsm.9168](https://doi.org/10.5664/jcsm.9168)] [Medline: [33590821](https://pubmed.ncbi.nlm.nih.gov/33590821/)]
35. Bishop CM. *Pattern Recognition and Machine Learning: All "just the Facts 101" Material*. New Delhi, India: Springer; 2006.
36. Claude J. *Morphometrics with R*. New York: Springer; 2008.
37. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Per Med* 2011 Mar;8(2):161-173 [FREE Full text] [doi: [10.2217/pme.11.7](https://doi.org/10.2217/pme.11.7)] [Medline: [21695041](https://pubmed.ncbi.nlm.nih.gov/21695041/)]
38. To the EU Commission on DST. *European Sleep Research Society*. URL: https://esrs.eu/wp-content/uploads/2019/03/To_the_EU_Commission_on_DST.pdf [accessed 2021-12-31]

Abbreviations

- DST:** daylight saving time
REM: rapid eye movement
TST: total sleep time

Edited by A Mavragani; submitted 06.07.21; peer-reviewed by M de Zambotti, M Salimi, Y Zhang; comments to author 25.08.21; revised version received 17.10.21; accepted 28.11.21; published 22.02.22

Please cite as:

Óskarsdóttir M, Islind AS, August E, Arnardóttir ES, Patou F, Maier AM

Importance of Getting Enough Sleep and Daily Activity Data to Assess Variability: Longitudinal Observational Study

JMIR Form Res 2022;6(2):e31807

URL: <https://formative.jmir.org/2022/2/e31807>

doi: [10.2196/31807](https://doi.org/10.2196/31807)

PMID:

©María Óskarsdóttir, Anna Sigridur Islind, Elias August, Erna Sif Arnardóttir, François Patou, Anja M Maier. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 22.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.