SLEEPJ, 2021, 1–6

doi: 10.1093/sleep/zsaa141
Advance Access Publication Date: 27 July 2020
Original Article

ORIGINAL ARTICLE

A multitrait, multimethod matrix approach for a consumer-grade wrist-worn watch measuring sleep duration and continuity

Donald L. Bliwise[1,*], Christopher Chapple[2], Lena Maislisch[3], Eva Roitmann[4] and Teodor Burtea[3]

[1]Emory Sleep Center, Emory University School of Medicine, Atlanta, GA, [2]Royal Hallamshire Hospital, Sheffield, UK, [3]Ferring International Center, Saint-Prex, Switzerland and [4]Paris, France

*Corresponding author. Donald L. Bliwise, Emory Sleep Center, Emory University School of Medicine, Brain Health Center, Room 435, 12 Executive Park Drive, Atlanta, GA 30329. Email: dbliwis@emory.edu.

Abstract

Study Objectives: We examined associations between self-reports about typical sleep patterns and sleep data derived from a wearable device worn on a nightly basis for a prolonged period (mean = 214 nights). We hypothesized that sleep characteristics would correlate better across different methods of assessment (self-report versus wearable) than they would correlate within the same method, a classic psychometric approach (multitrait, multimethod matrix).

Methods: A cross-national sample of 6,230 adult wearable users completed a brief sleep questionnaire collecting data on sleep duration and number of awakenings (NAW) and provided informed consent to link their responses to data from their wearable watches. The data collection for the wearable occurred over 12 months and the sleep questionnaire was completed subsequent to that.

Results: Results indicated a large ($r = .615$) correlation between sleep duration as assessed with the wearable and by self-report. A medium-to-large correlation ($r = .406$) was also seen for NAW. The multitrait, multimethod matrix suggested minimal method variance, i.e. similar “traits” (sleep duration and NAW) correlated across methods but within a given method, and such “traits” were generally unrelated.

Conclusions: The results suggest that the longer period of data collection with the wearable generates more stable estimates of sleep than have been reported in most studies of actigraphy. Alternatively, the data might imply that individuals modify their self-reports about sleep via daily feedback to align their perceptions to the output of the wearable.

Statement of Significance

Millions of people use consumer-grade wearable devices to derive information about how they sleep. Most of these devices lack requisite validation; however, their widespread usage necessitates some examination of how well they perform. We examined correlations between self-reports about typical sleep duration and continuity in a large, cross-national data set of individuals whose usage exceeded on average 6 months of nightly use. The magnitude of the associations was somewhat larger than those typically reported over much shorter periods of time with research-grade actigraphy. These results may be interpreted as indicating that longer intervals of data collection generate more stable estimates of an individual’s sleep or, conceivably, that the users of wearable may modify their perceptions of their own sleep to match the output of the wearable device.

Key words: sleep duration; self-reports; wearables

Submitted: 30 January, 2020; **Revised:** 20 June, 2020

Introduction

Consumer-grade wearable devices now see ubiquitous use, particularly with applications in the domain of sleep. The Sleep Research Society (SRS) has recently published guidelines regarding not only the potential utility but also the substantial barriers for application of such devices to sleep and circadian science [1]. This statement indicated that the rate-limiting step in the implementation of wearables for population-based research has been and continues to be the absence of gold standard measurements of polysomnography as their validation. The range of different and often-competing manufacturers of such devices, as well as the seemingly constant shifting of industry-driven developments of hardware, firmware, and software, often serving a company's proprietary interests, makes for challenging implementation of research within this space. Yet, the prospects of data generated from wearable devices, particularly within the context of "big data," have intuitive appeal for sleep and chronobiology research at the population level.

We describe in this study the analyses of sleep-related data derived from a particular company's (Withings [formerly Nokia Health], Issy-les-Moulineaux, France: www.withings.com) wearable device and its relation to self-reports about typical sleep. Polysomnographic (PSG) data were not available, but a feature of these data was the relatively large number of nights of wearable data for each participant for the analyses. In order to examine the correlates of the wearable, we employed a classic psychometric approach: the multitrait, multimethod matrix [2].

Evidence evaluating the validity of wearable devices for the measurement of sleep varies widely in quality [3]. Withings manufactures a number of different wearables that generate sleep data. Published sleep studies using wearables by this manufacturer report on the "Pulse O2" [4–6], the "Pulse Activity tracker" [7], or simply the "Pulse" [8]. Relevant for the current work, these studies report only somewhat limited PSG validation or, in some studies, only research-grade actigraphic validation for total sleep duration [4–8], and even less so for time awake during the night [4, 6] or number of awakenings (NAW) [5], thus not meeting the proposed SRS standards [1] (see Discussion for further review).

Methods

Overview

The data reported here derived from a previously published observational study that focused on sleep issues in relation to nocturnal urination [9]. The study was conducted in a multinational population not specifically selected for urinary symptoms but selected on the basis of having purchased a particular brand of wearable device (see next paragraph). The study was sponsored by Ferring Pharmaceuticals (Saint-Prex, Switzerland) and data extracted with the support of the manufacturer of the wearable watches (Withings). Details related to questionnaire items dealing with nocturia and quality of life are presented elsewhere [9]. In this report, we focus here on the associations between self-reports regarding sleep and data about sleep derived from the wearables.

Participants

Participants at least 18 years of age (mean age = 47.4 [SD = 13.9] years; 59.3% male) were obtained from an international registry of users derived from a company that manufactures wrist-worn wearables that record data related to sleep. Upon purchasing a wearable unit, owners routinely are requested to register their unit with the company and initially indicate whether they would be willing to be approached electronically in the future about possible completion of unspecified surveys related to health issues. Purchases originated from the consumers, and neither Ferring nor Withings bought or otherwise provided wearables for these individuals. This may suggest that the population under study here had a relatively high degree of interest in self-monitoring their behaviors. The research followed the 2017 Code of Conduct Guidelines established by the European Pharmaceutical Market Research Association. For the nighttime urination survey (consistent with the aforementioned interest in nocturia), an additional, more specific informed consent was obtained via smartphone app for the wearables. Because the data were sampled internationally, not linked to any specific geographic site, and subject to multiple jurisdictions varying widely by country, formal Institutional Review Board approval was not obtained; however, handling of all personal identifiers and health care data was in accord with the principles of the Declaration of Helsinki. Informed consent was obtained for only those individuals who completed the nighttime questionnaire, as that was the only identifiable link in the data set to Personal Health Information. Data originated from a sample of wearable users encompassing 138,674 individuals (about 75% European, representing 12 countries) whose watch data were derived from the period of November 1, 2017 and October 31, 2018, all of whom then received a request on December 15, 2018 to complete the questionnaire. The questionnaire was completed between December 15, 2018 and January 15, 2019 by 7,141 individuals, representing a 5.15% response rate.

Sleep questionnaire and wearable data

Data from the following questionnaire items were analyzed in this study: (1) habitual sleep duration (HSD) ("How many hours do you on average sleep at night?") and (2) total number of awakenings (NAW) ("How many times do you wake up during a typical night of sleep for whatever reason?"). For HSD, free-field responses were converted to hours and minutes for analyses. For TNA, categorical response allowed were: 0, 1, 2, or 3 or more. Cases missing data for all items on the questionnaire (10-items total) were subject to listwise deletion, as were cases with ambiguous responses in the free field format (e.g. an HSD of "7h5," which could be interpreted as "7 hours and 5 minutes," "7 hours and 50 minutes," or "7 hours and 30 minutes"). Data cleaning for questionnaire responses occurred prior to further data analyses related to the wearables' data.

Three different versions of the Withings watch generated the users' data in this study: Steel HR watch (firmware 3351) (50% of cases), Steel watch (firmware 1690) (45% of cases), and Pulse Ox tracker (firmware 1761) (5% of cases). As a precondition for obtaining data, specific data were not aligned with a particular hardware/firmware model. The following sleep measures were derived from the wearable watches for each case for each night: (1) HSD and (2) NAW. For the current analyses, we used mean nightly data from individuals having at least five nights of data; however, most of these individuals used the watch for considerably longer periods of time, with a mean length of use of 214 nights. A frequency distribution showing the number of valid nights of usable data for those individuals is shown in Figure 1.

All versions of the Withings watch can detect (via proprietary algorithm) when the watch is not being worn and also determine sleep by use of an algorithm examining successive intervals with a lack of detectable movement. Nights with apparent sleep durations of greater than 20 hours were also excluded from logical consideration. Application of the exclusion criteria for both the questionnaire and wearables resulted in a total of 6,230 participants with data available for analyses.

Analyses

The multitrait, multimethod matrix is a classic approach to the psychometric evaluation of convergent and discriminant validity [2]. It presupposes that two different methods should correlate more strongly when measuring the same "trait" than when measuring different "traits" using a given method. Conversely, different "traits" measured within a given method should reflect partially method variance and correlate modestly with each other. We applied this fundamental approach to the examination of data reflecting different methods (self-reports about typical sleep and sleep metrics derived from wearables) and different "traits" (HSD and NAW). When referring to correlation coefficient values, we rely upon the terminology suggested by Cohen [10], who refers to *r* values of .10, .30, and .50, as small, medium, and large, respectively.

Results

The normative data for the study sample are shown in Table 1. The mean sleep duration from the wearables is nearly 40 minutes higher than indicated by self-report, and the median is nearly 25 minutes higher than indicated by self-report. The mean NAW based on the wearables was comparable to the NAW assessed via self-report.

Table 2 shows the multitrait, multimethod matrix of questionnaire and watch-derived sleep measures. The "traits" of sleep duration and NAW showed medium-to-large correlations across methods (r = .615 for HSD and r = .406 for NAW), though "traits" within a given method less so. Within data generated from the wearables, HSD showed only a small correlation with NAW.

To further understand whether the correlations between questionnaire and watch-derived sleep measures were influenced by the duration of nights of wearable use, we subdivided participants by the duration of their usage (Table 3). These results showed minimal impact in the HSD difference (computed as the difference of wearable minus questionnaire) as a function of length of usage, although the magnitude of the correlation between the two measures increased as the duration of usage became longer (r = .453 for 5–50 nights; r = .588 for 51–100 nights; r = .597 for 101–150 nights; r = .624 for 151–200 nights; r = .612 for 201–250 nights; r = .654 for 251–300 nights; and r = .635 for 301–367 nights). We tested for a difference in the magnitude of these correlations, using Cohen's q statistic, defined as $q = z_1 - z_2$, following r to z transform [10]. These analyses showed statistically computed medium-sized effects for differences between the correlations for 5–50 nights and for those with 301–367 (q = .261) and 251–300 (q = .293) nights of wearable use, respectively.

We also examined the wearable-minus-questionnaire difference as a function of mean HSD as recorded by the wearable (Table 4). Here, we found statistically significant relationships, suggesting that for participants sleeping less than 6 hours a night, the wearable underestimated their sleep duration relative to self-report. The largest positive difference (i.e. wearable overstating the questionnaire estimate) occurred for participants whose wearable-derived HSD was 8 hours or more.

Discussion

As in many other studies employing research-grade actigraphy [e.g. 11, 12], the mean HSD derived from the wearable used here was notably discrepant from self-reported HSD, the magnitude of the difference approaching 40 minutes. Relative to polysomnography, wearable devices usually tend to record longer sleep duration [1]. Our reference data were self-reports, not polysomnography, but still indicated that the wearable

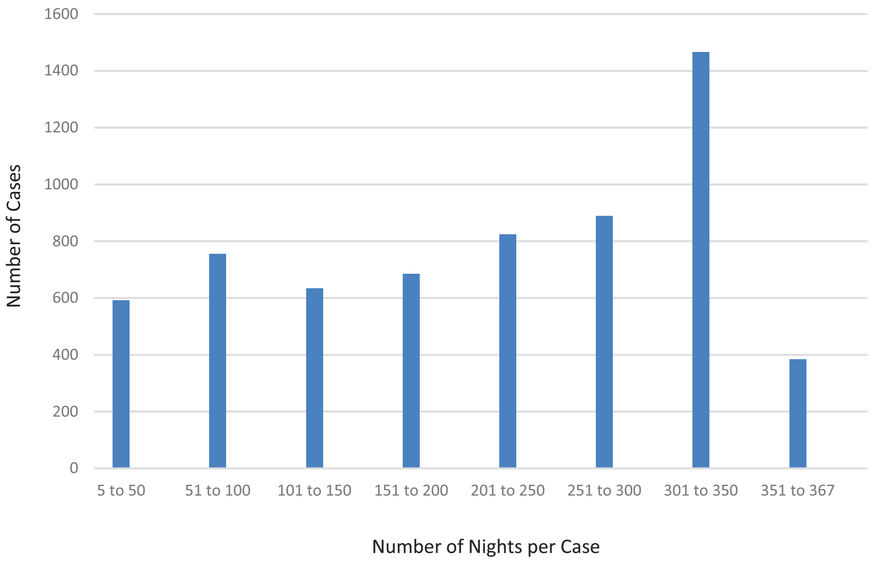

Figure 1. Frequency distribution of the number of usable nights per case derived from the wearables (N = 6,230). Numbers below each bar represent the number of nights per interval range.

Table 1. Descriptive data describing nightly sleep ($n = 6{,}230$)

Measure	Mean	Standard deviation	Median
HSD (wearable) (min)	442.1 min	46.1 min	444.6 min
NAW (wearable)	1.88 awakenings	0.95 awakenings	1.17 awakenings
HSD (self-report) (min)	404.3 min	59.5 min	420.0 min
NAW (self-report)	Frequency of awakenings	Number of cases	% of cases
	0–1 awakenings	3,064	49.2
	2 awakenings	2,092	33.6
	≥3 awakenings	1,074	17.2

Mean data for wearable based on all available nights.

Table 2. Multitrait, multimethod correlation matrix (Spearman correlations) between self-report and wearable-derived data

	Self-reported HSD	Self-reported NAW	Wearable-derived HSD	Wearable-derived NAW
Self-reported HSD	—	−.108	.615	−.065
Self-reported NAW		—	.031	.406
Wearable-derived HSD			–	.001
Wearable-derived NAW				–

R values > 0.0465 were statistically significant at $p < .0001$ (2-tailed) ($n = 6{,}230$).

Table 3. Within individual (signed) mean objective minus subjective differences in HSD in relation to number of nights of wearable use

Number of nights of wearable use	N	Mean of within individual (signed) mean objective minus subjective differences in HSD (in min)	Standard deviation of mean within individual objective minus subject differences in HSD (in min)
5–50	592	(+) 38.55	63.73
51–100	756	(+) 40.61	52.28
101–150	634	(+) 38.63	52.33
151–200	685	(+) 38.24	49.41
201–250	824	(+) 38.93	50.07
251–300	889	(+) 38.25	45.20
>300	1,850	(+) 35.39	42.79

One-way analysis of variance for category of the mean number of nights of wearable was nonsignificant ($F = 1.60$, $p = .16$).

Table 4. Within individual (signed) mean objective minus subjective differences in HSD in relation to mean habitual sleep duration based on wearable

Mean HSD from wearable (in min)	N	Mean of within individual (signed) mean objective minus subjective differences in HSD (in min)	Standard deviation of mean within individual objective minus subject differences in HSD (in min)
<300	28	(–) 30.45	91.39
300–359	264	(+) 19.24	47.22
360–419	1,513	(+) 29.39	44.92
420–479	3,178	(+) 37.73	45.96
≥480	1,247	(+) 53.87	56.14

One-way analysis of variance for the category of mean wearable-derived HSD was statistically significant ($F = 93.1$, $p < .0001$). All pairwise group contrasts were statistically significant using the Bonferroni-adjusted threshold (i.e. 0.05/10 = 0.005).

records greater self-reported HSD. The results also indicated that this was a graded association and most clearly apparent at longer HSDs than at shorter HSDs, where the opposite was true. The importance of the predominantly overestimation of sleep duration from the wearable relative to self-report is notable since self-reported sleep durations often exceed those seen with research actigraphy [11, 12] and cannot be considered a "gold standard" for evaluating a wearable. However, the importance of self-reported HSD cannot be discounted entirely, as such self-reports have been shown in many studies to be related to various morbidities and even mortality. For example, we have shown that even 30-minute incremental individual differences in self-reported estimates of HSD were relevant for glycemic control [13], suggesting that discrepancies of this magnitude of self-report may indeed be relevant for some aspects of health.

Perhaps the most surprising and unexpected aspect of these data was the large correlation between self-reported typical sleep duration and sleep duration as assessed with the wearable watch. The magnitude of this relationship is several orders of magnitude larger than the correlations reported between subjective appraisals of sleep duration and those recorded with objective measures in population-based work; although within the sleep laboratory, correlations are often large [14, 15], possibly because of the more delineated and less ambiguous boundaries for the start and end of the sleep period. In field studies examining associations between self-reported typical HSDs and actigraphically assessed HSDs over periods of 3 to 14 nights, correlations in the range of 0.30 [16], 0.43 [17], 0.45 [12], and 0.49 [18] are commonly encountered. As a statistical computation, the effect size (r^2) that we have demonstrated here (nearly 38% of the variance in reported sleep duration accounted for by HSD derived from the wearable) approaches nearly twice the magnitude of the variance accounted for in these studies. It is nearly 10 times the magnitude of the association between a one-night self-reported sleep duration and in-home PSG noted in an epidemiologic field study [19] in a population of comparable size to what we report here. Although one can postulate many possible reasons for the current larger magnitude of association that we have observed, two prime hypotheses appear likely.

The first explanation is that the much longer period of data collection coverage with the watch (Figure 1), with a mean exceeding 6 months, allows for a more accurate appraisal of an individual's customary amount of sleep. A large number of nights enhancing the reliability of estimates of sleep duration via actigraphy have many precedents in the current literature. Matthews et al. [20] reported that at least four nights of actigraphic measurements provided as stable an estimate of sleep duration as nine nights of actigraphy. However, other studies extending to 14 nights of data collection reported essentially that incremental stability occurred with every successive night of actigraphy [21], and yet, other work demonstrated the value of even longer periods of data collection (mean of 23 nights) when intermittent comorbidities (e.g. such as pain) may impact the duration of sleep [22]. Thus, in the current analyses, the prolonged period of data collection afforded by the wearable (Figure 1) may have yielded far more stable estimates of sleep duration, although, unlike this wearable, research-grade actigraphy does not allow easily obtainable feedback on a nightly basis to the user.

An alternative, more provocative, and indeed more subtle possibility is that by their very nature with providing the user daily feedback, the wearables might be postulated to be modifying the perceptions and experiences of the users. To borrow usage from social science [23], the device may serve an "effectance" function for the user, allowing them to modify and edit their beliefs about their own well-being (in this case, how much they sleep), until the user's perception is changed (perhaps irrevocably) so that they have learned what the wearable "teaches" them. What was once "intrinsic" [23] (the experience of sleep), thus becomes "extrinsic" (environmental determination via the wearable). This may have implications for future work. For example, studies demonstrating associations between self-reported short and long sleep durations and various adverse outcomes might need to take into account whether a respondent uses a wearable to track their sleep. Clinically, the implications of wearables are no less profound and perhaps even more difficult with which to deal. For example, one might envision scenarios in which it could become incumbent upon a clinician prescribing medication to improve sleep or a psychologist engaging in cognitive behavior therapy for insomnia to modify the output of the wearable to the satisfaction of the user to demonstrate longer (or better) sleep, rather than merely to alter the perception of the user that their sleep was longer (or better). As the availability of wearables increase and their cost decreases, such challenges may well confront sleep researchers and sleep clinicians in the future [24].

There are clearly a very large number of limitations associated with this study, involving both technical considerations and sampling issues. The study was limited to wearable data from a single company, and, furthermore, the hardware, firmware, and software that generated the data provided by the company for use in these analyses varied across the period of data collection. Although one might assume that various versions of hardware/software for a specific company would be more similar to earlier versions from that same company than when compared with those from a competitor, this is an unexamined presupposition. As a condition for obtaining these data, source code was not made available to the current investigators, thus leaving the potential for replicability of the data analyses in doubt.

In addition to these issues, only scant evidence suggests that sleep data from this particular manufacturer's wearables meet the gold standard of PSG-based validation, as recommended by the recent SRS report [1]. One study [7] reported on the use of the Withings watch to estimate sleep duration in a population of users and reported that 26.3% of over 15,000 users averaged 6 hours of sleep or less over a 7-day period but otherwise presented no concurrent validation for it. By contrast, Mantua et al. [6] demonstrated large ($r > .80$) correlations between PSG-based HSD and the Withings PulseO2 watch in 36 users, but that estimate was based on a single night in-home recording and was subject to a 10% technical failure rate. Additionally, the wearables were associated with a 6% absolute discrepancy (about 24 minutes) from PSG-based TST [6]. Another study [8] using research-grade, wrist-worn actigraphy as the gold standard employed a Withings wearable worn on the hip for a single night in 21 individuals and noted a large correlation ($r = .92$) for sleep duration but a wide range of differences from the gold standard (–17 minutes to +124 minutes). A study of 20 healthy young adults [4] undergoing a single night of simultaneous in-home PSG demonstrated a moderate-to-large association with the Withings product ($r = .48$), and a follow-up study from that same research group [5] in 22 patients with sleep apnea recorded in a clinical sleep laboratory showed a 9% technical failure rate with an absolute difference in HSD of 87 minutes. Across all of these studies, estimates for other sleep measures (e.g. time awake during the sleep period and NAW), when available at all, showed still larger discrepancies. Taken together, these data do not constitute rigorous scientific validation [1] of the watch as a measure of sleep duration or sleep continuity when used in population-based studies.

From the perspective of how this population was sampled, there are likely strong socioeconomic and demographic biases introduced by the fact that the data were obtained from individuals who had purchased the wearables for their personal use, which may have been driven by any number of financial, health, or other social considerations. Additionally, the response rate for completing the questionnaire was extremely low (about 5%) relative to traditional, survey-based research, suggesting that we may be examining associations with sleep in a highly selected subsample of the population, biased in ways that we cannot assess or fully appreciate. It may also be that individuals agreeing to complete the questionnaire were those individuals most concerned about urinary symptoms. Yet, another procedural limitation is that the subjective estimates of sleep were made retrospectively during a 1-month time frame that did not match the much longer window of exposure from which the watch data were generated. It remains unclear how retrospective judgment of an individual's habitual sleep patterns might be subject to unassessed recall bias. Finally, we had no specific assessments of sleep disorders, such as sleep apnea or a periodic limb movement disorder, which might be expected to impact the triaxial accelerometer that constitutes the Withings watch hardware. As mentioned previously, at least one small-scale study has shown that the PSG validation of sleep duration with this particular wearable becomes quite dubious in the presence of a common sleep disorder such as sleep apnea [5].

These may all constitute serious limitations for examining data derived from wearable devices worn by and (in all likelihood) trusted by hundreds of thousands, if not millions, of individuals [3]. Unquestionably, the wearables are aligned strongly with the proprietary interests of the companies that make them. However, to the extent that future scientific efforts to study sleep [1] and indeed the future practice of sleep medicine [24] may be influenced by the perceptions of sleep generated by such devices, it may be equally important and fundamentally essential to understand the meanings imbued to them by their users, perhaps even before the operational details of the instrumentation are fully understood or analyzed.

Funding

Conflict of interest statement. Financial Disclosures: D.L.B. is a consultant to Ferring, Merck, Huxley, and Eisai. C.C. is a consultant to Ferring, Astellas, Allergan, Bayer Schering Pharma AG, Contura, Galvani, Pierre Fabre, Symimetics, Taris Biomedical, and Urovant Sciences; has conducted Clinical Trials for Astellas and Ipsen; has been a speaker for Pfizer and Astellas; and holds a Patent with Symimetrics. L.M. is employed by Ferring. E.R. was previously employed by Withings and worked on this project as an independent researcher under contract with Ferring. T.B. is employed by Ferring. Non-Financial Disclosures: The authors report no non-financial disclosures.

References

1. Depner CM, *et al*. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep*. 2020;**43**(2):zsz254. doi:10.1093/sleep/zsz254 [Epub ahead of print]
2. Campbell DT, *et al*. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull*. 1959;**56**(2):81–105.
3. Baron KG, *et al*. Feeling validated yet? A scoping review of the use of consumer-targeted wearable and mobile technology to measure and improve sleep. *Sleep Med Rev*. 2018;**40**:151–159.
4. Gruwez A, *et al*. Reliability of commercially available sleep and activity trackers with manual switch-to-sleep mode activation in free-living healthy individuals. *Int J Med Inform*. 2017;**102**:87–92.
5. Gruwez A, *et al*. The validity of two commercially-available sleep trackers and actigraphy for assessment of sleep parameters in obstructive sleep apnea patients. *PLoS One*. 2019;**14**(1):e0210569.
6. Mantua J, *et al*. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors (Basel)*. 2016;**16**(5):E646.
7. Fagherazzi G, *et al*. An international study on the determinants of poor sleep amongst 15,000 users of connected devices. *J Med Internet Res*. 2017;**19**(10):e363.
8. Ferguson T, *et al*. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act*. 2015;**12**:42.
9. Chapple C, *et al*. Night-time voids, level of bother and sleep characteristics in a non-patient population of wearable devices users. *Int J Clin Pract*. 2020;**74**(7):e13495.
10. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Orlando, FL: Academic Press; 1977.
11. Van Den Berg JF, *et al*. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. *J Sleep Res*. 2008;**17**(3):295–302.
12. Lauderdale DS, *et al*. Self-reported and measured sleep duration: how similar are they? *Epidemiology* 2008;**19**(6):838–845.
13. Bliwise DL, *et al*. Habitual and recent sleep durations: graded and interactive risk for impaired glycemic control in a biracial population. *Am J Med*. 2017;**130**(5):564–571.
14. Lichstein KL, *et al*. Actigraphy validation with insomnia. *Sleep*. 2006;**29**(2):232–239.
15. Kushida CA, *et al*. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. *Sleep Med*. 2001;**2**(5):389–396.
16. Blackwell T, *et al*. Associations of objectively and subjectively measured sleep quality with subsequent cognitive decline in older community-dwelling men: the MrOS Sleep Study. *Sleep*. 2014;**37**(4):655–663.
17. Cespedes EM, *et al*. Comparison of self-reported sleep duration with actigraphy: results from the hispanic community health study/study of latinos Sueño Ancillary Study. *Am J Epidemiol*. 2016;**183**(6):561–573.
18. Landry GJ, *et al*. Measuring sleep quality in older adults: a comparison using subjective and objective methods. *Front Aging Neurosci*. 2015;**7**:166.
19. Silva GE, *et al*. Relationship between reported and measured sleep times: the Sleep Heart Health Study (SHHS). *J Clin Sleep Med*. 2007;**3**(6):622–630.
20. Matthews KA, *et al*. Similarities and differences in estimates of sleep duration by polysomnography, actigraphy, diary, and self-reported habitual sleep in a community sample. *Sleep Health*. 2018;**4**(1):96–103.
21. VAN Someren EJ. Improving actigraphic sleep estimates in insomnia and dementia: how many nights? *J Sleep Res*. 2007;**16**(3):269–275.
22. Kravitz HM, *et al*. An actigraphy study of sleep and pain in midlife women: the Study of Women's Health Across the Nation Sleep Study. *Menopause*. 2015;**22**(7):710–718.
23. Brewster Smith M. Competence and socialization. In: Clausen JA, ed. *Socialization and Society*. Boston, MA: Little, Brown and Company; 1968: 270–320.
24. Kirsch DB. Disruption in health care (and sleep medicine): "It's the End of the World as We Know it…and I Feel Fine.". *J Clin Sleep Med*. 2019;**15**(9):1185–1188.