Original Article

Multidimensional sleep health domains in older men and women: an actigraphy factor analysis

Meredith L. Wallace[1,*,], Lan Yu[2], Daniel J. Buysse[1,], Katie L. Stone[3,],
Susan Redline[4], Stephen F. Smagula[1], Marcia L. Stefanick[5],
Donna Kritz-Silverstein[6] and Martica H. Hall[1]

[1]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, [2]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, [3]California Pacific Medical Center Research Institute, San Francisco, CA, [4]Departments of Medicine, Brigham and Women's Hospital and Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, [5]Department of Medicine, Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA and [6]Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA

*Corresponding author. Meredith L. Wallace, Department of Psychiatry, University of Pittsburgh, 3811 O'Hara Street, Pittsburgh, PA 15213. Email: lotzmj@upmc.edu.

Abstract

The multidimensional sleep health framework emphasizes that sleep can be characterized across several domains, with implications for developing novel sleep treatments and improved prediction and health screening. However, empirical evidence regarding the domains and representative measures that exist in actigraphy-assessed sleep is lacking. We aimed to establish these domains and representative measures in older adults by examining the factor structure of 28 actigraphy-derived sleep measures from 2,841 older men from the Osteoporotic Fractures in Men Sleep Study and, separately, from 2,719 older women from the Study of Osteoporotic Fractures. Measures included means and standard deviations of actigraphy summary measures and estimates from extended cosine models of the raw actigraphy data. Exploratory factor analyses revealed the same five factors in both sexes: Timing (e.g. mean midpoint from sleep onset to wake-up), Efficiency (e.g. mean sleep efficiency), Duration (e.g. mean minutes from sleep onset to wake-up), Sleepiness/Wakefulness (e.g. mean minutes napping and amplitude of rhythm), and Regularity (e.g. standard deviation of the midpoint). Within each sex, confirmatory factor analyses confirmed the one-factor structure of each factor and the entire five-factor structure (Comparative Fit Index and Tucker–Lewis Index ≥ 0.95; Root Mean Square Error of Approximation 0.08–0.38). Correlation magnitudes among factors ranged from 0.01 to 0.34. These findings demonstrate the validity of conceptualizing actigraphy sleep as multidimensional, provide a framework for selecting sleep health domains and representative measures, and suggest targets for behavioral interventions. Similar analyses should be performed with additional measures of rhythmicity, other age ranges, and more racially/ethnically diverse samples.

Statement of Significance

Sleep is understood to be multidimensional; however, the exact domains captured through actigraphy are not established, leading to a lack of clarity regarding which aspects are relevant for clinical and research purposes. Using actigraphy measures from two epidemiologic samples of older adults, we empirically identified five domains (Timing, Efficiency, Duration, Alertness/Sleepiness, and Regularity) and determined which measures are most representative of each. This actigraphy framework could be used clinically to identify targets for behavioral interventions (e.g. improving specific sleep health deficits through "precision therapeutics") and can guide sleep health research by providing a template for variable selection, thereby enhancing reproducibility. Subsequent research should examine the multidimensionality of sleep in more diverse samples and with additional sleep measures and data types.

Introduction

In recent years, there has been a paradigmatic shift toward the study of sleep health, a "multidimensional pattern of sleep-wakefulness…that promotes mental and physical well-being" [1]. Research and practice based on a multidimensional sleep health framework emphasize that all individuals' sleep patterns can be characterized across several domains, regardless of the presence or absence of a sleep disorder. Thus, a multidimensional sleep health framework aims to promote healthy sleep for all individuals, rather than focusing on the relatively smaller proportion of the population with a diagnosed sleep disorder.

A key conceptual model for multidimensional sleep health is called Ru-SATED [1]. Ru-SATED refers to six domains of sleep health based on scientific and clinical rationale: (1) *Regularity*: the consistency of sleep timing, (2) *Satisfaction*: the subjective assessment of "good" or "poor" sleep (defined here by self-report only), (3) *Alertness/Sleepiness*: the ability to maintain attentive wakefulness during the day; (4) *Timing*: the placement of sleep within the 24-h day; (5) *Efficiency*: the ability to achieve consolidated sleep; and (6) *Duration*: the total amount of sleep obtained per 24 h. Rhythmicity has also been suggested as a potential seventh sleep health domain [2], defined as the robustness of the overall sleep–wake rhythm within a 24-h cycle. However, the exact number and type of domains have not been empirically established and may differ by data type (e.g. self-report, actigraphy, or polysomnography [PSG]) and/or population characteristics (e.g. age, sex, or race/ethnicity).

The Ru-SATED framework promotes the consideration of multidimensional sleep health for both clinical practice and research. Ru-SATED has already informed the development of new sleep treatments (e.g. the Transdiagnostic Sleep and Circadian Intervention [3]) and there have been numerous research studies that used an Ru-SATED framework to examine the importance of sleep health for health outcomes. For example, several research studies have developed ad hoc sleep health scores by counting the number of "potentially adverse" self-reported or behavioral sleep characteristics across Ru-SATED sleep health domains. Wallace et al. [2] measured overall sleep health in older men using either an actigraphy or self-report measure to represent each of the Ru-SATED domains plus sleep/wake rhythmicity. Brindle et al. [4] examined overall sleep health in men and women using actigraphy and daily diary Ru-SATED domains. Furihata et al. [5] measured overall sleep health in older women using self-reported measures reflecting SATED domains (i.e. without Regularity). These and several other recent applications [6–11] of Ru-SATED all employ slightly different representative measures and/or domains, making it difficult to assess the reproducibility of findings or converge on the importance of sleep health and its components within and across studies and health outcomes.

Because of the growing interest in a multidimensional sleep health framework for both clinical and research use, it is critical to formally establish which domains exist empirically and which measures are most representative of each domain. This information will promote the consideration of sleep health as a multidimensional construct, guide treatment development, and provide a template for selecting actigraphy measures in research, thereby helping to standardize the methods of the growing number of studies utilizing the Ru-SATED conceptual model. Our primary objective is to examine the factor structure of actigraphy-derived sleep health measures in two large epidemiological cohorts of older men and women, respectively: the Osteoporotic Fractures in Men (MrOS) Sleep study [12, 13] and the Study of Osteoporotic Fractures (SOF) [14, 15]. Based on Ru-SATED, prior research suggesting sleep–wake rhythmicity may be a separate actigraphy domain in older adults [3], and increased consideration of rest–activity rhythms for health research [16–19], we hypothesized that the actigraphy data would reveal six domains reflecting measures of rhythmicity, regularity, alertness/sleepiness, timing, efficiency, and duration. Satisfaction was not hypothesized as it is not amenable to measurement by actigraphy.

Methods

Participants

SOF was originally designed to determine risk factors for osteoporotic fractures in community-dwelling older women (https://sofonline.epi-ucsf.org) [14, 15]. MrOS was originally designed to assess risk factors for osteoporotic fractures in community-dwelling older men (http://mrosdata.sfcc-cpmc.net) [12, 13]. Participants in both studies provided written informed consent to participate in longitudinal studies of sleep health. Study design differences in MrOS and SOF precluded analyzing the two full samples simultaneously. Most notably, the SOF actigraphy data were collected in year 16 of the study ("Visit 8"), whereas the MrOS actigraphy data were collected in year 3 of the study ("Sleep Visit 1"). This resulted in SOF women being several years older than MrOS men on average. Therefore, to avoid confounding of age and sex, we examined the two samples separately in all analyses.

Participants self-reported sociodemographic and clinical information via a questionnaire. This included history of physician-diagnosed medical conditions, health behaviors (e.g. alcohol and tobacco use), self-rated health, measures of physical functioning, depressive symptoms (Geriatric Depression Scale [20]), and anxiety symptoms (Goldberg Anxiety Index [21]). During an in-clinic visit, all self-reported responses were reviewed with an examiner. Cognition was measured using the Teng 3S Modified Mini Mental State Exam [22] and Mini Mental State Exam [23] (MMSE) in MrOS and SOF, respectively. For comparability, we report a 26-item modified MMSE score that includes the overlapping items from these two measures. Participants were also asked to bring in all current medications used within the preceding 30 days. All prescription and nonprescription medications were entered into an electronic database and each medication was matched to its ingredient(s) based on the Iowa Drug Information Service Drug Vocabulary (College of Pharmacy, University of Iowa, Iowa City, IA) [24].

SOF and MrOS participants were instructed to wear a Sleepwatch-O actigraph (Ambulatory Monitoring, Inc, Ardsley, NY) on their wrist for a full 4 days. However, in some situations valid actigraphy data were not captured (e.g. if the participant removed the actigraph), resulting in less than 4 days of valid data. Because a minimum of 3 days is required to compute some of the actigraphy measures, we included participants in our analytic samples if they had at least three valid "in-bed" and three valid "out-of-bed" intervals. These criteria resulted in $n = 2,841$ men in the MrOS analytic sample (92.9% of $n = 3,058$ with any actigraphy) and $n = 2,719$ women in the SOF analytic sample (87.0% of $n = 3,127$ with any actigraphy). The MrOS analytic sample had a range of 3–13 days of actigraphy (median of 5 days). The SOF analytic sample had a

range of 3–9 days of actigraphy (median of 4 days). Demographic and clinical characteristics of the analytic samples are provided in Table 1. Full details regarding the derivations of the analytic samples are provided in Supplemental Figure S1.

Actigraphy processing

The Sleepwatch-O continuously monitors acceleration and stores the signal in 1-min epochs using three different storage techniques [25–27]. The data analyzed herein were stored using Proportional Integration Mode, a high-resolution measurement of the area under the curve. Action W-2 software (Ambulatory Monitoring, Inc.) and the University of California, San Diego scoring algorithm [28] were then applied to automatically score each epoch as "sleep" or "wake". Trained

actigraphy scorers at the San Francisco Coordinating Center further edited and scored the data using each participant's sleep diary to mark the times when the participant got in and out of bed and/or indicate times when the record should be deleted because the participant removed the watch. These markers, in conjunction with the scored "sleep" and "wake" epochs, facilitated the computation of key daily/nightly indices of sleep (e.g. sleep onset/offset time, minutes awake after sleep onset, and duration and frequency of napping). These daily/nightly indices were then summarized using means and standard deviations (SDs) across the monitoring period. Parametric extended cosine models [29] were also applied to the epoch-by-epoch data to produce circadian rest-activity rhythm variables that summarize other types of temporal features of the monitoring period.

Table 1. Clinical and sociodemographic characteristics of the analytic samples

	Females (N = 2,719)	Males (N = 2,841)
Age, mean (standard deviation)	83.53 (3.68)	76.32 (5.5)
Race, % (n)		
White	91.39 (2,485)	90.21 (2,563)
Non-White	8.61 (234)	9.79 (278)
Education, % (n)		
≥College education	66.46 (1,807)	56.28 (1,599)
HS or some college education	19.49 (530)	38.40 (1,091)
<HS education	14.05 (382)	5.32 (151)
Marital status, % (n)		
Married	27.07 (736)	84.55 (2,402)
Widowed	61.35 (1,668)	7.60 (216)
Other	11.58 (315)	7.85 (223)
Smoking status, % (n)		
Current	2.69 (73)	1.94 (55)
Past	33.14 (900)	58.70 (1,667)
Never	64.21 (1,746)	39.39 (1,119)
Any alcohol use, % (n)	41.72 (1,133)	34.24 (968)
Body mass index, mean (standard deviation)	26.86 (4.45)	27.18 (3.79)
Anxiety symptoms (GADS),* mean (standard deviation)	2.37 (2.6)	0.97 (1.89)
Depressive symptoms (GDS),† mean (standard deviation)	1.34 (2.2)	1.77 (2.16)
Cognition (26-item mMMSE),‡ mean (standard deviation)	24.26 (1.98)	24.23 (1.97)
Cognitive impairment (26-item mMMSE‡ < 21), % (n)	5.50 (137)	4.61 (131)
Self-rated health, % (n)		
1 (Excellent)	19.62 (533)	33.77 (959)
2 (Good)	56.52 (1,535)	53.17 (1,510)
3 (Fair)	21.76 (591)	11.87 (337)
4 (Poor/very poor)	2.10 (75)	1.2 (34)
Number of instrumental activities of daily living that could not be performed, % (n)		
0	53.26 (1,415)	69.39 (1,970)
1	18.74 (498)	18.56 (527)
2	11.82 (314)	6.73 (191)
3	8.51 (226)	3.80 (108)
4	7.68 (204)	1.51 (43)
Number of chronic conditions§ % (n)		
0	18.3 (480)	29.58 (840)
1	36.26 (951)	35.35 (1,004)
2	25.89 (679)	21.41 (608)
3	11.67 (306)	8.98 (255)
≥4	7.89 (207)	4.68 (133)
Total number of prescription medications, mean (SD)	4.49 (3.05)	3.86 (3.02)

*Goldberg Anxiety Index.

†15-item Geriatric Depression Scale.

‡26-item Modified Mini Mental State Exam (mMMSE). A score < 21 corresponds to an MMSE < 24.

§The following conditions were considered: stroke, angina, heart failure, heart attack, high blood pressure, diabetes, chronic obstructive pulmonary disease, osteoporosis, and arthritis.

Table 2. Actigraphy variables considered in the factor analysis

Variable name	Variable description
Number of naps	Mean of number of naps per day of duration ≥5 min
Minutes napping	Mean of minutes napping per day, considering only naps ≥5 min
Mesor*	The middle of the peak, computed as Minimum + Amplitude/2
Pseudo-F*	Pseudo-F goodness of fit statistic from ECM. Higher values indicate better conformity to extended cosine curve
Amplitude*	Amplitude from the extended cosine model
Beta*	Determines whether the function rises and falls more steeply than the cosine curve. Large values produce nearly square curves
Alpha*	Width of peaks relative to troughs from ECM. Large values indicate the peaks are narrow and the troughs are wide; small values indicate the peaks are wide and the troughs are narrow
Time in bed	Mean of minutes from bed to wake-up time
Time from onset to wake-up (TOW)	Mean of minutes from sleep onset to wake-up time
TST	Mean of minutes of sleep from bedtime to wake-up time
Bedtime	Mean of bedtime
Sleep onset time	Mean of sleep onset time
Midpoint (bed interval)	Mean of midpoint of bed to wake-up time
Midpoint (onset interval)	Mean of midpoint of sleep onset to wake-up time
Wake-up time	Mean of wake-up time
Acrophase*	Time of maximum activity from the extended cosine model
Up-mesor*	Time of switch from low to high activity (above to below mesor) from the ECM
Down-mesor*	Time of switch from high to low activity (below to above mesor) from the ECM
Wake after sleep onset	Mean of minutes awake after sleep onset
Sleep efficiency	TST/TIB × 100
Sleep maintenance	TST/TOW × 100
Sleep latency	Mean of minutes from bed to sleep onset time
Minimum*	Minimum value of activity function from ECM
SD wake-up time	Standard deviation of wake-up time
SD midpoint (onset interval)	Standard deviation of midpoint of sleep onset to wake-up time
SD midpoint time (bed interval)	Standard deviation of midpoint of bed to wake-up time
SD sleep onset	Standard deviation of sleep onset time
SD bedtime	Standard deviation of bedtime

*Computed from extended cosine model (ECM).

Variable selection and coding

Although actigraphy scoring algorithms produce sleep–wake characteristics for each day and night, we focused on measures that summarize information across the observation period because they allow examining features of habitual sleep that align with our hypothesized domains (e.g. rhythmicity and regularity). To date, these summary measures have been most commonly used in sleep health research, largely because of their interpretability and ability to be used as predictors of key prospective health outcomes.

We developed an initial pool of 28 actigraphy summary variables that: (1) provided a thorough representation of actigraphy variables currently used in research for sleep health in older adults; (2) were clinically and/or scientifically meaningful; and (3) were available to us in the MrOS and SOF samples. When uncertain as to whether to include a variable, we erred on the side of being over-inclusive. Because obtaining statistical convergence on large factor structures can be challenging, we coded variables as ordinal (five categories) based on quintiles from the respective MrOS or SOF sample. This approach facilitated convergence by standardizing the range of variables and reducing the effects of extreme outliers and/or highly skewed distributions, all of which are common in actigraphy data [30]. Table 2 lists and describes the selected actigraphy measures. Supplemental Table S1 displays the quintiles for each actigraphy variable in each sample.

Factor analysis methods

We randomly split each sample into two subsamples: one for exploratory factor analysis (EFA) and the other for confirmatory factor analysis (CFA). EFA and CFA were conducted using R version 3.5.1. For EFA we used the fa function within the psych package and for CFA, we used the cfa function within the lavaan package.

For EFA, we used polychoric correlations, a Promax rotation, and median imputation (default within the R function) to handle a very small amount of missing actigraphy data (0.7% in MrOS, 2.4% in SOF). No specific domains were imposed on the data. To determine the number of factors, we examined the eigenvalues, visually inspected scree plots, and considered our hypotheses. We empirically assigned variable to a factor if it had a factor loading >0.40 [31]. Because our main interest was in understanding the factor structure, if a variable loaded on multiple factors, we made the primary assignment to the factor on which it had the highest loading and noted any secondary assignments in our interpretation.

We then performed CFAs for ordered data on the independent male and female confirmatory samples to validate the factor structure identified in the EFAs. As the CFA methods required complete data samples, we omitted a small number of cases with missing data (0.9% in MrOS; 3.3% in SOF). We fit two types of CFAs for each sex: (1) separate CFA models for each factor to determine whether a single factor structure was present and (2) a full-model CFA to determine whether the multifactor

structure identified in the EFA would hold and to examine correlations among factors. For each CFA model, we examined the Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA) to assess goodness of fit [32, 33]. If a variable did not have a loading >0.40 on any factor in the CFA, it was removed and the models were rerun without it.

Finally, to further assess factor cohesion, we examined spearman correlations of the original continuous variables within each identified factor. Correlations were examined separately in the full MrOS and SOF analytic samples.

Results

Sample partition

We randomly split each of the male and female samples into two non-overlapping subsamples for EFA (n = 1,359 for female; n = 1,420 for male) and CFA (n = 1,360 for female; n = 1,421 for male). Within the CFA samples, 96.7% of females (n = 1,315) and 99.1% of males (n = 1,408) had complete actigraphy data for use in the CFA analysis, which does not allow missing observations. (See Supplemental Figure S1.)

Exploratory factor analysis

The patterns of eigenvalues were nearly identical for the female and male samples (Figure 1). After examining the eigenvalues, we chose to more closely inspect the 4-, 5-, 6-, and 7-factor structures based on three main observations: (1) there was clearly a multiple-factor solution as evidenced by the fact that the ratio of the first and second eigenvalue was >4; (2) the steep decrease in the first four eigenvalues suggested that two or three-factor solutions were not upheld by the data; and (3) the first 7 factors had eigenvalues larger than 1 for both samples (range of first 7 eigenvalues = 6.08–1.18 for females and 6.15–1.36 for males). After inspecting the results of the 4-, 5-, 6-, and 7-factor structures, we selected the 5-factor structure because it provided the most interpretable scenario, showed clear and consistent factor

Figure 1. Scree plot displaying eigenvalues for the MrOS and SOF samples.

loadings, and largely aligned with our hypothesized structure. Because all 28 actigraphy variables had loadings > 0.40 in the 5-factor structure for both samples, we retained all 28 variables.

Table 3 displays the 5-factor structures for the EFAs in the male and female samples. Across samples, the 28 actigraphy variables were grouped identically, the only difference being the order of Factors 3 and 4. Therefore, for interpretability and consistency across samples, we reordered Factors 3 and 4 in the results for men. When selecting labels for the factors, we considered factor loadings, interpretation, and our prior hypotheses based on the Ru-SATED domains.

Factor 1: Timing (8 variables)
All measures reflected either the means of sleep timing variables (e.g. bed, sleep onset, and wake-up) or timing-related estimates from the extended cosine model (acrophase, up-mesor, and down-mesor).

Factor 2: Efficiency (5 variables)
Four of the five measures reflected either minutes awake at night (minutes awake after sleep onset [WASO] and sleep latency) or the percentage of time spent asleep (e.g. sleep efficiency). The minimum from the extended cosine model also loaded on this factor, which is consistent with the factor interpretation because an individual with very efficient sleep and/or very little time awake at night would tend to have a low estimated minimum of the rest/activity function. For both males and females, total sleep time (TST) had a primary loading on Factor 3 (Duration) but had secondary loadings on Factor 2 (loadings of –0.44 for males and –0.49 for females). This finding reflects the relationship between TST and WASO in community-dwelling older adults whose sleep schedules were not constrained by study methods. It suggests a continuum of sleep propensity: as sleep increases, wakefulness decreases, and vice versa. However, the relationship between TST and WASO can vary depending upon study or participant characteristics (e.g. if a behavioral treatment/intervention [34] is employed that restricts or extends TIB).

Factor 3: Duration (4 variables)
Three of the four measures related to the durations of sleep and/or time spent in bed. The alpha parameter from the extended cosine model also loaded on this factor. In a well-fitting extended cosine model (i.e. with a high pseudo-F), large values of alpha indicate that the peaks of the rest/activity rhythm are narrower than the troughs. An individual with longer sleep duration would tend to have wider troughs (times of inactivity) relative to their peaks (times of activity) and thus a larger alpha. Some measures with primary assignments to the timing factor loaded secondarily on the duration factor, albeit with relatively low factor loadings. These included wake-up time and up-mesor in both samples (loadings of 0.44 for males and 0.46 for females) and bedtime in the male sample (loading of –0.42). As earlier bedtimes and later wake-up-times tend to be related to longer duration in an unconstrained sleep setting, their secondary loadings within this factor are consistent with its interpretation.

Factor 4: Alertness/sleepiness (5 variables)
This factor included measures we originally hypothesized to load on an Alertness/Sleepiness domain (number of naps per day, minutes napping per day, and mesor from the extended
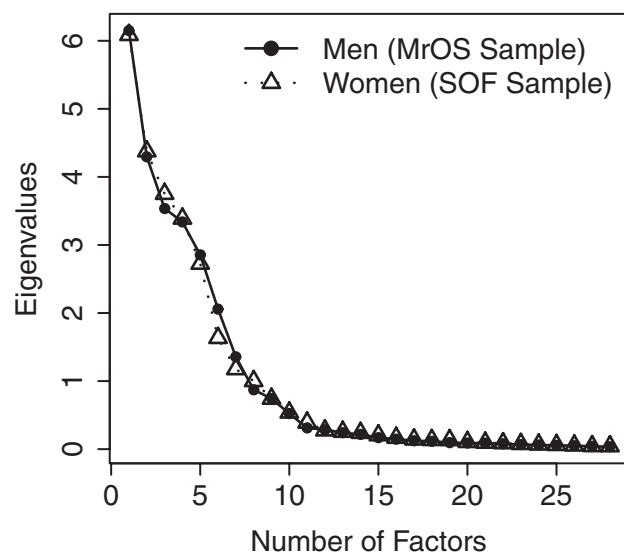
Table 3. Factor structure from separate exploratory factor analyses of females and males

	Females					Males				
	F1	F2	F3	F4	F5	F1	F2*	F3*	F4	F5
Factor 1: Timing										
Midpoint (Bed to Wake-Up)	0.87	–0.07	0.01	0.03	0.01	0.88	–0.04	0.04	0.03	0.04
Midpoint (Sleep Onset to Wake-Up)	0.87	0.06	0.03	0.03	0.03	0.86	0.09	0.05	0.05	0.06
Acrophase†	0.84	0.01	0.03	–0.07	–0.06	0.84	–0.02	0.01	–0.12	–0.07
Wake-Up Time	0.80	0.03	0.46	0.04	0.01	0.82	0.03	0.44	0.04	0.01
Up-Mesor†	0.73	0.04	0.46	–0.02	0.08	0.75	0.04	0.44	–0.03	0.10
Sleep Onset Time	0.70	0.06	–0.39	0.00	0.08	0.71	0.09	–0.33	0.06	0.08
Down-Mesor†	0.67	–0.03	–0.28	–0.04	–0.13	0.62	–0.05	–0.28	–0.09	–0.16
Bedtime	0.67	–0.17	–0.42	0.04	0	0.72	–0.12	–0.35	0.04	0.04
Factor 2: Efficiency										
Sleep Maintenance	0.01	–0.81	0.01	0.02	–0.01	–0.01	–0.84	0.09	0.06	–0.04
Sleep Efficiency	–0.02	–0.8	0.11	0.07	–0.09	–0.01	–0.83	0.06	0.06	0.00
Wake After Sleep Onset	–0.02	0.77	0.19	0.00	0.00	0.04	0.82	0.13	–0.08	–0.02
Minimum†	–0.06	0.66	–0.02	0.18	–0.11	–0.05	0.51	0.08	0.32	–0.17
Sleep Latency	0.16	0.45	–0.01	–0.11	0.14	0.05	0.54	0.03	0.01	0.11
Factor 3: Duration										
Time from Sleep Onset to Wake-Up	0.03	–0.03	0.84	0.01	–0.09	0.05	–0.08	0.80	–0.03	–0.10
Time from Bed to Wake-Up	0.08	0.23	0.79	–0.03	0.01	0.08	0.16	0.81	–0.02	–0.05
Total sleep time	0.02	–0.44	0.73	0.00	–0.08	0.05	–0.49	0.65	0.01	–0.09
Alpha†	–0.12	0.08	0.57	0.02	0.16	–0.13	0.07	0.55	0.07	0.19
Factor 4: Alertness/Sleepiness										
Amplitude†	–0.06	–0.18	–0.07	–0.75	0.09	–0.06	–0.07	–0.06	–0.75	0.13
Number of Naps per Day	–0.08	–0.09	–0.05	0.75	0.11	–0.06	–0.12	0.01	0.74	0.12
Pseudo-F†	–0.06	–0.08	0.05	–0.74	–0.14	–0.05	–0.06	0.13	–0.71	–0.13
Minutes Napping per Day	–0.06	–0.08	–0.07	0.73	0.11	–0.08	–0.11	–0.03	0.73	0.14
Mesor†	–0.08	0.24	–0.12	–0.68	0.05	–0.1	0.28	–0.05	–0.62	0.07
Factor 5: Regularity (SD = Standard Deviation)										
SD Midpoint (Bed to Wake-Up)	0.02	–0.05	–0.02	0.09	0.76	0.06	–0.09	–0.05	0.05	0.78
SD Midpoint (Sleep Onset to Wake-Up)	0.02	0.09	–0.04	0.06	0.73	0.05	0.09	–0.08	0.06	0.75
SD Bedtime	–0.08	–0.02	–0.06	0.02	0.64	–0.02	–0.05	–0.01	0.03	0.66
SD Sleep Onset Time	–0.03	0.23	–0.05	0.01	0.62	0.00	0.23	–0.03	0.06	0.61
SD Wake-Up Time	0.12	–0.03	–0.02	0.08	0.59	0.07	–0.02	–0.04	0.02	0.61
Beta†	0.00	0.30	–0.08	0.27	–0.49	0.05	0.16	–0.07	0.31	–0.46

*F3 and F4 were reversed for consistent interpretation with females.

†Computed using extended cosine model.

cosine model) as well as variables, we hypothesized to load on a separate rhythmicity domain (the pseudo-F statistic and amplitude). Although we considered both Alertness/Sleepiness and Rhythmicity to be reasonable labels for Factor 4, we ultimately selected Alertness/Sleepiness because all of the measures reflect this dichotomy to some degree: mesor represents the middle of the sleep–wake peak, amplitude is the difference between sleep and wake, pseudo-F reflects the level of separation of sleep and wake, and duration/frequency of napping reflects the fragmentation of sleep–wake periods throughout the day. The Alertness/Sleepiness label also better reflects the Ru-SATED framework, which was the motivation for this work.

Factor 5: Regularity of sleep timing (6 variables)
This factor included the standard deviations of sleep-related timing measures and the beta parameter from the extended cosine model. Beta reflects whether the function rises and falls more steeply than the cosine curve, with large values producing curves that are nearly square waves. In a well-fitting model with a high pseudo-F, waves that are more square-like could indicate a faster (i.e. more regular) shift between periods of "sleep" and "wake".

Finally, because we originally hypothesized six factors, we also examined the 6-factor structure closely. In both samples,

it suggested the same five factors as in the five-factor structure, plus a sixth factor that included only alpha and down-mesor. Therefore, the five-factor structure provided a clearer and more cohesive interpretation.

Confirmatory factor analysis

Prior to performing CFAs, we reverse coded variables so that interpretations were consistent within each factor and, where relevant, higher values would reflect better sleep health. All CFAs were performed on the independent CFA samples. Across CFAs, all variables loaded well on their corresponding factors for both samples except beta, with loadings <0.40. Therefore, beta was removed from Factor 5 and the CFAs were rerun. The final CFA loadings within each factor were very similar across samples and CFAs. CFA results are provided in Table 4.

The correlations among the 5 factors from the CFA model are displayed in Table 5. The magnitudes of most correlations were small (<0.30), indicating relatively independent domains, and were largely comparable across male and female samples. For women, only two correlation magnitudes were ≥0.30 (Efficiency and Regularity, $r = 0.34$; Timing and Duration, $r = –0.30$). For men, only Efficiency and Duration had a magnitude correlation ≥0.30 ($r = 0.32$).

Table 4. Confirmatory factor analysis results from males and females

	Female		Male	
	All factors on single run	Each factor on separate run	All factors on single run	Each factor on separate run
Factor 1: Timing				
Midpoint (Bed to Wake-Up)	0.95	0.95	0.97	0.97
Midpoint (Sleep Onset to Wake-Up)	0.96	0.96	0.97	0.97
Acrophase†	0.86	0.87	0.86	0.86
Wake-Up Time	0.90	0.91	0.90	0.9
Up-Mesor†	0.88	0.89	0.86	0.87
Bedtime	0.88	0.88	0.92	0.92
Sleep Onset Time	0.91	0.90	0.92	0.92
Down-Mesor†	0.85	0.81	0.82	0.80
Factor 2: Efficiency				
Sleep Efficiency	0.91	0.91	0.96	0.96
Sleep Maintenance	1.01	1.01	1.01	1.01
WASO‡	0.95	0.95	0.96	0.96
Sleep Latency‡	0.59	0.54	0.57	0.53
Minimum†‡	0.52	0.53	0.42	0.45
Factor 3: Duration				
Time from Sleep Onset to Wake-Up	1.01	1.03	1.00	1.04
Time from Bed to Wake-Up	0.81	0.83	0.84	0.85
Total sleep time	0.80	0.78	0.88	0.78
Alpha†	0.72	0.55	0.52	0.42
Factor 4: Alertness/Sleepiness				
Amplitude†	0.89	0.88	0.81	0.82
Number of Naps per Day‡	0.95	0.95	0.95	0.95
Minutes Napping per Day‡	0.94	0.94	0.93	0.93
Pseudo-F†	0.80	0.79	0.77	0.75
Mesor†	0.79	0.80	0.71	0.72
Factor 5: Regularity				
SD Midpoint (Bed to Wake-Up)‡	0.85	0.87	0.93	0.94
SD Midpoint (Sleep Onset to Wake-Up)‡	0.92	0.91	0.93	0.92
SD Bedtime‡	0.68	0.71	0.81	0.82
SD Sleep Onset Time‡	0.81	0.78	0.86	0.84
SD Wake-Up Time‡	0.65	0.63	0.70	0.70

*F2 and F3 were switched for consistent interpretation with females.
†Computed using extended cosine model.
‡Reverse coded.

Table 5. Correlations among factors

		Females				
		Timing	Efficiency	Duration	Alertness/Sleepiness	Regularity
Males	Timing	–	−0.07 (0.01)	−0.30 (<0.01)	0.03 (0.29)	−0.12 (<0.01)
	Efficiency	−0.07 (0.01)	–	0.11 (<0.01)	0.10 (<0.01)	0.34 (<0.01)
	Duration	−0.22 (<0.01)	0.32 (<0.01)	–	−0.06 (0.03)	−0.01 (0.86)
	Alertness/Sleepiness	−0.03 (0.30)	0.01 (0.64)	0.02 (0.50)	–	0.06 (0.03)
	Regularity	−0.17 (<0.01)	0.27 (<0.01)	0.07 (0.01)	0.09 (<0.01)	–

Cells show estimated correlation between factors (*p*-value).

CFA fit indices are displayed in Table 6. For both the single- and multifactor CFAs across men and women, the CFIs and TLIs were ≥0.95, indicating that the average correlation among variables on each factor was sufficiently high [32, 33]. The RMSEA ranged from 0.08 to 0.38 in the single-factor CFAs and was 0.22 for both male and female multifactor CFAs. Although these RMSEA values are outside of the range of what is typically considered a "good fit" (RMSEA < 0.06) [32, 33], they are consistent with RMESA values observed in other studies of multidimensional sleep [35].

The CFA factor loadings in Table 4 indicate the representativeness of each measure for its respective factor. For Factor 1 (Timing), the mean midpoint from bed to wake and mean midpoint from sleep onset to wake-up had the highest loadings across samples (0.95–0.97). For Factor 2 (Efficiency), mean sleep maintenance had the highest loading across samples (1.00–1.02). For Factor 3 (Duration), mean time from sleep onset to wake-up had the highest loadings across samples (1.00–1.04). For Factor 4 (Alertness/Sleepiness), mean minutes napping per day and mean number of naps per day had the highest loadings

Table 6. Fit statistics from CFAs

	Female			Male		
	CFI	TLI	RMSEA (90% CI)	CFI	TLI	RMSEA (90% CI)
Factor 1: Timing	0.98	0.98	0.36 (0.35, 0.37)	0.96	0.98	0.37 (0.36, 0.38)
Factor 2: Efficiency	1.00	1.00	0.16 (0.14, 0.18)	1.00	1.00	0.09 (0.07, 0.11)
Factor 3: Duration	1.00	1.00	0.08 (0.05, 0.11)	1.00	1.00	0.08 (0.05, 0.11)
Factor 4: Alertness/Sleepiness	0.98	0.97	0.33 (0.31, 0.35)	0.97	0.95	0.38 (0.36, 0.40)
Factor 5: Regularity	0.98	0.96	0.22 (0.20, 0.24)	0.99	0.97	0.26 (0.24, 0.28)
Full 5-Factor Model	0.96	0.95	0.22 (0.22, 0.22)	0.97	0.97	0.22 (0.21, 0.22)

CFI, comparative fit index; TLI, Tucker–Lewis Index; RMSEA, root mean square error of approximation; CI, confidence interval.

across samples (0.93–0.95). For Factor 5 (Regularity), the *SD* of the midpoint from sleep onset to wake-up had the highest loading across samples (0.92–0.93) and the *SD* of the midpoint from bed to wake-up had high factor loadings in the male sample (0.93–0.94). A total of 7 factor loadings across all CFA analyses ranged from 1.01 to 1.04. Loadings >1 indicate an overfitted item, likely due to redundant information among the variables. However, we retained these measures in the model because of their content importance and our focus on understanding the factor structure of actigraphy rather than developing an instrument.

Correlations within factors

As an additional check of factor cohesion, we computed Spearman correlations among the original continuous variables within each factor using the full male and female samples. The majority of pairwise correlations were moderate-to-large. For males, the median (minimum, maximum) Spearman correlations were 0.70 (0.22, 0.96) for Factor 1 (Timing), 0.44 (0.22, 0.96) for Factor 2 (Efficiency), 0.52 (0.28, 0.88) for Factor 3 (Duration), 0.59 (0.36, 0.91) for Factor 4 (Alertness/Sleepiness), and 0.70 (0.30, 85) for Factor 5 (Regularity; excluding Beta). For females, they were 0.71 (0.26, 0.94) for Factor 1 (Timing), 0.53 (0.21, 0.96) for Factor 2 (Efficiency), 0.55 (0.38, 0.84) for Factor 3 (Duration), 0.65 (0.46, 0.90) for Factor 4 (Alertness/Sleepiness), and 0.62 (0.22, 0.76) for Factor 5 (Regularity; excluding Beta). Within each sex and for a given factor, no more than two pairs of measures had correlation magnitudes <0.30. Full correlation matrices are provided in Supplemental Tables S3–S7.

Discussion

We empirically identified and validated five actigraphy-assessed sleep health domains in large community-based samples of older men and women: Timing, Efficiency, Duration, Alertness/Sleepiness, and Regularity. The domains were remarkably consistent across male and female samples and largely aligned with our hypothesized domains based on the self-report Ru-SATED sleep health framework comprised of *Regularity, Satisfaction, Alertness/Sleepiness, Timing, Efficiency,* and *Duration* domains (because Satisfaction is subjective, it was not a hypothesized actigraphy domain). Findings from our study also suggest a focused set of measurements for clinical and research use. While dozens of actigraphy variables exist, only five are needed to represent underlying domains. For example, the mean midpoint from sleep onset to wake-up (Timing), mean sleep maintenance (Efficiency), mean time from sleep onset to wake-up (Duration),

mean minutes napping per day (Alertness/Sleepiness), and the *SD* of the midpoint from seep onset to wake-up (Regularity) can be selected as a representative, relatively independent set of actigraphy measures with strong empirical justification in older adults.

We also hypothesized a separate Rhythmicity domain based on increased consideration of actigraphy rest-activity rhythms for health research. However, it was not upheld by the data. Hypothesized measures of rhythmicity (pseudo-F and amplitude) loaded with measures of napping (duration/frequency of naps) on the domain we named Alertness/Sleepiness. This finding underscores the fact that these measures are highly interrelated. More fragmentation of the activity rhythm during the day (i.e. more napping) decreases the difference in magnitude between the "sleep" and "wake" phases (i.e. lower amplitude) and results in less conformity of the rest-activity rhythm to the parametric extended cosine function (i.e. lower pseudo-F). Their inter-related nature can also be directly observed through the moderate-to-large pairwise correlations among these measures in the MrOS and SOF samples (Spearman $r = -0.62$ to -0.68 for pseudo-F and duration/frequency of napping; $r = -0.47$ to -0.56 for amplitude and duration/frequency of napping). However, there are several reasons why this finding may be specific to our sample and/or study design. First, MrOS and SOF are samples of older adults, who tend to have both dampened circadian rhythms and increased napping relative to younger adults [36, 37]. Second, roughly 5% of MrOS and SOF participants had impaired cognition based on the MMSE. As dampened rest–activity rhythms and increased napping are associated with worse cognition [38–42], it is possible that applying similar methods in a sample of older adults without any impaired cognition could produce different findings. Third, the parametric measures of rhythmicity that were used may not fully capture the nuances of between- and within-day rhythmicity in older adults, especially given the limited number of days of monitoring. We may have been able to see a unique rhythmicity factor with alternative (e.g. nonparametric) measures of rhythmicity and/or more days of monitoring. Thus, it will be important to apply similar methods in samples of different ages and with more nuanced measures of rhythmicity to further assess this finding.

Results from our study will help to extend the sleep health framework to answer important research and clinical questions. Most directly, our findings can guide sleep health research by suggesting a set of relatively independent measures for analysis. This set of measures opens possibilities for several different analyses. For example, the measures can be simultaneously be included in within a single predictive model, categorized and used to compute a sleep health score, and/or examined in

clustering models to reveal common phenotypes of sleep health characteristics [10]. Regardless of the particular statistical model (which should ideally be guided by the specific research question), our findings can enhance reproducibility and generalizability across studies by clarifying the multidimensionality of actigraphy-assessed sleep health and providing a template for the selection of representative measures. Our findings can also inform clinical practice. The sleep health framework has already proven useful in understanding several key outcomes, including depression [5], mortality [3, 10, 11], and adolescent health [8], and has also informed novel treatments (e.g. the Transdiagnostic Sleep and Circadian Intervention [2]). The identified actigraphy factors and representative measures will further guide and accelerate this work.

To our knowledge, this is the first formal attempt to empirically evaluate the factor structure of actigraphy assessed sleep health. We focused our analysis on actigraphy because it is relatively easy and inexpensive to collect over a long period of time, making it a key tool for larger epidemiological studies of sleep. Actigraphy also uniquely allows for monitoring of habitual, long-term behavioral sleep–wake patterns including regularity and rhythmicity [43]. However, there are limitations of actigraphy. Actigraphy captures physical rest–activity patterns rather than sleep per se, making it is less accurate for measuring actual sleep than the "gold standard" PSG. Also, the standard summary measures we examined (means and SDs of various sleep characteristics across days/nights) are derived in part through manual actigraphy scoring. Manual actigraphy scoring is labor intensive, especially for a large study, and it interjects some subjectivity into the actigraphy measures, although a study conducted in a subset of SOF participants showed high inter-subject scorer agreement [27]. Given the increasing availability of raw accelerometry data and advances in machine learning, our findings may help to guide the derivation of new metrics that further inform sleep health measurement. As these metrics are validated, they may also provide automated techniques for more rapidly and consistently analyzing actigraphy data.

It is important to consider our findings from actigraphy in the context of findings observed using other sleep data types, each of which provides unique information. Most notably, factor analytic methods have been used to examine the dimensionality of retrospective self-report sleep with the aim of developing new self-report sleep indices including SATED/Ru-SATED [1, 44, 45], the National Sleep Foundation Sleep Health Index (SHI) [46], and the National Institute of Health Patient-Reported Outcomes Information System (PROMIS) Sleep measure [35]. Our own analyses were based on the Ru-SATED framework and thus our findings align closely with this measure. In contrast, the SHI includes only three domains (Sleep Quality, Sleep Duration, and Disordered sleep) and the PROMIS Sleep measure includes only two domains (Sleep Disturbance and Sleep Related Impairments). These differences can largely be explained by the investigators' distinct conceptual and/or methodological frameworks. For example, SHI and PROMIS allowed sleep disorder symptoms (e.g. difficulty falling asleep), while Ru-SATED explicitly considered only quantifiable sleep characteristics (e.g. minutes to fall asleep). Moreover, PROMIS included only likert-scale self-report measures that could be validated using Item Response Theory, and for this reason did not include any quantitative

self-report data (e.g. sleep timing or duration). Overall, the more "behavioral" and quantifiable nature of Ru-SATED makes it an ideal framework on which to base our actigraphy sleep health factor analysis.

PSG and self-report daily sleep diary also play important roles in the study of sleep health, and factor analyses of measures captured using these data types may suggest different domains. For instance, a previous factor analysis of PSG sleep in depressed and healthy adults identified factors reflecting slow-wave sleep, rapid eye movement (REM) sleep, sleep continuity, and REM latency/delta sleep ratio [47]. While PSG can uniquely measure sleep architecture and is considered the "gold standard" of sleep measurement, its utility across many cohorts is limited because it is much more expensive and burdensome (e.g. requiring an overnight lab visits and/or set-up of in-home monitoring equipment). Furthermore, PSG is often based on a single night, and thus, the resulting metrics do not necessarily reflect an individual's usual sleep [48], nor can measures of regularity or rhythmicity be obtained. In contrast, sleep diary provides habitual, daily monitoring of subjective sleep and thus could plausibly capture all six proposed Ru-SATED domains [49, 50]. Sleep diaries are also potentially more available and less expensive than actigraphy, making them a key tool in sleep health research. However, sleep diaries require ongoing effort of participants to complete, which can result in missing data, and are subjective and limited in their ability to capture wakefulness during the night. Regardless, it will be important to apply consistent factor analytic methods to self-report sleep diary and PSG and compare findings with those from actigraphy.

The strengths of our study include the consistency of our findings across exploratory and confirmatory analyses of men and women and our use of large and well-characterized samples of community-dwelling older adults from multi-site epidemiological studies. However, there are some study limitations to note. First, the MrOS and SOF samples captured a median of five and four days/nights of actigraphy, respectively, with a minimum of three nights each. While 3 days/nights is sufficient for estimating the mean of actigraphy measures in older adults, up to 7 days/nights may be required for accurately measuring regularity using standard deviations [51]. We also cannot discount the possibility that additional days and nights of actigraphy might have improved estimation of rhythmicity measures and resulted in the observance of a separate rhythmicity domain as hypothesized. Even so, these same extended cosine features are highly predictive of key prospective health outcomes in the MrOS and SOF samples [3, 16, 18, 19, 38–40, 42, 52], indicating that they capture important aspects of older adults' sleep–wake activity despite the limited numbers of days/nights of measurement. Second, our findings are generalizable primarily to older, white adults. It will be important to apply similar methods to other samples with more diversity of age and race/ethnicity. Third, some goodness of fit statistics from the CFAs were not within the traditional range of "good fit". Optimizing fit indices is less relevant for our work in understanding factor structures [53]. They are only one piece of information within the larger context of our analyses, which ultimately showed remarkable consistency, clarity, and interpretability across analyses. Still, it will be important to perform further validations of our findings in additional samples. Fourth, direct comparisons of our sleep health findings between males and female are somewhat confounded by sample differences. For example, women from SOF

were almost 10 years older than males from MrOS and were also substantially more likely to be widowed rather than married. Finally, in future studies it will be important to consider nonparametric rest activity rhythms that were not currently available for our use in the MrOS and SOF cohorts (e.g. inter-daily stability or intra-daily variability [54]). Such metrics provide flexibility for examining within- and between-day patterns without imposing a specific parametric model (i.e. the extend cosine curve).

Conclusion

We empirically identified and validated five actigraphy-measured sleep health domains in older men and women (Timing, Efficiency, Duration, Alertness/Sleepiness, and Regularity) and determined which specific measures are most representative of each domain. These findings have the potential to guide innovative sleep health research and impact clinical practice and public health by improving health screening and suggesting novel treatments. Critical next steps are to examine the factor structure in more diverse samples and additional data types. Developing and validating factor scores to represent each factor may also serve to enhance subsequent analyses of sleep and health.

Supplementary material

Supplementary material is available at *SLEEP* online.

Disclosure statements

Financial disclosure: KS receives grant funding from Merck. SR reports grant and consulting fees from Jazz Pharma and consulting fees from Eisai Pharma and Respircardia, unrelated to this work. MHH reports a financial conflict of interest related to EISAI Inc. DJB has served as a paid consultant to Bayer, BeHealth Solutions, Emmi Solutions, Weight Watchers International, and Pear Therapeutics. He has served as a paid consultant for professional educational programs developed by the American Academy of Physician Assistants, CME Institute and Emmi Solutions, and received payment for a professional education program sponsored by Eisai. He is an author of the Pittsburgh Sleep Quality Index, Pittsburgh Sleep Quality Index Addendum for PTSD (PSQI-A), Brief Pittsburgh Sleep Quality Index (B-PSQI), Daytime Insomnia Symptoms Scale, Pittsburgh Sleep Diary, Insomnia Symptom Questionnaire, and RU_SATED (copyright held by University of Pittsburgh). These instruments have been licensed to commercial entities for fees. He is also coauthor of the Consensus Sleep Diary (copyright held by Ryerson University), which is licensed to commercial entities for a fee. MLW, LY, SFS, MLS, and DKS report no financial conflict of interest.

Nonfinancial disclosure: No authors have non-financial conflict of interest to report.

References

1. Buysse DJ. Sleep health: can we define it? Does it matter? *Sleep.* 2014;**37**(1):9–17.
2. Wallace ML, *et al.*; Osteoporotic Fractures in Men (MrOS) Study Research Group. Which sleep health characteristics predict all-cause mortality in older men? An application of flexible multivariable approaches. Sleep. 2018;**41**(1):zsx189.
3. Harvey AG, *et al. Treating Sleep Problems: A Transdiagnostic Approach.* 1st ed. New York, NY: The Guilford Press; 2017.
4. Brindle RC, *et al.* The relationship between childhood trauma and poor sleep health in adulthood. *Psychosom Med.* 2018;**80**(2):200–207.
5. Furihata R, *et al.* An aggregate measure of sleep health is associated with prevalent and incident clinically significant depression symptoms among community-dwelling older women. Sleep. 2017;**40**(3):zsw075.
6. Dalmases M, *et al.* Assessing sleep health in a European population: results of the Catalan Health Survey 2015. *PLoS One.* 2018;**13**(4):e0194495.
7. Dalmases M, *et al.* Impact of sleep health on self-perceived health status. *Sci Rep.* 2019;**9**(1):7284.
8. Dong L, *et al.* A composite measure of sleep health predicts concurrent mental and physical health outcomes in adolescents prone to eveningness. *Sleep Health.* 2019;**5**(2):166–174.
9. DeSantis AS, *et al.* A preliminary study of a composite sleep health score: associations with psychological distress, body mass index, and physical functioning in a low-income African American community. *Sleep Health.* 2019;**5**(5):514–520.
10. Wallace ML, *et al.* Heightened sleep propensity: a novel and high-risk sleep health phenotype in older adults. *Sleep Health.* 2019;**5**(6):630–638.
11. Wallace ML, *et al.* Multidimensional sleep and mortality in older adults: a machine-learning comparison with other risk factors. J Gerontol A Biol Sci Med Sci. 2019;**74**(12):1903–1909.
12. Blank JB, *et al.* Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials.* 2005;**26**(5):557–568.
13. Orwoll E, *et al.* Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study–a large observational study of the determinants of fracture in older men. *Contemp Clin Trials.* 2005;**26**(5):569–585.

14. Cummings SR, *et al*. Bone density at various sites for prediction of hip fractures. The Study of Osteoporotic Fractures Research Group. *Lancet*. 1993;**341**(8837):72–75.

15. Cummings SR, *et al*. Appendicular bone density and age predict hip fracture in women. The Study of Osteoporotic Fractures Research Group. *JAMA*. 1990;**263**(5):665–668.

16. Tranah GJ, *et al*. Circadian activity rhythms and risk of incident dementia and mild cognitive impairment in older women. *Ann Neurol*. 2011;**70**(5):722–732.

17. Zuurbier LA, *et al*. Fragmentation and stability of circadian activity rhythms predict mortality: the Rotterdam study. *Am J Epidemiol*. 2015;**181**(1):54–63.

18. Smagula SF, *et al*. Circadian rest-activity rhythms predict future increases in depressive symptoms among community-dwelling older men. *Am J Geriatr Psychiatry*. 2015;**23**(5):495–505.

19. Rogers-Soeder TS, *et al*. Rest-activity rhythms and cognitive decline in older men: the osteoporotic fractures in men sleep study. *J Am Geriatr Soc*. 2018;**66**(11):2136–2143.

20. Sheikh JI, *et al*. (1986) Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol*. **5**:165–173.

21. Goldberg D, *et al*. Detecting anxiety and depression in general medical settings. *BMJ (Clinical research ed)*. 1988;**297**(6653):897–899.

22. Teng EL, *et al*. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry*. 1987;**48**(8):314–318.

23. Folstein MF, *et al*. The mini-mental state examination. *Arch Gen Psychiatry*. 1983;**40**(7):812.

24. Pahor M, *et al*. Drug data coding and analysis in epidemiologic studies. *Eur J Epidemiol*. 1994;**10**(4):405–411.

25. Blackwell T, *et al*. Factors that may influence the classification of sleep-wake by wrist actigraphy: the MrOS Sleep Study. *J Clin Sleep Med*. 2011;**7**(4):357–367.

26. Mehra R, *et al*. Prevalence and correlates of sleep-disordered breathing in older men: osteoporotic fractures in men sleep study. *J Am Geriatr Soc*. 2007;**55**(9):1356–1364.

27. Blackwell T, *et al*. Actigraphy scoring reliability in the study of osteoporotic fractures. *Sleep*. 2005;**28**(12):1599–1605.

28. Jean-Louis G, *et al*. Sleep estimation from wrist movement quantified by different actigraphic modalities. *J Neurosci Methods*. 2001;**105**(2):185–191.

29. Marler MR, *et al*. The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes. *Stat Med*. 2006;**25**(22):3893–3904.

30. Wallace ML, *et al*. Variable selection for skewed model-based clustering: application to the identification of novel sleep phenotypes. *J Am Stat Assoc*. 2018;**113**(521):95–110.

31. Stevens JP. Exploratory factor analysis. In: *Applied Multivariate Statistics for the Social Sciences*. 2 ed. Hillsdale, NJ: Erlbaum; 1992:339–390.

32. Kline RB. *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press; 1998.

33. Bentler P. Comparative fit indices in structural models. *Psychol Bull*. 1990;**107**:238–246.

34. Buysse DJ, *et al*. Clinical management of insomnia disorder. *JAMA*. 2017;**318**(20):1973–1974.

35. Buysse DJ, *et al*. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep*. 2010;**33**(6):781–792.

36. Hood S, *et al*. The aging clock: circadian rhythms and later life. *The Journal of Clinical Investigation*. 2017;**127**(2):437–446.

37. Buysse DJ, *et al*. Napping and 24-hour sleep/wake patterns in healthy elderly and young adults. *J Am Geriatr Soc*. 1992;**40**(8):779–786.

38. Leng Y, *et al*. Objective napping, cognitive decline, and risk of cognitive impairment in older men. *Alzheimers Dement*. 2019;**15**(8):1039–1047.

39. Rogers-Soeder TS, *et al*. Rest-activity rhythms and cognitive decline in older men: the osteoporotic fractures in men sleep study. *J Am Geriatr Soc*. 2018;**66**(11):2136–2143.

40. Tranah GJ, *et al*. Circadian activity rhythms and risk of incident dementia and mild cognitive impairment in older women. *Ann Neurol*. 2011;**70**(5):722–732.

41. Yaffe K, *et al*. Connections between sleep and cognition in older adults. *Lancet Neurol*. 2014;**13**(10):1017–1028.

42. Walsh CM, *et al*. Weaker circadian activity rhythms are associated with poorer executive function in older women. *Sleep*. 2014;**37**(12):2009–2016.

43. Ancoli-Israel S, *et al*. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*. 2003;**26**(3):342–392.

44. Brandolim BN, *et al*. Sleep health assessment: a scale validation. *Psychiatry Res*. 2018;**259**:51–55.

45. Benitez I, *et al*. Validation of the satisfaction alertness timing efficiency and duration (SATED) questionnaire for sleep health measurement. *Ann Am Thorac Soc*. 2020;**17**(3):20338–203343.

46. Knutson KL, *et al*. The national sleep foundation's sleep health index. *Sleep Health*. 2017;**3**(4):234–240.

47. Buysse DJ, *et al*. Latent structure of EEG sleep variables in depressed and control subjects: descriptions and clinical correlates. *Psychiatry Res*. 1998;**79**(2):105–122.

48. Wohlgemuth WK, *et al*. How many nights are enough? The short-term stability of sleep parameters in elderly insomniacs and normal sleepers. *Psychophysiology*. 1999;**36**(2):233–244.

49. Carney CE, *et al*. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep*. 2012;**35**(2):287–302.

50. Monk TH, *et al*. The pittsburgh sleep diary. *J Sleep Res*. 1994;**3**(2):111–120.

51. Rowe M, *et al*. Actigraphy in older adults: comparison of means and variability of three different aggregates of measurement. *Behav Sleep Med*. 2008;**6**(2):127–145.

52. Smagula SF, *et al*. Circadian rest-activity rhythms predict future increases in depressive symptoms among community-dwelling older men. *Am J Geriatr Psychiatry*. 2015;**23**(5):495–505.

53. Cook KF, *et al*. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Qual Life Res*. 2009;**18**(4):447–460.

54. Goncalves BS, *et al*. Nonparametric methods in actigraphy: an update. *Sleep Science*. 2014;**7**(3):158–164.