



Published in final edited form as:

Sleep Health. 2021 December ; 7(6): 699–707. doi:10.1016/j.sleh.2021.08.005.

Multidimensional Sleep Health in a Diverse, Aging Adult Cohort: Concepts, Advances, and Implications for Research and Intervention.

Joon Chung^{1,2,*}, Matthew Goodman^{1,2}, Tianyi Huang^{1,2}, Suzanne Bertisch^{1,2}, Susan Redline^{1,2}

¹Brigham and Women's Hospital

²Harvard Medical School

Abstract

Purpose: To illustrate two frame-shifts of multidimensional sleep health: i) use of composite sleep metrics; and ii) the inter-correlations among sleep dimensions.

Participants: 735 adults of diverse backgrounds aged <65 years who participated in the Multi-Ethnic Study of Atherosclerosis.

Measures: In-home polysomnography (PSG), 7-day wrist actigraphy, and validated questionnaires.

Methods: The Buysse Ru SATED model – sleep **R**egularity, **S**atisfaction, **A**lertness, **T**iming, **E**fficiency, **D**uration – was operationalized, then extended by including additional measures of sleep architecture and sleep apnea from PSG and difficulties initiating sleep from questionnaire and sleep onset latency and duration [ir]regularity from actigraphy. We dichotomized sleep variables, operationalizing optimal and non-optimal ranges as ‘1’ and ‘0’, respectively, summed into a Sleep Health Score, and computed global sleep health scores via Principal Components Analysis (PCA).

Findings: Participants showed low prevalence of sleep regularity in timing (<30 minutes Standard deviation [sd]; 21.4% favorable) and duration (<60 minutes sd; 36.9%). Although 62.7% of participants demonstrated favorable sleep duration by actigraphy, few met criteria for favorable levels of % N3 (11.4%) or % REM (34.1%). The average Sleep Health Score was 5.6 of 13 (higher is better). Sleep variables were variably inter-correlated ($r=0$ to $r=-0.72$). The first Principal Component (PC) for each operationalization of sleep health was interpretable as a ‘health’ score; all summary scores captured variable but systematic shifts towards more favorable sleep in each sleep variable.

* **Corresponding author:** Correspondence to Dr. Joon Chung, Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Harvard Medical School, BL-252, 221 Longwood Ave., Boston, MA 02115 (jchung26@bwh.harvard.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conclusions: Multidimensional sleep health can be measured by complementary composite scores as well as consideration of multiple individual dimensions.

Keywords

sleep; sleep health; actigraphy; polysomnography; multidimensional

Introduction

Although it would be convenient if available datasets contained a comprehensive variable called ‘sleep’, what datasets typically include are multiple metrics *about* sleep: estimates of specific features of sleep as well as the impact of sleep on daytime symptoms and function. Each feature or *dimension* of sleep is often modeled separately and interpreted independently – as exposure or outcome – in distinct models: an effect size for total sleep time, one for sleepiness, and another for sleep efficiency.

A variety of sleep dimensions influence diverse outcomes: academic performance, mental and physical health, and mortality. Sleep dimensions are sensitive to exposures at the individual level (*e.g. age; development; financial strain; caffeine, tobacco, and other substance use; physical activity, etc.*) and social and environmental levels (*e.g. neighborhoods and the built environment; social cohesion; social relationships; marital status; loneliness; discrimination; and work schedules, etc.*). However, sleep dimensions are not independent of each other (1). Just as food is not consumed as singular nutrients (2), sleep is not experienced as singular dimensions. Adults who meet national recommendations of 7–9 hours of duration may evince heterogeneity in regularity, continuity, macro- and micro-architecture, and satisfaction. Accordingly, Buysse (2014) has defined sleep health as: “a multidimensional pattern of sleep-wakefulness, adapted to individual, social, and environmental demands, that promotes physical and mental well-being” (3).

A multidimensional paradigm is intrinsic to our understanding of sleep. Clinicians, for instance, routinely consider multiple aspects of sleep in diagnosis and management of sleep disorders. The literature has long reflected interest in the combined, including interactive, aspects of sleep, although has focused more on sleep disturbances rather than sleep health, and focused on clinical disorders rather than population health. Thus: What is distinctive and useful about the paradigm of ‘multidimensional sleep health’?

We provide empirical data to support multidimensional sleep health as a distinctively useful approach for characterizing ‘sleep health’ across the population. One innovation is the quantification of sleep *health*, which ranges beyond disorder or insufficiency – acknowledging the existence of gradients of “healthy sleep” beyond meeting a minimum duration of sleep or an absence of disorders such as insomnia or obstructive sleep apnea (OSA) (3). Sleep health also parallels broader paradigm shifts which acknowledge the World Health Organization’s definition of health as more than the absence of illness (4).

Sleep health also has other features that contribute to its utility. First, sleep health approaches multiple sleep dimensions (conceptually, operationally, analytically) as contributors to a *composite* concept – a metric of global sleep health. Extant scales include

Buyse's Ru SATED scale and the National Sleep Foundation's (NSF) Sleep Health Index (3, 5). Consistent and valid scales are necessary for scalable, longitudinal data collection that enables quantification of population shifts in global sleep health. Such work does not preclude analyzing each dimension in its own right, although statistical dependence may complicate model interpretation.

Second, sleep health emphasizes that sleep dimensions do not exist in isolation (1). There are physiological reasons why certain dimensions are correlated, perhaps causally, with others, for instance, duration and sleepiness or OSA and continuity. Sleep dimensions also may show statistical correlation when derived from common measurement tools. Thus, interpretation of individual dimensions as if they were independent may lead to erroneous, or partial conclusions about 'sleep': what kind of continuity, quality, alertness, and regularity is likely to be seen among 'sufficient' sleepers (however defined) as opposed to 'insufficient' sleepers?

In this paper, we illustrate these two frameshifts: i) composite sleep scores; and ii) consideration of correlations. To the Buyse Ru SATED model of sleep health—sleep **Regularity, Satisfaction, Alertness, Timing, Efficiency, Duration**—we incorporated several additional metrics associated with later health: sleep stage information (% N3, % Rapid Eye Movement [%R]), the Apnea-Hypopnea Index[AHI]; self-reported frequency of difficulties initiating sleep (DIS) or sleep onset latency from actigraphy; and (ir)regularity in duration; and mapped these dimensions to a conceptual sleep health model. We considered how to model multiple, often correlated, dimensions and found merit in both analyzing individual dimensions as well as a global sleep health metric. We constructed a Sleep Health Score (SHS: sum of indicator variables); and 4 data-driven sleep health scores from Principal Components Analysis (PCA). Finally, we showed how these composite and individual dimensions map across race/ethnic groups, an extension of prior work where we reported in detail differences in sleep health by race/ethnicity across a larger age range but did not evaluate the impact of differences in composites and their psychometric features (6).

Methods

Sample:

The sample was from the Multi-Ethnic Study of Atherosclerosis (MESA), a 6-community cohort of aging adults in the United States (Forsyth County, NC; Northern Manhattan and the Bronx, NY; Baltimore City and County, MD; St. Paul, MN; Chicago, IL; and Los Angeles, CA), diverse in race/ethnicity (White, Black, Hispanic, Chinese) (7). Details on this cohort are published (8). In brief, at Exam 5 (2010–2013), participants were invited to the MESA-Sleep ancillary study and underwent single-night at-home polysomnography (PSG), 7-day wrist actigraphy (Actiwatch Spectrum; Philips Respironics, PA; Actiware-Sleep v 5.59), and validated sleep questionnaires (Epworth Sleepiness Scale [ESS], Women's Health Initiative Insomnia Rating Scale [WHIIRS]) (9, 10). Analyses were performed on secondary, de-identified, publicly available data; participants provided informed consent for the original Multi-Ethnic Study of Atherosclerosis (MESA) study; and MESA and MESA-Sleep complied with the Declaration of Helsinki. Actigraphy-assessed sleep onset and offset (rest interval) were determined by trained scorers based on a sleep

diary completed concurrently, participant-actuated event markers, actigraph light sensors, and activity counts. Analyses were restricted to adults 54–64 years because there are more normative data available for this age range (11).

Sleep health conceptual model and mapping concepts to variables: As described previously (6), sleep health domains were drawn primarily from Ru SATED but also the National Sleep Foundation's Sleep Health Index (3, 5, 12). Selection of additional variables was informed by prior knowledge of a high prevalence of OSA and sleep fragmentation in middle-age adults (13). Thus, several continuity metrics were chosen *a priori*: sleep maintenance efficiency (SME; i.e., sleep efficiency from sleep onset through sleep offset), the Fragmentation Index (FI), and wake after sleep onset (WASO). Other additions to Ru SATED such as % R, % N3, AHI (3% desaturation to define hypopneas), duration (ir)regularity, and self-reported frequency of difficulties initiating sleep were also selected prior to analysis, using expert knowledge or evidence from the literature indicating their relevance for health outcomes.

Our conceptual model of sleep focused on four features (Figure 1): 1) the transition from wake to sleep (DIS) and daytime sequelae (quality, alertness) from the participants' perspective; 2) the period between sleep onset; 3) inter-daily variability/irregularity in timing and duration; and 4) the entire sleep experience both during and across nights.

Derivation of global sleep health composite scores: We first created a summary score of dichotomized sleep variables (Sleep Health Score; SHS). As described in (6) following (14, 15), cut-points defined optimal ranges, with coding of 'favorable' sleep as '1' and non-optimal ranges as '0'. Optimal ranges were drawn from the literature, the NSF's objective sleep quality report (eg, for WASO), expert consensus, or sample characteristics (11, 13, 16–18). These dichotomous indicators of favorable sleep were summed into a global metric (SHS), additionally categorized as: least favorable (SHS 0–3), less (4–6), more (7–9), and most favorable (10–12) sleep health.

PCA also was conducted on continuous metrics: a) Ru SATED variables (midpoint variability/irregularity [MPsd], quality, timing (log-transformed), maintenance efficiency, and duration); b) Ru SATED, expanded with selected variables from PSG, adding AHI, % N3, %R; c) all variables chosen *a priori* (adding several variables measuring similar dimensions: duration irregularity/variability, frequency of trouble sleeping, WASO, fragmentation). For PCA, actigraphy-assessed onset latency was used rather than Difficulties Initiating Sleep given the focus on analysis of continuous variables for PCA. Each PC1 score was standardized (mean=0, sd=1) and coded such that higher is better.

Statistical Methods: Pearson correlations were computed for continuous sleep metrics (log-transformed where necessary). Principal Components and loadings were evaluated for three composites: i) Ru SATED; ii) Ru SATED + OSA (AHI) and architecture (%N3, %R); and iii) the Sleep Health Score (the most comprehensive set of indices across domains that mapped to the SHS constructed using dichotomous cutoffs). Sensitivity analyses assessed a parsimonious version of the Sleep Health Score (PC1) that eliminated potentially redundant measures. Composite score internal reliability was assessed by alpha Cronbach. Consistency

in PC weights-both direction and magnitude-was evaluated across each component. Trends among individual sleep health variables in relation to SHS categories were reported. Outcomes of composite sleep health scores (linear regression), and their dichotomized components (Zou's modified Poisson regression), were regressed on the exposure of race-ethnicity (White=ref), with adjustment for age and sex. Type I errors were allocated to tests on global scores; component-level analyses were evaluated qualitatively. Analyses were conducted in R 3.6.3.

Results

The sample had an average age of 59.4 ± 3.0 years, 55.6% were female, 31.8% were currently employed, and most obtained at least a high school degree (Table 2). The majority met criteria for actigraphy-assessed sufficient sleep (6–8 hrs; 62.7%; Table 1) but did not meet favorability thresholds for: continuity (WASO [6.8% favorable], FI [27.1%]), sleep architecture (% N3 [11.4%], %R [34.1%]), both types of (ir)regularity/variability (midpoint [21.4%], duration [36.9%]), AHI (48.2%), and quality (21.4%).

Table 2 suggests that when the components in Table 1 are summed into the SHS, increasing scores capture variable yet systematic positive shifts across all sleep dimensions. Differences in most-least favorable sleep health were notable in duration (+1.6 hours), midpoint irregularity/variability (−43.8 minutes), duration irregularity/variability (−48.7 minutes), AHI (−23.1 events/hr), % R (+8.4%), % N3 (+9.9%), and WASO (−58.5 minutes).

Figure 2 shows correlations among continuous sleep variables within and across domains. Within the continuity domain, fragmentation and sleep maintenance efficiency are highly correlated ($\rho=0.72$). Across domains, there are additional statistical dependencies: timing regularity (MPsd(log)) and sleep duration (TST) are moderately correlated ($\rho=0.40$), as is timing regularity with sleep timing (Timing(log)) ($\rho=0.32$). AHI correlates with continuity metrics (FI, SME, WASO) and % R (ρ 's range: −0.34 to 0.32).

Table 3 shows PC weights for the first three PCs for each sleep composite. The first PC is interpretable in all models as a “sleep health score”: the directionality of PC1 weights for individual variables concur with *a priori* knowledge of “better” sleep (e.g., higher sleep duration contributes positively). As expected, the PC1 weights varied depending on variables included but did not substantively change the interpretation of the PC1 composites. The simplest Ru SATED score shows that the measures of timing regularity, timing, and duration loaded on PC1 while satisfaction/alertness and efficiency loaded on PC2. With the addition of PSG and actigraphy measures to the Ru SATED score, we observe that AHI loaded with efficiency and regularity, %R loaded with satisfaction/alertness, duration regularity loaded with timing regularity, and fragmentation loaded with efficiency. Across the extended (SHS) scores, timing and regularity tended to load on PC2 while self-reported sleep indices loaded on PC3. Less than 30% of the variation in the data was explained by each of the first PCs, consistent with a multi-dimensional basis for sleep health. Moderate internal consistency was observed for each sleep health composite, with alpha Cronbach varying from 0.42 for Ru SATED to 0.61 for the full SHS.

Figure 3A shows the Sleep Health Score (SHS), SHS-PC1, Ru SATED + OSA and architecture (PC1), Ru SATED (PC1), and their components by race-ethnicity, adjusting for age and sex. Consistent with prior analyses (47), the sample averaged 5.7 ± 2.1 of 13 possible favorable dimensions on the SHS. Black adults averaged 1.34 fewer favorable domains than White adults ($p < 0.001$). Hispanic adults averaged 0.74 fewer favorable domains ($p < 0.001$); however, most of the individual dimension-level disparities overlap with the reference. Ru SATED (PC1) and Ru SATED + OSA and architecture (PC1) showed similar variation by race/ethnicity as the SHS and SHS-PC1. Figure 3B suggests that the largest drivers of global racial-ethnic disparities were sleep regularity (timing and duration), continuity (fragmentation, maintenance efficiency), and total sleep time.

Discussion

There is growing interest in considering sleep health as not merely the absence of a disorder but as a summary of the positive attributes of healthy sleep. Similarly, there is movement towards using a conceptual framework that articulates sleep as multidimensional, (3). We showed that sleep characteristics were commonly correlated, underscoring the limitations of focusing on single aspects of sleep: alternate intervention targets can be missed, and population variation in sleep health will not be adequately described. The inter-correlation among sleep variables may reflect intrinsic physiological inter-relationships and potential responsiveness to common stressors. One approach for summarizing the combined effects of multiple sleep dimensions is to create composite scores. We constructed a Sleep Health Score (SHS) by summing favorable indicators across 13 aspects of sleep health. Within the SHS, which varied by race-ethnicity, we found several key drivers of sleep health disparities: regularity, duration, and continuity. The validity of this approach was supported by PCA performed on continuous Sleep Health Score variables, the subset of Ru SATED variables, and Ru SATED extended with OSA and architecture. Those analyses showed that regardless of the specific model (Ru-SATED vs the expanded SHS), the conditional distribution of each composite score consistently showed racial-ethnic disparity. Moreover, the PCs reflected consistent correlations across sleep dimensions; the first PC (which represents a linear combination of weights for each sleep characteristic that explains the largest variance in the data) reflects large contributions of regularity and duration, and additional PCs further describe attributes more strongly influenced by sleep quality and sleep disorders. Our study was not designed to identify the optimal number of dimensions to include in a composite; nor to construct a universally-applicable, optimal composite (which will vary according to the purpose of the research and availability of measures). Rather, we highlight the (a) feasibility of using a multi-dimensional approach for evaluating sleep health and its value in identifying sleep health disparities; (b) overall consistency of scores when components such as regularity, duration and quality are minimally considered; and (c) complementary value of considering both global scores and multiple individual components

Below we elaborate on our specific approach to sleep dimensions as: i) potential intervention targets; ii) correlated constructs; iii) contributors to a composite concept; and as iv) indicators that systematically track with global scores.

i) Public health perspective: sleep dimensions are potential intervention targets

Dichotomous indicators lend themselves to prevalence reporting, public health goal setting, and provide threshold targets for interventions. A public health perspective suggested the importance of an *a priori* approach, relying on concepts/theory (Figure 1) and prior knowledge of meaningful clinical thresholds (Table 1). This perspective views sleep characteristics as candidate targets for public health intervention or biomarkers which may signal later disease.

ii) Internal perspective of Sleep Health: sleep dimensions do not exist in isolation

A different yet complementary perspective was to understand the inter-relationship among sleep dimensions, which provides insight into common etiologies and inter-dependencies. This approach is of interest because not all aspects of sleep are easily modified directly (eg % R) but may be modified via targeted correlated features (via AHI); thus, correlations may give clues as to intervention targets.

iii) Holistic perspective of Sleep Health: sleep health can be viewed as a composite concept

A third perspective suggests that sleep health can be viewed as a composite or ‘global’ concept. We observed similar patterns of association for the SHS compared to global assessments derived using PCA and alternative sleep health frameworks. Prior work using questionnaires or actigraphy found similar evidence of patterning or latent factors (5, 12, 19). While the SHS was informed by expert/prior knowledge (placement of cut-points), the PC1 sleep scores were a data-driven approach relying instead on covariance. That the two approaches were derived differently yet produced nearly identical pictures of racial-ethnic disparity suggested the robustness of various composites for estimating sleep disparities estimates. Additional research is needed to examine consistency of alternative scores in other settings (tracking temporal trends, responses to global exposures such as the pandemic, and public health interventions, etc).

iv) Composite-component perspective of Sleep Health

A final perspective was to examine the relationship between composite sleep scores and their individual components. The SHS—because of its dichotomous basis—is akin to asking: “In how many important aspects of sleep does the participant have a ‘passing’ or ‘favorable’ score?” Increasing scores on the global SHS captured systematic shifts towards favorability in each sleep variable and in each PC1 score: we observed increases in variables in which ‘higher is better [eg duration]’ and decreases in variables in which ‘lower is better [eg AHI]’. When derived from PCA, the weights of the individual variables were consistently patterned in all PCA operationalizations of sleep health; i.e., a 1-unit *increase* on the composite score may also indicate that the individual-level metrics are trending towards favorability. Thus, an advantage of a global score is the consideration of multiple dimensions together.

Practical uses of global sleep scores

Global scores can provide overall assessments when dimension-level associations are under-powered (eg Hispanic disparity; Figure 3). For health outcomes, the effects of ‘poor sleep’

on outcomes such as cardiovascular disease and mortality may be the result of aggregate effects of co-occurring aspects of unfavorable sleep. Thus, a global score may capture small 'effects' among the individual variables that cumulate to a larger composite effect in higher-dimensional data, akin to individual nutrients which inform a larger dietary pattern (2). Second, analysis of this global score can reduce multiple testing, which may be beneficial in settings involving limited sample size and multiple available sleep metrics.

Dimension-level analyses, basis for dichotomization, and cohort sleep health assessment

Despite their benefits, global scores do not provide specific sleep targets for intervention. One approach is to perform post-hoc analyses of individual dimensions only if global dimensions show variation by the risk factor under consideration. We focused on a SHS constructed using a simple sum of the *dichotomous* versions of sleep variables, which provides readily interpretable metrics for public health purposes and is supported by the availability of clinical cut-points for many aspects of sleep (11, 16). While continuous sleep exposures or outcomes may offer greater statistical power, there are advantages for using cutoffs (20). First, cut-points aid clinicians in making decisions. Second, prevalence can be defined and then used for needs assessment: e.g., approximately one third of Americans do not meet sleep quantity recommendations (21). Third, cut-points enable quantification of prevalence trends over time and across groups. Additionally, cut-points may help avoid issues of non-linearities in exposure-response relationships, and cut-points are useful for setting goals for public health initiatives. Categorical assessments may be appropriate even if there is no evidence of a latent, *internal* discrete structure (taxon) for that dimension. As Kessler (2002) notes, "there appears to be no taxon for high blood pressure," and yet cut-points for blood pressure based on *external* criteria such as risk of stroke guide clinical and public health decisions (20).

Correlations facilitate interpretation and can suggest common biology and risk factors

Sleep dimensions are not isolated phenomena and attention to their correlations may shed light on sleep health. For example, compared to White adults, Black adults in MESA showed i) lower sleep health composite scores, on average by more than a full SHS component difference; and ii) the largest dimension-level disparities for Black adults are in regularity (timing and duration), duration, % N3, and continuity (Figure 3). Viewed from the sleep health paradigm, their correlations suggest the likelihood of observing multiple sleep disparities given a large disparity in, and high correlation with, a single sleep dimension. It is also important to consider the bases for correlation: measurement of similar constructs (e.g., SME and FI each measure sleep continuity); biological association (e.g., possibly irregularity and timing); similar sensitivity to external factors (e.g., light, shiftwork that influence sleep regularity, duration, and continuity). Here, inter-correlation is cast as a source of research hypotheses.

While sleep duration has been a key focus of epidemiological research, findings on the detrimental effects of insufficient sleep duration may implicitly capture effects of poor sleep in correlated dimensions – such as regularity, continuity, sleep architecture, OSA, placement in the 24-hour day, and sleepiness/alertness – and erroneously attribute this omnibus effect to duration. Investigating dimensions other than duration may thus inform interventions on a

population scale. It may prove more accurate to say: ‘Sufficient sleep duration (*and all that it implies*) is protective of later health.’ Alternatively, ‘Sleep health, of which average sleep duration is one vital component, is protective of later health.’ Novel sleep health analyses using machine learning support this conclusion (1, 22).

Tensions and limitations

A limitation concerns the dependencies among dimensions, which also are among the more interesting and useful features of multidimensional sleep health. The correlations observed may be particular to the sample or attributable to confounding factors such as the socioecological context, age or other factors. A second limitation concerns evolving definitions and criteria of sleep health. For instance, there are not yet consensus optimal ranges for % N3 in late adulthood (11). While PCA revealed interpretable principal components, objective and subjective sleep measures each loaded on distinct PCs, which could reflect measurement bias separate from dimensional differences. The correspondence between self-reported composite scales and those involving objective measures is not characterized, and research indicates systematic differences in estimates between self-report and objective measures that vary by socio-demographics as well as by underlying sleep characteristics (23, 24).

The core definition of Ru SATED dimensions is evolving, with sleep regularity a recent addition. A canon of sleep health parameters for measures of sleep micro-architecture (e.g. spindles, k-complexes) has yet to be established (25, 26). Future work may utilize commercial health and activity monitors for eg sleep regularity (wearable activity monitors); sleep apnea (portable oximetry); and EEG via streamlined multi-day in-home recordings (27).

Finally, the components of global sleep health scores will likely vary according to the questions at hand (e.g., interest in a pediatric versus an aging population; cardiovascular disease versus cancer) and measurements available for study. In our study, the availability of PSG allowed us to comprehensively include quantitative sleep measures that are predictive of health. However, as shown by our PCA, restricting to simpler measures also provided interpretable global scores that also varied by race/ethnicity. A key construct that is difficult to assess is sleep onset latency, which is subject to recall bias (by diary) or misclassification by actigraphy due to challenges in correctly determining ‘lights off’ time. Improvement in measuring sleep onset latency may allow inclusion of more informative assessments of this key feature of insomnia and sleep quality.

Because a single, canonical operationalization of sleep health may be suboptimal for any particular project, adaptations may be warranted. For example, Meltzer and colleagues recently proposed an adaptation of Buysse’s Ru SATED framework for pediatric use, the Peds B-SATED model (B stands for Behavior), in which sleep hygiene *behaviors* are nominated as an additional dimension in lieu of ‘regularity’ for children (28). However, the multidimensionality of sleep behaviors themselves may pose a challenge to integrating sleep behaviors into multidimensional sleep health definitions. Future work might draw from prior literature on composites or complex systems in other fields which also deal

with multi-dimensional data in choosing how to weight and aggregate the complex set of phenomena comprising sleep health in a context-specific way (29).

Sleep health is undergoing conceptual advances and scale development (3, 5), scale validation (5, 12), implementations of sleep health in cohort studies and community samples (14, 30, 31), innovations in methodological approaches (1, 22, 32, 33), and optimal range and threshold determination (11, 16, 34). We suggest that statistical dependencies are an additional factor to contend with. In this study, the dependencies were cross-sectional; a fuller understanding of sleep health may be aided by longitudinal and comprehensive sleep data collection.

Summary

Multidimensional sleep health represents a paradigm shift in sleep science. Two useful frameshifts are that sleep health is both conceptually and operationally a composite concept and that sleep dimensions do not exist in isolation. These frameshifts logically implicate sleep disordered breathing and sleep regularity as targets for further research and intervention.

Acknowledgements

This research was supported by National Institute of Health grants 5T32HL007901, R01HL098433; and by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169, and contract 5N92019C00011 from the National Heart, Lung, and Blood Institute. SR was partly funded by the NIH National Heart, Lung, and Blood Institute (NHLBI) R35HL135818. The National Sleep Research Resource was supported by the NHLBI (R24HL114473, 75N92019R002).

Appendix A.: Multidimensional sleep health: conceptualization and operationalization

A1. Overview – how we operationalized multidimensional sleep measures for hypothesis testing

Step 1: Variable selection

Step 1 was dimension/variable selection. The basis for inclusion/exclusion of dimensions was prior concepts (Ru SATED, NSF), adapted to i) fit our population (diverse adults, mean age = 59.4 years); ii) leverage available sleep measures, which included polysomnography and actigraphy; iii) to assess racial-ethnic disparity. Our conceptual model and available measures suggested comprehensive inclusion of variables.

Step 2: Operationalizing selected variables

The second step was operationalization. For public health use, we dichotomized each variable selected in Step 1 and assessed sample prevalence. Optimal or favorable sleep ranges were coded '1', and less favorable ranges were coded '0'. For instance, actigraphy-assessed total sleep time was coded '1' if the individual averaged 6–8 objective hours of sleep, and '0' otherwise. If sampling weights existed, we would assess population prevalence in weighted analyses (we did not have such weights in MESA, which is not

intended to be a national probability sample). Dichotomized sleep variables were summed into the Sleep Health Score (SHS).

The Sleep Health Score (PC1) and other PC scores were constructed by PCA. As described previously, sleep variables were centered, scaled, and log-transformed where necessary. As noted, we substituted actigraphy-assessed sleep onset latency (SOL_{act}) for Difficulties Initiating Sleep (DIS), because the ordinal variable DIS was not improved by log-transformation, yet our conceptual model (Figure 1) suggested that the wake-sleep transition may be important to assess. We operationalized canonical Ru SATED (6 variables), Ru SATED + OSA/PSG (9 variables), via PCA. Figure A2 shows the distribution of each global sleep health metric: Ru SATED (PC1), Ru SATED + OSA and architecture (PC1), Sleep Health Score, Sleep Health Score (PC1), Sleep Health Score (PC1) parsimonious (eliminating potential redundant measures). Resulting scores were standardized (scaled, centered) and coded so that higher is better and a 1-unit increase/decrease is a 1-sd increase/decrease in sleep health. Each score is approximately normally-distributed. We will primarily discuss the SHS-PC1 due to its higher internal reliability, comprehensiveness, and distributional similarity to other PC sleep scores. The SHS is under-dispersed (mean=5.6, sd=2.1), but is unimodal, may have less skew than PC-derived scores, and is amenable to linear regression.

Step 3: Hypothesis testing

The final step was hypothesis tests: i) global Sleep Health Scores using Ordinary Least Squares regression (OLS); and ii) the dichotomized dimensions which comprise the SHS.

Below, we i) detail the mapping of prior concepts to our conceptual model, then to sleep measures; ii) remark on population-specific thresholds; iii) and compare/contrast global sleep scores.

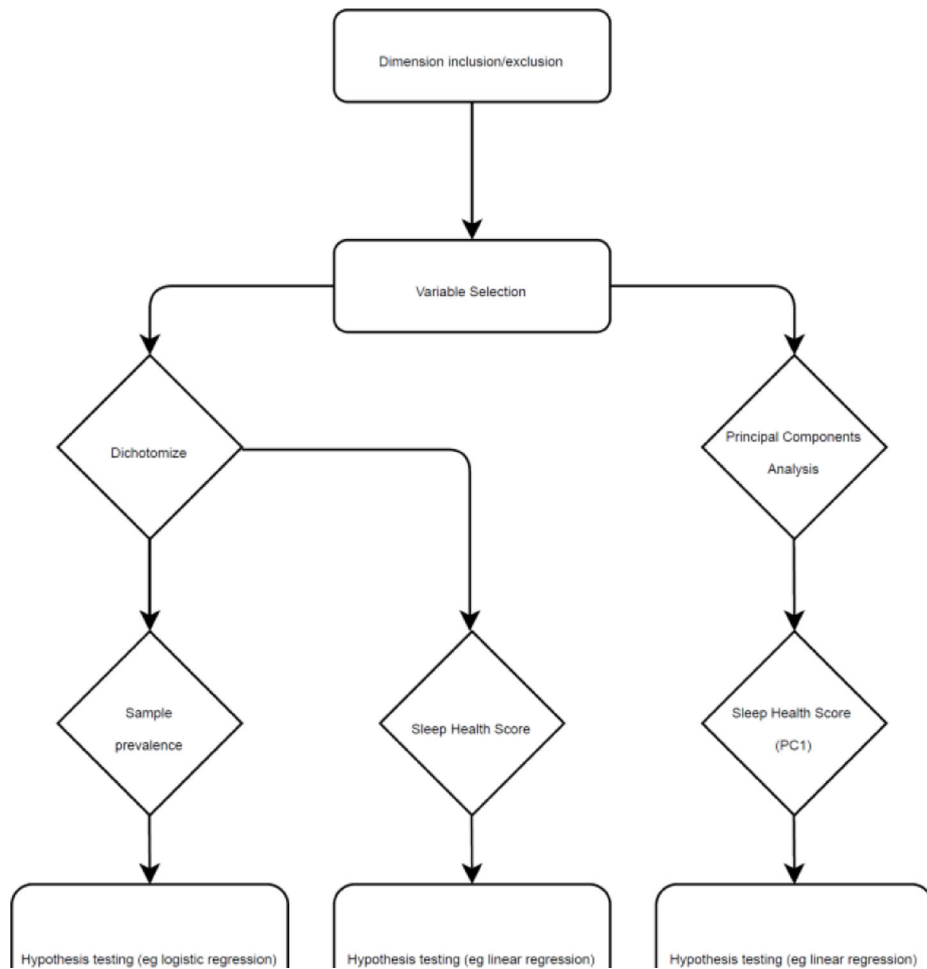


Figure A1.
Work flow

A2. Conceptual model: mapping concepts to variables

Before and after sleep: Participants' point of view

The SHS includes the self-reported difficulties initiating sleep (DIS) item from the Women's Health Initiative Insomnia Rating Scale (WHIIRS) to assess the transition from wake to sleep (10). The SHS includes the insomnia phenotype of difficulties initiating sleep (DIS) – a symptom of high clinical utility. Self-reported measures of insomnia complaints may be of higher clinical utility than objective measures because one of the more striking characteristics of insomnia is its apparent lack of a straightforward objective analogue. At present, there is no accurate objective basis for the classification of insomnia. The SHS includes the WHIIRS item for habitual quality (Quality) and Epworth Sleepiness Scale (ESS) scores as after-sleep metrics of subjective satisfaction and participant-reported functioning.

The SHS-PC1 uses actigraphy assessed sleep onset latency (SOL_{act}) due to intractable skew of DIS, and thus has one more objective dimension but lacks a participant-oriented measure

for this period. Conversely, the SHS lacks objective characterization of the wake-sleep transition in favor of participant evaluations of this period.

Between sleep onset-offset

The median adult MESA-Sleep respondent experienced 62 minutes of Wake After Sleep Onset (WASO), which could reflect a single 62-minute consolidated chunk of wakefulness or could reflect a highly fragmented period, with alternating 1-minute wake and sleep bouts. Sleep Maintenance Efficiency (SME) values would not differ in these two extreme cases, whereas the Fragmentation Index values would. We chose Sleep Maintenance Efficiency (SME) over Sleep Efficiency (SE). Sleep efficiency includes Time in Bed in the denominator, which includes sleep onset latency. As described, we already used DIS to characterize this period of sleep onset latency, from the participants' own viewpoint, as a measure of difficulties initiating sleep and SOL_{act} as an objective measure. Thus, for greater precision in characterizing the sleep period (sleep onset – sleep offset), we favored SME over SE.

Finally, from actigraphy, we used average total sleep time to assess the overall quantity of sleep. From PSG, we used % N3 (slow wave sleep), % Rapid Eye Movement sleep (% REM), and the Apnea-Hypopnea Index (AHI) as measures of sleep architecture and sleep disordered breathing.

Across nights

To characterize (ir)regularity in timing and duration across nights, we used the standard deviation of midpoint and total sleep time (TST) across the study period. To characterize average placement of sleep in the 24-hour day, we used midpoint timing/placement (Timing); for MPsd and Timing, the data in MESA-Sleep arrive in HH:MM:SS format which we converted to minutes by calculating difference in minutes from midnight to obtain a numerical value. We coded 02:00:00–04:00:00 (or 120–240 minutes from midnight) as favorable sleep midpoint placement in the 24-hour day (Timing) for the SHS.

Timing is natively a continuous and 'circular' concept in that, for instance, 23:59:30 and 00:00:30 are both 30 seconds from midnight. For PCA, then, we required a transformation of the variable we calculated above to ensure 23:59:30 was not coded as the most extreme value from midnight. Our solution was to first constrain timing to less than 12 hours or 720 minutes from midnight by coding values greater than 720 minutes from midnight as: $Timing = Timing - 1440$ minutes. We then calculated a metric (in minutes) based on the distance in midpoint between each participant and that of the median participant. Because this variable is skewed and included 0, we added 1 to this variable and log-transformed it:

1. $Timing = \text{Difference minutes of timing HH:MM:SS from midnight (00:00:00)}$
2. $Timing = Timing$ if < 720 minutes; otherwise, $Timing = Timing - 1440$ minutes.
3. $Timing = |Timing - Timing_{median}|$
4. $\log Timing = \log(Timing + 1)$

We treated MPsd as differences from midnight, in standard deviation (minutes) of midpoint. MPsd is right-skewed and is improved by log-transformation; TST regularity (TSTsd) is likewise skewed but not overly so, and we left this variable as is.

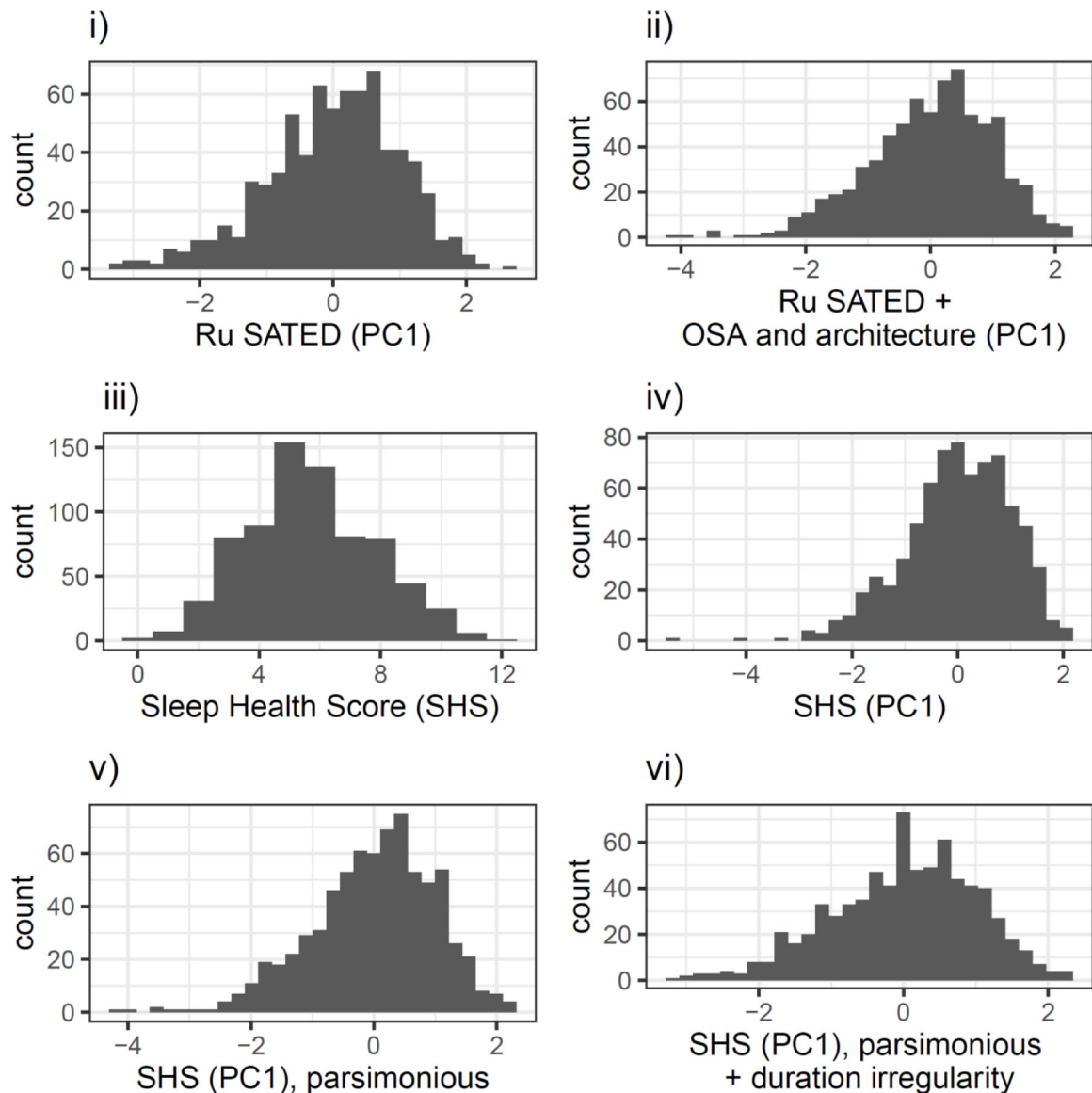


Figure A2.
composite score distns (140mm)

A3. Defining optimal ranges according to the population

The Sleep Health Score is the sum of dichotomized dimensions. We used thresholds from the literature, expert consensus, or sample characteristics (Table 2). We remark on several thresholds of note, which highlight the importance of considering the population of interest and suggest future research.

Thresholds

First, we used AHI ≥ 15 events/hr as favorable. This may be important in aging cohorts, such as MESA, where low level OSA is common (38). The estimated U.S. prevalence of OSA near our threshold (AHI ≥ 15) in midlife-older men and women (aged 50–70 years) was 17.4% and 9.1%, respectively, and at a more stringent criterion (AHI ≥ 5) was estimated to be 43.2% and 27.8%.

For %N3 and %REM, normative values for %N3 and %REM do not yet exist for older adults (65+). For this reason, the National Sleep Foundation’s panel could not reach consensus (80%) for sleep architecture values for this age group. In our analyses we therefore applied adult thresholds universally to adults (<65). Establishing normative sleep ranges for late adulthood may provide guidelines for creating consistent scores across populations.

We coded the nearest whole number to the lowest quartile as favorable for the Fragmentation Index (FI ≤ 15). FI is defined, in essence, by the number of sleep periods less than 1 minute in length, with wake periods before and after this sleep period – thus, a Fragmentation Index (see Figure A2). From the MESA-Sleep data dictionary, FI is the “sum of Percent Mobile [wake] epochs and Percent Immobile Bouts Less than 1-Minute Duration to the number of Immobile Bouts, for the given interval.” A validated threshold does not yet exist for the Fragmentation Index (FI). Sleep Maintenance Efficiency was dichotomized at 90%. The most favorable quartile was approximately 93–94%, which we regarded as too extreme for an older adult population with a possibly elevated normative OSA prevalence and mid-sleep awakenings. Further research is needed on more precise cut points for these dimensions by validating against external criteria.

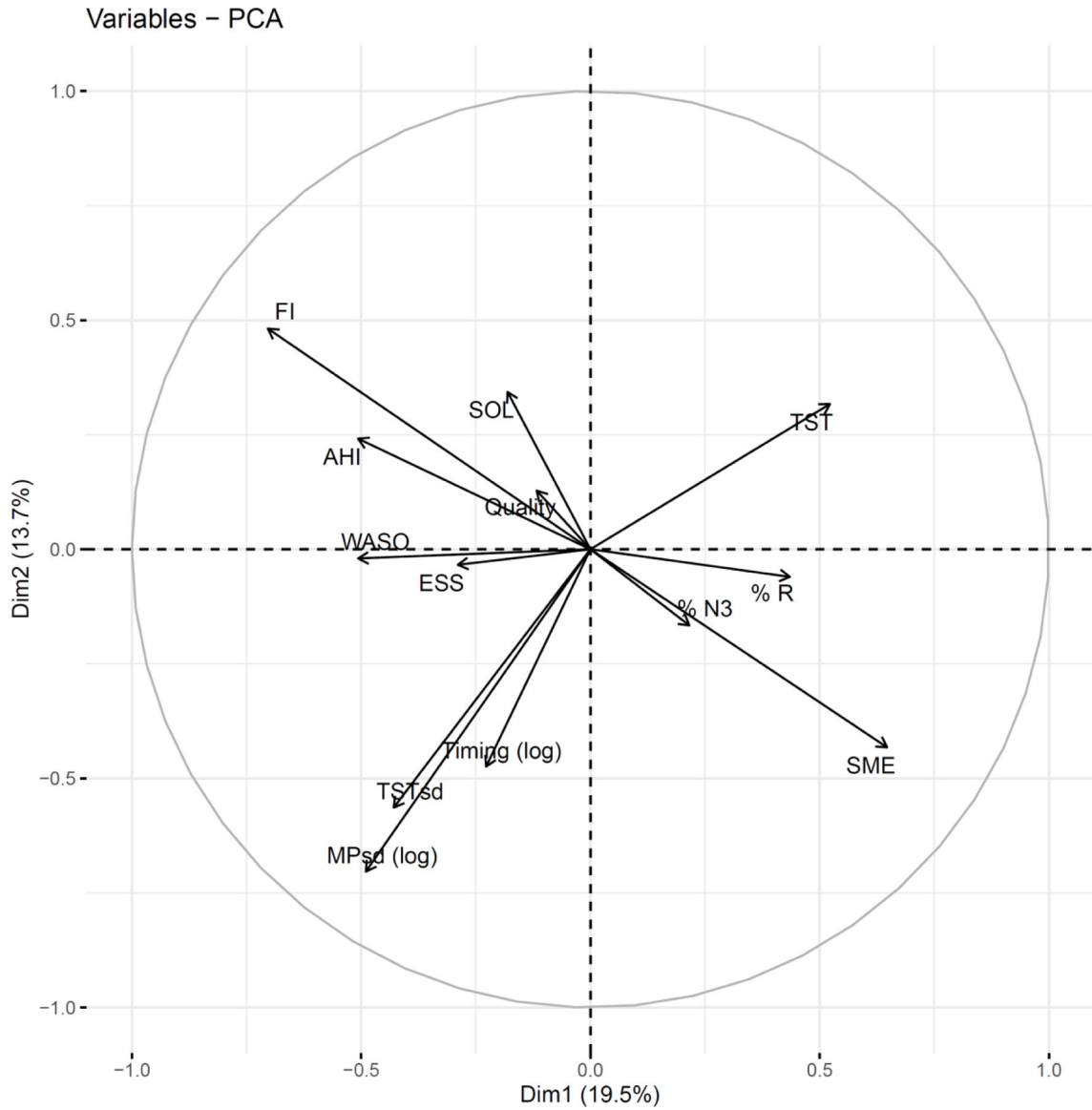


Figure A3. variable loading (190mm)

A4. SHS-PC1

We aimed to create a data-driven unidimensional sleep score and used PCA. Of 13 sleep variables, “Timing” and “Midpoint standard deviation” (MPsd) were log-transformed for normality. For PCA-derived Sleep Scores, we **replaced** Difficulties Initiating Sleep that we used in the Sleep Health Score with actigraphy-estimated Sleep Onset Latency (SOL_{act}) for PC scores because its distribution was more amenable to inclusion for PCA. Our measure of the wake-sleep transition–frequency of Difficulties Initiating Sleep (DIS) assessed by the ‘trouble sleeping’ item from the WHIIRS–was overly skewed and not improved by log-transformation.

Table 3 described the 13 variables used to calculate PC scores for: Ru SATED (PC1), Ru SATED + OSA and architecture (PC1), and Sleep Health Score (PC1). Each variable was standardized, mean=0, standard deviation=1. PC1 scores are therefore interpreted relative to the sample, in our case, an aging adult sample (mean age = 59.4 years). We conducted PCA on sets of standardized variables: 6 variables for Ru SATED, 9 variables for Ru SATED + OSA and architecture, and 13 variables for Sleep Health Score. We ran a sensitivity analysis on a 10-variable version of the SHS (PC1) that eliminated potentially redundant measures such as TST regularity, and WASO (because we already assessed maintenance efficiency, SME). The results are not substantively different, and we focus on the main PC scores.

We evaluated the first 3 Principal Components for each run of PCA. We first evaluated variable loadings in the first three PCs for Ru SATED. We then evaluated variable loadings for the first three PCs for Ru SATED + OSA and architecture as well as the first three PCs for the SHS-PC1. These were reported in Table 3. We extracted the first Principal Component for each PC analysis, resulting in 3 global sleep scores derived from PCA: Ru SATED (PC1), Ru SATED + OSA and architecture (PC1), and the Sleep Health Score (PC1). Each global score was standardized, mean=0, sd=1.

We regarded the first dimension to be interpretable as a unidimensional sleep health score. As Table 3 and Figure A3 show, higher values on the PC dimensions given sleep dimensions – such as TST, %N3, %R, and SME – are generally regarded as favorable for later health along dimension 1 but not along dimension 2. Similarly, higher irregularities in timing (MPsd) and duration (TSTsd), and AHI are regarded as unfavorable for later health in dimension 1 but not 2. Quality (quality) is a Likert scale, higher is worse. We aimed for a unidimensional scale relevant for sleep and later health and used and interpreted the first dimension as our composite sleep score.

Figure 2, the correlation matrix of sleep dimensions, further supports the interpreted directionality of our measure: dimensions in which higher values are accepted as favorable for health tend to positively correlate (e.g. TST, SME), and dimensions of the contrary interpretation also tend to show varying degrees of similarly directed correlation (e.g. MPsd, TSTsd, FI, AHI, %REM).

The strength of the SHS-PC1 is that it represents a single, reduced dimension that explains the highest proportion of variance (19.5%) and whose directionality allows a ‘health’ or ‘favorability’ interpretation. The limitation of the SHS-PC1 is that the highest proportion of variance is 19.5%, leaving 80.5% of the variance unexplained by dimension 1. Examining the potential value of this unexplained variance was outside of the scope of this article, and dimensions 2–13 may contain useful composite sleep health information.

References

1. Wallace ML, Stone K, Smagula SF, Hall MH, Simsek B, Kado DM, et al. Which sleep health characteristics predict all-cause mortality in older men? An application of flexible multivariable approaches. *Sleep*. 2018;41(1):zxx189.
2. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Current opinion in lipidology*. 2002;13(1):3–9. [PubMed: 11790957]

3. Buysse DJ. Sleep health: can we define it? Does it matter? *Sleep*. 2014;37(1):9–17. [PubMed: 24470692]
4. Larson JS. The World Health Organization's definition of health: Social versus spiritual health. *Social indicators research*. 1996;38(2):181–92.
5. Knutson KL, Phelan J, Paskow MJ, Roach A, Whiton K, Langer G, et al. The National Sleep Foundation's sleep health index. *Sleep health*. 2017;3(4):234–40. [PubMed: 28709508]
6. Chung J, Goodman MO, Johnson DA, Wallace MW, Huang TI, Bertisch SM, Redline SR. . Racial/ethnic Differences in Actigraphy, Questionnaire, and Polysomnography Indicators of Healthy Sleep: The Multi-Ethnic Study of Atherosclerosis. *American journal of epidemiology* (in press). 2021.
7. Zhang GQ, Cui L, Mueller R, Tao S, Kim M, Rueschman M, et al. The National Sleep Research Resource: towards a sleep data commons. *J Am Med Inform Assoc*. 2018;25(10):1351–8. [PubMed: 29860441]
8. Bertisch SM, Sillau S, De Boer IH, Szklo M, Redline S. 25-hydroxyvitamin D concentration and sleep duration and continuity: multi-ethnic study of atherosclerosis. *Sleep*. 2015;38(8):1305–11. [PubMed: 25669179]
9. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *sleep*. 1991;14(6):540–5. [PubMed: 1798888]
10. Levine DW, Lewis MA, Bowen DJ, Kripke DF, Kaplan RM, Naughton MJ, et al. Reliability and validity of Women's Health Initiative Insomnia Rating Scale. *Psychological assessment*. 2003;15(2):137. [PubMed: 12847774]
11. Ohayon M, Wickwire EM, Hirshkowitz M, Albert SM, Avidan A, Daly FJ, et al. National Sleep Foundation's sleep quality recommendations: first report. *Sleep health*. 2017;3(1):6–19. [PubMed: 28346153]
12. Ravits SG, Dzierzewski JM, Perez E, Donovan EK, Dautovich ND. Sleep Health as Measured by RU SATED: A Psychometric Evaluation. *Behavioral Sleep Medicine*. 2019:1–9.
13. Peppard PE, Young T, Barnett JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *American journal of epidemiology*. 2013;177(9):1006–14. [PubMed: 23589584]
14. Brindle RC, Cribbet MR, Samuelsson LB, Gao C, Frank E, Krafty RT, et al. The relationship between childhood trauma and poor sleep health in adulthood. *Psychosomatic medicine*. 2018;80(2):200. [PubMed: 29215455]
15. Dong L, Martinez AJ, Buysse DJ, Harvey AG. A composite measure of sleep health predicts concurrent mental and physical health outcomes in adolescents prone to eveningness. *Sleep health*. 2019;5(2):166–74. [PubMed: 30928117]
16. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's updated sleep duration recommendations: final report. *Sleep Health*. 2015;1(4):233–43. [PubMed: 29073398]
17. Huang T, Mariani S, Redline S. Sleep irregularity and risk of cardiovascular events: the multi-ethnic study of atherosclerosis. *Journal of the American College of Cardiology*. 2020;75(9):991–9. [PubMed: 32138974]
18. Huang T, Redline S. Cross-sectional and prospective associations of actigraphy-assessed sleep regularity with metabolic abnormalities: The multi-ethnic study of atherosclerosis. *Diabetes care*. 2019;42(8):1422–9. [PubMed: 31167888]
19. Wallace ML, Yu L, Buysse DJ, Stone KL, Redline S, Smagula SF, et al. Multidimensional sleep health domains in older men and women: an actigraphy factor analysis. *Sleep*. 2020.
20. Kessler RC. The categorical versus dimensional assessment controversy in the sociology of mental illness. *Journal of health and social behavior*. 2002:171–88. [PubMed: 12096698]
21. Liu Y, Wheaton AG, Chapman DP, Cunningham TJ, Lu H, Croft JB. Prevalence of healthy sleep duration among adults—United States, 2014. *Morbidity and Mortality Weekly Report*. 2016;65(6):137–41. [PubMed: 26890214]
22. Wallace ML, Buysse DJ, Redline S, Stone KL, Ensrud K, Leng Y, et al. Multidimensional sleep and mortality in older adults: a machine-learning comparison with other risk factors. *The Journals of Gerontology: Series A*. 2019;74(12):1903–9.

23. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? *Epidemiology*. 2008;838–45. [PubMed: 18854708]
24. Jackson CL, Patel SR, Jackson WB, Lutsey PL, Redline S. Agreement between self-reported and objectively measured sleep duration among white, black, Hispanic, and Chinese adults in the United States: Multi-Ethnic Study of Atherosclerosis. *Sleep*. 2018;41(6):zsy057.
25. Purcell S, Manoach D, Demanuele C, Cade B, Mariani S, Cox R, et al. Characterizing sleep spindles in 11,630 individuals from the National Sleep Research Resource. *Nature communications*. 2017;8(1):1–16.
26. Djonlagic I, Mariani S, Fitzpatrick AL, Van Der Klei VM, Johnson DA, Wood AC, et al. Macro and micro sleep architecture and cognitive performance in older adults. *Nature human behaviour*. 2020:1–23.
27. Arnal PJ, Thorey V, Ballard ME, Hernandez AB, Guillot A, Jourde H, et al. The Dreem headband as an alternative to polysomnography for EEG signal acquisition and sleep staging. *bioRxiv*. 2019:662734.
28. Meltzer LJ, Williamson AA, Mindell JA. Pediatric sleep health: It matters, and so does how we define it. *Sleep Medicine Reviews*. 2021:101425. [PubMed: 33601324]
29. Greco S, Ishizaka A, Tasiou M, Torrisi G. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*. 2019;141(1):61–94.
30. Kubala AG, Buysse DJ, Brindle RC, Krafty RT, Thayer JF, Hall MH, et al. The association between physical activity and a composite measure of sleep health. *Sleep and Breathing*. 2020:1–8.
31. Mireia D, Benítez I, Esther S-B, Garcia-Codina O, Medina-Bustos A, Joan E, et al. Impact of sleep health on self-perceived health status. *Scientific Reports (Nature Publisher Group)*. 2019;9(1).
32. Wallace ML, Buysse DJ, Germain A, Hall MH, Iyengar S. Variable selection for skewed model-based clustering: application to the identification of novel sleep phenotypes. *Journal of the American Statistical Association*. 2018;113(521):95–110. [PubMed: 31086426]
33. Wallace ML, Lee S, Hall MH, Stone KL, Langsetmo L, Redline S, et al. Heightened sleep propensity: a novel and high-risk sleep health phenotype in older adults. *Sleep Health*. 2019;5(6):630–8. [PubMed: 31678177]
34. Brindle RC, Yu L, Buysse DJ, Hall MH. Empirical derivation of cutoff values for the sleep health metric and its relationship to cardiometabolic morbidity: results from the Midlife in the United States (MIDUS) study. *Sleep*. 2019;42(9):zsz116. [PubMed: 31083710]

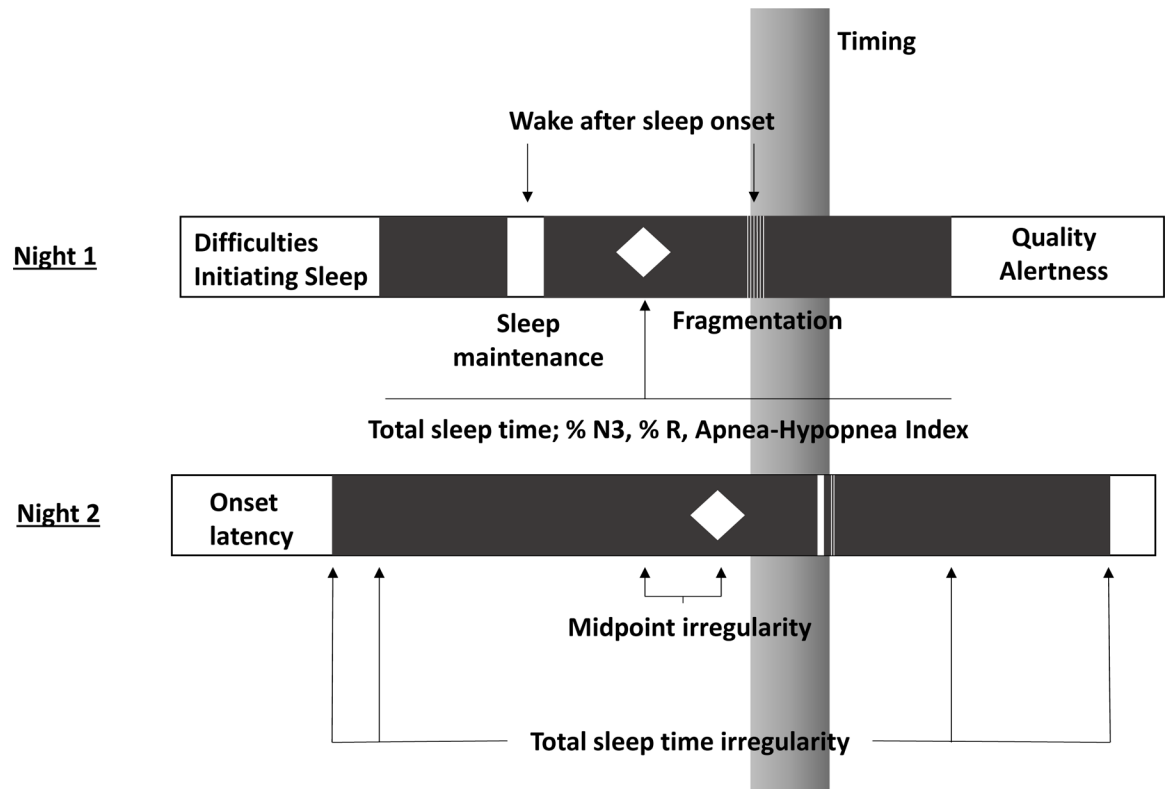


Figure 1.
Conceptual model of sleep health

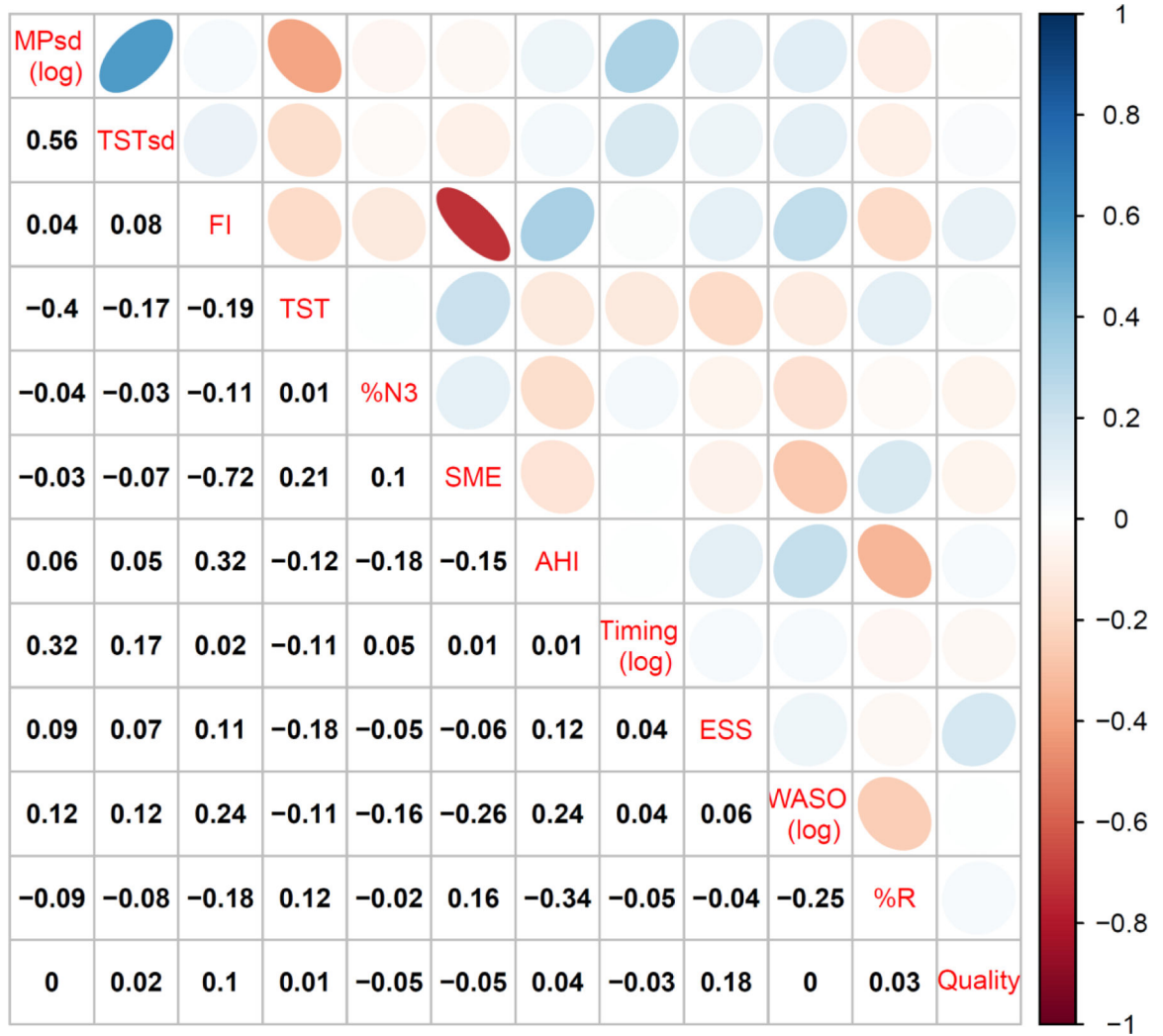


Figure 2.
Correlation matrix (190mm)

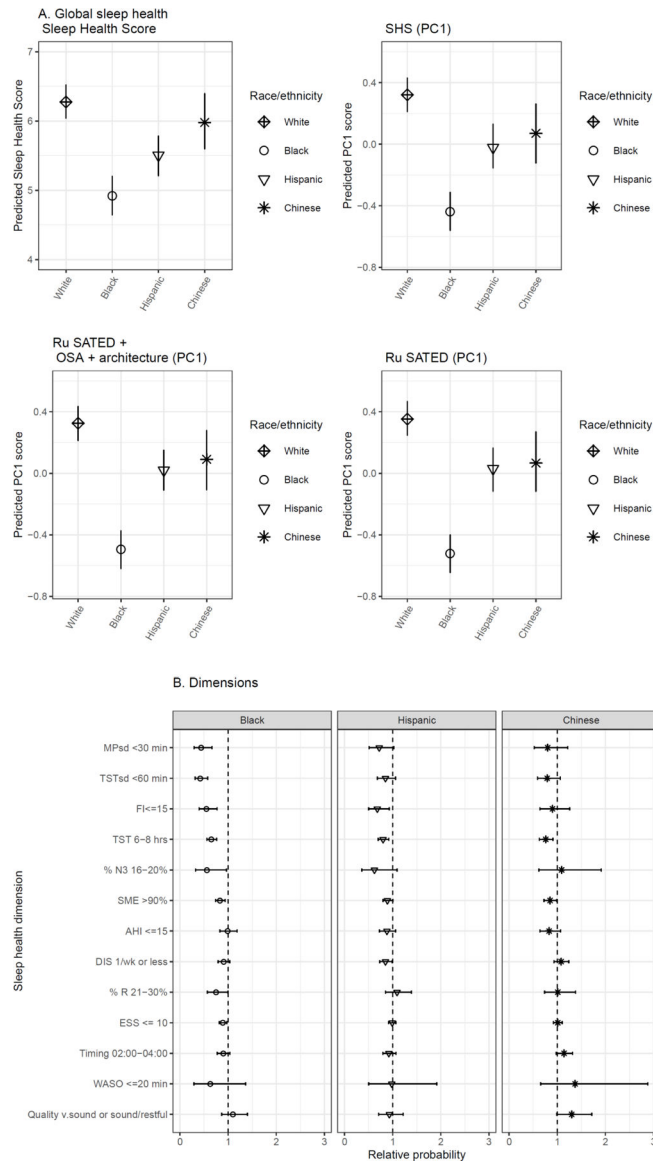


Figure 3.
 (A) Composite disparity (140mm)
 (B) race-eth dimension disparities (140mm)

Table 1.

Sleep variables, dimensions, and dichotomization rules. The Multi-Ethnic Study of Atherosclerosis, age <65 years (n=735).

Sleep variable	Sleep health dimension	Measure	Dichotomization rule (coded as '1')	Sample prevalence of favorability ('1'), n (%)	Abbreviation
Midpoint variability (sd, min) ^d	Regularity ^d	Actigraphy	<30 minutes	157 (21.4%)	MPsd
Duration variability (sd, min) ^f	Regularity	Actigraphy	<60 minutes	271 (36.9%)	TSTsd
Quality (WHIIRS Likert subscale; higher is increasing complaints) ^d	Satisfaction ^d	Questionnaire	“Sound and restful” or “very sound or restful”	267 (36.3%)	Quality
Sleepiness ^d	Alertness ^d	Questionnaire	10	617 (83.9%)	ESS
Timing (midpoint) ^{b,d}	Timing ^d	Actigraphy	02:00–04:00 ^b	466 (63.4%)	Timing
Sleep maintenance efficiency (%) ^d	Efficiency ^d	Actigraphy	>90%	524 (71.3%)	SME
Fragmentation Index ^c	Efficiency	Actigraphy	15 ^a	199 (27.1%)	FI
Wake after sleep onset (min) ^f	Efficiency	Actigraphy	20 minutes	50 (6.8%)	WASO
Total sleep time ^{b,d} (min)	Duration ^d	Actigraphy	6–8 hours ^b	461 (62.7%)	TST
%R ^{b,e}	Architecture	PSG	21%–30% ^b	251 (34.1%)	% R
%N3 ^{b,e}	Architecture	PSG	16%–20% ^b (50)	84 (11.4%)	% N3
Apnea-Hypopnea Index (events/hr) ^e	Obstructive sleep apnea	PSG	15 events/hr	354 (48.2%)	AHI
Sleep onset latency (difficulties initiating sleep – WHIIRS subscale) ^f	Wake-sleep transition	Questionnaire	“Less than once a week” or “no, not in the past 4 weeks”	476 (64.8%)	DIS
Sleep onset latency (min) ^g	Wake-sleep transition	Actigraphy	-		SOL _{act}

^aWHIIRS: Women’s Health Initiative Insomnia Rating Scale. PSG: polysomnography. MPsd: midpoint standard deviation. TSTsd: total sleep time standard deviation. FI: the Fragmentation Index. TST: total sleep time. % N3: % non-rapid eye movement, stage 3 sleep. SME: sleep maintenance efficiency. AHI: the Apnea-Hypopnea Index. ESS: Epworth Sleepiness Scale. WASO: wake after sleep onset. % R: % rapid eye movement sleep.

^bThese dimensions have optimal ranges in which overexpression of this dimension is considered problematic.

^cThe cut-point for FI was defined as the nearest whole number to the most favorable quartile (lowest quartile).

^dVariables used in our operationalization of canonical Ru SATED (PC1)

^eIn addition to Ru SATED (PC1), variables used in Ru SATED + OSA/Architecture (PC1)

^fIn addition to Ru SATED + OSA/Architecture (PC1), variables used in the Sleep Health Score

^gActigraphy-assessed onset latency used in all Principal Components-based sleep scores, in place of self-reported difficulties initiating sleep (see Appendix, and Table 3).

Table 2.

MESA-Sleep Participant socio-demographics, global sleep health, and sleep health metrics by Sleep Health Score. The Multi-Ethnic Study of Atherosclerosis, ages <65 years (n=735).

	Sleep Health Score (SHS) categorized: least to most favorable SHS					p
	Overall	least (0–3)	less (4–6)	more (7–9)	most (10–12)	
N	735	120	378	205	32	
Socio-demographics						
Race-ethnicity (%)						<0.001
White	265 (36.1%)	26 (21.7%)	123 (32.5%)	96 (46.8%)	20 (62.5%)	
Chinese	93 (12.7%)	11 (9.2%)	49 (13.0%)	30 (14.6%)	3 (9.4%)	
Black	200 (27.2%)	52 (43.3%)	108 (28.6%)	38 (18.5%)	2 (6.2%)	
Hispanic	177 (24.1%)	31 (25.8%)	98 (25.9%)	41 (20.0%)	7 (21.9%)	
Female (%)	409 (55.6%)	55 (45.8%)	205 (54.2%)	126 (61.5%)	23 (71.9%)	0.011
Age	59.4 (3.0)	59.3 (3.0)	59.6 (3.0)	59.2 (2.9)	59.4 (2.9)	0.638
Education						0.082
Less than high school	71 (9.7%)	11 (9.2%)	43 (11.4%)	14 (6.8%)	3 (9.4%)	
High school or some college	352 (47.9%)	57 (47.5%)	191 (50.5%)	95 (46.3%)	9 (28.1%)	
College degree	174 (23.7%)	34 (28.3%)	79 (20.9%)	52 (25.4%)	9 (28.1%)	
Graduate	138 (18.8%)	18 (15.0%)	65 (17.2%)	44 (21.5%)	11 (34.4%)	
Married	480 (65.3%)	65 (54.2%)	248 (65.6%)	141 (68.8%)	26 (81.2%)	0.010
Employed	234 (31.8%)	42 (35.0%)	124 (32.8%)	59 (28.8%)	9 (28.1%)	0.615
Global sleep health						
Sleep Health Score [†]	5.7 (2.1)	2.6 (0.7)	5.1 (0.8)	7.8 (0.8)	10.3 (0.5)	<0.001
Sleep Health Score (PC1) [†]	0 (1)	-1.1 (0.9)	-0.2 (0.8)	0.8 (0.5)	1.4 (0.4)	<0.001
Ru SATED (PC1) [†]	0 (1)	-1.0 (0.9)	-0.1 (0.9)	0.6 (0.7)	1.2 (0.5)	<0.001
Ru SATED + OSA and architecture (PC1) [†]	0 (1)	-1.2 (0.9)	-0.1 (0.8)	0.7 (0.6)	1.4 (0.4)	<0.001
Sleep health variables						
Regularity						
Midpoint variability ^e (sd, min) [‡]	47.2 [31.8, 70.6]	67.8 [47.9, 95.2]	53.0 [36.9, 76.1]	35.3 [25.5, 50.8]	23.2 [16.7, 26.7]	<0.001
Duration variability ^e (sd, min) [‡]	76.4 (36.8)	90.1 (29.9)	83.2 (38.0)	61.6 (32.0)	39.6 (12.8)	<0.001
Satisfaction						
Quality ^c (Likert subscale ^d ; higher=increasing complaints) [‡]	2.8 (0.9)	3.3 (0.8)	2.8 (0.9)	2.4 (0.8)	2.3 (0.7)	<0.001
Alertness/sleepiness						
Epworth Sleepiness Scale ^c ↓	6.3 (4.2)	9.0 (4.9)	6.2 (4.1)	5.4 (3.4)	4.1 (2.3)	<0.001
Timing						
Timing ^e (midpoint minutes from midnight) [‡]	196.1 [155.1, 242.5]	230.8 [151.2, 278.8]	201.1 [157.9, 254.6]	182.1 [151.2, 222.3]	177.5 [151.2, 205.1]	0.001

	Sleep Health Score (SHS) categorized: least to most favorable SHS					p
	Overall	least (0–3)	less (4–6)	more (7–9)	most (10–12)	
Efficiency						
Sleep Maintenance Efficiency ^e (%) [†]	91.4 (3.4)	89.2 (4.3)	91.1 (3.1)	93.0 (2.2)	93.2 (1.8)	<0.001
Fragmentation Index ^e ↓	19.3 (6.8)	23.3 (7.4)	20.3 (6.7)	15.7 (4.7)	14.7 (4.0)	<0.001
Wake after Sleep Onset ^e (min) ↓	76.8 (56.3)	102.8 (63.2)	81.9 (58.5)	57.3 (40.0)	45.5 (33.1)	<0.001
Duration						
Total Sleep Time ^a (hrs) [↗]	6.4 (1.2)	5.6 (1.4)	6.4 (1.2)	6.9 (0.9)	7.2 (0.6)	<0.001
Sleep architecture						
% R ^b ↗	19.0 (6.6)	15.5 (6.9)	18.4 (6.5)	21.2 (5.7)	23.8 (4.0)	<0.001
% N3 ^b ↗	9.7 [3.2, 16.7]	8.2 [3.2, 14.7]	8.6 [1.9, 15.9]	10.9 [4.9, 17.6]	18.6 [11.9, 20.0]	<0.001
Obstructive sleep apnea						
Apnea-Hypopnea Index ^b (events/hr) ↓	15.5 [7.3, 30.35]	29.5 [16.3, 48.7]	17.8 [9.1, 32.0]	9.6 [5.3, 16.5]	6.8 [3.7, 11.2]	<0.001
Wake-sleep transition – insomnia complaint						
Onset latency ^c (freq. of difficulties initiating sleep)						<0.001
No, not in the past 4 weeks	405 (55.1%)	33 (27.5%)	203 (53.7%)	142 (69.3%)	27 (84.4%)	
Yes, less than once a week	71 (9.7%)	6 (5.0%)	28 (7.4%)	34 (16.6%)	3 (9.4%)	
Yes, 1–2 times a week	132 (18.0%)	37 (30.8%)	74 (19.6%)	21 (10.2%)	0 (0.0%)	
Yes, 3–4 times a week	66 (9.0%)	17 (14.2%)	41 (10.8%)	7 (3.4%)	1 (3.1%)	
Yes, 5+ times a week	61 (8.3%)	27 (22.5%)	32 (8.5%)	1 (0.5%)	1 (3.1%)	
Wake-sleep transition (actigraphy)						
Onset latency ^e (actigraphy)	4.2 (3.0)	4.6 (3.8)	4.1 (2.9)	4.1 (2.7)	3.6 (2.1)	0.326

^a↑ = Higher is better; ↓ = lower is better; ↗ = generally higher is better, but overexpression can be problematic (eg long sleepers, excessive % R).
 PCI: principal component 1. sd: standard deviation. OSA: obstructive sleep apnea.

^b polysomnography

^c self-report

^d Women's Health Initiative Insomnia Rating Scale, subscale

^e actigraphy

Table 3.

Principal Components Analysis of sleep health metrics: Ru SATED, Ru SATED + OSA and architecture, Sleep Health Score (SHS), SHS-parsimonious

		Ru SATED			Ru SATED + OSA and architecture			SHS-PC1			SHS-PC1 (parsimonious)		
		$\alpha= 0.42$			$\alpha= 0.47$			$\alpha= 0.61$			$\alpha= 0.45$		
Percent of Variance		27.72%	20.01%	17.06%	20.40%	15.09%	13.01%	19.55%	13.71%	9.32%	18.39%	14.06%	11.75%
	Also known as	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
Ru SATED													
(Ir)regularity [↓] (timing)	MPSd	-0.59	0.26	-0.12	-0.45	-0.45	0.06	-0.31	-0.53	0.11	-0.45	-0.43	0.10
Satisfaction/quality [↓] (increasing insomnia symptoms)	Quality	-0.07	-0.61	-0.45	-0.08	0.22	0.63	-0.07	0.10	0.50	-0.08	0.23	0.60
Alertness [↓] (increasing ESS)	ESS	-0.31	-0.51	-0.31	-0.29	0.10	0.53	-0.18	-0.03	0.37	-0.29	0.10	0.50
Timing [↓] (increasing delay)	Timing	-0.39	0.41	-0.35	-0.26	-0.49	-0.04	-0.14	-0.36	0.00	-0.25	-0.46	0.01
Efficiency [↑] (maintenance)	SME	0.24	0.36	-0.69	0.32	-0.29	0.06	0.41	-0.32	-0.09	0.33	-0.32	0.05
Duration [↗]	TST	0.58	0.05	-0.29	0.50	0.21	-0.07	0.33	0.24	-0.03	0.49	0.24	-0.07
PSG (architecture)													
Apnea-Hypopnea Index ⁺	AHI				-0.38	0.43	-0.23	-0.32	0.18	-0.29	-0.38	0.38	-0.27
% N3 [↗]	% N3				0.14	-0.34	-0.18	0.14	-0.12	0.08	0.15	-0.31	-0.14
% R [↗]	% R				0.35	-0.25	0.48	0.27	-0.05	0.53	0.35	-0.17	0.51
SHS additions													
Irregularity [↓] (duration)	TS Tsd							-0.27	-0.42	0.15			
Onset latency [↓]	SO Lact							-0.11	0.26	0.33	-0.06	0.34	0.12
Fragmentation [↓]	FI							-0.44	0.36	0.11			
Wake after sleep onset [↓]	WASO							-0.32	-0.01	-0.28			

^aRu SATED: **R**egularity **S**atisfaction **A**lertness **T**iming **E**fficiency **D**uration. SHS: Sleep Health Score. SHS (PC1): Sleep Health Score principal component 1. Ru SATED + OSA and architecture: **R**egularity **S**atisfaction **T**iming **E**fficiency **D**uration + obstructive sleep apnea + sleep architecture. ESS: Epworth Sleepiness Scale. % N3: % non-rapid eye movement, stage 3 sleep. % R: % rapid eye movement sleep. PSG: polysomnography.

^b↑ = Higher is better; ↓ = lower is better; ↗ = generally higher is better, but overexpression can be problematic (eg long sleepers, excessive % R)