

Multidimensional Sleep Health: 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}.

1. Brigham and Women's Hospital
2. Harvard Medical School

Abstract

The new paradigm of multidimensional sleep health ('sleep health') offers both challenges and opportunities for sleep science. Buysse (2014) has described sleep health to be multidimensional, framed as positive attributes, operationalizable into composite measures of global sleep health, sensitive to upstream exposures, and consequential for downstream health. We highlight two paradigm-shifting effects of a multidimensional sleep health perspective. The first is the use of composite sleep metrics which i) enable quantification of population shifts in sleep health, ii) with possibly reduced measurement error, iii) greater statistical stability, and iv) reduced multiple-testing burdens. The second is that sleep dimensions do not occur in isolation, that is, they are commonly biologically or statistically dependent. These dependencies complicate hypothesis tests yet can be leveraged to inform scale construction, model interpretation, and inform targeted interventions. To illustrate these points, we i) extended Buysse's Ru SATED model; ii) constructed a conceptual model of sleep health; and iii) showed exemplar analyses from the Multi-Ethnic Study of Atherosclerosis (n=735). Our findings support that sleep health is a distinctively useful paradigm to facilitate interpretation of a multitude of sleep dimensions. Nonetheless, the field of sleep health is still undergoing rapid development and is currently limited by: i) a lack of evidence-based cut-offs for defining optimal sleep health; ii) longitudinal data to define utility for predicting health outcomes; and iii) methodological research to inform how to best combine multiple dimensions for robust and reproducible composites.

Key words: sleep, sleep health, actigraphy, polysomnography, multidimensional

1 **Multidimensional Sleep Health: Concepts, Advances, and Implications for Research and** 2 **Intervention.**

3 *Introduction*

4 Although it would be convenient if we had a single variable in our data sets called ‘sleep’
5 that encompasses the complex neurophysiologic process of sleep, what data sets typically
6 include is a wide range of metrics *about* sleep: estimates of specific features of sleep as well as
7 the impact of sleep on daytime symptoms and function. Each feature or *dimension* of sleep is
8 often modeled separately and interpreted independently – as exposure or outcome – in distinct
9 models: an effect size for total sleep time, one for sleepiness, and yet another for sleep
10 efficiency.

11 However, sleep dimensions do not exist in isolation, and sleep health is multidimensional
12 (1). Just as food is not consumed as singular nutrients (2), sleep is not experienced as singular
13 dimensions. Adults who meet national recommendations of 7-9 hours of duration may evince
14 vast heterogeneity in sleep latency, continuity, macro- and micro-architecture, and satisfaction.
15 Moreover, a substantial body of evidence suggests that a variety of sleep dimensions influence
16 academic outcomes, mental and physical health, illness, and mortality; and that these dimensions
17 are sensitive to exposures at the individual level (*e.g. age; physiological shifts over the life-*
18 *course, such as menopause; financial strain; caffeine, tobacco, and other substance use;*
19 *physical activity, genetic risk loci, etc.*) and social and environmental levels (*e.g. neighborhoods*
20 *and the built environment; social cohesion; social relationships; socioeconomic status in*
21 *childhood; marital status; loneliness; discrimination; and work schedules, etc.*). Accordingly,
22 Buysse (2014) has defined sleep health as: “a multidimensional pattern of sleep-wakefulness,

1 adapted to individual, social, and environmental demands, that promotes physical and mental
2 well-being” (1).

3 A multidimensional paradigm is, however, already intrinsic to our understanding of sleep.
4 Clinicians, for instance, routinely consider multiple aspects of sleep in diagnosis and
5 management of sleep disorders. Moreover, the literature has long reflected interest in the
6 combined, including interactive, aspects of sleep, although has focused more on sleep
7 disturbances rather than sleep health, and focused on clinical disorders rather than population
8 health. Thus: What is distinctive and useful about the paradigm of ‘multidimensional sleep
9 health’? How does it differ from prior and current work? Does ‘multidimensional sleep health’
10 constitute a paradigm shift, that is, a new *way* of thinking about sleep?

11 We provide empirical data to support multidimensional sleep health as a distinctively
12 useful approach for characterizing ‘sleep health’ across the population. One innovation is the
13 quantification of sleep **health**, which ranges beyond disorder or insufficiency – acknowledging
14 the existence of gradients of “healthy sleep” beyond meeting a minimum duration of sleep or an
15 absence of disorders such as insomnia or obstructive sleep apnea (OSA) (1). Sleep health also
16 parallels broader paradigm shifts which acknowledge the World Health Organization’s definition
17 of health as more than the absence of illness (3).

18 While these are important frameshifts, sleep health has other distinctive but less
19 emphasized features, which render it a more radical and useful paradigm shift than would seem
20 at first glance. First, sleep health approaches multiple sleep dimensions (conceptually,
21 operationally, analytically) as contributors to a unitary or a *composite* concept – a metric of
22 global sleep health (1, 4). Extant scales include Buysse’s Ru SATED scale and the National
23 Sleep Foundation’s Sleep Health Index (5, 6). Consistent and valid scales are necessary for

1 scalable, longitudinal data collection that enables quantification of population shifts in global
2 sleep health. Such work does not preclude analyzing each dimension in its own right, although
3 statistical dependence among dimensions complicates model interpretation because sleep
4 dimensions do not exist in isolation.

5 Second, sleep health emphasizes that sleep dimensions do not exist in isolation (4, 7).
6 There are physiological reasons why certain dimensions are correlated, perhaps causally, with
7 others, for instance, duration and sleepiness, OSA and continuity, timing irregularity and
8 duration, circadian body-temperature nadir and sleep onset latency, et al. (8-11). Sleep
9 dimensions also may show statistical correlation when derived from common measurement tools.
10 Thus, interpretation of individual dimensions as if they were independent may lead to erroneous,
11 or partial conclusions about ‘sleep’: what kind of continuity, quality, alertness, and regularity is
12 likely to be seen among ‘sufficient’ sleepers (however defined) as opposed to ‘insufficient’
13 sleepers?

14 The purpose of this study was to illustrate these two frameshifts: i) unidimensional
15 composites; and ii) consideration of correlations. We explored the utility of expanding Buysse’s
16 model by incorporating sleep stage information (% N3, % Rapid Eye Movement [NR]), the
17 Apnea-Hypopnea Index[AHI], self-reported frequency of long sleep onset latency (an insomnia
18 complaint), and regularity in duration (12, 13) by mapping these dimensions to a conceptual
19 sleep health model. We considered how to model multiple dimensions and found merit in both
20 analyzing individual dimensions as well as a global sleep health metric: from a public health
21 perspective, it was important to know which sleep dimensions are the drivers of a global sleep
22 effect. Finally, we showed how these composite and individual dimensions map across
23 race/ethnic groups, an extension of prior work (14).

1 **Methods**

2 Sample: The sample was from the Multi-Ethnic Study of Atherosclerosis (MESA), a 6-
3 community cohort of aging adults, diverse in race/ethnicity (White, Black, Hispanic, Chinese).
4 Details on this cohort are published (15). In brief, at Exam 5 (2010-2013), participants were
5 invited to the MESA-Sleep ancillary study and underwent single-night at-home
6 polysomnography (PSG), 7-day wrist actigraphy (Actiwatch Spectrum; Philips Respironics, PA;
7 Actiware-Sleep v 5.59), and validated sleep questionnaires (Epworth Sleepiness Scale, Women's
8 Health Initiative Insomnia Rating Scale) (16, 17). A previous racial-ethnic sleep disparities
9 project included adults 54-93 years old (14); in following analyses, only adults 54-64 years were
10 included, because there are more normative data available for this age range (18).

11 Sleep health conceptual model: As described previously (14), sleep health domains were
12 drawn primarily from Ru SATED but also the National Sleep Foundation's Sleep Health Index
13 (6). Selection of additional variables was informed by prior knowledge of a high prevalence of
14 OSA and sleep fragmentation in middle-age and older adults (19). Thus, several continuity
15 metrics were chosen *a priori*: sleep maintenance efficiency (SME), the Fragmentation Index (FI),
16 and Wake after Sleep Onset (WASO). Other additions to Ru SATED such as % NR, % N3, AHI
17 (3% desaturation to define hypopneas), duration regularity, and self-reported frequency of
18 un/desirable onset latency (freq. of difficulties initiating sleep) were also selected *a priori*, using
19 expert knowledge or evidence from the literature indicating their relevance for health outcomes
20 (20). Specifically, the mapping of each sleep variable to our conceptual model of sleep focused
21 on these four periods (Figure 1): 1) the transition from wake to sleep (onset latency, self-report)
22 and daytime sequelae (quality, alertness) from the participants' perspective; 2) the period
23 between sleep onset and offset evaluated objectively by PSG (eg, sleep architecture and/or

1 actigraphy (eg, fragmentation); 3) inter-daily variability in timing and duration; and 4) the entire
2 sleep experience both during and across nights (14).

3 Two strategies were used to construct sleep health composites: i) Principal Components
4 Analysis (PCA); and ii) a summary score of dichotomized sleep variables. PCA was conducted
5 on a) Ru SATED variables (midpoint regularity [MPsd], quality, timing (log-transformed),
6 maintenance efficiency, and duration); b) Ru SATED, expanded with selected variables from
7 PSG, adding AHI, % N3, %NR; c) all variables chosen *a priori* for the Sleep Health Score
8 (adding several variables measuring similar dimensions: duration regularity, sleep onset latency
9 (frequency of trouble sleeping), WASO, fragmentation) (14). The resulting PC scores were
10 standardized and coded so that higher values represent better health, and a 1-unit increase
11 corresponds to a 1-sd increase in sleep health. For PCA, actigraphy-assessed onset latency was
12 used rather than self-report due to the high skewness of the latter ordinal variable.

13 As described in (14), for the summary score (21-23), cut-points defined optimal ranges,
14 with coding of ‘favorable’ or ‘healthier’ sleep as ‘1’ and non-optimal ranges as ‘0’. Optimal
15 ranges were drawn from the literature, the NSF’s objective sleep quality report (eg, for WASO),
16 expert consensus, or sample characteristics (12, 13, 18, 19, 21). These dichotomous indicators of
17 favorable sleep were summed into a global metric, a Sleep Health Score (SHS), additionally
18 categorized as: least favorable (SHS 0-3), less (4-6), more (7-9), and most favorable (9-12) sleep
19 health.

20

21 Statistical Methods: Pearson correlations were computed for continuous sleep metrics
22 (log-transformed where necessary). Correlations were further investigated by inspecting
23 Principal Components for three composites: (1) Ru SATED; Ru SATED + OSA/Sleep

1 Architecture; and the Sleep Health Score (comprehensive indices across domains). Sensitivity
2 analyses assessed a parsimonious version of the Sleep Health Score (PC1) that eliminated
3 potentially redundant measures. Composite score internal reliability was assessed by alpha
4 Cronbach. Consistency in PC weights - both direction and magnitude - was evaluated across
5 each component. Trends among individual sleep health variables (e.g., duration, sleepiness) in
6 relation to SHS categories were reported. Global sleep health variations by racial/ethnic group
7 were assessed for potential utility for sleep health disparities research in the non-elderly.
8 Outcomes of composite sleep health scores (linear regression), and their dichotomized
9 components (modified Poisson regression (24)), were regressed on the exposure of race-ethnicity
10 (White=ref), with adjustment for age and sex. Analyses were conducted in R 3.6.3.

11

12 **Results**

13 The sample of 735 ethnically/racially diverse participants had an average age of $59.4 \pm$
14 3.0 years, 55.6% were female, 31.8% were currently employed, and most obtained at least a high
15 school degree (n=664, 90.3%) (Table 2). The majority met criteria for actigraphy-assessed
16 sufficient sleep (6-8 hrs; n=461, 62.7%; Table 1) but did not meet favorability thresholds for:
17 continuity (WASO (6.8% favorable), fragmentation (27.1%)), sleep architecture (% N3 [11.4%],
18 % NR [34.1%]), regularity (midpoint (21.4%) and duration (36.9%)), AHI (48.2%), and quality
19 (21.4%).

20 Figure 2a shows the Sleep Health Score, Sleep Health Score (PC1), Ru SATED +
21 OSA/Sleep Architecture (PC1), Ru SATED (PC1), and their components by race-ethnicity,
22 adjusting for age and sex. Consistent with a prior report (14), the sample averaged 5.7 ± 2.1 of 13
23 possible favorable dimensions on the SHS. Blacks averaged 1.34 fewer favorable domains than

1 Whites in global sleep health ($p < 0.001$). Hispanics averaged 0.74 fewer favorable domains
2 ($p < 0.001$); however, most of the dimension-level disparities in Hispanics overlap with the
3 reference. Ru SATED (PC1), Ru SATED + OSA/Sleep Architecture (PC1) showed similar
4 variation by race/ethnicity as the Sleep Health Score, and Sleep Health Score (PC1). Figure 2b
5 suggests that the largest drivers of global racial-ethnic disparities were sleep regularity (timing
6 and duration), continuity (fragmentation, maintenance efficiency), and total sleep time.

7 Figure 3 shows correlations among sleep variables within and across domains. Within the
8 continuity domain, fragmentation and sleep maintenance efficiency are highly correlated
9 ($\rho = 0.72$). Across domains, there are non-trivial statistical dependencies: timing regularity
10 (MPsd(log)) and sleep duration (TST) are moderately correlated ($\rho = 0.40$), as are timing
11 regularity with sleep timing (Timing(log)) ($\rho = 0.32$). AHI correlates with continuity metrics (FI,
12 SME, WASO) and % NR (ρ 's range: -0.34 to 0.32).

13 Table 2 suggests that when the components in Table 1 are summed into the SHS,
14 increasing Sleep Health Scores capture variable yet systematic positive shifts across all sleep
15 dimensions. Mean (median) differences in most-least favorable sleep health were notable in
16 duration (+1.6 hours), midpoint regularity (-43.8 minutes), duration regularity (-48.7 minutes),
17 AHI (-23.1 events/hr), % NR (+8.4%), % N3 (+9.9%), and WASO (-58.5 minutes).

18 Table 3 shows PC weights for the first three PCs for each composite score (variance
19 explained ranged from 43.0% for the SHS (PC1) to 66.3% for Ru SATED). The first PC for each
20 score is consistently interpretable in all models as a “sleep health score”: the directionality of
21 PC1 weights for individual variables concur with *a priori* knowledge of “better” sleep (e.g.,
22 higher AHI contributes negatively, higher TST contributes positively). As expected, the PC1
23 weights varied depending on variables included but did not substantively change the

1 interpretation of the PC1 composites. The simplest Ru SATED score shows that the measures of
2 irregular timing, absolute timing, and duration loaded on PC1 while satisfaction/alertness and
3 efficiency loaded on PC2. When measures from PSG and actigraphy were added to the Ru
4 SATED score, AHI tended to load with efficiency and irregularity, %NR loaded with
5 satisfaction/alertness, irregular duration loaded with irregular timing, and fragmentation loaded
6 with efficiency. Across the extended (SHS) scores, timing and regularity tended to load on PC2
7 while self-reported sleep indices loaded on PC3. Moderate internal consistency was observed for
8 different sleep health composites, with alpha Cronbach varying from 0.42 for Ru SATED to 0.61
9 for the full SHS.

10 **Discussion**

11 There is growing interest in considering sleep health as not merely the absence of a
12 disorder but as a summary of the positive attributes of healthy sleep. Similarly, there is
13 movement towards using a conceptual framework that articulates sleep as multidimensional,
14 potentially summarized as a composite measure (1). We further suggest the value in considering
15 the intrinsic correlations and interactions among sleep measurements that comprise global sleep
16 scores; the challenges in developing optimal composite indices; and the potential value in
17 extending sleep health scores to include additional dimensions that describe common sleep
18 disorders (especially in middle-aged and older populations), quantify sleep architecture, and
19 consider regularity in not only sleep timing but sleep duration.

20 In evaluating alternative ways to summarize sleep health in an aging multi-ethnic
21 community sample, we found evidence to support the utility of composite sleep health scores
22 based on the Ru SATED framework as well as scores extended to include measures from PSG
23 and actigraphy. In extending the Ru SATED framework with the additional consideration of

1 sleep architecture (particularly % NR), OSA (AHI), and duration regularity, as well as several
2 measures of sleep efficiency/continuity, we observed that the new composite yielded
3 interpretable principal components with more comprehensive scores evincing higher internal
4 reliability than simpler scores. PCA on original and extended Ru SATED frameworks showed
5 that the individual indices of sleep health for each score aggregated in consistent dimensional
6 patterns. The face validity of the extended SHS as a composite was supported by the empirical
7 finding that it summarized broad systematic shifts (most-least favorable categories; 0-3 vs 9-12)
8 across multiple sleep health metrics, including shifts in the initial components of Ru SATED as
9 well as in duration regularity, AHI, and sleep architecture. Moreover, each composite score
10 varied with race/ethnicity, supporting its utility for describing and monitoring sleep health
11 disparities. Adding additional sleep information also suggested important patterns of
12 association. For instance, timing and duration irregularity showed sleep disparities, which might
13 suggest that assessing determinants or barriers to (ir)regular bed and wake times may be
14 beneficial in reducing racial-ethnic sleep disparities rather than a focus on regular bed or wake
15 times alone (25). Although timing and duration regularity are moderately correlated, only about
16 31% of the variance of one accounts for the other, suggesting these regularity metrics to be
17 patterned yet distinct phenomena. For example, rotating shift-workers may have consistent
18 duration, yet inconsistent timing. Therefore, in middle-age or older cohorts where sleep disorders
19 and circadian disruption are prevalent, there is value in extending global sleep health assessments
20 with data from PSG and actigraphy. Moreover, metrics of OSA, altered sleep architecture, and
21 sleep duration variability predict adverse health outcomes, cognitive decline and mortality,
22 underscoring the potential utility of these data for informing sleep health assessments (26-28).
23

1 Our analyses not only supported use of composite indices, but also concurrent
2 consideration of multiple individual dimensions. For example, we found that the individual
3 “drivers” of differences in sleep health composites differed across race/ethnic groups, consistent
4 with prior work in both adults and the elderly (14). For example, compared to Whites, scores in
5 Blacks were influenced by sleepiness and %NR, while scores in Chinese were influenced by
6 AHI. Identifying individual drivers of sleep health may be particularly important for designing
7 targeted interventions. However, post-hoc analysis of the individual components of a summary
8 score requires cautious interpretation, and confidence intervals and p-values for individual
9 dimensions may be less informative than the pattern and trend among point estimates (2). Below
10 we further discuss some of the challenges and opportunities presented by multidimensional sleep
11 health, and analysis of both composite and individual dimensions.

12 *Statistical correlation among sleep variables and implications for a composite sleep*
13 *health approach*

14 We demonstrated that many sleep dimensions (both within and across domains) are inter-
15 correlated, reflecting their intrinsic physiological inter-relationships and potential responsiveness
16 to common stressors. Imagine, for instance, that each dichotomized SHS dimension in Table 1 is
17 represented by a light in a circuit. In the *absence* of dependence patterns, sleep health dimensions
18 would function like lights that burn out in parallel circuits, where health in one dimension does
19 not inform the others. However, if there are dependencies, much like a complex circuit with both
20 serial and parallel components, there is increased likelihood of some lights burning out (% NR,
21 sleepiness, sleep fragmentation) if certain other lights burn out (AHI). For instance, in our data,
22 higher sleep irregularity (timing and duration) and less sleep duration tended to co-occur. If this
23 pattern indicates the effects of an underlying common cause or latent factor, including measures

1 of both irregularity and duration should better characterize this driver within the distribution of a
2 composite.

3 In scales such as the SHS, therefore, the distribution of the composite score may reflect
4 relationships among the individual items that are informative. However, it is challenging to not
5 only select which measures to include in composites, but to determine the optimal way to
6 aggregate and weight each item, and whether to use continuous metrics or cutoff values
7 (discussed later). In choosing sleep variables for a composite, interpretability and ease of data
8 collection argue for a more parsimonious approach; on the other hand, there are potential
9 benefits of more comprehensive inclusion. First suppose that sleep health dimensions have
10 independent errors (e.g. from different measurement modalities, or measurements on
11 independent nights). Analogous to the case that adding test items improves measurement in
12 classical test theory, including additional dimensions within the same domain would reduce
13 measurement error, increasing signal and stability for an underlying latent target. On the other
14 hand, if multiple variables representing the same domain are calculated from the same
15 underlying data, the errors will not be independent; but if it is not known in advance which
16 function/non-linear transformation results in a measure more correlated with other variables in
17 the composite, including both forms may improve the chance of discovering these
18 relationships. Data-driven techniques may allow room for new dimensions/patterns to
19 emerge (29), perhaps better representing underlying drivers, etiological factors, or groups for
20 stratified interventions.

21 *Public health: Population sleep health assessment by defining optimal ranges and*
22 *characterizing individual dimensions and patterns across dimensions*

1 For public health utility, however, a categorical basis using prior knowledge to define
2 optimal ranges for each sleep dimension may also be appropriate. Early work operationalized Ru
3 SATED domains by dichotomizing daily diary, questionnaire, or actigraphy measures and
4 summing them into an overall sleep score (21-23). While continuous sleep exposures or
5 outcomes offer greater statistical power, categorical assessments – those involving a cut-point –
6 provide three practical benefits (30). First, cut-points aid clinicians in making decisions. Second,
7 prevalence can be defined and then used for needs assessment: e.g., approximately one third of
8 Americans do not meet sleep quantity recommendations (31). Third, cut-points enable
9 quantification of prevalence trends over time and across groups. Additionally, cut-points may
10 help avoid issues of non-linearities in exposure-response relationships, and cut-points are useful
11 for setting goals for public health initiatives. Categorical assessments may be appropriate even if
12 there is no evidence of a latent, *internal* discrete structure (taxon) for that dimension. As Kessler
13 (2002) notes, “there appears to be no taxon for high blood pressure,” and yet cut-points for blood
14 pressure based on *external* criteria such as risk of stroke guide clinical and public health
15 decisions (30).

16 Implications of the categorical approach for population sleep health are illustrated in
17 Table 1. At the dimension level, needs assessment becomes clearer. That WASO ≤ 20 minutes,
18 for instance, only has 6.8% favorability in MESA may suggest a calibration of the NSF’s optimal
19 ranges through extra data collection or may indicate an unmet need in middle-aged and older
20 adults (or both). Both cases are of public health import because age-appropriate optimal sleep
21 ranges can provide a set of public health targets to be validated against external criteria, whereas
22 unmet needs in a community suggest a need for investigation and intervention.

23

1 *Leveraging information from both composite scores and individual dimensions: patterns*
2 *of association illuminating potential sources of sleep health disparities*

3 Our results indicate how composites and individual dimensions can be mutually
4 informative. Compared to Whites, and consistent with prior evidence in MESA which included
5 both adults (aged 54-64 years) and the elderly (65-93 years) (14), adult Blacks in MESA show i)
6 lower sleep health composite scores, on average by more than a full component's difference for
7 the SHS; and ii) the largest dimension-level disparities for Blacks are in regularity (timing and
8 duration), duration, % N3, and continuity (Figure 2). Viewed from the sleep health paradigm,
9 Figure 3 suggests the likelihood of observing multiple disparities given a large disparity in, and
10 high correlation with, a single sleep dimension. We know why maintenance efficiency (SME)
11 and fragmentation (FI) show a similar pattern of disparity: each measures different facets of
12 sleep continuity. The bases for inter-relationships among sleep irregularity, duration, and
13 continuity and their racial-ethnic patterning are less clear. If timing irregularity is accompanied
14 by circadian misalignment (32), the two-process model of sleep suggests that circadian
15 misalignment and duration variability may interact, suggesting *internal* (to sleep health) causal
16 relationships. On the other hand, the observed disparities may reflect *external*, socioecological
17 causal factors such as light, shiftwork, etc. that influence sleep regularity, duration, and
18 continuity. Here, inter-correlation is cast as a source of research hypotheses about patterns of
19 multiple dimensions of sleep (2).

20 Consequently, we offer an interpretation of Buysse's definition that sleep health is a
21 *pattern* of sleep-wakefulness, illustrated by the dimensions of the SHS. Sleep health researchers
22 could estimate independent models for each dimension and assemble the interpretation of these
23 models in light of i) a global sleep health effect; and ii) correlated sleep dimensions. For

1 instance, individuals of Hispanic ethnicity showed average global sleep health (as indexed by the
2 SHS) that was significantly lower than that of Whites (Figure 2a). However, few of the
3 dimension-level disparities for Hispanics are distinct from the reference. The old paradigm that
4 focused on single dimensions may have under-estimated Hispanic sleep disparities. Rather,
5 under the sleep health paradigm, it is notable that the Hispanic sleep health disparity *pattern*
6 mostly resembles that of Blacks.

7 *Statistical correlation among sleep variables: patterns of association and implications*
8 *for intervention targeting and evaluation*

9 Changing a single sleep dimension may have “carry over” effects on other sleep
10 attributes. For example, an intervention which improves sleep regularity may improve sleep
11 continuity. Conversely, if sleep deteriorates in one dimension (e.g. duration), deteriorations in
12 other dimensions of sleep health may be expected (e.g. sleepiness, satisfaction, duration
13 variability, timing variability, etc.). For dimensions more difficult to directly target (% NR),
14 interventions could focus on correlated dimensions that are amenable to intervention (e.g.,
15 irregularity, AHI). Examining multiple dimensions also may be useful in evaluating the effects
16 of interventions such as zolpidem, which may positively influence features such as sleep latency
17 and WASO but negatively influence sleepiness/alertness and NR sleep. Interventions that affect
18 only one sleep dimension may be the exception rather than the rule.

19 Multidimensionality opens the possibility of ‘indirect’ or ‘root-cause’ targeting, such as
20 targeting OSA to improve multiple sleep health metrics such as alertness, quality, continuity, and
21 % NR on a population level. Indeed, the high prevalence of OSA, its underdiagnosis, and causal
22 effects on other sleep dimensions suggest OSA as a driver of poor sleep health (19, 33-35), and

1 thus an attractive intervention target to improve multiple sleep health dimensions (36-38).

2 Notably, multidimensionality does not preclude focus on multiple dimensions. (39, 40)

3 While sleep duration has been a key focus of epidemiological research, our analyses
4 support the importance in examining the contributions of multiple dimensions. Findings on the
5 detrimental effects of insufficient sleep duration may implicitly capture effects of poor sleep in
6 correlated dimensions – such as regularity, continuity, sleep architecture, OSA, placement in the
7 24-hour day, and sleepiness/alertness – and erroneously attribute this omnibus effect to duration.
8 Investigating dimensions other than duration may thus inform interventions on a population
9 scale. It may prove more accurate to say: ‘Sufficient sleep duration (*and all that it implies*) is
10 protective of later health.’ Alternatively, ‘Sleep health, of which average sleep duration is one
11 vital component, is protective of later health.’ Novel sleep health analyses using machine
12 learning support this conclusion (7, 41).

13 **Limitations**

14 The first limitation concerns the dependencies among dimensions, which may be one of
15 the more interesting and useful features of sleep health. The source(s) of these dependencies,
16 however, remain poorly understood. It is possible that the correlations observed are particular to
17 the sample or attributable to confounding factors such as the socioecological context, age or
18 other factors.

19 A second limitation concerns evolving or uncertain definitions and criteria of sleep
20 health. Consensus measures to assess sleep health with objective sleep data are lacking. We
21 leveraged comprehensive sleep data in MESA, and in doing so traded the parsimony of the
22 original Ru SATED framework for flexibility in adding dimensions empirically linked to race-
23 ethnicity (42, 43) as well as downstream health (26, 44) – at the cost of scalability. Future work

1 might estimate weights for each dimension in relation to a health outcome, with the limitation
2 that weights may be particular to the population and outcome.

3 The composites used self-reported and objective sleep indices. Even for the simple Ru
4 SATED score, which we derived by using objective data to define timing, duration and
5 efficiency, and from questionnaire data on quality and sleepiness/alertness, the PCA showed that
6 the objective and subjective measures each loaded on distinct components. These distinct loading
7 could reflect measurement bias rather than dimensional differences. The correspondence
8 between purely self-reported composite scales and those involving objective measures is not
9 characterized, and prior research indicates systematic differences in estimates between self-
10 report and objective measures according to short/long duration, sleep efficiency, health, and
11 socio-demographics (45).

12 Moreover, explicit guidance on population specific thresholds is limited. The core
13 definition of Ru SATED dimensions is evolving as well, with sleep regularity a recent addition.
14 A canon of sleep health parameters for evolving measures of sleep micro-architecture (e.g.
15 spindles, k-complexes) has yet to be established despite emerging data on the unique information
16 contained in these measures (27, 46).

17 These limitations highlight that sleep health is undergoing conceptual advances and scale
18 development (1, 6), scale validation (5, 6), implementations of sleep health in cohort studies and
19 community samples (21, 47), innovations in methodological approaches (4, 7, 41, 48), optimal
20 range and threshold determination (18, 22), and many practical choices on how best to analyze
21 multidimensional sleep data. We suggest that statistical dependencies are an additional factor to
22 contend with. As sleep health continues to develop as a field, we suggest two solutions that may
23 help sleep health's future development.

1

2 *Solution 1: Studying challenges and solutions in disciplines which have already shifted to*
3 *multidimensional paradigms*

4 From multidimensional/symptom psychiatric epidemiology (30), it is instructive to
5 observe the outcome of the tension between epidemiological theories of ‘distribution shifts vs
6 targeted intervention’ and policy making (30). Although shifting distributions is a powerful
7 public health approach, the current policy paradigm favors targeted intervention, prevalence
8 estimates, reporting prevalence trends over time, etc. (31). Thus, identifying an initial threshold
9 target, with monitoring of impact of interventions on this target and related health outcomes,
10 provides an example of how sleep health research can be translated to policy. This approach is
11 similar to the long-standing practices of using blood pressure cut-offs as public health and
12 individual targets despite evolving data on specific thresholds that confer disease (30). Future
13 research to develop normative sleep ranges will further support the utility of this approach for
14 policy makers and the public health. Logically implied are similar needs for developing and
15 validating normative or optimal ranges for composite sleep metrics.

16 Advantages to using composite scores include the reduction of multiple testing burden (as
17 compared to analysis of multiple individual dimensions) and the ability to detect the effects of
18 multiple small effects. An analogy can be drawn between sleep health and nutrition, in which
19 most nutrients appear to contribute in small ways to a larger composite effect (and sometimes
20 this composite effect better describes nutritional profiles than when individual nutrients are
21 examined) (2). Given their simplicity, composites also have ready public health application,
22 similar to promoting diets such as the Mediterranean diet rather than recommending specific
23 nutrient targets. (2). However, while summary scales may be selected *a priori* based on extant

1 empirical or conceptual work, prospective data are needed to demonstrate that they predict
2 clinically meaningful outcomes.

3 Methodological advances in sleep health composites are anticipated with advances in
4 sleep knowledge (and vice versa), perhaps leveraging increasingly larger datasets (49).
5 Nonetheless, there is a need to recognize that the components of global sleep health scores will
6 likely vary according to the questions at hand (e.g., interest in a pediatric versus aging
7 population; interest in cardiovascular disease or cancer). Future work might draw from prior
8 literature on composites in other fields which also deal with complex systems with high-
9 dimensional data. This may be of high utility in choosing how to weight and aggregate the
10 complex set of phenomena comprising sleep health in a context-specific way (50).

11 Another tension appears to depend on sleep health's complexity and wide-ranging
12 importance. Because sleep health is multidimensional, whose dimensions may be variably
13 more/less responsive to a wide set of individual and social factors, and whose different
14 components may be more/less consequential for different outcomes, a single, canonical
15 operationalization of sleep health may be suboptimal for any particular project. It may be useful
16 to construct at least two sleep health indices: i) canonical Ru SATED (one dimension per
17 domain); and ii) modified Ru SATED for a specific project (if necessary). This may help
18 preserve consistency in the field while also affording flexibility in answering varied research
19 questions. For assessing racial-ethnic disparities, the Sleep Health Score showed adequacy in
20 distinguishing between groups; analysis of individual effects pointed to the strongest and
21 modifiable drivers of the composite association, for public health targeting (regularity, duration).

22

23 *Solution 2: Objective and longitudinal sleep data.*

1 A recent NIH workshop report on sleep health disparities emphasized that
2 multidimensional sleep health must be experienced on a *regular* basis (51). However, there are
3 little extant cohort data to assess whether sleep health is experienced on a regular basis, and what
4 factors produce regular sleep health. Moreover, there are few data on within-individual
5 correlations of sleep dimensions across time, and thus little understanding of the ‘natural
6 histories’ of good or poor multidimensional sleep health, and a lack of longitudinal data limits
7 causal modeling through natural experiments. Two solutions to these problems are 1) analysis of
8 multi-day sleep studies as a time series (52); 2) longitudinal, comprehensive cohort sleep data
9 collection.

10 First, Hale, Troxel, and Buysse (2020) argued that daily sleep assessments, modeled as a
11 time series, may unpack bi-directional relationships between sleep and regular exposures such as
12 diet and stress (52). Such work can help to capture the effect of short-term exposures. Moreover,
13 a single night’s poor sleep health in a particular dimension may cascade to affect other
14 dimensions the same or next night or may suggest common exposures that affect multiple
15 dimensions (e.g. rotating shift work) (9). Multi-day comprehensive sleep assessment with daily
16 physical and psychosocial exposure measurements should elucidate how sensitive sleep health
17 dimensions may be to acute, fluctuating exposures. Thus, future cohort studies might consider
18 daily psychosocial instruments or an expanded sleep diary(now possible through electronic data
19 collection) to capture daily exposures of interest across the sleep study period and consider how
20 such data are summarized as composites or individual dimensions.

21 Consumer wearables are a promising alternative to traditional actigraphs. Consumer
22 wearables are relatively inexpensive and are amenable to data collection over long periods. Sleep
23 data can be paired with physical activity measurement and geo-coding to get a fuller picture of

1 sleep-wake patterns.(53) As these data are collected on larger numbers of individuals, it will be
2 important to consider whether they provide novel metrics of sleep timing, variability and others
3 features not apparent in shorter-term measurements.

4 Longitudinal sleep data are critically needed to better understand the ‘natural history’ of
5 multidimensional sleep health, and may help identify which dimensions are the most variable
6 over long time scales and how favorable patterns of sleep health may become distorted by
7 changes in behavior or environment over the life course (e.g. major life events) (54, 55).
8 Accordingly, longitudinal data may help better characterize critical features and etiology of
9 consequential sleep health phenotypes. For instance, Vgontzas et al. (2012) reported that
10 objective short sleep duration (<6 hours) at baseline was associated with higher odds of
11 persistent insomnia 7.5 years later compared to normal sleepers and those with fully remitted
12 insomnia (56). Longitudinally collected sleep phenotypes can clarify time-dependent features
13 whose proximal cause and age of onset and time-dependent modifying factors are otherwise
14 difficult to characterize.

15 Longitudinal data may be particularly useful to clarify specific sleep drivers and levers.
16 We suggest sleep regularity (in timing and duration) as candidates for further longitudinal
17 research due to their i) potentially modifiable nature (as reflected in sleep hygiene
18 recommendations for consistent bed and wake times), ii) consistent correlation with many other
19 sleep dimensions (Figure 3), and iii) ability to forecast incident metabolic dysfunction and
20 cardiovascular events (12, 13). If regularity as a candidate driver of sleep health is supported by
21 longitudinal studies, then targeted interventions can be tested, including the impact on sleep
22 health composites and clinical endpoints.

1 In summary, comprehensive and repeated sleep assessment may inform the etiology and
2 ‘natural history’ of multidimensional sleep health, identify candidate drivers and phenotypes, and
3 help establish normative sleep health ranges for scientific knowledge and as targets for public
4 health. To create robust sleep health scores, additional research showing their value in predicting
5 health outcomes is needed.

6 *Incorporate qualitative data*

7 Finally, qualitative research complements other research methods, providing information,
8 from the participant’s point of view, about: i) barriers and facilitators to achieving sleep health
9 and ii) impact of sleep on overall health and well-being. For instance, in a small Boston sample
10 of racially-ethnically diverse low-income adults, commonly stated barriers to “good sleep” were
11 work (or multiple) work schedules, consuming large quantities of soda in the evening, reluctance
12 to stop using personal electronics at bed-time, child-care, and financial worries (25). Further
13 investigation of factors such as these can be thoughtfully integrated into future data collection
14 efforts that inform multidimensional sleep. Qualitative work is vital to retrospectively understand
15 changes in sleep health and is thus of high value to understanding etiology. For instance, one of
16 the richest pieces of evidence in support of the ‘precipitation’ component of the 3-P model of
17 insomnia (predisposition, precipitation, perpetuation) comes from Healey et al. (1981) who used
18 a set of retrospective interviews assessing stressful life events as precipitants of insomnia
19 complaints (55).

20 *Summary*

21 Multidimensional sleep health represents a paradigm shift in sleep science. Two useful
22 frameshifts are that sleep health is both conceptually and operationally a composite concept and
23 that sleep dimensions do not exist in isolation. These frameshifts logically implicate sleep

- 1 disordered breathing and sleep regularity as targets for further research and intervention. To fully
- 2 explore the implications of this new paradigm, comprehensive longitudinal assessment of sleep
- 3 is necessary.

1 **Acknowledgements**

2 This research was supported by National Institute of Health grants 5T32HL007901-14,
3 R35HL135818, R01HL098433; and by contracts HHSN268201500003I, N01-HC-95159, N01-
4 HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165,
5 N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart,
6 Lung, and Blood Institute.

7 **Declaration of conflicts of interest**

8 JC, MW, TH, SB, and SR report funding from the NIH. SB reports consultant fees from Eisai,
9 Inc, and Merck Sharpe & Dohme. The authors declare no other conflict.

10 **Human Subjects Protection**

11 This research used secondary, de-identified, publicly available data (National Sleep Research
12 Resource). Participants provided informed consent for the original study, which complied with
13 the Declaration of Helsinki.

Tables and Figures

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)	Regularity	Actigraphy	<30 minutes (13)	157 (21.4%)	MPsd
Duration variability (sd, min)	Regularity	Actigraphy	<60 minutes (13)	271 (36.9%)	TSTsd
Quality (WHIIRS Likert subscale; higher is increasing complaints)	Satisfaction	Questionnaire	“Sound and restful” or “very sound or restful”	157 (21.4%)	Quality
Sleepiness	Alertness	Questionnaire	≤10 (17)	199 (27.1%)	ESS
Timing (midpoint) ^b	Timing	Actigraphy	02:00-04:00 ^b	466 (63.4%)	Timing
Sleep Maintenance Efficiency (%)	Efficiency	Actigraphy	>90%	524 (71.3%)	SME
Fragmentation Index ^a	Efficiency	Actigraphy	≤15 ^a	199 (27.1%)	FI
Wake after Sleep Onset (min)	Efficiency	Actigraphy	≤20 minutes (18)	50 (6.8%)	WASO
Total Sleep Time ^b (min)	Duration	Actigraphy	6-8 hours ^b	461 (62.7%)	TST
%NR ^b	Architecture	PSG	21%-30% ^b (18)	251 (34.1%)	% NR
%N3 ^b	Architecture	PSG	16%-20% ^b (50)	84 (11.4%)	% N3
Apnea-Hypopnea Index (events/hr)	SDB	PSG	≤15 events/hr (19)	354 (48.2%)	AHI
Sleep onset latency (difficulties initiating sleep – WHIIRS subscale)	Wake-sleep transition	Questionnaire	“Less than once a week” or “no, not in the past 4 weeks”	476 (64.8%)	SOLsr
Sleep onset latency (min)	Wake-sleep transition	Actigraphy	-		SOL

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

b: These dimensions have optimal ranges in which overexpression of this dimension is considered problematic.

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 categorized: least to most favorable SHS					
	Overall	least (0-3)	less (4-6)	more (7-9)	most (9-12)	p
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 + 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 ^a (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 ^a (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 ^a (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
Efficiency						
Sleep Maintenance Efficiency ^a (%) [↑]	91.4 (3.4)	89.2 (4.3)	91.1 (3.1)	93.0 (2.2)	93.2 (1.8)	<0.001
Fragmentation Index ^{a↓}	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 ^a (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						
% NR ^{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) [↓]	22.0 (20.2)	34.6 (23.7)	23.7 (20.0)	13.6 (13.7)	8.3 (6.7)	<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 ^a (actigraphy)	4.2 (3.0)	4.6 (3.8)	4.1 (2.9)	4.1 (2.7)	3.6 (2.1)	0.326

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

a: actigraphy

b: polysomnography

c: self-report

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

p-values for skewed variables by Kruskal-Wallis; otherwise ANOVA.

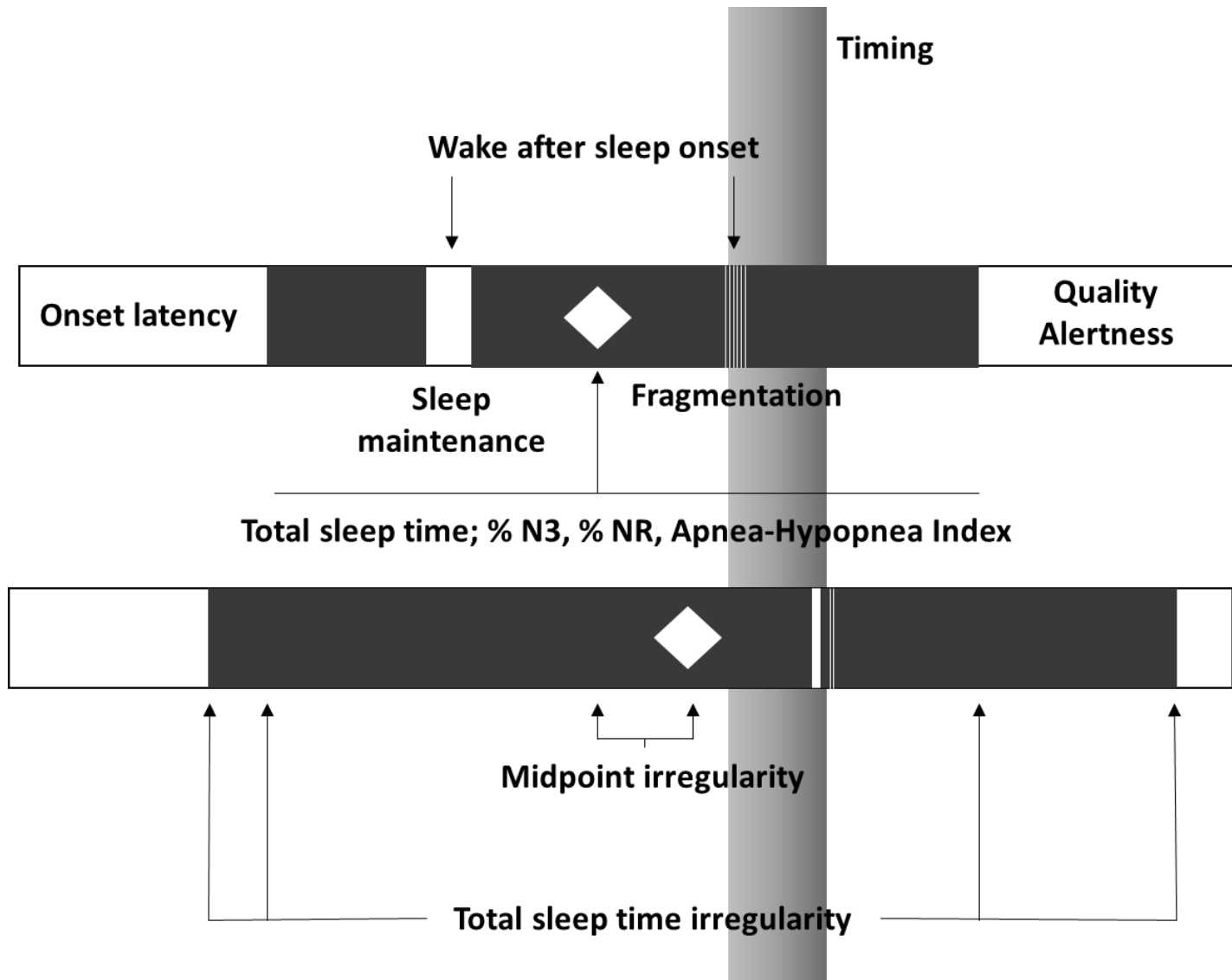
Table 3. Principal Components Analysis of Sleep Health dimensions: Ru SATED, Ru SATED + OSA and Architecture, SHS, SHS-parsimonious

	Ru SATED			Ru SATED + OSA and Sleep Architecture			SHS-PC1			SHS-PC1 (parsimonious)		
	$\alpha= 0.42$			$\alpha= 0.47$			$\alpha= 0.61$			$\alpha= 0.45$		
Proportion 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%
	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
Ru SATED												
Irregularity [↓] (timing)	-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)	-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)	-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)	-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)	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 [↗]	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 [↓]				-0.38	0.43	-0.23	-0.32	0.18	-0.29	-0.38	0.38	-0.27
% N3 [↗]				0.14	-0.34	-0.18	0.14	-0.12	-0.08	0.15	-0.31	-0.14
% NR [↗]				0.35	-0.25	0.48	0.27	-0.05	0.53	0.35	-0.17	0.51
SHS additions												
Irregularity [↓] (duration)							-0.27	-0.42	0.15			
Onset latency [↓]							-0.11	0.26	0.33	-0.06	0.34	0.12
Fragmentation [↓]							-0.44	0.36	0.11			

WASO [↓]							-0.32	-0.01	-0.28			
-------------------	--	--	--	--	--	--	-------	-------	-------	--	--	--

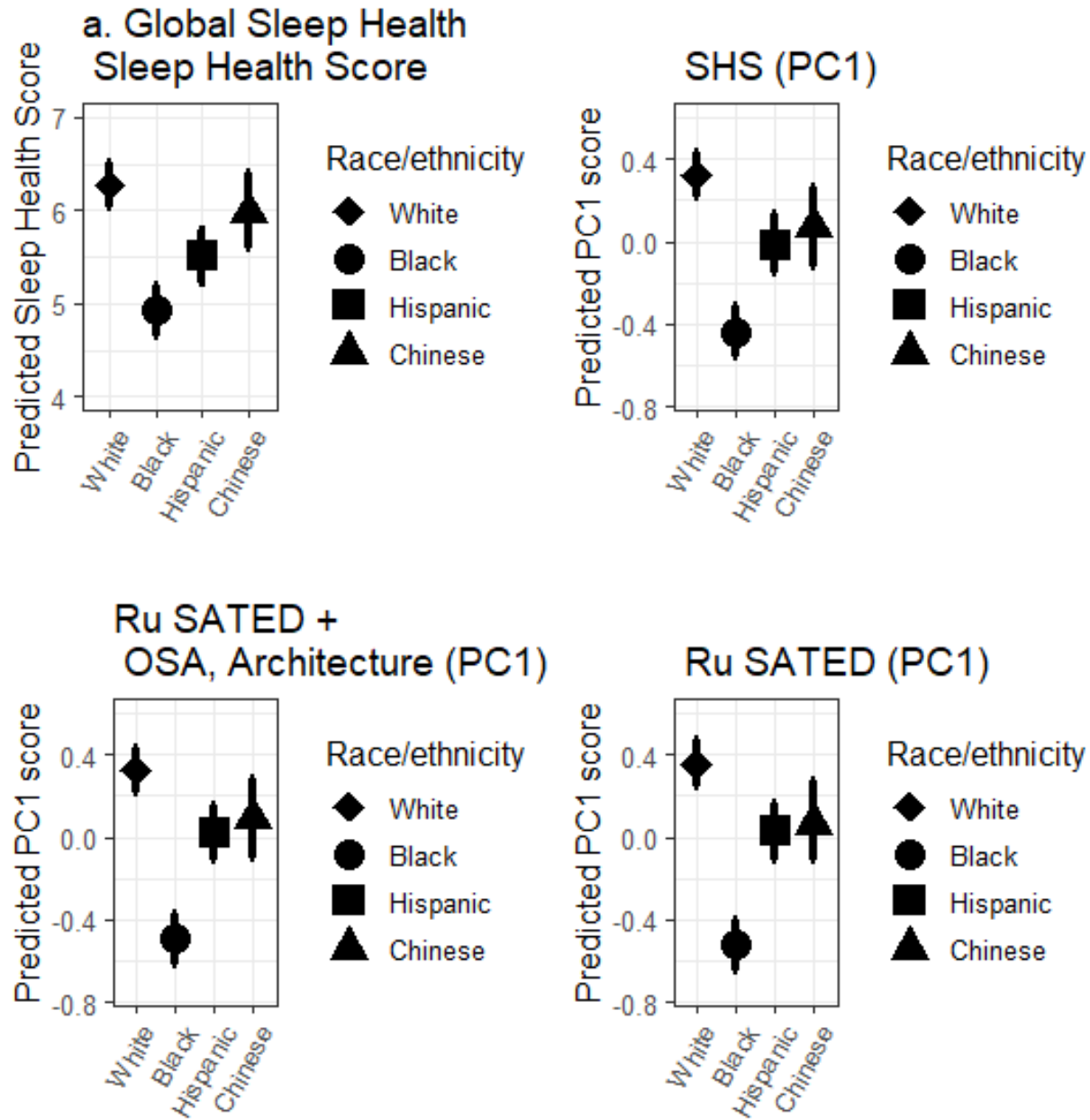
↑ = Higher is better; ↓ = lower is better; ↗ = generally higher is better, but overexpression can be problematic (eg long sleepers, excessive % NR)

Figure 1. Conceptual Sleep Health dimensions across time.



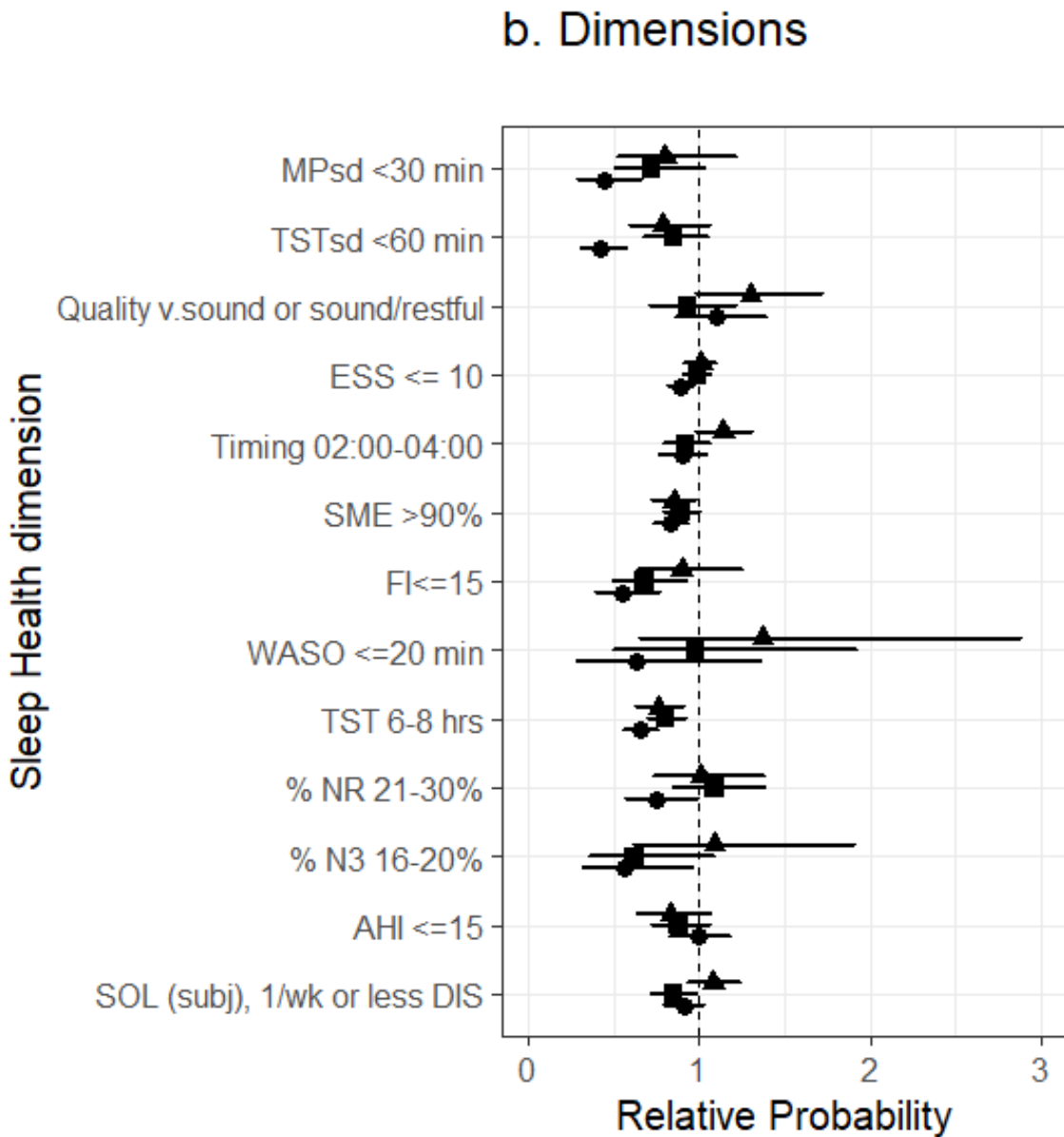
Note: Diagram of two consecutive days/nights of sleep. Dark bars indicate sleep; white bars indicate wake. This diagram is conceptual and does not represent a particular individual in MESA-Sleep. Diamonds indicate midpoint. Onset latency, quality, and alertness are evaluated during wake (self-report).

- 1 Figure 2. Age and sex-adjusted estimates of a) global sleep health and b) its components
- 2 regressed on race-ethnicity. The Multi-Ethnic Study of Atherosclerosis (n=735).
- 3 Figure 2a



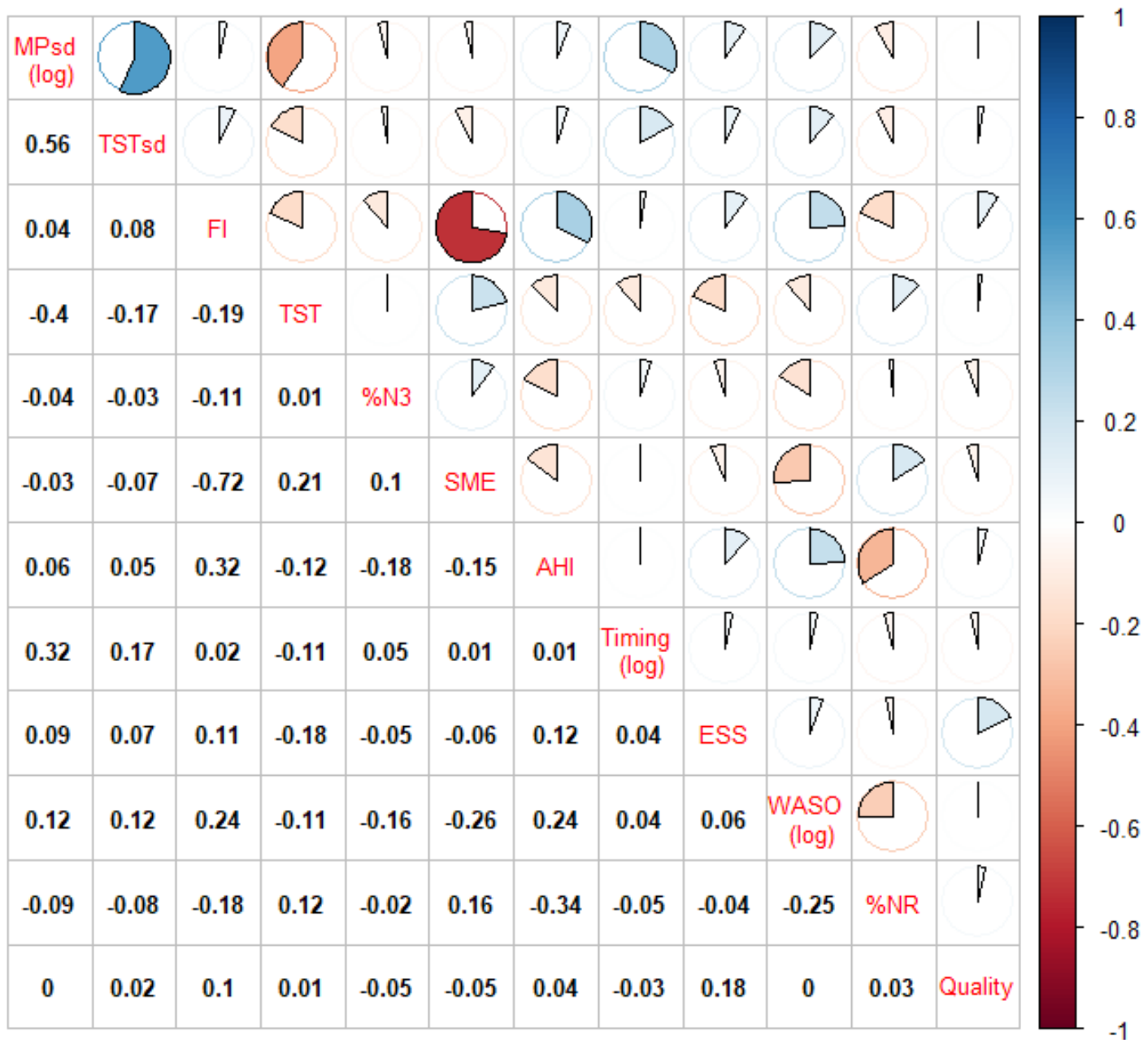
- 4
- 5
- 6
- 7
- 8

1 Figure 2b:



2

1 Figure 3. Pearson correlation matrix of objective sleep variables and Epworth scores. The Multi-
 2 Ethnic Study of Atherosclerosis (n=735)^a.



3

4 a. onset latency, self-report was omitted due to unimprovable skew. Quality is a Likert-
 5 scale, approximately normally distributed; lower is better.

6 b. Pie size and shading indicate degree of correlation.

7

8

9

10

11

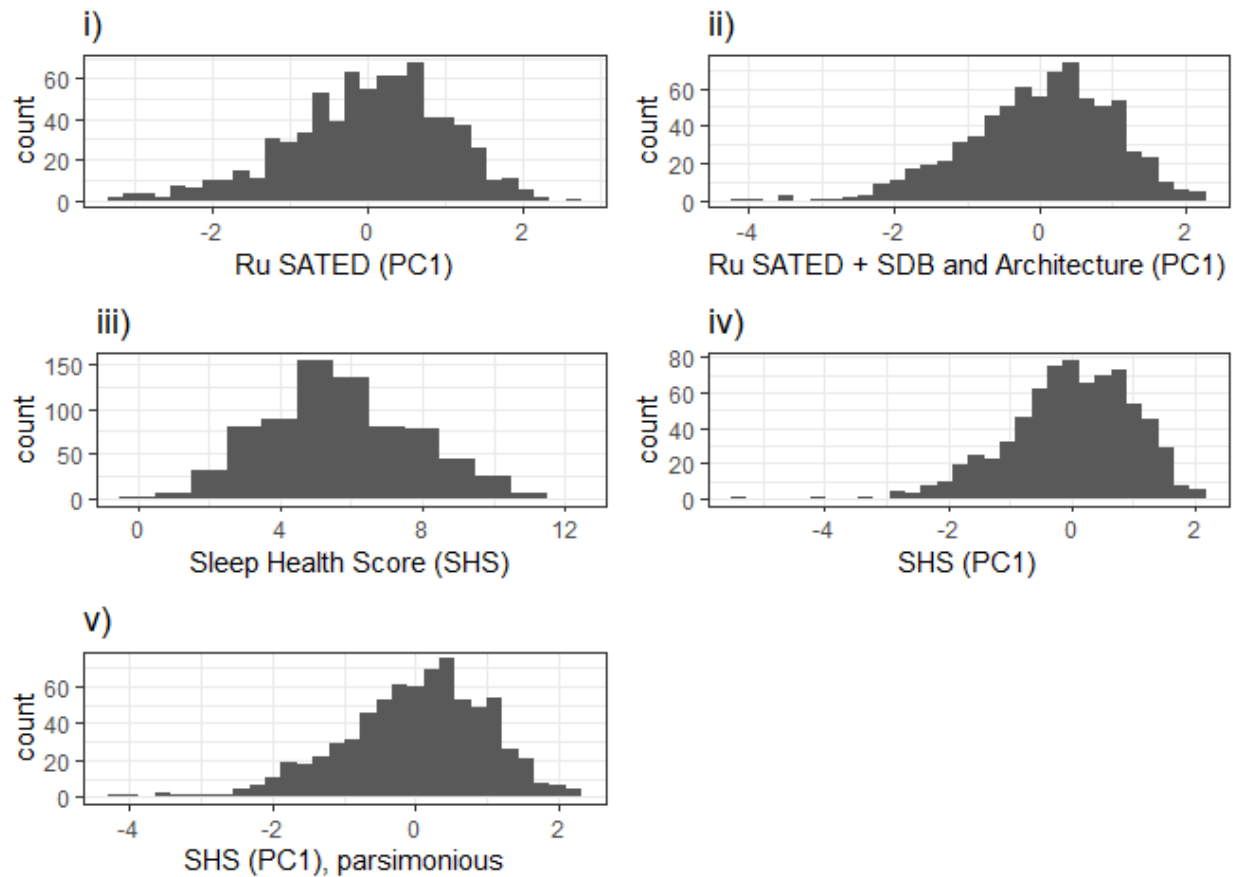
12

Appendix A. Comparison of Ru SATED and the Sleep Health Score (SHS)

Table A1. Compare/contrast Ru SATED and Sleep Health Scores

Sleep dimension	Ru SATED	SHS
Regularity	x	x
Satisfaction	x	x
Alertness	x	x
Timing	x	x
Efficiency	x	x
Duration	x	x
Sleep architecture		x
Duration regularity		x
Onset latency		x
Dimensions	1 measure per domain (parsimony)	Multiple related measures allowed (see discussion of composite scores)

Figure A1. Distributions of sleep health composites: i) Ru SATED (PC1), ii) Ru SATED + SDB and Architecture (PC1), iii) Sleep Health Score (SHS), iv) Sleep Health Score (PC1), v) Sleep Health Score (PC1) parsimonious (one measure per domain)



Appendix B.

Table B1. Age and sex-adjusted regression results for Ru SATED (PC1), Ru SATED + SDB/Architecture (PC1), SHS (PC1), and the SHS.

	Ru SATED (PC1)		Ru SATED + OSA/Architecture (PC1)		SHS (PC1)		SHS	
	β [95% CI]	p	β [95% CI]	p	β [95% CI]	p	β [95% CI]	p
Chinese	-0.28 [-0.50, -0.06]	0.013	-0.24 [-0.46, -0.02]	0.029	-0.25 [-0.46, -0.03]	0.027	-0.29 [-0.77, 0.19]	0.235
Black	-0.87 [-1.05, -0.70]	<0.001	-0.82 [-0.99, -0.65]	<0.001	-0.76 [-0.92, -0.59]	<0.001	-1.34 [-1.72, -0.97]	<0.001
Hispanic	-0.32 [-0.50, -0.14]	<0.001	-0.31 [-0.48, -0.13]	0.001	-0.33 [-0.51, -0.16]	<0.001	-0.77 [-1.16, -0.38]	<0.001
Female	0.29 [0.15, 0.43]	<0.001	0.54 [0.41, 0.67]	<0.001	0.53 [0.40, 0.67]	<0.001	0.55 [0.25, 0.84]	<0.001
Age (years)	0.02 [-0.00, 0.04]	0.077	0.00 [-0.02, 0.02]	0.945	-0.02 [-0.04, 0.01]	0.161	-0.01 [-0.06, 0.04]	0.796
(Intercept)	-1.04 [-2.41, 0.33]	0.138	-0.02 [-1.36, 1.32]	0.976	0.98 [-0.37, 2.34]	0.155	6.36 [3.36, 9.36]	<0.001

Table B2. Variable-level regression results for the components of the SHS. N=735.

Variable:	Black RR [95% CI]	Hispanic RR [95% CI]	Chinese RR [95% CI]	Age RR [95% CI]	Female RR [95% CI]	(Intercept) RR [95% CI]
MPsd	0.44 [0.31, 0.62]	0.72 [0.54, 0.97]	0.80 [0.56, 1.14]	1.02 [0.98, 1.06]	0.97 [0.77, 1.22]	0.12 [0.01, 1.19]
TSTsd	0.42 [0.33, 0.55]	0.85 [0.71, 1.02]	0.79 [0.62, 1.01]	1.00 [0.97, 1.02]	0.92 [0.79, 1.07]	0.61 [0.13, 2.96]
Fragmentation	0.55 [0.42, 0.73]	0.68 [0.52, 0.88]	0.90 [0.68, 1.19]	0.97 [0.93, 1.00]	1.57 [1.27, 1.95]	2.09 [0.29, 15.32]
TST	0.65 [0.57, 0.74]	0.80 [0.72, 0.90]	0.76 [0.65, 0.88]	0.99 [0.98, 1.01]	1.12 [1.02, 1.23]	0.99 [0.39, 2.52]
% N3	0.56 [0.35, 0.88]	0.62 [0.39, 1.00]	1.09 [0.68, 1.75]	0.97 [0.92, 1.03]	1.80 [1.25, 2.59]	0.45 [0.02, 12.82]
SME	0.83 [0.76, 0.92]	0.89 [0.81, 0.98]	0.85 [0.74, 0.96]	1.00 [0.99, 1.02]	1.17 [1.08, 1.27]	0.60 [0.28, 1.28]
AHI	0.99 [0.86, 1.15]	0.88 [0.75, 1.03]	0.83 [0.67, 1.03]	0.97 [0.95, 0.99]	1.71 [1.48, 1.97]	2.44 [0.74, 7.99]
SOL	0.91 [0.81, 1.01]	0.85 [0.75, 0.96]	1.08 [0.95, 1.21]	1.02 [1.00, 1.03]	0.83 [0.76, 0.91]	0.29 [0.12, 0.73]
% NR	0.75 [0.60, 0.95]	1.09 [0.89, 1.33]	1.01 [0.77, 1.31]	1.00 [0.97, 1.03]	1.29 [1.08, 1.54]	0.34 [0.07, 1.82]
ESS	0.89 [0.83, 0.96]	0.99 [0.93, 1.05]	1.01 [0.93, 1.09]	1.01 [1.00, 1.02]	1.01 [0.96, 1.06]	0.45 [0.26, 0.78]
Timing	0.90 [0.79, 1.01]	0.92 [0.81, 1.04]	1.14 [1.00, 1.29]	0.99 [0.97, 1.00]	1.02 [0.93, 1.12]	1.31 [0.51, 3.35]
WASO	0.63 [0.33, 1.20]	0.98 [0.56, 1.72]	1.37 [0.74, 2.56]	0.90 [0.83, 0.97]	1.71 [1.05, 2.77]	31.38 [0.34, 2909.80]
Quality	1.10 [0.90, 1.35]	0.93 [0.74, 1.17]	1.30 [1.03, 1.64]	1.05 [1.02, 1.08]	0.91 [0.77, 1.06]	0.02 [0.00, 0.12]

References

1. Buysse DJ. Sleep health: can we define it? Does it matter? *Sleep*. 2014;37(1):9-17.
2. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Current opinion in lipidology*. 2002;13(1):3-9.
3. Larson JS. The World Health Organization's definition of health: Social versus spiritual health. *Social indicators research*. 1996;38(2):181-92.
4. 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.
5. Ravyts 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.
6. 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.
7. 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.
8. Dijk D-J, Czeisler CA. Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. *Journal of Neuroscience*. 1995;15(5):3526-38.
9. Ohayon MM, Smolensky MH, Roth T. Consequences of shiftworking on sleep duration, sleepiness, and sleep attacks. *Chronobiology international*. 2010;27(3):575-89.
10. Åkerstedt T, Anund A, Axelsson J, Kecklund G. Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of sleep research*. 2014;23(3):242-54.
11. Roehrs T, Zorick F, Sicklesteel J, Wittig R, Roth T. Excessive daytime sleepiness associated with insufficient sleep. *Sleep*. 1983;6(4):319-25.
12. 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.
13. 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.
14. 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*. Under review (2021).
15. 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.
16. 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.
17. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-5.
18. 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.
19. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *American journal of epidemiology*. 2013;177(9):1006-14.
20. Song M-K, Lin F-C, Ward SE, Fine JP. Composite variables: when and how. *Nursing research*. 2013;62(1):45.
21. 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.

22. 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.
23. 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.
24. Zou G. A modified poisson regression approach to prospective studies with binary data. *American journal of epidemiology*. 2004;159(7):702-6.
25. Rottapel RE, Zhou ES, Spadola CE, Clark CR, Kontos EZ, Laver K, et al. Adapting sleep hygiene for community interventions: a qualitative investigation of sleep hygiene behaviors among racially/ethnically diverse, low-income adults. *Sleep Health*. 2020.
26. Leary EB, Watson KT, Ancoli-Israel S, Redline S, Yaffe K, Ravelo LA, et al. Association of Rapid Eye Movement Sleep With Mortality in Middle-aged and Older Adults. *JAMA Neurology*. 2020.
27. 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.
28. Azarbarzin A, Sands SA, Stone KL, Taranto-Montemurro L, Messineo L, Terrill PI, et al. The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic Fractures in Men Study and the Sleep Heart Health Study. *European heart journal*. 2019;40(14):1149-57.
29. 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.
30. Kessler RC. The categorical versus dimensional assessment controversy in the sociology of mental illness. *Journal of health and social behavior*. 2002:171-88.
31. 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.
32. Phillips AJ, Clerx WM, O'Brien CS, Sano A, Barger LK, Picard RW, et al. Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. *Scientific Reports*. 2017;7(1):1-13.
33. Kapur V, Strohl KP, Redline S, Iber C, O'Connor G, Nieto J. Underdiagnosis of sleep apnea syndrome in US communities. *Sleep and Breathing*. 2002;6(2):49-54.
34. Costa LE, Uchôa CHG, Harmon RR, Bortolotto LA, Lorenzi-Filho G, Drager LF. Potential underdiagnosis of obstructive sleep apnoea in the cardiology outpatient setting. *Heart*. 2015;101(16):1288-92.
35. Johnson DA, Guo N, Rueschman M, Wang R, Wilson JG, Redline S. Prevalence and correlates of obstructive sleep apnea among African Americans: the Jackson Heart Sleep Study. *Sleep*. 2018;41(10).
36. McARDLE N, DOUGLAS NJ. Effect of continuous positive airway pressure on sleep architecture in the sleep apnea-hypopnea syndrome: a randomized controlled trial. *American journal of respiratory and critical care medicine*. 2001;164(8):1459-63.
37. Quan SF, Budhiraja R, Kushida CA. Associations Between Sleep Quality, Sleep Architecture and Sleep Disordered Breathing and Memory After Continuous Positive Airway Pressure in Patients with Obstructive Sleep Apnea in the Apnea Positive Pressure Long-term Efficacy Study (APPLES). *Sleep Sci*. 2018;11(4):231-8.
38. Loredó JS, Ancoli-Israel S, Kim E-J, Lim WJ, Dimsdale JE. Effect of Continuous Positive Airway Pressure Versus Supplemental Oxygen on Sleep Quality in Obstructive Sleep Apnea: A Placebo-CPAP-Controlled Study. *Sleep*. 2006;29(4):564-71.

39. Harvey AG, Hein K, Dong L, Smith FL, Lisman M, Yu S, et al. A transdiagnostic sleep and circadian treatment to improve severe mental illness outcomes in a community setting: study protocol for a randomized controlled trial. *Trials*. 2016;17(1):606.
40. Spadola CE, Rottapel RE, Zhou ES, Chen JT, Guo N, Khalsa SBS, et al. A sleep hygiene and yoga intervention conducted in affordable housing communities: Pilot study results and lessons for a future trial. *Complementary Therapies in Clinical Practice*. 2020;39:101121.
41. 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.
42. Redline S, Kirchner HL, Quan SF, Gottlieb DJ, Kapur V, Newman A. The effects of age, sex, ethnicity, and sleep-disordered breathing on sleep architecture. *Archives of internal medicine*. 2004;164(4):406-18.
43. Mokhlesi B, Pannain S, Ghods F, Knutson KL. Predictors of slow - wave sleep in a clinic - based sample. *Journal of sleep research*. 2012;21(2):170-5.
44. Dew MA, Hoch CC, Buysse DJ, Monk TH, Begley AE, Houck PR, et al. Healthy older adults' sleep predicts all-cause mortality at 4 to 19 years of follow-up. *Psychosomatic medicine*. 2003;65(1):63-73.
45. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? *Epidemiology*. 2008:838-45.
46. 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.
47. 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.
48. 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.
49. Dean DA, Goldberger AL, Mueller R, Kim M, Rueschman M, Mobley D, et al. Scaling up scientific discovery in sleep medicine: the National Sleep Research Resource. *Sleep*. 2016;39(5):1151-64.
50. 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.
51. Jackson CL, Walker JR, Brown MK, Das R, Jones NL. A workshop report on the causes and consequences of sleep health disparities. *Sleep*. 2020.
52. Hale L, Troxel W, Buysse DJ. Sleep Health: An Opportunity for Public Health to Address Health Equity. *Annual Review of Public Health*. 2020;41.
53. Fore R, Hart JE, Choirat C, Thompson JW, Lynch K, Laden F, et al. Embedding mobile health technology into the Nurses' Health Study 3 to study behavioral risk factors for cancer. *AACR*; 2020.
54. Holmes TH, Rahe RH. The social readjustment rating scale. *Journal of psychosomatic research*. 1967.
55. Healey ES, Kales A, Monroe LJ, Bixler EO, Chamberlin K, Soldatos CR. Onset of insomnia: role of life-stress events. *Psychosomatic medicine*. 1981.
56. Vgontzas AN, Fernandez-Mendoza J, Bixler EO, Singareddy R, Shaffer ML, Calhoun SL, et al. Persistent insomnia: the role of objective short sleep duration and mental health. *Sleep*. 2012;35(1):61-8.