Research Practice

Multidimensional Sleep and Mortality in Older Adults: A Machine-Learning Comparison With Other Risk Factors

Meredith L. Wallace, PhD,[1,2,]*, Daniel J. Buysse, MD,[1,] Susan Redline, MD, PhD,[3] Katie L. Stone, PhD,[4,5] Kristine Ensrud, MD, MPH,[6,7] Yue Leng, PhD,[8] Sonia Ancoli-Israel, PhD,[9] and Martica H. Hall, PhD[1]

[1]Department of Psychiatry, University of Pittsburgh, Pennsylvania. [2]Department of Biostatistics, University of Pittsburgh, Pennsylvania. [3]Departments of Medicine, Brigham and Women's Hospital and Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts. [4]California Pacific Medical Center, Research Institute, San Francisco. [5]Department of Epidemiology and Biostatistics, University of California, San Francisco. [6]Department of Medicine and Division of Epidemiology and Community Health, University of Minnesota, Minneapolis. [7]Center for Chronic Disease Outcomes Research, Veterans Affairs Health Care System, Minneapolis, Minnesota. [8]Department of Psychiatry, University of California, San Francisco. [9]Department of Psychiatry, University of California, San Diego.

*Address correspondence to: Meredith L. Wallace, PhD, Department of Psychiatry, University of Pittsburgh, 3811 O'Hara Street, Pittsburgh PA 15213. E-mail: lotzmj@upmc.edu

Decision Editor: David Melzer, MBBCh

Abstract

Background: Sleep characteristics related to duration, timing, continuity, and sleepiness are associated with mortality in older adults, but rarely considered in health recommendations. We applied machine learning to: (i) establish the predictive ability of a multidimensional self-reported sleep domain for all-cause and cardiovascular mortality in older adults relative to other established risk factors and (ii) to identify which sleep characteristics are most predictive.

Methods: The analytic sample includes $N = 8,668$ older adults (54% female) aged 65–99 years with self-reported sleep characterization and longitudinal follow-up (≤15.5 years), aggregated from three epidemiological cohorts. We used variable importance (VIMP) metrics from a random survival forest to rank the predictive abilities of 47 measures and domains to which they belong. VIMPs > 0 indicate predictive variables/domains.

Results: Multidimensional sleep was a significant predictor of all-cause (VIMP [99.9% confidence interval {CI}] = 0.94 [0.60, 1.29]) and cardiovascular (1.98 [1.31, 2.64]) mortality. For all-cause mortality, it ranked below that of the sociodemographic (3.94 [3.02, 4.87]), physical health (3.79 [3.01, 4.57]), and medication (1.33 [0.94, 1.73]) domains but above that of the health behaviors domain (0.22 [0.06, 0.38]). The domains were ranked similarly for cardiovascular mortality. The most predictive individual sleep characteristics across outcomes were time in bed, hours spent napping, and wake-up time.

Conclusion: Multidimensional sleep is an important predictor of mortality that should be considered among other more routinely used predictors. Future research should develop tools for measuring multidimensional sleep—especially those incorporating time in bed, napping, and timing—and test mechanistic pathways through which these characteristics relate to mortality.

Keywords: Sleep health, Mortality, Elderly, Machine learning, Random forest

Sleep characteristics are associated with mortality in older adults. For example, long and short sleep duration (1,2), increased daytime sleepiness (3,4), misaligned sleep timing (5), and poor continuity (3,6) are each associated with increased mortality risk in this population. Nevertheless, sleep is not typically considered in high-profile recommendations for measuring health (7), possibly because these individual sleep characteristics sometimes show weaker associations with mortality relative to other established risk factors (1,8,9).

A limitation of many studies of sleep and mortality is that they examine the importance of each individual sleep characteristic

separately. But the clinical reality is that sleep is multidimensional, with characteristics representing duration, timing, continuity, and sleepiness existing concurrently as well as in the presence (or absence) of sleep disorders and medications (10). Considering sleep as a single multidimensional construct, rather than a series of separate characteristics, could provide a more comprehensive understanding of its predictive ability and enhance recommendations for measuring public health (8,10).

To date, efforts to quantify "multidimensional sleep" primarily have been limited to counting the number of adverse sleep characteristics (6,8,11). However, a count is not likely an accurate representation of the relationship of multiple aspects of sleep to health (8), as sleep characteristics may interact with one another and contribute different weights in relation to specific health outcomes. Moreover, sleep–health associations need to be considered in the context of nonsleep predictors and confounders, including sociodemographic factors (12), physical and mental health (12–16), health behaviors (12,17), and medication use (12,18,19).

The aims of this study are to: (i) determine the predictive abilities of multidimensional self-reported sleep for mortality relative to other established self-reported risk factors and (ii) to identify which self-reported sleep characteristics are most predictive. These aims call for flexible methods that empirically organize numerous interrelated predictors and a sample big enough to support such methodology. Therefore, we compiled a large (*N* = 8,668) multicohort sample of black and white older adults aged 65–99 years (mean = 78.7, median = 80) with self-reported sleep characteristics and nonsleep risk factors captured during an initial assessment, followed for ≤ 15 years for mortality. We applied random survival forest machine learning to quantify and rank the predictive abilities of each risk-factor and prespecified "domains" to which they belong (20).

Methods

Sample
Our analytic sample consists of 8,668 older adults aged ≥65 years, aggregated from three epidemiological cohorts: the Sleep Heart Health Study (SHHS) (21), Study of Osteoporotic Fractures (SOF) (22,23), and Outcomes of Sleep Disorders in Older Men Study (MrOS Sleep; http://mrosdata.sfcc-cpmc.net) (24,25). All participants had complete self-reported sleep characteristics, longitudinal follow-up, and provided written informed consent to participate in longitudinal sleep studies. There were less than 5% missing data for each nonsleep predictor; *N*s of measures with missing data are provided in Table 1. To reduce cohort-related confounding, we included only participants who reported black or white race, excluding *N* = 247 of the available sample. Further cohort details are provided in the Supplementary Figure and Text.

Predictors
Predictors were grouped into six domains: Sleep, Sociodemographic Factors, Health Behaviors, Mental Health, Medications, and Physical Health. We included predictors that had previously been associated with mortality, were clinically meaningful, and which could be harmonized across cohorts. Harmonization details are provided in Supplementary Tables S1 and S2.

Sleep
Self-reported habitual sleep characteristics
We considered Total Sleep Time (usual total hours of reported sleeping), Bed Time (Time an individual usually goes to bed at night),

Wake-up Time (Time an individual usually wakes up in the morning), Time in Bed (time elapsed between Bed and Wake-up), Sleep Efficiency (Total Sleep Time/Time in Bed × 100), Sleep Latency (Usual time reported to fall asleep), Napping (number of hours napping per week), and the Epworth Sleepiness Scale (26). In cohort-specific MrOS and SOF analyses, we also included the sleep quality item from the Pittsburgh Sleep Quality Index (PSQI; "Rate your usual sleep quality the past month") (27).

Self-reported symptoms of sleep disorders
Sleep apnea symptoms were represented by "Frequent Snoring" (≥3 night per week) and "Ever Stopped Breathing During Sleep". Insomnia symptoms were represented by perceived problems related to "Difficulty Falling Asleep" and "Waking Too Early."

Medications with recognized effects on sleep
We included three indicators of medications with recognized effects on sleep: (i) Frequent ("often" or "almost always") use of prescription or nonprescription sleep aids; (ii) Recent use of prescription sedative-hypnotics (benzodiazepines, benzodiazepine receptor agonists) or other benzodiazepines; and (iii) Recent use of tricyclic antidepressants, trazodone, mirtazapine, or nefazodone. "Recent use" was defined as the past 2 weeks for MrOS/SOF and the past 30 days for SHHS. Some individuals may have used these medication classes for indications other than sleep problems. However, the intended use would not alter these medications' effects on sleep.

Sociodemographic
Age, sex, education, race, and marital status were included.

Health Behaviors
We included indicators for smoking status and alcohol use. In MrOS analyses, we included the Physical Activity Scale for the Elderly (28). In SOF analyses, we included the estimated number of calories burned per week from walking. Measures of physical activity were not publicly available in SHHS.

Mental Health
Depressive symptoms
We included a three-item measure of depressive symptoms, harmonized from three overlapping items in the Geriatric Depression Scale (GDS; available in MrOS and SOF) (29) and the SF-36 (available in SHHS) (30). These three items reflected both the depressed mood and life satisfaction component subscales on the GDS (31) and were summed such that scores ranged from 0 to 3. This three-item score had a correlation of *r* = .80 with the full GDS in MrOS and SOF. In cohort-specific analyses, we included the full GDS score (MrOS and SOF) or the mental health component of the SF-36 (SHHS) instead of the three-item depression scale.

Anxiety and cognition
In MrOS and SOF analyses, we included anxiety symptoms (Goldberg Anxiety and Depression Scale [GADS]) (32) and cognition (Teng 3S Modified Mini-Mental State Exam (33) and Mini-Mental State Exam (34), respectively).

Medications
We included indicators for recent use (past 2 weeks for MrOS/SOF; past 30 days for SHHS) of nontricyclic antidepressants, nonsteroidal anti-inflammatory drugs, aspirin, corticosteroids, beta blockers,

Table 1. Sample Nonsleep Characteristics (*N* = 8,668)

Sociodemographic, %(*N*) or Mean (*SD*)	
Age	78.7 (6.7)
Black (vs White)	8.3 (718)
Female (vs Male)	54.0 (4,682)
≥College education (*N* = 8,639)	34.9 (3,017)
Marital status (*N* = 8,651)	
Married	57.5 (4,970)
Widowed	32.8 (2,837)
Other	9.8 (844)
Health behaviors, %(*N*)	
Smoking status (*N* = 8,665)	
Current	3.4 (298)
Past	45.7 (3,961)
Never	50.9 (4,406)
Alcoholic drinks per week (*N* = 8,609)	
≥14 drinks/wk	10.5 (905)
7–13 drinks/wk	38.2 (3,290)
<7 drinks/wk	51.3 (4,414)
Mental health symptoms, mean (*SD*)	
Depressive symptoms (range 0–3; *N* = 8,555)	0.7 (0.78)
Physical health conditions, %(*N*) or mean (*SD*)	
Body mass index (*N* = 8,380)	27.2 (4.33)
Number of functional limitations (*N* = 8,459) (range 0–4)	1.0 (1.3)
Self-rated health (*N* = 8,587) (1 = Excellent, 4 = Poor/Very Poor)	1.9 (0.7)
History of:	
Stroke (*N* = 8,639)	8.4 (724)
Angina (*N* = 8,626)	13.4 (1,155)
Heart failure (*N* = 8,615)	6.4 (551)
Heart attack (*N* = 8,630)	13.4 (1,157)
High blood pressure (*N* = 8,632)	53.5 (4,616)
Diabetes (*N* = 8,643)	11.7 (1,007)
Chronic obstructive pulmonary disease (*N* = 8,611)	13.3 (1,149)
Nonsleep medications, %(*N*) or mean (*SD*)	
Number of prescription medications	3.9 (3.0)
Antidepressants (nontricyclic)	6.1 (524)
Corticosteroids (*N* = 8,645)	7.5 (644)
Nonsteroidal anti-inflammatory drugs	20.8 (1,801)
Aspirin	51.1 (4,432)
Hypoglycemic agents	9.0 (776)
Lipid-lowering medications	27.1 (2,348)
Diuretics	32.1 (2,779)
Calcium channel blockers	18.9 (1,637)
Ace inhibitors	21.7 (1,878)
Beta blockers	27.0 (2,336)

angiotensin-converting enzyme (ACE) inhibitors, calcium channel blockers, diuretics, and lipid-lowering medications. These medications were selected because they either reflected treatment for underlying physical comorbidities related to mortality and/or had themselves been associated with health outcomes. We also included a count of the total number of prescription medications per participant to characterize polypharmacy.

Physical Health

We included measured body mass index, self-reported health status, number of functional limitations, and indicators for self-reported history of stroke, heart attack, angina, heart failure, high blood pressure, diabetes, or chronic obstructive pulmonary disease (COPD). In MrOS and/or SOF analyses, we also included indicators for history of osteoporosis, arthritis (rheumatoid or osteoarthritis), and cancer (SOF only).

Outcomes

Our primary outcome is time to all-cause mortality, selected because it is an unequivocal "hard" outcome that is of importance to patients and society. Our secondary outcome is time to cardiovascular mortality. Full outcome details, including adjudication processes, are provided in the Supplementary Text.

In the full analytic sample, 41% (*N* = 3,552) participants are known to have died of any cause. The remainder were either lost to follow up or had not yet died at the time of the last follow-up assessment. The median (25%, 75%) years to all-cause mortality was 12.3 (6.5, >15.5) years. Thirteen percent of participants (*N* = 1,079 of *N* = 8,130 with cause-specific mortality data) are known to have died of cardiovascular disease.

Statistical Analysis

Our primary analytic method was a random survival forest, a machine learning algorithm generated by creating thousands of bootstrap samples (each omitting 37% of the full sample, referred to as "out of bag" [OOB] data) and fitting a tree-structured survival model to each one. Tree-structured survival models empirically identify the predictor and cut-point that will optimally split a sample into two subsamples with different survival outcomes; splitting continues iteratively on each subsample until a predefined stopping rule is met. A new individual's survival outcome is predicted by aggregating the tree-level estimates. Random survival forests can accommodate numerous predictors with complex, nonlinear associations; they also accommodate missing covariate data through built-in imputation (20).

Using the full sample, we fit random survival forests for time to all-cause and cardiovascular mortality, considering 47 sleep and non-sleep predictors (Tables 1 and 2; categorical measures were included as a set of indicator variables with a reference group omitted). From each model, we extracted the variable importance (VIMP) of each individual predictor and the joint VIMP for a priori domains of predictors (sleep, physical health, health behaviors, medications, and sociodemographic factors). We used bootstrapping to estimate 99.9% confidence intervals of the VIMP for each predictor/domain, adjusted for multiple comparisons (35). The VIMP is the difference in prediction error between a model including the original predictor(s) versus a model including a randomly permuted version of the predictor(s). Prediction error is computed as 1—the Concordance (C) Index (36). The C-Index is related to the area under the curve and estimates the probability that, for two randomly selected participants, the one who died first had a worse predicted outcome (20). Variables/domains with larger VIMPs have greater predictive abilities; variables/domains with VIMPs ≤ 0 are nonpredictive. In secondary analyses, we applied similar methods to stratified samples (Males < 80, Males ≥ 80, Females < 80, Females ≥ 80, MrOS, SOF, SHHS).

VIMPs can produce unexpected results if predictors are highly correlated (20). Our 47 predictors had low correlation overall (Median [Q1, Q3] Spearman |*r*| = 0.05 [0.02, 0.11]) with only one pair having a large correlation (*r* = .56 between Wake-up Time and Time in Bed). Thus, we expect the VIMPs to accurately reflect the predictive abilities of the variables/domains. However, as a more rigorous assessment of the predictive utility of sleep relative to other domains, we also compared the prediction errors (1-C-Index, computed on OOB data) from random survival forests derived with and without each domain.

Random survival forests analyses were implemented through the randomForestSRC package in R (37). Additional methodological details, including specific parameters settings, are provided in the Supplementary Text.

Results

Characteristics of the full analytic sample are provided in Tables 1 and 2. Characteristics of stratified samples are provided in Supplementary Tables S3–S6.

All-Cause Mortality

Figure 1 displays the VIMP (99.9% CI) of each predictor and domain for all-cause mortality. The joint VIMP of the sleep domain was significant (0.94 [0.60, 1.29]) with confidence limits overlapping those of functional limitations (1.66 [1.24, 2.08]), the medication domain (1.33 [0.95, 1.72]), self-rated health status (0.51 [0.35, 0.68]), and prior heart failure (0.42 [0.23, 0.62]). The top significant individual sleep characteristics were Time in Bed (0.21 [0.12, 0.30]) and Napping (0.18 [0.07, 0.30]); these were of similar importance as other well-established domains/predictors including body mass index (0.24 [0.15, 0.33]), the health behaviors domain (0.22 [0.06, 0.38]), and prior stroke (0.17 [0.04, 0.31]).

The prediction error for the full model including all predictors was 27.02%. The change in prediction error after removing each domain separately was: 1.20 (physical health), 0.44 (medications), 0.25 (sleep), 0.21 (sociodemographic factors, excluding age), and 0.14 (health behaviors). The prediction error for a base model including only age and cohort was 32.04%. The change in prediction error after adding each domain separately was: −2.13 (physical health), −0.82 (sleep), −0.08 (medications), 0.31 (health behaviors), and 1.10 (sociodemographic factors, excluding age).

Supplementary Tables S7 and S8 provide results from stratified analyses. After adjusting for multiple comparisons, multidimensional sleep was a significant predictor of all-cause mortality among females age ≥80 and in the MrOS and SOF cohorts. (It was significant without adjustment across all stratified analyses.) The sleep domain was consistently ranked as less predictive than the physical health, sociodemographic, mental health (MrOS and SOF), and medications domains. However, it was consistently ranked as more predictive than the health behaviors domain, even after including physical activity (MrOS and SOF). Time in Bed and a measure of timing (Bed or Wake-up) were among the top sleep predictors across all stratified

analyses. Napping was also among the top sleep predictors in all stratified analyses except males ≥80.

Cardiovascular Mortality

Figure 2 displays the VIMP (99.9% CI) of each individual predictor and domain for cardiovascular mortality. The multidimensional sleep domain was a significant predictor (1.98 [1.31, 2.64]) with confidence limits overlapping the sociodemographic factors domain (3.59 [2.52, 4.66]), the medication domain (2.82 [1.99, 3.65]), age (2.63 [1.87, 3.38]), and functional limitations (1.09 [0.76, 1.42]). The top significant individual sleep characteristics were Time in Bed (0.36 [0.22, 0.50]), Napping (0.29 [0.12, 0.45]), and Wake-up Time (0.26 [0.14, 0.39]); these had VIMPs similar to other well-established domains/predictors including the health behaviors domain (0.32 [0.09, 0.55]) and depressive symptoms (0.24 [0.11, 0.37]). The prediction error for the model was 24.80%.

Discussion

In a large, heterogeneous sample of older adults, we used machine learning to establish the predictive ability of multidimensional sleep for mortality relative to other self-reported established risk factors. Based on a Variable Importance metric, the multidimensional sleep domain was a significant predictor of both all-cause and cardiovascular mortality in the full sample. Across full and stratified analyses, its predictive ability ranked lower than that of the physical health, sociodemographic (including age), mental health, and medication domains, but higher than that of the health behaviors domain and several well-established individual nonsleep predictors (eg, self-rated health status, heart failure). Analyses directly comparing models with and without the multidimensional sleep domain further established that the improvement in predictive accuracy offered by self-reported sleep is in line with that offered by our medication, sociodemographic factors (excluding age), and health behaviors domains.

These findings highlight the importance of developing multidimensional measures of sleep—especially those incorporating time in bed, timing, and napping, which were among the most predictive individual sleep characteristics. Such multidimensional sleep measures could improve and standardize longitudinal follow-up of health in epidemiological studies of older individuals. Moreover, unlike some nonsleep measures with similar predictive abilities (eg, prior stroke, diabetes), time in bed, timing, and napping are modifiable in older adults by behavioral measures (38). The combination of modifiability and predictive ability makes multidimensional sleep a promising and novel target for improving health and reducing mortality risk in older adults.

This study is the first to incorporate a comprehensive set of sleep characteristics, put them in the context of well-established nonsleep predictors, and link them to mortality in a large heterogeneous sample using machine learning. While machine learning provides a powerful tool for organizing numerous interrelated predictors to examine their relative importance, it does not provide information about direction/magnitude of effects or potential pathways to mortality. Nevertheless, our finding that time in bed, timing (with potential for circadian misalignment), and napping were the most predictive sleep characteristics is consistent with a growing literature suggesting that these characteristics influence inflammation and metabolism, key pathways associated with health and disease (1,39,40).

Combining three cohorts improved our study by increasing sample size, enhancing generalizability, and facilitating stratified
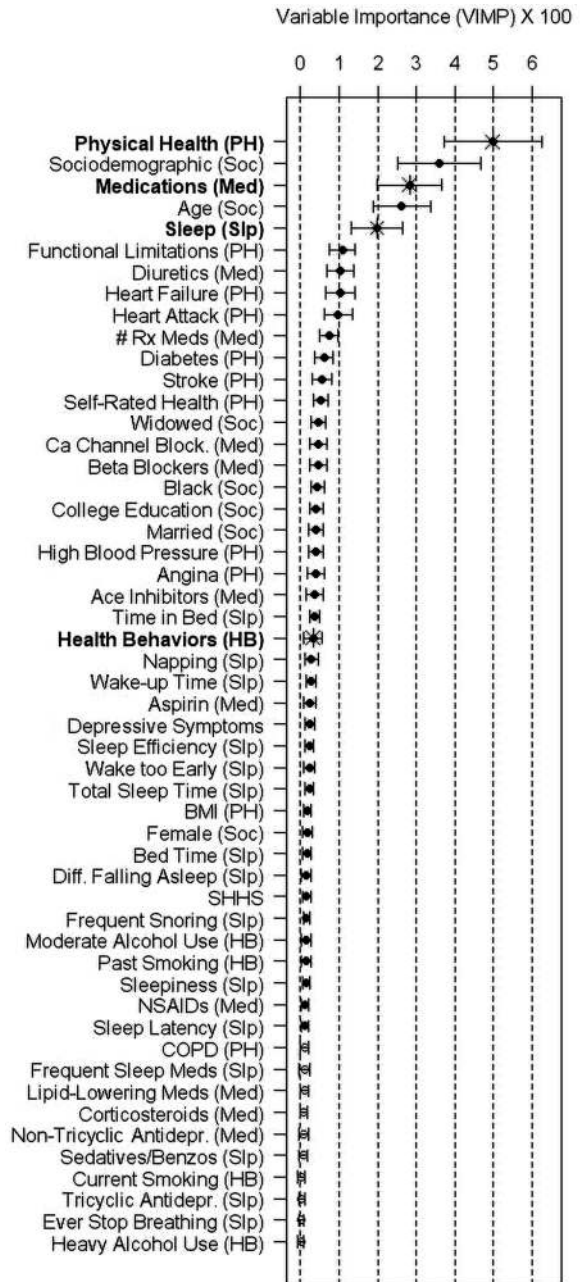
Table 2. Sample Sleep Characteristics (*N* = 8,668)

Habitual Sleep Characteristics	%(N) or Mean(SD)
Epworth sleepiness scale	6.3 (4.0)
Hours of napping per week	1.8 (3.3)
Bed time (HH:MM)	22:41 (01:14)
Wake-up time (HH:MM)	06:55 (01:19)
Total sleep time (hours)	7.0 (1.3)
Time in bed (hours)	8.2 (1.4)
Sleep latency (minutes)	22.1 (24.5)
Sleep efficiency	85.2 (13.6)
Symptoms of sleep disorders	
Frequent snoring	15.6 (1,352)
Ever stop breathing during sleep	8.1 (706)
Difficulty falling asleep	21.5 (1,863)
Wake up too early	38.6 (3,342)
Medications with recognized effects on sleep	
Sedatives and other hypnotics or benzodiazepines	7.2 (623)
Frequent use of sleep medications	15.1 (1,311)
Tricyclic antidepressants, mirtazapine, nefazodone, or trazodone	4.0 (343)
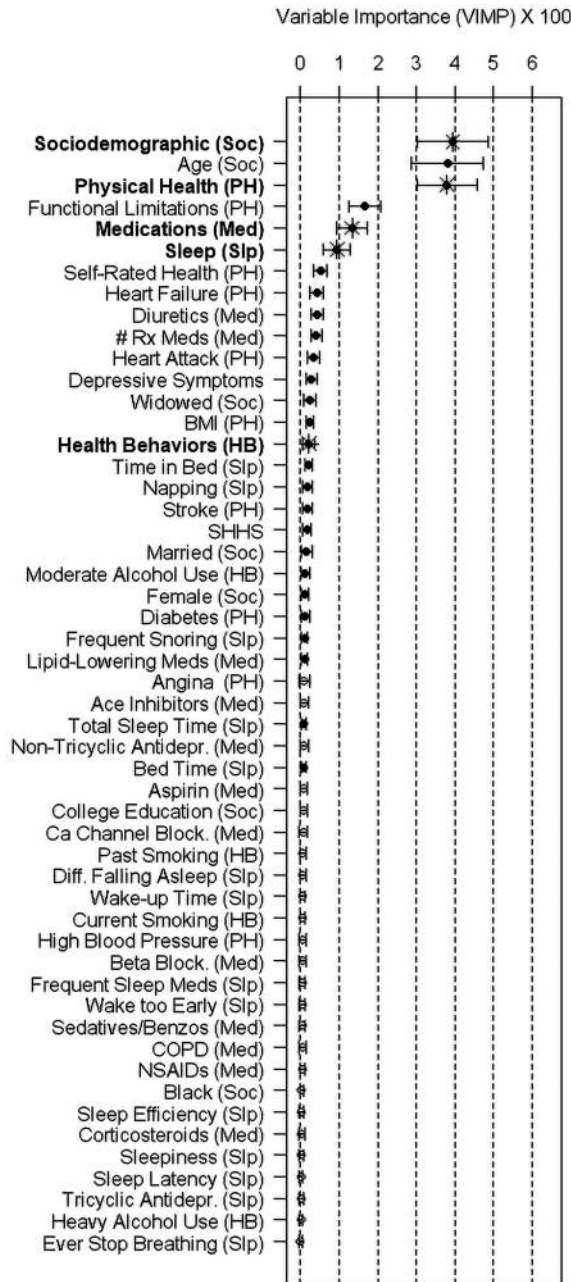
Figure 1. Variable importance values × 100 (VIMPs; 99.9% confidence intervals) for all-cause mortality. Domains of multiple measures are bolded. Abbreviations following each individual (non-bolded) measure reflect the domain in which they are grouped.

Figure 2. Variable importance values × 100 (VIMPs; 99.9% confidence intervals) for time to cardiovascular mortality. Domains of multiple measures are bolded. Abbreviations following each individual (non-bolded) measure reflect the domain in which they are grouped.

analyses. However, this may have produced residual confounding, especially related to age and gender. Another limitation is that some domains contained fewer measures than other domains, potentially limiting their predictive abilities. Although adding informative measures to a domain could improve its predictive abilities, adding noninformative measures could induce noise and worsen prediction. Also, the physical health conditions in the full harmonized sample were primarily limited to cardiometabolic-related diseases. Although these diseases are generally the most prevalent and strongest risk factors for mortality in older adults (12), some important

noncardiometabolic conditions were excluded in the full sample. Despite these limitations, several sensitivity analyses probing their potential impact produced relatively consistent results. Thus, we expect that our main findings are robust.

In future research, it will be important to develop tools for measuring multidimensional sleep, especially those incorporating time in bed, timing, and napping. These tools should be used to examine nuanced associations between sleep and mortality, probe mechanistic pathways (eg, inflammation or metabolism), and establish effect modifiers (eg, pain, postmenopausal vasomotor symptoms,

or marital status, among others). Finally, we focused this analysis on self-reported data given their wide availability and scalability. However, future research should consider objective sleep measures (eg, captured through polysomnography or actigraphy), which provide more granularity and may be even stronger predictors of mortality in older adults than self-reported sleep (8).

Supplementary Material

Supplementary data is available at *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences* online.

Funding

Conflict of Interest

K.E. and M.L.W. receive grant support from the NIH (and supporting agencies) as listed under Funding Sources. K.S. receives grant funding from Merck and NIH (and supporting agencies) as listed under Funding Sources. D.J.B. reports receiving consulting fees from American Academy of Physician Assistants, Bayer HealthCare, BeHealth Solutions, CME Institute, Emmi Solutions, and grants from NIH, outside the submitted work. In addition, D.J.B. receives licensing fees (royalties) for the Pittsburgh Sleep Quality Index (PSQI), which is copyrighted by the University of Pittsburgh. S.R. reports grant support from JAZZ Pharma, ASMF, and NIH, outside the submitted work. In addition, S.R. reports consulting for Jazz Pharma, outside the submitted work. S.A.-I. is a consultant for Eisai, Purdue, Merck, Acadia, and Jazz Pharmaceuticals. M.H.H. and Y.L. report no conflicts of interest.

References

1. Hall MH, Smagula SF, Boudreau RM, et al. Association between sleep duration and mortality is mediated by markers of inflammation and health in older adults: the Health, Aging and Body Composition Study. *Sleep*. 2015;38:189–195. doi:10.5665/sleep.4394

2. da Silva AA, de Mello RG, Schaan CW, Fuchs FD, Redline S, Fuchs SC. Sleep duration and mortality in the elderly: a systematic review with meta-analysis. *BMJ Open*. 2016;6:e008119. doi:10.1136/bmjopen-2015-008119

3. Newman AB, Spiekerman CF, Enright P, et al. Daytime sleepiness predicts mortality and cardiovascular disease in older adults. The Cardiovascular Health Study Research Group. *J Am Geriatr Soc*. 2000;48:115–123. doi:10.1111/j.1532–5415.2000.tb03901.x

4. Zhong G, Wang Y, Tao T, Ying J, Zhao Y. Daytime napping and mortality from all causes, cardiovascular disease, and cancer: a meta-analysis of prospective cohort studies. *Sleep Med*. 2015;16:811–819. doi:10.1016/j.sleep.2015.01.025

5. Lin X, Chen W, Wei F, Ying M, Wei W, Xie X. Night-shift work increases morbidity of breast cancer and all-cause mortality: a meta-analysis of 16 prospective cohort studies. *Sleep Med*. 2015;16:1381–1387. doi:10.1016/j.sleep.2015.02.543

6. Rod NH, Vahtera J, Westerlund H, et al. Sleep disturbances and cause-specific mortality: results from the GAZEL cohort study. *Am J Epidemiol*. 2011;173:300–309. doi:10.1093/aje/kwq371

7. Diehr PH, Thielke SM, Newman AB, Hirsch C, Tracy R. Decline in health for older adults: five-year change in 13 key measures of standardized health. *J Gerontol A Biol Sci Med Sci*. 2013;68:1059–1067. doi:10.1093/gerona/glt038

8. Wallace ML, Stone K, Smagula SF, et al. Which sleep health characteristics predict all-cause mortality in older men? An application of flexible multivariable approaches. *Sleep*. 2018;41: zsx189. doi:10.1093/sleep/zsx189

9. Wingard DL, Berkman LF. Mortality risk associated with sleeping patterns among adults. *Sleep*. 1983;6:102–107. doi:10.1093/sleep/6.2.102

10. Buysse DJ. Sleep health: can we define it? Does it matter? *Sleep*. 2014;37:9–17. doi:10.5665/sleep.3298

11. Furihata R, Hall MH, Stone KL, et al. An aggregate measure of sleep health is associated with prevalent and incident depression among community-dwelling older women. *Sleep*. 2017;40:zsw075. doi:10.1093/sleep/zsw075

12. Fried LP, Kronmal RA, Newman AB, et al. Risk factors for 5-year mortality in older adults: the Cardiovascular Health Study. *JAMA*. 1998;279:585–592. doi:10.1001/jama.279.8.585

13. Van Aller, C, Lara J, Stephan BCM, et al. Sarcopenic obesity and overall mortality: results from the application of novel models of body composition phenotypes to the National Health and Nutrition Examination Survey 1999–2004. *Clin Nutr*. 2018;38:264–270. doi:10.1016/j.clnu.2018.01.022

14. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav*. 1997;38:21–37. doi:10.2307/2955359

15. Nybo H, Petersen HC, Gaist D, et al. Predictors of mortality in 2,249 nonagenarians–the Danish 1905-Cohort Survey. *J Am Geriatr Soc*. 2003;51:1365–1373. doi:10.1046/j.1532-5415.2003.51453.x

16. Diez-Quevedo C, Lupón J, González B, et al. Depression, antidepressants, and long-term mortality in heart failure. *Int J Cardiol*. 2013;167:1217–1225. doi:10.1016/j.ijcard.2012.03.143

17. White J, Greene G, Kivimaki M, Batty GD. Association between changes in lifestyle and all-cause mortality: the Health and Lifestyle Survey. *J Epidemiol Community Health*. 2018;72:711–714. doi:10.1136/jech-2017-210363

18. Brouwers C, Christensen SB, Damen NL, et al. Antidepressant use and risk for mortality in 121,252 heart failure patients with or without a diagnosis of clinical depression. *Int J Cardiol*. 2016;203:867–873. doi:10.1016/j.ijcard.2015.11.032

19. Bonaga B, Sánchez-Jurado PM, Martínez-Reig M, et al. Frailty, polypharmacy, and health outcomes in older adults: the frailty and dependence in Albacete study. *J Am Med Dir Assoc*. 2018;19:46–52. doi:10.1016/j.jamda.2017.07.008

20. Ishrawan H, Kogalur UB, Blackston EH, et al. Random survival forests. *Ann Appl Stat*. 2008;2:841–860. doi:10.1214/08-AOAS169

21. Quan SF, Howard BV, Iber C, et al. The sleep heart health study: design, rationale, and methods. *Sleep*. 1997;20:1077–1085. doi:10.1093/sleep/20.12.1077

22. Cummings SR, Black DM, Nevitt MC, et al. Appendicular bone density and age predict hip fracture in women. The Study of Osteoporotic Fractures Research Group. *JAMA*. 1990;263:665–668. doi:10.1001/jama.1990.03440050059033

23. Cummings SR, Black DM, Nevitt MC, et al. Bone density at various sites for prediction of hip fractures. The Study of Osteoporotic Fractures Research Group. *Lancet*. 1993;341:72–75. doi:10.1016/0140-6736(93)92555-8

24. Blank JB, Cawthon PM, Carrion-Petersen ML, et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials*. 2005;26:557–568. doi:10.1016/j.cct.2005.05.005

25. Orwoll E, Blank JB, Barrett-Connor E, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study–a large observational study of the determinants of fracture in older men. *Contemp Clin Trials*. 2005;26:569–585. doi:10.1016/j.cct.2005.05.006

26. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14:540–545. doi:10.1093/sleep/14.6.540

27. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28:193–213. doi:10.1016/0165-1781(89)90047-4

28. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol*. 1993;46:153–162. doi:10.1016/0895-4356(93)90053–4

29. Aikman GG, Oehlert ME. Geriatric depression scale: long form versus short form. *Clin Gerontol*. 2001;22. doi:10.1300/J018v22n03_07

30. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). *Med Care*. 1992;30:473–483.

31. Friedman B, Heisel MJ, Delavan RL. Psychometric properties of the 15-item geriatric depression scale in functionally impaired, cognitively intact, community-dwelling elderly primary care patients. *J Am Geriatr Soc*. 2005;53:1570–1576. doi:10.1111/j.1532-5415.2005.53461.x

32. Goldberg DP, Bridges K, Duncan-Jones P, Grayson D. Dimensions of neuroses seen in primary-care settings. *Psychol Med*. 1987;17:461–470. doi:10.1017/S0033291700025022

33. Teng EL, Chui HC. The modified mini-mental state (3MS) examination. *J Clin Psychiatry*. 1987;48:314–318. doi:10.1007/springerreference_183334

34. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189–198. doi:10.1016/0022-3956(75)90026-6

35. Ishwaran H, Lu M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Stat Med*. 2019;38:558–582. doi:10.1002/sim.7803

36. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61:92–105. doi:10.1111/j.0006-341X.2005.030814.x

37. Ishwaran H, Kogalur UB. *Random Forests for Survival, Regression, and Classification (RF-SRC), R package version 2.5.1*, 2017. https://CRAN.R-project.org/package=randomForestSRC

38. Ancoli-Israel S. Sleep and aging: prevalence of disturbed sleep and treatment considerations in older adults. *J Clin Psychiatry*. 2005;66(Suppl 9):24–30; quiz 42.

39. Okun ML, Reynolds CF 3rd, Buysse DJ, et al. Sleep variability, health-related practices, and inflammatory markers in a community dwelling sample of older adults. *Psychosom Med*. 2011;73:142–150. doi:10.1097/PSY.0b013e3182020d08

40. Li X, Pang X, Zhang Q, et al. Long-term single and joint effects of excessive daytime napping on the HOMA-IR index and glycosylated hemoglobin. *Medicine (Baltim)..* 2016;95:e2734. doi:10.1097/MD.0000000000002734