Assessing Risk of Health Outcomes From Brain Activity in Sleep
A Retrospective Cohort Study

Haoqi Sun, PhD*, Noor Adra, MD*, Muhammad Abubakar Ayub, MD*, Wolfgang Ganglberger, PhD, Elissa Ye, MSc, Marta Fernandes, PhD, Luis Paixao, MD, Ziwei Fan, MSc, Aditya Gupta, MSc, Manohar Ghanta, MSc, Valder F. Moura Junior, MSc, Jonathan Rosand, MD, MSc, M. Brandon Westover, MD, PhD, and Robert J. Thomas, MD†

Neurology: Clinical Practice 2024;14:e200225. doi:10.1212/CPJ.0000000000200225

Correspondence
Dr. Thomas
rthomas1@bidmc.harvard.edu

Abstract

Background and Objectives
Patterns of electrical activity in the brain (EEG) during sleep are sensitive to various health conditions even at subclinical stages. The objective of this study was to estimate sleep EEG-predicted incidence of future neurologic, cardiovascular, psychiatric, and mortality outcomes.

Methods
This is a retrospective cohort study with 2 data sets. The Massachusetts General Hospital (MGH) sleep data set is a clinic-based cohort, used for model development. The Sleep Heart Health Study (SHHS) is a community-based cohort, used as the external validation cohort. Exposure is good, average, or poor sleep defined by quartiles of sleep EEG-predicted risk. The outcomes include ischemic stroke, intracranial hemorrhage, mild cognitive impairment, dementia, atrial fibrillation, myocardial infarction, type 2 diabetes, hypertension, bipolar disorder, depression, and mortality. Diagnoses were based on diagnosis codes, brain imaging reports, medications, cognitive scores, and hospital records. We used the Cox survival model with death as the competing risk.

Results
There were 8673 participants from MGH and 5650 from SHHS. For all outcomes, the model-predicted 10-year risk was within the 95% confidence interval of the ground truth, indicating good prediction performance. When comparing participants with poor, average, and good sleep, except for atrial fibrillation, all other 10-year risk ratios were significant. The model-predicted 10-year risk ratio closely matched the observed event rate in the external validation cohort.

Discussion
The incidence of health outcomes can be predicted by brain activity during sleep. The findings strengthen the concept of sleep as an accessible biological window into unfavorable brain and general health outcomes.

Introduction
Good sleep and healthy life are closely associated. For example, people with dementia have difficulty falling asleep and have reduced delta oscillations (1–4 Hz) in their brain waves (EEG) during deep sleep compared with matched controls without cognitive problems.¹

¹These authors contributed equally to this work as co-first authors.
†These authors contributed equally to this work as co-senior authors.
People with sleep disorders or anxiety tend to have insomnia and increased sleep fragmentation compared with people without depression or anxiety while controlling for covariates. People with atrial fibrillation or congestive heart failure have fragmented sleep beyond the frequently present central and obstructive sleep apnea compared with controls. Normal nocturnal dipping of blood pressure during sleep is associated with overall better cardiovascular and cerebrovascular outcomes compared with people who lack or have diminished nocturnal dipping. In a large-scale international study, multiple aspects of sleep disturbance were found to be positively associated with existing conditions of stroke, although sleep microstructures such as EEG were not examined.

The ability to use physiologic measurements of sleep, such as EEG, to predict future incident health outcomes is significant because it could allow early interventions to prevent unfavorable outcomes. Because sleep is not only a window into sleep health but also a causal determinant, such interventions could also aim to improve sleep quality, which is a major goal of sleep medicine and the pharmaceutical industry. Yet, how to measure sleep quality remains unclear. Conventional metrics including sleep stages and arousals seem limited and are highly down-sampled summaries of sleep physiology and pathology. One approach is to develop measures of sleep quality that are associated with clinical outcomes by design. Along these lines, there are several recently introduced sleep-based biomarkers associated with clinical outcomes by design. Along these lines, there are several recently introduced sleep-based biomarkers associated with clinical outcomes by design.

We describe measures of sleep quality that are explicitly designed to predict future health outcomes. We characterize the risk of these outcomes following baseline sleep measurements and investigate sleep features that make these predictions possible. As a comparison, we also compute the risk ratios of these outcomes following baseline sleep measurements to quantify the risk of future health outcomes. We characterize the baseline recording. The flowchart is shown in eFigure 2.

For external validation, we used the Sleep Heart Health Study (SHHS) data, available from the National Sleep Research Resource. SHHS is a community-based cohort. We used the PSGs from sleep visit 1. We studied 3 outcomes available in this data set: ischemic stroke, myocardial infarction, and death. Inclusion criteria were (1) EEG signals and annotations available, (2) at least 3 sleep stages were present in the EEG, and (3) the sleep recording was at least 2.5 hours long.

**Standard Protocol Approvals, Registrations, and Patient Consents**

The Mass General Brigham Institutional Review Board (IRB) approved the analysis of PSG acquired in the Massachusetts General Hospital Sleep Clinic, with a waiver of informed consent for this retrospective study.

**Sleep EEG Preprocessing and Feature Extraction**

For the MGH data set, each PSG includes 6 EEG channels: F3-M2, F4-M1, C3-M2, C4-M1, O1-M2, and O2-M1. The sampling frequency was 200 Hz. For SHHS, each PSG includes 2 EEG channels: C3-M2 and C4-M1. The sampling frequency was 125 Hz. EEG signals were first notch filtered at 60 Hz and then band-pass filtered from 0.3 to 35 Hz. Every 30-second epoch was staged into one of W, N1, N2, N3, or R following American Academy of Sleep Medicine (AASM) guidelines. We combined N1, N2, and N3 into NREM (1) to accommodate the relatively low number of N1 epochs and (2) to avoid the artificial boundary between N2 and N3. We also provided a score computation table (eTable 1) and a score-to-risk conversion chart (eFigure 1) to compute the risks. We validate the results in an external large community-based sleep data set.

**Methods**

**Study Design**

This was an observational retrospective cohort study from 2008 to January 2020. As a clinic-based cohort, polysomnogram (PSG) was indicated because of suspected sleep disorders. Participants were followed up by querying the electronic health records from the Research Patient Data Registry in February 2020, sourced from all hospitals under Massachusetts General (MGB) including both inpatient and outpatient records. Inclusion criteria were those who (1) underwent a diagnostic study (not positive airway pressure (CPAP) treatment) and (2) were 18 years and older at the time of the PSG. Exclusion criteria were as follows: (1) developed the outcome before the baseline sleep recording, (2) PSGs with duration shorter than 2.5 hours (300 × 30 second epochs), and (3) took sodium oxybate on the night of the sleep study because this drug markedly increases slow-wave sleep and delta power. For participants with multiple PSGs, we used only the PSG from the earliest visit, which is referred to as the baseline recording. The flowchart is shown in eFigure 2.

Spindle and slow oscillation patterns were detected using Luna during NREM sleep. Spindles were detected based on a wavelet method, with a central frequency of 13.5 Hz and a wavelet cycle number set to 12. Slow oscillations were...
detected by first band-pass filtering between 0.2 Hz and 4.5 Hz, followed by detecting positive-to-negative zero crossings in the filtered signal, and then choosing intervals between 0.8 and 2 seconds having negative peak higher than 1.5 times the median voltage across all zero crossings and peak-to-peak amplitude higher than 1.5 times the median.

When using sleep EEG as the exposure, the signals were first preprocessed as discussed in the eMethods (links.lww.com/CPJ/A484), including spindle and slow oscillation detection. Then for each 30-second epoch, we extracted features from both spectral and temporal domains of sleep EEG. Spectral features included mean and kurtosis (a measure of extreme values to represent bursts) of band powers across 2-second sub-epochs from delta (1 Hz ≤ f < 4 Hz), theta (4 Hz ≤ f < 8 Hz), alpha (8 Hz ≤ f < 12 Hz), and sigma (12 Hz ≤ f < 16 Hz) bands and their band power ratios. Temporal features included density, amplitude, frequency, and duration of spindles; amplitude and density of slow oscillations (0.5–1.25 Hz); and coupling of spindles and slow oscillations. For each 30-second NREM epoch, we first extracted 57 features, and then each feature was averaged across all NREM epochs to represent the whole night. Similarly, for each 30-second REM epoch, we first extracted 21 features, and then each feature was averaged across all REM epochs. In addition, we adjusted for 8 covariates in each outcome prediction model as discussed in the “Covariates” subsection. In total, we had 86 features to summarize the information over the whole night of sleep. A full list of sleep EEG features is provided in eTable 2. The univariate hazard ratio of each sleep EEG feature after adjusting for covariates (feature importance) is shown in eFigure 3.

Outcomes
MCI and dementia were ascertained using problem lists, medications (donepezil, rivastigmine, memantine, galantamine), and cognition scores when available, as described in prior work. In brief, dementia was defined as at least one of the following: taking at least 1 dementia-related medication and having an ICD diagnosis code containing at least 1 dementia key phrase (“dementia” or “Alzheimer”), a problem list containing at least 1 dementia key phrase, MoCA score ≤19, or MMSE score ≤25. MCI was defined as first not qualifying for dementia and then having the problem list contain an MCI key phrase (“MCI,” “mild cognitive impairment,” or “minimal cognitive impairment”) or MoCA score between 20 and 25. Because the diagnosis of MCI or dementia prevents another (competing risk), to simplify the analysis, we combined MCI and dementia using whichever was documented first as a composite outcome. Note that 333 participants (3.8%) had a MoCA score and 343 participants (4.0%) had a MMSE score. We had 1 independent neurologist review the clinical notes and provide gold standard assessments for 160 participants, including 68 dementia, 50 MCI, 16 symptomatic, and 26 no dementia participants. Definitions were developed to maximize agreement between the gold standard and the ICD and text-based outcomes. The rules were also extensively vetted by a neurologist specializing in memory disorders.

One author (M.A.A.) independently reviewed the cases for other outcomes other than dementia, MCI, and mortality: first, randomly selected 10 predicted positive and 10 predicted negative cases, adjusted the keywords and medications to make sure the 20 cases were all correct, and then applied the rules to all the cases. Specifically, ischemic stroke and intracranial hemorrhage were ascertained using ICD diagnostic codes (ICD-10 and ICD-9 converted from ICD-10) and brain imaging reports. The ICD codes are listed in eTable 3 (links.lww.com/CPJ/A484). For the brain imaging reports, we used regular expressions to identify brain-related imaging reports (see eTable 4), then extracted the “impression” part of the report, and finally used regular expressions to identify reports describing ischemic stroke or intracranial hemorrhage (eTable 4). For atrial fibrillation, myocardial infarction, type 2 diabetes, hypertension, bipolar disorder, and depression, we used disease-related medications combined with ICD codes because the medications can be broad and nonspecific. The ICD codes are listed in eTable 3. The generic and brand names of disease-related medications are given in eTable 5. Finally, mortality was determined from the MGB Enterprise Data Warehouse (EDW), which covers all hospitals under MGB.

When participants were lost to follow-up (leaving the hospital system, never developing the outcome, or death before developing the outcome of interest, etc), their outcome statuses were included in the time-to-event analysis as censored. We used the last active date in all hospital records as the date of censoring.

For the external validation set SHHS, 3 of the 11 outcomes were available, including ischemic stroke, myocardial infarction, and death. Stroke was broadly defined as a constellation of neurologic symptoms with a sudden onset that lasts at least 24 hours or until death. The ischemic type was reviewed by physicians. Myocardial infarction was defined by a combination of chest pain, ECG tracings, and myocardial enzyme profiles. Date of death was determined using multiple approaches as described previously where all known contacts for the participant were called to determine the participant’s vital status and both local death registries and the National Death Index were searched for their name or social security number. If the participant had not developed the outcome or was deceased at the last contact, the time to event was censored.

Exposure
Exposure was defined as 3 levels: good, average, and poor sleep. When we used sleep EEG, the 3 levels were based on the sleep EEG-predicted score (see Survival Analysis below). Good sleep was defined as having a score lower than the 25% percentile of the training set, average sleep being between...
25% and 75% percentiles, and poor sleep being higher than the 75% percentile of the training set. Other alternative exposures were similarly analyzed, including AHI (scored at 3% desaturation threshold), hypoxic burden, respiratory event duration, brain age index, sleep efficiency, and WASO. For AHI, good sleep was defined as <15/hour, average sleep 15–30/hour, and poor sleep >30/hour. We also obtained the Framingham Risk Score (FRS) \(^1\) for a subset of the MGH data set. For myocardial infarction, we checked the Spearman correlation of the sleep EEG-predicted score vs FRS and the prediction performance of FRS using the metrics introduced below.

### Covariates

Eight covariates were included to address potential confounding, including age; sex; body mass index (BMI) at the baseline sleep study; and 5 categories of medications that can affect sleep: benzodiazepines, antidepressants, sedatives, antiseizure drugs, and stimulants. Medications refer to those taken on the night of the sleep study if available from the pre-sleep questionnaire or, when not available, documented in the electronic medical record within 1 month before the PSG. The brand and generic medication names of the 5 categories of medications are available in eTable 6 (links.lww.com/CPJ/A484). We stratified results by sex because it is a possible effect modifier.

For SHHS, benzodiazepine use was determined based on self-report within 2 weeks before the PSG. Medications were later categorized by physician review. Antidepressant medication use was determined similarly, including both tricyclic and non-tricyclic antidepressants (excluding monoamine oxidase inhibitors). Stimulant use was determined similarly. Information on sedative and antiseizure medication use was not available for the SHHS.

### Survival Analysis

For each outcome prediction model other than death, we modeled death as a competing outcome in a Cox proportional hazard model. To account for the large number of features and to improve interpretability, we selected features using ElasticNet. To account for the collinearity among the same EEG features from different brain regions (frontal, central, and occipital), for each outcome, we fit 3 separate models with features from each brain region. The output hazard was then averaged across the 3 brain regions. To summarize the risk of developing an outcome over time, we calculated the cumulative incidence curves (CICs). The models were adjusted for the covariates described in the Covariates subsection. The model was trained using nested 5-fold cross-validation, detailed in the eMethods (links.lww.com/CPJ/A484). The sleep EEG-predicted score of a participant is defined as \( \beta^\top x \), which is a term in the hazard function of the Cox model, \( \lambda(t) = \exp(\beta^\top x) \). The score is a weighted sum of all sleep EEG features \( x \) of that participant, where the weight is the Cox model coefficient \( \beta \).

### Evaluation Metrics

#### Cumulative Incidence Curves

To estimate the CIC for each stratum, a Cox model was first fitted using the stratified hazard in the training set as the only input and then a CIC was obtained for a specific hazard stratum. To obtain out-of-sample estimates, 5 CICs were obtained from held-out testing sets and then averaged. To get the ground truth CIC for each stratum, we calculated the Aalen-Johansen (AJ) estimate to the held-out testing sets using the stratified hazard, where the 25% and 75% percentiles came from corresponding in-sample training sets.

### Effect Sizes

We used 2 risk ratios at 10 years post-sleep study as the clinically meaningful metric: one comparing poor sleep vs average sleep, and another one comparing average sleep vs good sleep. To evaluate prediction performance, we checked whether the 95% confidence intervals of the ground truth 10-year risk from the AJ estimator and Cox model–predicted 10-year risk overlap. We also evaluated C-index in the supplemental material, including C-index for all outcomes (eTable 7, links.lww.com/CPJ/A484), in different AHI strata (eFigure 4), in people with insomnia, hypersomnia, and restless leg syndrome (eFigure 5, participant numbers of these comorbidities in eTable 8), and using covariates only (eTable 9).

### Statistical Analysis

For estimating effect sizes, covariates were adjusted for confounding by holding them to constant values and only varying the sleep EEG measures. For multiple comparison correction, we performed Bonferroni correction when investigating the significances of individual sleep EEG features (univariate analysis; eFigure 3, links.lww.com/CPJ/A484), but not for other analyses because other analyses were intended to estimate effect size rather than for hypothesis testing. We defined statistical significance as a 2-sided \( p < 0.05 \). All confidence intervals were obtained by bootstrapping 1000 times and taking the 2.5% and 97.5% percentiles as the lower and upper bounds, respectively, to form a 95% percentile confidence interval. The analyses were performed using Python 3.7 and R 4.2.

### Data Availability

The MGH data set can be requested from the corresponding author with a written data sharing agreement. The SHHS data set can be requested from sleepdata.org/datasets/shhs. The codes to reproduce the results are available at github.com/mghcDAC/sleep-outcome-prediction. The online tool to predict the risk given a sleep recording (EDF format and sleep stage annotations) will be hosted on Brain Data Science Platform at bdsp.io/.

### Results

#### Cohort Characteristics

The MGH cohort characteristics stratified by outcomes are summarized in Table 1. The detailed characteristics are provided in eTable 8 (links.lww.com/CPJ/A484). In total,
there were 8,673 participants. The average age at the baseline sleep EEG was 51.0 years; 51.5% were female. The average BMI was 29.9 kg/m². The average AHI was 9.4/hour. There were 6652 White (77%), 527 Black (6%), 414 Hispanic (5%), 332 Asian (4%), 17 American native (0.2%), 1 Middle Eastern (0.01%), and 730 unknown (8%) participants. The median follow-up time ranged from 4.5 to 5.4 years across outcomes. The average time to event ranged from 2.4 to 4.1 years across outcomes. By a maximum of 12 years after the baseline sleep study, depression, hypertension, and MCI/dementia were the top 3 outcomes with the highest incidence rate; intracranial hemorrhage, ischemic stroke, and bipolar disorder were the least. For MGH, there were 4558 of 8673 participants (53%) with AHI ≥5/hour. For SHHS, there were 2900 of 5650 participants (51%) with AHI ≥5/hour.

Cumulative Incidences

In Figure 1, we show the cumulative incidence curves of 11 outcomes for 3 strata: poor, average, and good sleep. These curves were obtained by fixing values of the covariates so that the estimated differences in risk were only related to the differences in characteristics of sleep brain activity. Overall, depression had the highest cumulative incidence at 10 years after the baseline sleep recording, followed by hypertension, and MCI/dementia were the top 3 outcomes with the highest incidence rate; intracranial hemorrhage, ischemic stroke, and bipolar disorder were the least. For MGH, there were 4558 of 8673 participants (53%) with AHI ≥5/hour. For SHHS, there were 2900 of 5650 participants (51%) with AHI ≥5/hour.

Ten-Year Risk Ratios

In Table 2, we show the sleep EEG-predicted 10-year risk ratios for poor vs average and average vs good sleep in both female and male participants. Except for atrial fibrillation in poor vs average sleep, all others showed risk ratios significantly higher than 1, indicating that sleep is predictive of future health outcomes. The average 10-year risk ratio across all outcomes using sleep EEG was 3.8 when comparing poor vs average sleep and 3.5 when comparing average vs good sleep, and male and female strata had the same results. The average 10-year risk ratios across all outcomes using alternative exposures (AHI, hypoxic burden, respiratory event duration, brain age index, sleep efficiency, and WASO) were all lower than that of using sleep EEG (eTable 10, links.lww.com/CPJ/A484).

For myocardial infarction, we additionally checked the FRS. The FRS was positively correlated with sleep EEG-predicted score at Spearman rho 0.42 (p < 0.001; n = 2161). In female
Figure 1  Cumulative Incidence of the 11 Outcomes in Female (Left of Each Subplot) and Male (Right of Each Subplot) Participants

Each subplot shows the cumulative incidence for one outcome. Cumulative incidence is the proportion of a population at risk that develops the outcome over a specified time. For each outcome, there are 3 cumulative incidence curves: poor sleep (risk score is within the upper 75th percentile, red), average sleep (risk score between the 25th and 75th percentiles, black), and good sleep (risk score lower than the 25th percentile, blue). The dashed lines represent ground truth risks from a nonparametric estimator. Note the different y-axis maximum values. The shaded areas indicate 95% confidence intervals. To ensure these estimates reflect out-of-sample performance, each curve is the average of the 5 curves from the testing sets in cross-validation.

Figure 2  Ground Truth (Black) and Model-Predicted 10-Year Risk (Red) for the Average Sleep Group for Each Outcome, Stratified by Sex: Female (A) and Male (B)

Numbers indicate the 10-year risk as percentage. Error bars indicate the 95% confidence interval. AFib = atrial fibrillation; BD = bipolar disorder; Dem = dementia; Dep = depression; HTN = hypertension; ICH = intracranial hemorrhage; IS = ischemic stroke; MCI = mild cognitive impairment; MI = myocardial infarction; T2D = type 2 diabetes.
participants, the FRS had a higher 10-year risk ratio at 34.4 (7.7–46.8) than the sleep EEG at 9.8 (7.2–56.7) when comparing poor vs good sleep. In male participants, the FRS had a higher 10-year risk ratio at 33.6 (7.9–46.5) than the sleep EEG at 9.2 (7.0–57.4). Owing to the reduced sample size with available FRS, the confidence interval was wider. Although the sleep EEG had a lower 10-year risk ratio than FRS, our objective was to show that sleep EEG also contains information about incident myocardial infarction, not to compare sleep EEG with FRS.

External Validation
The SHHS cohort characteristics stratified by outcomes are provided in eTable 11 (links.lww.com/CPJ/A484). In total, there were 5650 participants included from sleep visit 1. The average age at the time of the sleep study was 63.1 years; 52.3% were female. The average BMI was 28.2 kg/m². The average AHI was 10.2/hour. In eTable 12, we quantify the 10-year risk ratios after the baseline sleep recording. The results indicate that the confidence intervals overlap between ground truth and prediction for 3 outcomes (ischemic stroke, myocardial infarction, and mortality); hence, these risk prediction models developed on the MGH cohort generalize to the external community-based SHHS cohort.

Discussion
The results show that the content of brain electrical activity during sleep is predictable of a wide range of future health outcomes. The risk scores developed in this work can be viewed as outcome-associated measures of sleep quality scores (eTable 1 and eFigure 1, links.lww.com/CPJ/A484). We discuss the significance of the ability of sleep EEG to predict development of mild cognitive impairment, dementia, and cardiovascular outcomes and mortality. We discuss other outcomes, including subjective measures, and potential mechanisms in eDiscussion.

The finding that incident MCI and dementia are predictable from the sleep EEG is concordant with observations in prior work that many neurodegenerative diseases, including Alzheimer disease (AD) and Parkinson disease (PD), often exhibit sleep disturbances many years before other clinical symptoms develop.

Sleep fragmentation and changes in NREM delta power are early features of several transgenic mice models of AD. APP/PS1 mutant mice already exhibit a different spectral profile than wild-type mice as early as 3 months, which is the pre-plaque stage. In concordance, several studies have shown that in people already diagnosed with AD, sleep is more fragmented and the NREM sleep has reduced slow waves, spindles, and K-complexes while having increased low-frequency oscillatory activity during REM sleep. Sleep disturbances occur at early phases of AD. For example, in a study of elderly participants including 25 healthy participants and 25 with MCI, sleep was found to be...
Table 2 Ten-Year Risk Ratios Predicted by Sleep EEG or AHI Stratified by Sex

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Poor vs Average sleep</th>
<th>Average vs Good sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sleep EEG</td>
<td>AHI</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
<td>7.8 (4.8-21.3)*</td>
<td>1.3 (0.0-5.4)</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>3.4 (2.9-6.7)*</td>
<td>0.9 (0.4-2.0)</td>
</tr>
<tr>
<td>Dementia</td>
<td>6.2 (4.6-9.6)*</td>
<td>1.5 (0.8-2.7)</td>
</tr>
<tr>
<td>MCI or Dementia</td>
<td>4.1 (3.2-5.0)*</td>
<td>0.9 (0.6-1.4)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>1.6 (1.0-2.5)</td>
<td>1.3 (0.8-1.9)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>3.7 (2.8-5.8)*</td>
<td>1.5 (0.9-2.4)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>2.6 (2.3-3.5)*</td>
<td>2.0 (1.4-2.9)*</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1.8 (1.7-2.2)*</td>
<td>1.2 (0.9-1.5)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>3.3 (1.8-4.0)*</td>
<td>1.6 (0.4-4.2)</td>
</tr>
<tr>
<td>Depression</td>
<td>1.7 (1.6-1.9)*</td>
<td>1.0 (0.8-1.3)</td>
</tr>
<tr>
<td>Death</td>
<td>5.6 (4.9-7.5)*</td>
<td>1.3 (0.9-1.8)</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
<td>7.9 (4.8-21.0)*</td>
<td>1.3 (0.0-5.2)</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>3.4 (2.9-6.6)*</td>
<td>0.9 (0.4-1.9)</td>
</tr>
<tr>
<td>Dementia</td>
<td>6.1 (4.6-9.6)*</td>
<td>1.5 (0.8-2.6)</td>
</tr>
<tr>
<td>MCI or Dementia</td>
<td>4.0 (3.1-4.9)*</td>
<td>0.9 (0.6-1.3)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>1.6 (1.0-2.4)</td>
<td>1.2 (0.8-1.9)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>3.6 (2.7-5.7)*</td>
<td>1.5 (0.9-2.4)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>2.6 (2.3-3.5)*</td>
<td>2.1 (1.4-2.9)*</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1.8 (1.7-2.2)*</td>
<td>1.2 (0.9-1.5)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>3.2 (1.8-4.0)*</td>
<td>1.6 (0.4-4.3)</td>
</tr>
<tr>
<td>Depression</td>
<td>1.8 (1.6-2.1)*</td>
<td>1.0 (0.8-1.3)</td>
</tr>
<tr>
<td>Death</td>
<td>5.5 (4.8-7.4)*</td>
<td>1.3 (0.9-1.8)</td>
</tr>
</tbody>
</table>

*p < 0.05.

Significantly impaired in participants with MCI on both PSG-derived and subjective measures. The lower alpha peak frequency was also related to the early phase of AD at resting wake state. In our cohort, the reduced alpha band power for predicting dementia could be related to the lowering of the alpha peak frequency, which can enter the theta band range. However, inconsistent results are found where, in a study of 85 elderly women who were newly diagnosed with MCI or dementia 5 years after the baseline sleep recording, increased alpha and theta band powers during NREM and increased alpha and sigma band powers during REM were observed in their baseline sleep recordings. Other sleep measures such as actigraphy have also shown evidence of sleep-wake fragmentation in people at early stages of AD.

Sleep is tightly related to normal functioning of the cardiovascular system. Protocols such as partial or total sleep deprivation and extended sleep restriction have been widely used to simulate sleep disturbances in daily life and to study their cardiovascular outcomes. Lack of sleep leads to deterioration of vascular structure and function, including increased arterial stiffness, impaired coronary microcirculation, and endothelial dysfunction. Lack of sleep also increases heart rate and sympathetic activity and decreases parasympathetic activity.

Besides sleep duration, other more refined measures, such as sleep efficiency, sleep fragmentation, and circadian fragmentation or shift, have been found to predict incident cardiovascular outcomes. On the other hand, data-driven approaches that combine multimodal sleep features have also been developed to predict hypertension or composite cardiovascular outcomes, where the prediction is better than that of AHI, consistent with our comparison with AHI in Table 2.

Among cardiovascular outcomes, atrial fibrillation was the only one with an insignificant poor-to-average 10-year risk ratio when using sleep EEG (Table 2). Perhaps this should not be surprising because the association of sleep EEG with
cardiovascular outcomes is less clearly described in the literature, as compared with those of AHI or hypoxia. The associations between AHI and cardiovascular outcomes are expected because sleep apnea results in fluctuations in autonomic function and intermittent hypoxia has pleotropic undesirable cardiac and vascular effects. During sleep, parasympathetic modulation is dominant. However, sleep apnea disturbs this quiescence and leads to nocturnal arrhythmias, loss of blood pressure dipping, and blood pressure surges associated with arousals. CPAP treatment that reduces AHI has been shown to reduce nocturnal arrhythmias including atrial fibrillation.

There have been many scores developed to measure future cardiovascular risk. For example, the FRS uses age, sex, smoker, total cholesterol, high-density lipoprotein cholesterol, systolic BP, and BP being treated with medicines. The result that sleep-predicted score correlates with the FRS shows the construct validity of our approach.

Various aspects of sleep have been shown to be associated with mortality. For example, a large cohort study in 322,721 participants revealed that 7 hours of sleep is associated with a lower rate of all-cause, cardiovascular disease, and other-cause mortality, supported by other reviews. Sleep efficiency and the percentage of sleep time with oxygen saturation lower than 90% also predict mortality. Among all sleep stages, reduction in REM sleep has been found to be associated with increased all-cause, cardiovascular, and other non–cancer-related mortality in 2 independent cohorts.

Our study has limitations in terms of noise and biases. Determination of outcomes may have included noise because ICD codes may only reflect an encounter problem that may not necessarily be the true pathology. Determination of medications is also affected by different time frames from heterogeneous sources, i.e., medication of MGH participants are from a questionnaire or within 1 month before PSG from the electronic medical record and medications of SHHS participants are from 2 weeks before PSG. Another important source of noise is the night-to-night variability of sleep. This can be overcome by averaging the derived risk across multiple nights to reduce its variance as the wearable sleep devices are becoming more feasible. Translating our results to individual-level decision-making aids will need a lot more work but may first be considered on some larger scale only. Our cohort is also biased because it is from a clinical sleep laboratory; hence, it is possible that the sleep captured in the laboratory is not typical compared with sleep at home. Therefore, our results may only generalize to the sleep clinic population with suspected sleep disorders in the United States. Outcomes are almost certainly affected by the levels of available and affordable health services, and thus, these results may not generalize to other parts of the United States, underserved areas or populations, and other countries. Another source of bias is unmeasured confounding bias where variation in genes can contribute to both the changes in sleep EEG and incidence of unfavorable outcomes, creating spurious noncausal association. Furthermore, it is unclear how to best modify possible mediator sleep metrics, either at the population level or individually. Possible ways to improve sleep include cognitive behavioral therapy, optimizing circadian alignment, CPAP, oxygen supplementation for sleep hypoxia, sedatives which improve desirable sleep features, and potentially noninvasive brain stimulation. Therefore, it is plausible that some of these improvements will improve brain health.

The sleep EEG contains predictive and generalizable information about key adverse health outcomes. Sleep EEG provides an accessible biological window into brain and body health. Our study provides a theoretical basis for future changes in clinical care informed by sleep EEG-predicted risks of unfavorable outcomes. Further work is needed to validate clinical usefulness based on treatment effects on the predicted risks of outcomes.

**Study Funding**
During this research, Dr. Westover was supported by the Glenn Foundation for Medical Research and the American Federation for Aging Research through a Breakthroughs in Gerontology Grant; the American Academy of Sleep Medicine through an AASM Foundation Strategic Research Award; and grants from the NIH (R01NS102190, R01NS102574, R01NS107291, RFIAG064312, RFI1NS120947, R01AG073410) and NSF (2014431). Dr. Thomas was supported by an AASM Foundation Strategic Research Award and by the NIH (RF1AG064312, R01NS102190).

**Disclosure**
M.B. Westover is the co-founder of Beacon Biosignals and Director for Data Science for the McCance Center for Brain Health. R.J. Thomas discloses (1) patent and license/royalties from MyCardio, LLC, for the ECG-spectrogram; (2) patent and license/royalties from DeVilbiss-Drive for an auto-CPAP algorithm; and (3) consulting for Jazz Pharmaceuticals, Guidepoint Global, and GLG Councils. Other authors declare that they have no conflict of interest. Full disclosure form information provided by the authors is available with the full text of this article at Neurology.org/cp.

**Publication History**
Received by Neurology: Clinical Practice October 10, 2022. Accepted in final form October 4, 2023. Submitted and externally peer reviewed. The handling editor was Editor Luca Bartolini, MD, FAAN.

**Appendix Authors**

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haoqi Sun, PhD</td>
<td>Department of Neurology,</td>
<td>Drafting/revision of the manuscript for content,</td>
</tr>
<tr>
<td></td>
<td>Massachusetts General Hospital</td>
<td>including medical writing for content; drafting/</td>
</tr>
<tr>
<td></td>
<td>and Allison McCance Center for</td>
<td>revision of the manuscript for content, including</td>
</tr>
<tr>
<td></td>
<td>Brain Health at Mass General</td>
<td>medical writing for content; major role in the</td>
</tr>
<tr>
<td></td>
<td>Department of Neurology,</td>
<td>acquisition of data; analysis or interpretation</td>
</tr>
<tr>
<td></td>
<td>Beth Israel Deaconess Medical Center, Boston, MA</td>
<td>of data</td>
</tr>
</tbody>
</table>

Copyright © 2023 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.
<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noor Adra, MD</td>
<td>Department of Neurology, Massachusetts General Hospital, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; drafting/revision of the manuscript for content; major role in the acquisition of data; analysis or interpretation of data</td>
</tr>
<tr>
<td>Muhammad Abubakar Ayub, MD</td>
<td>Department of Neurology, Louisiana State University Health Sciences Center, Shreveport, LA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
<tr>
<td>Wolfgang Gangberger, PhD</td>
<td>Department of Neurology, Massachusetts General Hospital, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data</td>
</tr>
<tr>
<td>Elissa Ye, MSc</td>
<td>Department of Neurology, Massachusetts General Hospital, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
<tr>
<td>Marta Fernandes, PhD</td>
<td>Department of Neurology, Massachusetts General Hospital, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; drafting/revision of the manuscript for content; medical writing for content</td>
</tr>
<tr>
<td>Luis Paixao, MD</td>
<td>Department of Neurology, Washington University School of Medicine in St. Louis, MO</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; drafting/revision of the manuscript for content; medical writing for content</td>
</tr>
<tr>
<td>Ziwei Fan, MSc</td>
<td>Department of Neurology, Massachusetts General Hospital, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
<tr>
<td>Aditya Gupta, MSc</td>
<td>Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
<tr>
<td>Manohar Ghanbar, MSc</td>
<td>Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
</tbody>
</table>

References
