Published in final edited form as:

Neurobiol Aging. 2019 February ; 74: 112–120. doi:10.1016/j.neurobiolaging.2018.10.016.

Brain Age from the Electroencephalogram of Sleep

Haoqi Sun[1], **Luis Paixao**[1], **Jefferson T. Oliva**[1,2], **Balaji Goparaju**[1], **Diego Z. Carvalho**[1], **Kicky G. van Leeuwen**[1,3], **Akeju Oluwaseun**[4], **Robert J. Thomas**[5], **Sydney S. Cash**[1], **Matt T. Bianchi**[1,*], and **M. Brandon Westover**[1,*]

[1]Department of Neurology, Massachusetts General Hospital, Boston, MA, USA [2]Bioinspired Computing Laboratory, Computer Science Department, University of São Paulo, Brazil [3]University of Twente, Enschede, the Netherlands [4]Department of Anesthesiology, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA, USA [5]Department of Medicine, Division of Pulmonary, Critical Care & Sleep, Beth Israel Deaconess Medical Center, Boston, MA, USA

Abstract

The human electroencephalogram (EEG) of sleep undergoes profound changes with age. These changes can be conceptualized as “brain age”, which can be compared to chronological age to reflect the degree of deviation from normal aging. Here, we develop an interpretable machine learning model to predict brain age based on two large sleep EEG datasets: the Massachusetts General Hospital sleep lab dataset (MGH, N = 2,532; ages 18 to 80); and the Sleep Hearth Health Study (SHHS, N = 1,974; ages 40 to 80). The model obtains a mean absolute deviation of 7.6 years between brain age (BA) and chronological age (CA) in healthy participants in the MGH dataset. As validation, a subset of SHHS containing longitudinal EEGs 5.2 years apart shows an average of 5.4 years increase in BA. Participants with significant neurological or psychiatric disease exhibit a mean excess brain age, or “brain age index” (BAI = BA - CA) of 4 years relative to healthy controls. Participants with hypertension and diabetes have a mean excess brain age of 3.5 years. The findings raise the prospect of using the sleep EEG as a potential biomarker for healthy brain aging.

Keywords

brain age; sleep; EEG; machine learning

Introduction

Human sleep undergoes robust and predictable changes with age, reflected in both overall sleep architecture and electroencephalogram (EEG) oscillations/waveforms. At the level of sleep architecture (macro-structure), older participants sleep earlier and wake earlier; have

Corresponding Author: M. Brandon Westover, M.D. Ph.D., Department of Neurology, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, mwestover@mgh.harvard.edu, Phone: +1 650-862-1154.
*These authors contributed equally to the work as co-senior authors

Disclosure
The authors declare that they have no conflict of interest.

shorter sleep duration, increased sleep fragmentation and reduced percentages of rapid eye movement (REM) sleep, as well as (at least in males) deep non-REM (NREM) sleep (Mander et al., 2017; Scullin, 2017). At the level of EEG micro-structure, older participants exhibit reduced slow waves during deep sleep (Carrier et al., 2001; Larsen et al., 1995), decreased sleep spindle amplitude, density, and duration (Purcell et al., 2017), and less phase coupling between slow oscillations and sleep spindles (Helfrich et al., 2017). However, no study has yet addressed the inverse problem: how accurately can a person's age be predicted from the sleep EEG? What factors make a person's "brain age" (BA) older or younger than their chronological age (CA)?

Brain age serves as a potential aging biomarker where the variation of brain age between individuals of the same chronological age may carry important information about the risk of cognitive impairment, neurological or psychiatric disease, or death. Various biomarkers of aging have been proposed to better predict lifespan and functional capability, ranging from molecular and cellular levels to structural level biomarkers. At the molecular and cellular level, aging-related biomarkers include leucocyte telomere length (Kruk et al., 1995), Ink4/Arf locus expression (Krishnamurthy et al., 2004), N-glycan profile (Vanhooren et al., 2010), mitochondrial DNA deletions (Eshaghian et al., 2006), and DNA-methylation status at CpG sites across the genome (Petkovich et al., 2017), among others. At the structural level, brain anatomy changes dramatically throughout life (Raz and Rodrigue, 2006). For example, cortical volume (Cole, 2017), thalamic volume (Redline et al., 2004), and white matter integrity (Mander et al., 2017), each decreases with aging. Using magnetic resonance images (MRI), chronological age can be predicted with mean absolute deviation (MAD) of 5 years in healthy participants (Franke et al., 2010). Various diseases, including Alzheimer's disease, schizophrenia, epilepsy, traumatic brain injury, bipolar disorder, major depression, cognitive impairment, diabetes mellitus, and HIV, are associated with excess brain age, i.e. older MRI brain age than chronological age (Cole, J. et al., 2017; Cole, 2017; Cole and Franke, 2017; Cole, J.H. et al., 2017; Franke et al., 2012; Franke et al., 2010).

Alongside these biomarkers, EEG-based brain age is a complement with several potential advantages: (1) EEG-based brain age could reflect functional changes rather than structural changes; (2) EEG is more participant friendly, has less costs and contraindications, and in principle could be measured by home-based devices; (3) EEG-based brain age could facilitate repeated within-participant measures to assess the effectiveness of interventions, such as medications (Roehrs and Roth, 2010) or brain stimulation (Tasali et al., 2008) that aim to preserve and/or improve brain function and life span.

Here we describe an EEG-based brain age and "brain age index" (BAI), i.e. the difference between brain age and chronological age, developed using two large sleep EEG datasets: the Massachusetts General Hospital sleep lab (MGH) dataset (Biswal et al., 2017; Sun et al., 2017), and the Sleep Hearth Health Study (SHHS) (Dean et al., 2016; Quan et al., 1997; Redline et al., 1998). We identify participants with significant neurological or psychiatric disease in the MGH dataset, and train the model only on EEGs from participants without these diseases. The effects of different EEG electrode choices are also investigated. The model is then validated on a longitudinal cohort from a subset of the SHHS dataset without neurological or cardiovascular disease. We correlate the mean deviation of brain age from

chronological age with clinical covariates and diseases. Finally, to interpret the BAI, we derive a model-based age norm, which can be compared to show how EEG features contribute to BAI.

Material and Methods

Dataset

The Partners Institutional Review Board approved retrospective analysis of the de-identified polysomnogram (PSG) dataset, acquired in MGH between 2008 and 2012, without requiring additional consent for its use in this study. The EEG signals contain 6 channels: frontal (F3-M2 and F4-M1), central (C3-M2 and C4-M1) and occipital (O1-M2 and O2-M1), each referenced to the contralateral mastoid. EEG signals are sampled at 200Hz. The signals are segmented into non-overlapping 30-second epochs. Each PSG is scored by one of seven EEG technicians according to the American Academy of Sleep Medicine (AASM) standards (AASM, 2007). Epochs are scored as one of five stages: wake (W), rapid eye movement (REM), Non-REM stage 1 (N1), Non-REM stage 2 (N2) and Non-REM stage 3 (N3). The inclusion criteria include participants (1) age between 18 to 80 years; (2) who underwent diagnostic PSG (not split night study); and (3) with a clinical diagnosis available from five years before to one year after the PSG recording from the hospital database. To remove the potentially confounding effect of imputing missing sleep stages, we exclude EEGs with missing sleep stages and investigate their effect on the brain age in the supplementary material (Figure S1).

We define a participant as “having significant neurological or psychiatric disease” if they have at least one neurological or psychiatric diagnosis in Table S1 (supplementary material) in the medical record from five years before to one year after the PSG recording. This definition is similar to those used in prior studies of MRI brain age, e.g. Brain Images of Normal Participants (BRAINS) (Job et al., 2017), Open Access Series of Imaging Studies (OASIS) (Marcus et al., 2010), (Steffener et al., 2016) and (Cole et al., 2015). Although some evidence suggests sleep apnea may be a risk factor for dementia, the relationship remains uncertain (Ding et al., 2016). We therefore do not exclude participants with sleep apnea. However, we compare BAIs for participants without sleep apnea (apnea-hypopnea index (AHI) <5) vs. those with moderate to severe apnea (AHI>15) using a t-test. We refer to participants without significant neurological or psychiatric disease as “healthy”.

In total, we identify 2,532 participants, 167 of whom have significant neurological or psychiatric diseases. Table 1 provides summary statistics for the dataset. Self-reported medication intake collected at the time of PSG recording is reported in supplemental Table S2.

We also use a subset of the SHHS dataset (Dean et al., 2016; Quan et al., 1997; Redline et al., 1998), which contains repeated EEGs from the same participant in two visits about 5 years apart, making it possible to evaluate the longitudinal reliability of our model at the population level. The participant inclusion criteria are: (1) having EEGs from both visits; (2) chronological age between 40 to 80 years at both visits (minimum age in SHHS is 40); (3) having EEG and sleep stage scoring of high quality according to SHHS specifications (Table

S3); and (4) having no neurological or cardiovascular disease (Table S3). We also exclude EEGs with missing sleep stages. As a result, 987 EEGs from SHHS visit 1 and 987 paired EEGs from visit 2 are used. The average difference in the chronological age between the two visits is 5.2 years. Unlike the MGH dataset, the SHHS dataset uses two EEG channels (C3-M2 and C4-M1) and is scored according to R&K standard (Rechtschaffen, 1968). To make scoring consistent across datasets, we combine S3 and S4 in SHHS to match the scoring of N3 in the MGH data. Summary statistics for the SHHS dataset are shown in Table 2.

EEG Preprocessing and Artifact Removal

EEG signals are notch-filtered at 60Hz to reduce line noise, and bandpass filtered from 0.5Hz to 20Hz to reduce myogenic artifacts. For 30s-epochs, those with absolute amplitude larger than 500uV are removed to minimize movement artifacts. Epochs containing flat EEG for more than 2 seconds are also removed. We also exclude EEGs contaminated by electrocardiogram (ECG), indicated by 1Hz harmonic in the EEG spectrogram. To reduce inter-participant variance, the amplitude of each EEG channel is normalized to have zero median and unit interquartile range across the whole night. The total amount of data removed by these preprocessing procedures is 7% in the MGH dataset and 9% in the SHHS dataset.

Feature Extraction

For age prediction, we extract features from sleep EEG used in sleep staging in our previous work (Sun et al., 2017). We extract 102 features from each 30-second epoch covering both time and frequency domains, as summarized in Table S4. For each EEG recording, we average the features in each of the 5 sleep stages over time, yielding 102 × 5 = 510 features per EEG. The features are log-transformed to render feature distributions approximately Gaussian. Subsequently, these features are z-transformed to have zero mean and unit standard deviation in the training set. The same z-transformation is applied to the testing set.

Brain Age Prediction

The model minimizes an objective function $J(w, b)$ with two terms: (1) mean squared prediction error; and (2) magnitude of the covariance between CA and BAI:

$$J(w,b) = \left\langle BAI^2 \right\rangle + \lambda \mid Cov(CA, BAI) \mid$$

where $BA_i = \text{softplus}(w^T x_i + b) = \ln[1 + e^{w^T x_I + b}]$, i.e. a linear combination of EEG features followed by a softplus function to ensure positivity of BA;
$\left\langle BAI^2 \right\rangle = \sum_{i=1}^{N} (BAI_i)^2 / N = \sum_{i=1}^{N} (BA_i - CA_i)^2 / N$ is the mean squared prediction error; and $Cov(CA, BAI) = \sum_{i=1}^{N} \mid (CA_i - \left\langle CA \right\rangle)(BAI_i - \left\langle BAI \right\rangle) \mid$ is the average absolute correlation (covariance) between the chronological age CA_i and brain age index BAI_i. Minimizing the first term $\langle BAI^2 \rangle$ encourages the model to produce predictions BA that are accurate (close to CA). Minimizing the second component $cov(CA, BAI)$ encourages deviations of BA from CA to be uncorrelated with CA. The second component is weighted by a hyperparameter λ which tradeoff between the two components.

To determine the optimal hyperparameter λ, we randomly select 300 EEGs from the training set to serve as internal validation data. We perform grid search on λ ranging from [0, 1, 5, 10]. We find the hyperparameter with the highest performance on the internal validation set. The performance metric for each hyperparameter is evaluated using corr(*CA, BA*) – |corr(*CA, BAI*)|, where corr(·) denotes Pearson's correlation. A larger corr(*CA, BA*) and a smaller |corr(*CA, BAI*)| indicate better model performance. Here the metric is correlation instead of covariance or prediction error, since correlation is normalized and comparable among different hyperparameter values. Once the optimal hyperparameter with the best validation performance is determined, the validation set is combined with the rest of the training set, and the model is re-trained using the optimal hyperparameter.

We maintain strict separation of training and testing participants to achieve statistically unbiased estimates of model's performance. The reported results are based on the testing set, unless stated.

Results

EEG-Based Brain Age Prediction

The MGH dataset (N = 2,532) was partitioned into a healthy training set (N = 1,343) used to train the model, and a testing set (N = 1,189) to evaluate model performance. In the testing set, 167 had significant neurological or psychiatric disease (Table S1); the remaining 1,022 were healthy. A comparison between these two groups is presented in subsection "Correlation between Disease and Brain Age Index".

We first compared age prediction performance using the sleep macro-structure features (Mander et al., 2017; Redline et al., 2004) (30 features, Table S5) vs. using the 510 sleep micro-structure EEG features, on the testing set. As shown in Figure 1, the mean absolute deviation (MAD) of BA from CA using macro-structure features was much higher (23.3 years) than that using EEG features (7.8 years), and the correlation was weaker (0.46, 95% CI 0.42 – 0.50) than that using EEG features (0.82, 95% CI 0.80 – 0.84). The results suggest that the effects of age on brain activity are more consistently reflected in sleep EEG (micro-structure) than in sleep stage composition (macro-structure). The progression of EEG features with age is further illustrated by a two-dimensional visualization of the feature space in Figure S2.

In Figure 2, we show spectrograms of sleep EEGs from 6 typical individuals, arranged as a "confusion matrix", where rows show participants with CA in young (18 – 30 years), intermediate (40 – 50 years), and older groups (60 – 80 years) from top to bottom, and columns with BA in three groups from left to right. The participants positioned along the diagonal line are examples of participants with matched CA and BA. As an example, the middle right panel indicated by * shows a 48-year-old female diagnosed with hypertension, seizure disorder, kidney stone and arteriosclerotic heart disease, whose predicted brain age was 65.8 years (BAI = BA – CA = 65.8 – 48 = 17.8 years). One visually apparent abnormality is the relatively weaker sleep spindle density during N2 compared to healthy participants of similar chronological age (middle left and center panels).

Effect of EEG Electrode Placement

Most home-based EEG devices use fewer electrodes than lab-based PSG. We therefore explored how EEG-based age prediction depends on the number and location of EEG electrodes. Prediction performance on the 1,022 healthy testing participants using different subsets of EEG electrodes is shown in Table 3. Using all six electrodes (2 frontal, 2 central 2 occipital) provided the lowest prediction error (7.6 years) and the highest correlation with CA (0.83) (Kruskal-Wallis p-value < 0.01). If limited to recording EEG using two electrodes only, the frontal electrodes provided the best performance with non-significant difference compared to all six electrodes (p-value = 0.20); the central and occipital electrodes led to significantly reduced performance (p-value < 0.01). Therefore it is suggested to use frontal EEG electrodes if required to use fewer EEG electrodes.

Longitudinal Reliability using Two Central EEG Channels

To assess the longitudinal reliability of the brain age model at the population level, we used a subset of the SHHS dataset where each participant underwent two study visits, referred as SHHS1 and SHHS2. The dataset contains 987 adults for a total of 1,974 nights of EEG recorded from C3-M2 and C4-M1 only. The average increase of chronological age between the two visits was 5.2 years (standard deviation, SD 0.5 year) (Table 2). We trained the model in two ways.

First, we trained the model on 752 participants with paired EEGs (1,504 EEGs) from both visits and tested on the held out 235 participants with paired EEGs (470 EEGs) in both visits. Testing results are shown in Figure 3A and 3B. The MAD was 8.1 years on SHHS1 and 7.8 years on SHHS2. The Pearson's correlation was 0.66 (95% CI 0.58 – 0.73) on SHHS1 and 0.67 (95% CI 0.59 – 0.73) on SHHS2. The average difference in predicted BA between the two visits, $\Delta BA = BA2 - BA1$, was 5.4 years (SD 9.3 years) (Figure 3C); hence the magnitude of the observed difference between the progression of mean chronological age $\Delta CA = CA2 - CA1$ and mean brain age ΔBA was $|\Delta CA - \Delta BA| = |5.2 - 5.4| = 0.2$ years. A t-test revealed that this difference was not statistically significant (p-value = 0.7).

Second, we investigated the effect of the training population. To do this, we trained the model on the 2,365 EEGs from healthy participants in MGH dataset using the C3-M2 and C4-M1 leads only (to be consistent with SHHS EEGs). We then predicted BA on the 1,974 paired EEGs in SHHS as the testing set. Results on the testing set are shown in Figure 3D and 3E. The MAD was 8.2 years on SHHS1 and 9.4 years on SHHS2. The Pearson's correlation was 0.61 (95% CI 0.57 – 0.65) on SHHS1 and 0.56 (95% CI 0.52 – 0.60) on SHHS2. The average difference in predicted BA between the two visits ΔBA was 4.3 years (SD 9.6 years) (Figure 3F). In this case the difference between mean progression of chronological age was statistically significant (p-value <0.01, t-test), indicating that different populations (datasets) did introduce some systematic error. However, the magnitude of this difference was relatively small, $|\Delta CA - \Delta BA| = |5.2 - 4.3| = 0.9$ years.

The results suggest that, at the population level, the model longitudinally tracked brain aging when trained on either the SHHS or MGH cohorts, with a marginal error of 0.2 to 0.9 years. We note that these results were obtained using a restricted set of EEG electrodes (C3-M2

and C4-M1). Given limited channels in the SHHS EEG data, we were not able to investigate whether even higher accuracy is possible using a full set of channels.

Correlation between Covariates and Brain Age Index

We collected various covariates from the MGH dataset, including sleep macro-structure (i.e. sleep stages), the Epworth Sleepiness Scale, and measures of sleep disruption, and then obtained the Spearman's correlation with BAI. For healthy participants, the covariates with statistically significant ($p < 0.05$) Spearman's correlation are shown in ascending order in Table 4. The full list of all covariates is shown in Table S6 in the supplementary material. The amount of deep sleep (N3) showed the most negative correlation with BAI, although the correlation is weak. Wake time after sleep onset also showed weak but significant positive correlation with BAI.

One noteworthy lack of association is that BAI is statistically independent of AHI (apnea-hypopnea index; number of apnea events per hour of sleep, used to quantify severity of sleep apnea). A scatter plot between AHI and CA/BA/BAI shown in Figure S3 in the supplementary material reveals a correlation between AHI and CA and BA, but not BAI. The histogram of BAI for AHI<5 and AHI>15 are shown in Figure S4, showing no significant difference between the two groups (t-test, $p = 0.70$).

The weak correlations in Table 4 indicate that BAI was relatively independent of these covariates. Therefore BAI can be interpreted as a relatively orthogonal metric for measuring sleep and brain health. In other words, BAI was robust in participants with different characteristics.

Correlation between Disease and Brain Age Index

We next investigated some population-level health correlates of "excess brain age", BAI = BA – CA > 0. In Materials and Methods we describe criteria for significant neurological or psychiatric disease (Table S1), which we hypothesized to have older BA than CA. This hypothesis was verified in Figure 4A. BAI for the healthy group (2,365 participants) was around 0 years, while BAI for the group with significant neurological and psychiatric disease (167 participants) increased up to 4 years with a significant difference (t-test p-value < 0.01). Therefore, at the population level, participants with significant neurological or psychiatric disease had on average a brain age 4 years older than their chronological age. Further, in the healthy group, the MAD was 7.6 years and correlation between CA and BA was 0.83; while in the group with significant neurological and psychiatric disease, the MAD was larger at 9.5 years and correlation smaller at 0.74.

Though often not associated with any formal neurological or psychiatric or psychiatric diagnosis, there is biological plausibility and accumulating evidence for the idea that hypertension and diabetes mellitus accelerate brain aging (Gorelick et al., 2017; Gorelick et al., 2011). We therefore explored whether EEG brain age is sensitive to hypertension and diabetes. These participants were identified using the diagnosis from five years before to one year after the PSG recording, as well as the self-reported medication intake form collected at the time of PSG recording. As compared in Figure 4B, the control group was defined as participants taking no medications, having no diagnoses of hypertension or diabetes, and no

significant neurological or psychiatric disease. The control group had a mean excess brain age of BAI = −1.1 years (n = 472). The "contrast" group was defined as participants with both hypertension and diabetes, which had BAI = 2.4 years (n = 137). The attributable excess brain age was 3.5 years (p = 0.0002).

Model-Derived Age Norm

The parametric formulation of our brain age model allows studying individual differences at the level of EEG features. To do this, we first derived an "age norm", i.e. typical EEG feature values at different CAs with matched BAs. Specifically, the age norm of age x was obtained by averaging EEG features from all healthy participants who had both CA and BA within [x-5, x+5] years.

The age norm showed that when following a normal aging process, the features most negatively correlated with age were delta band (< 4Hz) powers in stage N3 sleep in the occipital and central areas; and features that most positively correlate with age were theta band (4 – 8Hz) powers in stage N1 sleep in frontal, occipital and central areas. The details are described in Table S6.

Using this model-derived age norm, we could examine reasons for deviation from CA quantitatively. For example, Figure 5A shows the EEG spectrogram of a 29-year-old female with multiple medical problems, including diabetes mellitus, obesity, smoking, a cerebrovascular accident, congestive heart failure and acute renal failure. Her brain activity resembled that of a much older adult. Her brain age was 63 years. As shown in Figure 5B, when compared to the age norm, her sleep EEG exhibited (1) less delta to theta ratio during deep N3 sleep; and (2) less theta and alpha band bursts, i.e. more continual theta and alpha, during N2 sleep.

Note that our results only showed that BAI reflects age and pathology at the population level. Nevertheless this example suggested the possibility that BAI may, with further studies, serve as a personalized biomarker of brain health. Other examples are shown in Figure S5 – S7 in the supplementary material.

Discussion

The concept of differential aging – the idea that a 50-year-old can "have the heart of a 20-year-old" or that the lungs of a 30-year-old smoker "work like they're 80" – is useful clinically, and is gaining scientific traction. Similarly, while cognitive decline is a "normal" part of aging, some individuals clearly age better than others. In this work we have developed a machine learning model that leverages statistical changes across the adult lifespan in the pattern of brain activity during sleep. We call the predicted age "brain age", or BA. Model predictions are highly correlated with chronological age (CA) (r = 0.83, MAD = 7.6 years) for healthy participants, and at the population level the model accurately tracks advancing age. More interestingly, this study presents preliminary evidence that excess BA (EEG brain age in excess than chronological age) reflects underlying brain pathology. In particular, our findings show that significant neurological or psychiatric disease accelerate brain aging, as do hypertension and diabetes. These findings suggest that features of brain

activity during sleep, reflected in EEG, can be harnessed using machine learning to provide an easily accessible and low-cost biomarker of brain health.

One strength of our study is the use of large datasets: 2,532 sleep EEGs from the MGH Sleep Lab, and 1,974 sleep EEGs from the SHHS dataset (Dean et al., 2016; Quan et al., 1997; Redline et al., 1998). The number of EEGs involved in this study is large among relevant “brain age” studies (Cole and Franke, 2017). The large size of the datasets helps to ensure the statistical power as well and to minimize selection bias. A large training set helps ensure that the trained model does not overfit to a particular dataset, improving the ability to generalize when applied beyond the training set. A large testing set allows accurate statistical measurement of how accurately the model performs.

Another strength of our study is the use a parametric model, which improves model interpretation by inspecting each EEG feature and comparing to the age norm for each feature explicitly. In contrast, nonparametric or semi-parametric models such as kernel methods and Gaussian process models (Franke et al., 2010) estimate age based on the similarity to the training examples, where the contribution of each feature is involved implicitly. An example of interpretability is demonstrated in Figure 5.

The deviation of EEG-based BA from CA arises from multiple sources, which can be divided into technical and physiological sources. Potential technical sources include (1) first-night effect, where a participant’s quality of sleep is impaired by the effect of the environment and polysomnographic recording; (2) constraints of the current model, which assumes brain age is a linear combination (though approximately due to the softplus function to ensure positivity) of sleep EEG features averaged separately for each sleep stage over the night; and (3) imputation of missing sleep stages based on Figure S1. Neurophysiological sources, our central interest, might arise from (1) night-to-night variability; (2) underlying diagnosed or undiagnosed neurological or psychiatric diseases. For example, sleep pathology enhances amyloidogenesis, implicated in the development of Alzheimer’s dementia (Ju et al., 2014). Moreover, normal glymphatic flow, required for clearance of amyloid and other metabolic wastes that accumulate during wakefulness, requires undisturbed sleep (Benveniste et al., 2017; Hladky and Barrand, 2017); (3) general medical health, including diabetes and cerebrovascular disease; (4) genetically-determined inter-individual differences in the EEG; and (5) exposure to environmental insults. Further studies are needed to clarify the relative contributions of the various types of neurophysiological sources, and to measure the association between EEG-based brain age and various outcomes, such as cognitive performance (Steffener et al., 2016), intelligence (Ujma et al., 2017) and survival (Cole, J. et al., 2017). Additional unobserved sources, such as diet, exercise and diseases, might also further explain the variance. An additional important question for future investigation is whether interventions can change the brain aging process and life span, and whether improvements in brain health can be effectively tracked using the sleep EEG.

Our study has important limitations. (1) Although we have demonstrated that the change of brain age after 5 years of follow up is close to 5 years at population level, our dataset does not allow studying variability in brain age at the individual level. Understanding the factors

governing changes in brain age over time and test-retest reliability at the individual level are important topics for further research. (2) Although we attempted to exclude participants with significant neurological or psychiatric disease (criteria in Table S1 and Table S3), it is possible that both the MGH and SHHS datasets still include a variety of health conditions that could affect brain health. Obtaining a large, “truly healthy” cohort is an important future direction that will require a prospectively controlled study. (3) We only include ages 18 to 80. The distribution of ages in the dataset is not uniform, where most of the participants are at middle age. This could lead to less accurate prediction of brain age for older and younger participants. (4) Our model formulation is limited. Advanced algorithms such as convolutional neural networks, trained end-to-end from raw EEG data rather than on hand-engineered physiologically-based features as in our model, might predict age and capture biologically important variability more accurately (Biswal et al., 2017; Cole, J.H. et al., 2017). In addition, whereas we average features within the same sleep stage, representations learned by recurrent neural network models might provide a more powerful summary of the overnight EEG that utilizes more information (e.g. fragmentation of sleep stages due to high AHI) and more accurately reflects brain age. (5) Finally, conclusions from our present results are primarily limited to the population level. Although selected cases raise the possibility that BAI is meaningful for individuals. Our data provide no way to definitively determine the sources of variation at the level of individuals (e.g. biological effects related to brain health vs. random night-to-night variation). This is of course a classic issue in the clinical/biological prediction literature, where evidence for a correlation at a population level does not necessarily translate into a usable algorithm for individual use. Nevertheless, our findings form the basis for further studies and refinements. To reach individual level, first, further studies with more granular detail about individuals’ medical conditions (e.g. medication doses), neurological function (e.g. neuropsychiatric test scores) and brain structure (e.g. cortical thickness, total brain volume, white matter integrity, etc.) will be required to better understand the biological mechanism of deviations between EEG-based brain age and chronological age.

Conclusions

Using a machine learning model, a measure of brain age can be inferred from the pattern of brain activity during sleep. Our data suggest that neurological, psychiatric, and medical illnesses that adversely affect brain health result in the model predicting an older brain age. Using participants with matched chronological age and brain age, we developed an age norm to provide a direct interpretation of an individual’s deviation from normal aging in terms of EEG features. In summary, the EEG-based brain age serves as a potential biomarker which sets the stage for future EEG-based studies of brain health.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

MBW reports grants from NIH-NINDS (NIH-NINDS 1K23NS090900, 1R01NS102190, 1R01NS102574, 1R01NS107291). MTB has received funding from the Center for Integration of Medicine and Innovative

Technology, the Milton Family Foundation, the MGH-MIT Grand Challenge, and the American Sleep Medicine Foundation, and the Department of Neurology. MTB has a patent pending on a home sleep monitoring device, has research agreements with MC10 and Insomnisolv, and consulting agreements with McKesson, International Flavors and Fragrances, and Apple Inc., serves as a medical monitor for Pfizer, and has provided expert testimony in sleep medicine. RJT reports the following: (1) Patent, license and royalties from MyCardio, LLC, for an ECG-based method to phenotype sleep quality and sleep apnea; (2) Grant support, license and intellectual property (patent submitted) from DeVilbiss Healthcare; (3) GLG consulting for general sleep medicine; (4) Intellectual Property (patent) for a device using CO2 for central / complex sleep apnea. This is not an industry supported study, and none of these entities had any role in the study.

References

AASM, 2007 The AASM manual for the scoring of sleep and associated events: Rules. Terminology and Technical Specifications. Westchester: AASM.

Benveniste H, Lee H, Volkow ND, 2017 The glymphatic pathway: waste removal from the CNS via cerebrospinal fluid transport. The Neuroscientist 23(5), 454–465. [PubMed: 28466758]

Biswal S, Kulas J, Sun H, Goparaju B, Westover MB, Bianchi MT, Sun J, 2017 SLEEPNET: Automated Sleep Staging System via Deep Learning. arXiv preprint arXiv:1707.08262.

Carrier J, Land S, Buysse DJ, Kupfer DJ, Monk TH, 2001 The effects of age and gender on sleep EEG power spectral density in the middle years of life (ages 20–60 years old). Psychophysiology 38(2), 232–242. [PubMed: 11347869]

Cole J, Ritchie S, Bastin M, Hernández MV, Maniega SM, Royle N, Corley J, Pattie A, Harris S, Zhang Q, 2017 Brain age predicts mortality. Molecular Psychiatry.

Cole JH, 2017 Neuroimaging-derived brain-age: an ageing biomarker? Aging (Albany NY) 9(8), 1861. [PubMed: 28858849]

Cole JH, Franke K, 2017 Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. Trends in Neurosciences.

Cole JH, Leech R, Sharp DJ, Initiative A.s.D.N., 2015 Prediction of brain age suggests accelerated atrophy after traumatic brain injury. Annals of neurology 77(4), 571–581. [PubMed: 25623048]

Cole JH, Poudel RP, Tsagkrasoulis D, Caan MW, Steves C, Spector TD, Montana G, 2017 Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. NeuroImage 163, 115–124. [PubMed: 28765056]

Dean DA, Goldberger AL, Mueller R, Kim M, Rueschman M, Mobley D, Sahoo SS, Jayapandian CP, Cui L, Morrical MG, 2016 Scaling up scientific discovery in sleep medicine: the National Sleep Research Resource. Sleep 39(5), 1151–1164. [PubMed: 27070134]

Ding X, Kryscio RJ, Turner J, Jicha GA, Cooper G, Caban-Holt A, Schmitt FA, Abner EL, 2016 Self-Reported Sleep Apnea and Dementia Risk: Findings from the Prevention of Alzheimer's Disease with Vitamin E and Selenium Trial. Journal of the American Geriatrics Society 64(12), 2472–2478. [PubMed: 27801937]

Eshaghian A, Vleugels RA, Canter JA, McDonald MA, Stasko T, Sligh JE, 2006 Mitochondrial DNA deletions serve as biomarkers of aging in the skin, but are typically absent in nonmelanoma skin cancers. Journal of investigative dermatology 126(2), 336–344. [PubMed: 16374452]

Franke K, Luders E, May A, Wilke M, Gaser C, 2012 Brain maturation: predicting individual BrainAGE in children and adolescents using structural MRI. Neuroimage 63(3), 1305–1312. [PubMed: 22902922]

Franke K, Ziegler G, Klöppel S, Gaser C, Initiative A.s.D.N., 2010 Estimating the age of healthy subjects from T 1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. Neuroimage 50(3), 883–892. [PubMed: 20070949]

Gorelick PB, Furie KL, Iadecola C, Smith EE, Waddy SP, Lloyd-Jones DM, Bae H-J, Bauman MA, Dichgans M, Duncan PW, 2017 Defining optimal brain health in adults: a presidential advisory from the American Heart Association/American Stroke Association. Stroke 48(10), e284–e303. [PubMed: 28883125]

Gorelick PB, Scuteri A, Black SE, DeCarli C, Greenberg SM, Iadecola C, Launer LJ, Laurent S, Lopez OL, Nyenhuis D, 2011 Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 42(9), 2672–2713. [PubMed: 21778438]

Helfrich RF, Mander BA, Jagust WJ, Knight RT, Walker MP, 2017 Old brains come uncoupled in sleep: slow wave-spindle synchrony, brain atrophy, and forgetting. Neuron.

Hladky SB, Barrand MA, 2017 Metabolite Clearance During Wakefulness and Sleep.

Job DE, Dickie DA, Rodriguez D, Robson A, Danso S, Pernet C, Bastin ME, Boardman JP, Murray AD, Ahearn T, 2017 A brain imaging repository of normal structural MRI across the life course: Brain Images of Normal Subjects (BRAINS). NeuroImage 144, 299–304. [PubMed: 26794641]

Ju Y-ES, Lucey BP, Holtzman DM, 2014 Sleep and Alzheimer disease pathology—a bidirectional relationship. Nature reviews Neurology 10(2), 115. [PubMed: 24366271]

Krishnamurthy J, Torrice C, Ramsey MR, Kovalev GI, Al-Regaiey K, Su L, Sharpless NE, 2004 Ink4a/Arf expression is a biomarker of aging. Journal of Clinical Investigation 114(9), 1299. [PubMed: 15520862]

Kruk PA, Rampino NJ, Bohr VA, 1995 DNA damage and repair in telomeres: relation to aging. Proceedings of the National Academy of Sciences 92(1), 258–262.

Larsen LH, Moe KE, Vitiello MV, Prinz PN, 1995 Age trends in the sleep EEG of healthy older men and women. Journal of sleep research 4(3), 160–172.

Mander BA, Winer JR, Walker MP, 2017 Sleep and Human Aging. Neuron 94(1), 19–36. [PubMed: 28384471]

Marcus DS, Fotenos AF, Csernansky JG, Morris JC, Buckner RL, 2010 Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. Journal of cognitive neuroscience 22(12), 2677–2684. [PubMed: 19929323]

Petkovich DA, Podolskiy DI, Lobanov AV, Lee S-G, Miller RA, Gladyshev VN, 2017 Using DNA methylation profiling to evaluate biological age and longevity interventions. Cell metabolism 25(4), 954–960. e956. [PubMed: 28380383]

Purcell S, Manoach D, Demanuele C, Cade B, Mariani S, Cox R, Panagiotaropoulou G, Saxena R, Pan J, Smoller J, 2017 Characterizing sleep spindles in 11,630 individuals from the National Sleep Research Resource. Nature Communications 8, ncomms15930.

Quan SF, Howard BV, Iber C, Kiley JP, Nieto FJ, O'connor GT, Rapoport DM, Redline S, Robbins J, Samet JM, 1997 The sleep heart health study: design, rationale, and methods. Sleep 20(12), 1077–1085. [PubMed: 9493915]

Raz N, Rodrigue KM, 2006 Differential aging of the brain: patterns, cognitive correlates and modifiers. Neuroscience & Biobehavioral Reviews 30(6), 730–748. [PubMed: 16919333]

Rechtschaffen A, 1968 A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Public health service.

Redline S, Kirchner HL, Quan SF, Gottlieb DJ, Kapur V, Newman A, 2004 The effects of age, sex, ethnicity, and sleep-disordered breathing on sleep architecture. Archives of Internal Medicine 164(4), 406–418. [PubMed: 14980992]

Redline S, Sanders MH, Lind BK, Quan SF, Iber C, Gottlieb DJ, Bonekat WH, Rapoport DM, Smith PL, Kiley JP, 1998 Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep 21(7), 759–767. [PubMed: 11300121]

Roehrs T, Roth T, 2010 Drug-related sleep stage changes: functional significance and clinical relevance. Sleep medicine clinics 5(4), 559–570. [PubMed: 21344068]

Scullin MK, 2017 Do Older Adults Need Sleep? A Review of Neuroimaging, Sleep, and Aging Studies. Current sleep medicine reports 3(3), 204–214. [PubMed: 29226069]

Steffener J, Habeck C, O'Shea D, Razlighi Q, Bherer L, Stern Y, 2016 Differences between chronological and brain age are related to education and self-reported physical activity. Neurobiology of aging 40, 138–144. [PubMed: 26973113]

Sun H, Jia J, Goparaju B, Huang G-B, Sourina O, Bianchi MT, Westover MB, 2017 Large-Scale Automated Sleep Staging. Sleep 40(10).

Tasali E, Leproult R, Ehrmann DA, Van Cauter E, 2008 Slow-wave sleep and the risk of type 2 diabetes in humans. Proceedings of the National Academy of Sciences 105(3), 1044–1049.

Ujma PP, Konrad BN, Gombos F, Simor P, Pótári A, Genzel L, Pawlowski M, Steiger A, Bódizs R, Dresler M, 2017 The sleep EEG spectrum is a sexually dimorphic marker of general intelligence. Scientific reports 7(1), 18070. [PubMed: 29273758]

Vanhooren V, Dewaele S, Libert C, Engelborghs S, De Deyn PP, Toussaint O, Debacq-Chainiaux F, Poulain M, Glupczynski Y, Franceschi C, 2010 Serum N-glycan profile shift during human ageing. Experimental gerontology 45(10), 738–743. [PubMed: 20801208]

Highlights

- Brain age based on two large sleep EEG datasets (N_{MGH}=2,532 and N_{SHHS} = 1,974).
- Brain age tracks chronological age at the population level.
- Chronological age is predicted with mean absolute deviation of 7.6 years in healthy participants.
- Significant neurological and psychiatric diseases lead to a 4-year increase in brain age.
- Interpretation using age norms derived from matched brain age and chorological age.

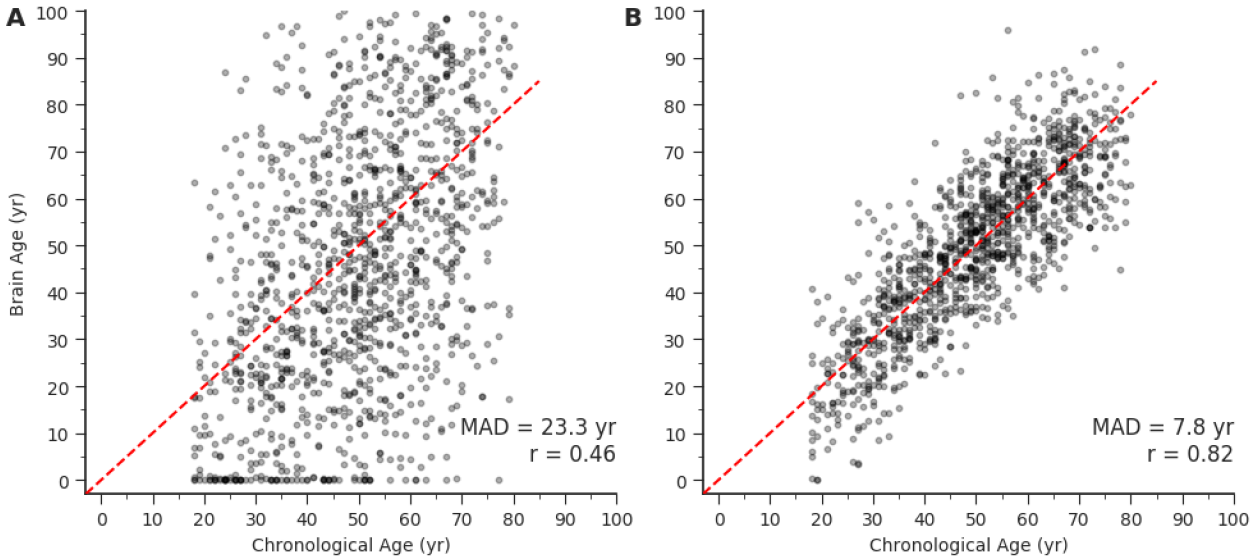

Figure 1.

Scatter plot of predicted BA vs. CA using **(A)** sleep EEG macro-structure features (described in Table S5) and **(B)** sleep micro-structure features (described in Table S4). The red dashed diagonal line is the identity line where BA is equal to CA. Both plots are obtained from the same testing participants, including both healthy and with significant neurological or psychiatric disease.

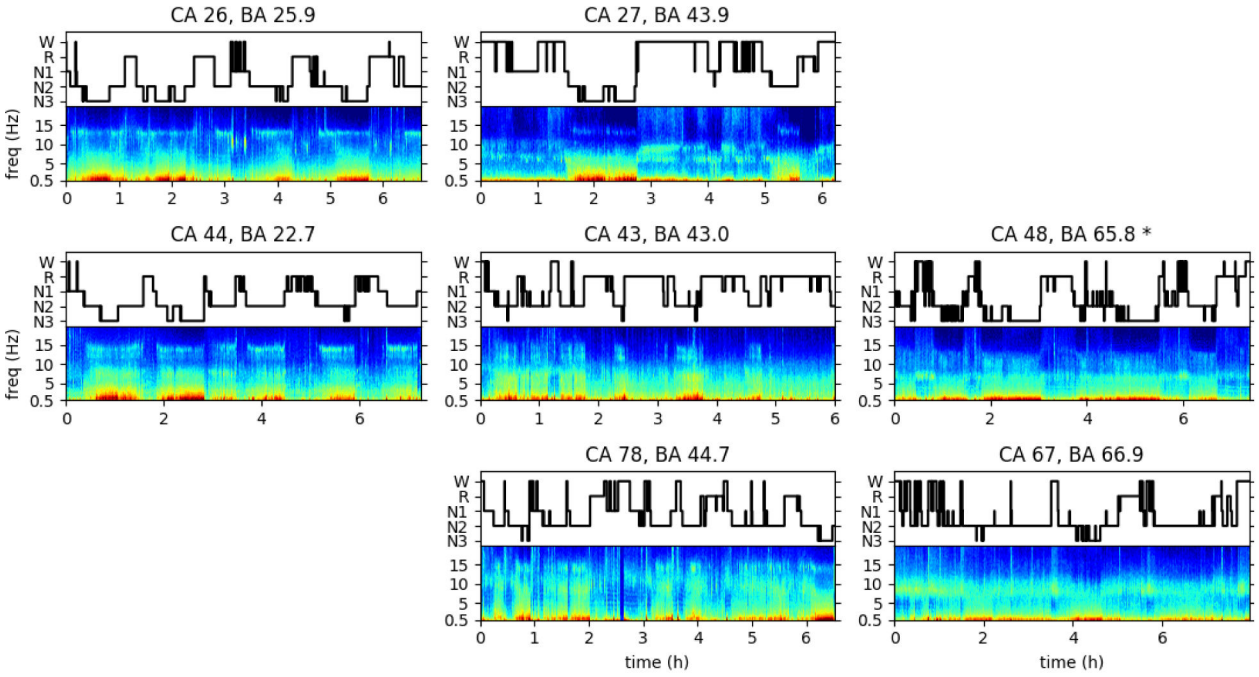

Figure 2.

Cases in which BA matches (diagonal) or is younger or older than CA (off diagonal). Each panel consists of a hypnogram and the corresponding spectrogram. Spectrograms are calculated as the average across the 6 EEG channels. The horizontal axis is time in hours. The case indicated by * is described in the main text.

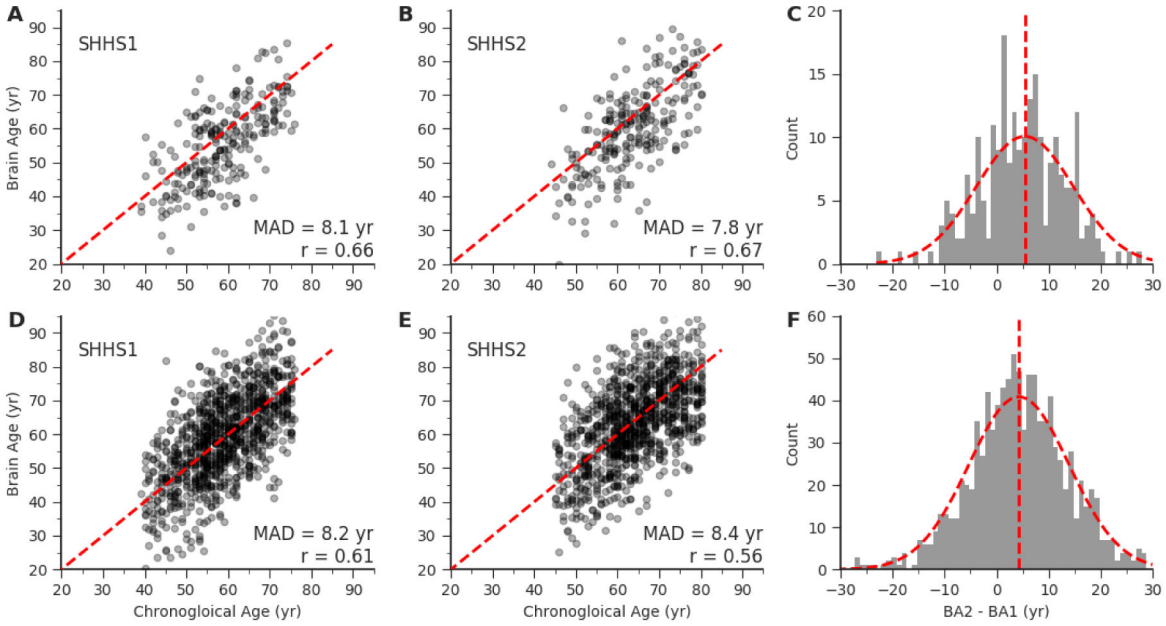

Figure 3.

(A, B) Chronological age (CA) vs. Brain age (BA) from the 470 testing EEGs in SHHS when trained on the other part of SHHS. The red dash line indicates identity. **(D, E)** CA vs. BA from the 1,974 testing EEGs in SHHS when trained on the healthy participants in MGH dataset. **(C, F)** Histogram of BA differences between two SHHS visits, showing tracking of CA at the population level. The mean difference in predicted BA was 5.4 years in (C); and 4.3 years in (F). The dashed red line is the Gaussian fit to the histogram.

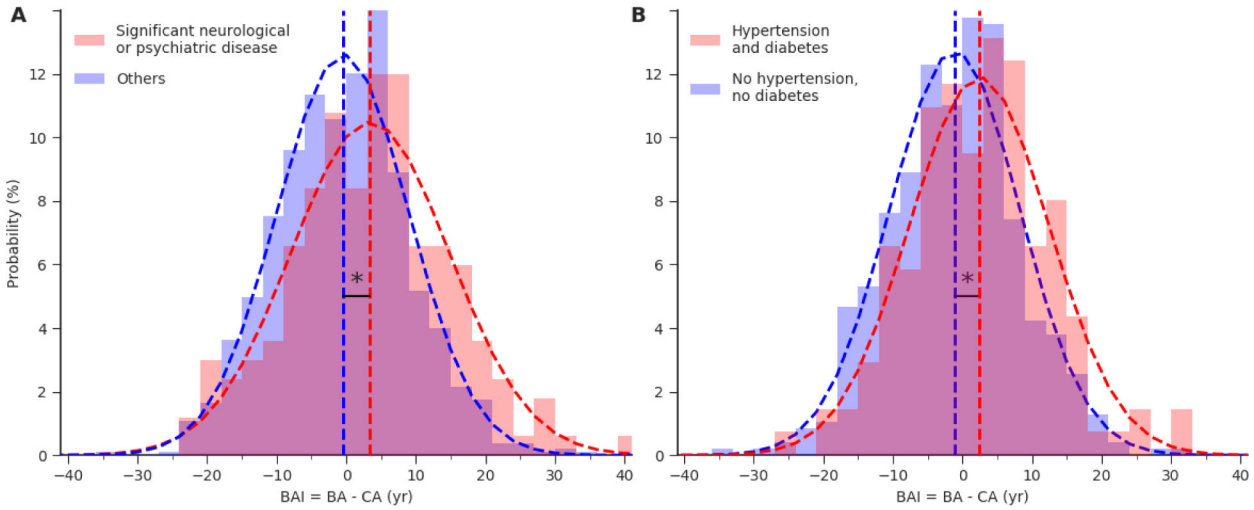

Figure 4.

(A) Histograms of BAI s for participants with (red) and without (blue) significant neurological or psychiatric disease defined in Table S1. The dashed lines show the Gaussian fit to these distributions. The difference of group means is 4 years. **(B)** Histograms of BAIs for participants with (red) and without (blue) hypertension and diabetes. The difference of group means is 3.5 years.

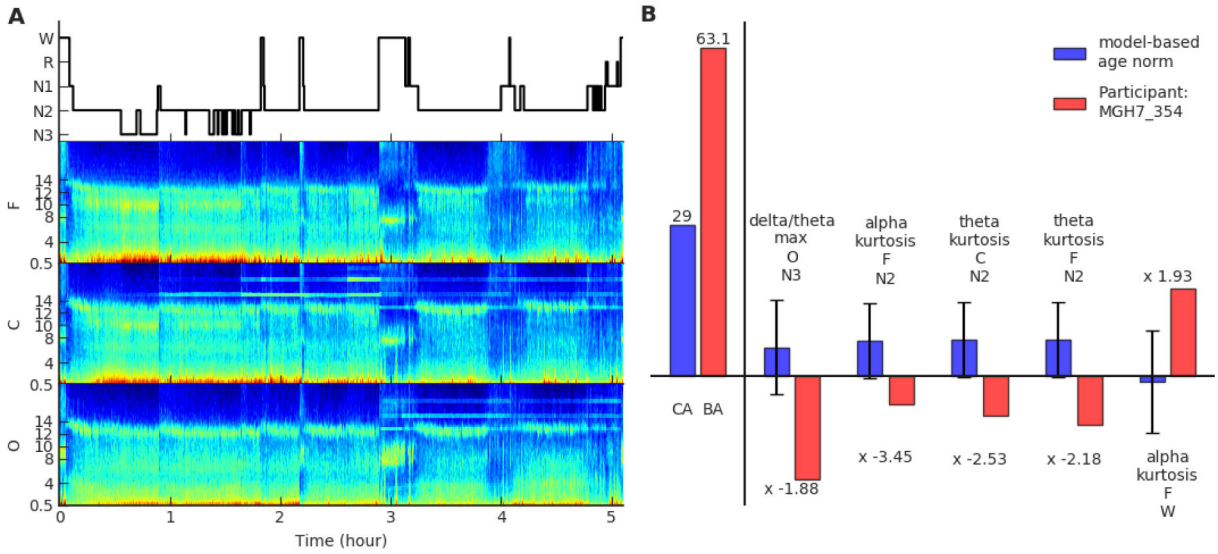

Figure 5.

(A) The hypnogram and EEG spectrogram from frontal (F), central (C) and occipital (O) channels. **(B)** The left two bars show the chronological age and brain age for this participant. The right parts show the top five features that contribute most positively to the older brain age. The blue bar is the feature value based on the model-based age norm at age 29, with the error bar indicating the standard deviation. The red bar is the feature value for this participant. The number at the top/bottom indicates the model weight associated with this feature.

Table 1.

MGH Sleep Dataset Summary.

Characteristics	Value
Number of EEGs	2,532 (including 167 with significant neurological or psychiatric disease *)
Age (year), median (IQR)	50 (38 – 61)
Gender	Female 1,266 (50%), Male 1,266 (50%)
BMI (kg/m^2), median (IQR)	29 (25 – 34)
Overall AHI (per hour), median (IQR)	5 (2 – 12)
Normal (<5)	1,232 (48.7%)
Mild sleep apnea (5<=AHI<15)	813 (32.1%)
Moderate sleep apnea (15<=AHI<30)	373 (14.7%)
Severe sleep apnea (AHI>=30)	114 (4.5%)

IQR: interquartile range; BMI: body mass index; AHI: apnea-hypopnea index (# apnea events per hour of sleep) at 4% oxygen desaturation for hypopnea.

* See Table S1 in the supplementary material.

Table 2.

Sleep Heart Health Study Dataset Summary.

Characteristics	**SHHS visit 1 (SHHS1)**	**SHHS visit 2 (SHHS2)**
Number of EEGs	987	987 (same participants as SHHS1)
Age (year), median (IQR)	58 (53 – 66)	64 (58 – 71)
Gender	Female 609 (62%), Male 378 (38%)	
BMI (kg/m^2), median (IQR)	27 (25 – 30)	28 (25 – 31)
Overall AHI (per hour), median (IQR)	3 (1 – 7)	5 (2 — 11)

Table 3.

Brain age using subsets of EEG electrodes in healthy testing participants.

Electrode	Mean absolute deviation (years)	Pearson's correlation
Frontal (F3 and F4)	9.6	0.80 (0.77 – 0.82)
Central (C3 and C4)	12.2	0.78 (0.75 – 0.80)
Occipital (O1 and O2)	10.0	0.76 (0.73 – 0.78)
All electrodes	7.6	0.83 (0.81 – 0.85)

Table 4.

Covariates that have significant correlation with BAI (BA - CA).

Covariate	Spearman's correlation
N3 time	−0.11
N3 percentage	−0.087
Sleep efficiency	−0.082
Total sleep time	−0.076
N2 time	−0.065
Respiratory Disturbance Index	0.043
N1 percentage	0.058
Wake Time After Sleep Onset	0.081