Artificial Intelligence and Sleep – The Next Frontier

HARBORVIEW MEDICAL CENTER UW Medicine

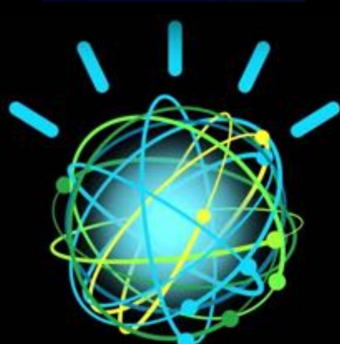
SASM

Nathaniel F. Watson, MD, MSc
Professor of Neurology
University of Washington (UW) School of Medicine
Director, Harborview Medical Center Sleep Clinic
Co-director, UW Medicine Sleep Center
@SleepDocWatson

UNIVERSITY OF WASHINGTON SCHOOL OF MEDICINE

Full Disclosure: No Relation/No COI

IBM Watson



≠

University of Washington Watson

What I will tell you

- What is Artificial Intelligence (AI)?
- Sleep medicine and AI
- Reading tea leaves – how can/will AI transform sleep medicine?
- What does AI mean for us more generally in medicine/life?

What I won't tell you

- How to feel about it (but I WILL challenge you to think about it)

Infinite Monkey Theorem

What do we mean by Artificial Intelligence?

- Artificial Intelligence
  - A branch of computer science dealing with the simulation of intelligent behavior in computers
  - The capability of a machine to imitate intelligent human behavior
- Machine Learning
  - The process by which a computer is able to improve its own performance (as in analyzing image files) by continuously incorporating new data into an existing statistical model
- Computational Phenotyping
  - A biomedical informatics method for identifying certain patient populations

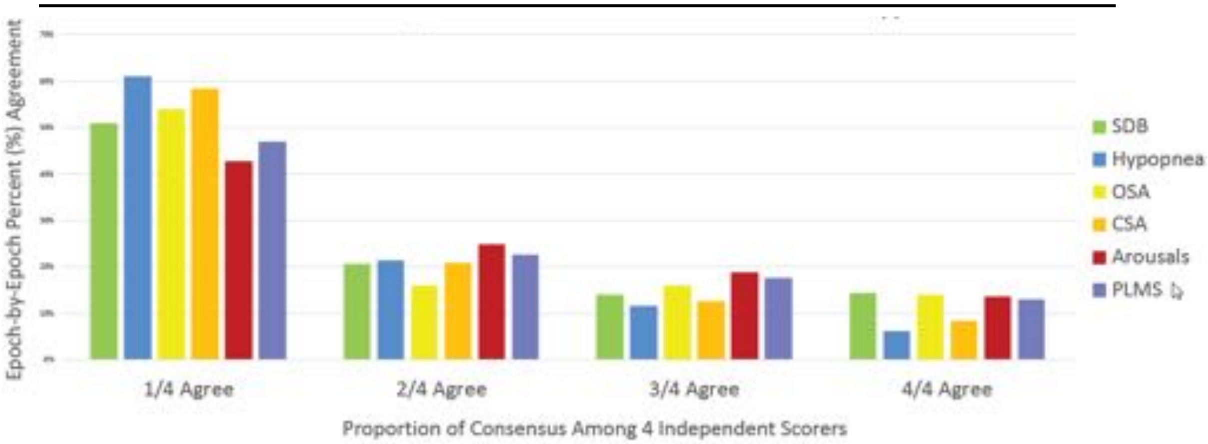
Merriam Webster; https://www.coursera.org/lecture/computational-phenotyping/introduction-to-computational-phenotyping-s0GJ8

PSG and the Dynamic Multivariate Human Physiological State

- Each of the 3.2 billion DNA base pairs in a human genome can be encoded by two bits – 800 megabytes for the entire genome
- Sequence of nucleotides comprising DNA is relatively static… while environment within each cell is highly variable
- Genome sequence does not indicate exposure to toxic water, how badly injured in a fall, how a recent surgery or change in medication affected health, etc.
- By some estimates, your physiological state at any point in time contains roughly 10^{18} (a million trillion) times more information than resides in your genetic code

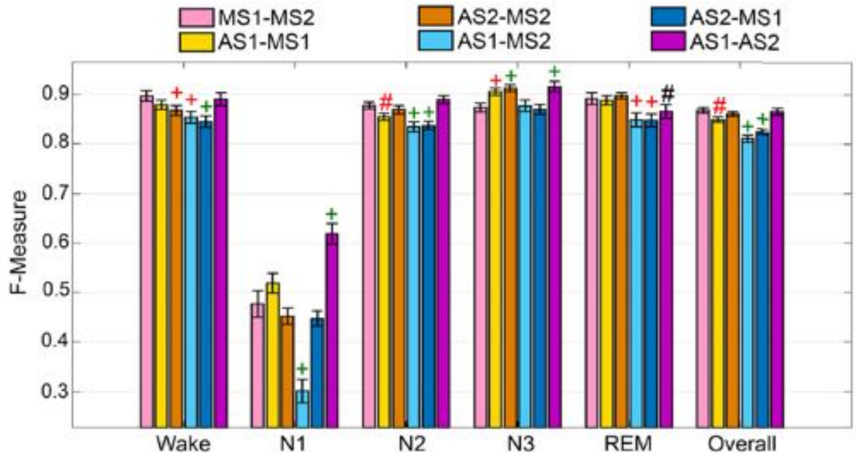
Courtesy of Chris Fernandez ensodata

Statistical Analysis of Individual Versus Consensus Scoring Agreement



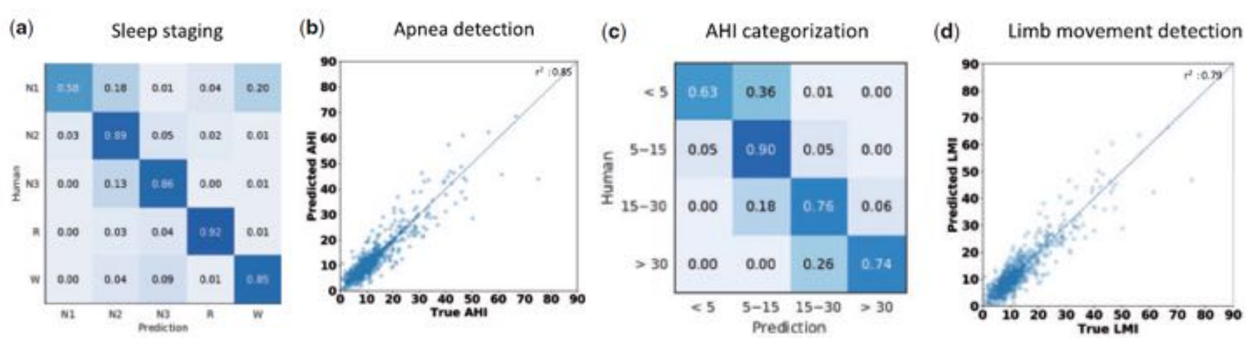
Fernandez et al. A cross validation approach to inter-scorer reliability assessment. SLEEP 2018;41:A122-123.

Machine Learning Algorithm for Automated Scoring and Analysis of Polysomnography Data



Allocca et al. Validation of 'Somnivore', a Machine Learning Algorithm for Automated Scoring and Analysis of Polysomnography Data. Frontiers in Neuroscience 2019;13: 1-18

Deep Neural Network Sleep Scoring



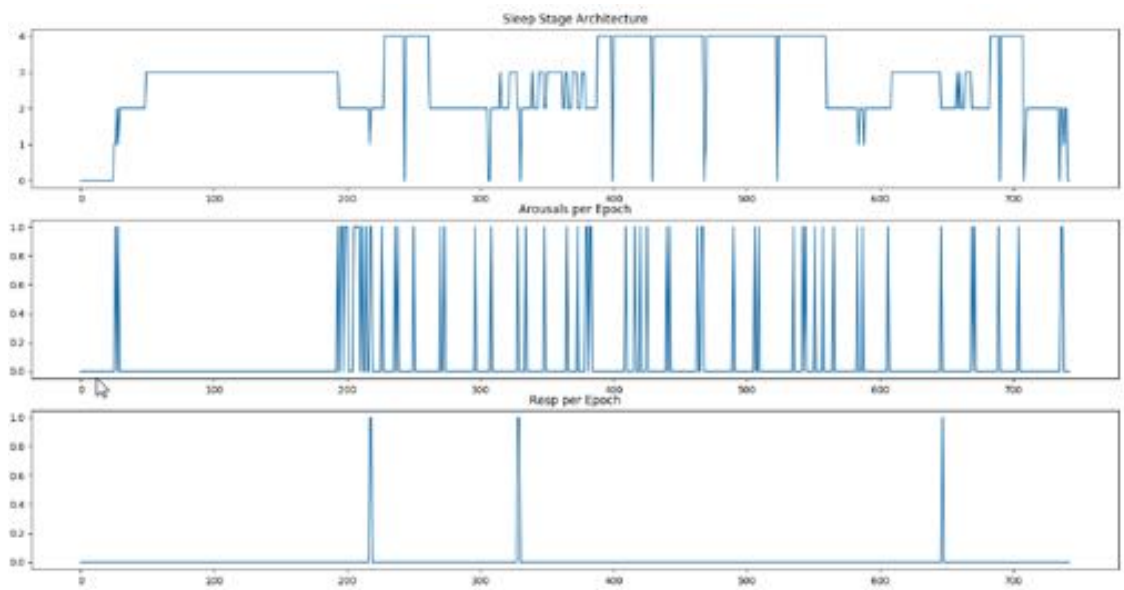
Biswal et al. Expert-level sleep scoring with deep neural networks. Journal of the American Medical Informatics Association, 25(12), 2018, 1643–1650

Can AI Estimate OSA Severity by EEG Alone?

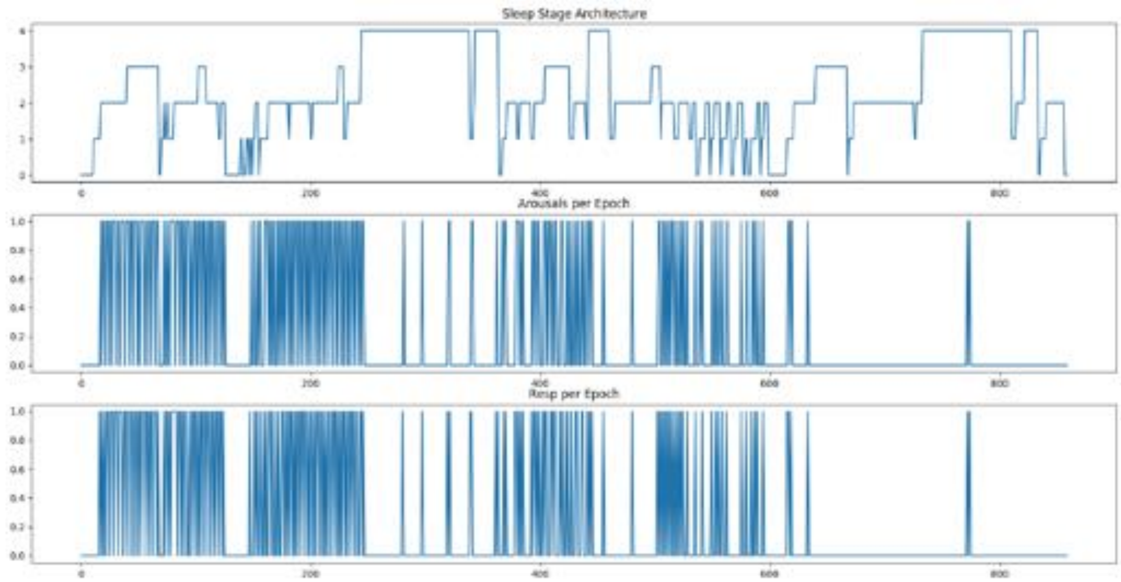
- Adult patients (N = 4,650) who completed an overnight PSG study
- All signals were excluded from analysis except the 10/20 EEG sensor array
- Global phenotypic features were derived from EEG study sleep architecture and fragmentation profiles
- Local phenotypic features were derived by analyzing biomarker patterns and respiratory cycle-related EEG changes exhibited in the EEG signals directly
- AI methods including Bidirectional-LSTM, Deep-CNN, and a combination of both were trained, optimized, and evaluated to model the relationship between global and local EEG phenotypes and OSA severity
- Performance for predicting moderate and severe OSA (AHI ≥ 15) was evaluated using randomized 10-fold cross-validation

Fernandez et al. Using novel EEG phenotypes and AI to estimate OSA severity. SLEEP 2019;42:A375

Sleep-Arousal Architecture Estimated Low-Risk for OSA (AHI ≤ 15)

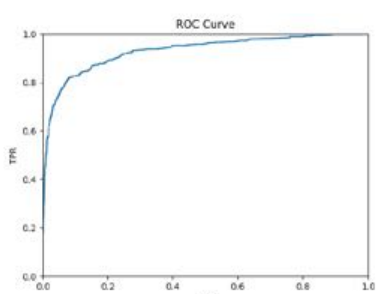


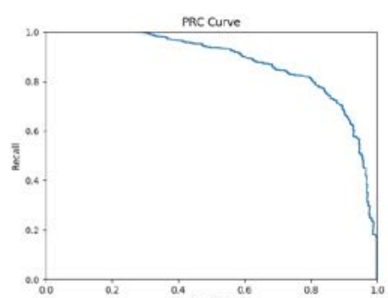
Sleep-Arousal Architecture Estimated High-Risk OSA (AHI ≥ 15)



Performance and Statistical Significance of EEG based OSA Severity Estimation

Machine Learning Model	Sensitivity	Specificity	Accuracy
DCNN-BLSTM	86.9%	99.5%	91.1%
Deep Convolutional Neural Network	84.2%	87.9%	86.8%
Bi-directional LSTM Network	70.6%	98.3%	87.6%





$$\textbf{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \text{ or } \frac{\text{True Positive}}{\text{True Positive + False Positive}}$$

$$\textbf{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \text{ or } \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

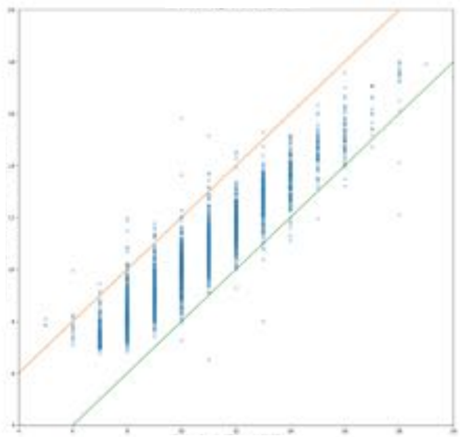
$$\textbf{Accuracy} = \frac{\text{True Positive + True Negative}}{\text{Total}}$$

https://towardsdatascience.com/precision-vs-recall-386cf9f89488

Machine Learning Predicts CPAP Rx Pressure ± 2 cm H_2O

CPAP Starting Pressure Comparison

Machine Learning Predicted Therapeutic CPAP Pressure



Physician Rx CPAP Pressure

Machine Learning Based PAP Compliance Assessment:

- Snoring time
- Heart rate
- Longest apnea
- ESS score
- Percent time under SpO_2 of 85%
- Number of apneas/hour

Munafo et al. Computational phenotyping in CPAP therapy: Using interpretable physiology-based machine learning models to predict therapeutic CPAP pressures. SLEEP 2019;42:A217.

Random Forest Analysis: 5-year All Cause Mortality

- 1,541 interpretable physiological and clinical features computationally derived from the SHHS dataset (N=5,803)
  - 435 clinical observational variables (e.g., smoking, blood pressure, cholesterol)
  - 1,170 PSG variables (e.g., sleep architecture, AHI, SpO_2 trends)
- Compumedics P-series type II PSG
- Machine learning models were trained, optimized, and evaluated: Ordinary Least Squares, Random Forest, Deep MLP, Kernel SVM, Naïve Bayes, KNN, Gaussian Process, QDA, LASSO, Logistic Regression, AdaBoost

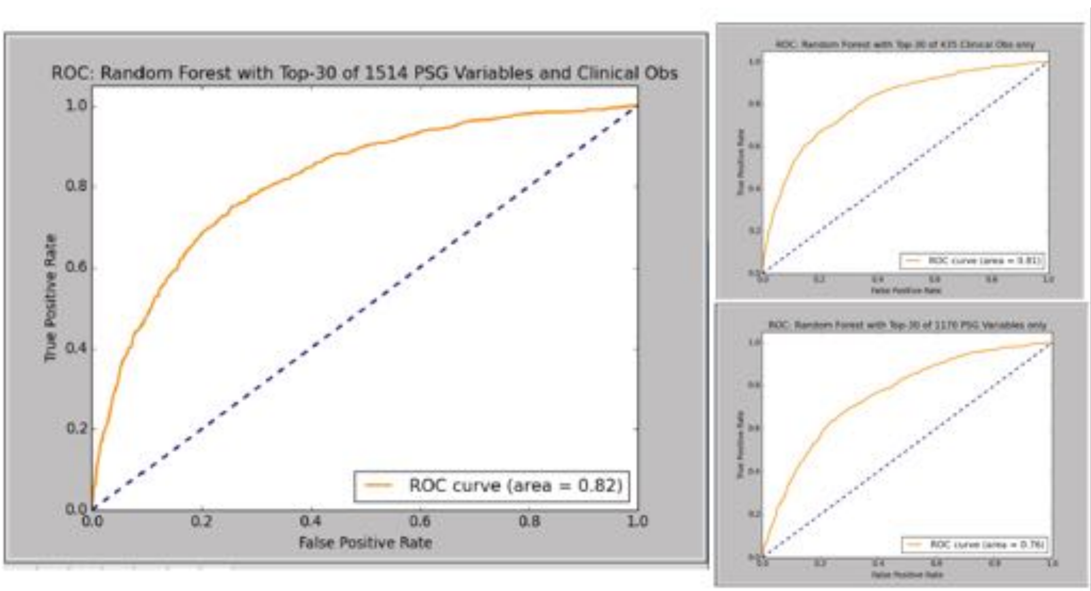
Fernandez et al. Computational phenotyping in PSG: Using interpretable physiology-based machine learning models to predict health outcomes. SLEEP 2017;40:A26.

Predictive Utility Ranking

Table of Top-30 PSG variable and clinical observation features ranked by Gini Importance:

Gini Importance (Mean Decrease Impurity)	Feature Definition
0.067	Supine arm systolic blood pressure
0.044	Forced Expiratory Volume in One Second at SHHS1
0.034	Has ECG data (SHHS1)
0.014	Ventricular rate
0.014	PSG Report (SHHS2): Sleep Efficiency
0.010	Quality of Life (SHHS1): General health
0.008	Cigarette pack-years (SHHS1)
0.008	Percent of sleep time SaO2 is below 95%
0.007	HDL cholesterol
0.006	SF-36 Calculated (SHHS1): Physical Functioning Standardized Score
0.006	Number of days since the baseline PSG until collected: ECG (SHHS1)
0.006	SF-36 Calculated (SHHS1): Physical Component Scale Standardized Score
0.005	Minimum Heart Rate (REM, Other, all oxygen desaturations)
0.005	Has SHHS1 Quality of Life form
0.005	SF-36 Calculated (SHHS1): Physical Functioning Raw Score
0.005	Forced Vital Capacity at SHHS1
0.004	Wake After Sleep Onset
0.004	Systolic BP: reading 3 of 3 (SHHS1)
0.004	Average Systolic BP (SHHS1)
0.004	Cholesterol
0.004	Minimum HR with arousal (REM, Other, 3% oxygen desaturation)
0.004	Triglycerides
0.004	Neck Circumference (SHHS1)
0.004	Sleep Time
0.004	Sleep onset time
0.003	Gender
0.003	Ankle-arm BP Index (SHHS1)
0.003	Number of oxygen desaturation with at least 2% oxygen desaturation
0.003	REM Latency II - excluding wake
0.003	Sleep time used in calculations

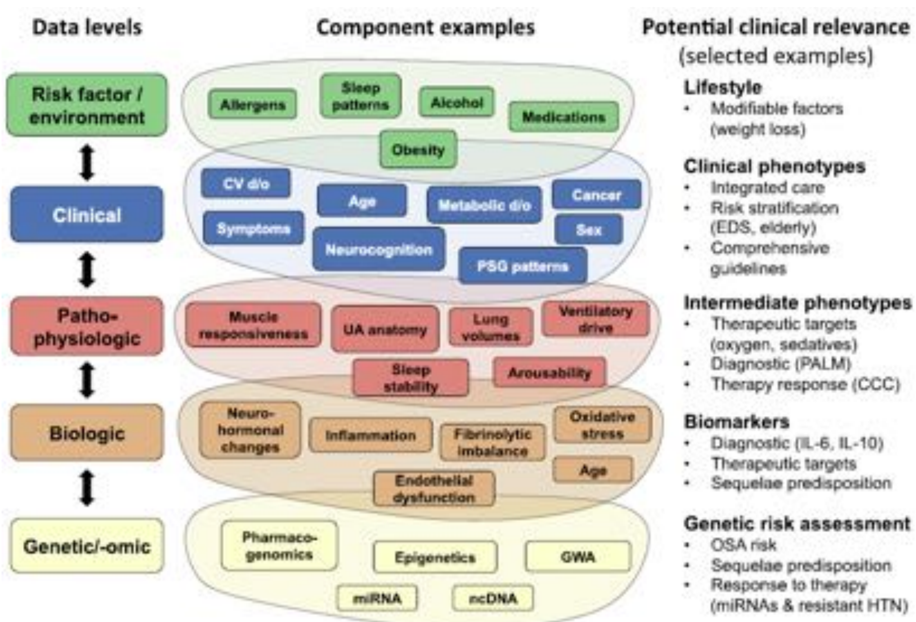
PSG-only, Obs-only, and Combined Random Forest Analysis



Computational Phenotyping Model Comparison

Table 1: Multivariate Model Comparison for Predicting All-Cause 5-Year Mortality					
N = 5,803 subjects	ROC-AUC	Accuracy	Precision	Recall	Support
Random Forest: PSG and Clinical Obs	*0.82*	*77.4%*	*86%*	*78%*	*4497*
Random Forest: Clinical Obs only	0.81	75.1%	85%	75%	4497
OLS: PSG and Clinical Obs	0.79	72.9%	85%	73%	4497
Deep MLP: PSG and Clinical Obs	0.78	77.9%	84%	78%	4497
Random Forest: PSG only	0.76	70.3%	84%	70%	4497

Fernandez et al. Computational phenotyping in PSG: Using interpretable physiology-based machine learning models to predict health outcomes. SLEEP 2017;40:A26.

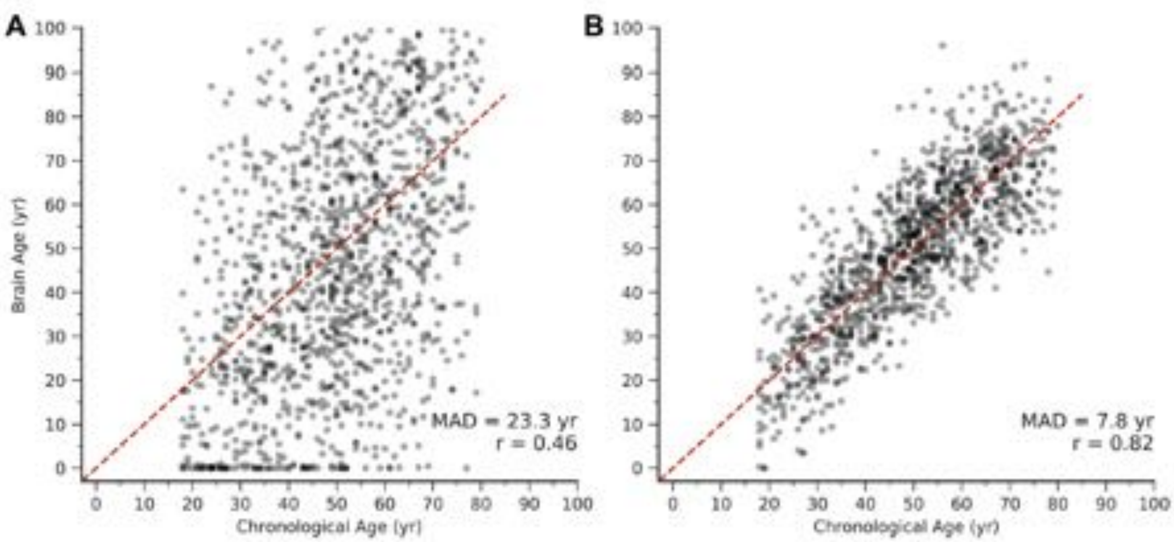


Zinchuk et al. Phenotypes in obstructive sleep apnea: A definition, examples and evolution of approaches. Sleep Medicine Reviews 35 (2017) 113-123

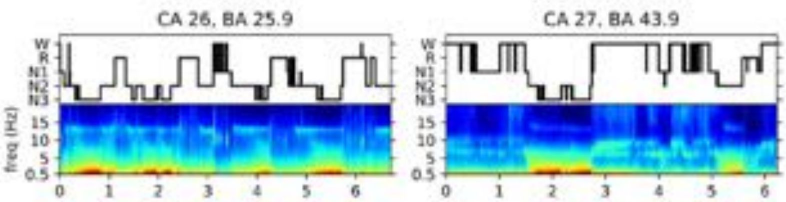
Can “Brain Age” Be Predicted by the EEG?

- Brain age (BA) serves as a potential aging biomarker where the variation of BA between individuals of the same chronological age may carry important information about the risk of cognitive impairment, neurological or psychiatric disease, or death
- Alzheimer’s disease, schizophrenia, epilepsy, traumatic brain injury, bipolar disorder, major depression, cognitive impairment, diabetes mellitus, and HIV, are associated with excess BA (on MRI)
- Machine learning model developed to predict BA based on 2 large sleep EEG data sets: the Massachusetts General Hospital (MGH) sleep lab data set (N = 2,532; ages 18-80); and the Sleep Heart Health Study (SHHS, N = 1,974; ages 40-80).

Sun et al. Brain age from EEG of sleep. Neurobiology of Aging 2019;74:112-120

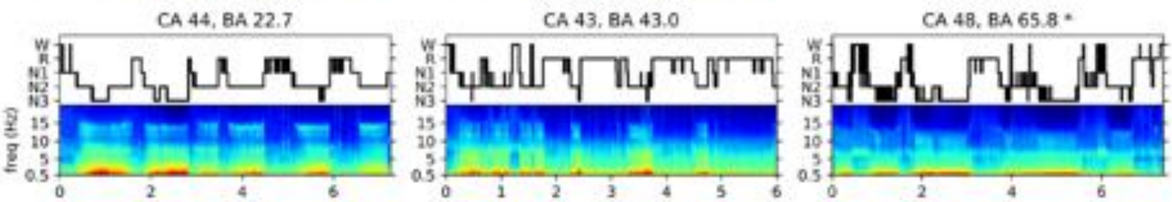


Sun et al. Brain age from EEG of sleep. Neurobiology of Aging 2019;74:112-120

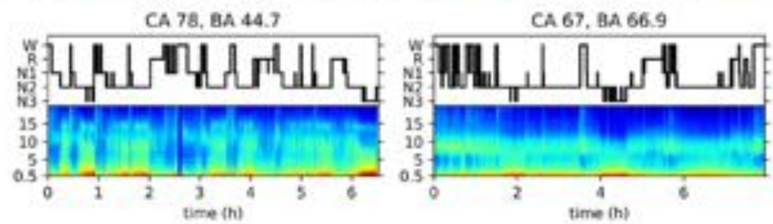


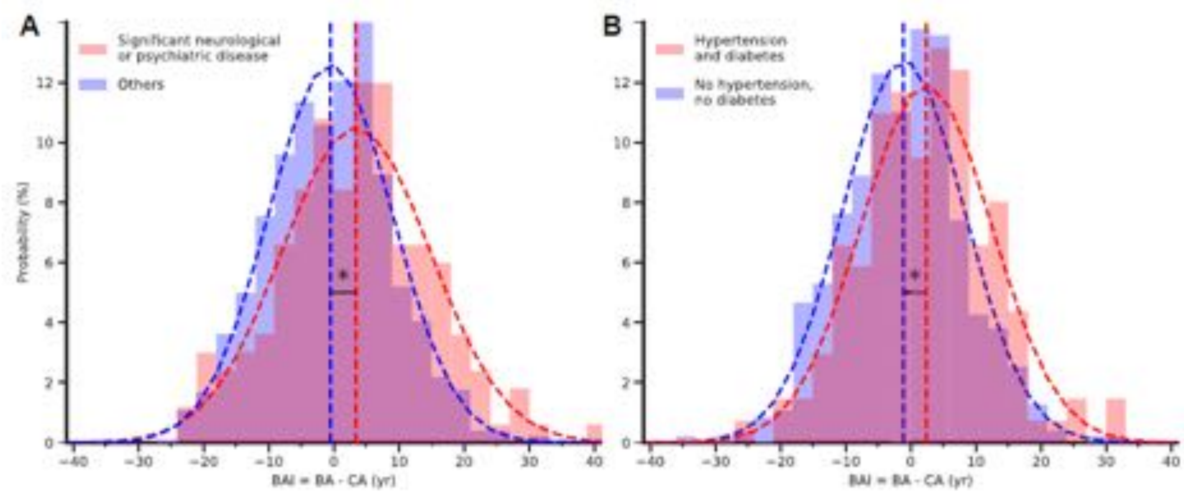
Covariates that have significant correlation with BAI (BA-CA)

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



Sun et al. Brain age from EEG of sleep. Neurobiology of Aging 2019;74:112-120





Sun et al. Brain age from EEG of sleep. Neurobiology of Aging 2019;74:112-120

Time to Think About it

- What does it mean:
  - For sleep medicine?
  - For medicine in general?
  - For us as members of society?

Embrace the Future

Protect the Sleep Medicine Status Quo

- **Genotype** - the genetic constitution of an individual organism
- **Phenotype** - the set of observable characteristics of an individual resulting from the interaction of its genotype with the environment
- **Genetic determinism** - the belief that human behavior is controlled by an individual's genes or some component of their physiology, generally at the expense of the role of the environment, whether in embryonic development or in learning
- **The quantified self** - refers both to the cultural phenomenon of self-tracking with technology and to a community of users and makers of self-tracking tools who share an interest in "self-knowledge through numbers."
- **Phenotype determinism** - ???

Comparison of Accuracy of CSTs Versus PSG

	Sleep	Wake	Light	Deep	REM
SleepScore	88%	66%	58%	62%	57%
Fitbit Charge 2	96%	61%	81%	49%	74%
Ōura Ring	96%	48%	65%	51%	61%
Beddit	N/A	42%	56%	37%	N/A

De Zambotti et al. Chronobiology International 2018;35(4):465-476
Tuominen et al. *J Clin Sleep Med.* 2019;15(3):483–487
Zaffaroni et al. Engineering in Medicine and Biology 2019, Berlin, Germany, July 23-27
de Zambotti et al. Behav Sleep Med 2018 1–15. doi:10.1080/15402002.2017.1300587

Back to the Future?

https://www.nuance.com/healthcare/ambient-clinical-intelligence; https://www.engineering.com/DesignerEdge/DesignerEdgeArticles/ArticleID/17664/A-Healthy-Future-for-Artificial-Intelligence-in-Healthcare.aspx

THE WALL STREET JOURNAL.

Latest News U.S. Politics World Business

Expedia uses five of the seven emotions that the facial-recognition system tracks: joy, anger, surprise, sadness and disgust. The other two—contempt and fear—don't really factor into travel decisions (they hope). The systems detect changes in emotional state and can capture that the second it happens. The eye tracker tells researchers what the person was looking at when the emotion changed.

Accessed 10/5/2019

What I told you

- Defined AI, machine learning, and computational phenotyping
- AI and sleep study scoring
- EEG and OSA severity estimation (diagnosis?)
- AI and PAP Rx accuracy
- Computational phenotyping and 5-year survival with PSG/clinical variables
- “Brain Age” and AI
- Consumer sleep technology accuracy comparison
- AI and implications for the future of health care