Screening for Cognitive Impairment During Sleep Studies: The George B. Moody PhysioNet Challenge 2026

Matthew A. Reyna[1], Audrey Weigle[1], Qichen Li[1], Zuzana Koscova[1], Haoqi Sun[2,3], Shenghan Wen[1,4], Umakanth Katwa[3,5], Robert Thomas[2,3], Dennis Hwang[6], Lynn Marie Trotti[7], Emmanuel Mignot[8], Reza Sameni[1,9], Samaneh Nasiri[1], M. Brandon Westover[8,*], Gari D. Clifford[1,9,*]

[1] Department of Biomedical Informatics, Emory University, Atlanta, Georgia, USA
[2] Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA, USA
[3] Harvard Medical School, Boston, MA, USA
[4] Department of Biomedical Engineering, Tsinghua University, Beijing, China
[5] Boston Children's Hospital, Boston, MA, USA
[6] Kaiser Permanente, San Bernardino County Sleep Disorders Center, San Bernardino, CA, USA
[7] Department of Neurology and Emory Sleep Center, Emory School of Medicine, Atlanta, GA, USA
[8] Department of Neurology, Stanford University, Palo Alto, CA, USA
[9] Department of Biomedical Engineering, Emory University and the Georgia Institute of Technology, Atlanta, Georgia, USA
[*] co-senior authors

Abstract

The George B. Moody PhysioNet Challenge 2026 invited teams to develop algorithmic approaches for using physiological data from sleep studies to predict future cognitive impairment diagnoses.

Sleep is a vital physiological process that affects many aspects of human health. Polysomnography (PSG) is a type of sleep study that records physiological signals during sleep, often to diagnose sleep disorders. However, the PSG can reveal far more.

For the PhysioNet Challenge 2026, we invited participants to design and implement working, open-source algorithms for using PSG recordings from the Human Sleep Project to predict future cognitive impairment diagnoses. The team with the best score on the hidden test set will win the Challenge.

Hundreds of participants from over 100 teams submitted over 1000 entries during the Challenge, representing diverse approaches from academia and industry worldwide.

1. Introduction

The George B. Moody PhysioNet Challenges are annual competitions that support the development of open-source approaches to complex physiological and clinical problems [1]. For the PhysioNet Challenge 2026, we invited teams to develop algorithms for using physiological signals from overnight sleep studies to predict future cognitive impairment diagnoses.

Sleep is a vital physiological process that is essential for brain, cardiovascular, metabolic, immune, and mental health. Quality sleep supports health and well-being, while disrupted or insufficient sleep can signal and contribute to a wide range of illnesses [2]. Polysomnography (PSG), which records multiple physiological signals during sleep, is traditionally used to diagnose sleep disorders such as insomnia and obstructive sleep apnea. However, the value of the PSG extends far beyond sleep-specific diagnoses: changes in brain activity, heart rate, respiration, and other signals during sleep can reveal underlying or emerging chronic neurological, cardiovascular, and respiratory conditions [3]. By identifying these changes, sleep studies provide not only a snapshot of sleep health but an early warning for broader neurological and physiological dysfunction. The George B. Moody PhysioNet Challenge 2026 leverages this potential by inviting teams to use PSG recordings to identify individuals at risk of cognitive impairment, supporting earlier interventions and improved long-term outcomes.

2. Methods

We shared physiological, demographic, and diagnostic data from the Human Sleep Project [4,5], devised and implemented a clinically relevant evaluation metric, and evaluated over 1000 algorithmic entries from over 100 teams for the PhysioNet Challenge 2026.

2.1. Challenge Data

The Challenge data included physiological data and annotations, demographic data, and diagnostic data from the Human Sleep Project [4].

The physiological data include PSG recordings with electroencephalography (EEG), electrooculography (EOG), electromyography (EMG), electrocardiography (ECG), chest and abdominal respiratory effort sensors, and thermal or pressure-based airflow sensors. Algorithmic and human annotations of the PSGs include sleep staging and events [6]. The demographic and diagnostic data include age, sex, body mass index, and International Classification of Diseases (ICD) codes and dates. Some channels and other variables were not available for some patients or encounters.

The Human Sleep Project contains six distinct institutional settings. We included data from three sites in the public training set, one site in the hidden validation set, and one site in the hidden test set; we excluded data from a pediatric site. We included all available variables in the training set, but we only included PSG data, algorithmic annotations, and non-diagnostic demographic data in the validation and test sets; we excluded data requiring human annotation and diagnosis from the hidden data.

We preserved the data in European Data Format (EDF) from the original sites. We replaced all ages above 89 with a single age of 90 according to the Safe Harbor method for de-identification of protected health information [7]. Teams could choose which signals to use and how to use them, but we provided optional code for renaming and re-referencing the signals. We shared “small” and “large” versions of the training set, which were approximately 200 GB and 1.2 TB, respectively; the small version was sampled from the larger version to approximately preserve the demographic and diagnostic distributions of the large version of the training set for teams that preferred a smaller training set [8,9].

2.2. Challenge Objective

The George B. Moody PhysioNet Challenge 2026 invited teams to develop algorithmic approaches for using physiological data from sleep studies to predict future cognitive impairment diagnoses.

2.2.1. Challenge Timeline

This year’s Challenge was the 27$^{\text{th}}$ George B. Moody PhysioNet Challenge [1]. As in previous years, the Challenge had an unofficial phase and an official phase.

The unofficial phase (10 February 2026 to 9 April 2026) introduced the teams to the Challenge. We publicly shared the Challenge objective, training data, example algorithms, and evaluation metric and invited the teams to submit their code for training and evaluation, scoring at most five entries from each team on the hidden validation set. Between the unofficial and official phases, we took a hiatus (10 April 2026 to 2 June 2026) to improve the Challenge. The official phase (3 May 2026 to 20 August 2026) continued the Challenge. We updated the Challenge data, example algorithms, and evaluation metric and invited teams to continue to submit their code for training and evaluation, scoring at most ten entries from each team on the hidden validation set.

After the end of the official phase, we attempted to evaluate one entry from each team on the hidden test set to determine the winners of the Challenge.

We will announce the results at the end of the Computing in Cardiology (CinC) 2026 conference, where the teams presented, defended, and published their work. Participation in CinC was a requirement for rankings and prize eligibility. We will publicly release the algorithms after the end of the Challenge and the publication of these papers.

2.2.2. Challenge Evaluation

We evaluated the ability of algorithms to prioritize patients with future cognitive impairment diagnoses over patients without future cognitive impairment diagnoses. To do so, we defined positive and negative patient groups for cognitive impairment and an age-conditioned area under the receiver-operating characteristic curve (AUROC) metric to evaluate the algorithms. By conditioning on age, the age-conditioned AUROC metric better reflects the ability of algorithms to use physiological data and annotations to predict cognitive impairment.

We defined labels for cognitive impairment as any label from a list of 115 ICD-9 and ICD-10 codes for mild cognitive impairment, dementia, and Alzheimer’s disease [10]. With these ICD codes and dates, we defined three lists of patients:

- **Positive patients:** Two or more cognitive impairment diagnoses with a first diagnosis at least 1 year and no more than 6 years after the sleep study and at least one week between cognitive impairment diagnoses.

• **Negative patients:** No cognitive impairment diagnoses at any time and at least 6 years of encounters after the sleep study.

• **Other patients:**, Two or more cognitive impairment diagnoses with a first diagnosis less than 1 years after the sleep study; two or more cognitive impairment diagnoses with a first diagnosis more than 6 years after the sleep study; only one cognitive impairment diagnosis; no cognitive impairment diagnoses and less than 6 years of encounters after the sleep study. These patients were not included in the comparison.

Formally, let $\mathcal{C}$ be a cohort of positive patients $\mathcal{P}$ and negative patients $\mathcal{N}$. For each patient $c_k \in \mathcal{C}$, let $a_k \in \mathbb{R}$ be the patient's age, and, for a given model, let $z_k \in \mathbb{R}$ be the model's cognitive impairment prediction for the patient. We define the age-conditioned AUROC

$$s_{\mathcal{C}}^{\delta} = \Pr(z_i > z_j \mid c_i \in \mathcal{P},\ c_j \in \mathcal{N},\ |a_i - a_j| \leq \delta) + \frac{1}{2}\Pr(z_i = z_j \mid c_i \in \mathcal{P},\ c_j \in \mathcal{N},\ |a_i - a_j| \leq \delta) \quad (1)$$

as the probability that the model prioritizes a random positive patient over a random negative patient of approximately the same age, i.e., within $\delta = 2$ years. We compare similarly aged patients to reduce dataset artifacts.

3. Challenge Results

During the Challenge, over 100 teams submitted over 1000 algorithms. We will update this section after the conclusion of the Challenge.

4. Discussion

We will update this section after the conclusion of the Challenge.

5. Conclusions

We will update this section after the conclusion of the Challenge.

Acknowledgements

This research was supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB: R01EB030362); the National Heart, Lung, and Blood Institute (NHLBI: R01HL161253); the National Center for Advancing Translational Sciences of the National Institutes of Health (NCATS: UL1TR002378); as well as AliveCor under unrestricted gifts. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the above entities.

References

[1] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–e220.

[2] Ramar K, Malhotra RK, Carden KA, Martin JL, Abbasi-Feinberg F, Aurora RN, et al. Sleep is essential to health: an American Academy of Sleep Medicine position statement. Journal of Clinical Sleep Medicine October 2021; 17(10):2115–2119. ISSN 1550-9397.

[3] Miller MA. The role of sleep and sleep disorders in the development, diagnosis, and management of neurocognitive disorders. Frontiers in Neurology October 2015;6. ISSN 1664-2295.

[4] Li Q, Wen S, Sun H, Ganglberger W, Tripathi A, Turley N, et al. The Human Sleep Project: A multi-center clinical polysomnography dataset across the human lifespan. SLEEP August 2026;ISSN 1550-9109.

[5] Sun H, Ganglberger W, Nasiri S, Gupta A, Ghanta M, Moura V, et al. The Human Sleep Project (version 2.0), 2023.

[6] Nasiri S, Ganglberger W, Nassi T, Meulenbrugge EJ, Moura Junior V, Ghanta M, et al. CAISR: achieving human-level performance in automated sleep analysis across all clinical sleep metrics. SLEEP 2025;48(8). ISSN 1550-9109.

[7] U.S. Department of Health and Human Services. Guidance on methods for de-identification of protected health information in accordance with the HIPAA privacy rule. Technical report, Office for Civil Rights, November 2012. URL https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html.

[8] George B. Moody PhysioNet Challenge. PhysioNet Challenge 2026 Data, 2026. URL https://www.kaggle.com/datasets/physionet/physionetchallenge2026data. Kaggle dataset. Accessed August 26, 2026.

[9] George B. Moody PhysioNet Challenge. PhysioNet Challenge 2026 Data (Large Version), 2026. URL https://www.kaggle.com/datasets/physionet/physionetchallenge2026datalargeversion. Kaggle dataset. Accessed August 26, 2026.

[10] George B. Moody PhysioNet Challenge. Screening for Cognitive Impairment During Sleep Studies: The George B. Moody PhysioNet Challenge 2026, 2026. URL https://physionetchallenges.org/2026/. Accessed August 26, 2026.

Address for correspondence:
Matthew A Reyna
DBMI, 101 Woodruff Circle, 4th Floor East, Atlanta, GA 30322
matthew.a.reyna@emory.edu