118

Enhancement of a Machine Learning Algorithm to Alert Sleep Clinicians of Patients at Risk for Narcolepsy, Using Nocturnal Polysomnography in General Sleep Medicine Clinics

World Sleep 2022
11-16 March 2022
Rome, Italy

Hyatt Moore IV, PhD[1,2]; Alex Zheng, PhD[3]; Alyssa Cairns, PhD[4]; Prasheel Lillaney, PhD[5]; Jed Black, MD[2,5]

[1]InformAton Inc., Salinas, CA, USA; [2]Stanford University Center for Sleep Sciences and Medicine, Palo Alto, CA, USA; [3]Huneo, Fremont, CA, USA; [4]BioSerenity, Columbia, SC, USA; [5]Jazz Pharmaceuticals, Palo Alto, CA, USA

Introduction

- Narcolepsy frequently remains undiagnosed for many years following symptom onset, likely due to a combination of clinician lack of narcolepsy-specific education/experience and substantial medical comorbidity (eg, sleep apnoea)[1]
- Previous work has shown that polysomnography (PSG) contains quantitative information that, using machine learning algorithms, may aid in the identification of narcolepsy type 1 (NT1)[2]
  - Stephansen et al[2] introduced the hypnodensity graph, which is a machine learning scoring method that estimates the probability of each sleep stage for each epoch of sleep, thereby conveying more information about sleep trends than classical hypnograms
  - Applying deep learning–derived hypnodensity features to the diagnosis of NT1, Stephansen et al[2] showed that analysis of a single nocturnal PSG can perform as well as the PSG–multiple sleep latency test (MSLT) gold standard; however, when applied to a real-world sleep clinic population, the algorithm performed inadequately

Objective

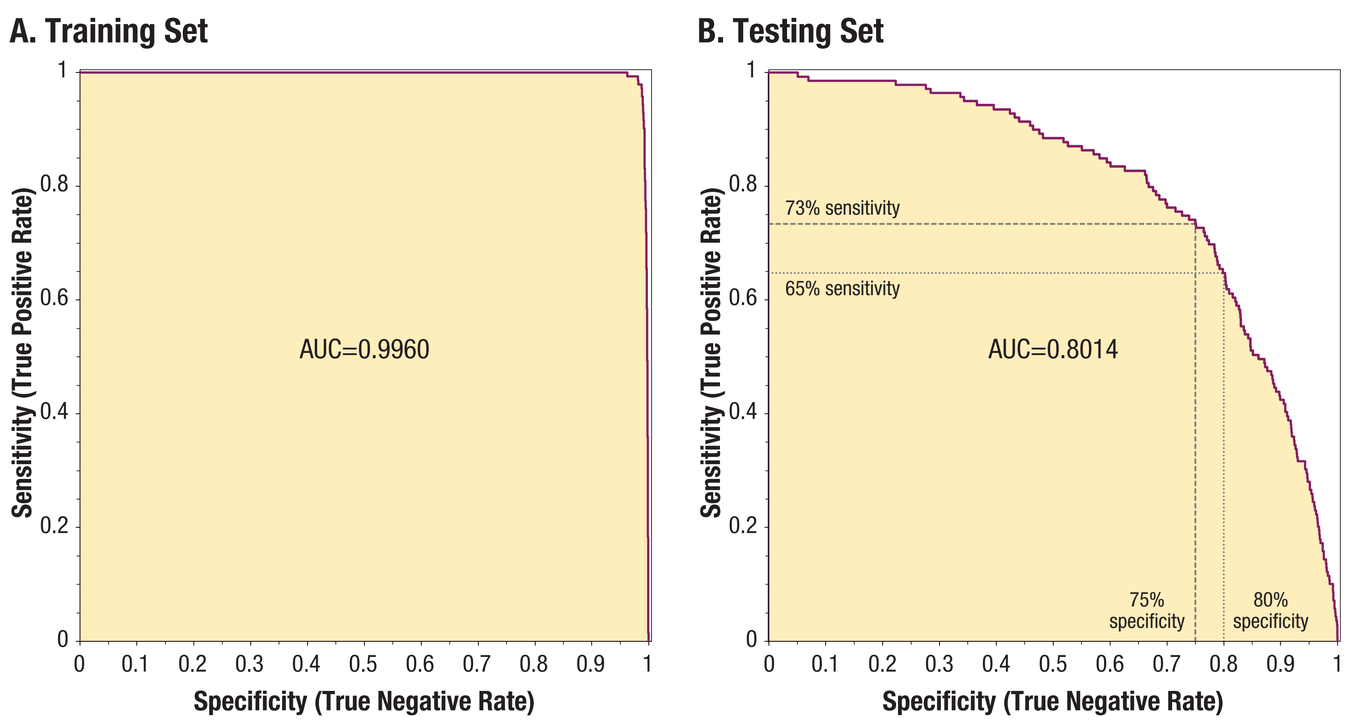
- The current study aimed to further develop and enhance the performance of the algorithms developed by Stephansen et al[2] using data from a large sleep clinic population to create a tool with high sensitivity and specificity to alert sleep clinicians about patients at risk for narcolepsy

Methods

- PSG studies were obtained in collaboration with SleepMed Research (now BioSerenity, USA) from a random sample of sleep clinic patients who were being evaluated for various sleep conditions
  - Studies were randomly split (1:1) into a training set (used to develop/train a Gaussian process [GP] model to identify patients with a high probability of having narcolepsy) and a testing set (used to validate the GP model)
  - MSLT results (performed the morning following the nocturnal PSG for some individuals) were used as a proxy for categorising high probability of NT1 (MSLT with ≥3 sleep onset rapid eye movement periods [SOREMPs] and mean sleep latency ≤5 minutes) and narcolepsy type 2 (NT2; MSLT with ≥2 SOREMPs and mean sleep latency ≤8 minutes [and did not meet criteria for NT1]); these criteria were at least as stringent as historical standards[3,4]
- Hypnodensity graphs were estimated from the PSGs on 15-second epochs using a previously developed convolutional neural network[2]
- Additional feature engineering was applied to the hypnodensities to create a feature vector that was used to train the GP model
  - 3 approaches were considered for scaling the features: scaled to the 85th percentile, scaled to zero mean and unit variance, and unscaled
  - A recursive feature elimination scheme was compared with training the GP kernel's length scale for determining the subset of features that best discriminate narcolepsy and controls in the GP model
  - A synthetic minority oversampling technique was applied in combination with random undersampling to balance the distribution of cases and controls in the training set
- The model's performance (ability to distinguish between narcolepsy and healthy controls based on sensitivity and specificity) considered receiver operating characteristics (ROC) with the goal of achieving an area under the curve (AUC) ≥0.80 when plotting specificity versus sensitivity
- An additional goal, based on sleep medicine clinician feedback, was to confirm narcolepsy diagnosis in at least 3 of each 4 algorithm-identified patients with narcolepsy (ie, ≥75% specificity) while having the system identify at least half of the patients with narcolepsy in the overall sample (ie, ≥50% sensitivity)

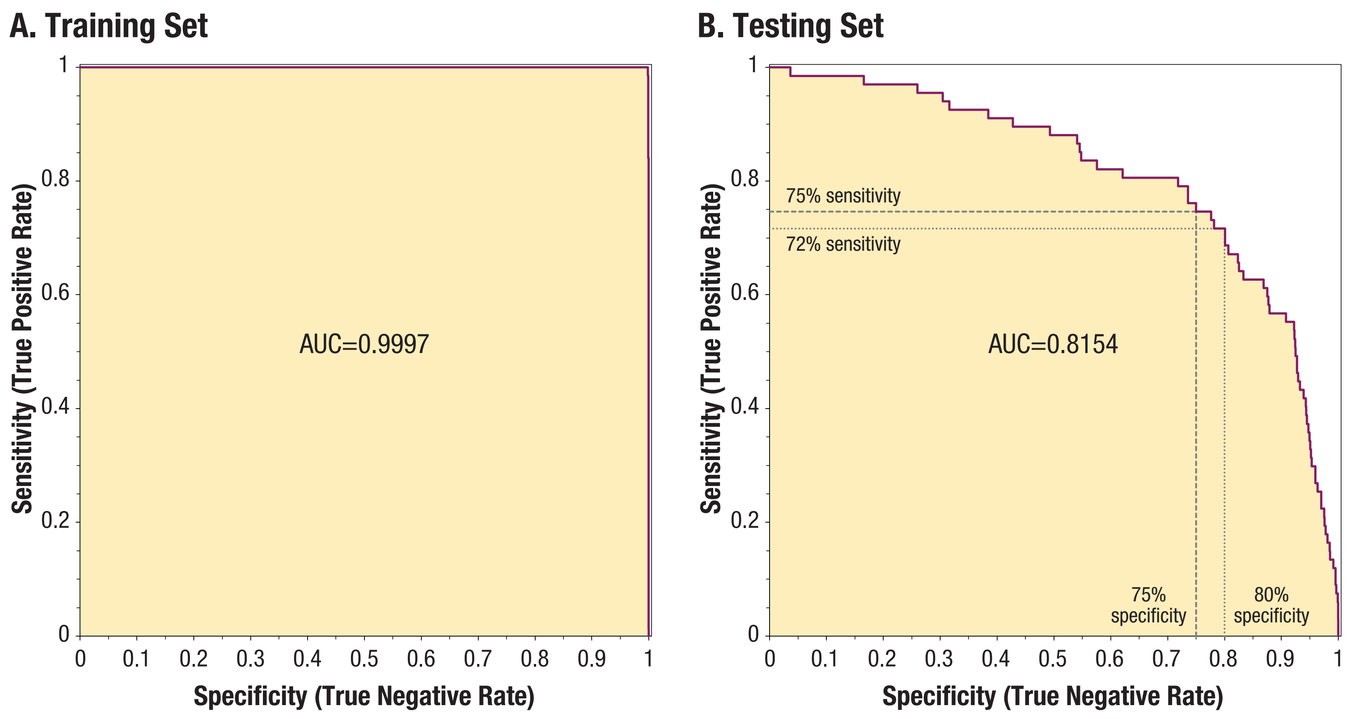
Results

Figure 1. Algorithm Classified Patients Who Have a High Probability of Narcolepsy (NT1 or NT2) With High Sensitivity and Specificity



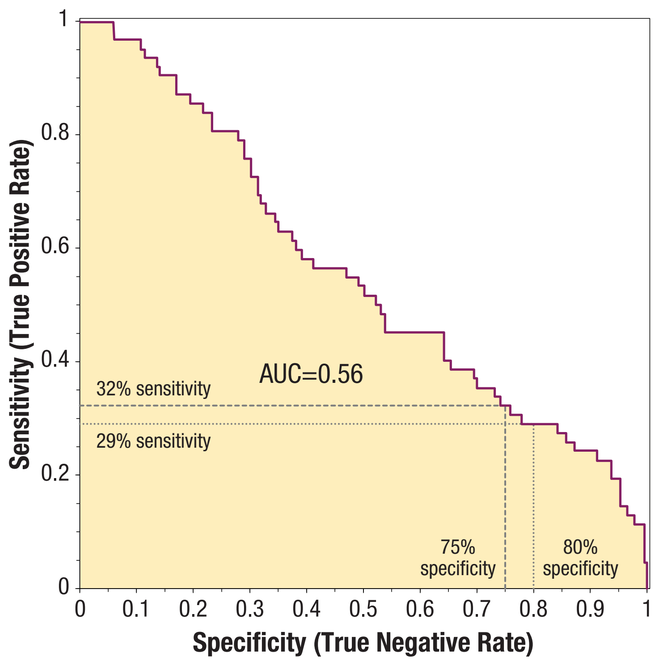
AUC, area under the curve; NT1, narcolepsy type 1; NT2, narcolepsy type 2.

Figure 2. Algorithm Performance Was Best (Greatest Sensitivity and Specificity) When Narcolepsy Was Defined as NT1 Only



AUC, area under the curve; NT1, narcolepsy type 1.

Figure 3. Inadequate Stephansen et al Algorithm Performance (Low Sensitivity for Given Specificity) When Applied to a Real-World Dataset With Narcolepsy Defined as NT1 Only



AUC, area under the curve; NT1, narcolepsy type 1.

Table 1. Dataset Characteristics

	Training Set (n=10,938)	Testing Set (n=10,899)	Total (N=21,837)
Non-narcolepsy control PSG studies (n)	10,787	10,748	21,535
Narcolepsy-related PSG studies (n)	151	151	302
NT1[a] (n)	69	68	137
NT2[b] (n)	82	83	165

MSLT, multiple sleep latency test; NT1, narcolepsy type 1; NT2, narcolepsy type 2; PSG, polysomnography; SOREMP, sleep onset rapid eye movement period.
[a]Defined as MSLT with ≥3 SOREMPs and mean sleep latency ≤5 minutes.
[b]Defined as MSLT with ≥2 SOREMPs and mean sleep latency ≤8 minutes (and did not meet criteria for NT1).

- The final GP model used a Matérn 5/2 covariance kernel with the length scale hyperparameter trained to determine the feature subset selection
- Input features were normalised to zero mean and unit variance

Conclusions

- Results of this study support the development of a machine learning–based algorithm that can offer an objective, sensitive, and specific tool to alert sleep clinicians about patients at risk for narcolepsy, using nocturnal polysomnography in general sleep medicine clinics
- The present model improves upon the algorithm of Stephansen et al, which performed strongly in a research study population but did not translate to a real-world clinical population[2]; additional validation is ongoing

References: 1. Thorpy MJ, Krieger AC. *Sleep Med.* 2014;15:502-7. **2.** Stephansen JB, et al. *Nat Commun.* 2018;9:5229. **3.** Mitler MM, et al. *Electroencephalogr Clin Neurophysiol.* 1979;46:479-81. **4.** Richardson GS, et al. *Electroencephalogr Clin Neurophysiol.* 1978;45:621-7.

Support and Acknowledgments: This study is supported by Jazz Pharmaceuticals. Under the direction of the authors, Hannah Ritchie, PhD, of Peloton Advantage, LLC, an OPEN Health company, provided medical writing and editorial support for this poster, which was funded by Jazz Pharmaceuticals.

Disclosures: H Moore IV is a consultant for InformAton Inc. **A Zheng** is an employee of Huneo who works with and receives financial support from Jazz Pharmaceuticals. **A Cairns** is an employee of BioSerenity and, during the course of this project, received grant funding from Jazz Pharmaceuticals. **P Lillaney** is an employee of Jazz Pharmaceuticals who, in the course of his employment, has received stock options exercisable for, and other stock awards of, ordinary shares of Jazz Pharmaceuticals plc. **J Black** is a part-time employee of Jazz Pharmaceuticals and shareholder of Jazz Pharmaceuticals plc.

Scan this code to access this poster online. This code is not for promotional purposes.