053

External Validation of an Enhanced Machine Learning Algorithm: Polysomnography-Based Narcolepsy-Like Feature Assessment and Clinician Notification in Routine Sleep Medicine Clinics

SLEEP 2022, the 36th Annual Meeting of the Associated Professional Sleep Societies (APSS)
June 4-8, 2022 • Charlotte, NC

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 little specific education for clinicians and few opportunities to gain experience, as well as substantial patient medical comorbidities (eg, sleep apnea) that may be associated with similar symptoms[1]
- Polysomnography (PSG) contains quantitative information that, with use of machine learning (ML) algorithms, may help identify type 1 narcolepsy (NT1)[2]
  - Stephansen et al. introduced the hypnodensity graph, which is an ML 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[2]
  - Applying deep learning–derived hypnodensity features to the diagnosis of NT1, Stephansen et al. showed that analysis of a single nocturnal PSG can perform as well as the gold standard PSG–multiple sleep latency test (MSLT)[2]; however, the algorithm performed inadequately in a real-world sleep clinic population
- The algorithms of Stephansen et al. were further developed to enhance performance by 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[3]

Objective

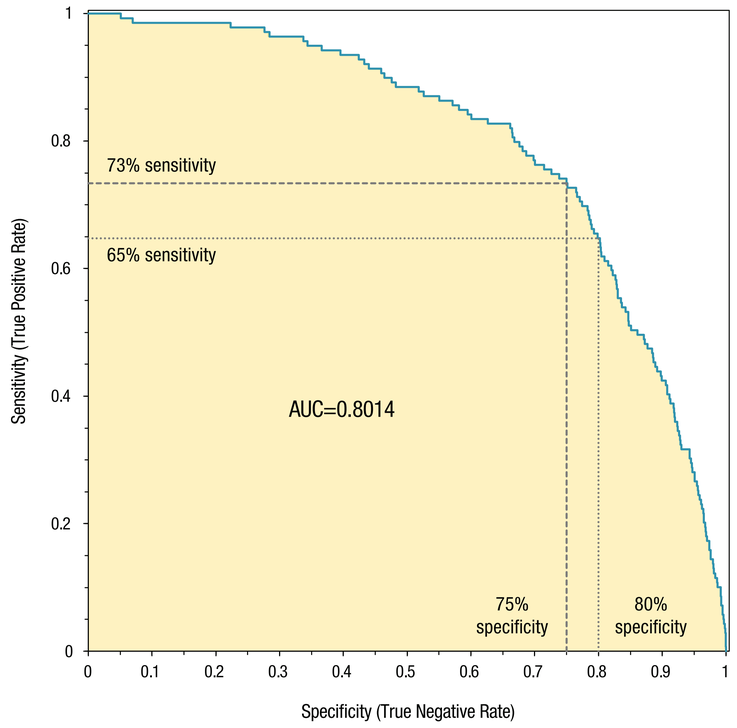
- This study aimed to further validate the utility of the previously developed ML evaluation of PSG data for detecting narcolepsy in a "real-world" sleep clinic population

Methods

- PSG studies were obtained in collaboration with SleepMed Research (now BioSerenity, USA) from a random sample of sleep clinic patients being evaluated for various sleep conditions
  - For the initial algorithm testing and training, nocturnal PSG studies (narcolepsy, n=302; controls, n=21,535) were randomly split (1:1) into a training set and a validation set for testing
    - Results of MSLTs (performed the morning following the nocturnal PSG) were used as a proxy for categorizing 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 not meeting criteria for NT1]); these criteria were at least as stringent as historical criteria[4,5] and patients with MSLTs not meeting these criteria were excluded
  - To further validate algorithm performance, a separate, external PSG data set (narcolepsy, n=82; controls, n=6948) was used for additional testing of the final model
- Hypnodensities were estimated from PSGs on 15-second epochs using a previously developed convolutional neural network
- Feature engineering was applied to hypnodensities to create a feature vector that was used to train a Gaussian process (GP) model to identify patients with a high probability of having narcolepsy
  - 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 patients with narcolepsy and controls
  - 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 considered receiver operating characteristics (ROCs) with the goal of achieving an area under the curve (AUC) ≥0.80 when plotting specificity versus sensitivity
- An additional goal was to achieve ≥75% specificity while having the system identify at least half of the patients with narcolepsy in the overall sample (ie, ≥50% sensitivity)

Results

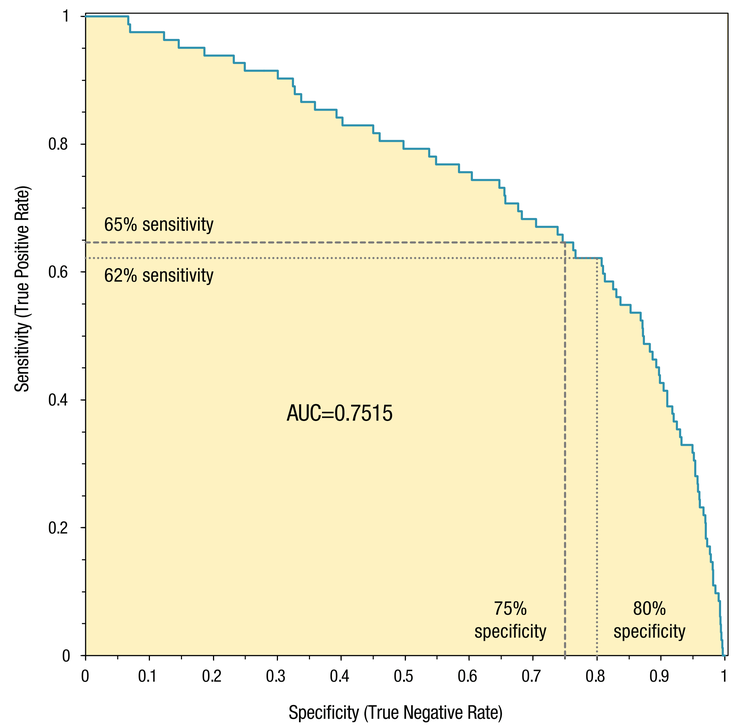
Figure 1. Algorithm-Classified Patients Who Have a High Probability of Narcolepsy (NT1 or NT2) With High Sensitivity and Specificity



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

- The model had AUCs of 0.9960 and 0.8014 for classifying narcolepsy in the training and initial testing sets, respectively
- Sensitivity ranged from 73% to 65% when specificity was between 75% and 80%

Figure 2. Validation Set Confirmed Performance of Algorithm for Classifying Narcolepsy (NT1 or NT2) With High Sensitivity and Specificity



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

- The model had an AUC of 0.7515 for classifying narcolepsy in the validation set
- Sensitivity ranged from 65% to 62% when specificity was between 75% and 80%

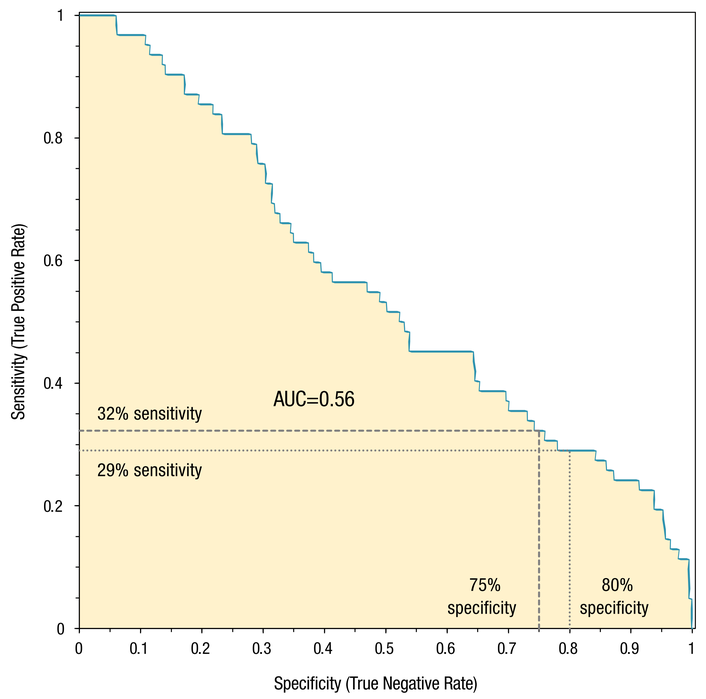
Table 1. Data Set Characteristics

	Training Set (n=10,938)	Testing Set (n=10,899)	Total (N=21,837)	Validation Set (N=7030)
Non-narcolepsy control PSG studies (n)	10,787	10,748	21,535	6948
Narcolepsy-related PSG studies (n)	151	151	302	82
NT1[a] (n)	69	68	137	33
NT2[b] (n)	82	83	165	49

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 not meeting 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 normalized to zero mean and unit variance

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.

Conclusions

- The utility and performance of an ML-based model to detect narcolepsy were validated using nocturnal PSG studies from a real-world sleep clinic population
  - The model exceeded the predefined goal of specificity ≥75% with sensitivity ≥50% in the original testing set and the separate validation set
- The present model contrasts with the algorithm of Stephansen et al.,[2] which performed strongly in a research study population but did not translate to a real-world clinical population
- Results of this study support efforts to develop an ML-based algorithm using nocturnal PSG in general sleep medicine clinics that can offer an objective, sensitive, and specific tool for alerting sleep clinicians about patients at risk for narcolepsy

References: 1. Thorpy MJ, Krieger AC. *Sleep Med.* 2014;15:502-7. **2.** Stephansen JB, et al. *Nat Commun.* 2018;9:5229. **3.** Moore IV H, et al. Presented at: World Sleep; March 11-16, 2022; Rome, Italy (poster 118). **4.** Mitler MM, et al. *Electroencephalogr Clin Neurophysiol.* 1979;46:479-81. **5.** Richardson GS, et al. *Electroencephalogr Clin Neurophysiol.* 1978;45:621-7.

Support and Acknowledgments: This study was supported by Jazz Pharmaceuticals. Under the direction of the authors, Shivani Vaidya, PharmD, 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 and works with and receives financial support from Jazz Pharmaceuticals on this work. **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.