Use of Artificial Intelligence for Early Characterization of Patients with RBD

Sam Rusk, BS1, Joseph R. Cataline, MD2, Yoav N. Nygate, MS1, Jiansan Gu, MD2, Natalie Witek, MD2, Chris R. Fernandez, MS1, Nathaniel F. Watson, MD, MS3, James J. Herdegen, MD2

1 EnsoData Research, Ensodata, Madison, WI
2 Rush University Medical Center, Chicago, IL
3 Department of Neurology, University of Washington School of Medicine, Seattle, WA

Introduction

What is REM Behavior Disorder (RBD)?

RBD is a chronic sleep disorder where patients exhibit violent behavior during sleep due to the loss of muscle atonia that typically occurs during REM sleep. RBD can be considered idiopathic, or manifest in individuals using antidepressant, those with obstructive sleep apnea (OSA), and in individuals with other sleep disorders such as narcolepsy [1].

Prevalence of RBD:

RBD has a prevalence of 0.38% to 2% among individuals over 60 years old [2]. The prevalence of polysomnography (PSG) confirmed RBD has been estimated at 0.68% of the general population, and that of probable RBD at 5.65%. Idiopathic RBD is considered a pre-clinical marker of neurodegeneration with strong predictive value. Large longitudinal cohort studies demonstrate 81–91% of idiopathic RBD patients, followed for ≥14-years, will develop a definite neurodegenerative disease or mild cognitive impairment.

Diagnosis of RBD:

The diagnosis of RBD typically involves a combination of clinical evaluation and PSG with video monitoring. PSG is necessary to confirm the diagnosis by demonstrating the loss of muscle atonia during REM sleep. Additional tests, such as neuroimaging and neurological evaluation, may be needed to rule out other causes of RBD-like symptoms [1]. High prevalence of OSA exists in patients with RBD, which can confound the analysis of EMG tone and arousal activity. Accurate diagnosis of RBD is crucial for predicting the likelihood of developing neurodegenerative disease and implementing appropriate treatment strategies.

Study Objective:

Considering the complexities and potential confounders in diagnosing RBD, there is a need for straightforward, accessible, reliable, and affordable diagnostic approaches. This study investigates different techniques for automated RBD detection using single-night PSG, and compares the performance on patients with and without OSA. The performance of these methods is demonstrated on a dataset collected from Rush University Medical Center.

The Dataset

- RBD Dataset:
  - PSG sleep studies collected from Rush University Medical Center.
  - N=69 RBD patients:
    - N=44 OSA, N=25 Non-OSA
  - N=326 negative controls:
    - N=261 OSA, N=65 Non-OSA
  - Patient demographics, PSG characteristics, and co-morbid conditions were reviewed.
- Historical Database:
  - Historical database of over 1 million sleep studies
  - N=50,000 PSG sleep studies sampled from over 300 clinics across the U.S. and 10 different recording devices, used to train a sleep staging model.

Methodology

Models:

- Hypnodensity Random Forest (Hypno-RF):
  - We trained a sleep staging model using the N=50,000 dataset.
  - We then extracted the hypnodensity [3] using the output probabilities of the trained sleep staging model.
  - We trained a random forest model using 400 hand engineered features (as defined in reference [4]) extracted from the hypnodensity.
- PSG Sleep Report Data Random Forest (Sleep-RF):
  - Trained a random forest model using 15 different PSG-based and sleep-based report data calculated for each subject in the RBD dataset.
- PSG-EEG Based Deep Learning Model (PSG-DL):
  - Trained a deep learning model on raw EEG signals derived from the RBD dataset

Evaluation Methods:

- All models were trained using 10-fold cross-validation
- Performance was evaluated using area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, and F1-Score.
- When possible, we evaluated feature importance using the Gini Index.
- We also performed statistical analysis on the report data variables.

Results

RBD Detection Performance

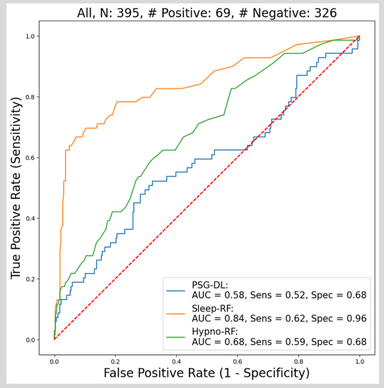

Figure 1. ROC curve comparing the performance of each model. Sensitivity and specificity were calculated by choosing the threshold that maximizes the F1-Score.

RBD Detection Performance on OSA Patients

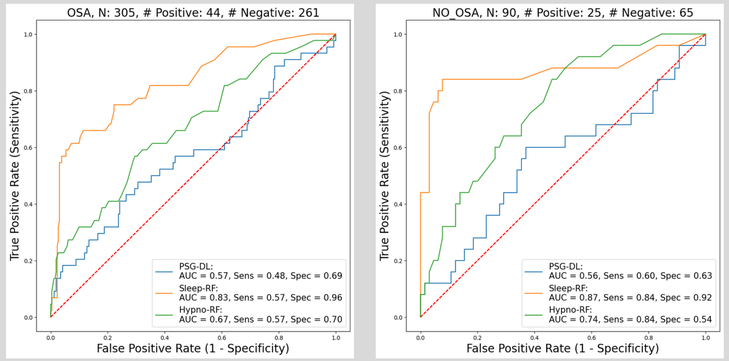

Figure 2. ROC curves comparing the performance of each model and demonstrates the difference in performance between patients with OSA and patients without OSA.

Feature Importance Analysis for Sleep-RF

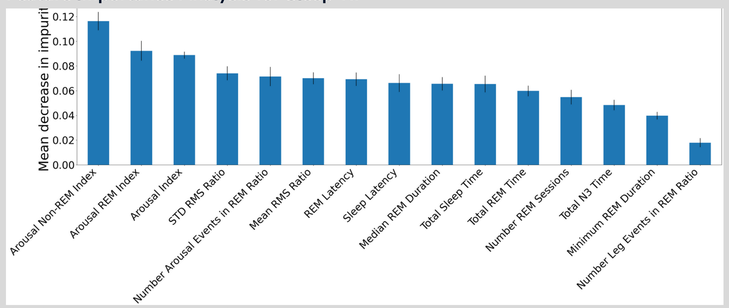

Figure 3. Feature importance analysis for the Sleep-RF model. The values in the y-axis represent the importance of each feature. The higher the mean decrease in impurity is, the more important the feature is to the overall prediction. RMS ratios compare between the level of chin tone in REM vs non-REM sections, which is hypothesized to be associated with the level of atonia in patients with and without RBD.

PSG Report Data Variables Statistical Analysis

Feature Name	Positive Population	Negative Population	OLS P-Value	OLS Coefficient	Gini
Number Arousal Events in REM Ratio	0.073 ± 0.282	0.108 ± 0.204	0.001884	-0.841861	0.072
STD RMS Ratio	1.275 ± 1.746	0.862 ± 1.856	0.008487	0.06169	0.074
Minimum REM Duration	0.078 ± 0.316	0.112 ± 0.37	0.035908	0.462131	0.04
Number REM Sessions	6.188 ± 7.812	4.482 ± 6.136	0.039539	0.01971	0.055
Mean RMS Ratio	0.899 ± 1.016	0.729 ± 0.648	0.041311	0.121314	0.07
Arousal REM Index	8.781 ± 32.34	16.018 ± 32.002	0.041677	-0.006266	0.092
Median REM Duration	0.172 ± 0.332	0.24 ± 0.39	0.069918	-0.469911	0.066
Arousal Non-REM Index	2.679 ± 10.802	3.493 ± 7.416	0.098039	0.084289	0.116
Arousal Index	1.974 ± 7.722	2.735 ± 5.572	0.304021	-0.078111	0.089
Total REM Time	1.025 ± 1.07	0.998 ± 1.19	0.350448	-0.07033	0.06
Sleep Latency	1.116 ± 2.09	1.258 ± 1.418	0.405768	-0.020864	0.066
REM Latency	2.25 ± 3.62	2.562 ± 3.72	0.542636	-0.008406	0.069
Total N3 Time	0.536 ± 0.898	0.649 ± 0.868	0.557272	-0.026277	0.048
Total Sleep Time	5.489 ± 2.414	5.535 ± 2.06	0.801469	0.006306	0.065
Number Leg Events in REM Ratio	0.034 ± 0.206	0.041 ± 0.328	0.81402	0.029094	0.018

Table 1. Ordinary least squares (OLS) summary. We ran all variables through an OLS model where each time one variable was varied while all other variables were controlled for.

Conclusions

- Sleep-RF outperformed all other methods . This demonstrates the potential of utilizing simple PSG report data along with machine learning as a screening tool for the detection of RBD.
- The lower performance of the other two approaches might be attributed to the dataset size. Since both approach depend on either more complex features or more sophisticated machine learning methods, a larger dataset is usually required in order to observe an increase in performance compared to simpler methods.
- When comparing between the performance on patients with OSA and patients without OSA, it is clear that the model's performance is superior when tested on the non-OSA population. We hypothesize the performance difference attributable to similarity and additional complexity in sleep disturbance characteristics between RBD and OSA.
- The feature importance analysis indicated that the level of sleep fragmentation plays an important role in the detection of RBD, as expected. Furthermore, the chin tone of the patients during REM surfaced as an important feature which is expected when differentiating between patients with and without RBD.
- The statistical analysis resulted in a statistically significant association between chin tone and the presence of RBD, where patients with RBD often demonstrated a higher chin tone compared to non-RBD patients. In addition, patients with RBD displayed larger number of REM sessions and less arousal events in REM.

Future Work

- Future work will be dedicated to the collection of more data, which might unlock the performance increase that is expected from the more advanced machine learning methods.
- More development is needed in order to isolate the similarities between RBD patients and OSA patients, to allow for a more robust RBD detector that is able to distinguish between patients with and without RBD in both OSA and non-OSA populations.
- Broad implementation of AI methods show potential to expand early detection and diagnosis of RBD. Future work should be dedicated to the testing of this detector on large, real-world data, and asses its potential for the detection of neurodegenerative conditions through expanded analysis, and follow-up.
- Additional research is underway to help improve accuracy and ensure these methods are generalizable across platforms and clinical datasets.

References

[1] Sateia, M. J. (2014). International classification of sleep disorders-third edition: highlights and modifications. Chest, 146(5), 1387-1394. doi: 10.1378/chest.14-0970. PMID: 25367475

[2] Dauvilliers, Y., et al. (2018). REM sleep behaviour disorder. Nature Reviews Disease Primers, 4, 19. https://doi.org/10.1038/s41572-018-0016-5

[3] Stephansen, Jens B., et al. "Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy." Nature communications 9.1 (2018): 5229.