SLEEP, 2024, XX, 1–10
https://doi.org/10.1093/sleep/zsae031
Advance access publication 8 February 2024
Original Article

Original Article

EEG-based machine learning models for the prediction of phenoconversion time and subtype in isolated rapid eye movement sleep behavior disorder

El Jeong[1,‡], Yong Woo Shin[2,‡], Jung-Ick Byun[3,*], Jun-Sang Sunwoo[4], Monica Roascio[5,6], Pietro Mattioli[7,8], Laura Giorgetti[7], Francesco Famà[7,8], Gabriele Arnulfo[5,6], Dario Arnaldi[7,8], Han-Joon Kim[2] and Ki-Young Jung[9,10,*]

[1]Interdisciplinary Program in Bioengineering, College of Engineering, Seoul National University, Seoul, South Korea,
[2]Department of Neurology, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, South Korea,
[3]Department of Neurology, Kyung Hee University Hospital at Gangdong, Seoul, South Korea,
[4]Department of Neurology, Kangbuk Samsung Hospital, Seoul, South Korea,
[5]Department of Informatics, Bioengineering, Robotics and System engineering (DIBRIS), University of Genoa, Genoa, Italy,
[6]RAISE (Robotics and AI for Socio-economic Empowerment) Ecosystem, Genoa, Italy,
[7]Department of Neuroscience (DINOGMI), University of Genoa, Genoa, Italy,
[8]Neurophysiopathology Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy
[9] Seoul National University Hospital, Seoul, South Korea, and
[10]Seoul National University Medical Research Center Neuroscience Research Institute, Sensory Organ Research Institute, Seoul National University College of Medicine, Seoul, South Korea

[‡]These authors equally contributed to this work.
[*]Corresponding authors. Jung-Ick Byun, Department of Neurology, Kyung Hee University Hospital at Gangdong, 892, Dongnam-ro, Gangdong-gu, Seoul, 05278, South Korea. Email: jbyun@khu.ac.kr; Ki-Young Jung, Department of Neurology, Seoul National University Hospital, Sensory Organ Research Institute, Seoul National University Medical Research Center Neuroscience Research Institute, Seoul National University College of Medicine, 101, Daehak-ro, Jongno-gu, Seoul, 03080, South Korea. Email: jungky@snu.ac.kr.

Abstract

Study Objectives: Isolated rapid eye movement sleep behavior disorder (iRBD) is a prodromal stage of α-synucleinopathies and eventually phenoconverts to overt neurodegenerative diseases including Parkinson's disease (PD), dementia with Lewy bodies (DLB), and multiple system atrophy (MSA). Associations of baseline resting-state electroencephalography (EEG) with phenoconversion have been reported. In this study, we aimed to develop machine learning models to predict phenoconversion time and subtype using baseline EEG features in patients with iRBD.

Methods: At baseline, resting-state EEG and neurological assessments were performed on patients with iRBD. Calculated EEG features included spectral power, weighted phase lag index, and Shannon entropy. Three models were used for survival prediction, and four models were used for α-synucleinopathy subtype prediction. The models were externally validated using data from a different institution.

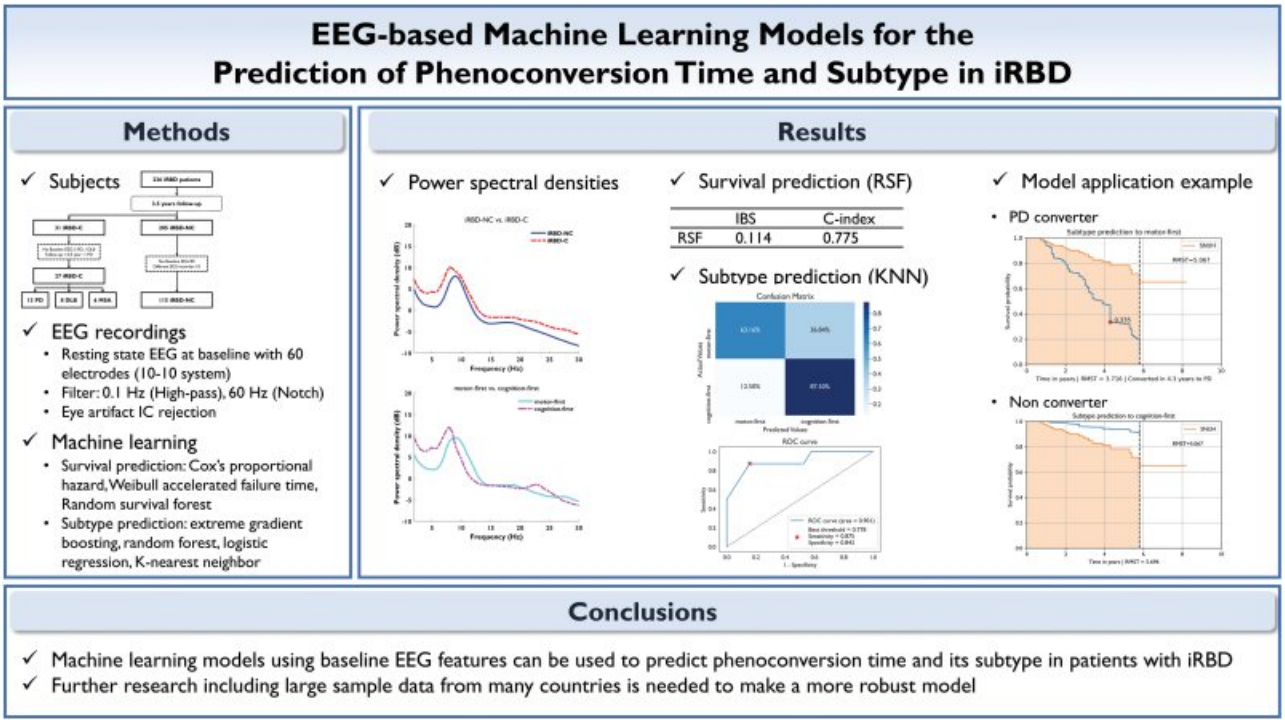
Results: A total of 236 iRBD patients were followed up for up to 8 years (mean 3.5 years), and 31 patients converted to α-synucleinopathies (16 PD, 9 DLB, 6 MSA). The best model for survival prediction was the random survival forest model with an integrated Brier score of 0.114 and a concordance index of 0.775. The K-nearest neighbor model was the best model for subtype prediction with an area under the receiver operating characteristic curve of 0.901. Slowing of the EEG was an important feature for both models.

Conclusions: Machine learning models using baseline EEG features can be used to predict phenoconversion time and its subtype in patients with iRBD. Further research including large sample data from many countries is needed to make a more robust model.

Key words: REM sleep behavior disorder; Parkinson's disease; dementia with lewy bodies; multiple system atrophy; phenoconversion; EEG; machine learning

Submitted for publication: November 5, 2023; **Revised:** January 20, 2024

Graphical Abstract



Statement of Significance

This study's significance lies in its focus on early α-synucleinopathy detection in patients with Isolated rapid eye movement sleep behavior disorder (iRBD), which can evolve into debilitating neurodegenerative diseases like Parkinson's disease, Dementia with Lewy bodies, and Multiple system atrophy. Using baseline EEG data, machine learning models were developed to predict phenoconversion onset and α-synucleinopathy subtype. The random survival forest model effectively predicted survival outcomes, while the K-nearest neighbor model identified subtypes. Notably, EEG slowing played a pivotal role in these models. This research highlights the potential for early intervention and improved patient care by identifying high-risk individuals. However, additional research with larger, diverse datasets is crucial to strengthen model robustness and applicability.

Introduction

Rapid eye movement sleep behavior disorder (RBD) is characterized by dream enactment and loss of atonia during rapid eye movement sleep [1]. Isolated RBD (iRBD) is known as a prodromal stage of α-synucleinopathies, specifically Parkinson's disease (PD), dementia with Lewy bodies (DLB), and multiple system atrophy (MSA) [2, 3]. The risk of developing α-synucleinopathy among patients with iRBD is approximately 18% after 3 years, 31% after 5 years, and 74% after 12 years [2]. In short, most iRBD patients eventually develop an α-synucleinopathy, preceded by a decline in either motor or cognitive function [4].

Age, olfactory function, cognitive function, and motor function have been reported as clinical biomarkers to predict phenoconversion with a hazard ratio of up to 3.16 [5, 6]. Neuroimaging of dopamine transporters provides a promising biomarker to predict phenoconversion [7]; however, it is expensive and has limited accessibility. Electroencephalography (EEG) is a safe and easy method to objectively measure brain activity. In patients with iRBD, EEG findings have demonstrated a slowing pattern in the occipital region, a reduction in delta-band weighted phase lag index (wPLI) in frontal region, and an elevation in alpha wPLI accompanied by a decrease in delta orthogonalized Correlation Coefficient (oCC) [8–11]. Two longitudinal studies have evaluated the value of EEG in predicting phenoconversion in iRBD [12, 13]. These studies, however, described phenoconversion dichotically, i.e. converted or not converted.

However, it is demanded for both patients and clinicians to predict *when and how fast* phenoconversion to α-synucleinopathy will occur and *to which subtype* of synucleinopathy the patient will phenoconvert. An individualized model that can predict the time from diagnosis until each patient develops a α-synucleinopathy is important in understanding their prognosis. Moreover, estimating whether the iRBD patients will first develop motor or cognitive symptoms is also crucial.

Machine learning is being extensively explored for potential applications in various diseases and has achieved excellent performance compared with conventional methods [14]. Thus, machine learning methods can be considered for the application of predicting the complex survival time and subtype of iRBD phenoconversion. The aim of this study was to propose EEG-based machine learning models that can predict the time to phenoconversion and the subtype of phenoconversion for each patient. Various survival analyses and classification models were compared to select the best model.

Materials and Methods

Participants

Patients with iRBD who visited the sleep clinic of Seoul National University Hospital were enrolled and followed up every year. RBD was diagnosed according to the International Classification of Sleep Disorders—Third Edition (ICSD-3) criteria using overnight video-polysomnography (vPSG) [15]. Two neurologists specialized in sleep disorders (JK) and movement disorders (KH) examined each patient at baseline to evaluate them for dementia, cerebellar ataxia, parkinsonism, or other neurodegenerative diseases.

Participants with a neurodegenerative disease, neurological disorder, severe medical illness, or severe obstructive sleep apnea (apnea–hypopnea index ≥30) were excluded. This study was authorized by the Institutional Review Board (IRB) of the Seoul National University Hospital (IRB Number 1406-100-589). Written informed consent was obtained from each participant.

For external validation, we used clinical and EEG data of patients with iRBD provided by the University Neurology Clinics at Policlinico San Martino in Genoa. Clinical and EEG data have been described in detail elsewhere [11]. In brief, inclusion/exclusion criteria, clinical and EEG assessments substantially overlapped with the Korean cohort. All patients completed routine clinical follow-ups during which systematic assessments for parkinsonism and dementia were performed, including a semistructured interview with patients and caregivers (IRB Number 703, from the Genoa IRB).

Clinical evaluation

The Korean version of the Mini-Mental Status Examination (K-MMSE) and the Korean version of the Montreal Cognitive Assessment (MoCA-K) were used to evaluate general cognitive function [16, 17]. The Korean version of the RBD Screening Questionnaire-Hong Kong (RBDQ-KR) was used to assess the RBD symptom severity [18]. The Korean Version of Sniffing Sticks (KVSS) was applied to test olfactory symptoms [19]. The Scales for Outcomes in Parkinson's Disease for Autonomic Symptoms (SCOPA-AUT) questionnaire was used to examine the symptoms of autonomic dysfunction [20]. The Movement Disorder Society—Unified Parkinson's Disease Rating Scale (MDS-UPDRS) part III was used to assess motor symptoms [21]. Additionally, subjective sleep quality and excessive daytime sleepiness were assessed using the Pittsburgh Sleep Quality Index (PSQI) and the Epworth Sleepiness Scale (ESS), respectively [22, 23].

During follow-up, cognitive function (K-MMSE, MoCA-K), motor function (MDS-UPDRS part III), autonomic function (SCOPA-AUT), self-reported sleep propensity and quality (PSQI, ESS), RBD symptom severity (RBDQ-KR), and olfactory function (KVSS) were evaluated every year. Phenoconversion in patients with iRBD was assessed every 6 to 12 months by the same two neurologists (JK and KH). Finally, patients with iRBD who developed PD, DLB, or MSA were classified as converters (iRBD-C), while the remaining patients were classified as non-converters (iRBD-NC). The diagnoses of PD, DLB, and MSA were made according to standard criteria [24–26].

EEG recordings and preprocessing

Scalp EEGs were obtained using a 60-channel EEG cap (Wave-Guard EEG cap, Advanced NeuroTechnology, Enschede, Netherlands) arranged according to the international 10–10 system. The reference electrode was positioned on an ear and the ground electrode was placed on the AFz. Impedances were kept under 10 kΩ. To detect and eliminate eye movement artifacts, two EOG electrodes were attached to the left and right outer canthi. The sampling rate was 400 Hz. The resting-state EEG was recorded for a total of 5 minutes in all patients while they were awake and alternating opening and closing their eyes every 30 seconds. To preprocess the data, a 0.5 Hz high-pass filter and a 60 Hz notch filter were applied. Only the EEG data recorded while the participant's eyes were closed were extracted and analyzed in this study. EEG segments with severe artifacts or poor signal quality were removed by visually inspecting the data. Then, independent component analysis was applied, and the EEGLAB plugin ICLabel was used to automatically remove eye artifacts [27, 28]. The threshold for eye artifact probability was set to 90%.

EEG data for the external validation set used a system with 61 electrodes according to the international 10–10 system. The reference and the ground electrode were Fpz and Oz, respectively, and the signals were sampled at 512 Hz. We simultaneously recorded electrooculogram to monitor eye movements. The acquisition protocol consisted of approximately 25 resting state recordings subdivided into 2–3 minutes with eyes open, 3–4 minutes during hyperventilation, and 17–18 minutes with eyes closed. We conducted the extraction of a total of 100 windows from the pre-hyperventilation segment, thoroughly focusing on the region that remained unaffected by the ensuing hyperventilation. Impedances were kept under 5 kΩ. The same preprocessing procedures used for our dataset were implemented for the external validation set.

For both centers' data, a total of 101 seconds of EEG data for each patient were eventually included in this study. Preprocessing of EEG data were performed using the EEGLAB package (version 2019.1) for MATLAB (version 9.8.0, The MathWorks, Natick, MA, USA) [29].

Experimental procedures

EEG features.

To make the model more robust and reduce overfitting, data augmentation were performed for the training set. To augment the total data size the first 100 2-second EEG epochs were extracted by the sliding window method with 50% overlap. Thus, one patient's EEG data was augmented to one hundred EEG epochs.

For each EEG epoch, the fast Fourier transforms using the Hanning window were applied with a frequency of interest range of 1–50 Hz in 0.5 Hz steps. In our study, four frequency bands were used: delta (2–3.5 Hz), theta (4–7.5 Hz), alpha (8–12.5 Hz), and beta (13–30 Hz). Absolute power was averaged across all electrodes and converted decibels. Relative power was calculated by expressing the percentage of power for each frequency band over the total power in the 2–30 Hz range. The dominant occipital frequency (DOF) was defined as the average peak frequency with the maximum power between 4 and14 Hz in the two occipital channels (O1, O2). The slow-to-fast power ratio (STF) was calculated using the absolute power values averaged for all electrodes as follows: [(delta + theta)/(beta)]. Recently it has been suggested that metrics describing network interactions of large-scale inter-areal synchronization between brain oscillations could improve classification accuracy in iRBD patients [30]. Accordingly, overall functional connectivity for each frequency band was extracted by averaging the wPLI values of all 1770 electrode pairs [8, 31]. Furthermore, Shannon entropy (SE) was defined with 10 bins of amplitude values [32]. In total, 15 EEG features were calculated for analysis (Supplementary Table S1). All spectral analyses were performed using the FieldTrip toolbox version 20200607 [33].

Modeling process (1) - prediction of phenoconversion time.

Eighty percent of the overall dataset was assigned as the training set, and the remaining 20% of the data were designated as

the testing set. All patients with iRBD data, which were divided into patients living with iRBD not phenoconverted at follow-up (iRBD-NC) and patients with phenoconverted at follow-up (iRBD-C), were used in this survival prediction analysis. All duration information was calculated by the difference between the EEG acquisition date and the last visit date of the patient. To identify the most relevant features for predicting phenoconversion time, a two-step feature selection process was used. First, univariable Cox proportional hazard (CPH) regression was performed to identify features with a p-value less than 0.1. Next, backward multivariable CPH regression was performed to eliminate features with a p-value greater than 0.1. To address the imbalanced nature of the data, with a larger number of nonconverter samples, the synthetic minority oversampling technique (SMOTE) was applied to the training set [34]. The CPH, Weibull-accelerated failure time (wAFT), and random survival forest (RSF) models were used to train and test these data. To evaluate the models, stratified group 5-fold cross-validation was implemented for internal validation. During the training process, stratification was employed specifically for iRBD-NC and iRBD-C subgroups. Harrel's concordance index (C-index) and the integrated Brier score (IBS) were used to evaluate the performance of survival prediction analysis [35, 36]. The C-index is a commonly used metric in survival analysis that measures the ability of the model to correctly rank the survival times of patients. The IBS, on the other hand, is a measure of the overall accuracy of the model's predictions, considering both the predicted survival times and the observed survival times. For all models, hyperparameter optimization was performed using the training set while the testing set was used for model performance. The final model was fitted with the augmented and resampled total dataset by using the best prediction model. As there were no patients with MSA in the external validation set, we additionally fitted the model excluding the patients with MSA. Permutation feature importance was employed to represent the importance of each variable in the model. All analyses were conducted using Python 3.8.5 (scikit-learn: v.1.1.1; lifelines: v.0.27.0; scikit-survival: v.0.18.0; hyperopt: v.0.2.7).

Modeling process (2) - prediction of phenoconversion subtype.

Only iRBD-C data were used for subtype prediction analysis to classify subtypes of phenoconversion. A training set made up of 80% of the dataset and a testing set using the remaining 20% were assigned. Following a previous study, we classified phenoconversion into two subtypes according to the first presenting symptom: motor-first subtype, which includes PD and MSA, and cognition-first subtype, which includes DLB [13]. Recursive feature elimination was applied using multiple models to select the most relevant and predictive features for subtype prediction. The features selected by recursive feature elimination using XGBoost showed the best performance and were therefore used for the models. The other models were trained and tested on the selected features. Due to data imbalance, the SMOTE was applied in the training set. The data were trained and tested using the XGBoost, random forest, logistic regression with elastic net regularization, and k-nearest neighbor (KNN) models. Hyperparameter optimization was performed using the training set for all models while evaluating the testing set. A 10-fold cross-validation was performed 10 times for internal validation to obtain a robust estimate of the performance of the models. The area under the receiver operating characteristic curve (AUC), accuracy, precision, recall, and F1 score were utilized to evaluate the performance of the subtype prediction analysis. The best prediction model was used to fit the final model to the augmented and resampled entire dataset. For the same reason described above for survival prediction, a dataset without data from MSA patients was additionally analyzed. Classification into PD, MSA, and DLB was also performed. Python 3.8.5 was used to conduct each analysis (scikit-learn: v.1.1.1; xgboost: v.0.90; hyperopt: v.0.2.7).

Statistical analysis

All data are shown as the mean ± standard deviation [range]. The Kolmogorov–Smirnov test was used to test the normality of all variables before analysis. Independent sample t-tests were employed to evaluate differences in continuous data. Categorical data were analyzed with Fisher's exact test. Nonnormally distributed variables were compared using the Mann–Whitney U test. Survival curves were plotted using the Kaplan–Meier method. The log-rank test was used to compare survival distributions between our dataset and the external validation dataset. The Restricted Mean Survival Time (RMST) is obtained to estimate the average survival time up to 5.8 years, which is the last observed event time in our dataset [37]. The significance threshold was set to 0.05. All statistical evaluations were performed with Python 3.8.5 using the SciPy library (scipy: v.1.5.2).

Results

Participant characteristics

A total of 236 iRBD patients were included in the internal validation dataset. During a mean follow-up duration of 3.5 years (range: 0.9–8.6 years), 31 patients converted to overt neurodegenerative diseases, and 205 remained in an isolated state of RBD (Figure 1A). The mean time to phenoconversion was 2.66 ± 1.48 years. Eighty-three patients with no baseline EEG data, 1 patient who phenoconverted within 6 months after EEG acquisition, and 10 patients whose data were recorded with another EEG system were excluded from further analysis. Of the remaining 142 patients, 27 patients were phenoconverted during follow-up (13 to PD, 8 to DLB, and 6 to MSA). All patients with MSA were of the cerebellar type (MSA-C).

There were no significant differences in sex, RBDQ-KR, K-MMSE, KVSS, SCOPA-AUT, ESS, and PSQI between the iRBD-NC and iRBD-C groups. However, the patients in the iRBD-C group were older, had lower education levels, lower MoCA-K scores, and higher MDS-UPDRS-III scores (Table 1). When comparing motor- and cognition-first subtypes, the cognition-first group was older and had lower baseline K-MMSE and MoCA-K scores (Supplementary Table S2).

The external validation dataset included 62 patients with iRBD who were followed up for 2.17 ± 1.53 years (Supplementary Table S3). Seven patients were excluded: five because of poor data quality, 1 because the data does not have conversion date information and 1 because the data were recorded after phenoconversion. Seventeen of the iRBD patients in the external validation dataset were phenoconverted during follow-up (7 to PD and 10 to DLB). Compared to our dataset, patients in the external validation dataset were older, were more likely to be male, and had lower MMSE scores (Supplementary Table S4). In addition, the log-rank test showed that the two datasets showed significantly different survival rates (Figure 1B, $p < 0.005$). The RMST of our dataset was 5.067 years and the RMST of each patient was calculated from the survival curve. In addition, a positive correlation was shown between the RMST and time to conversion for the converted

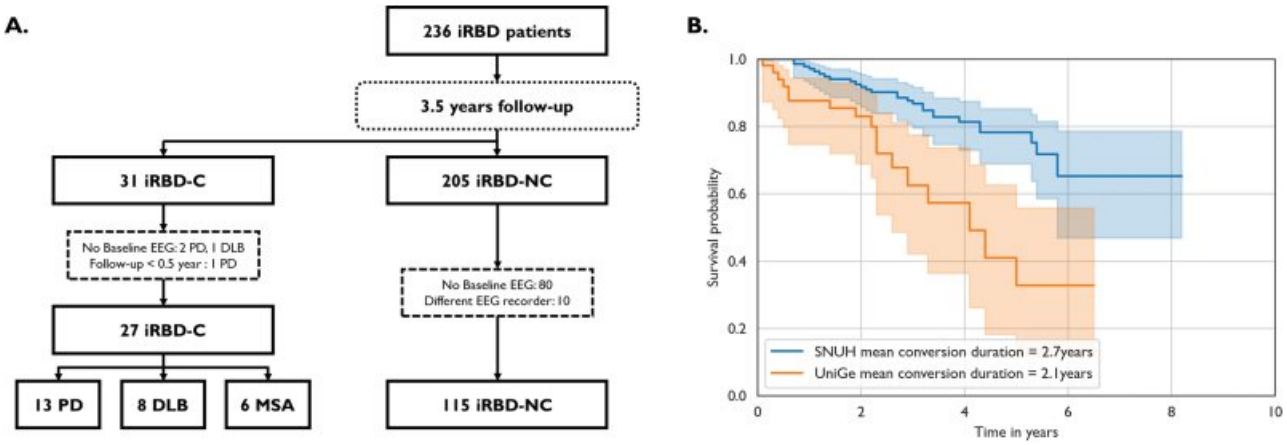


Figure 1. Flowchart and survival curve. (A) Flowchart. (B) Survival curves of Seoul National University Hospital and University of Genoa. Abbreviations: iRBD, isolated REM sleep behavior disorder; iRBD-C, iRBD converters; iRBD-NC, iRBD non-converters; PD, Parkinson's disease; DLB, dementia with Lewy bodies; MSA, multiple system atrophy; SNUH, Seoul National University Hospital; UniGe, University of Genoa.

Table 1. Participant Characteristics of Patients With iRBD Who Further Converted or Not From Seoul National University Hospital

	iRBD-NC (*n* = 115)	iRBD-C (*n* = 27)	*P* value
Age (y)	66.61 ± 6.44 [50–82]	69.85 ± 7.30 [57–82]	0.023
Sex (Male %)	M: 75, F: 40 (65.2)	M: 15, F: 12 (55.6)	0.380[a]
Education (y)	12.83 ± 4.08 [0–18]	10.89 ± 4.28 [0–18]	0.029[b]
RBDQ-KR	49.74 ± 19.88 [4–100] (*n* = 106)	48.26 ± 16.20 [5–70] (*n* = 23)	0.740
Conversion duration (y	—	2.66 ± 1.48 [0.7–5.8]	
K-MMSE	27.77 ± 1.77 [21–30]	27.15 ± 2.21 [20–30]	0.166[b]
MoCA-K	25.83 ± 2.79 [16–30]	23.15 ± 4.64 [7–29]	<0.001
KVSS	17.43 ± 5.32 [7–31] (*n* = 104)	17.74 ± 6.08 [7–27] (*n* = 21)	0.815
SCOPA-AUT	12.57 ± 7.12 [1–30] (*n* = 107)	14.91 ± 9.08 [2–39] (*n* = 23)	0.176
MDS-UPDRS-III	0.91 ± 1.95 [0–11] (*n* = 97)	2.35 ± 2.85 [0–8] (*n* = 17)	0.008[b]
ESS	5.62 ± 3.50 [0–16]	5.96 ± 4.19 [1–20]	0.658
PSQI	7.02 ± 4.23 [1–18]	6.04 ± 4.12 [1–18]	0.274[b]

Italics font indicates statistical significance. Abbreviations: iRBD, isolated REM sleep behavior disorder; iRBD-NC, iRBD nonconverters; iRBD-C, iRBD converters; RBDQ-KR, Korean version of the RBD screening Questionnaire-Hong Kong; K-MMSE, Korean version of the Mini-Mental Status Examination; MoCA-K, Korean version of the Montreal Cognitive Assessment; KVSS, Korean Version of Sniffing Sticks; SCOPA-AUT, Scales for Outcomes in Parkinson's Disease for Autonomic Symptoms; MDS-UPDRS-III, Movement Disorder Society—Unified Parkinson's Disease Rating Scale Part III; ESS, Epworth Sleepiness Scale; PSQI, Pittsburgh Sleep Quality Index.
[a]Fisher's exact test.
[b]Mann–Whitney U test.

patients (Pearson correlation $r = 0.525$, $p = 0.005$, Supplementary Figure S1).

Phenoconversion time prediction

Delta wPLI was excluded through univariable CPH regression, and DOF, relative delta power, relative beta power, and SE were further excluded through multivariable CPH regression. Finally, 10 features were included in this survival prediction analysis. The power spectral densities of iRBD-NC and iRBD-C are shown in Figure 2B.

We compared the three survival analysis methods using our dataset (Supplementary Table S5). For the internal validation using 5-fold cross-validation, the RSF model was the best, with an IBS of 0.114 and a C-index of 0.775. The five most important features of RSF were absolute theta power, absolute delta power, STF, beta wPLI, and absolute alpha power (Figure 2A and Supplementary Table S6). The iRBD-C group showed higher absolute delta power and absolute theta power but also higher alpha power than the iRBD-NC group. For the external validation dataset, the RSF model showed an IBS of 0.128 and a C-index of 0.561.

Additionally, model results excluding data from patients with MSA are listed in Supplementary Table S7.

Phenoconversion subtype prediction

Through recursive feature elimination, eight features were excluded due to their low feature importance. As a result, seven features were used in this subtype prediction analysis. The selected features in subtype prediction were DOF, STF, absolute theta power, absolute beta power, relative beta power, beta wPLI, and SE (Supplementary Table S8). The power spectral densities of the motor- and cognition-first subtypes are shown in Figure 2C.

The scores for internal validation of motor- and cognition-first are shown in Table 2. For internal validation, the KNN model's performance was the best among the models, with an AUC of

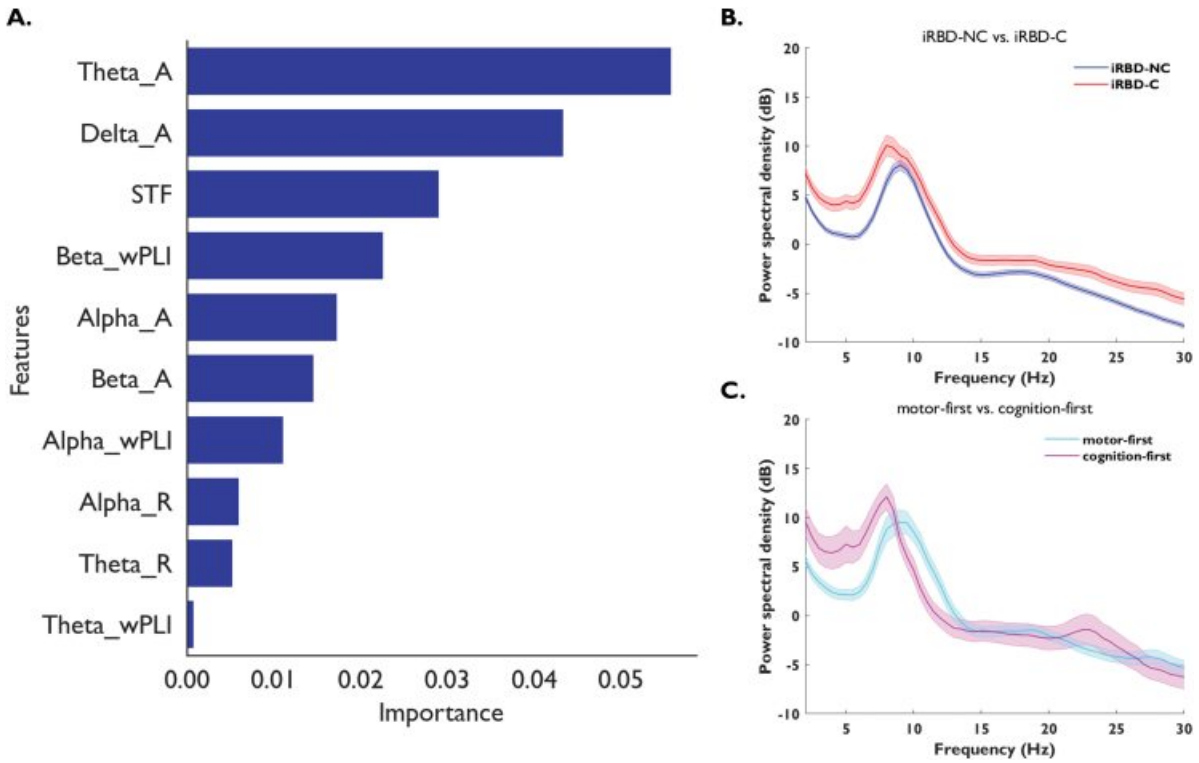


Figure 2. (A) Feature importance of random survival forest model and (B) comparison of power spectral densities with standard error between the iRBD-NC and iRBD-C groups, and (C) comparison of motor-first and cognition-first presentations. Abbreviations: _A, absolute power; STF, slow-to-fast power ratio; wPLI, weighted phase lag index; _R, relative power; SE, Shannon entropy; iRBD, isolated REM sleep behavior disorder; iRBD-C, converters to a neurodegenerative disorder from iRBD; iRBD-NC, nonconverters from iRBD.

Table 2. Subtype Prediction Results

	AUC	Accuracy	Precision	Recall	F1
XGBoost	0.717	0.741	0.556	0.625	0.588
RF	0.737	0.704	0.500	0.750	0.600
LR	0.770	0.778	0.667	0.500	0.571
KNN	0.901	0.704	0.500	0.875	0.636

Abbreviations: XGBoost, extreme gradient boosting; RF, random forest; LR, logistic regression; KNN, K-nearest neighbor; AUC, area under the receiver operating characteristic curve.

0.901, accuracy of 0.704, precision of 0.500, recall of 0.875, and F1 of 0.636 (Figure 3). External validation of the KNN model resulted in an AUC of 0.536, accuracy of 0.527, precision of 0.304, recall of 0.412, and F1 of 0.350 (Supplementary Table S9).

In addition, evaluation results without data from patients with MSA and classification into PD, MSA, and DLB are shown in Supplementary Tables S10 and S11, respectively. Example plots using both phenoconversion time and subtype prediction models are shown in Supplementary Figures S2 and S3.

Discussion

In this study, we aimed to predict phenoconversion time and subtype in patients with iRBD using resting-state EEG features collected at baseline. Our models, which were based on machine learning algorithms, showed promising results in predicting phenoconversion time and subtype. The RSF model showed acceptable performance in predicting phenoconversion time, while the KNN model was able to predict the conversion subtype (motor-first or cognition-first) with good AUC. Our models may provide a practical solution for predicting individualized phenoconversion time and subtype in patients with iRBD. These predictions are important for better management of the disease and to help patients to be better prepared for their future.

Two previous studies have attempted to predict phenoconversion from iRBD patients using EEG. In the first, researchers used EEG slowing features to predict neurodegeneration in patients with iRBD [13]. The focus was to classify iRBD and whether patients would phenoconvert without considering phenoconversion time or specific subtypes. Later, the same research group used deep learning techniques with EEG spectrograms recorded from patients with iRBD to differentiate them from healthy controls [38]. In contrast, our study aimed to predict not only the phenoconversion time but also the conversion subtype of patients with iRBD. We predicted phenoconversion time in patients with iRBD using only baseline EEG features. Absolute theta power, absolute delta power, beta wPLI, STF, and absolute alpha power were the most important features for phenoconversion time prediction. In previous studies, it was shown that the absolute EEG power of recordings during both sleep and resting state was significantly different not only between iRBD patients and controls but also between patients with iRBD who converted to neurodegenerative diseases and those who had not yet converted [12, 13, 38–41]. In particular, the increases in absolute theta power and delta power were prominent in converted patients. Higher low-frequency power and lower high-frequency power, called EEG slowing, have already been shown in iRBD patients by various neurodegenerative studies [13, 39, 42–44]. Therefore, EEG slowing is known to be common in patients with iRBD, particularly in those who convert to neurodegenerative diseases. Our results demonstrate that EEG features can be applied as biomarkers for predicting phenoconversion time in patients with iRBD.

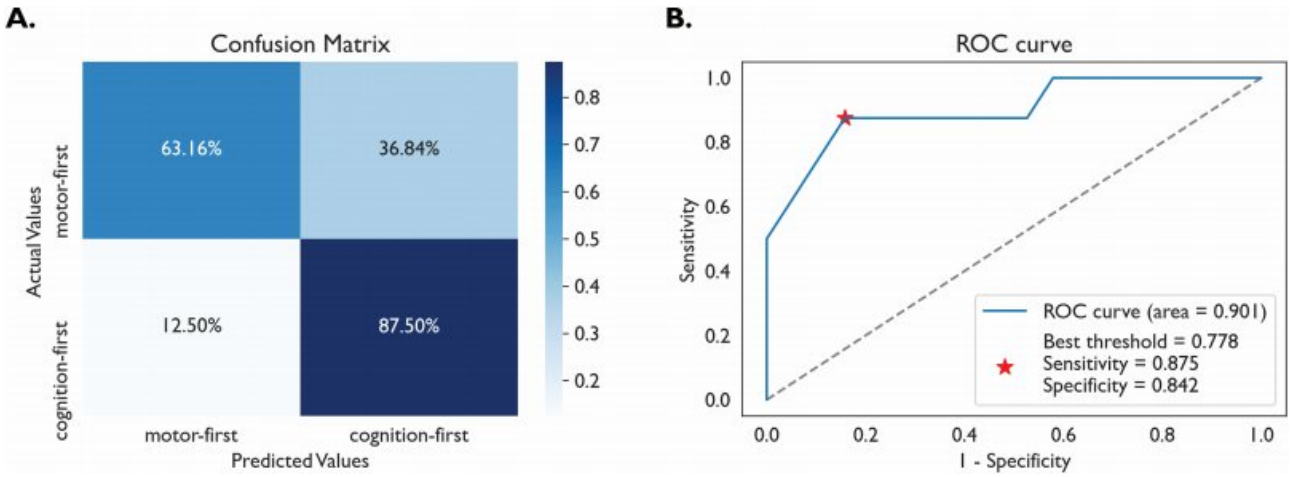


Figure 3. K-nearest neighbor model prediction results. These results are obtained by internal validation using repeated 10-fold cross-validation. (A) Confusion matrix. (B) Receiver operating characteristic curve.Abbreviations: ROC, receiver operating characteristic curve.

Phenoconversion subtype prediction from patients with iRBD was also feasible using baseline EEG features. Differences in EEG between the motor- and cognition-first subtypes had been shown in a previous study. The cognition-first subtype (patients with DLB) showed increased delta and theta power, higher STF, and lower DOF [45, 46]. Indeed, the selected features for subtype prediction in our study were DOF, STF, absolute theta, and beta power, which are consistent with previous studies. Additionally, EEG slowing is correlated with cognitive impairment [47]. In previous studies comparing the motor- and cognition-first subtypes, the main difference at baseline was cognitive function, which was significantly decreased in the cognition-first subtype [2, 48, 49].

In response to the observation regarding subjects with low cognitive scores, we analyzed the association between these scores and prediction accuracy. For the True Positive (TP) cognition-first predictions using the KNN model, the average MoCA-K score was 21.4 ± 2.8, encompassing 7 subjects. In contrast, the single subject with a False Negative (FN) cognition-first prediction had a MoCA-K score of 7. For motor-first predictions, the average MoCA-K score was 24.1 ± 3.1 among 12 subjects with correct predictions, while it was 25.6 ± 3.6 among 7 subjects with incorrect predictions. These findings suggest that lower MoCA-K scores tend to be associated with being predicted in the cognition-first group as compared to the motor-first group. However, it is important to note that our model, primarily based on EEG data, does not incorporate clinical variables including cognitive scores. Therefore, subjects with lower MoCA scores may still be inaccurately predicted in our model.

Additionally, our model demonstrated enhanced efficacy in predicting the cognition-first subtype over the motor-first subtype. This finding aligns with previous research indicating that certain EEG characteristics, such as slowing and increased theta power, are associated with cognitive impairment and the likelihood of future cognitive decline [43, 44]. Notably, these EEG findings have not been similarly correlated with motor dysfunction. This distinction in EEG markers between cognitive and motor impairments provides a plausible explanation for our model's differential predictive performance between these two subtypes.Exclusion of patients with MSA slightly improved the performance of the phenoconversion time prediction model (Supplementary Table S7). The reduced heterogeneity of the sample may have made it easier for the model to identify the relevant features for phenoconversion time prediction. However, it reduced the performance of the KNN model for phenoconversion subtype prediction. The decrease in the number of converters in the dataset from 27 to 21 following the exclusion of 6 MSA patients could have contributed to a significant decrease in the model performance.

It is notable that the performance of external validation was not as good as expected for both survival prediction and subtype prediction. Significantly lower performance for the external validation dataset indicates that our model is overfitted to our dataset. Although we took measures to prevent overfitting by conducting cross-validation and adjusting parameters known to affect overfitting, we could not fully escape overfitting. Compared to the external validation dataset, our dataset showed a higher proportion of females, younger age, lower MDS-UPDRS-III scores, and more MSA-converted patients. Moreover, lower number of participants in the external validation dataset may have also affected the performance of our model. These significant differences between the two cohorts might have contributed to the poor performance in the external validation of this study.

Studies conducted in cohorts from Asian countries have found a slower phenoconversion rate than those from European and American cohorts, suggesting ethnic differences in the prognosis of iRBD [50, 51]. If this is true, ethnic differences should be taken into consideration in the prediction model. Further study is mandatory to confirm any ethnic and/or regional differences in the prognosis of iRBD in the future. Expanding global datasets from multiple clinics is critical for constructing ML/DL models that generalize across varied acquisition situations, including equipment, protocols, geographies, ethnicity, and age.

There are a few limitations to note. First, age and cognitive function scores, which may affect EEG findings were not accounted for [52, 53]. Second, due to the small sample size, we were forced to apply data augmentation. The number of patients with iRBD in this study was 143, which was relatively small for the use of machine learning methods. However, as many studies have used EEG sliding window data augmentation, this method of data augmentation is likely to be reliable enough to achieve the goal of our study [54–56].

In conclusion, we were able to create a useful RSF model and KNN model for predicting the time of phenoconversion and its subtype, respectively, in patients with iRBD simply using resting-state EEG features at baseline. We believe our prediction

model and method contribute to opening new horizons in the management and counseling of patients with iRBD. Furthermore, our model can be implemented in clinical EEG machines or can be developed as a stand-alone device that can be used in outpatient clinics. A future multicenter study with a larger number of patients is needed to elucidate the predictive value of baseline EEG features.

Supplementary Material

Supplementary material is available at *SLEEP* online.

Acknowledgments

This work was supported by the Brain Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2017M3C7A1029688) and the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (2017R1A2B2012280 and NRF-2020R1C1C1013160, 2022R1H1A2092329). This work was partially supported by a grant from the Italian Ministry of Health to IRCCS Ospedale Policlinico San Martino (Fondi per la Ricerca Corrente 2019/2020, and Italian Neuroscience Network (RIN)). This work was developed within the framework of the DINOGMI Department of Excellence of MIUR 2018-2022 (legge 232 del 2016). This work was carried out within the framework of the project "RAISE - Robotics and AI for Socio-economic Empowerment" and has been supported by European Union—NextGenerationEU.

Disclosure Statement

Financial disclosure: none. Nonfinancial disclosure: none.

Author Contribution

El Jeong: Methodology, Formal analysis, Software, Writing—original draft preparation. Yong Woo Shin: Methodology, Formal analysis, Software, Writing—original draft preparation. Jung-Ick Byun: Supervision, Data curation, Validation, Writing—review & editing. Jun-Sang Sunwoo: Data curation, Validation, Investigation. Monica Roascio: Data curation, Validation, Investigation, Writing—review & editing. Pietro Mattioli: Data curation, Validation, Investigation. Laura Giorgetti: Data curation, Validation, Investigation. Francesco Famà: Data curation, Validation, Investigation. Gabriele Arnulfo: Data curation, Validation, Investigation, Writing—review & editing. Dario Arnaldi: Data curation, Validation, Investigation, Writing—review & editing. Han-Joon Kim: Resources, Validation, Investigation. Ki-Young Jung: Supervision, Writing—review & editing, Project administration, Funding acquisition.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Preprint Repositories

A preprint version of this study has been uploaded to the preprint repository medRxiv and can be accessed through the provided DOI: https://doi.org/10.1101/2023.09.04.23294964. The preprint, titled "EEG-based Machine Learning Models for the Prediction of Phenoconversion Time and Subtype in iRBD," was openly released on September 5, 2023, to expedite the early distribution of our research to the broader research community.

References

1. Schenck CH, Bundlie SR, Ettinger MG, Mahowald MW. Chronic behavioral disorders of human REM sleep: a new category of parasomnia. *Sleep*. 1986;**9**(2):293–308. doi: 10.1093/sleep/9.2.293
2. Postuma RB, Iranzo A, Hu M, *et al*. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain*. 2019;**142**(3):744–759. doi: 10.1093/brain/awz030
3. Zhou J, Zhang J, Lam SP, Tang X, Wing YK. Clinical biomarkers of neurodegeneration in REM sleep behavior disorder. *J Sleep Med*. 2015;**12**(2):27–33. doi: 10.13078/jsm.15006
4. Martí MJ, Tolosa E, Campdelacreu J. Clinical overview of the synucleinopathies. *Mov Disord*. 2003;**18**(S6):21–27. doi: 10.1002/mds.10559
5. Miglis MG, Adler CH, Antelmi E, *et al*. Biomarkers of conversion to α-synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder. *Lancet Neurol*. 2021;**20**(8):671–684. doi: 10.1016/S1474-4422(21)00176-9
6. Wang C, Chen F, Li Y, Liu J. Possible predictors of phenoconversion in isolated REM sleep behaviour disorder: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry*. 2022;**93**(4):395–403. doi: 10.1136/jnnp-2021-328062
7. Arnaldi D, Chincarini A, Hu MT, *et al*. Dopaminergic imaging and clinical predictors for phenoconversion of REM sleep behaviour disorder. *Brain*. 2021;**144**(1):278–287. doi: 10.1093/brain/awaa365
8. Sunwoo JS, Lee S, Kim JH, *et al*. Altered functional connectivity in idiopathic rapid eye movement sleep behavior disorder: a resting-state EEG study. *Sleep*. 2017;**40**(zsx058). doi: 10.1093/sleep/zsx058
9. Byun JI, Cha KS, Kim M, *et al*. Altered insular functional connectivity in isolated REM sleep behavior disorder: a data-driven functional MRI study. *Sleep Med*. 2021;**79**:88–93. doi: 10.1016/j.sleep.2020.12.038
10. Byun JI, Kim HW, Kang H, *et al*. Altered resting-state thalamo-occipital functional connectivity is associated with cognition in isolated rapid eye movement sleep behavior disorder. *Sleep Med*. 2020;**69**:198–203. doi: 10.1016/j.sleep.2020.01.010
11. Roascio M, Canessa A, Trò R, *et al*. Phase and amplitude electroencephalography correlations change with disease progression in people with idiopathic rapid eye-movement sleep behavior disorder. *Sleep*. 2021;**45**. doi: 10.1093/sleep/zsab232
12. Gong SY, Shen Y, Gu HY, *et al*. Generalized EEG slowing across phasic REM sleep, not subjective RBD severity, predicts neurodegeneration in idiopathic RBD. *Nat Sci Sleep*. 2022;**14**:407–418. doi: 10.2147/NSS.S354063
13. Rodrigues Brazète J, Gagnon JF, Postuma RB, Bertrand JA, Petit D, Montplaisir J. Electroencephalogram slowing predicts neurodegeneration in rapid eye movement sleep behavior disorder. *Neurobiol Aging*. 2016;**37**:74–81. doi: 10.1016/j.neurobiolaging.2015.10.007
14. Ben-Israel D, Jacobs WB, Casha S, *et al*. The impact of machine learning on patient care: a systematic review. *Artif Intell Med*. 2020;**103**:101785. doi: 10.1016/j.artmed.2019.101785
15. American Academy of Sleep Medicine. *The International Classification of Sleep Disorders*. Darien, IL.2014.
16. Lee J, Lee DW, Cho S, *et al*. Brief screening for mild cognitive impairment in elderly outpatient clinic: validation of the Korean version of the Montreal cognitive assessment. *J Geriatr Psychiatry Neurol*. 2008;104–110.

17. Kang YW, Na DL, Hahn SH. A validity study on the korean mini-mental state examination (K-MMSE) in dementia patients. *J Korean Neurol Assoc*. 1997;**15**(2):300–308.
18. You S, Moon HJ, Do SY, *et al*. The REM sleep behavior disorder screening questionnaire: validation study of the korean version (RBDQ-KR). *J Clin Sleep Med*. 2017;**13**(12):1429–1433. doi: 10.5664/jcsm.6840
19. Cho JH, Jeong YS, Lee YJ, Hong SC, Yoon JH, Kim JK. The Korean version of the Sniffin' stick (KVSS) test and its validity in comparison with the cross-cultural smell identification test (CC-SIT). *Auris Nasus Larynx*. 2009;**36**(3):280–286. doi: 10.1016/j.anl.2008.07.005
20. Visser M, Marinus J, Stiggelbout AM, Hilten JJV. Assessment of autonomic dysfunction in Parkinson's disease: The SCOPA-AUT. *Mov Disord*. 2004;**19**(11):1306–1312. doi: 10.1002/mds.20153
21. Ramaker C, Marinus J, Stiggelbout AM, van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. *Mov Disord*. 2002;**17**(5):867–876. doi: 10.1002/mds.10248
22. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;**14**(6):540–545. doi: 10.1093/sleep/14.6.540
23. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The pittsburgh sleep quality index: a new instrument psychiatric practice and research. *Psychiatry Res*. 1989;**28**(2):193–213. doi: 10.1016/0165-1781(89)90047-4
24. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry*. 1992;**55**(3):181–184. doi: 10.1136/jnnp.55.3.181
25. McKeith IG, Dickson DW, Lowe J, *et al*.; Consortium on DLB. Diagnosis and management of dementia with Lewy bodies: third report of the DLB consortium. *Neurology*. 2005;**65**(12):1863–1872. doi: 10.1212/01.wnl.0000187889.17253.b1
26. Gilman S, Wenning GK, Low PA, *et al*. Second consensus statement on the diagnosis of multiple system atrophy. *Neurology*. 2008;**71**(9):670–676. doi: 10.1212/01.wnl.0000324625.00404.15
27. Jung TP, Makeig S, Humphries C, *et al*. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*. 2000;**37**(2):163–178.
28. Pion-Tonachini L, Kreutz-Delgado K, Makeig S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage*. 2019;**198**:181–197. doi: 10.1016/j.neuroimage.2019.05.026
29. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004;**134**:9–21.
30. Roascio M, Turrisi R, Arnaldi D, *et al*. Large-scale network metrics improve the classification performance of rapid-eye-movement sleep behavior disorder patients. *bioRxiv* 2022:2022.08.16.504129. doi: 10.1101/2022.08.16.504129
31. Vinck M, Oostenveld R, van Wingerden M, Battaglia F, Pennartz CMA. An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage*. 2011;**55**(4):1548–1565. doi: 10.1016/j.neuroimage.2011.01.055
32. Shannon CE. "A mathematical theory of communication. *The Bell System Technical J*. 1948;**27**(3):623–656. doi: 10.1002/j.1538-7305.1948.tb00917.x
33. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2010;**2011**:e156869. doi: 10.1155/2011/156869
34. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;**16**(1):321–357. doi: 10.1613/jair.953
35. Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med*. 1999 Sep 15-30;**18**(17-18):2529–2545. doi: 10.1002/(sici)1097-0258(19990915/30)18:17/18<2529::aid-sim274>3.0.co;2-5. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0258(19990915/30)18:17/18%3C2529::AID-SIM274%3E3.0.CO;2-5
36. Harrell FE, Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;**247**(18):2543–2546. doi: 10.1001/jama.1982.03320430047030
37. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;**13**(1):152. doi: 10.1186/1471-2288-13-152
38. Ruffini G, Ibañez D, Castellano M, *et al*. Deep learning with EEG spectrograms in rapid eye movement behavior disorder. *Front Neurol*. 2019;**10**:806. doi: 10.3389/fneur.2019.00806.
39. Livia Fantini M, Gagnon JF, Petit D, *et al*. Slowing of electroencephalogram in rapid eye movement sleep behavior disorder. *Ann Neurol*. 2003;**53**(6):774–780. doi: 10.1002/ana.10547
40. Soikkeli R, Partanen J, Soininen H, Pääkkönen A, Riekkinen P. Slowing of EEG in Parkinson's disease. *Electroencephalogr Clin Neurophysiol*. 1991;**79**(3):159–165. doi: 10.1016/0013-4694(91)90134-p
41. Klassen BT, Hentz JG, Shill HA, *et al*. Quantitative EEG as a predictive biomarker for Parkinson disease dementia. *Neurology*. 2011;**77**(2):118–124. doi: 10.1212/WNL.0b013e318224af8d
42. Sasai T, Matsuura M, Inoue Y. Electroencephalographic findings related with mild cognitive impairment in idiopathic rapid eye movement sleep behavior disorder. *Sleep*. 2013;**36**(12):1893–1899. doi: 10.5665/sleep.3224
43. Rodrigues Brazète J, Montplaisir J, Petit D, *et al*. Electroencephalogram slowing in rapid eye movement sleep behavior disorder is associated with mild cognitive impairment. *Sleep Med*. 2013;**14**(11):1059–1063. doi: 10.1016/j.sleep.2013.06.013
44. Iranzo A, Isetta V, Molinuevo JL, *et al*. Electroencephalographic slowing heralds mild cognitive impairment in idiopathic REM sleep behavior disorder. *Sleep Med*. 2010;**11**(6):534–539. doi: 10.1016/j.sleep.2010.03.006
45. Bonanni L, Thomas A, Tiraboschi P, Perfetti B, Varanese S, Onofrj M. EEG comparisons in early Alzheimer's disease,. dementia with Lewy bodies and Parkinson's disease with dementia patients with a 2-year follow-up. *Brain*. 2008;**131**(3):690–705. doi: 10.1093/brain/awm322
46. van der Zande JJ, Gouw AA, van Steenoven I, *et al*. Diagnostic and prognostic value of EEG in prodromal dementia with Lewy bodies. *Neurology*. 2020;**95**(6):e662–e670. doi: 10.1212/WNL.0000000000009977
47. Geraedts VJ, Boon LI, Marinus J, *et al*. Clinical correlates of quantitative EEG in Parkinson disease: a systematic review. *Neurology*. 2018;**91**(19):871–883. doi: 10.1212/WNL.0000000000006473
48. Terzaghi M, Toscano G, Casoni F, *et al*. Assessment of cognitive profile as a prodromal marker of the evolution of rapid eye movement sleep behavior disorder. *Sleep*. 2019;**42**(8). doi: 10.1093/sleep/zsz103
49. Génier Marchand D, Montplaisir J, Postuma RB, Rahayel S, Gagnon JF. Detecting the cognitive prodrome of dementia with

lewy bodies: a prospective study of REM sleep behavior disorder. *Sleep*. 2017;**40**(1). doi: 10.1093/sleep/zsw014

50. Youn S, Kim T, Yoon IY, *et al*. Progression of cognitive impairments in idiopathic REM sleep behaviour disorder. *J Neurol Neurosurg Psychiatry*. 2015;**87**:jnnp-2015-311437. doi: 10.1136/jnnp-2015-311437
51. Wing YK, Li SX, Mok V, *et al*. Prospective outcome of rapid eye movement sleep behaviour disorder: psychiatric disorders as a potential early marker of Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2012;**83**(4):470–472. doi: 10.1136/jnnp-2011-301232
52. Scally B, Burke MR, Bunce D, Delvenne JF. Resting-state EEG power and connectivity are associated with alpha peak frequency slowing in healthy aging. *Neurobiol Aging*. 2018;**71**:149–155. doi: 10.1016/j.neurobiolaging.2018.07.004
53. van der Hiele K, Vein AA, Reijntjes RHAM, *et al*. EEG correlates in the spectrum of cognitive decline. *Clin Neurophysiol*. 2007;**118**(9):1931–1939. doi: 10.1016/j.clinph.2007.05.070
54. Lashgari E, Liang D, Maoz U. Data augmentation for deep-learning-based electroencephalography. *J Neurosci Methods*. 2020;**346**:108885. doi: 10.1016/j.jneumeth.2020.108885
55. Kwak NS, Müller KR, Lee SW. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PLoS One*. 2017;**12**(2):e0172578. doi: 10.1371/journal.pone.0172578
56. Schirrmeister RT, Springenberg JT, Fiederer LDJ, *et al*. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum Brain Mapp*. 2017;**38**(11):5391–5420. doi: 10.1002/hbm.23730