Identification of REM Sleep Behavior Disorder by Structural Magnetic Resonance Imaging and Machine Learning

Jie Mei[1*†], Shady Rahayel[2,3*], Christian Desrosiers[4], Ronald B. Postuma[2,5], Jacques Montplaisir[2,6], Julie Carrier[2,7,8], Oury Monchi[8,9], Johannes Frasnelli[1,2#], Jean-François Gagnon[2,10#]

[1]Department of Anatomy, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada
[2]Centre for Advanced Research in Sleep Medicine, Hôpital du Sacré-Coeur de Montréal, Montreal, QC, Canada
[3]Montreal Neurological Institute and Hospital, McGill University, Montreal, QC, Canada
[4]École de technologie supérieure Montréal, Montreal, QC, Canada
[5]Department of Neurology, Montreal General Hospital, Montreal, QC, Canada
[6]Department of Psychiatry, Université de Montréal, Montreal, QC, Canada
[7]Department of Psychology, Université de Montréal, Montreal, QC, Canada
[8]Research Centre, Institut universitaire de gériatrie de Montréal, Montreal, QC, Canada
[9]Departments of Clinical Neurosciences, Radiology, and Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada
[10]Department of Psychology, Université du Québec à Montréal, Montreal, QC, Canada

*These authors contributed equally to this work.
#These authors share senior authorship.
†Current affiliation: 1. The Brain and Mind Institute, University of Western Ontario, London, ON, Canada. 2. Department of Computer Science, University of Western Ontario, London, ON, Canada

Contact information:
Dr. Jean-François Gagnon
Département de psychologie
Université du Québec à Montréal
C.P. 8888 succ. Centre-ville,
Montréal (Québec), Canada
H3C 3P8
Email: gagnon.jean-francois.2@uqam.ca

Jie Mei:
Email: jmei47@uwo.ca

Abstract

Background

Idiopathic rapid eye movement sleep behavior disorder (iRBD) is a major risk factor for synucleinopathies, and patients often present with clinical signs and morphological brain changes. However, there is an important heterogeneity in the presentation and progression of these alterations, and the brain regions that are more vulnerable to neurodegeneration remain to be determined.

Objectives

To assess the feasibility of morphology-based machine learning approaches in the identification and subtyping of iRBD.

Methods

For the three classification tasks [iRBD (n=48) vs controls (n=41); iRBD vs Parkinson's disease (n=29); iRBD with mild cognitive impairment (n=16) vs without mild cognitive impairment (n=32)], machine learning models were trained with morphometric measurements (thickness, surface area, volume, and deformation) extracted from T1-weighted structural magnetic resonance imaging. Model performance and the most discriminative brain regions were analyzed and identified.

Results

A high accuracy was reported for iRBD vs controls (79.6%, deformation of the caudal middle frontal gyrus and putamen, thinning of the superior frontal gyrus, and reduced volume of the inferior parietal cortex and insula), iRBD vs Parkinson's disease (82%, larger volume and surface area of the insula, thinning of the entorhinal cortex and lingual gyrus, and reduced volume of the fusiform gyrus), and iRBD with vs without mild cognitive impairment (84.8%, thinning of the pars triangularis, superior temporal gyrus, transverse temporal cortex, larger surface area of the superior temporal gyrus, and deformation of isthmus of the cingulate gyrus).

Conclusions

Morphology-based machine learning approaches may allow for automated detection and subtyping of iRBD, potentially enabling efficient preclinical identification of synucleinopathies.

Introduction

Isolated/idiopathic rapid eye movement (REM) sleep behavior disorder (iRBD) is a parasomnia characterized by loss of muscle atonia and abnormal motor manifestations during REM sleep (1). iRBD is a prodromal alpha-synucleinopathy, with most patients developing Parkinson's disease (PD) or dementia with Lewy bodies (DLB) over time (2). iRBD patients already show clinical and brain changes that are reminiscent of overt synucleinopathies (3–5), including cognitive impairment and brain atrophy (1,4-7). However, the identification and progression of these changes vary widely in the iRBD population (4-7). In particular, mild cognitive impairment (MCI), present in about one third of iRBD patients (5,6), is associated with a higher likelihood of phenoconversion to a dementia- compared to a parkinsonism-first phenotype (8).

Studies using structural magnetic resonance imaging (MRI) in iRBD have reported extensive cortical and subcortical changes (4,5,7,9–14), which are more pronounced in patients with MCI (15). Findings have included frontal, temporal and occipital cortical thinning (9–11), subcortical shape contraction in the basal ganglia and hippocampus (10,12), and gray matter volume alterations, in the prefrontal cortex, caudate, brainstem, cerebellum, and parahippocampal gyrus (9,13,14). Recently, a brain volume deformation signature that could predict development of DLB at the individual level has been identified in iRBD (16). However, most of these studies have only looked at morphological changes one metric at a time, thereby preventing identifying sets of regions that may discriminate between iRBD patients and healthy individuals (controls) or between iRBD subtypes (presence of MCI), while taking into consideration several different morphological features that can be derived from every brain region.

Machine learning and deep learning have been increasingly applied to the identification and prognosis of neurodegenerative diseases including PD and DLB (17–20). Studies have validated machine learning in the identification of iRBD using polysomnogram (PSG), electroencephalogram (EEG), motor, and olfaction measures (21–25). To our knowledge, only one study has applied machine learning to structural MRI (diffusion tensor imaging) in iRBD and achieved a high accuracy (87.5%) (26). However, this study has a small sample size and no distinction of the cognitive profile. Given the ability of brain changes, mainly those associated with gray matter alterations, to identify MCI and predict phenoconversion in iRBD (15,16), an assessment of morphometric changes, along with a thorough evaluation of the discriminative power of single versus multiple brain region(s) in the identification and subtyping of iRBD, is of relevance. Moreover, machine learning has been applied to the extraction of relevant pre-clinical and clinical features, and can be used in inter-modality registration of different types of data (17,27), potentially enabling further improvements on the identification of iRBD.

In this study, we derived regional brain morphological measurements including cortical thickness, volume, surface area and tissue deformation to assess whether machine learning models differentiate between **(a)** iRBD patients and controls, **(b)** iRBD and PD patients, and **(c)** iRBD patients with and without MCI. We hypothesized that cortical and subcortical morphology can significantly differentiate between controls, PD patients and iRBD subtypes with high accuracy (≥80%)

Methods

Participants

iRBD patients were enrolled at the Centre for Advanced Research in Sleep Medicine of the *Centre Intégré universitaire de santé et de services sociaux du Nord-de-l'Île-de-Montréal – Hôpital du Sacré-Cœur de Montréal (CIUSS-NÎM-HSCM)*. All patients met the diagnostic criteria for iRBD based on the International Classification of Sleep Disorders, Third Edition and PSG (28,29). iRBD patients were excluded if they presented with parkinsonism or dementia (30,31). A history of brain injury, head trauma, stroke, claustrophobia, EEG abnormalities suggesting epilepsy, encephalitis, or other neurological disorders also led to exclusion. Controls without PD, iRBD, and MCI were recruited through newspaper advertisements or by word of mouth. They were also subjected to the same exclusion criteria as iRBD patients. All these participants were part of previous studies on neuroimaging in iRBD (9,12,15,16,32).

PD patients were recruited from the Department of Neurology of the Montreal General Hospital and the *Unité des troubles du mouvement André-Barbeau* of the *Centre Hospitalier de l'Université de Montréal*. Inclusion criteria were: (1) a diagnosis of idiopathic PD as the likeliest cause (30), (2) being 45-85 years old, (3) disease duration ≤10 years, (4) Hoehn & Yahr stage ≤3, (5) absence of dementia or any major psychiatric disorder, (6) respiratory event index ≤20, and (7) absence of a history of head injury, stroke, brain tumor, cerebrovascular disease, and chronic obstructive pulmonary disease or abnormal EEG features suggesting epilepsy. These patients were part of previous studies on neuroimaging in PD (16,33).

All participants were part of research protocols approved by local ethics committees (CIUSSS-NÎM-HSCM) and CIUSSS *du Centre-Sud-de-l'Île-de-Montréal – Comité d'éthique de la recherche vieillissement-neuroimagerie,* Montreal, Canada) and provided written informed consent.

Neuropsychological assessment

Participants underwent neurological and neuropsychological assessments including the Unified Parkinson's Disease Rating Scale part III (34), and, in iRBD patients only, the Montreal Cognitive assessment (MoCA) for screening (35). MCI was diagnosed with the neuropsychological assessment and a consensus between the neurologist and neuropsychologist based on the following criteria: (1) subjective cognitive complaints by the patient, the spouse or the caregiver, as measured using the Cognitive Failures Questionnaire (36) or as assessed during the semi-structured interview, (2) the presence of objective cognitive impairment, as defined by a performance score at least 1.5 SD

below the standardized mean on at least 2 tasks within a single cognitive domain, namely attention/executive functions, learning and verbal memory, or visuospatial abilities, (3) preserved daily life functioning, (4) no dementia, and (5) cognitive deficits not being explained solely by a medication or another medical condition (8,15).

MRI

Acquisition

T1-weighted scans were acquired in all participants using a 3T Siemens TrioTIM scanner with a 12-channel head matrix coil and an MP-RAGE sequence with the following parameters: repetition time: 2.3s, echo time: 2.91ms, inversion time: 900ms, flip angle: 9°, field of view: 256×240mm, matrix resolution: 256×240mm, voxel size: 1×1×1mm, and bandwidth: 240Hz/Px, 160 slices.

Morphological processing

Cortical surface processing was first conducted using FreeSurfer (version 6.0.0) (37,38). Processing steps included non-brain tissue removal, Talairach transformation, segmentation of subcortical white matter and gray matter subcortical volumetric structures, intensity normalization, and cortical reconstruction. Quality control was performed visually at a slice level and pial and white matter surface errors were corrected manually by a trained rater (S.R.). Cortical thickness, surface area, and volume maps were then parcellated into 34 cortical regions per hemisphere using the Desikan-Killiany atlas (39). Regional surface area and volume values were normalized by the estimated total intracranial volume since these measurements scale with head size.

Deformation-based morphometry (DBM) processing was applied to the T1-weighted images using CAT12 (http://www.neuro.uni-jena.de/cat/) to generate gray matter tissue deformation maps. DBM quantified volume differences by computing the deformations needed to perform nonlinear transformation of an individual brain to a template space (40). The Jacobian determinant maps were smoothed using a 12-mm FWHM kernel and used as a marker of local brain deformation. Maps were parcellated into 83 regions, namely 34 cortical and 7 subcortical regions (i.e., putamen, caudate, pallidum, thalamus, hippocampus, amygdala, accumbens area) for each hemisphere plus the brainstem, according to the Desikan-Killiany atlas (39). The deformation value of each region represented the mean deformation of all voxels assigned to the region.

Feature selection and machine learning

We first performed feature selection to identify morphometric measurements with the highest discriminative power. To ensure no data would be used both for feature selection and model validation, the dataset was randomly split into a training set and a test set in a stratified manner. To understand whether feature selection and model performance were affected by patient distribution of the train-test split, splitting was performed for 25 times using a train:test split ratio of 9:1 (for iRBD vs controls and iRBD vs PD) or 4:1 (for iRBD with MCI vs iRBD without MCI), to generate 25 unique splits for obtaining the averaged model performance (Supplementary Figure 1). Given the large number of features (i.e., 68, 68, 68, and 83 for thickness, surface area, volume and deformation respectively), we used a random forest classifier (n_estimators = 10) on the training set with 5-fold cross-validation to select features based on relative importance averaged over 1,000 iterations.

Machine learning models were trained with 1, 2, 3, or 5 features out of the 15-20 most relevant features, to investigate the effects of combining measurements across modalities (i.e., thickness, surface area, volume and deformation)/brain regions on iRBD detection and subtyping. To examine whether the use of multiple morphometric modalities is related to higher model performance, we performed the same procedures after combining measurements across modalities (n=287).

We implemented an expanded grid search with 5-fold cross-validation to search through models including (1) logistic regression, (2) decision tree, (3) random forest, (4) k-nearest neighbors (KNN), (5) support vector machine (SVM), and (6) boosting models, and their hyperparameters. For feature selection, model training and testing, all morphometric measurements were centered to the mean and scaled to unit variance. All procedures were implemented using Scikit-learn (version 0.17) (41).

Model evaluation and statistical analysis

To identify brain regions showing atrophy, we performed Mann-Whitney U tests followed by the Benjamini-Hochberg procedure for controlling false discovery for multiple comparisons (42). Given the imbalanced dataset, in addition to accuracy, metrics including recall, precision, and F1-score were used to assess the model performance. Recall, also known as sensitivity or true positive rate (TPR), is defined as the number of true positives divided by the sum of true positives and false negatives and assesses the model's ability to identify relevant cases. Precision is defined as the number of true positives divided by the sum of true positives and false positives, representing how many relevant cases identified by the model are relevant. The F1-score is the harmonic mean of precision and recall, and accuracy indicates the ratio of number of correct predictions

to number of all predictions. Data are shown as mean (SD). Statistical analyses were performed using pandas (version 0.25.3) (43) and scikit-posthocs (version 0.6.7) (44).

Results

Demographics and clinical data

Of the 59 iRBD patients initially recruited, 11 were excluded for parkinsonism or dementia at presentation or unclear clinical presentation/borderline electromyogram thresholds for REM sleep muscle atonia. All participants survived quality control except for one PD patient who had abnormal surface reconstruction in the occipital lobe, yielding 48 iRBD patients, 29 PD patients [mean PD duration since diagnosis (SD): 3.8 (2.7) years; mean levodopa equivalent dosage (SD): 536.5 (283.1) mg/daily; mean Hoehn & Yahr stage (SD): 2.2 (0.8); 41% (12/29) with MCI; 48% (14/29) with RBD], and 41 controls for MRI processing (Table 1). Sixteen (33%) iRBD patients had MCI. There were no significant differences in age and education between the groups. The proportion of men was higher in iRBD than PD patients. PD had a higher UPDRS-III total score than iRBD patients. Moreover, iRBD patients with MCI performed poorer than those without MCI on the MoCA.

Insert Table 1 here

Morphometric measurements

Of the 287 morphometric features, 37 (12.9%) significantly differed between iRBD patients and controls, 20 (7.0%) between iRBD and PD patients, and 27 (9.4%) between iRBD patients with or without MCI (for detailed morphometric modalities and brain regions, see Supplementary Tables 1-5). Compared to controls, iRBD patients had cortical thinning, reduced volume, and tissue deformation in several brain regions (Supplementary Table 1). Compared to PD, iRBD patients had mixed results for cortical thickness, but with reduced volume in PD (bilateral fusiform gyrus) and tissue deformation, mainly in left subcortical regions, in iRBD (Supplementary Tables 3 and 4). Compared to patients without MCI, those with MCI had cortical thinning, larger surface area, and tissue deformation, mainly in posterior regions (Supplementary Tables 3 and 5).

Feature importance and machine learning-based classification

Although 5 brain regions that contributed to the highest overall model performance were reported, it should be noted that in all classification tasks, feature importance of the most relevant brain regions (top 5%-10%) were often comparable (Supplementary Tables 2 and 3) and could lead to classification performance that was only 2%-5% lower than the reported results.

iRBD vs Controls

The highest accuracy (79.6%) was obtained when measurements of all morphometric modalities were merged to select 5 features for model training, namely deformation of the left caudal middle frontal gyrus and right putamen, thinning of the left superior frontal gyrus, and reduced volume of the right inferior parietal cortex and insula (Table 2 and Figure 1a-e). When using a single modality for model training, a classification accuracy of 79.1% was achieved with 5 features selected from DBM-based measurements, namely tissue deformation in the left caudal middle frontal gyrus, paracentral lobule, thalamus, and bilateral putamen. When training the model with 1 feature out of all features derived from DBM or across all modalities, deformation of the left caudal middle frontal gyrus led to an accuracy of 71.1%.

Among all machine learning models evaluated, depending on the morphometric modality used, linear SVM or SVM with a radial basis function (RBF) was the best performing model in most, if not all, of the 25 randomly generated train-test splits (Table 2; n=24.95 (0.22); range=24-25). As indicated by the relative importance, cortical regions, including the left caudal middle frontal cortex, paracentral lobule, and superior frontal cortex, as well as the right superior frontal cortex, inferior parietal cortex, and insula, yielded high discriminative power (Table 2; Figure 2; Supplementary Table 2). As model training with merged measurements across morphometric modalities led to the highest accuracy in the classification of iRBD and controls, for iRBD vs PD patients and iRBD patients with MCI vs patients without MCI, we merged across all morphometric modalities for model training without examining classification performance of individual modalities.

Insert Table 2 here

Insert Figure 1 here

Insert Figure 2 here

iRBD vs PD

With morphometric measurements of 5 brain regions, including the larger volume and surface area of the right insula, thinning of the right entorhinal cortex and lingual gyrus, and reduced volume of the right fusiform gyrus in PD, an accuracy of 82% was achieved (Table 2; Figure 1f and 3). When morphometric measurements of single brain regions were used to train the machine learning model, the reduced cortical volume of the right insula in iRBD led to an accuracy of 70% (Table 2). SVM with an RBF kernel demonstrated the highest overall performance in a majority of the 25 randomly generated train-test splits (n=24.75 (0.5); range=24-25).

Insert Figure 3 here

iRBD with MCI vs iRBD without MCI

An accuracy of 84.8% was obtained from a model trained with morphometric measurements of 5 brain regions, namely thinning of the left transverse temporal cortex, pars triangularis and superior temporal gyrus, larger surface area of the left superior temporal gyrus, and deformation of isthmus of the right cingulate gyrus (Table 2; Figures 1g and 3). When using a single brain region (thinning of the left pars triangularis in iRBD with MCI) or 2 regions (thinning of the left pars triangularis and superior temporal gyrus in iRBD with MCI), a decision tree classifier of depth 1 achieved an accuracy of 72.8% and 73.2%, respectively (Table 2; Figure 1g and 3). When more brain regions were used in model training, a linear SVM yielded the highest overall performance in most of the train-test splits (n=22.5 (0.71); range=22-23).

Discussion

General observations

In this study, we derived a comprehensive set of morphological regional brain measurements and used machine learning models to identify the sets of regions that best discriminate between iRBD patients and controls, iRBD and PD patients, and cognitive subtypes of iRBD. The main findings showed that machine learning models trained with morphometric measurements efficiently differentiate between iRBD patients and controls (accuracy=79.6%), iRBD and PD patients (accuracy=82%), and iRBD patients with or without MCI (accuracy=84.8%) and may prove useful for guiding future automated algorithms that may discriminate patients based on structural MRI.

When discriminating iRBD patients from controls with all morphometric modalities, an accuracy of 79.6% was achieved by a model trained using deformation of the left caudal middle frontal gyrus and right putamen, thinning of the left superior frontal gyrus, and reduced volume of the right inferior parietal cortex and insula. With single modalities, a highly comparable accuracy of 79.1% was obtained from models trained with subcortical and cortical DBM-derived data, i.e., tissue deformation in the left caudal middle frontal gyrus, paracentral lobule, thalamus, and bilateral putamen. For the other modalities, namely cortical surface-derived measures of thickness, surface area or volume, a lower but significantly above-chance-level accuracy was also observed. This supports the importance of tissue deformation in the structural pattern of atrophy seen in iRBD as well as the importance of taking into consideration the changes occurring in subcortical regions in iRBD and early synucleinopathy. The model trained with greater volume and surface area of the right insula, thinning of the right entorhinal cortex and lingual gyrus, and reduced volume of the right fusiform gyrus led to an accuracy of 82% in the discrimination of iRBD from PD patients. Furthermore, an accuracy of 84.8% was obtained in the identification of iRBD patients with MCI, from a model trained with thinning of the left transverse temporal cortex, pars triangularis and superior temporal gyrus, larger surface area of the left superior temporal gyrus, and deformation of isthmus of the right cingulate gyrus.

Few studies have used machine learning to identify iRBD patients. One study performed machine learning models trained with diffusion tensor imaging measures and found structural differences between 20 iRBD patients and 20 controls, enabling identification of iRBD with an accuracy of 87.5% (26). Other studies used machine learning with other clinical markers for identification of iRBD patients. In one study (23), resting EEG data were analyzed using a convolutional neural network and a recurrent neural network. The two architectures achieved comparable results, leading to an

accuracy of 80% when predicting conversion to PD in iRBD. This study also revealed relevant features including bursts in the theta band and a decrease of bursting in the alpha band in iRBD patients who have converted to PD or DLB, when compared with controls. Another study derived features from EEG signals during sleep to train a random forest model that identified iRBD vs controls with an accuracy of 96%, a sensitivity of 98%, and a specificity of 94% (24). In another study that used PSG data, electrooculogram signals were acquired for deriving features relevant to micro-sleep structure, which gave rise to identification of iRBD patients in PD patients through micro-sleep instability, with accuracy, sensitivity, and specificity over 80% (22). Using sensor data recorded during motor tasks, another study trained random forest classifiers to distinguish controls, iRBD and PD patients and obtained a sensitivity ranging from 84.6% to 91.9%, and a specificity ranging from 88.3% to 90.1% (25). One study made use of olfaction data (21) and reported an area under the curve value of ≥0.90 when classifying between iRBD and PD. While these studies demonstrated a high accuracy in the identification of iRBD or PD, many are based on motor features that may appear at a later stage of neurodegeneration as in patients at risk of DLB. Moreover, in motor tasks and EEG/PSG, data acquisition can take up to days, making quality control and follow-ups challenging. In addition, previous studies have not attempted to subtype between iRBD patients with or without MCI, for which we have achieved an accuracy of 84.8%. This is essential given the large heterogeneity of cognitive impairment reported in iRBD patients, and the higher risk of developing DLB and PD in iRBD patients with MCI. Comparing with these studies, the present study highlights morphological changes in cortical and subcortical regions that allow the identification and subtyping of iRBD patients.

The deformation of putamen and thalamus as well as the thinning of caudal middle frontal and paracentral cortices best distinguished between iRBD patients and controls. This is in line with previous studies performed on the same cohorts but with a single morphometric measure (12,15,16). Putamen dopamine depletion underlies parkinsonism (45) and the putamen is a region that is already structurally and functionally abnormal in iRBD (12,46–50), as observed by reduced dopamine reuptake imaging (5,47) and perfusion/metabolism imaging (48–50). The smaller volume of the insula in iRBD was also found to discriminate between iRBD and PD patients. In the present study, greater thinning in the left insula was observed in iRBD patients with MCI, and the presence of MCI in iRBD is a risk factor for DLB (5,6,8), which is in agreement with studies reporting the insula as highly associated with prodromal and clinical DLB (51–54). Its ability to discriminate between iRBD and PD may therefore be due to some iRBD patients being on a dementia-first trajectory. Also, in accordance with studies that used morphometric measures of the fusiform gyrus to detect PD (55,56), our study highlighted the reduced volume of the fusiform gyrus in PD patients as a discriminative feature.

Compared to surface-derived cortical metrics assessing local thickness, surface area or volume, DBM-derived tissue deformation measurements generated higher discriminative power between iRBD patients and controls. This is important, as different morphological metrics were shown to follow different developmental trajectories (57), to relate to different genetic determinants (58,59), and to be affected differentially in neurodegenerative prodromes (15,60). Several reasons can explain the fact that DBM is most discriminative: First, the DBM maps included information about subcortical structures, which are known to be affected in iRBD (9,10). Secondly, white matter changes are an increasingly recognized feature of neurodegeneration in PD and DLB (61,62); although tissue deformation measurements were extracted from the gray matter only, it is possible that gray matter deformation may have been influenced by dynamics occurring in the underlying white matter. This is even more important as some studies have reported white matter abnormalities in iRBD (14,63,64). Recently, we identified in the same cohort of iRBD patients a DBM-derived structural signature that predicts the development of DLB in iRBD; interestingly, white matter tissue contraction was an important feature of this signature (16). Future studies should investigate the association between changes in gray and white matter in iRBD.

We also found that some of the discriminative brain regions in the classification between iRBD patients and controls, and the identification of PD patients, such as the medial orbitofrontal cortex and pars orbitalis, are involved in olfactory processing. Morphometric measures of olfactory processing areas in the orbitofrontal cortex such as the olfactory sulcus, the gyrus rectus, and the medial orbitofrontal cortex are associated with olfactory sensitivity (65,66). Further, olfactory function is impaired in both PD and iRBD patients (3,67). We have shown in a previous study that a convolutional neural network can distinguish PD patients from patients with non-parkinsonian olfactory dysfunction with high sensitivity and specificity (68). In the future, olfaction and other early markers of motor or cognitive performance could be therefore combined with morphometric measures to increase the specificity and sensitivity of automated diagnostic approaches (5).

Limitations

From a clinical perspective, brain imaging data used in this study were collected from a considerable number of participants. However, these data represent a relatively small sample size when compared with deep learning studies using multicentric data, or with studies applying machine learning and deep learning approaches to iRBD detection using PSG or EEG features. To enable computer-aided clinical decision-making in the identification and subtyping of iRBD, validation studies using a larger prospective-longitudinal cohort are required. In addition, it has been shown that the version of the preprocessing pipeline may become a confounding factor in the assessment of

neurological disorders (69), therefore, it would be crucial to understand to what extent the detection of iRBD is affected by the preprocessing pipeline, and/or the software version used.

Author roles

J. Mei: Conception and execution of research project, design and execution of statistical analysis, writing the first draft of manuscript.
SR: Conception of research project, design, review and critique of statistical analysis, writing the first draft of manuscript.
CD: Review and critique of statistical analysis, review of manuscript.
RBP: Funding, organization of research project, data collection, review of manuscript.
J. Montplaisir: Funding, organization of research project, data collection, review of manuscript.
JC: Organization of research project, data collection, review of manuscript.
OM: Organization of research project, data collection, review of manuscript.
JF: Funding, organization and supervision of research project, review of manuscript.
JFG: Funding, organization and supervision of research project, data collection, writing and review of manuscript.

Acknowledgements

We would like to thank Amélie Pelletier and Jessie De Roy for their assistance in organizing data, and Sylvain Chouinard and Michel Panisset for clinical evaluations of PD patients. Shady Rahayel receives a scholarship from the Fonds de recherche du Québec – Santé. JFG holds a Canada Research Chair in Cognitive Decline in Pathological Aging.

Financial Disclosures

This work was supported by the Canadian Institutes of Health Research (CIHR), the *Fonds de recherche du Québec – Santé* (FRQ-S), the *Réseau intersectoriel de recherche en santé de l'Université du Québec* (RISUQ), the Centre de recherche en neurosciences cognitives de l'Université du Québec à Montréal (NeuroQAM), and the W. Garfield Weston Foundation. JFG received research grant outside this work from the National Institutes of Health/National Institute on Aging.

J. Mei:
Stock Ownership in medically-related fields: Oyama Health AB
Consultancies: deepkapha.ai, Oyama Health AB
Employment: Université du Québec à Trois-Rivières, University of Western Ontario
Grants: BrainsCAN Postdoctoral Fellowship Award through Canada First Research Excellence Fund (CFREF)
Partnerships: Charité Global Health

JF:
Employment: Université du Québec à Trois-Rivières
Royalties: Styriabooks
Grants: FRQS, NSERC, CIHR

References

1. Högl B, Stefani A, Videnovic A. Idiopathic REM sleep behaviour disorder and neurodegeneration - an update. Nat Rev Neurol. 2018 Jan;14(1):40–55.

2. Postuma RB, Iranzo A, Hu M, Högl B, Boeve BF, Manni R, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain. 2019 Mar 1;142(3):744–59.

3. Fereshtehnejad S-M, Yao C, Pelletier A, Montplaisir JY, Gagnon J-F, Postuma RB. Evolution of prodromal Parkinson's disease and dementia with Lewy bodies: a prospective study. Brain. 2019 Jul 1;142(7):2051–67.

4. Bourgouin P-A, Rahayel S, Gaubert M, Arnaldi D, Hu M, Heidbreder A, et al. Neuroimaging of Rapid Eye Movement Sleep Behavior Disorder. Int Rev Neurobiol. 2019;144:185–210.

5. Miglis MG, Adler CH, Antelmi E, Arnaldi D, Baldelli L, Boeve BF, et al. Biomarkers of conversion to α-synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder. The Lancet Neurology. 2021 Aug 1;20(8):671–84.

6. Gagnon J, Bourgouin P, Roy J, Marchand DG. Neuropsychological Aspects: Cognition in RBD. In 2019.

7. Campabadal A, Segura B, Junque C, Iranzo A. Structural and functional magnetic resonance imaging in isolated REM sleep behavior disorder: A systematic review of studies using neuroimaging software. Sleep Med Rev. 2021 Apr 22;59:101495.

8. Génier Marchand D, Montplaisir J, Postuma RB, Rahayel S, Gagnon J-F. Detecting the Cognitive Prodrome of Dementia with Lewy Bodies: A Prospective Study of REM Sleep Behavior Disorder. Sleep. 2017 Jan 1;40(1).

9. Rahayel S, Montplaisir J, Monchi O, Bedetti C, Postuma RB, Brambati S, et al. Patterns of cortical thinning in idiopathic rapid eye movement sleep behavior disorder. Mov Disord. 2015 Apr 15;30(5):680–7.

10. Campabadal A, Segura B, Junque C, Serradell M, Abos A, Uribe C, et al. Cortical Gray Matter and Hippocampal Atrophy in Idiopathic Rapid Eye Movement Sleep Behavior Disorder. Front Neurol. 2019 Apr 5;10:312.

11. Pereira JB, Weintraub D, Chahine L, Aarsland D, Hansson O, Westman E. Cortical thinning in patients with REM sleep behavior disorder is associated with clinical progression. NPJ Parkinsons Dis. 2019 May 3;5:7.

12. Rahayel S, Postuma RB, Montplaisir J, Bedetti C, Brambati S, Carrier J, et al. Abnormal Gray Matter Shape, Thickness, and Volume in the Motor Cortico-Subcortical Loop in Idiopathic Rapid Eye Movement Sleep Behavior Disorder: Association with Clinical and Motor Features. Cereb Cortex. 2018 Feb 1;28(2):658–71.

13. Hanyu H, Inoue Y, Sakurai H, Kanetaka H, Nakamura M, Miyamoto T, et al. Voxel-based magnetic resonance imaging study of structural brain changes in patients with idiopathic REM sleep behavior disorder. Parkinsonism Relat Disord. 2012 Feb;18(2):136–9.

14. Holtbernd F, Romanzetti S, Oertel WH, Knake S, Sittig E, Heidbreder A, et al. Convergent patterns of structural brain changes in rapid eye movement sleep behavior disorder and Parkinson's disease on behalf of the German rapid eye movement sleep behavior disorder study group. Sleep. 2021 Mar 12;44(3):zsaa199.

15. Rahayel S, Postuma RB, Montplaisir J, Génier Marchand D, Escudier F, Gaubert M, et al. Cortical and subcortical gray matter bases of cognitive deficits in REM sleep behavior disorder. Neurology. 2018 May 15;90(20):e1759–70.

16. Rahayel S, Postuma RB, Montplaisir J, Mišić B, Tremblay C, Vo A, et al. A Prodromal Brain-Clinical Pattern of Cognition in Synucleinopathies. Ann Neurol. 2021 Feb;89(2):341–57.

17. Mei J, Desrosiers C, Frasnelli J. Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature. Front Aging Neurosci [Internet]. 2021 [cited 2021 Jul 19];0. Available from: https://www.frontiersin.org/articles/10.3389/fnagi.2021.633752/full

18. Iizuka T, Fukasawa M, Kameyama M. Deep-learning-based imaging-classification identified cingulate island sign in dementia with Lewy bodies. Sci Rep. 2019 Jun 20;9(1):8944.

19. Shigemizu D, Akiyama S, Asanomi Y, Boroevich KA, Sharma A, Tsunoda T, et al. A comparison of machine learning classifiers for dementia with Lewy bodies using miRNA expression data. BMC Medical Genomics. 2019 Oct 30;12(1):150.

20. Nemoto K, Sakaguchi H, Kasai W, Hotta M, Kamei R, Noguchi T, et al. Differentiating Dementia with Lewy Bodies and Alzheimer's Disease by Deep Learning to Structural MRI. Journal of Neuroimaging. 2021;31(3):579–87.

21. Lo C, Arora S, Ben-Shlomo Y, Barber TR, Lawton M, Klein JC, et al. Olfactory Testing in Parkinson Disease and REM Behavior Disorder: A Machine Learning Approach. Neurology. 2021 Apr 13;96(15):e2016–27.

22. Cesari M, Christensen JAE, Muntean M-L, Mollenhauer B, Sixel-Döring F, Sorensen HBD, et al. A data-driven system to identify REM sleep behavior disorder and to predict its progression from the prodromal stage in Parkinson's disease. Sleep Med. 2021 Jan;77:238–48.

23. Ruffini G, Ibañez D, Castellano M, Dubreuil-Vall L, Soria-Frisch A, Postuma R, et al. Deep Learning With EEG Spectrograms in Rapid Eye Movement Behavior Disorder. Front Neurol [Internet]. 2019 [cited 2021 Jul 19];0. Available from: https://www.frontiersin.org/articles/10.3389/fneur.2019.00806/full

24. Cooray N, Andreotti F, Lo C, Symmonds M, Hu MTM, De Vos M. Detection of REM sleep behaviour disorder by automated polysomnography analysis. Clin Neurophysiol. 2019 Apr;130(4):505–14.

25. Arora S, Baig F, Lo C, Barber TR, Lawton MA, Zhan A, et al. Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. Neurology. 2018 Oct 16;91(16):e1528–38.

26. Lee DA, Lee H-J, Kim HC, Park KM. Application of machine learning analysis based on diffusion tensor imaging to identify REM sleep behavior disorder. Sleep Breath. 2021 Jul 8;

27. Huang S-C, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. npj Digit Med. 2020 Oct 16;3(1):1–9.

28. Montplaisir J, Gagnon J-F, Fantini ML, Postuma RB, Dauvilliers Y, Desautels A, et al. Polysomnographic diagnosis of idiopathic REM sleep behavior disorder. Mov Disord. 2010 Oct 15;25(13):2044–51.

29. American Academy of Sleep Medicine. International Classification of Sleep Disorders. Diagnostic and Coding Manual [Internet]. 2014 [cited 2021 Sep 17]; Available from: https://ci.nii.ac.jp/naid/20001061569/

30. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2015 Oct;30(12):1591–601.

31. McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor J-P, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. Neurology. 2017 Jul 4;89(1):88–100.

32. Bourgouin P-A, Rahayel S, Gaubert M, Postuma RB, Montplaisir J, Carrier J, et al. Gray matter substrates of depressive and anxiety symptoms in idiopathic REM sleep behavior disorder. Parkinsonism Relat Disord. 2019 May;62:163–70.

33. Rahayel S, Gaubert M, Postuma RB, Montplaisir J, Carrier J, Monchi O, et al. Brain atrophy in Parkinson's disease with polysomnography-confirmed REM sleep behavior disorder. Sleep [Internet]. 2019 Jun 11 [cited 2021 Jul 19];42(6). Available from: https://doi.org/10.1093/sleep/zsz062

34. Fahn S, Elton R, Members of the UPDRS Development Committee. Recent developments in Parkinson's disease. Macmillan health care information. 1987;2, 153–163:293–304.

35. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005 Apr;53(4):695–9.

36. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. Br J Clin Psychol. 1982 Feb;21(1):1–16.

37. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage. 2012 Jul 16;61(4):1402–18.

38. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage. 1999 Feb;9(2):179–94.

39. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006 Jul 1;31(3):968–80.

40. Ashburner J, Hutton C, Frackowiak R, Johnsrude I, Price C, Friston K. Identifying global anatomical differences: deformation-based morphometry. Hum Brain Mapp. 1998;6(5–6):348–57.

41. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(Oct):2825–30.

42. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289–300.

43. McKinney W. Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference. 2010;56–61.

44. Terpilowski MA. scikit-posthocs: Pairwise multiple comparison tests in Python. Journal of Open Source Software. 2019 Apr 14;4(36):1169.

45. Politis M. Neuroimaging in Parkinson disease: from research setting to clinical practice. Nat Rev Neurol. 2014 Dec;10(12):708–22.

46. Rolinski M, Griffanti L, Piccini P, Roussakis AA, Szewczyk-Krolikowski K, Menke RA, et al. Basal ganglia dysfunction in idiopathic REM sleep behaviour disorder parallels that in early Parkinson's disease. Brain. 2016 Aug;139(8):2224–34.

47. Bauckneht M, Chincarini A, De Carli F, Terzaghi M, Morbelli S, Nobili F, et al. Presynaptic dopaminergic neuroimaging in REM sleep behavior disorder: A systematic review and meta-analysis. Sleep Med Rev. 2018 Oct;41:266–74.

48. Vendette M, Gagnon J-F, Soucy J-P, Gosselin N, Postuma RB, Tuineag M, et al. Brain perfusion and markers of neurodegeneration in rapid eye movement sleep behavior disorder. Mov Disord. 2011 Aug 1;26(9):1717–24.

49. Holtbernd F, Gagnon J-F, Postuma RB, Ma Y, Tang CC, Feigin A, et al. Abnormal metabolic network activity in REM sleep behavior disorder. Neurology. 2014 Feb 18;82(7):620–7.

50. Kim R, Lee J-Y, Kim YK, Kim H, Yoon EJ, Shin JH, et al. Longitudinal Changes in Isolated Rapid Eye Movement Sleep Behavior Disorder-Related Metabolic Pattern Expression. Mov Disord. 2021 Mar 31;

51. Zhong J, Pan P, Dai Z, Shi H. Voxelwise meta-analysis of gray matter abnormalities in dementia with Lewy bodies. Eur J Radiol. 2014 Oct;83(10):1870–4.

52. Blanc F, Colloby SJ, Philippi N, Pétigny X de, Jung B, Demuynck C, et al. Cortical Thickness in Dementia with Lewy Bodies and Alzheimer's Disease: A Comparison of Prodromal and Dementia Stages. PLOS ONE. 2015 Jun 10;10(6):e0127396.

53. Roquet D, Noblet V, Anthony P, Philippi N, Demuynck C, Cretin B, et al. Insular atrophy at the prodromal stage of dementia with Lewy bodies: a VBM DARTEL study. Sci Rep. 2017 Aug 25;7(1):9437.

54. Firbank MJ, Durcan R, O'Brien JT, Allan LM, Barker S, Ciafone J, et al. Hippocampal and insula volume in mild cognitive impairment with Lewy bodies. Parkinsonism & Related Disorders. 2021 May 1;86:27–33.

55. Peng B, Wang S, Zhou Z, Liu Y, Tong B, Zhang T, et al. A multilevel-ROI-features-based machine learning method for detection of morphometric biomarkers in Parkinson's disease. Neurosci Lett. 2017 Jun 9;651:88–94.

56. Mei J, Tremblay C, Stikov N, Desrosiers C, Frasnelli J. Differentiation of Parkinson's disease and non-Parkinsonian olfactory dysfunction with structural MRI data. In: Medical Imaging 2021: Computer-Aided Diagnosis [Internet]. International Society for Optics and Photonics; 2021 [cited 2021 Jul 19]. p. 115971E. Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11597/115971E/Differentiation-of-Parkinsons-disease-and-non-Parkinsonian-olfactory-dysfunction-with/10.1117/12.2581233.short

57. Wierenga LM, Langen M, Oranje B, Durston S. Unique developmental trajectories of cortical thickness and surface area. Neuroimage. 2014 Feb 15;87:120–6.

58. Winkler AM, Kochunov P, Blangero J, Almasy L, Zilles K, Fox PT, et al. Cortical Thickness or Grey Matter Volume? The Importance of Selecting the Phenotype for Imaging Genetics Studies. Neuroimage. 2010 Nov 15;53(3):1135–46.

59. Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, et al. Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex. 2009 Nov;19(11):2728–35.

60. Rahayel S, Bocti C, Sévigny Dupont P, Joannette M, Lavallée MM, Nikelski J, et al. Subcortical amyloid relates to cortical morphology in cognitively normal individuals. Eur J Nucl Med Mol Imaging. 2019 Oct;46(11):2358–69.

61. Sarasso E, Agosta F, Piramide N, Filippi M. Progression of grey and white matter brain damage in Parkinson's disease: a critical review of structural MRI literature. J Neurol [Internet]. 2020 May 6 [cited 2021 Jul 27]; Available from: https://doi.org/10.1007/s00415-020-09863-8

62. Donaghy PC, Firbank M, Petrides G, Lloyd J, Barnett N, Olsen K, et al. Diffusion imaging in dementia with Lewy bodies: Associations with amyloid burden, atrophy, vascular factors and clinical features. Parkinsonism & Related Disorders. 2020 Sep 1;78:109–15.

63. Unger MM, Belke M, Menzler K, Heverhagen JT, Keil B, Stiasny-Kolster K, et al. Diffusion tensor imaging in idiopathic REM sleep behavior disorder reveals microstructural changes in the brainstem, substantia nigra, olfactory region, and other brain regions. Sleep. 2010 Jun;33(6):767–73.

64. Scherfler C, Frauscher B, Schocke M, Iranzo A, Gschliesser V, Seppi K, et al. White and gray matter abnormalities in idiopathic rapid eye movement sleep behavior disorder: a diffusion-tensor imaging and voxel-based morphometry study. Ann Neurol. 2011 Feb;69(2):400–7.

65. Seubert J, Freiherr J, Frasnelli J, Hummel T, Lundström JN. Orbitofrontal Cortex and Olfactory Bulb Volume Predict Distinct Aspects of Olfactory Performance in Healthy Subjects. Cerebral Cortex. 2013 Oct 1;23(10):2448–56.

66. Frasnelli J, Lundström JN, Boyle JA, Djordjevic J, Zatorre RJ, Jones-Gotman M. Neuroanatomical correlates of olfactory performance. Exp Brain Res. 2010 Feb;201(1):1–11.

67. Rahayel S, Frasnelli J, Joubert S. The effect of Alzheimer's disease and Parkinson's disease on olfaction: a meta-analysis. Behav Brain Res. 2012 May 16;231(1):60–74.

68. Tremblay C, Mei J, Frasnelli J. Olfactory bulb surroundings can help to distinguish Parkinson's disease from non-parkinsonian olfactory dysfunction. NeuroImage: Clinical. 2020 Jan 1;28:102457.

69. Chepkoech J-L, Walhovd KB, Grydeland H, Fjell AM, Alzheimer's Disease Neuroimaging Initiative. Effects of change in FreeSurfer version on classification accuracy of patients with Alzheimer's disease and mild cognitive impairment. Hum Brain Mapp. 2016 May;37(5):1831–41.

Tables

Table 1. Demographic and clinical in participants.

Variable	iRBD (A)	iRBD with MCI (B)	iRBD without MCI (C)	PD (D)	Controls (E)	p value, A vs. E	p value, A vs. D	p value, B vs. C
Age at MRI, years	65.8 (6.4)	67.9 (4.6)	64.8 (7.0)	64.8 (8.3)	63.2 (8.2)	0.09[b]	0.56[b]	0.080[b]
Men, n (%)	37 (77)	10 (63)	27 (84)	15 (50)	25 (61)	1.00[c]	0.014[c]	0.089[c]
Education, years	13.3 (3.7)	11.9 (3.8)	14.0 (3.4)	14.9 (3.8)	14.6 (4.1)	0.18[b]	0.061[b]	0.058[b]
RBD duration since symptom onset, years	12.1 (12.2)	13.3 (13.4)	11.5 (11.7)	-	-	-	-	0.47[d]
UPDRS-III, total[a]	4.3 (3.6)	5.6 (5.0)	3.7 (2.6)	20.6 (9.2)	-	-	<0.001[d]	0.24[d]
MoCA score	25.7 (2.8)	24.1 (3.3)	26.9 (1.8)	-	-	-	-	0.005[d]

Values are presented as mean (SD).

PD: Parkinson's disease, iRBD: idiopathic rapid eye movement sleep behavior disorder, MCI: mild cognitive impairment, UPDRS-III: Unified Parkinson's Disease Rating Scale part III, MoCA: Montreal Cognitive Assessment.

[a]The UPDRS-III was measured in the 'on' state in PD patients.

[b]Student t test.

[c]Chi-squared test.

[d]Mann-Whitney U test.

Table 2. Results of the best performing models, brain regions and associated morphometric measurements used in the differentiation between iRBD patients vs controls, iRBD vs PD patients, and iRBD patients with MCI vs without MCI.

Modality	**Number of features**	**Recall**	**Precision**	**F1-score**	**Accuracy**	**Machine learning model**	**Brain region and characteristics in iRBD vs controls**
Cortical thickness	1	0.640 (0.173)	0.702 (0.182)	0.661 (0.153)	0.640 (0.155)	Linear SVM, C = 1	thinning of paracentral lobule (lh)
	5	0.712 (0.183)	0.725 (0.166)	0.710 (0.160)	0.684 (0.159)	Linear SVM, C = 1	thinning of paracentral lobule (lh) thinning of medial orbital frontal cortex (lh) thicker banks of the superior temporal sulcus (lh) thinning of temporal pole (lh) thinning of frontal pole (rh)
Surface area	1	0.920 (0.183)	0.573 (0.069)	0.703 (0.108)	0.582 (0.108)	Linear SVM, C = 1	larger surface area of isthmus of cingulate cortex (lh)
	5	0.800 (0.141)	0.721 (0.093)	0.754 (0.103)	0.716 (0.107)	Linear SVM, C = 1	larger surface area of isthmus of cingulate cortex (lh) reduced surface area of rostral anterior cingulate cortex (lh) larger surface area of precentral gyrus (lh) reduced surface area of rostral middle frontal gyrus (lh) larger surface area of temporal pole (lh)
Volume	1	0.784 (0.152)	0.680 (0.150)	0.720 (0.124)	0.658 (0.157)	SVM with RBF kernel, C = 1	reduced volume of caudal middle frontal gyrus (lh)
	5	0.832 (0.160)	0.735 (0.126)	0.775 (0.121)	0.733 (0.140)	SVM with RBF kernel, C = 1	reduced volume of caudal middle frontal gyrus (lh) reduced volume of paracentral lobule (lh) reduced volume of rostral anterior cingulate cortex (lh) larger volume of fusiform gyrus (rh) reduced volume of inferior parietal cortex (rh)
Deformation	1	0.816 (0.140)	0.714 (0.113)	0.758 (0.112)	0.711 (0.132)	Linear SVM, C = 1	deformation of caudal middle frontal gyrus (lh)
	5	0.768 (0.149)	0.852 (0.164)	0.803 (0.142)	0.791 (0.152)	Linear SVM, C = 1	deformation of caudal middle frontal gyrus (lh) deformation of paracentral lobule (lh) deformation of thalamus proper (lh) deformation of putamen (lh) deformation of putamen (rh)
All modalities	1	0.816 (0.140)	0.714 (0.113)	0.758 (0.112)	0.711 (0.132)	Linear SVM, C = 1	deformation of caudal middle frontal gyrus (lh)
	5	0.776 (0.176)	0.855 (0.170)	0.806 (0.154)	0.796 (0.163)	Linear SVM, C = 1	deformation of caudal middle frontal gyrus (lh) thinning of superior frontal gyrus (lh) reduced volume of inferior parietal cortex (rh) reduced volume of insula (rh) deformation of putamen (rh)
	Number of features	**Recall**	**Precision**	**F1-score**	**Accuracy**	**Machine learning model**	**Brain region and characteristics in iRBD vs PD**

All modalities	1	0.848 (0.145)	0.724 (0.073)	0.776 (0.086)	0.700 (0.102)	SVM with RBF kernel, C = 1	larger volume of insula (rh)
All modalities	2	0.880 (0.129)	0.786 (0.088)	0.828 (0.099)	0.775 (0.125)	SVM with RBF kernel, C = 1	larger surface area of insula (rh) thinning of entorhinal cortex (rh)
All modalities	3	0.896 (0.131)	0.801 (0.114)	0.842 (0.105)	0.790 (0.134)	SVM with RBF kernel, C = 1	larger surface area of insula (rh) thinning of entorhinal cortex (rh) thinning of lingual gyrus (rh)
All modalities	5	0.904 (0.131)	0.832 (0.120)	0.863 (0.111)	0.820 (0.140)	SVM with RBF kernel, C = 1	larger volume of insula (rh) larger surface area of insula (rh) thinning of entorhinal cortex (rh) thinning of lingual gyrus (rh) reduced volume of fusiform gyrus (rh)
	Number of features	**Recall**	**Precision**	**F1-score**	**Accuracy**	**Machine learning model**	**Brain region and characteristics in iRBD with MCI vs without MCI**
All modalities	1	0.907 (0.205)	0.525 (0.094)	0.660 (0.127)	0.728 (0.079)	Decision tree, depth = 1	thinning of pars triangularis (lh)
All modalities	2	0.907 (0.181)	0.537 (0.092)	0.667 (0.104)	0.732 (0.075)	Decision tree, depth = 1	thinning of pars triangularis (lh) thinning of superior temporal gyrus (lh)
All modalities	3	0.733 (0.236)	0.701 (0.188)	0.692 (0.160)	0.812 (0.093)	Linear SVM, C = 1	thinning of pars triangularis (lh) thinning of transverse temporal cortex (lh) deformation of isthmus of cingulate gyrus (rh)
All modalities	5	0.787 (0.190)	0.749 (0.171)	0.754 (0.147)	0.848 (0.087)	Linear SVM, C = 1	thinning of pars triangularis (lh) thinning of superior temporal gyrus (lh) larger surface area of superior temporal gyrus (lh) thinning of transverse temporal cortex (lh) deformation of isthmus of cingulate gyrus (rh)

Data are shown as mean (SD). All values reported are the average of model performance obtained from the test sets of the 25 randomly generated, unique train-test splits. iRBD: idiopathic rapid eye movement sleep behavior disorder, lh: left hemisphere, MCI: mild cognitive impairment, PD: Parkinson's disease, RBF: radial basis function, SVM: support vector machine, rh: right hemisphere.

Figures

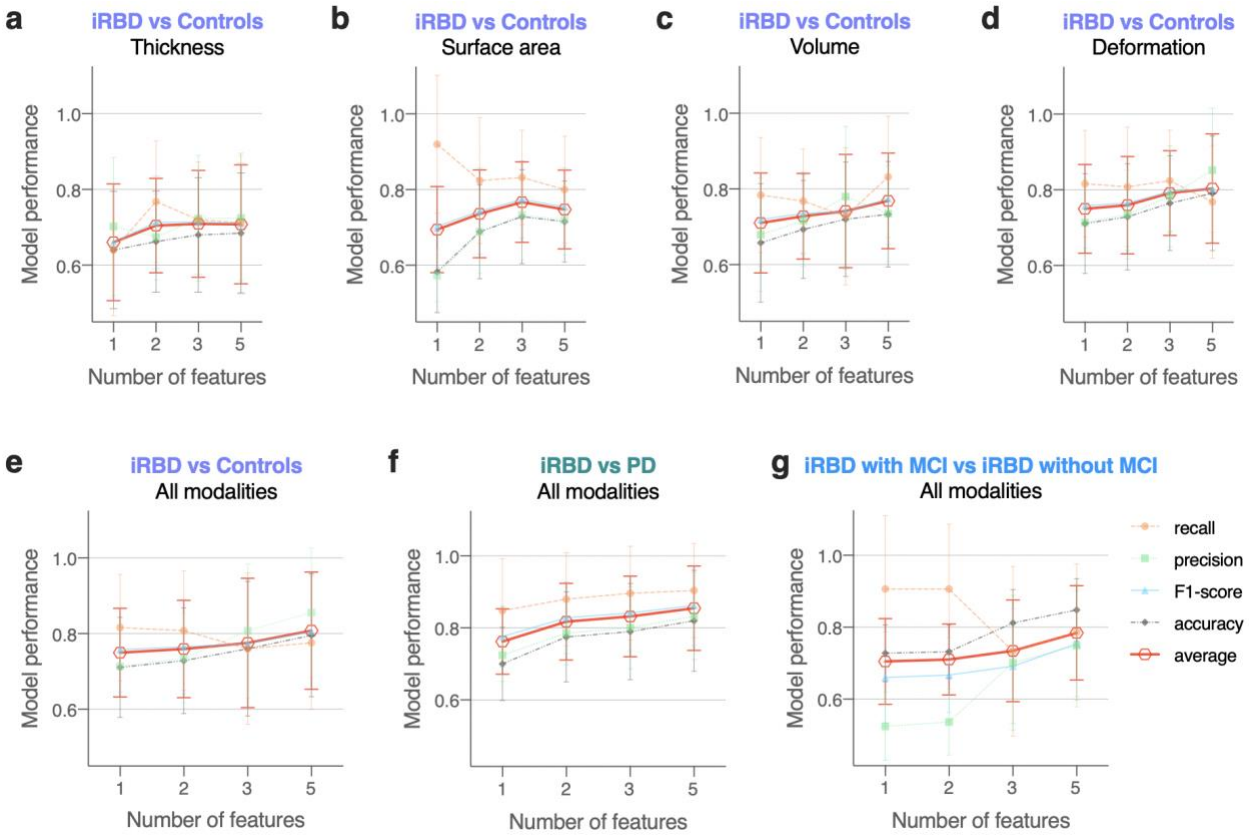

Figure 1. Results of machine learning models with the highest overall performance as measured by recall, precision, F1-score, accuracy, and average of the four metrics. (a-d) Outcomes of individual morphometric modalities in the differentiation between iRBD patients and controls. (e-g) Outcomes when all four morphometric modalities were merged and used in model training, for iRBD patients vs controls (e), iRBD vs PD patients (f) and iRBD patients with MCI vs without MCI (g). iRBD: idiopathic rapid eye movement sleep behavior disorder, PD: Parkinson's disease, MCI: mild cognitive impairment. Error bars indicate standard deviation.

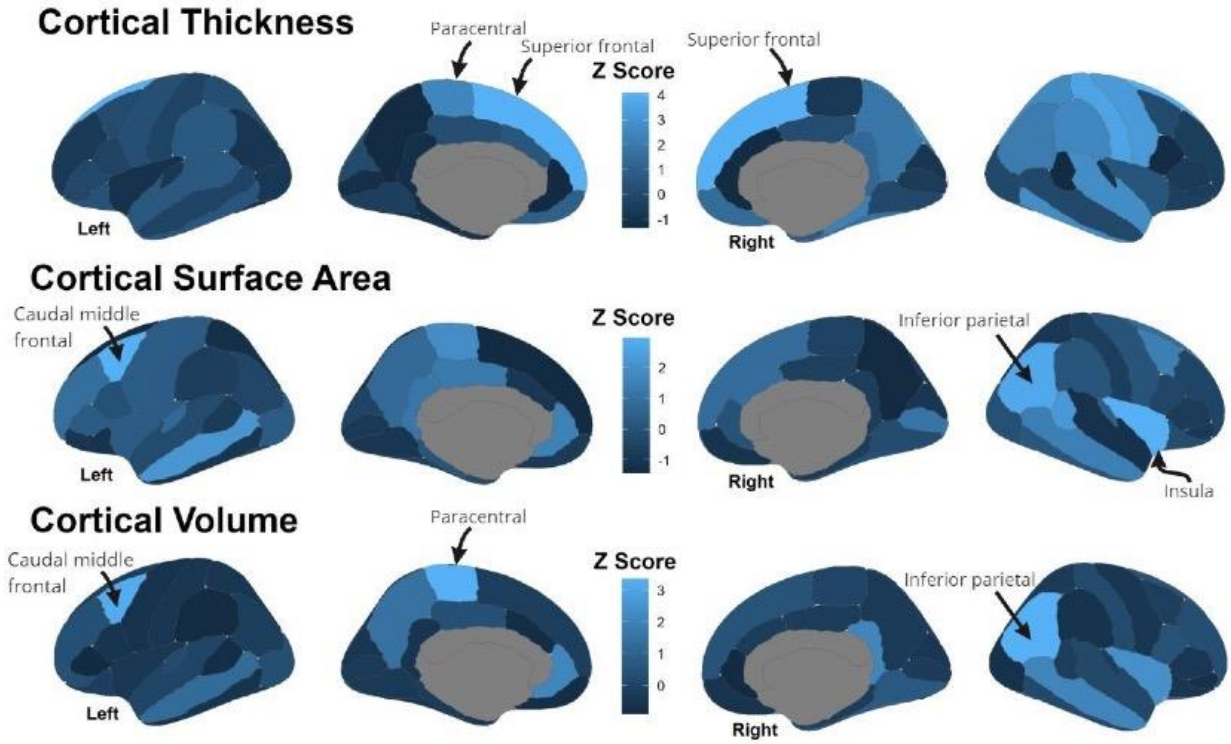

Figure 2. Relative importance of the top 20 most relevant cortical regions in terms of thickness, surface area, and volume in the discrimination between iRBD patients and controls. The dark-to-light blue color bar indicates the Z-scored relative importance of the cortical feature, with lighter regions representing more relevant regions in the discrimination between the two groups. The top three regions in terms of relative importance for each metric are labelled. iRBD: idiopathic rapid eye movement sleep behavior disorder.

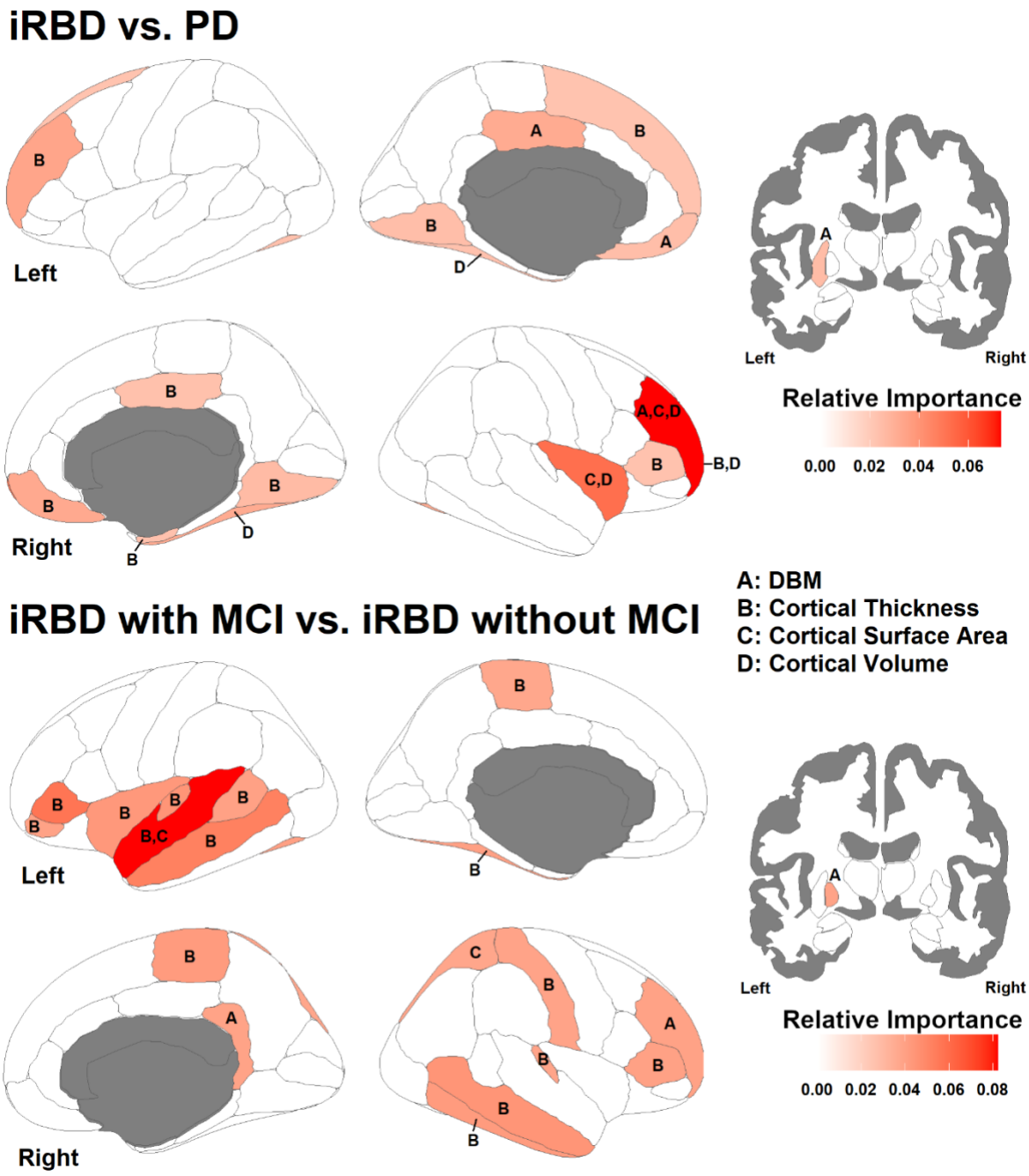

Figure 3. Relative importance of the top 20 most relevant brain regions in the discrimination between iRBD vs PD (top) and iRBD patients with vs without MCI (bottom). The white-red color bar represents the relative importance of the feature in discriminating groups, with redder regions representing a more relevant feature. Every region from the top 20 is labeled from A-D in terms of the contributing structural metric (legend on the right). DBM: deformation-based morphometry, iRBD: idiopathic rapid eye movement sleep behavior disorder, MCI: mild cognitive impairment, PD: Parkinson's disease.