Investigating the power of eyes open resting state EEG for assisting in dementia diagnosisComputing

Jack Jennings (j.l.jennings2@newcastle.ac.uk)
Newcastle University School of Computing Science https://orcid.org/0000-0002-6108-9155
Luis R Peraza
IXICO plc
Mark Baker
Newcastle University Institute for Ageing and Health
Kai Alter
Newcastle University Faculty of Medical Sciences
John-Paul Taylor
Newcastle University Institute for Ageing and Health
Roman Bauer
University of Surrey Department of Computer Science

Research

Keywords: EEG, eyes open (EO), Alzheimer's disease (AD), Dementia with Lewy bodies (DLB)

Posted Date: February 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-245206/v1

Abstract

Introduction: The differentiation of Lewy body dementia from other common dementia types clinically is difficult, with a considerable number of cases only being found post-mortem. Consequently, there is a clear need for inexpensive and accurate diagnostic approaches for clinical use. Electroencephalography (EEG) is one potential candidate due to its relatively low cost and non-invasive nature. Previous studies examining the use of EEG as a dementia diagnostic have focussed on the eyes closed (EC) resting state; however, eyes open (EO) EEG may also be a useful adjunct to quantitative analysis due to clinical availability.

Methods: We extracted spectral properties from EEG signals recorded under research study protocols (1024 Hz sampling rate, 10:5 EEG layout). The data stems from a total of 40 dementia patients with an average age of 74.42, 75.81 and 73.88 years for Alzheimer's disease (AD), Dementia with Lewy bodies (DLB) and Parkinson's disease dementia (PDD) respectively and 15 healthy controls (HC) with an average age of 76.93 years. We utilised k-nearest neighbour, support vector machine and logistic regression machine learning to differentiate between groups utilising spectral data from the delta, theta, high theta, alpha and beta EEG bands.

Results: We found that the combination of EC and EO resting state EEG data significantly increased inter-group classification accuracy compared to methods not using EO data. Secondly, we observed a distinct increase in the dominant frequency variance for HC between the EO and EC state, which was not observed within any dementia subgroup. For inter-group classification we achieved a specificity of 0.87 and sensitivity of 0.92 for HC vs Dementia classification and 0.75 specificity and 0.91 sensitivity for AD vs DLB classification, with a k nearest neighbour machine learning model which outperformed other machine learning methods.

Conclusions: The findings of our study indicate that the combination of both EC and EO quantitative EEG features improves overall classification accuracy when classifying dementia types in older age adults. In addition, we demonstrate that healthy controls display a definite change in dominant frequency variance between the EC and EO state. In future, a validation cohort should be utilised to further solidify these findings.

Introduction

As of 2018 dementia has reached an estimated prevalence of 50 million people worldwide, an expected increase to 75 million cases by 2030 [1]. Dementia with Lewy body (DLB) pathologies being found in 25-45% of dementia cases post-mortem likely showing an under representation in current data [2, 3]. Notably most cases of DLB are misdiagnosed as the more common Alzheimer's disease (AD) [4] .

Alzheimer's disease currently represents the largest proportion of dementia cases globally [5] and is characterised by an irregular build-up of the polypeptide beta amyloid, which form plaques within brain regions such as the hippocampus leading to the damage of neurons [6]. Another common group is

characterised by the Lewy body disease (LBD) pathology; being comprised of Parkinson's disease dementia (PDD) and DLB with a common pathology of Lewy bodies protein building up within specific brain regions[5, 7].

The initial overlap symptomatically of some dementia types is problematic, particularly due to the differential treatment required. For example, DLB it is often misdiagnosed as the more common AD due to presenting similar symptoms early in disease progression such as memory loss and language pathology[4, 8]. Those with DLB are however much more sensitive to commonly prescribed neuroleptics and require more specific and targeted treatment [9]. DLB also has symptomatic overlap in later disease progression stages with the dementia type PDD. This manifests with physical tremors, problems with balance and a shuffling gait in addition to memory loss; differentiation between these two diseases currently is highly subjective as it normally relates to which symptoms manifest first. If cognitive impairment develops first a patient will often be diagnosed with DLB, however if tremors develop first a Parkinson's diagnosis is usually given. New research though has shown the presence of Mild Cognitive Impairment (MCI) in Parkinson's Disease patients (PD) with no current dementia diagnosis[10], which presents a possible further blurring of the line between PDD and DLB. Multiple other similarities displayed between the PDD and DLB neuropathologically compound the belief that these diseases represent regions of a larger spectrum. Commonly both being grouped into the Lewy body dementia group [11]. However, despite similarities the more aggressive disease progression of DLB leads to a shorter life expectancy, likely necessitating the development of a differential treatment between the two LBD dementias in the future.

Electroencephalography (EEG) is a tool for the detection of biomarkers related to surface brain activity and is often used in the clinical diagnosis of other neurological conditions such as epilepsy [12]. EEG records brain activity through use of nodes placed across the scalp in specific locations. The brain activity is recorded in the form of spectral data. EEG node density can vary based upon the specific layout which is quantified via distance. The 10-5 system which was utilised in this study, comprises nodes that are densely packed on the scalp. Data can then be interpreted as a visual spectrum of activity through which quantitative results can be gathered, often referred to as quantitative EEG (qEEG). Some studies have already shown promise for the ability of qEEG to classify between different dementia types[13-15]. Importantly clinical EEG is a non-invasive and cost-effective data acquisition technique at around $300 USD [16] and is already widely available. Clinical EEG systems have a significantly lower cost than current gold standards such as brain MRI or DatScan, with respective costs of $600 USD and $1200 USD each [16], (all prices are based on UK private patient healthcare costs as of 2015). Techniques such as DatScan also require small dosing with radioactive tracers which limits repeat data-acquisition in addition to limited global availability.

Currently, EEG is not used for the diagnosis of dementia as a whole and is instead primarily an interest of research; however EEG has become popular as an auxiliary tool for the clinical diagnosis of DLB patients displaying its utility as a diagnostic tool [17]. The possibility of further expanding EEG as an auxiliary

biomarker detection method in addition to its overall affordability would allow for more frequent screening which could assist in targeted treatment for patients on an individual basis.

As a tool for assisting in dementia research, EEG is most often recorded in an eyes closed resting state (EC) were participants remain awake while performing no task or movements. It has been shown that, during EC resting state EEG dementia patients display a definite decrease in power for the alpha frequency range within the EEG spectrum when compared to healthy participants [18, 19].

A similar procedure can be performed with eyes kept open (EO), commonly used for the diagnosis of conditions such as epilepsy or other seizure related neurological conditions. In this case, participants are asked by their clinician to open and close their eyes [20]. This is currently not widely utilised for dementia research despite the routine EO EEG data acquisition in clinical neurophysiology, as during EO EEG the alpha peak of dementia patients does not display the same significant decrease when compared to healthy participants. Recently however, a significant impairment has been found within the EO resting state alpha reactivity of LBD patients when compared to those with AD; thus representing the need for further investigation into the EO data set [21].

Previous studies that utilised qEEG features as a tool for the classification of AD, DLB and PDD dementia groups have focused on the EC resting state EEG of patients [13-15]. In this study we investigate the hypothesis that the inclusion of EO resting state EEG data represents a valuable source of statistically significant qEEG features, improving classification accuracies in dementia diagnosis.

Methods

The data set used in this study is the same as that used by Luis Peraza et al. (2018) [13], hence the assessments used for participants as well as the process of cleaning EEG-recordings and data acquisition are the same.

Participant Assessment

In total 98 participants were recruited from the North East Region of England. Those with dementia were recruited from a population that had been referred to old age neurology services and clinics. In total 80 dementia patients were recruited for the study; this included 32 AD patients (22 Male, 10 Female), 26 DLB patients (21 Male, 5 Female) and 22 PDD patients (20 Male, 2 Female). Along with these dementia patients 18 age matched healthy controls (11 Male, 7 Female) were also recruited for between group comparisons [13]. A diagnosis of dementia was made by two experienced old-age psychiatrists who followed the clinical diagnosis criteria for DLB [4, 22], the diagnostic criteria for PDD [23] and the AD National institute on Aging-Alzheimer's Association criteria for the diagnosis of AD [8].

For this study participants underwent in-depth neurological and neuropsychiatric testing. The Cambridge Cognition Examination (CAMCOG) and Mini-Mental state exam (MMSE) were used to assess cognitive function in patients, with both tests being commonly used for assessing the extent of a participant's

dementia symptoms. Additionally, the Neuropsychiatric inventory test was performed to assess the severity and frequency of hallucinations for participants (NPI hal).

Electroencephalography and Signal Processing

150 seconds of resting state EEG was acquired using a 128 sintered Ag/AgCl electrode Waveguard cap (ANT Neuro, The Netherlands) placed in a 10-5 positioning system [24] for each participant. Channels were recorded at a sample rate of 1024 Hz with an electrode impedance of no more than 5kΩ [13]. Before the analysis of qEEG features all EEG pre-processing and cleaning was carried out blinded to group membership using EEGLAB[25] MATLAB functions (R2012; MathWorks, Natick Massachusetts) for both EC and EO recordings [13]. After pre-processing and cleaning, the following 7 processing steps were conducted: 1) Baseline components were subtracted from EEG channels and a phase invariant band pass was applied between 0.3 and 54 Hz with a 2^{nd} order Butterworth filter. 2) Bad channels were deleted from EEG recordings based on noise, bad contact or zero amplitude, the Fz reference channel was also deleted for each EEG recording. 3) The full EEG recording was inspected for noisy time bound artefacts that contaminated all channels such as chewing and swallowing. These bad segments were then deleted, the EO state required the deletion of additional segments due to greater noise compared to the EC state. 4) EEG with ICA implemented using fast ICA [26] through EEGLAB default parameters and used to reduce noise via the removal of eye artefacts, the 50 Hz power line component and heartbeats were identified and deleted, with no more than 12 artefactual independent components being deleted. 5) After the initial use of ICA some EEG recordings still showed artefacts and so ICA was repeated, no more than 12 ICs were deleted. 6) Previously deleted channels were then interpolated back using information from surrounding electrodes using spherical interpolation in EEGLAB. 7) EEG data was then referenced to the average reference using EEGLAB[13] with recordings being segmented into 2 second windows with a 1 second overlap.

From the original 98 subjects 65 were used for analysis: including 15 HC, 12 AD, 21 DLB and 17 PDD subjects. The other 33 patients (2 HC, 13 AD, 4 DLB, 4 PDD) were removed from the dataset due to participants not having at least 20 seconds of combined resting state eyes closed or eyes open EEG after cleaning[13], This criterion was employed because 20 seconds of continuous resting state EEG has been shown to be the required amount to account for the inherent variability in EEG [27].

Spectral Analysis

EEG segments were analysed over 5 cortical regions: Frontal (F), Central (C), Temporal (T), Parietal (P) and Occipital (O). Within each of these regions the relative power for the delta (0.5-4 Hz), theta (4-5.5 Hz), high theta (5.5-8Hz), alpha (8-13 Hz) and beta (13-30 Hz) frequency bands in addition to each subject's dominant frequency (DF) and dominant frequency variance (DFV) were calculated using a Welch's Periodogram; 2048 sample segments, using a Hamming window with a 50% overlap between segments as well as a fast Fourier transform size of 2^13; giving a frequency resolution of 0.125 Hz. Here DF is defined as the power spectrum with the highest power frequency bin between 4-15 Hz with an average

taken across the time series frequency bins for the mean, and DFV is the standard deviation across these segments [13]. Relative power spectral density was used as to account for inter-patient variability that is present in non-normalised data, such as neurophysiology and tissue properties [13, 18, 28].

Statistics

The statistical analysis of qEEG features were conducted using the MATLAB statistical toolbox feature (R2018a; MathWorks, Natick Massachusetts) ,WEKA[29] and SPSS (SPSS Statistics for Windows, Version 23.0. IBM). qEEG measures were analysed using a one-way ANOVA for the four groups using a post-hoc unpaired Bonferroni t-test for between group comparison, any features found not to have normally distributed data were assessed similarly using the Kruskal Wallis non-parametric test. Normality of data was tested for by quartile-quartile plots. A p value of less than 0.05 was considered statistically significant ($p < 0.05$), the significance value of the post-hoc tests were automatically corrected so that a p value of 0.05 represented a statistically significant difference between group means.

Feature Selection

Feature selection is critical to the reliability of machine learning algorithms as it prevents the overfitting of data while removing redundant features, such that the overall generalizability of the created classifier is maximised. Feature selection was performed in WEKA [29] and MatLab (R2018a; MathWorks, Natick Massachusetts) using Neighbourhood Component Analysis (NCA). NCA utilises a gradient ascent technique as well as a leave-one-out cross validation to maximise the classification accuracy and to remove redundant features from the data set [30]. For feature selection, participant data was split into training and testing to avoid overestimation of cross-validation accuracy. This feature selection method produced a weight for each feature allowing it to be ranked against all other features for future classification as well as removal of redundant features. We performed wrapper style feature selection, utilising 100 simulated runs to ascertain the consistency of our feature selection method. Wrapped feature selection provides as subset of provided features that improve classification accuracy without adding redundant information. Two bar charts displaying the number of times features improved classification accuracy can be found in Figure 1a and 1b, for differentiation between HC vs dementia groups and AD vs DLB groups, respectively.

Classification and Machine Learning Algorithms

Several supervised machine learning were evaluated using k-fold cross validation (CV) method in MATLAB and WEKA [29] for HC - Dementia, AD – DLB and DLB - PDD classification.

Selected features were used to train machine learning classifiers utilising 10-fold cross validation. The k-Nearest-Neighbour model was employed for classification between different patient groups. Moreover, both logistic regression and support vector machines were used to assess the impact of the EO data set on classification accuracy. For each classifier, a receiver operating curve (ROC) was also examined to determine the area under the curve value (AUC) and the confidence interval (CI) of classification results.

Additionally, logistic regression and support vector machine models were used to compare the effect that inclusion of EO data had on overall classification accuracy; these machine learning methods are compared due to their common use in dementia classification based on EEG data.

Results

Demographics

The demographic and clinical variables of this study are given in Table 1. No significant differences were observed between groups for both age and gender. Similar overall global cognitive impairment was displayed between the dementia groups. However, the AD group displayed significant differences with the DLB and PDD groups in lower CAMCOG memory impairment and NPI hallucinations when utilising a Bonferroni correction ($p < 0.05$), which is displayed in supplementary Table 2. Additionally, it was found that AD patients displayed a significant difference in CAMCOG total score when compared to DLB patients which is not seen when comparing the AD and PDD patients, which can also be found in supplementary Table 2.

Table 1 Demographic and clinical variables for HC, AD, DLB and PDD groups, including descriptive statistics for each variable.

	HC (N = 15)	AD (N = 12)	DLB (N =21)	PDD (N = 17)	p-value
Age	76.93 ± 4.57	74.42 ± 8.97	75.81 ± 6.77	73.88 ± 5.01	F(3,64) = 0.725, p-value = 0.541 [T]
Male/Female	8 \ 7	8 \ 4	18 \ 3	15 \ 2	X^2 (3, N = 65) = 22.71 , p-value = 0.418 [‡]
MMSE	29.13 ± 0.83	20.08 ± 4.94	23.33 ± 4.35	23.18 ± 4.76	F(2,49) = 2.15, p-value = 0.128*
CAMCOG total	96.93 ± 3.67	64.50 ± 21.41	78.57 ± 12.31	74.94 ± 13.47	F(2,49) = 3.291, p-value = 0.046*
CAMCOG executive	22.40 ± 2.10	14.67 ± 6.67	13.90 ± 5.08	12.65 ± 2.87	F(2,49) = 0.638, p-value = 0.533*
CAMCOG memory	23.87 ± 1.18	10.50 ± 5.95	18.67 ± 4.81	17.82 ± 5.08	F(2,49) = 10.41, p-value < 0.001*
CAMCOG attention	6.80 ± 0.56	3.67 ± 2.67	0.67 ± 0.48	0.88 ± 0.33	F(2,49) = 1.68, p-value = 0.198*
NPI hall	0.00 ± 0	0.00 ± 0	0.67 ± 0.48	0.88 ± 0.33	F(2,49) = 7.94, p-value = 0.001*
CAF total	0.00 ± 0	0.64 ± 1.43	3.52 ± 3.95	6.38 ± 4.46	t(36) = 1.95, p-value = 0.058 [‡]
Achel (yes/no)	0 \ 15	11 \ 1	19 \ 2	13 \ 4 [†]	X^2 (3, N = 50) = 2.91 , p-value = 0.573
Years since diagnosis	0 ± 0	1.96 ± 1.23	0.90 ± 0.63	1.47 ± 2.07	H(2) = 9.33, p-value = 0.009 [ᴲ]

[T] Four group ANNOVA
[*] Three group ANNOVA (AD, DLB, PDD)
[‡] X^2 test four groups
[‡] Unpaired t-test (DLB vs PDD)
[ᴲ] Kruskal-Wallis three groups (AD, DLB, PDD)
† One PDD patient was on Memantine

Importantly, there were no significant differences in the usage of cholinesterase inhibitors across dementia groups. The Bonferroni corrected values are displayed in supplementary Table 3.

Neighbourhood component analysis (NCA):

NCA was performed on all data sets to remove redundant features (that could lead to decreased classification accuracies for machine learning due to overfitting). Notably NCA transforms and

maximises the performance of features for utilisation in k-nearest-neighbour [31]. We performed these simulation 100 times to ascertain model consistency, and as can be seen in Figure 1a and 1b the overall model consistently chooses the same features for classification across multiple runs, with 4 features chosen in >95% of runs for HC-Dementia and 3 features for AD-DLB. For HC vs dementia the EC Frontal High Theta, EC Central Theta, EC Occipital Delta, and the ratio between the EC and EO dominant frequency variance (DFV) in the parietal region were chosen; for AD-DLB EC Frontal Beta, EO Frontal High Theta and EO Parietal High Theta; for AD-DLB/PDD the EC Temporal DFV, EO Frontal High Theta, EO Parietal High Theta, EO Occipital High Theta and EO Occipital DFV and finally for DLB-PDD classification only the EC Parietal Alpha and EO Occipital Delta features were chosen. Features were utilised for machine learning if they were picked in 95% of simulated runs, a full list of features tested for the HC-D group is displayed in supplementary Figure 1.

Dominant frequency, dominant frequency variance and theta alpha ratio.

From the one-way four group ANOVA it was found that there were significant differences for the eyes closed DF between the HC and dementia groups in the Parietal and Occipital regions, with dementia groups displaying a mean slowing in their DF towards the high-theta frequency range. Additionally, we found the same significant difference between the HC and dementia patients DF in the EO resting state within the same regions. These results being displayed in supplementary Table 4.

Similarly, we investigated the significance of DFV in the EC and EO state between groups. By computing the ratio between the two states we found that healthy controls displayed a significant decrease in variance in the EC compared to EO state. Notably, this change was found to be significant across all cortical regions for the healthy control groups when compared to dementia groups. For the AD, DLB and PDD groups we found no significant change in DFV between the EC and EO states. These results are displayed in supplementary Table 5. Additionally, we investigated the EEG data scroll of participants between the two states and found that HC participants displayed an apparent change in frequency that was not seen for dementia patients, with some exemplary examples shown in supplementary Figure 2.

The ratio of the theta and alpha relative power (TAR) also showed a significant difference between the HC and dementia groups within the occipital region in both the EC and EO resting state. The parietal region showed a significant difference between the HC and the dementia groups within the EC state. However, in the EO state only the DLB and PDD groups were found to have a significantly greater TAR than the HC groups.

Machine Learning Classification

We investigated inter-group group separability using machine learning classification for EC and EO spectral features. To this end, we only used features which had been chosen via feature selection. A k-Nearest-Neighbour model classifier was chosen for inter-group classification. A k value of 10 was selected for cross fold validation between different participant groups. A summary of the results for

classification between healthy controls and dementia patients in addition to classification between the AD, DLB and PDD groups can be found in Table 2 for EC and EO qEEG features.

Table 3 summarises the classification results between the AD and DLB patient groups when utilising only EC qEEG features and combining EC and EO qEEG features, for the k-nearest neighbour algorithm, logistic regression, and a quadratic support vector machine.

Additionally, we investigated the change in classification accuracy for AD vs DLB classification when one includes clinical scores. With CAMCOG memory score being chosen during feature selection alongside EC Frontal beta and EO parietal high theta. With a marked increase in accuracy for identifying patients with DLB. These results being presented in Table 4.

Table 2: Comparison of classification results between HCs and Dementia patients as well as AD and DLB using 10-fold cross validation. Utilising a K-Nearest-Neighbour machine learning model, in addition we present the Confidence interval (CI) for the specificity and sensitivity of each classification type.

Machine Learning classification accuracies				
ClassificationType	**Accuracy**	**Specificity (±CI)**	**Sensitivity (±CI)**	**Weighted AverageAUC**
HC – D	0.91±0.07	0.87±0.09	0.92±0.08	0.85
AD-LBD	0.86±0.10	0.75±0.19	0.9±0.13	0.76
AD – DLB	0.82±0.13	0.75±0.19	0.81±0.17	0.74
DLB-PDD	0.61±0.16	0.76±0.20	0.3±0.22	0.61

Table 3: AD – DLB Classification results comparison for EC and EC-EO classification using 10-fold cross validation, with comparisons across 3 separate machine learning models: k-nearest neighbour, logistic Regression, and support vector machine.

AD - DLB classificaiton accuracies				
EC + EO data				
Classifier	**Accuracy**	**Specificity (± CI)**	**Sensitivity (± CI)**	**Weighted AverageAUC**
Cosine KNN	0.82±0.13	0.75±0.19	0.81±0.17	0.74
LogisticRegression	0.76±0.15	0.67±0.20	0.86±0.15	0.84
QuadraticSVM	0.82±0.13	0.58±0.21	0.95±0.09	0.82
EC data only				
Cosine KNN	0.67±0.16	0.5±0.21	0.76±0.18	0.63
LogisticRegression	0.73±0.15	0.58±0.21	0.81±0.17	0.77
QuadraticSVM	0.73±0.15	0.58±0.21	0.81±0.17	0.72

Table 4: Improved AD – DLB Classification results when including CAMCOG memory total score for classification, done for both EC + EO and EC only classification.

AD - DLB classifications with CAMCOG memory inclusion				
Data	**Accuracy**	**Specificity**	**Sensitivity**	**Weighted AverageAUC**
EC + EO	0.88±0.12	0.75±0.19	0.95±0.09	0.94
EC	0.73±0.27	0.67±0.20	0.76±0.18	0.71

Discussion

Through spectral data analysis it was found that dementia participants, within our study, displayed a definite mean EEG slowing between 4-13 Hz. This slowing is most prominently seen as a decrease in dominant frequency in the occipital and parietal brain regions when comparing between healthy controls and dementia patients. This decrease in DF can be most prominently seen in the DLB and PDD groups, with the theta-alpha relative power ratio of both groups also being significantly different form the healthy control group for both regions. These findings are in agreement with already well documented literature on EC EEG slowing [13, 14, 18, 19]. Additionally, our results suggest that there is still a significant decrease in the dominant frequency of dementia patients when compared to healthy controls in the occipital and parietal regions in the EO state, this data is presented in supplementary Table 4.

In healthy controls, we found a significant increase in dominant frequency variance (DFV) in the EC state compared to the EO state as shown in Figure 2. This difference in DFV between EC and EO was not seen in any dementia group. This difference was yet undocumented in relevant literature. It is with possible that there may be further features within the EO data set that have yet to be investigated. We display the EEG data scrolls in the occipital regions for a HC, AD, DLB and PDD participant in supplementary Figure 2, with the occipital region being chosen as not only do we see the greatest difference in DFV between states for HC participants in this region, but these nodes also display prominent alpha rhythms for HC participants. We found that the HC participant displayed a decrease in wavelength in the EO state when compared to the EC state, which is not the case in the AD, DLB or PDD participants. This may indicate a reason for the change in DFV of HC participants.

For our first inter-group comparison we investigated the classification between HC and dementia participants when utilising features selected through feature reduction. Utilising the k-neatest-neighbour (KNN) machine learning method we achieved a specificity of 0.87 and sensitivity of 0.86. Good classification accuracies were achieved, in line with classification results being comparable with those of Peraza *et al* [13] who utilised more complex network analysis EEG features.

Secondly, we investigated the AD – DLB classification type, achieving results of 0.75 and 0.86 for sensitivity and specificity, respectively. These classification results are similar to those of Zande et al [14] who achieved a specificity of 0.87 and sensitivity of 0.83 and M.T Pascareli *et al* who achieved a diagnostic result of 0.75 sensitivity and 0.85 specificity [32] or R Mehraram *et al* whom achieved 0.47 sensitivity and 1.00 specificity [33]. Unlike our investigation other studies combined either the EC resting state qEEG of participants with network derived connectivity measures [14] or investigated inter-group differences using EEG cortical sources which are more difficult to acquire in clinical settings [32], without the inclusion of EO qEEG data. Notably, our results for AD-DLB classification are comparable to FP – CIT – SPECT, which currently acts as the gold standard in diagnosis of dementia patients [34-36]. However, unlike EEG, FP-CIT-SPECT scans are limited in repeat data-acquisition due to requiring the use of single gamma photon emission for imaging. Additionally, we investigated AD vs DLB/PDD classification, which achieved a specificity of 0.75 and a sensitivity of 0.90 (with all PDD patients being successfully

differentiated from those with AD). Finally, we investigated the classification accuracy between the DLB and PDD dementia groups with a specificity of 0.76 and a sensitivity of 0.30. The ability to achieve such classification results between AD and DLB through the utilisation of only qEEG features from the EC and EO EEG data set is highly applicable to assisting in dementia classification in a clinical environment due to the relative simplicity of data acquisition and interpolation without the need for a more complex method. Finally, we investigated the impact of standard dementia testing scores, such as CAMCOG, when combined with the spectral qEEG data. For AD vs DLB classification we achieved the same specificity of 0.75 but with a marked increase in sensitivity to 0.95, displaying the power of combining simple EC and EO qEEG features with standard clinical testing data to differentiate between these dementia groups.

To investigate the overall impact of the EO data on classification, we also compared the results of the three widely used classifiers: KNN, Support vector machines (SVM) and Logistic regression (LR). It was found that when only EC qEEG features were utilised, all three machine learning methods performed worse in overall classification results, summarised in Table 2. Overall, these results indicate that EO qEEG data represents an underutilised source of data for assisting in the diagnosis and classification of dementia, but a larger validation cohort is required to further cement the strength by which

EO features improve inter-group classification accuracies.

Further research into the utilisation of EC and EO qEEG for assisting in the diagnosis of dementia should likely investigate the combination of network derived functional connectivity methods, as this will likely further improve classification accuracies. EC data has already been utilised in studies for comparing derived and weighted network measures which show a significant difference between AD and LBD groups [13, 14, 33], thus investigating EO derived connectivity measures could further improve such inter-group classification methods. Moreover, the inclusion of structural information (MRI and DTI) is likely to provide additional valuable information to further improve diagnostic accuracy.

Limitations

While our work constitutes a proof-of-principle, future work leveraging data from larger participant cohorts will decrease overfitting and help to improve classification statistics.

Conclusions

Our study has found that EO qEEG data contains several features with significant inter-group variability for the classification between healthy controls and dementia groups, in addition to dementia inter-group analysis. By including such EO qEEG features, we achieved similar classification results to previous qEEG studies using more complex computational methods. Overall, the combination of EC and EO qEEG features improve the overall classification accuracy for HC-D and AD-DLB.

Additionally, we found a yet undocumented phenomenon between healthy controls and dementia patients when one compares the difference in DFV across brain regions in the EC and EO states. Healthy

controls displayed a significantly higher DFV in the EO state, which was not seen in the dementia patients. This is possibly representing an underlying biomarker for inter-group classification; this should be further investigated with functional connectivity methods.

Overall, our study presents EO qEEG as an underutilised significant data set for assisting in the diagnosis of dementia patients in combination with EC data, improving intergroup classification accuracy for classifying between HC-D and AD-DLB data sets. Additionally, we present undocumented differences between the HC and dementia groups which are not seen while utilising only the EC data set.

Abbreviations

Abbreviation	Full name
Participant Groups	
HC	Healthy Control
AD	Alzheimer's Disease
DLB	Dementia with Lewy Bodies
PDD	Parkison's Disease Dementia
LBD	Lewy Body Dementia (DLB + PDD)
Brain Region	
G	Global
F	Frontal
C	Central
T	Temporal
P	Parietal
O	Occipital
EEG Terms	
EEG	Electoencephalography
QEEG or qEEG	Quantitative EEG data
EC	Eyes closed resting state
EO	Eyes open resting state
HSR	1024Hz High sample rate
LSR	256Hz Low sample rate
DS	Down sampled
Feature Terms	
RDP	Relative Delta power
RTP	Relative Theta power
RAP	Relative Alpha power
RBP	Relative Beta power
TAR	Theta Alpha power ratio
DF	Dominant Frequency
DFV	Dominant Frequency Variance
EC/EO-X-DFV	Ration of eyes closed eyes open DFV in X brain region
Machine Learning	
PCA	Principal component analysis
KNN	K-Nearest-Neighbour
SVM	Support vector macine
LR	Logistic Regression
CI	Confidence Interval
AUC	Area under the curve
Sens	Sensitivity
Spec	Specificity
Machine Learning Groups	
HC-D	Healthy Controls against dementia groups
AD-LBD or AD-DLB/PDD	AD compared to DLB and PDD combined
AD-DLB	AD group compared to DLB
DLB-PDD	DLB group compared to PDD

Declarations

Author Contributions:

All authors in some way contributed to the work presented within this paper.

J.L.Jennings: executed the project, formal analysis, investigation and methodology, writing and editing of the manuscript.

L.R.Peraza: Performed the cleaning of all raw EEG data, provided code utilised within the project and reviewing the manuscript.

M.Baker: Assisted in research direction and reviewing of the manuscript.

J.P.Taylor: Funding for clinical cohort recruitment, design and delivery of original EEG study, Assisted in research direction and reviewing of the manuscript.

K.Alter: Assisted in research and development in addition to reviewing the manuscript.

R.Bauer: Provided funding for the study, designed the study, assisted in research and development of the work as well as reviewing the manuscript.

Funding:

R.B. and J.J. were supported by the Medical Research Council of the United Kingdom (MR/N015037/1) and the Engineering and Physical Sciences Research Council of the United Kingdom (EP/S001433/1). J.J. was also supported by a PhD studentship funded by the School of Computing at University of Newcastle. J-PT is supported by National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre (BRC) based at Newcastle upon Tyne Hospitals NHS Foundation Trust and Newcastle University. Participant recruitment and data collection was funded by an intermediate clinical Wellcome Trust Fellowship (WT088441MA) to J.P Taylor.

Acknowledgements:

We would additionally like to acknowledge Julia Schumacher, Peter Taylor and Nishant Sinha for participating in thought provoking discussions about the work presented within this paper.

Conflicts of Interest:

The authors declare no conflict of interest.

Availability of data and materials:

The dataset used and analysed during this study in addition to the relevant code are available from J.Jennings upon reasonable request.

Consent for publication:

Not applicable to this study.

Ethics approval and consent to participate:

Not applicable to this study, Ethical approval was already approved for data acquisition and study in the work of Peraza *et al* [13] whom data set was used for our investigation.

References

1. Patterson, C., *World Alzheimer Report 2018. The state of the art of dementia research: New frontiers.* Alzheimer's Disease International, 2018: p. 48.
2. Schneider, J.A., et al., *The Neuropathology of Older Persons with and Without Dementia from Community versus Clinic Cohorts.* Journal of Alzheimer's Disease, 2009. **18**: p. 691-701.
3. Jon B Toledo, N.J.C., Xiao Da, Kewei Chen, Deborah Carter, Adam Fleisher, Erin Householder, Napatkamon Ayutyanont, Auttawut Roontiva, Robert J Bauer, Paul Eisen, Leslie M Shaw, Christos Davatzikos, Michael W Weiner, Eric M Reiman, John C Morris, John Q Trojanowski, *Clinical and multimodal biomarker correlates of ADNI neuropathological findings.* Acta Neuropathologica Communications, 2013. **1:65**.
4. McKeith, I.G., et al., *Diagnosis and management of dementia with Lewy bodies,Fourth consensus report of the DLB Consortium.* Neurology, 2017. **89**(1): p. 88-100.
5. Prince, M., Knapp, M, Guerachet, M, McCrone, P, Prina, M, Comas-Herra, A, Wittenberg, R, Adelaja, B, Hu, B, King, D, Rehill, A, and Salimkumar, D., *Dementia UK: Second Edition -Overview.* Dementia UK: Second Edition, 2014: p. 24.
6. M. Paul Murphy , H.L., *Alzheimer's disease and the beta-Amyloid Peptide.* J Alzheimers Dis, 2010. **19**: p. 17.
7. McKeith, I., *Dementia with Lewy bodies and Parkinson's disease with dementia: where two worlds collide.* Practical Neurology, 2007. **7**: p. 374-382.
8. McKhann GM, K.D., Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH., *The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.* Alzheimers & Dementia, 2011. **7**(3): p. 263-269.
9. Taylor, J.-P., et al., *New evidence on the management of Lewy body dementia.* The Lancet Neurology, 2020. **19**(2): p. 157-169.
10. Weil.R.S, C.A.A., Schrag.A.E *Mild Cognitive Impairment in Parkinson's Disease : What is it ?* Current Neurology and Neuroscience Reports, 2018. **18**(17).
11. Jellinger, K.A. and A.D. Korczyn, *Are dementia with Lewy bodies and Parkinson's disease dementia the same disease?* BMC Med, 2018. **16**(1): p. 34.
12. Smith, S.J.M., *EEG In the Diagnosis, Classification and managment of patients with epilepsy.* Journal of Neurology, Neurosurgery and Psychiatry, 2005. **76**: p. 6.

13. Luis R. Peraza, et al., *Electroencephalographic derived network differences in Lewy body dementia compared to Alzheimer's disease patients.* Scientific Reports, 2018(8).
14. Jessica J. van der Zande, A.A.G., Inger van Steenoven, Philip Scheltens, Cornelis Jan Stam and Afina W. Lemstra, *EEG Characteristics of Dementia With Lewy Bodies, Alzheimer's Disease and Mixed Pathology.* Frontiers in Aging Neuroscience, 2018. **10**.
15. Heinrich Garn, C.C., Markus Waser, Georg Caravias, Gerhard Ransmayr, *Differential diagnosis between patients with probable Alzheimer's disease, Parkinson's disease dementia, or dementia with Lewy bodies and frontotemporal dementia, behavioral variant, using quantitative electroencephalographic features.* Journal of Neural Transmission, 2017. **124**(5): p. 569 - 581.
16. NHS, S.U.H., *Private patient price list*. 2014.
17. McKeith, I.G., et al., *Diagnosis and management of dementia with Lewy bodies.* Neurology, 2017. **89**(1): p. 88.
18. Myrto Stylianoua , N., Luis R.Perazaa , Sara Graziadiob , Ruth Cromarty , Alison Killen , John T.O' Brien , Alan J.Thomasa , Fiona E.N.LeBeau , John-PaulTaylor, *Quantitative electroencephalography as a marker of cognitive fluctuations in dementia with Lewy bodies and an aid to differential diagnosis.* Clinical Neurophysiology, 2018. **Volume 129**(Issue 6): p. Pages 1209-1220.
19. Christina Micanovic , S.P., *The diagnostic utility of EEG in early-onset dementia: a systematic review of the literature with narrative analysis.* Journal of Neural Transmission, 2014. **121**(1): p. 59-69.
20. NHS. *Diagnosis, Epilepsy.* 2020 [cited 2020; Available from: https://www.nhs.uk/conditions/epilepsy/diagnosis/.
21. Schumacher, J., et al., *EEG alpha reactivity and cholinergic system integrity in Lewy body dementia and Alzheimer's disease.* Alzheimer's Research & Therapy, 2020. **12**(1): p. 46.
22. McKeith IG, D.D., Lowe J, Emre M, O'Brien JT, Feldman H, Cummings J, Duda JE, Lippa C, Perry EK, Aarsland D, Arai H, Ballard CG, Boeve B, Burn DJ, Costa D, Del Ser T, Dubois B, Galasko D, Gauthier S, Goetz CG, Gomez-Tortosa E, Halliday G, Hansen LA, Hardy J, Iwatsubo T, Kalaria RN, Kaufer D, Kenny RA, Korczyn A, Kosaka K, Lee VM, Lees A, Litvan I, Londos E, Lopez OL, Minoshima S, Mizuno Y, Molina JA, Mukaetova-Ladinska EB, Pasquier F, Perry RH, Schulz JB, Trojanowski JQ, Yamada M., *Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium.* Neurology, 2005. **65**(12): p. 1863-1872.
23. Emre, M.e.a., *Clinical diagnostic criteria for dementia associated with Parkinson's disease.* Movement disorders: official journal of the Movement Disorder Society 2007. **22**: p. 1689-1707.
24. Oostenveld, R.P., P., *The five percent electrode system for high-resolution EEG and ERP measurements.* Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology, 2001. **112**: p. 713-719.
25. Neuroscience, S.C.f.C., *EEGLab Wiki.* 2018.
26. Hyvarinen, A., *Fast and robust fixed-point algorithms for independent component analysis.* IEEE Transactions on Neural Networks 1999. **10**(3): p. 626-634.

27. T. Gasser, P.B., H. Steinberg, *Test-Retest Reliability of Spectral Parameters of the EEG.* Electroencephalography and Clinical Neurophisiology, 1985. **60**: p. 7.
28. Guido Rodriguez , F.C., Paolo Vitali, Germana Perego, Flavio Nobili, *EEG spectral profile to stage Alzheimer's disease.* Clinical Neurophysiology, 1999. **110**(10): p. 1831 - 1837.
29. Eibe Frank, M.A.H., and Ian H. Witten, *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques".* 2016.
30. W. Yang, K.W., W. Zuo, *Neighborhood component feature selection fo High-Dimensional Data.* Journal of Computers, 2012. **7**(1): p. 8.
31. Goldberger, J., et al., *Neighbourhood Components Analysis*. Vol. 17. 2004.
32. Pascarelli, M.T., et al., *Abnormalities of resting-state EEG in patients with prodromal and overt dementia with Lewy bodies: Relation to clinical symptoms.* Clinical Neurophysiology, 2020. **131**(11): p. 2716-2731.
33. Mehraram, R., et al., *Weighted network measures reveal differences between dementia types: An EEG study.* Hum Brain Mapp, 2020. **41**(6): p. 1573-1590.
34. John T. O'Brien, I.G.M., Zuzana Walker, Klaus Tatsch, Jan Booij, Jacques Darcourt, Moritz Marquardt and Cornelia Reininger for the DLB Study Group*, *Diagnostic accuracy of 123I-FP-CIT SPECT in possible dementia with Lewy bodies.* The British Journal of Psychiatry, 2009. **194**: p. 5.
35. Timo S. Spehl, L.F., Sabine Hellwig, Cornelius Weiller, Michael Hüll, Philipp T. Meyer, Florian Amtage, *Role of Semiquantitative Assessment of Regional Binding Potential in 123I-FP-CIT SPECT for the Differentiation of Frontotemporal Dementia, Dementia With Lewy Bodies, and Alzheimer's Dementia.* Clinical Nuclear Medicine, 2015. **40**(1): p. 27-33.
36. McKeith I, O.B.J., Walker Z, Tatsch K, Booij J, Darcourt J, Padovani A, Giubbini R, Bonuccelli U, Volterrani D, Holmes C, Kemp P, Tabet N, Meyer I, Reininger C, *Sensitivity and specificity of dopamine transporter imaging with 123I-FP-CIT SPECT in dementia with Lewy bodies: a phase III, multicentre study.* The Lancet, Neurology, 2007. **6**(4): p. 305-313.

Figures

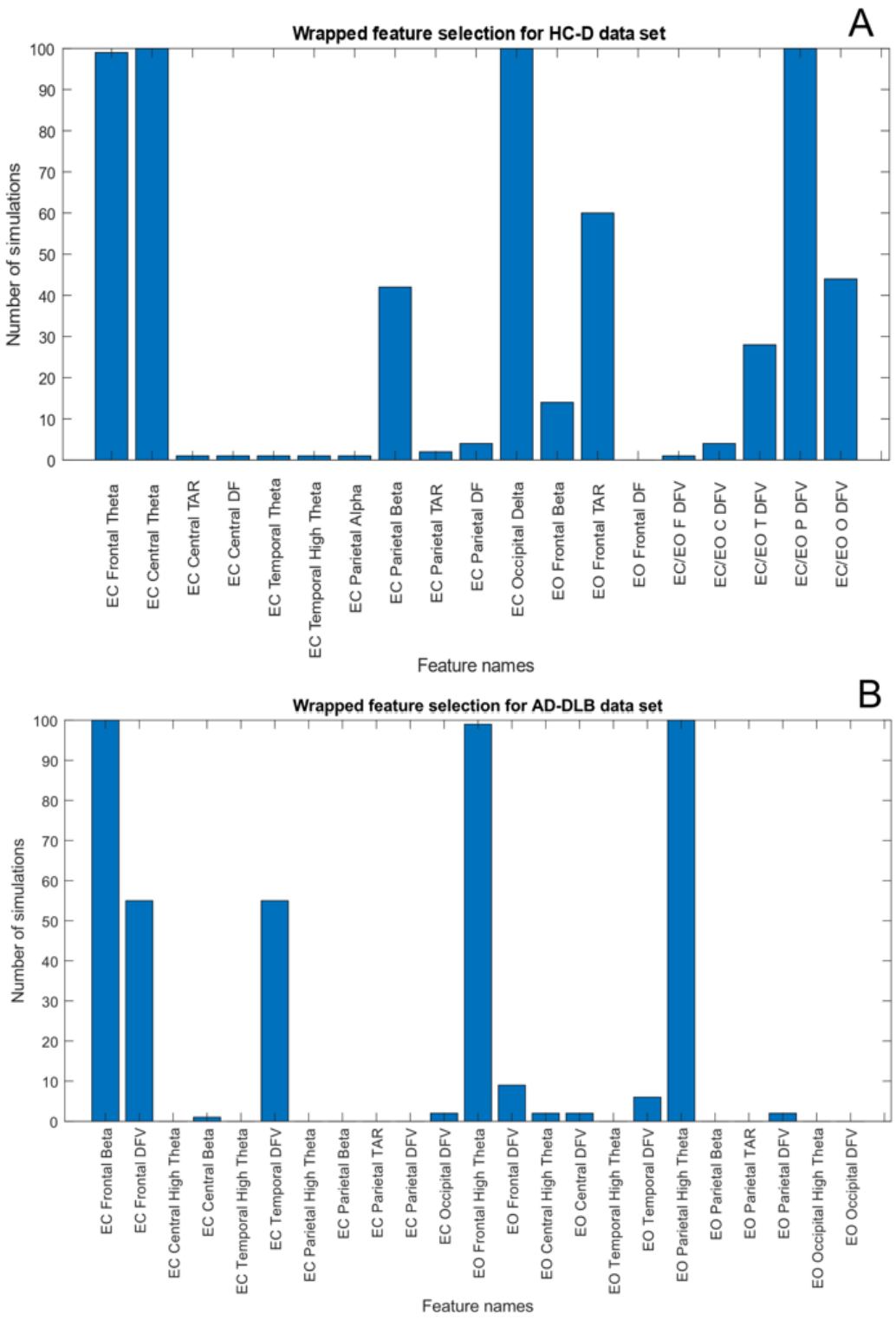

Figure 1

Figures showing the total number of times that features for HC-D (A) and AD-DLB (B) classification. Wrapped feature selection utilised training and testing data sets and was simulated 100 times such that model consistency could be ascertained. In both cases several features were selected consistently, with other features which were selected less adding redundant information that did not improve classification accuracy. With features consisting of the relative delta, theta, high theta, alpha and delta power in

addition to the ration of the hightheta-alpha relative power (TAR) dominant frequency (DF), dominant frequency variance (DFV) and the ratio of the dominant frequency variance between the EC and EO state (EC/EO).

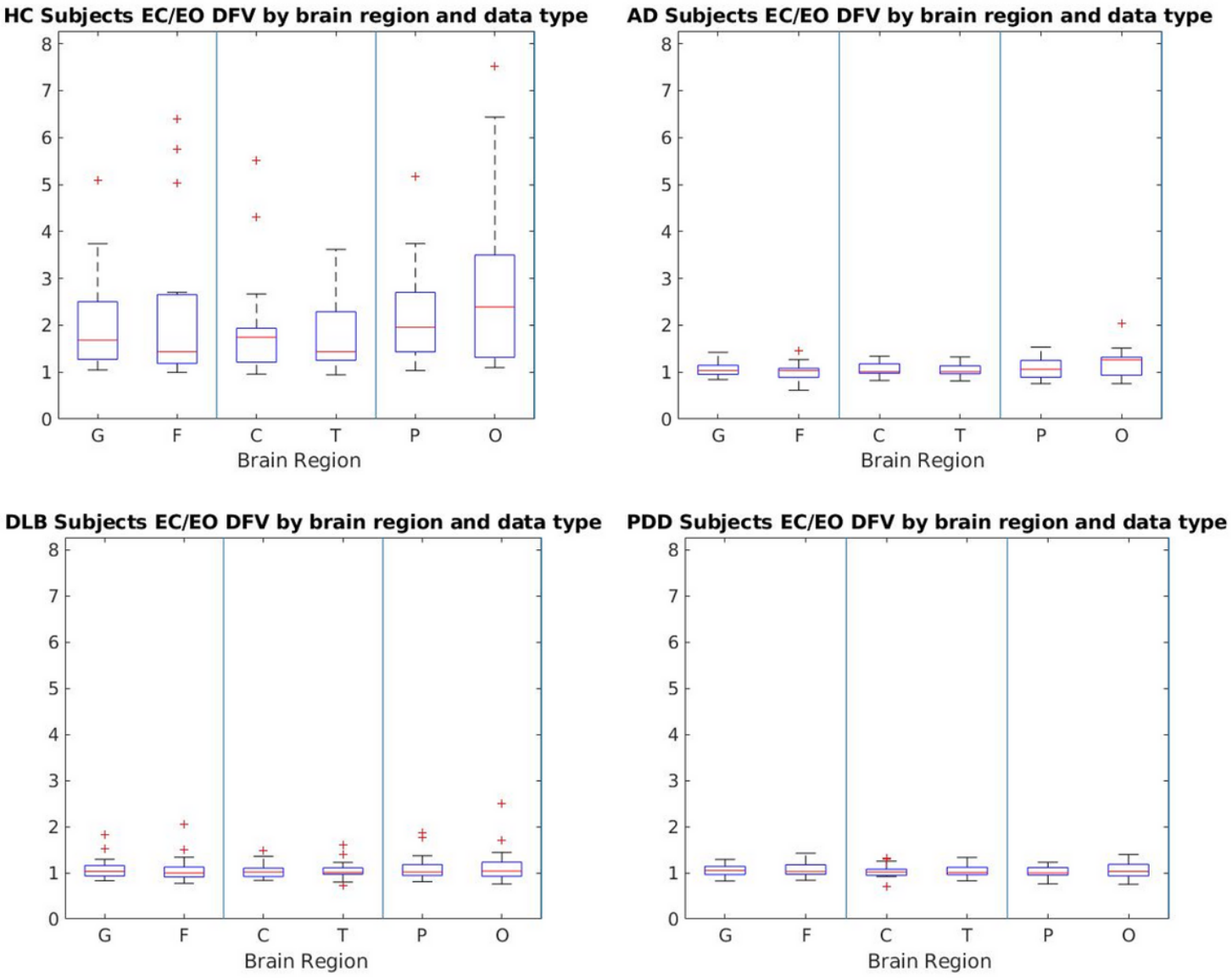

Figure 2

Box plots for the dominant frequency variance (DFV) ratio in the eyes closed (EC) and eyes open (EO) resting state for HC, AD, DLB and PDD participants across all cortical regions. These boxplots display a possible difference between healthy and dementia participants when comparing eyes closed and open states that has yet been uncommented upon in literature for inter-group differentiation and may be representative on an underlying biomarker. HC group showed a significant difference ($p < 0.05$) in comparison to all dementia groups for the ratio of EC and EO DFV. In addition, no dementia group was found to have a significant difference with any group other than HC, as shown in supplementary Table 5. – is the median DFV value of the group. With + representing outliers.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx