Regularized Linear Discriminant Analysis of EEG Features in Dementia Patients

Emanuel Neto[1,2]*, Felix Biessmann[3], Harald Aurlien[1], Helge Nordby[2] and Tom Eichele[1,2,4]

[1] Section for Clinical Neurophysiology, Haukeland University Hospital, Bergen, Norway, [2] Institute of Biological and Medical Psychology, University of Bergen, Bergen, Norway, [3] Amazon Development Center Germany, Berlin, Germany, [4] K.G. Jebsen Center for Neuropsychiatric Disorders, Bergen, Norway

The present study explores if EEG spectral parameters can discriminate between healthy elderly controls (HC), Alzheimer's disease (AD) and vascular dementia (VaD) using. We considered EEG data recorded during normal clinical routine with 114 healthy controls (HC), 114 AD, and 114 VaD patients. The spectral features extracted from the EEG were the absolute delta power, decay from lower to higher frequencies, amplitude, center and dispersion of the alpha power and baseline power of the entire frequency spectrum. For discrimination, we submitted these EEG features to regularized linear discriminant analysis algorithm with a 10-fold cross-validation. To check the consistency of the results obtained by our classifiers, we applied bootstrap statistics. Four binary classifiers were used to discriminate HC from AD, HC from VaD, AD from VaD, and HC from dementia patients (AD or VaD). For each model, we measured the discrimination performance using the area under curve (AUC) and the accuracy of the cross-validation (cv-ACC). We applied this procedure using two different sets of predictors. The first set considered all the features extracted from the 22 channels. For the second set of features, we automatically rejected features poorly correlated with their labels. Fairly good results were obtained when discriminating HC from dementia patients with AD or VaD (AUC = 0.84). We also obtained AUC = 0.74 for discrimination of AD from HC, AUC = 0.77 for discrimination of VaD from HC, and finally AUC = 0.61 for discrimination of AD from VaD. Our models were able to separate HC from dementia patients, and also and to discriminate AD from VaD above chance. Our results suggest that these features may be relevant for the clinical assessment of patients with dementia.

Keywords: Alzheimer's disease, vascular dementia, electroencephalogram, qEEG, quantitative analysis, spectral features, group classification, LDA

Abbreviations: AD, Alzheimer's disease; AUC, area under curve; CSF, cerebrospinal fluid; CT, computerized tomography; CV, cross-validation; cv-ACC, accuracy of the cross-validation; DSM-IV-TR, diagnostic and statistical manual of mental disorders-text revised; EEG, electroencephalogram; FFT, fast Fourier transform; FPR, false positive rate; HC, healthy controls; ICD, international classification of disease; LDA, linear discriminant analysis; MRI, magnetic resonance imaging; PC LDA, principal component linear discriminant analysis; PC LR, principal component logistic regression; PET, positron emission tomography; PLS, least square linear discriminant analysis; PLS LR, partial least square logistic regression; PSD, power spectral density; qEEG, quantitative electroencephalogram; RLDA, regularized linear discriminant analysis; ROC, receiver operator characteristic; SEM, standard error mean; SVM, support vector machines; TPR, true positive rate; VaD, vascular dementia.

INTRODUCTION

Alzheimer's disease is the most common form of dementia among the elderly population (Jeong, 2004; Kandimalla et al., 2011, 2013, 2014). The second most common form of dementia is VaD (Roman, 2002), affecting approximately 20% of all dementia cases worldwide (Dubois and Hebert, 2001). Presently, there are no specific *in vivo* tests for VaD and AD.

The clinical assessment of dementia is grounded in guidelines from the World Health Organization's ICD and the American Psychiatric Association's DSM-IV-TR. These guidelines have been criticized for their lack of specificity (Reilly et al., 2010; McKhann et al., 2011). If fact, many intruments are available to screen dementia, such as CSF measures, CT, MRI examinations, EEG, ultrasound, PET, as well as combination of neurological and psychological tests and laboratory blood analysis. However, many of these examinations are expensive, partially invasive or demand large resources. Furthermore, at early stages these dementias present several symptomatic similarities, creating diagnostic uncertainty (Gearing et al., 1995; Massoud et al., 1999). Consequently, finding alternative methods to detect dementia and classify subtypes is a relevant research topic.

EEG Based Dementia Diagnosis

Electroencephalogram is a widely available and non-invasive instrument (Rossini et al., 2007). The clinical assessment of an EEG is based on the visual expert interpretation of patient's electrophisiological activity in a spatio-temporal scale. Classical EEG biomarkers such as relative power and dominant activity rythms in conventional frequency bands and at specific brain regions were shown to be valuable measures to screen dementia (Dierks et al., 1995; Signorino et al., 1996; Jelic et al., 1998). Additionally, the literature has shown that spectral analysis of EEG may provide alternative markers to distinguish AD or VaD patients (Signorino et al., 1996; Besthorn et al., 1997; Bonanni et al., 2008). The spectra of VaD patients show increased power in delta and theta frequencies and decreased power in alpha and beta frequencies (Tsuno et al., 2004; Lou et al., 2011; van Straaten et al., 2012). This phenomenon has also been described by other studies and is associated with general cognitive decline (Dierks et al., 1995; Kwak, 2006; Babiloni et al., 2011b; Fraga et al., 2013). AD and VaD patients are commonly described as having reduced frequency of the posterior dominant rhythm compared with healthy subjects (Babiloni et al., 2004; Gawel et al., 2009). When comparing AD with VaD patients, the literature reports two core electrophysiological differences between these two groups: AD patients have lower posterior alpha power when compared with VaD; the lower frequencies power appears to be higher in VaD compared to AD (Signorino et al., 1995, 1996; Babiloni et al., 2004, 2011a).

Extracting Features from EEG Signals

There are many EEG features that one could potentially use for discrimination purposes. FFT and power spectral density (PSD) are two of the most widely used transforms that allow to extract potential markers such as frequency, power, coherence across the delta, theta, alpha, or beta bands (Dierks et al., 1991; Kwak, 2006; Fraga et al., 2013). However, the analysis of multi-channel EEG data results in high-dimensional data vectors including spatiotemporal information creating a high computational requirement for solving discrimination problems. Another important aspect is the contamination of EEG with muscular or ocular artifacts such as eye movement and eye blinks (Hillyard and Galambos, 1970; Nunez and Srinivasan, 2006). Artifacts are ubiquitous during the EEG recording and contribute to the non-stationarity of EEG signals (Guger et al., 2000; Shenoy et al., 2006). Non-stationarity of the EEG (von Bunau et al., 2009) constitutes one of the major challenges for data analysis and machine learning classification methods (Kaplan et al., 2005)

Classification Algorithms

Many different classification algorithms are available. In dementia, it has been shown that it is possible to discriminate patients with different levels of cognitive decline through the use of linear and non-linear classification algorithms (Cichocki et al., 2005; Lopez et al., 2009; Elgendi et al., 2011; Gallego-Jutgla et al., 2012; Gutman et al., 2013). Lehmann et al. (2007) tested several algorithms commonly used for identifying AD patients. Specifically, they employed PC LDA, PLS LDA, PC LR, PLS LR, bagging, random forests, SVM, and others, to identify dementia patients. This study concluded that such techniques may indeed provide remarkable performances (accuracies of 70–90%) for clinical diagnosis. However, they emphasized that the classification models should not be trained with small sample size when compared to their feature dimensionality, as this may result in overfitting the model. This problem is often referred in the literature as the curse of dimensionality (Bishop, 1995). In addition, the feature extraction and discrimination methods vary from study to study, making assessment of reproducibility difficult. One classification method that has been widely used in aging studies, specifically to identify AD is LDA (Dierks et al., 1995; Kwak, 2006; Fraga et al., 2013).

LDA

The purpose of the LDA classification is to assign observations to the corresponding class based on a set of measurements or predictors by finding an optimal linear transformation that maximizes the class separability (Fisher, 1936). Technically, LDA achieves optimal solutions when predictors or the feature vector is multivariate normally distributed in each group class and when the different group classes have similar covariance (Fisher, 1936). However, due to the non-stationary of EEG signals (Shenoy et al., 2006), this is rarely the case in practice, and models can easily be overfitted, and their predictability overestimated. This is typically the case when EEG data is derived from multi-channels carrying noise captured from adjacent channels, especially if channels are close enough to each other. Whereas this problem could be solved by applying spatial filters that maximize the variance of EEG signals of one class while minimizing the variance from the other class (Ramoser et al., 2000; Blankertz et al., 2008), at the same time, this may result on overfitting the model by changing predictor variances (Reuderink and Poel, 2008; Huang et al., 2010). In addition, large amount of predictors and

relatively small sample sizes dramatically induce misclassification errors contributing to overfitting the model. To circumvent this, many classifiers implement feature reduction based on screening and excluding the features that carry the less amount of information regarding the prediction problem. However, this procedure sometimes eliminates significant features from the data affecting the performance of the classification models (Dauwels et al., 2010). Many of these questions have been addressed in a recent review (Haufe et al., 2014) determining the importance of implementing complementary analysis methods to avoid the bias typically found in generalized linear models (GLM) implementations.

Regularization with Cross-Validation

The regularization technique is based on replacing the within-group sample covariance by a weighted average of the whole sample covariance using a shrinking intensity parameter (λ). Technically, this parameter increases larger eigenvalues of the covariance matrix while decreasing smaller ones, therefore creating a pooled-covariance matrix that is corrected for the bias when estimating sample-based eigenvalues. The optimal shrinkage parameter is determined by CV (Friedman, 1989; Ye et al., 2006). CV is a technique to estimate the classification error rate by splitting the original sample data into training, CV, and test datasets (Bishop, 2006). The prediction model is calibrated using the training sets, and model parameters are optimized by the CV sets, while the test sets are used primarily for empirical error estimation. An effective approach to address both the sensitivity and dimensionality problems is using regularized LDA (RLDA) with CV denoted in the literature as CV-RLDA. This approach has been used in several studies with high dimensional and non-stationary problems (Goulermas et al., 2005; Maggi et al., 2006; Huang et al., 2008) achieving very high accuracy levels. It has been shown that the choice of the regularization value or technique has great impact on the overall discrimination performance (Lotte and Guan, 2011).

Aim of this Study

The literature contains a number of EEG studies revealing several biomarkers that are relevant in dementia group differentiation. However, the methodologies of pre-processing and extracting EEG markers are not identical and sample sizes are typically small. In order to understand dementia at a wider level, we believe that the methodologies to extract EEG biomarkers should be unified across studies. Hence, in our former work, we proposed a feature extraction method that reveals six core EEG parameters at channel-level from any standard clinical EEGs (Neto et al., 2015). We have also shown significant differences between AD, VaD, and HC for several of those parameters at specific channels, in accordance with the present literature.

The aim of the present study was to evaluate the relevance of six EEG parameters on differentiating groups of patients with dementia. To assess this, we designed four discrimination models using LDA to distinguish *HC* from *AD*, *HC* from *VaD*, *AD* from *VaD* and *HC* from both *AD* and *VaD*. Moreover, we tested the performance of our four models using two different settings of parameters and measured the class separability performance of our models, which we refer to in the methods section.

MATERIALS AND METHODS

We applied LDA to generate four binary LDA classifiers on previously extracted EEG features (Neto et al., 2015). In particular, we used six electrophysiological markers that characterize the EEG spectrum and tested them in group discrimination between AD, VaD, and HC using CV-RLDA. In line with other studies, (Besserve et al., 2007; Zhdanov et al., 2007; Shenoy et al., 2008; Velu and de Sa, 2013), the present study used CV-RLDA methodology for group discrimination using six EEG spectral features. We explored differentiation performances between AD and VaD patients by repeatedly rejecting features with low correlation with their corresponding labels.

Sample

A total of 342 patients were selected for this study, based on the clinical ICD diagnoses. The first group ($n = 114$) was registered as probable AD patients, with ICD codes ICD-10, F00.x or G30.x. The second group ($n = 114$) was patients with probable VaD ($n = 114$) with ICD-10 F01.X. The HC group ($n = 114$) was selected from non-hospitalized individuals, free of any central nervous-system disease or any other brain disease. We controlled for clinical historic record of all subjects and excluded patients whose diagnose had been changed or removed. Only the most recent EEG from each patient was included. All samples were age-gender balanced with age of 72.9 ± 10.5 years and 48% females. We used these datasets to extract spectral features and test their significance for group discrimination. For further details, the reader is referred to Neto et al. (2015).

Pre-processing

All EEG datasets were acquired using 22 channels positioned in 10–20 system placements (Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, T3/T7, C3, Cz, C4, T4/T8, T5/P7, P3, Pz, P4, T6/P8, O1, O2, M1, M2), acquired at 128 Hz ($n = 86$), 256 ($n = 246$), and 500 Hz ($n = 10$) using NicoletOne[TM] EEG system. Input impedances were set to $Z > 100\text{M}\Omega$. Hardware single pole high-pass (0.16 Hz $\pm$ 10%) and low-pass (500 Hz $\pm$ 10%) filters were applied with individually to each channel before pre-amplification. EEGs were stored under raw format in the database. All the pre-processing and data analysis was done accordingly to our previous work (Neto et al., 2015), using Mathworks® Matlab environment. EEG raw files were imported to the EEGLab v.10.1.1.0b toolbox (Delorme and Makeig, 2004) using an in-house data-reader. Data were resampled to 256 Hz. From the standard clinical EEG recording protocol that lasts for 20 min and includes eye open/closed conditions, hyperventilation and provocations with photic stimulation, we restricted the input data for analysis to the first 9 min, which contained only the alternating eyes open/closed resting conditions. A 1536-points high-band filter was applied at cut off frequency of 0.5 and a low-pass filter to cut off of 50 Hz using a standard least square linear-phase FIR filter design.

EEGs were segmented into non-overlapping epochs of 1 s that were evaluated for possible rejection using automatic amplitude, power, and statistical thresholding. The remaining segments were subjected to an individual independent component analysis (ICA) using the Infomax algorithm with 15 components in order to identify and remove residual contributions from eye movements. The continuous data were reconstructed from the non-artifact components and then segmented into 2 s epochs with 1 s overlap, which is equivalent to the Welch's procedure (Welch, 1938) with a rectangular windows and 50% segment overlap. Subsequently, the data were transformed into the frequency domain using FFT. Since the frequency spectrum selected for the pre-processing was from 0.5 to 50 Hz, we obtained 100 frequency data points for the 22 channels and a variable number of epochs for each dataset subjected for analysis. The spatial standard deviation (sSTD) index of each epoch was calculated across the 22 channels in the frequency domain according to Lehmann and Skrandies (1980) and z-scored. Finally, in order to standardize the amount of data across subjects, we determined the minimum number of existing epochs across subjects that would maximize the inclusion of patients in this study. We considered a total of 334 epochs from each subject (~5 m 30 s), representing the artifact-free data for 22 channels (sampled at 256 Hz). We then represented each segment by its equivalent frequency domain data and further used them on the feature extraction method.

Feature Extraction and Feature Selection

We used a fit-curve model (Neto et al., 2015) which enables us to represent the spectra with six parameters ranging from 0.5 to 30 Hz (S, k, A, c, w, and b). For each spectrum, S represents low frequency power (delta), k indicates the 1/f decay from lower to higher frequencies where larger values of k denote a faster drop-off in power. Parameters A, c, and w relate to the amplitude, center and dispersion of the alpha power, respectively. Parameter b represents a global offset or *baseline* power of the entire frequency spectrum. We then applied the curve-fit model to the average spectral curve of each channel. Each final dataset was therefore represented by a total 132 parameters, which correspond to the six extracted spectral features from twenty-two channels.

Discrimination Analysis

To test the potential of such features in group discrimination, we tested two different feature sets to generate four binary classifiers which discriminated the classes HC *versus* AD *(Model 1)*, HC *versus* VaD *(Model 2)*, AD *versus* VaD *(Model 3)*, and HC *versus* all dementia patients AD or VaD *(Model 4)* as illustrated in **Figure 1**. These were all implemented by using 10-fold cv-RLDA. The two different set of features are referred as *Complete set of Features* and *Reduced set of Features*. The *Complete Set of Features* comprised all available parameters that were extracted from our fit-curve feature extraction procedure for the 22 channels of each dataset resulting on a total of 132 model predictors. The Reduced set of features is a sub-set of all 132 features. However, we decided to remove features that were poorly correlated with their labels. To determine this, before training the linear classifiers, we
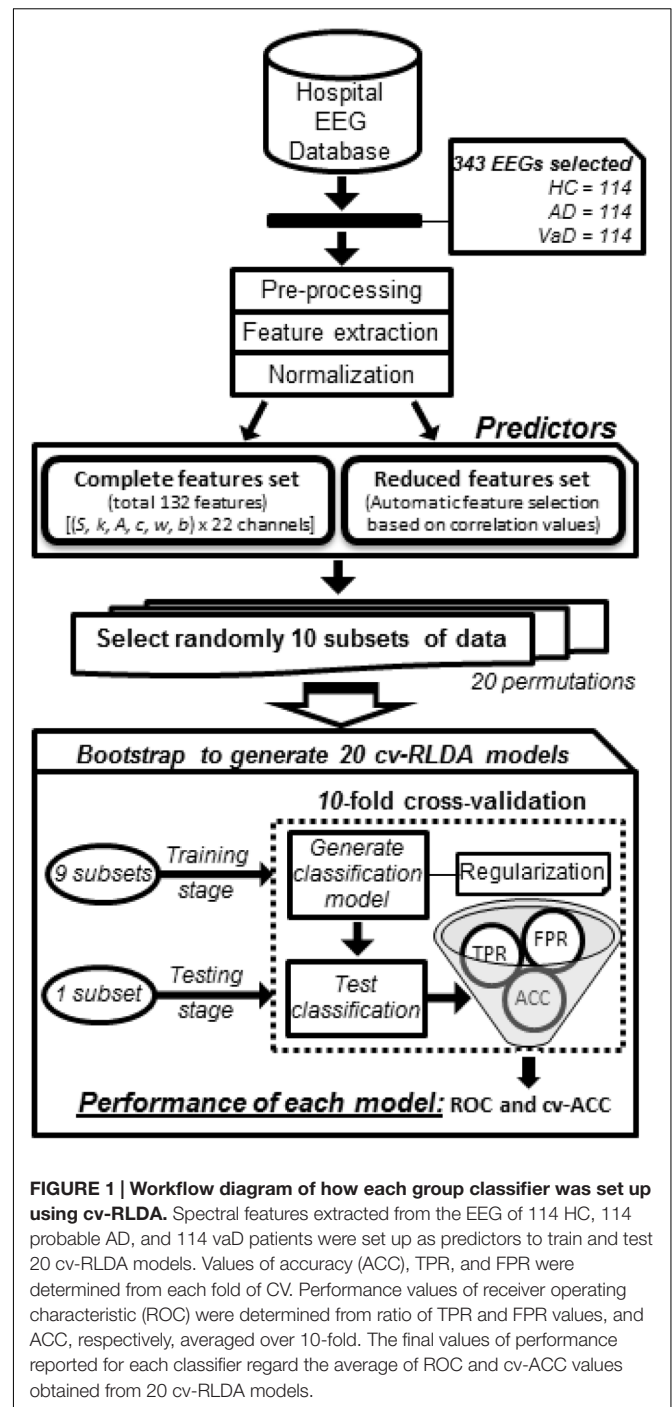
FIGURE 1 | Workflow diagram of how each group classifier was set up using cv-RLDA. Spectral features extracted from the EEG of 114 HC, 114 probable AD, and 114 vaD patients were set up as predictors to train and test 20 cv-RLDA models. Values of accuracy (ACC), TPR, and FPR were determined from each fold of CV. Performance values of receiver operating characteristic (ROC) were determined from ratio of TPR and FPR values, and ACC, respectively, averaged over 10-fold. The final values of performance reported for each classifier regard the average of ROC and cv-ACC values obtained from 20 cv-RLDA models.

computed the correlation of each feature with the target variable, we tested the correlation of each feature with their correspondent labels. We only considered data points in the respective training data set of each particular CV fold. The output of this test is a vector with correlation values (r) and their correspondent p-value (p). The exclusion criteria was to reject features with significantly low correlation index at $|r| < 0.15$ and $p < 0.01$ (two tailed t-test). The procedure of excluding features is highly dependent to which datasets are in use when generating each model. Since,

we used different sets of data to train each model, we also obtained different values of correlation and *p*-values. Hence, a variable number of features were rejected. The average numbers of used Reduced Set of Features were 74, 68, 18, and 76 for *Model 1*, *Model 2*, *Model 3*, and *Model 4*, respectively.

cv-RLDA Classifier Models

In order to use the extracted spectral features as predictors on our LDA models, we verified the fundamental assumptions of the LDA (Worth and Cronin, 2003). All predictors met the criteria of being scalar and positive. To transform the predictor matrix into an approximately normal distribution, we applied a log-transformation using box-cox transformation (Box and Cox, 1964) and measured the average, standard deviation and covariance values for the predictors on each group, which we reported as supplementary material showing the normal probability distribution of the predictors before and after the normalization process (Supplementary Table S1; Supplementary Figures S2–S5).

The predictor matrix for each model was constituted by 228 rows that referred to the subjects from each test and a variable number of columns, (132 or 29) that corresponded to the total predictors used in the *Total* or *Reduced Set of Features*, respectively.

To evaluate the linear classification model, we used 10-fold CV. We divided our 228 samples into 10 randomized subsets and determined a fold size of 22. For each fold of the CV, we considered nine subsets for the training and one subset for the testing phases. Therefore, 198 and 22 different datasets were considered on each fold as the training and testing sample, respectively. To set the regularization threshold constant of the regularized LDA model, we used the analytical shrinkage estimator (Ledoit and Wolf, 2004). We alternatively tried nested CV, but found the analytical estimator to yield slightly better results. The permutation order of the training sample may have an impact on the result of the *k*-fold CV (Triba et al., 2015). Therefore, to further validate the robustness of our model, we used bootstrap and generated each model 20 times, each one with the same procedure as described previously, but with a different random selection of data points.

Performance Tests

To benchmark the performance of our final models, we estimated the average of TPR and FPR from each group classifier. Using these metrics, we determined the ROC curve to represent the classification performance of each model (**Figures 2** and **3**). The ROC displays the trade between the sensitivity and specificity of classification from each model. The AUC for each group classifier provided a measure of the discrimination power which has been used as a gold standard diagnostic marker (Zhou et al., 2011). In addition, we determined the accuracies of the models measuring the CV accuracy. Each model's cv-ACC was estimated by averaging the accuracies obtained on each 10-fold classification tests. To avoid bias from the order of datasets used during the calibration stage, we iterated 20 times the generation of each of the four models using bootstrap of the *test* sample. For each iteration, a different permutation of the training

and testing datasets were chosen obtaining twenty different CV accuracies (cv-ACC) and 20 different areas under curve (AUC) for each model. The final performance results of each model were determined by averaging the twenty cv-ACC and AUC results obtained from each group classifier (**Table 1**).

RESULTS

We calculated the R^2 measure for each fit, reflecting the fraction of data variance captured by curve approximation. Additionally, we computed the correlations between model parameters to check for dependence between parameters, which can indicate model redundancies or instabilities in the fitting procedure. Histograms of the R^2 values were presented in Supplementary Figure S1 for each group, and show excellent model fits in nearly all cases. The median goodness-of-fit was 0.91 and the first and third quartiles were 0.92 and 0.98, respectively. We performed a Kruskal–Wallis one-way ANOVA test and no group differences were found in the R^2 values between the groups ($p < 5.15e^{-11}$), meaning that the model performed equally well fitting the datasets from each group. Only relatively weak correlations ($|r| \leq 0.4$) between the parameters were found, with the exception of a high correlation between parameters k and b ($r = 0.75$). This correlation was equally present within each group: $r = 0.80$ for HC, $r = 0.68$ for AD, and $r = 0.78$ for VaD.

Figures 2 and **3** show an overlay of the receiver operating characteristic (ROC) curves obtained for each group classifier using the Complete Set of Features and Reduced Set of Features.

The FPR and TPR results obtained from each group classifier were averaged across the 20 correspondent model's results with mean and corresponding SEM values (**Figures 2** and **3**). We found no significant differences of classification (cv-ACC and AUC) for the different permutations of subsets.
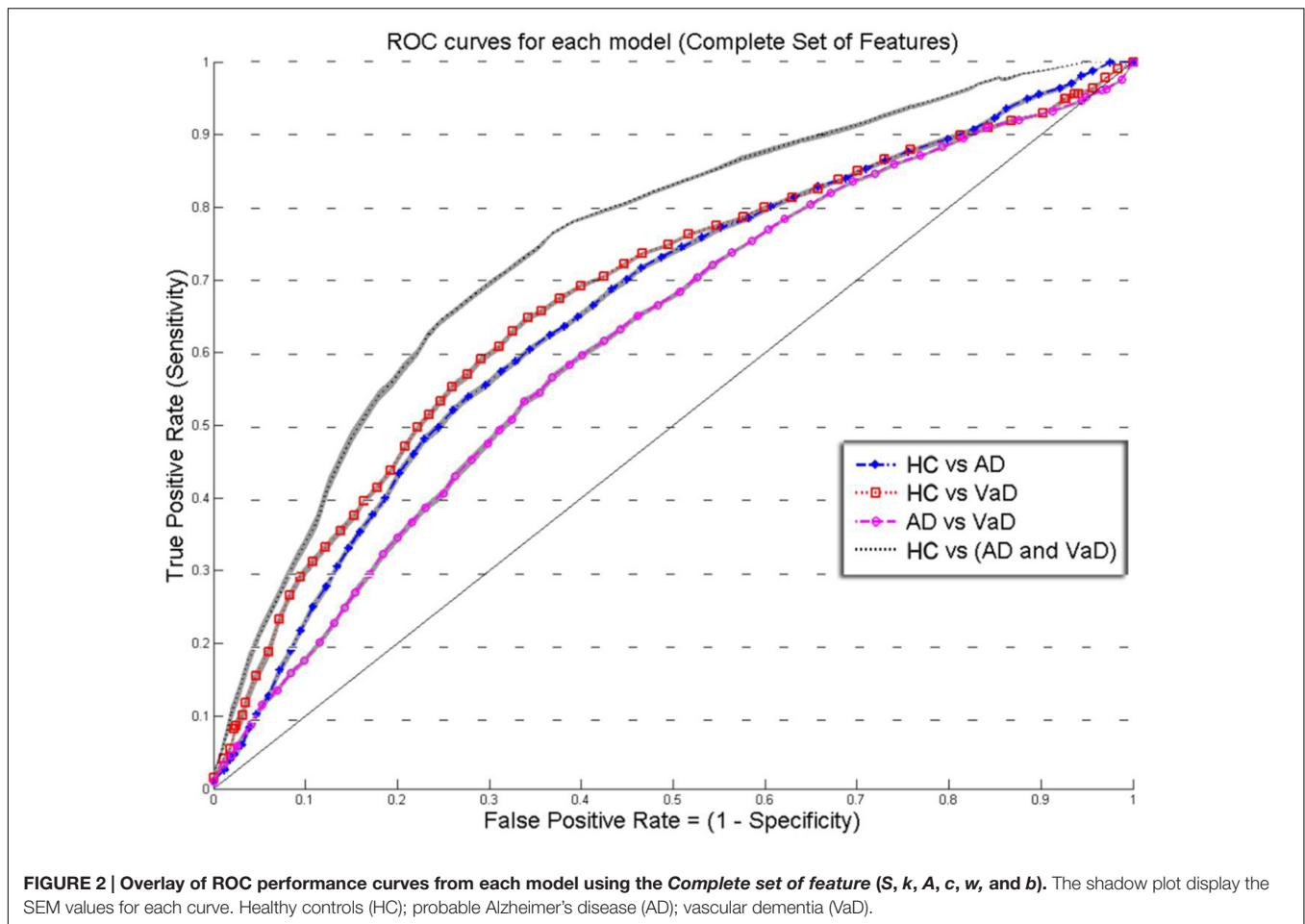
When using the Complete Set of Features, the first three RLDA models, *HC* vs *VaD*, *HC* vs *AD*, and *AD* vs *VaD*, achieved *poor* discrimination performance with AUC of 0.66, 0.66, and 0.62, respectively. The fourth model attained *fair* discrimination for *HC* from *AD* or *VaD* patients with AUC = 0.75. **Figure 2** displays an overlay of all four ROC curves with SEM values from each classifier for the *Complete Set of Features*.

When using the Reduced Set of Features, the first two models, *HC* vs *VaD*, *HC* vs *AD*, obtained *fair* discrimination performances with AUC of 0.74 and 0.77, respectively. The third model discriminated poorly *AD* vs *VaD* with AUC = 0.61. The fourth model discriminated good *HC* from *AD* or *VaD* patients with AUC = 0.83. **Figure 3** displays an overlay of all four ROC curves from each classifier for the *Reduced Set of Features*.

The corresponding cv-ACC values of each model and AUC median, first and third quartiles are provided as Supplementary Table S2.

DISCUSSION

Using features selected from routine EEG recordings and averaged across time for every channel, our results show that

FIGURE 2 | Overlay of ROC performance curves from each model using the *Complete set of feature* (*S, k, A, c, w,* and *b*). The shadow plot display the SEM values for each curve. Healthy controls (HC); probable Alzheimer's disease (AD); vascular dementia (VaD).
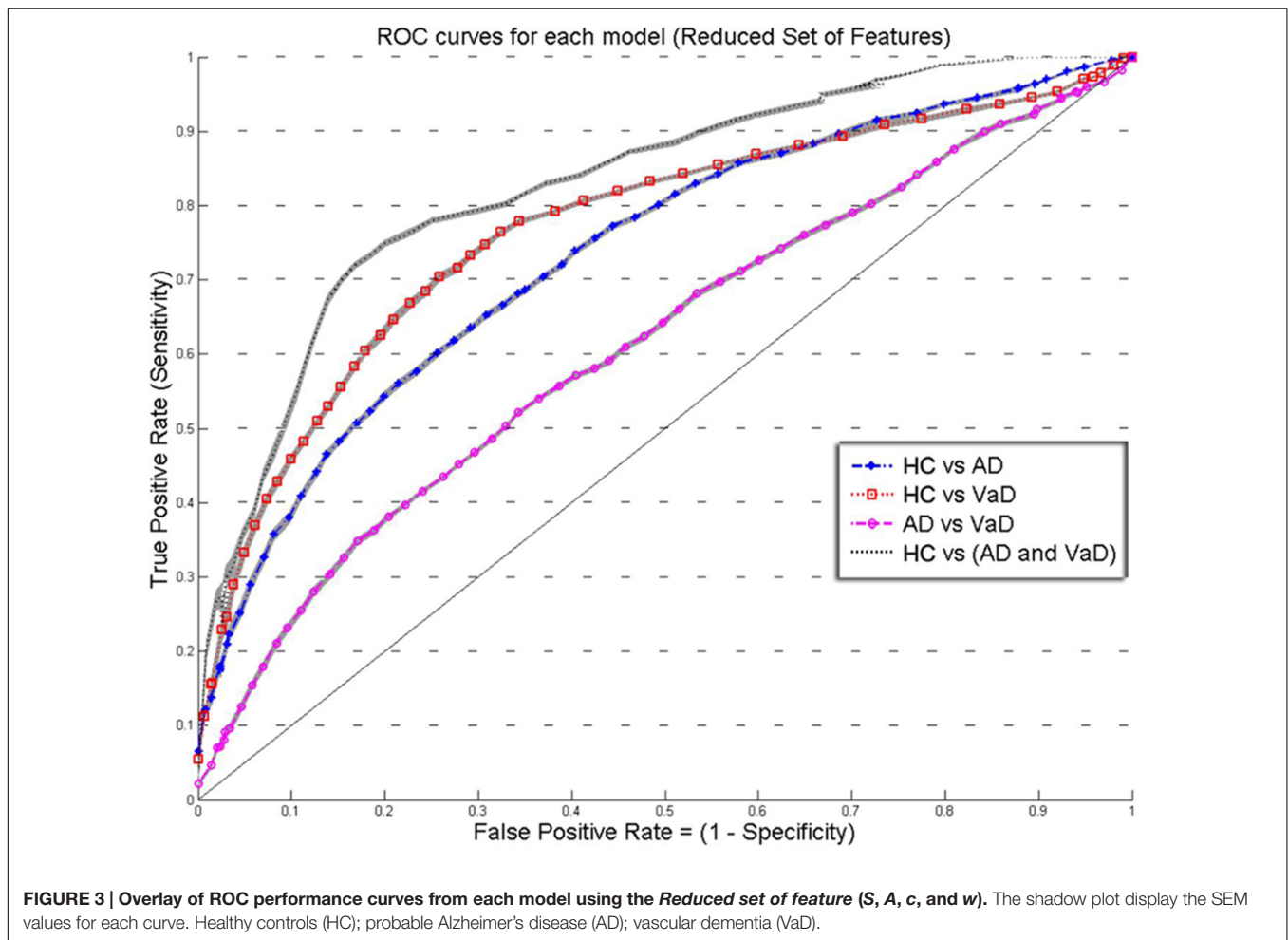
HC are separable from dementia patients. Furthermore, we were also able to differentiate AD from VaD above chance. Clinical practitioners may not always have the ability to assess the predominant biomarker values of an EEG recording. Quantitative EEG spectral analysis is an important tool that allows extracting such features. We have demonstrated that the tested features are reasonably robust biomarkers for the discrimination of dementia groups.

The extracted features were normalized across groups and with similar class covariance. Under these conditions, the LDA should have optimal discrimination performance for the set of predictors used. We extended our analysis by using a *Reduced Set of Features* excluding predictors which accounted with small or not statistically significant correlation for their corresponding classes. By removing these predictors, we expected better discrimination performance of our models as dimensionality was reduced. The performance of our classifier models was *Good* or *Fair* but not *Excellent*. Both results of performance (AUC) and accuracy (cv-ACC) of all models were improved when the same models were generated using the *Reduced Set of Features*.

Model 1: The discrimination between HC and AD, was initially determined with AUC = 0.66. As expected when generating the same classification model with the *Reduced*

set of features, the performance increased significantly to AUC = 0.74 and the cv-ACC model was also improved from cv-AUC = 0.62 to cv-AUC = 0.67, suggesting that these features contain good discriminatory information. The core features that have consistently been shown to describe AD patients is the decrease in alpha power frequency at posterior brain regions and a general power increase of delta and theta rhythms when compared with HC (Huang et al., 2000; Babiloni et al., 2004; Jeong, 2004; Prichep, 2005). Comparing to this model, similar discrimination performances were achieved by other studies, reporting accuracies ranging from 70 to 80% (Vecchio et al., 2013). Although this discrimination power appears to be reasonable in real clinical setting, these accuracies are not enough to be clinically useful alone. Based on spectroscopy and additional EEG variables, another study was able to discriminate AD from HC with 88% sensitivity and 89% specificity (Rodriguez et al., 1998). In fact, recent studies were able to demonstrate better classification performances (0.87 < AUC < 0.94) (Anghinah et al., 2011; Kanda et al., 2014; Buscema et al., 2015). However, the subjects sample, conditions of the EEG recording (e.g., subject awake, eyes open/closed, occurrence of external stimuli), the recording protocol, data processing and used methods induce complexity in the analysis and may compromise the conclusion

FIGURE 3 | Overlay of ROC performance curves from each model using the *Reduced set of feature* (*S, A, c,* and *w*). The shadow plot display the SEM values for each curve. Healthy controls (HC); probable Alzheimer's disease (AD); vascular dementia (VaD).

TABLE 1 | Performance values for each final classifier model.

	Complete set of features		Reduced set of features	
Classifier performances	**cv-ACC**	**AUC**	**cv-ACC**	**AUC**
Model 1 (HC vs. AD)	0.62	0.66	0.67	0.74
Model 2 (HC vs. VaD)	0.65	0.68	0.72	0.77
Model 3 (AD vs. VaD)	0.59	0.62	0.57	0.61
Model 4 (HC vs. AD&VaD)	0.70	0.75	0.77	0.83

Average of the CV accuracy (cv-ACC) and AUC performed for each final classifier using two different sets of features. The Complete Set of Features included a total of 132 predictors and the Reduced Set of Features included variable number of predictors based on automatic feature reduction. Values from [0.5–0.6] = Fail; [0.6–0.7] = Poor; [0.7–0.8] = Fair; [0.8–0.9] = Good; [0.9–1.0] = Excellent.

of such studies (Dauwels et al., 2010). We believe that the accessibility and low complexity of the tested features (*S, k, A, c, w, b*), provided satisfactory information for the discrimination process and may be further be validated as AD descriptors.

Model 2: The discrimination model for HC versus VaD, showed the same trend with slightly better results using the reduced set of features, obtaining an improvement of AUC from 0.68 to 0.77 and of cv-AUC from 0.65 to 0.72. Studies

reported that at specific brain regions, the correlation between the underlying structural changes and the EEG power is the marker that contributes most for the discrimination of VaD patients. (Szelies et al., 1994; Gawel et al., 2009). Nevertheless, it has repeatedly been demonstrated that VaD patients have increased delta power, increased diffused theta power, and decreased alpha rhythm (Signorino et al., 1995; Babiloni et al., 2004; Reed et al., 2004; Moretti et al., 2012). In fact, Moretti et al. (2012) suggested that theta/alpha ratio could be a reliable index for the estimation of the individual extent of CV damage. This model did not perform good or excellent. Hence the clinical utility of our features for the classification of VaD patients is yet unsatisfactory and should be complemented with other biomarkers such as neuroimaging.

Model 3: the discrimination of AD vs VaD, obtained the weakest classification results. In contradistinction to other models, when we used the Reduced Set of Features, this model performed poorly. Only poor performances values were obtained using both set of features with AUC = 0.62 and AUC = 0.61, respectively. The cv-ACC was also low with cv-ACC = 0.59 and 0.57 for each set of predictors, respectively. This is unexpected as one study reported performances of sensitivity 0.64% and specificity of 77% for distinguishing

VaD and AD (Walker et al., 2000). Additionally, it has been shown that 97% of VaD patients yielded a spectrum with dominant activity between 6.5 and 12 Hz and only 44% of AD when compared with HC (Signorino et al., 1995). When using the *Complete Set of Features*, our results showed poor performance results and that was due to a high correlation between features k and b ($r = 0.75$). This correlation was equally present within groups: $r = 0.80$ for HC, $r = 0.68$ for AD, and $r = 0.78$ for VaD. As mentioned above, this added redundancy and overfitted the LDA model, hence achieving low performance values. On the other hand, when we tested the *Reduced Set of Features*, the model appeared under-fitted. Our tested (and highly correlated) features k and b represent theta and beta bands which have been reported as very relevant for discrimination between AD and VaD (Babiloni et al., 2004). We speculate that in order to achieve better discrimination between these types of dementia our model lacked alternative features that characterize theta independently.

Model 4: The discrimination between HC and both groups of dementia patients (AD&VaD) obtained generally *good* discrimination performance for both conditions using the *Complete Set of Features* and the *Reduced Set of Features* with AUC = 0.75 and AUC = 0.83, respectively. The accuracy of the model was also good, presenting values of cv-ACC = 0.70 and cv-ACC = 0.77 for the *Complete* and *Reduced Set of Features*, respectively. AD and VaD patients may suffer occasionally from similar cognitive impairments that both result in neural degradation. The literature showed that these two diseases have many similarities at several electrophysiological markers such as increased power at lower frequencies, decreased and slower alpha or increase of theta power (Ernst Niedermeyer, 2010). Our previous work has demonstrated spectrum similarities between these groups of patients when compared with matched HC (Neto et al., 2015). The performance of model 4 revealed that the tested features characterized appropriately these spectrum differences. Therefore our results suggest that such features may be relevant for discrimination between dementia and healthy groups.

We obtained nearly the same performances measures for the variant models generated with permuted order of the learning and testing subsets and denoted no significant fluctuation on the models performances (cv-ACC and AUC) for the different permutations of subsets. The stable and low variations values obtained for cv-ACC and AUC across the permutations and the 10-fold CV are indicators of high robustness of the models here obtained.

CONCLUSION

The EEG features tested in this study for the discrimination between patients with dementia (*AD* or *VaD*) and HC performed generally well. The discrimination between *AD* and *VaD* showed above chance performance. Discrimination between *AD* from HC or *VaD* from HC was obtained with fair results.

Therefore, we conclude that the tested EEG features hold relevant discriminatory information, and, in combination with other markers and other known dementia diagnostic tools such as neuroimaging, may constitute necessary and valuable information to screen dementia in a clinical setting.

AUTHOR CONTRIBUTIONS

EN, study design, data analysis, wrote the article. FB, contributed with guidance and implementation code of discrimination analysis algorithms. HA, planning of this study and co-supervision; Selection of patient's datasets. HN, planning of this study and co-supervision. TE, planning of this study and supervision.

ACKNOWLEDGMENT

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnagi.2016.00273/full#supplementary-material

REFERENCES

Anghinah, R., Kanda, P. A., Lopes, H. F., Basile, L. F., Machado, S., Ribeiro, P., et al. (2011). Alzheimer's disease qEEG: spectral analysis versus coherence. Which is the best measurement? *Arq. Neuropsiquiatr.* 69, 871–874. doi: 10.1590/S0004-282X2011000700004

Babiloni, C., Binetti, G., Cassetta, E., Cerboneschi, D., Dal Forno, G., Del Percio, C., et al. (2004). Mapping distributed sources of cortical rhythms in mild Alzheimer's disease. A multicentric EEG study. *Neuroimage* 22, 57–67. doi: 10.1016/j.neuroimage.2003.09.028

Babiloni, C., De Pandis, M. F., Vecchio, F., Buffo, P., Sorpresi, F., Frisoni, G. B., et al. (2011a). Cortical sources of resting state electroencephalographic rhythms in Parkinson's disease related dementia and Alzheimer's disease. *Clin. Neurophysiol.* 122, 2355–2364. doi: 10.1016/j.clinph.2011.03.029

Babiloni, C., Infarinato, F., Marzano, N., Iacoboni, M., Dassu, F., Soricelli, A., et al. (2011b). Intra-hemispheric functional coupling of alpha rhythms is related to golfer's performance: a coherence EEG study. *Int. J. Psychophysiol.* 82, 260–268. doi: 10.1016/j.ijpsycho.2011.09.008

Besserve, M., Jerbi, K., Laurent, F., Baillet, S., Martinerie, J., and Garnero, L. (2007). Classification methods for ongoing EEG and MEG signals. *Biol. Res.* 40, 415–437. doi: 10.4067/S0716-97602007000500005

Besthorn, C., Zerfass, R., Geiger-Kabisch, C., Sattel, H., Daniel, S., Schreiter-Gasser, U., et al. (1997). Discrimination of Alzheimer's disease and normal aging by EEG data. *Electroencephalogr. Clin. Neurophysiol.* 103, 241–248. doi: 10.1016/S0013-4694(97)96562-7

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. New York, NY: Oxford University Press, Inc.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York, NY: Springer-Verlag.

Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., and Muller, K. R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Process. Mag.* 25, 41–56.

Bonanni, L., Thomas, A., Tiraboschi, P., Perfetti, B., Varanese, S., and Onofrj, M. (2008). EEG comparisons in early Alzheimer's disease, dementia with Lewy bodies and Parkinson's disease with dementia patients with a 2-year follow-up. *Brain* 131(Pt 3), 690–705. doi: 10.1093/brain/awm322

Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 26, 211–252.

Buscema, M., Vernieri, F., Massini, G., Scrascia, F., Breda, M., Rossini, P. M., et al. (2015). An improved I-FAST system for the diagnosis of Alzheimer's disease from unprocessed electroencephalograms by using robust invariant features. *Artif. Intell. Med.* 64, 59–74. doi: 10.1016/j.artmed.2015.03.003

Cichocki, A., Shishkin, S. L., Musha, T., Leonowicz, Z., Asada, T., and Kurachi, T. (2005). EEG filtering based on blind source separation (BSS) for early detection of Alzheimer's disease. *Clin. Neurophysiol.* 116, 729–737. doi: 10.1016/j.clinph.2004.09.017

Dauwels, J., Vialatte, F., and Cichocki, A. (2010). Diagnosis of Alzheimer's disease from EEG signals: where are we standing? *Curr. Alzheimer Res.* 7, 487–505. doi: 10.2174/156720510792231720

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Dierks, T., Frolich, L., Ihl, R., and Maurer, K. (1995). Correlation between cognitive brain function and electrical brain activity in dementia of Alzheimer type. *J. Neural Transm. Gen. Sect.* 99, 55–62. doi: 10.1007/BF01271469

Dierks, T., Perisic, I., Frolich, L., Ihl, R., and Maurer, K. (1991). Topography of the quantitative electroencephalogram in dementia of the Alzheimer type: relation to severity of dementia. *Psychiatry Res.* 40, 181–194. doi: 10.1016/0925-4927(91)90009-F

Dubois, M. F., and Hebert, R. (2001). The incidence of vascular dementia in Canada: a comparison with Europe and East Asia. *Neuroepidemiology* 20, 179–187. doi: 10.1159/000054785

Elgendi, M., Vialatte, F., Cichocki, A., Latchoumane, C., Jeong, J., and Dauwels, J. (2011). Optimization of EEG frequency bands for improved diagnosis of Alzheimer disease. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2011, 6087–6091. doi: 10.1109/IEMBS.2011.6091504

Ernst Niedermeyer, F. H. L. D. S. (2010). "Elecroencephalography," in *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 6th Edn, ed. F. H. L. D. S. Ernst Niedermeyer (New York, NY: Wolters Kluwer/Lippincott Williams & Wilkins Health), 375–387.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7, 179–188. doi: 10.1111/j.1469-1809.1936.tb02137.x

Fraga, F. J., Falk, T. H., Kanda, P. A., and Anghinah, R. (2013). Characterizing Alzheimer's disease severity via resting-awake EEG amplitude modulation analysis. *PLoS ONE* 8:e72240. doi: 10.1371/journal.pone.0072240

Friedman, J. H. (1989). Regularized discriminant-analysis. *J. Am. Stat. Assoc.* 84, 165–175. doi: 10.1080/01621459.1989.10478752

Gallego-Jutgla, E., Elgendi, M., Vialatte, F., Sole-Casals, J., Cichocki, A., Latchoumane, C., et al. (2012). Diagnosis of Alzheimer's disease from EEG by means of synchrony measures in optimized frequency bands. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2012, 4266–4270. doi: 10.1109/EMBC.2012.6346909

Gawel, M., Zalewska, E., Szmidt-Salkowska, E., and Kowalski, J. (2009). The value of quantitative EEG in differential diagnosis of Alzheimer's disease and subcortical vascular dementia. *J. Neurol. Sci.* 283, 127–133. doi: 10.1016/j.jns.2009.02.332

Gearing, M., Mirra, S. S., Hedreen, J. C., Sumi, S. M., Hansen, L. A., and Heyman, A. (1995). The Consortium to establish a registry for Alzheimer's disease (CERAD). Part X. Neuropathology confirmation of the clinical diagnosis of Alzheimer's disease. *Neurology* 45(3 Pt 1), 461–466.

Goulermas, J. Y., Findlow, A. H., Nester, C. J., Howard, D., and Bowker, P. (2005). Automated design of robust discriminant analysis classifier for foot pressure lesions using kinematic data. *IEEE Trans. Biomed. Eng.* 52, 1549–1562. doi: 10.1109/TBME.2005.851519

Guger, C., Ramoser, H., and Pfurtscheller, G. (2000). Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI). *IEEE Trans. Rehabil. Eng.* 8, 447–456. doi: 10.1109/86.895947

Gutman, B. A., Hua, X., Rajagopalan, P., Chou, Y. Y., Wang, Y. L., Yanovsky, I., et al. (2013). Maximizing power to track Alzheimer's disease and MCI progression by LDA-based weighting of longitudinal ventricular surface features. *Neuroimage* 70, 386–401. doi: 10.1016/j.neuroimage.2012.12.052

Haufe, S., Meinecke, F., Gorgen, K., Dahne, S., Haynes, J. D., Blankertz, B., et al. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* 87, 96–110. doi: 10.1016/j.neuroimage.2013.10.067

Hillyard, S. A., and Galambos, R. (1970). Eye movement artifact in CNV. *Electroencephalogr. Clin. Neurophysiol.* 28, 173–182. doi: 10.1016/0013-4694(70)90185-9

Huang, C., Wahlund, L., Dierks, T., Julin, P., Winblad, B., and Jelic, V. (2000). Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study. *Clin. Neurophysiol.* 111, 1961–1967. doi: 10.1016/S1388-2457(00)00454-5

Huang, G., Liu, G., Meng, J., Zhang, D., and Zhu, X. (2010). Model based generalization analysis of common spatial pattern in brain computer interfaces. *Cogn. Neurodyn.* 4, 217–223. doi: 10.1007/s11571-010-9117-x

Huang, Y., Erdogmus, D., Mathan, S., and Pavel, M. (2008). Detecting EEG evoked responses for target image search with mixed effect models. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2008, 4988–4991. doi: 10.1109/IEMBS.2008.4650334

Jelic, V., Blomberg, M., Dierks, T., Basun, H., Shigeta, M., Julin, P., et al. (1998). EEG slowing and cerebrospinal fluid tau levels in patients with cognitive decline. *Neuroreport* 9, 157–160. doi: 10.1097/00001756-199801050-00032

Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505. doi: 10.1016/j.clinph.2004.01.001

Kanda, P. A., Trambaiolli, L. R., Lorena, A. C., Fraga, F. J., Basile, L. F., Nitrini, R., et al. (2014). Clinician's road map to wavelet EEG as an Alzheimer's disease biomarker. *Clin. EEG Neurosci.* 45, 104–112. doi: 10.1177/1550059413486272

Kandimalla, R. J., Anand, R., Veeramanikandan, R., Wani, W. Y., Prabhakar, S., Grover, V. K., et al. (2014). CSF ubiquitin as a specific biomarker in Alzheimer's disease. *Curr. Alzheimer Res.* 11, 340–348. doi: 10.2174/1567205011666140331161027

Kandimalla, R. J., Prabhakar, S., Binukumar, B. K., Wani, W. Y., Gupta, N., Sharma, D. R., et al. (2011). Apo-Eepsilon4 allele in conjunction with Abeta42 and tau in CSF: biomarker for Alzheimer's disease. *Curr. Alzheimer Res.* 8, 187–196. doi: 10.2174/156720511795256071

Kandimalla, R. J., Prabhakar, S., Wani, W. Y., Kaushal, A., Gupta, N., Sharma, D. R., et al. (2013). CSF p-Tau levels in the prediction of Alzheimer's disease. *Biol. Open* 2, 1119–1124. doi: 10.1242/bio.20135447

Kaplan, A. Y., Fingelkurts, A. A., Fingelkurts, A. A., Borisov, S. V., and Darkhovsky, B. S. (2005). Nonstationary nature of the brain activity as revealed by EEG/MEG: methodological, practical and conceptual challenges. *Signal Process.* 85, 2190–2212. doi: 10.1016/j.sigpro.2005.07.010

Kwak, Y. T. (2006). Quantitative EEG findings in different stages of Alzheimer's disease. *J. Clin. Neurophysiol.* 23, 456–461. doi: 10.1097/01.wnp.0000223453.47663.63

Ledoit, O., and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.* 88, 365–411. doi: 10.1016/S0047-259x(03)00096-4

Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R. E., Wahlund, L. O., et al. (2007). Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *J. Neurosci. Methods* 161, 342–350. doi: 10.1016/j.jneumeth.2006.10.023

Lehmann, D., and Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48, 609–621. doi: 10.1016/0013-4694(80)90419-8

Lopez, M. M., Ramirez, J., Gorriz, J. M., Alvarez, I., Salas-Gonzalez, D., Segovia, F., et al. (2009). SVM-based CAD system for early detection of the Alzheimer's disease using kernel PCA and LDA. *Neurosci. Lett.* 464, 233–238. doi: 10.1016/j.neulet.2009.08.061

Lotte, F., and Guan, C. (2011). Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. *IEEE Trans. Biomed. Eng.* 58, 355–362. doi: 10.1109/TBME.2010.2082539

Lou, W., Xu, J., Sheng, H., and Zhao, S. (2011). Multichannel linear descriptors analysis for event-related EEG of vascular dementia patients during visual detection task. *Clin. Neurophysiol.* 122, 2151–2156. doi: 10.1016/j.clinph.2011.03.021

Maggi, L., Parini, S., Piccini, L., Panfili, G., and Andreoni, G. (2006). A four command BCI system based on the SSVEP protocol. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 1, 1264–1267. doi: 10.1109/IEMBS.2006.260353

Massoud, F., Devi, G., Stern, Y., Lawton, A., Goldman, J. E., Liu, Y., et al. (1999). A clinicopathological comparison of community-based and clinic-based cohorts of patients with dementia. *Arch. Neurol.* 56, 1368–1373. doi: 10.1001/archneur.56.11.1368

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). 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 Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Moretti, D. V., Zanetti, O., Binetti, G., and Frisoni, G. B. (2012). Quantitative EEG markers in mild cognitive impairment: degenerative versus vascular brain impairment. *Int. J. Alzheimers Dis.* 2012:917537. doi: 10.1155/2012/917537

Neto, E., Allen, E. A., Aurlien, H., Nordby, H., and Eichele, T. (2015). Electroencephalogram (EEG) spectral features discriminate between Alzheimer's (AD) and Vascular dementia (VaD). *Front. Neurol.* 6:25. doi: 10.3389/fneur.2015.00025

Nunez, P. L., and Srinivasan, R. (2006). *Electric fields of the brain : the neurophysics of EEG*, 2nd Edn. Oxford: Oxford University Press.

Prichep, L. S. (2005). Use of normative databases and statistical methods in demonstrating clinical utility of QEEG: importance and cautions. *Clin. EEG Neurosci.* 36, 82–87. doi: 10.1177/155005940503600207

Ramoser, H., Muller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* 8, 441–446. doi: 10.1109/86.895946

Reed, B. R., Eberling, J. L., Mungas, D., Weiner, M., Kramer, J. H., and Jagust, W. J. (2004). Effects of white matter lesions and lacunes on cortical function. *Arch. Neurol.* 61, 1545–1550. doi: 10.1001/archneur.61.10.1545

Reilly, J., Rodriguez, A. D., Lamy, M., and Neils-Strunjas, J. (2010). Cognition, language, and clinical pathological features of non-Alzheimer's dementias: an overview. *J. Commun. Disord.* 43, 438–452. doi: 10.1016/j.jcomdis.2010.04.011

Reuderink, B., and Poel, M. (2008). *Robustness of the Common Spatial Patterns algorithm in the BCI-pipeline*. Enschede: University of Twente.

Rodriguez, G., Nobili, F., Rocca, G., De Carli, F., Gianelli, M. V., and Rosadini, G. (1998). Quantitative electroencephalography and regional cerebral blood flow: discriminant analysis between Alzheimer's patients and healthy controls. *Dement. Geriatr. Cogn. Disord.* 9, 274–283. doi: 10.1159/000017072

Roman, G. C. (2002). Vascular dementia revisited: diagnosis, pathogenesis, treatment, and prevention. *Med. Clin. North Am.* 86, 477–499. doi: 10.1016/S0025-7125(02)00008-1

Rossini, P. M., Rossi, S., Babiloni, C., and Polich, J. (2007). Clinical neurophysiology of aging brain: from normal aging to neurodegeneration. *Prog. Neurobiol.* 83, 375–400. doi: 10.1016/j.pneurobio.2007.07.010

Shenoy, P., Krauledat, M., Blankertz, B., Rao, R. P., and Muller, K. R. (2006). Towards adaptive classification for BCI. *J. Neural. Eng.* 3, R13–R23. doi: 10.1088/1741-2560/3/1/R02

Shenoy, P., Miller, K. J., Ojemann, J. G., and Rao, R. P. (2008). Generalized features for electrocorticographic BCIs. *IEEE Trans. Biomed. Eng.* 55, 273–280. doi: 10.1109/TBME.2007.903528

Signorino, M., Pucci, E., Belardinelli, N., Nolfe, G., and Angeleri, F. (1995). EEG spectral analysis in vascular and Alzheimer dementia. *Electroencephalogr. Clin. Neurophysiol.* 94, 313–325. doi: 10.1016/0013-4694(94)00290-2

Signorino, M., Pucci, E., Brizioli, E., Cacchio, G., Nolfe, G., and Angeleri, F. (1996). EEG power spectrum typical of vascular dementia in a subgroup of Alzheimer patients. *Arch. Gerontol. Geriatr.* 23, 139–151. doi: 10.1016/0167-4943(96)00714-5

Szelies, B., Mielke, R., Herholz, K., and Heiss, W. D. (1994). Quantitative topographical EEG compared to FDG PET for classification of vascular and degenerative dementia. *Electroencephalogr. Clin. Neurophysiol.* 91, 131–139. doi: 10.1016/0013-4694(94)90034-5

Triba, M. N., Le Moyec, L., Amathieu, R., Goossens, C., Bouchemal, N., Nahon, P., et al. (2015). PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Mol. Biosyst.* 11, 13–19. doi: 10.1039/c4mb00414k

Tsuno, N., Shigeta, M., Hyoki, K., Faber, P. L., and Lehmann, D. (2004). Fluctuations of source locations of EEG activity during transition from alertness to sleep in Alzheimer's disease and vascular dementia. *Neuropsychobiology* 50, 267–272. doi: 10.1159/000079982

van Straaten, E. C., de Haan, W., de Waal, H., Scheltens, P., van der Flier, W. M., Barkhof, F., et al. (2012). Disturbed oscillatory brain dynamics in subcortical ischemic vascular dementia. *BMC Neurosci.* 13:85. doi: 10.1186/1471-2202-13-85

Vecchio, F., Babiloni, C., Lizio, R., Fallani Fde, V., Blinowska, K., Verrienti, G., et al. (2013). Resting state cortical EEG rhythms in Alzheimer's disease: toward EEG markers for clinical applications: a review. *Suppl. Clin. Neurophysiol.* 62, 223–236. doi: 10.1016/B978-0-7020-5307-8.00015-6

Velu, P. D., and de Sa, V. R. (2013). Single-trial classification of gait and point movement preparation from human EEG. *Front. Neurosci.* 7:84. doi: 10.3389/fnins.2013.00084

von Bunau, P., Meinecke, F. C., Kiraly, F. C., and Muller, K. R. (2009). Finding stationary subspaces in multivariate time series. *Phys. Rev. Lett.* 103, 214101. doi: 10.1103/PhysRevLett.103.214101

Walker, M. P., Ayre, G. A., Cummings, J. L., Wesnes, K., McKeith, I. G., O'Brien, J. T., et al. (2000). Quantifying fluctuation in dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. *Neurology* 54, 1616–1625. doi: 10.1212/WNL.54.8.1616

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika* 29, 350–362. doi: 10.2307/2332010

Worth, A. P., and Cronin, M. T. D. (2003). The use of discriminant analysis, logistic regression and classification tree analysis in the development of classification models for human health effects. *J. Mol. Struct. Theochem.* 622, 97–111. doi: 10.1016/S0166-1280(02)00622-X

Ye, J., Xiong, T., Li, Q., Janardan, R., Bi, J., Cherkassky, V., et al. (2006). "Efficient model selection for regularized linear discriminant analysis," in *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*, (New York, NY: ACM Press).

Zhdanov, A., Hendler, T., Ungerleider, L., and Intrator, N. (2007). Inferring functional brain states using temporal evolution of regularized classifiers. *Comput. Intell. Neurosci.* 2007:52609. doi: 10.1155/2007/52609

Zhou, X.-H., McClish, D. K., and Obuchowski, N. A. (2011). *Statistical Methods in Diagnostic Medicine*, 2nd Edn. Hoboken, NJ: Wiley.