Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 11 (2019) 588-598

Diagnostic Assessment & Prognosis

Robust automated computational approach for classifying frontotemporal neurodegeneration: Multimodal/multicenter neuroimaging

Patricio Andres Donnelly-Kehoe[a,b,1], Guido Orlando Pascariello[a,b,1], Adolfo M. García[c,d,e], John R. Hodges[f,g], Bruce Miller[h], Howie Rosen[i], Facundo Manes[c,d,f], Ramon Landin-Romero[f,j], Diana Matallana[k], Cecilia Serrano[l,m], Eduar Herrera[n], Pablo Reyes[o,p], Hernando Santamaria-Garcia[o,p], Fiona Kumfor[f,j], Olivier Piguet[f,j], Agustin Ibanez[c,d,f,q,r,*], Lucas Sedeño[c,d,**]

[a]*Multimedia Signal Processing Group - Neuroimage Division, French-Argentine International Center for Information and Systems Sciences (CIFASIS) - National Scientific and Technical Research Council (CONICET), Rosario, Argentina*
[b]*Laboratory of Neuroimaging and Neuroscience (LANEN), INECO Foundation Rosario, Rosario, Argentina*
[c]*Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCyT), INECO Foundation, Favaloro University, Buenos Aires, Argentina*
[d]*National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina*
[e]*Faculty of Education, National University of Cuyo (UNCuyo), Mendoza, Argentina*
[f]*Centre of Excellence in Cognition and its Disorders, Australian Research Council (ARC), Sydney, Australia*
[g]*The University of Sydney, Brain and Mind Centre and Clinical Medical School, Sydney, Australia*
[h]*Memory and Aging Center, University of California, San Francisco, San Francisco, CA, USA*
[i]*Department of Neurology, Memory Aging Center, University of California, San Francisco, CA, USA*
[j]*The University of Sydney, Brain and Mind Centre and School of Psychology, Sydney, Australia*
[k]*Medical School, Aging Institute, Psychiatry and Mental Health, Pontificia Universidad Javeriana (PUJ), Bogotá, Colombia*
[l]*Memory and Balance Clinic, Buenos Aires, Argentina*
[m]*Department of Neurology, Dr Cesar Milstein Hospital, Buenos Aires, Argentina*
[n]*Departamento de Estudios Psicológicos, Universidad Icesi, Cali, Colombia (eduar)*
[o]*Radiology, Hospital Universitario San Ignacio (HUSI), Bogotá, Colombia*
[p]*Pontificia Universidad Javeriana, Departments of Physiology and Psychiatry – Centro de Memoria y Cognición Intellectus, Hospital Universitario San Ignacio, Bogotá, Colombia*
[q]*Universidad Autónoma del Caribe, Barranquilla, Colombia*
[r]*Center for Social and Cognitive Neuroscience (CSCN), School of Psychology, Universidad Adolfo Ibáñez, Santiago, Chile*

Abstract

Introduction: Timely diagnosis of behavioral variant frontotemporal dementia (bvFTD) remains challenging because it depends on clinical expertise and potentially ambiguous diagnostic guidelines. Recent recommendations highlight the role of multimodal neuroimaging and machine learning methods as complementary tools to address this problem.

Methods: We developed an automatic, cross-center, multimodal computational approach for robust classification of patients with bvFTD and healthy controls. We analyzed structural magnetic resonance imaging and resting-state functional connectivity from 44 patients with bvFTD and 60 healthy controls (across three imaging centers with different acquisition protocols) using a fully automated processing pipeline, including site normalization, native space feature extraction, and a random forest classifier.

Results: Our method successfully combined multimodal imaging information with high accuracy (91%), sensitivity (83.7%), and specificity (96.6%).

[1]Both authors contributed equally to this work.

*Corresponding author. Tel./fax: +54 011 4812-0010.
**Corresponding author. Tel./fax: +54 011 4812-0010.
E-mail address: aibanez@ineco.org.ar (A.I.), lucas.sedeno@gmail.com (L.S.)

https://doi.org/10.1016/j.dadm.2019.06.002

Discussion: This multimodal approach enhanced the system's performance and provided a clinically informative method for neuroimaging analysis. This underscores the relevance of combining multimodal imaging and machine learning as a gold standard for dementia diagnosis.

Keywords: Dementia; bvFTD; Data-driven computational approaches; Classifiers; Neuroimaging

1. Background

Clinical neuroscience has failed to fully profit from the richness of neuroimaging data, and computational approaches represent the next frontier for developing more robust methods [1,2], in particular for neurodegenerative diseases [3]. Early diagnosis of these conditions remains challenging because it depends on clinical expertise to follow and interpret diagnostic guidelines [4], in a context of vast sociogeographical variability [5]. This is especially true for behavioral variant frontotemporal dementia (bvFTD) —the second most common dementia before age 65— which presents a relatively young onset, clinical overlap with other diseases, and variability in brain atrophy patterns [6]. Automatized multimodal neuroimaging analysis (mainly based on volumetric magnetic resonance imaging [MRI] and positron emission tomography measures), combined with computational decision-support methods, have been proposed as a potentially useful approach for early diagnosis [1,2]. However, few studies have evaluated the robustness of resting-state functional connectivity (FC) for discriminating patients with bvFTD from healthy controls (HCs) with automated methods, even though neurodegeneration is a complex process impacting brain structure, function, and network connectivity [7,8]. Moreover, previous studies have shown that disease-related network alterations precede brain atrophy [8,9]. Yet, the potential benefit of combining FC and atrophy measures in computational learning approaches remains untested in bvFTD. Data-driven approaches with feature-selection techniques are mostly neglected in the field, although they may successfully classify complex clinical conditions [2] because of their generalization power and reduction of overfitting. Here, we introduce a novel approach combining multimodal information associated with neurodegeneration (atrophy and FC), performing joint inferences over different centers (multicenter normalization process), and using robust machine learning methods to classify patients from HCs.

Different noninvasive neuroimaging modalities have been used to discriminate between patients with bvFTD and HCs. However, evidence for improved classification has been mixed when analyses rely on combined neuroimaging techniques (e.g., MRI and diffusion images) [10]. Moreover, despite yielding good accuracy rates, most studies presented several limitations including a relatively small sample size [11], the inclusion of heterogeneous FTD variants with different anatomo-clinical profiles [10–13], the absence of HCs [14], the lack of cross-center comparisons [10–13], the use of nonautomatic procedures [15], and the absence of data-driven computational approaches. More importantly, none of these reports have included FC data despite its proposed role as a biomarker for bvFTD diagnosis [7–9]. The only study using classification methods on FC data obtained high discrimination between HCs and patients with bvFTD [15]. Yet, its sample size was small (12 participants per group), it included unimodal information only, and resting-state images were acquired from two different scans without normalization procedures.

In sum, robust data-driven computational approaches combining atrophy and FC measures to differentiate between patients with bvFTD and HCs remain promising but unexplored. To bridge this gap, our main objective is to use a fully automated computational framework (Fig. 1A) to extract the most relevant multimodal MRI-based features (atrophy and FC) to identify patients with bvFTD across three countries. We expect that our multimodal approach will outperform classification results considering only one neuroimaging modality (atrophy or FC). Contrary to cognitive screening methods or other model-driven strategies, our approach is blind to any "a priori human" assumption. This study presents a feasible complementary tool for clinical diagnosis [1,2] representing a crucial initial step to consolidate multimodal and data-driven computational approaches as a gold standard method for dementia.

2. Materials and methods

2.1. Participants

The study comprised 104 participants from a current multicenter protocol [16–18]. Sixty were HCs with no history of psychiatric or neurological disease, and 44 were patients fulfilling revised consensus criteria for probable bvFTD [19] from three international clinics: INECO Foundation, Argentina (country-1); San Ignacio University Hospital-Javeriana University, Colombia (country-2); and the Frontotemporal Dementia Research Group (FRONTIER), Sydney, Australia (country-3). Both groups were matched on sex, age, and education (Table 1). The patients' clinical diagnosis was established in each center through a standard examination including extensive neurological, neuropsychiatric, and neuropsychological assessments. Each case was revised by a multidisciplinary

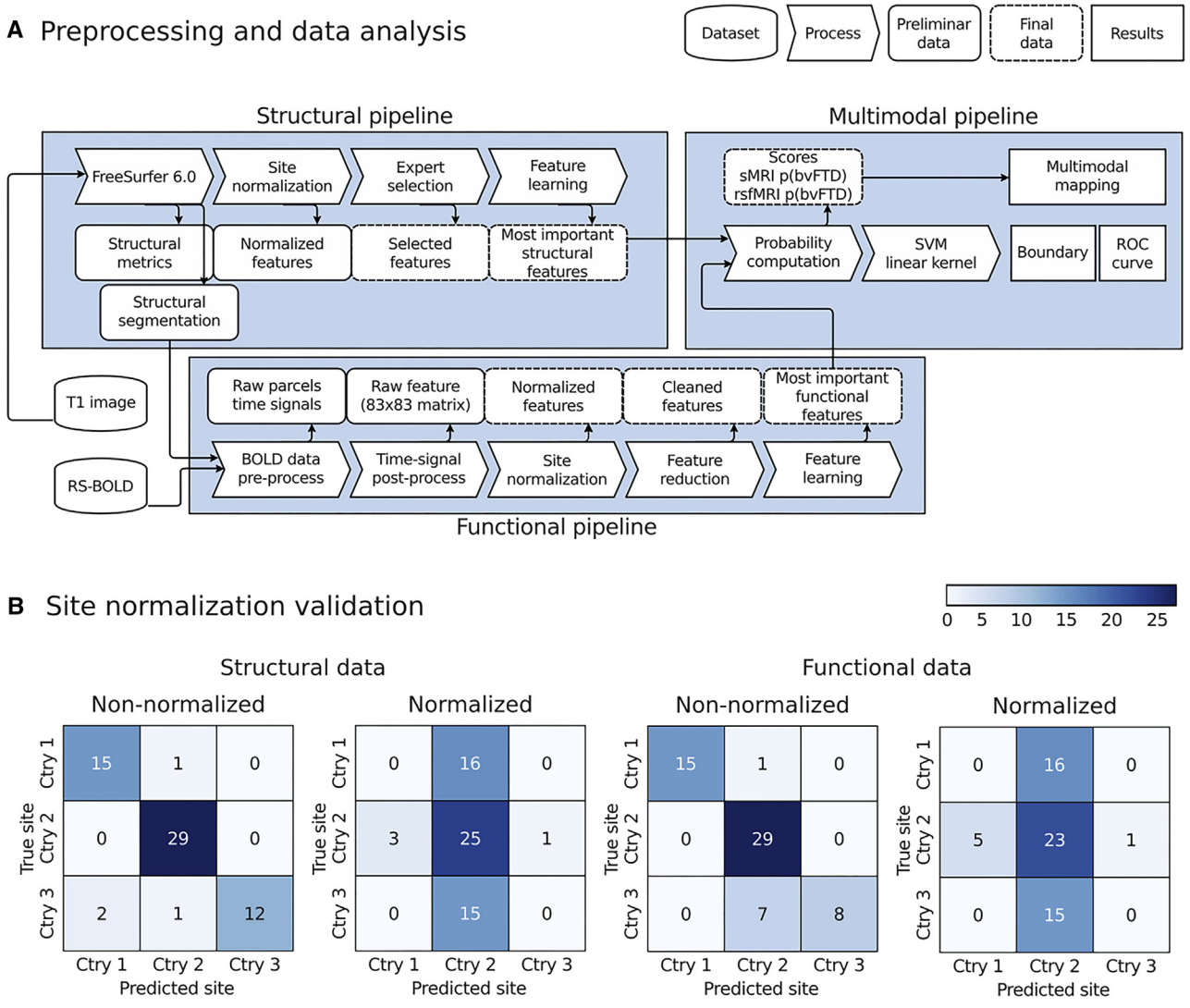

Fig. 1. Preprocessing, data analysis, and site normalization. (A) Main overview of the processing framework used for the multimodal integration of structural and functional information using machine learning techniques. p(bvFTD) refers to the probability of being classified as a behavioral variant frontotemporal dementia (bvFTD) patient, SVM refers to support vector machine classifier, ROC to receiver operating characteristic, RS to resting state, BOLD to blood oxygen level dependent, sMRI to structural MRI, and rsfMRI to resting-state functional MRI. (B) Confusion matrix before and after the site normalization process for both structural data (left) and functional data (right).

clinical meeting of bvFTD experts, as previously reported [16,18,20]. All patients were in early/mild disease stages. They did not fulfill criteria for specific psychiatric disorders, and they did not exhibit primary language deficits. Functional impairment and prominent changes in personality and social behavior were confirmed by caregivers for all cases. MRI revealed frontal atrophy in each patient, and those with single-photon emission computed tomography data exhibited frontal hypoperfusion (for further clinical and demographic details, see [18]).

Table 1
Demographic information

	Controls	bvFTD	Statistics	*P* values
Sex*	F = 33 M = 27	F = 25 M = 19	0.06	.79
Age†	63.91 (7.63)	66.72 (8.33)	3.10	.08
Education†	14.83 (4.27)	13.75 (4.06)	1.62	.20

Abbreviation: bvFTD, behavioral variant of frontotemporal dementia.
*Chi-square test.
†ANOVA test. Mean (standard deviation).

The institutional ethics committee of each center approved the study protocol. All participants (or their person responsible) provided signed informed consent in accordance with the Declaration of Helsinki.

2.2. Image acquisition

We followed a multicenter strategy [18] including structural and resting-state sequences from different MRI machine manufacturers, permanent magnetic fields, and varying acquisition parameters. This allowed us to test the reliability of our novel methods against the variability of

MRI devices and protocols. MRI acquisition and preprocessing steps are reported following guidelines from the Organization for Human Brain Mapping [21] (for acquisition details, see Supplementary Material 1-2, and Table 2).

2.3. Structural imaging preprocessing and analysis

2.3.1. Preprocessing for classification

All T1 brain volumes were processed to obtain a complete morphometric description, based on a surface-based morphometry analysis [22] using FreeSurfer (v 6.0) image analysis suite (see preprocessing details in Supplementary Material 3).

2.3.2. Site normalization for structural images

We performed a site normalization to avoid MRI-setup–dependent bias in the measurements. For this, we replicated a previous procedure [23], replacing *w*-scores by *z*-scores because of the narrow age range in our sample. To estimate the *z*-scores, we first removed nonuseful and nonphysiological features (including non–white matter hypointensities, white matter hypointensities, and several curvatures) presenting a nonnormal distribution after Shapiro-Wilk tests –no volumes or thickness measures were removed during this step. Then, for each center, each feature (volume and thickness) of both HCs and patients was *z*-scored based on the mean and standard deviation of the corresponding center's HCs.

To corroborate the effectiveness of procedure, we applied a Random Forest Classifier (RFC) with leave-one-out cross-validation (LOOCV) to identify each HC's center of origin. As in [23], results (Fig. 1B) showed high discrimination accuracy before site normalization and full confusion afterward (i.e., chance accuracy).

2.4. Functional imaging preprocessing

2.4.1. Image preprocessing

Data preprocessing was performed using a pipeline written in Python 2.7, following previous studies [24–26] (see preprocessing details in Supplementary Material 4).

2.4.2. Matrix construction

We performed a registration from the high-resolution anatomical space (T1 3D) to the low-resolution mean functional MRI (fMRI) image space using FSL FLIRT [24] with 6 degrees of freedom and mutual information score as cost function. Then, we used the parcels' masks from the native space obtained via Desikan-Killiany parcellation [27] in the structural processing to obtain mean time signals from 83 brain cortical and subcortical regions (by using FSL's fslmeants). Each resting-state fMRI recording was represented by 83 nodes with different timepoints depending on the acquisition center (Table 2).

In addition, using a sixth-order Butterworth band-pass filter [28], we estimated narrow-band time signals, filtered in a frequency band that has been mapped to gray matter activity (0.04-0.07 Hz) [26]. Fisher's correlation coefficient was used to construct an 83-node functional connectivity network for each subject. We integrated a quality assurance procedure in the pipeline to rapidly compare and explore time signals and correlation matrices before and after the narrow-band filtering to avoid unexpected artifacts (such as over peaks in the extremes of the signals, passband ripples, and unexpected phase effects). Features for the functional analysis (n = 3403) were extracted from the upper-right triangle of the connectivity matrix.

Table 2
fMRI acquisition parameters and scanning protocol in each center

	Country-1	Country-2	Country-3
A. Acquisition parameters			
Firm	Philips Intera	Philips Achieva	Philips Achieva
Tesla	1.5 T	3 T	3 T
Number of slices	33	40	29
Voxel size	3.6 x 3.6 × 4 mm	3 x 3 × 3 mm	1.88 x 1.88 × 4.5 mm
Flip angle	90		
Acquisition	Ascending. Parallel to the anterior and posterior commissures.		
Repetition time	2777 ms	3000 ms	2000 ms
Echo time	50 ms	30 ms	30 ms
Duration	10 min	5 min	7 min
Instruction	"Do not think about anything in particular"		
Number of volumes	209	120	208

	Controls	bvFTD*	*Z* values†	*P* values
B. Movement parameters				
Mean translational	0.07 (0.05)	0.08 (0.06)	−1.34	0.17
Mean rotational (°)	0.04 (0.02)	0.06 (0.05)	−1.82	0.07

*Behavioral variant frontotemporal dementia (bvFTD).
†Mann-Whitney U statistic.

2.4.3. Site normalization for FC matrices

We performed the same site normalization procedure and analysis used on anatomical data. Higher classification of the HCs' center of origin was obtained with nonnormalized data than with normalized data (Fig. 1B). No features were removed given that all of them exhibited a normal distribution at each center.

2.4.4. Feature reduction of FC for machine learning

Owing to the large number of features and the need to circumvent overfitting problems [29], we implemented a feature-reduction strategy that was integrated into all subsequent machine learning analyses with fMRI-based connections. At each iteration of the LOOCV scheme [30] (described at section 2.5.1), we followed a previously reported strategy [31]. However, to decrease computational costs, we replaced the linear regression [31] by the computation of the Pearson's correlation coefficient [32] between each feature and the binary class vector (the group of tags we wanted to predict). Then, we generated a histogram with all correlation coefficients, removed the two thirds with the lowest absolute values, and retained the most informative features to be processed in subsequent stages. This threshold was fixed ad-hoc to retain a number of features ($\approx$1000) that could be handled by the proposed classifier according to the training sample size ($\approx$100).

2.5. Machine learning methods

Fig. 1A shows the overall analysis pipeline used to integrate machine learning techniques with a multimodal and native space-based strategy.

2.5.1. Single MRI modality approach

2.5.1.1. Feature-selection procedure

Preprocessed features of each neuroimaging modality were analyzed via a progressive feature elimination procedure [33] with a LOOCV scheme. We optimized the accuracy of an RFC by varying the number of features from all to a single one according to its classificatory relevance. RFC quantifies the importance of a feature depending on how much the average Gini impurity index decreases in the forest because of its use as node in a tree. We used this score to progressively eliminate features by removing the feature with the lowest importance at each iteration. Finally, we kept the N first features in the ranking, where N is the optimal number of features such that using more than N features fail to improve the classifier's performance.

These analyses were performed with the RFC implemented in the Python's scikit-learn package, with a fixed number of trees (2000) and the recommended number of features (P) in each split, where P is the square root of the number of features.

The optimal number of features for each modality was selected visually by using the minimal quantity at which accuracy became constant. We used this fixed number of features to obtain the accuracy, the confusion matrix, and the receiver operating characteristic (ROC) curve, and to obtain each subject's intramodality probability of being in the disease group.

2.5.2. Multimodal analysis

We used each subject's intramodality probabilities (structural bvFTD probability and functional bvFTD probability) to generate a two-dimensional multimodal space. Then, we used a linear support vector machine (SVM) classifier [34] to obtain the multimodal decision boundaries for the classification among populations.

We used a LOOCV scheme to characterize the confusion matrix, ROC curves, and each subject's probability of being classified as bvFTD (p[bvFTD]) using this multimodal integration strategy. To select the optimal point of multimodal classification, we used the information provided by ROC curves and calculated the optimal working point, maximizing Youden's J statistic. In addition, we used 30 iterations on a bootstrap scheme with a proportion of 80%/20% (train/test) to study space regions where the decision became unstable.

2.5.2.1. Generalization across sites

We implemented a cross-center validation approach to test the generalization power of our multimodal classifier. One part of the data was used for training the predictive model, and the remaining data were kept only for validation. Two centers (country-1+country-2) were used for the predictive model because they include the largest number of participants (45 HCs and 33 bvFTD), whereas country-3 was used for testing (15 HCs and 11 bvFTD).

3. Results

3.1. Machine learning analysis using morphometric features

Structural MRI data yielded an optimal number of 29 features, with an accuracy of 90.3% and an area under the ROC curve (AUC) of 0.78 (Fig. 2B). The confusion matrix showed an almost balanced profile between false positives and false negatives, with a sensitivity of 86% and a specificity of 93.3%.

Fig. 2C shows the most relevant features from the LOOCV analysis. As the importance in RFC is a decimal value with a range of several orders of magnitude, we used decibels to compress the full scale of importance in a logarithmic way. In this, we used the last element of the optimal features group —in this case, feature 29— as reference. The most important features were mainly temporal (amygdala; hippocampus; inferior, middle, and superior temporal gyri; temporal pole) and frontal (middle and superior frontal gyri, anterior cingulate cortex) regions [6,35].

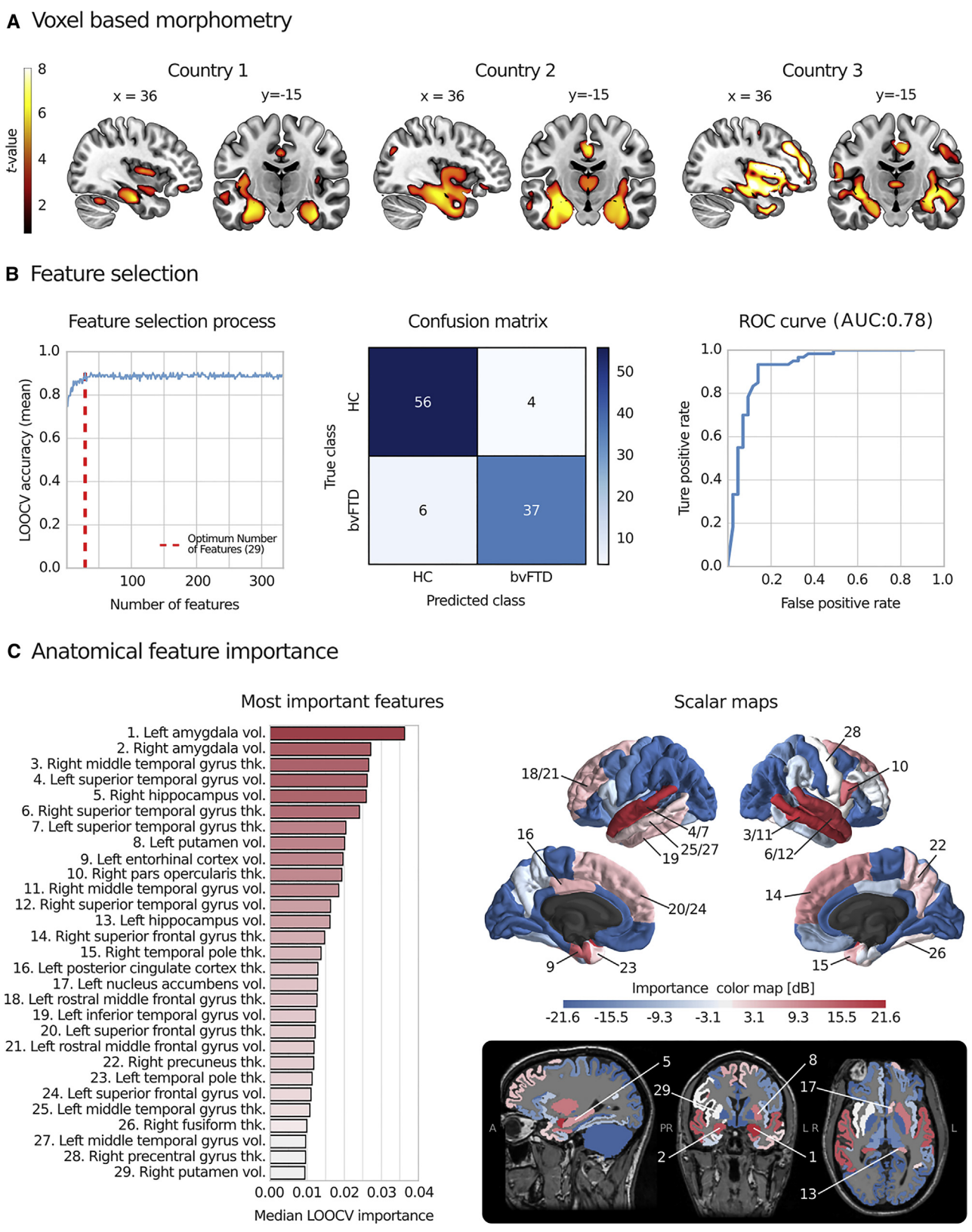

Fig. 2. Atrophy pattern and classification analysis based on morphometric features. (A) Damage extension in patients with behavioral variant of frontotemporal dementia (bvFTD) compared with the healthy controls (HCs) for each center via voxel-based morphometry (VBM) analysis (see Supplementary Material 1 for details of this analysis). (B) Overview of the selection of the optimal number of features, the confusion matrix, and the main descriptors of the classifier. LOOCV refers to leave-one-out cross validation, ROC to receiver operational characteristic curve, and AUC to area under the ROC curve. (C) On the left side, we list the main features in order of importance for the classification analysis; on the right side, we show the anatomical distribution of the features with a color code ranking their relevance: red-to-white features are those with the highest contribution (red) for the classification rate, while white-to-blue ones were dismissed after the selection of the optimal number of features. Numbers point to brain structures listed on the left.

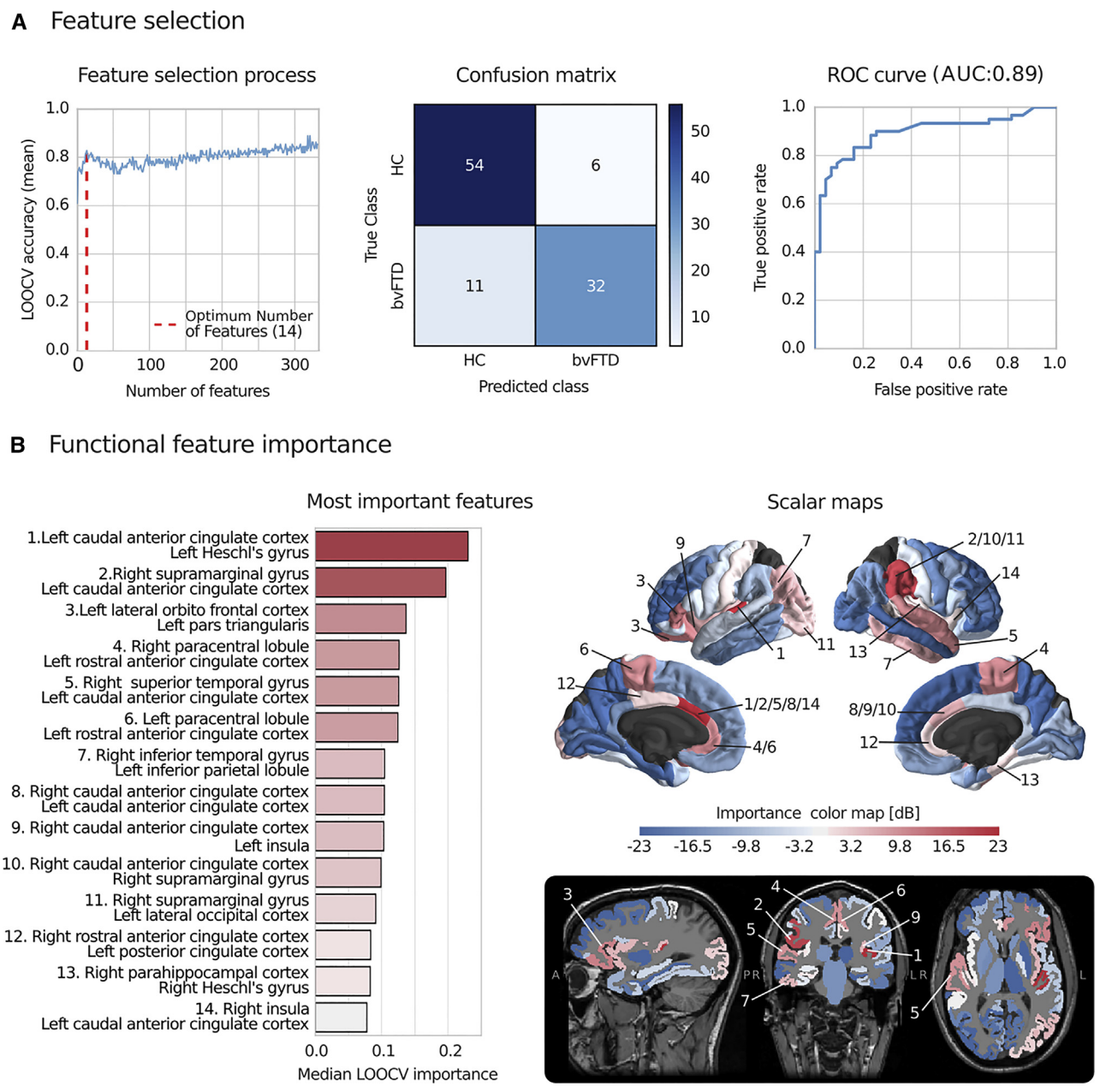

Fig. 3. Classification analysis based on functional connectivity features. (A) Overview of the selection of the optimal number of features, the confusion matrix, and the main descriptors of the classifier. LOOCV refers to leave-one-out cross-validation, ROC to receiver operational characteristic curve and AUC to area under the ROC curve. (B) On the left side, we list the main features in order of importance for the classification analysis; on the right side, we show the anatomical distribution of the features with a color code to ranking their relevance: red-to-white features are those with the highest contribution (red) for the classification rate, while white-to-blue ones were dismissed after the selection of the optimal number of features. Numbers point to brain structures listed on the left, and the same number points to both linked brain structures.

3.2. Machine learning analysis using functional connectivity

We found an optimal number of 14 features with an accuracy of 83.5% and an AUC of 0.89 (Fig. 3A). The confusion matrix showed an unbalanced profile between false positives and false negatives, with a tendency to yield more false negatives than false positives. The classifier performed worse than the previous one, with a sensitivity of 74.4% and a specificity of 90%.

The most important features (Fig. 3B) mainly include connections between frontofrontal (connections 3, 8, 9, and 14), frontotemporal (connections 1 and 5), and frontoparietal (connections 2, 4, 6, and 10) regions.

3.3. Multimodal integration

Multimodal analysis provided a robust classification (Fig. 4). We found an optimal probability value of 0.65, yielding an accuracy of 91% and an AUC of 0.95

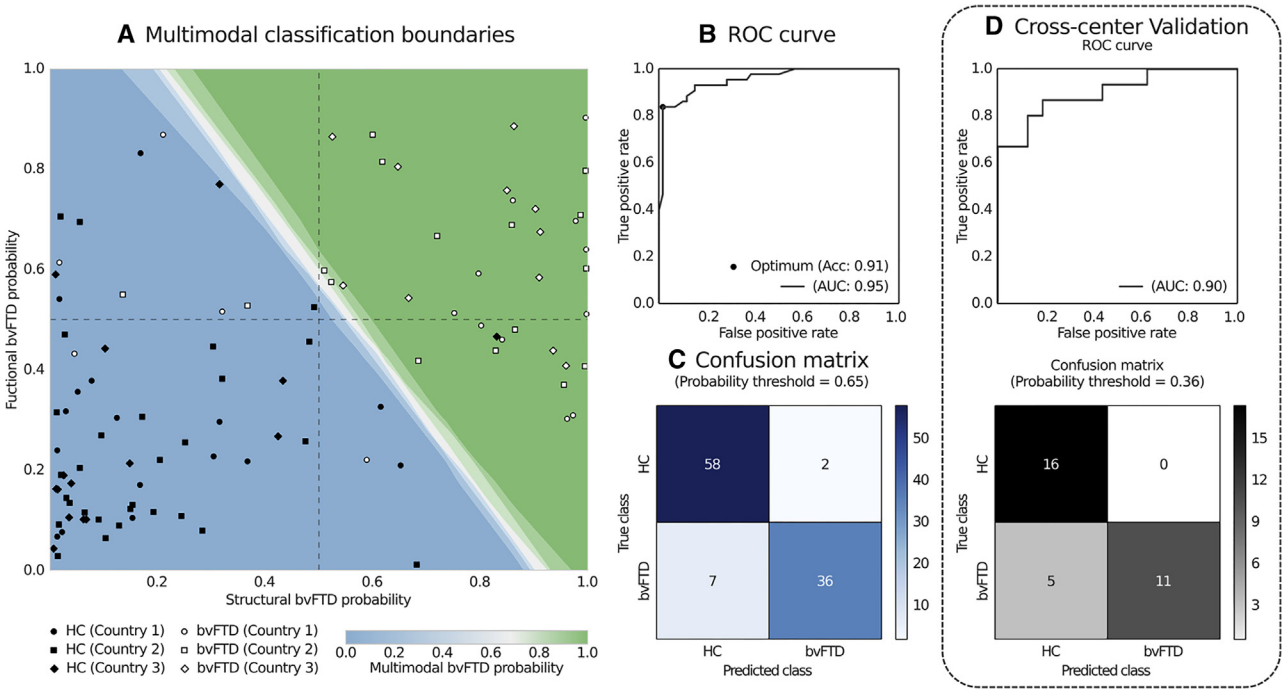

Fig. 4. Classification analysis based on integrated multimodal results. (A) Leave-one-out cross-validation: On the left side is the classification boundary between behavioral variant of frontotemporal dementia (bvFTD) and healthy control (HC) in the multimodal plane, with the structural bvFTD probability in the x-axis and the functional bvFTD probability in the y-axis. Points show the location for each participant in the multimodal plane while the colors represent the multimodal bvFTD probability based on a support vector machine (SVM) classifier over the 30 bootstrapping iterations. (B) The upper part of the right side shows the receiver operating characteristic (ROC) curve where the black dot indicates the optimal working point according to Youden's Index. The area under the ROC curve (AUC) and the accuracy for the optimal working point are written in the legend. (C) In the lower part of the right side is the confusion matrix for the optimal probability threshold according to Youden's Index. (D) Cross-center validation: on the left side is the ROC curve with the AUC written in the legend. On the right side is the confusion matrix obtained using an SVM.

(Fig. 4B). We also found clear boundaries between HCs and patients with bvFTD (Fig. 4A). As shown in Fig. 4C, the multimodal integration reduced false positives to two and yielded seven false negatives, with a sensitivity of 83.7% and a specificity of 96.6%.

In the cross-center validation approach, we obtained an accuracy of 92.3%, an AUC of 0.97, with a sensitivity of 100% and a specificity of 86.6%.

4. Discussion

This work introduced and validated a robust computational approach for automatic classification of patients with bvFTD despite multiple sources of variability across different clinical centers. A computational data-driven method combining structural connectivity and FC provided the best classification, supporting the multimodal nature of neurodegeneration [8,36]. These results highlight the potential of automatic feature extraction from MRI combined with decision-support methods to assess specific multidimensional patterns of disease.

To our knowledge, this is the best structural MRI-based classification (90%) to date (Fig. 2C), showing that data-driven features resemble the early patterns of neurodegeneration in bvFTD [6,35]. The main features corresponded to the temporal and frontal cortices. Although the amygdala, the hippocampus, and the temporal cortex showed the highest levels of discrimination, the highest accuracy levels were accomplished when frontal regions (left rostral middle frontal gyrus, right pars opercularis, superior frontal gyrus) were added to the classification because of the multivariate nature of the RFC. In previous multicenter studies, the best accuracy using only structural MRI reached 85% [35]. Main improvements compared with previous reports include (1) the use of native space to measure structural features [37] to remove errors/artifacts due to misregistration to an standard space [22,38], (2) the inclusion and validation of a cross-center normalization to allow the unbiased mixing of data from different protocols and centers, (3) a data-driven feature-selection process to find the optimum dimensionality for each feature space and enhance generalization; and (4) a classification performed by a nonlinear model (RFC). These characteristics make the proposed method especially robust for different MRI machines and protocols, enhancing its applicability in large multicenter studies.

FC results alone provided a sufficiently robust method to classify patients with bvFTD across multiple sources of variability (different MRI scanners, different recording centers, and inclusion subject's native space). Frontotemporal

networks, the hallmark of impaired connectivity in bvFTD [15,18], constituted the main features (Fig. 3B) elicited by our data-driven approach. The underperformance of this classifier's sensitivity compared with the structural MRI-based classifier may be explained by both the between-subject variability in the intrinsic FC profile [39] and the heterogeneous functional changes related to neurodegenerative diseases, which might vary over time depending on the deterioration stage at molecular and systemic levels [8]. Although a previous study has provided a higher classification rate in a multicenter setting [15], these results were limited by a very small sample size (12 HCs and 12 bvFTD patients), the nonnormalization of 2 MRI scan recordings, and the use of user-dependent methods (extraction of specific resting-state networks via independent component analysis). Here, we obtained an accuracy of 83.5% by using larger sample sizes, multicenter normalization of the signals, and a fully data-driven approach. Thus, our approach is more suitable for developing a nonhuman dependent and robust biomarker of neurodegeneration [1,2,18]. In addition, our methodology allowed for automatization of processing and also accounted for differences across centers. Additionally, we obtained the mean blood oxygen level dependent signal from region of interests in each subject's native space, avoiding artifacts produced by registration to a standard space [38]. Our fMRI processing method thus enhances the single-subject usability of brain network features across clinical centers.

By combining structural and functional probabilities to generate multimodal space, we obtained improved classification performance whose generalization power was validated by a cross-center approach. Although the multimodal improvement in accuracy (91%) was small, relative to structural data alone, the multimodal space allowed a more thoughtful analysis from a clinical viewpoint. For example, Fig. 4A shows one false positive in the green area with high structural bvFTD probability ($\approx$83%) and low functional bvFTD probability ($\approx$47%); this HC is 80 year old and, as most subjects are younger, the structural classifier's performance is affected by the changes related to normal aging [40], nevertheless the functional probability suggest a healthy pattern for this subject. In addition, Fig. 4A shows five false negatives in which the functional bvFTD probability is above 0.5 and that are misclassified because of a low structural bvFTD probability, suggesting that the functional metrics could add information not contained in the atrophy pattern [8,9]. These cases reflect the richness of a multimodal analysis, which could inform clinicians about different states of health and disease in a way that can be integrated with the patient's clinical history for diagnostic purposes. Neurodegeneration involves multiple dimensions of tissue changes. Local atrophy and apoptosis coexist or can be anticipated by synaptic connectivity abnormalities (e.g., autophagy, mitochondria, insulin/IGF-1, signaling, microRNAs) [41]. FC changes can be present before structural changes [8,9], which in turn seem closely related with changes in FC [41]. Neurodegeneration in bvFTD is generated by pathophysiological events [42] (e.g., axonal degeneration, synapse loss, dendritic retraction) and propagation of misfolded proteins [42,43], which impact on brain structure and connectivity. In turn, these changes induce neurotoxicity, which may activate compensatory mechanisms and/or disrupt other regions' structure and balance of damaged networks [7,14]. Thus, a data-driven computational approach combining both structural and connectivity features affords a more biologically plausible model than unimodal approaches [36].

Data-driven methods allowed us to introduce brain information from MRI (e.g., both structural and FC information) without considering any specific anatomic pattern of alterations, which resulted in a classification not biased by any a priori data selection. However, atrophy of frontotemporal regions and changes in frontotemporal connectivity emerged as a critical feature in our classification results. These findings indicate that (1) frontotemporal patterns of atrophy and connectivity alterations can provide probability metrics for the clinical assessment of bvFTD and (2) these metrics are robust enough to overcome the potential noise and variability introduced by the differences in recording site, in the diagnosis criteria made by different centers and in sociodemographic profiles.

The diagnosis of bvFTD is a complex process requiring consensus groups and multiple medical protocols/studies (visual inspection of MRI, patient's neuropsychological evaluation, relative's perspective). Therefore, the development of high sensitivity/specificity biomarkers has a major impact on clinical practice. Currently, clinical neuroimaging interpretation is reliant on visual inspection by an expert, whereas several studies (including this one) have shown that automatic analysis and use of functional networks can improve the utility of neuroimaging data by providing reliable and objective multimodal scores [1,2]. In this context, automated, nonhuman dependent data-driven method can help seek consensus independent from a certain criterion or a specific theoretical framework. This work shows that multimodal MRI analysis can be used in a fully automated way to assess for bvFTD with a similar accuracy to that reported in multimodal positron emission tomography studies [44], but with a simpler, noninvasive procedure.

4.1. Limitations and future directions

The results presented here have limitations to be covered in future works. First, sample patients were identified based on clinical diagnosis, as biological biomarkers/pathological/genetic confirmation was not available. Nevertheless, the research centers in this study are specialized in the diagnosis, assessment, and treatment of dementia, and followed validated protocols and diagnostic guidelines [19]. In fact, all previous studies in which neuroimaging and machine

learning tools were combined to study patients with bvFTD present the same limitation [10–15]. However, future research should include biological biomarkers of bvFTD to further evaluate our neuroimaging classification method. Second, although the bvFTD sample size is larger than previous related works [10,11,15,17,45,46] and that we used a cross-center validation approach (Fig. 4D), future studies should include even larger groups. In addition, other neurodegenerative conditions should be included in further analyses to test the power of the classifier regarding differential diagnosis. Additionally, further studies are needed to check for relationships between the probability metrics in the multimodal space and each patient's functional severity. The present approach could also be used in presymptomatic patients to look for markers in early stages of bvFTD. In addition, longitudinal assessments of the proposed metrics over time should be studied to evaluate their utility in follow-up studies of individual patients.

5. Conclusions

This work presented a fully automated, cross-site, multimodal, highly sensitive, and specific computational framework to assess bvFTD based on brain imaging data. Our approach provided a robust classification and assisted in the generation of a multimodal space in which each case can be assessed individually to obtain clinical insights via multimodal brain metrics. The method was validated over 103 subjects (60 HCs and 43 patients) from three different centers with high accuracy in the prediction of the state of health and disease, representing a major advance in the consolidation of multimodal image quantification and decision-support methods for clinical diagnosis. Thus, our study provides crucial support to the inclusion of data-driven methods and FC-based metrics as a gold standard in clinical neuroscience.

Acknowledgments

AI is partially supported by grants from CONICET, CONICYT / FONDECYT Regular [grant number 1170010], FONDAP [grant number 15150012], INECO Foundation, by the Inter-American Development Bank (IDB), by PICT [grant number 2017-1818, 2017-1820], and by the Global Brain Health Institute (GBHI). Also, this work was supported in part by funding to Forefront, a collaborative research group specialized in the study of frontotemporal dementia and motor neurone disease, from the National Health and Medical Research Council (NHMRC) of Australia program grant [grant number APP1037746] and the Australian Research Council (ARC) Centre of Excellence in Cognition and its Disorders Memory Program [grant number CE110001021]. FK is supported by an NHMRC-ARC Dementia Research Development Fellowship [grant number APP1097026]. RLR is supported by the ARC Centre of Excellence in Cognition and its Disorders Memory Program [grant number CE110001021] and by the Appenzeller Neuroscience Fellowship in Alzheimer's Disease. OP is supported by an NHMRC Senior Research Fellowship [grant number APP1103258]. COLCIENCIAS [grant number 371-2011, 697-2014], temporal union PUJ-HUSI, Bogotá Colombia.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.dadm.2019.06.002.

RESEARCH IN CONTEXT

1. Systematic review: Based on a review of high-quality, web of science (WoS)-indexed works, we assessed the advantages, limitations, and empirical inconsistencies of automated computational methods implemented to discriminate patients with behavioral variant frontotemporal dementia (bvFTD) from healthy controls using neuroimaging features such as atrophy and functional connectivity (FC) measures.
2. Interpretation: To our knowledge, this is the first study presenting an automatic, cross-center, computational approach which combines atrophy and FC measures for classifying patients with bvFTD and controls. Our method, validated over 103 subjects (60 controls and 43 patients) across three different countries, yielded high classification accuracy (>90). This level of performance underscores the potential of analyzing multimodal information through machine learning approaches for dementia diagnosis.
3. Future directions: To assess the potential clinical applications of our automatic computational approach, it is essential to study genetic forms of bvFTD during presymptomatic stages, include other neurodegenerative conditions, and perform follow-up studies of individual patients.

References

[1] Cohen JD, Daw N, Engelhardt B, Hasson U, Li K, Niv Y, et al. Computational approaches to fMRI analysis. Nat Neurosci 2017;20:304.
[2] Huys QJ, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 2016;19:404.
[3] Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. NeuroImage 2017;145:137–65.
[4] Forman MS, Farmer J, Johnson JK, Clark CM, Arnold SE, Coslett H, et al. Frontotemporal dementia: clinicopathological correlations. Ann Neurol 2006;59:952–62.

[5] Parra M, Baez M, Allegri F, Nitrini R, Lopera F, Slachevsky A, et al. Dementia in Latin America: assessing the present and envisioning the future. Neurology 2018;90:1–11.
[6] Piguet O, Hornberger M, Mioshi E, Hodges JR. Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. Lancet Neurol 2011;10:162–72.
[7] Pievani M, Filippini N, Van Den Heuvel MP, Cappa SF, Frisoni GB. Brain connectivity in neurodegenerative diseases from phenotype to proteinopathy. Nat Rev Neurol 2014;10:620–33.
[8] Palop JJ, Chin J, Mucke L. A network dysfunction perspective on neurodegenerative diseases. Nature 2006;443:768.
[9] Dopper EG, Rombouts SA, Jiskoot LC, den Heijer T, de Graaf JRA, de Koning I, et al. Structural and functional brain connectivity in presymptomatic familial frontotemporal dementia. Neurology 2013;80:814–23.
[10] Dukart J, Mueller K, Horstmann A, Barthel H, Möller HE, Villringer A, et al. Combined evaluation of FDG-PET and MRI improves detection and differentiation of dementia. PLoS One 2011;6:e18111.
[11] Bron EE, Smits M, Papma JM, Steketee RM, Meijboom R, de Groot M, et al. Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using structural and advanced MRI. Eur Radiol 2017;27:3372–82.
[12] Zhang Y, Schuff N, Camacho M, Chao LL, Fletcher TP, Yaffe K, et al. MRI markers for mild cognitive impairment: comparisons between white matter integrity and gray matter volume measurements. PloS One 2013;8:e66367.
[13] Kuceyeski A, Zhang Y, Raj A. Linking white matter integrity loss to associated cortical regions using structural connectivity information in Alzheimer's disease and fronto-temporal dementia: the Loss in Connectivity (LoCo) score. Neuroimage 2012;61:1311–23.
[14] Tahmasian M, Shao J, Meng C, Grimmer T, Diehl-Schmid J, Yousefi BH, et al. Based on the network degeneration hypothesis: separating individual patients with different neurodegenerative syndromes in a preliminary hybrid PET/MR study. J Nucl Med 2016;57:410–5.
[15] Zhou J, Greicius MD, Gennatas ED, Growdon ME, Jang JY, Rabinovici GD, et al. Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. Brain 2010;133:1352–67.
[16] Baez S, Couto B, Torralva T, Sposato LA, Huepe D, Montanes P, et al. Comparing moral judgments of patients with frontotemporal dementia and frontal stroke. JAMA Neurol 2014;71:1172–6.
[17] Melloni M, Billeke P, Baez S, Hesse E, de la Fuente L, Forno G, et al. Your perspective and my benefit: multiple lesion models of self-other integration strategies during social bargaining. Brain 2016;139:3022–40.
[18] Sedeño L, Piguet O, Abrevaya S, Desmaras H, García-Cordero I, Baez S, et al. Tackling variability: a multicenter study to provide a gold-standard network approach for frontotemporal dementia. Hum Brain Mapp 2017;38:3804–22.
[19] Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 2011;134:2456–77.
[20] Piguet O. Eating disturbance in behavioural-variant frontotemporal dementia. J Mol Neurosci 2011;45:589–93.
[21] Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M, et al. Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci 2017;20:299–303.
[22] Clarkson MJ, Cardoso MJ, Ridgway GR, Modat M, Leung KK, Rohrer JD, et al. A comparison of voxel and surface based cortical thickness estimation methods. Neuroimage 2011;57:856–65.
[23] Chung J, Yoo K, Lee P, Kim CM, Roh JH, Park JE, et al. Normalization of cortical thickness measurements across different T1 magnetic resonance imaging protocols by novel W-Score standardization. NeuroImage 2017;159:224–35.
[24] Jenkinson M, Bannister P, Brady JM, Smith SM. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. NeuroImage 2002;17:16.
[25] Laumann TO, Snyder AZ, Mitra A, Gordon EM, Gratton C, Adeyemo B, et al. On the stability of BOLD fMRI correlations. Cereb Cortex 2016:1–14.
[26] Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 1995;34:537–41.
[27] Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31:968–80.
[28] Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, et al. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. AJNR Am J Neuroradiol 2001;22:1326–33.
[29] Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. Bioinformatics 2007;23:2507–17.
[30] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai 1995:1137–45.
[31] Long D, Wang J, Xuan M, Gu Q, Xu X, Kong D, et al. Automatic classification of early Parkinson's disease with multi-modal MR imaging. PloS One 2012;7:e47714.
[32] Hall MA. Correlation-based feature selection of discrete and numeric class machine learning. San Francisco: Morgan Kaufmann Publishers Inc.; 2000.
[33] Donnelly-Kehoe PA, Pascariello GO, Gomez JC, Initiative ADN. Looking for Alzheimer's disease morphometric signatures using machine learning techniques. J Neurosci Methods 2017;302:24–34.
[34] Vapnik V. The nature of statistical learning theory. Springer science & business media; 2013.
[35] Meyer S, Mueller K, Stuke K, Bisenius S, Diehl-Schmid J, Jessen F, et al. Predicting behavioral variant frontotemporal dementia with pattern classification in multi-center structural MRI data. NeuroImage Clin 2017;14:656–62.
[36] Oxtoby NP, Alexander DC. Imaging plus X: multimodal models of neurodegenerative disease. Curr Opin Neurol 2017;30:371.
[37] Muñoz-Ruiz MÁ, Hartikainen P, Koikkalainen J, Wolz R, Julkunen V, Niskanen E, et al. Structural MRI in frontotemporal dementia: comparisons between hippocampal volumetry, tensor-based morphometry and voxel-based morphometry. PLoS One 2012;7:e52531.
[38] Zuo X-N, Xu T, Jiang L, Yang Z, Cao X-Y, He Y, et al. Toward reliable characterization of functional homogeneity in the human brain: preprocessing, scan duration, imaging resolution and computational space. Neuroimage 2013;65:374–86.
[39] Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci 2015;18:1664.
[40] Donnelly-Kehoe P, Pascariello G, Quaglino M, Nagel J, Gómez J. The changing brain in healthy aging: a multi-MRI machine and multicenter surface-based morphometry study. In: 12th International Symposium on Medical Information Processing and Analysis: International Society for Optics and Photonics; 2017. . 101600B.
[41] Bano D, Agostini M, Melino G, Nicotera P. Ageing, neuronal connectivity and brain disorders: an unsolved ripple effect. Mol Neurobiol 2011;43:124–30.
[42] Yates D. Neurodegenerative disease: Neurodegenerative networking. Nat Rev Neurosci 2012;13:288.
[43] Raj A, Kuceyeski A, Weiner M. A network diffusion model of disease progression in dementia. Neuron 2012;73:1204–15.
[44] Mosconi L, Tsui WH, Herholz K, Pupi A, Drzezga A, Lucignani G, et al. Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. J Nucl Med 2008;49:390–8.
[45] Day GS, Farb NA, Tang-Wai DF, Masellis M, Black SE, Freedman M, et al. Salience network resting-state activity: prediction of frontotemporal dementia progression. JAMA Neurol 2013;70:1249–53.
[46] Garcia-Cordero I, Sedeno L, de la Fuente LA, Slachevsky A, Forno G, Klein F, et al. Feeling, learning from and being aware of inner states: interoceptive dimensions in neurodegeneration and stroke. Philos Trans R Soc Lond B Biol Sci 2016;371:20160006.