Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Beyond group classification: probabilistic differential diagnosis of frontotemporal dementia and Alzheimer's disease with MRI and CSF biomarkers.

Agnès Pérez-Millan
Hospital Clínic de Barcelona
Bertrand Thirion
Inria Saclay - Île-de-France Research Centre
Neus Falgàs
Hospital Clínic de Barcelona
Sergi Borrego-Écija
Hospital Clínic de Barcelona
Beatriz Bosch
Hospital Clínic de Barcelona
Jordi Juncà-Parella
Hospital Clínic de Barcelona
Adrià Tort-Merino
Hospital Clínic de Barcelona
Jordi Sarto
Hospital Clínic de Barcelona
Josep Maria Augé
Hospital Clínic de Barcelona
Anna Antonell
Hospital Clínic de Barcelona
Nuria Bargalló
Hospital Clínic de Barcelona
Mircea Balasa
Hospital Clínic de Barcelona
Albert Lladó
Hospital Clínic de Barcelona
Raquel Sánchez-Valle
Hospital Clínic de Barcelona
Roser Sala-Llonch (roser.sala@ub.edu)
University of Barcelona

Research Article

Keywords: Alzheimer's disease, frontotemporal dementia, magnetic resonance imaging, machine learning, CSF biomarkers, SVM, individual probability

Posted Date: December 2nd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3627150/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Neuroimaging and fluid biomarkers are used in clinics to differentiate frontotemporal dementia (FTD) from Alzheimer's disease (AD) and other neurodegenerative and non-neurodegenerative disorders. We implemented a machine learning (ML) algorithm that provides individual probabilistic scores for these patients based on magnetic resonance imaging (MRI) and cerebrospinal fluid (CSF) data.

Methods

We used a calibrated classifier with a Support Vector Machine with MRI data. We obtained group classifications and individual probabilities associated with group correspondence. We used the individual probabilities to address the clinical problem of confidence in the diagnosis. We investigated whether combining MRI and CSF levels of Neurofilament light (NfL) and 14-3-3 could improve the diagnosis confidence.

Results

215 AD patients (65 ± 10 years, 137 women), 103 FTD patients (64 ± 8 years, 49 women), and 173 healthy controls (CTR) (59 ± 15 years, 106 women) were studied. With MRI data only, we obtained accuracies of 88% in the AD vs. healthy controls (CTR) classification, 87% for FTD vs. CTR, 82% for AD vs. FTD, and 80% when differentiating the three groups. A total of 74% of FTD and 73% of AD participants have a high (≥ 0.8) probability of accurate diagnosis in the FTD vs. AD comparison. Adding CSF-NfL and 14-3-3 levels slightly improved the accuracy and the number of patients in the high diagnosis confidence group.

Conclusion

We propose a ML algorithm that provides individual diagnostic probabilities, and we validate it using MRI and/or CSF data. Our solution holds promise towards clinical applications as support to clinical findings or in settings with limited access to expert diagnoses.

1. INTRODUCTION

Alzheimer's Disease (AD) is the most frequent neurodegenerative disorder. Frontotemporal dementia (FTD) is also frequent in people younger than 65 years old and is the main differential diagnosis with AD in this age group. AD and FTD are characterized by prototypical clinical features and patterns of progressive brain atrophy that constitute the disease's fingerprint or signature. An early and accurate diagnosis is essential for treatment, prognosis, and genetic counseling. However, there is considerable

individual variability in the clinical features, especially in the early stages. This limits the accuracy of the clinical diagnosis at this stage.

During the last two decades, fluid biomarker studies have substantially improved the diagnosis of neurodegenerative dementias. The current clinical criteria for AD diagnosis include cerebrospinal fluid (CSF) biomarkers, such as the amyloid-beta protein 42 (Aβ42), the total tau (t-tau), and phosphorylated tau (p-tau) (1, 2). However, currently, FTD criteria do not include biochemical markers. Neurofilament light chain (NfL) levels, a marker of neuroaxonal damage, and 14-3-3 protein levels, a marker of synaptic-neuronal loss, have been both proposed as nonspecific neurodegeneration markers that could support the diagnosis of FTD, although their levels are also increased in AD compared to controls (3–6).

Magnetic Resonance Imaging (MRI) is broadly used in the study of AD and FTD, both at the research and the clinical levels. Visual evaluation of the atrophy pattern is mainly used in the clinical setting (7, 8). Quantitative MRI studies have described patterns of cortical thickness and gray matter (GM) volume loss in AD and FTD at the group level when compared separately with healthy populations (9–14). However, quantitative MRI studies are only scarcely used in clinics due to technical difficulties and limited accuracy in performing the diagnosis at the individual level.

A growing body of evidence supports the role of machine learning (ML) techniques using brain MRI (15–17) to support the clinical diagnosis of these two dementias (18–22). Many studies have shown that a support vector machine (SVM) with neuroimaging data differentiates AD or FTD patients from healthy controls (22–27). However, fewer studies exist on the differential diagnosis of these two dementias, even though the clinical symptoms of FTD and AD can display a substantial overlap between them (28–30).

In this study, we aimed to develop a probabilistic computer-aided classification method for FTD and AD, using MRI data assuming that there will be overlapping and differential brain patterns in these two neurodegenerative disorders. Then, we addressed the clinical problem of diagnosis confidence using individual prediction probabilities. Finally, we proposed investigating whether combining MRI and CSF biomarkers could lead to better differentiation of these two dementias and gain more confidence in the diagnosis.

2. MATERIALS AND METHODS

2.1. Participants

We recruited the participants from the Alzheimer's disease and other cognitive disorders unit of the Hospital Clínic de Barcelona (HCB), Barcelona, Spain. All participants underwent a complete clinical and cognitive evaluation and a 3T high-resolution structural MRI scan. Participants with a history of stroke, traumatic brain injury, major psychiatric disorder, or alcohol abuse were excluded.

All AD participants fulfilled the criteria for mild dementia due to AD (1, 2) supported by the CSF biomarkers profile suggesting underlying AD neuropathology according to National Institute on

Aging/Alzheimer's Association Research Framework 2018 (31). The FTD participants fulfilled the diagnostic criteria for either behavioral variant frontotemporal dementia (bvFTD) or FTD-related primary progressive aphasia (PPA) phenotypes, including Semantic Variant Primary Progressive Aphasia (svPPA) and Nonfluent Variant Primary Progressive Aphasia (nfvPPA) (32, 33). All FTD patients included here showed a CSF profile not suggestive of AD. Healthy adults (CTR) had cognitive performance within the normative range (cutoff 1.5 SD from the normative mean).

The HCB Ethics Committee approved the study (HCB 2019/0105), and all participants gave written informed consent.

2.2 Biochemical markers

We used commercially available single-analyte enzyme-linked immunosorbent assay (ELISA) kits to determine levels of CSF NfL (IBL International, Hamburg, Germany) and CSF 14-3-3 (CircuLex, MBL International Corporation, Woburn, MA) at the Alzheimer's disease and other cognitive disorders unit laboratory, Barcelona, Spain.

2.3 MRI acquisition

We acquired a high-resolution 3D structural dataset (T1-weighted, MP-RAGE, repetition time = 2.300 ms, echo time = 2.98 ms, 240 slices, field-of-view = 256 mm, voxel size = 1 × 1 × 1 mm) for everyone at each time point in a 3T Magnetom Trio Tim scanner (Siemens Medical Systems, Germany) upgraded to a 3T Prisma scanner (Siemens Medical Systems, Germany) during the study.

2.4 MRI processing

We used the processing stream available in FreeSurfer version 6.0 (http://surfer.nmr.mgh.harvard.edu.sire.ub.edu/) to perform cortical reconstruction and volumetric segmentation of the T1-weighted acquisitions. FreeSurfer allowed us to obtain cortical thickness (CTh) maps and segment the subcortical structures (34,35). From reconstructed data, we got summary measures of mean CTh and GM volumes across the left and right hemispheres and summary measures of mean CTh in 68 cortical regions and GM volumes of 16 subcortical structures, all derived from atlases available in FreeSurfer (36, 37). The estimated intracranial volume estimated with FreeSurfer was used to normalize volume measures. All images and individual segmentations were visually inspected and manually corrected if needed.

2.5 MRI-based individual probabilistic classification algorithm

We used all CTh values, GM subcortical volumes, and the age of the participants to create our ML algorithm. We introduced the regional measures of both hemispheres separately, leading to a total of 84 features per subject (see Supplementary Material).

We first converted MRI data to z-scores. We implemented a calibrated classifier with an SVM as a base estimator to predict these values. For each classifier, we fitted a regression that distributes the classifier's output to calibrate the probability between 0 and 1. We created classifiers for each pair of diagnostic groups (AD vs. CTR, FTD vs. CTR, and AD vs. FTD) and across the three groups (AD vs. FTD vs. CTR). Then, we subdivided the FTD group into bvFTD and PPA, and we used them as independent groups in a new set (AD vs. bvFTD, AD vs. PPA, CTR vs. bvFTD, CTR vs. PPA, and bvFTD vs. PPA). All the comparisons were performed with a 5-fold cross-validation to evaluate the performance of the classification. Then, we analyzed the importance of each region for the decision of the classification through a permutation feature importance estimation (38) using the test data of each run. The higher the weight, the larger the importance of the feature in the classification.

We obtained individual probabilities associated with group correspondences as output values for each test data point given by the calibrated SVM. Notably, they had complementary values (i.e., the probability of one group is equal to 1 minus the probability of the other in the classification between two diagnostics), and they were directly associated with the output category (i.e., the final classification was the one with probability > 0.5). We conventionally set two levels of diagnosis confidence: an individual probability ≥ 0.8 (or ≤ 0.2) was considered to provide high diagnosis confidence, while probabilities between 0.2 and 0.8 were considered a “gray zone”, with lower or insufficient diagnosis confidence for the clinical decision. Thus, we estimated the accuracy and the number of individuals with a high probability of being from the group for each classification.

Finally, we aimed to explore if NfL and 14-3-3 levels could help diagnose the individuals of the gray zone of the MRI diagnosis for the following comparisons due to the available data: AD vs. CTR, FTD vs. CTR, and AD vs. FTD. Thus, we created a reduced dataset with subjects having MRI data, NfL, and 14-3-3 levels. We trained and tested the proposed algorithm in 3 situations: MRI-based algorithm, CSF-based algorithm, and MRI and CSF-based algorithm to study if the individual probabilities towards the actual class increased. We did not include Aβ42, t-tau, and p-tau levels to avoid circularity, as these markers were used in the clinical diagnosis according to current criteria.

We implemented the ML algorithm in Python version 3.10.6 (www.python.org) with the Scikit-learn library (39).

3. RESULTS

3.1. Sample demographics

The prospective study includes 491 subjects: 215 AD, 103 FTD (56 bvFTD, 24 svPPA, 21 nfvPPA, and 2 PPA), and 173 cognitively normal control (CTR) participants. A subset of the study participants had CSF measures available: NfL (N = 365) and 14-3-3 (N = 182). Table 1 shows demographic information, group statistics, and biomarker levels. As expected, CSF biomarkers levels showed significant differences between groups (corrected p-value < 0.05). There were differences in age and sex. As expected, based on

previous studies, AD and CTR groups had more women than men; meanwhile, the FTD groups were more harmonized. Regarding age, healthy controls were younger than AD and FTD participants.

Table 1

Group summaries written as each measure's mean and standard deviation. We calculated differences between groups using Fisher Test for sex or the Anova Test for the rest of the variables. We highlighted the significant group differences in bold. We measured pairwise differences with a Benjamini-Hochberg correction p-value). CTR: healthy subjects, AD: Alzheimer's disease, FTD: frontotemporal dementia, NfL: neurofilament light chain.

	CTR	AD	FTD	CTR-AD p-values	CTR-FTD p-values	AD-FTD p-values
N MRI	173	215	103	—	—	—
Sex at MRI, Men/Women	67/106	78/137	54/49	0.67	**0.049**	**0.022**
Age at MRI, years (SD)	59.4 (15.0)	65.0 (9.9)	63.7 (8.3)	**1.3e-5**	**0.0045**	0.39
N CSF NfL	112	175	78	—	—	—
CSF NfL, pg/mL (SD)	536.1 (312.6)	1134.7 (587.1)	2340.6 (1736.3)	**1.2e-07**	**< 2e-16**	**5.9e-06**
N CSF 14-3-3	50	68	64	—	—	—
CSF 14-3-3, pg/mL (SD)	2531.9 (748.2)	5727.3 (2303.5)	4234.9 (1869.1)	**< 2e-16**	**3.0e-06**	**5.9e-06**

3.2. MRI-based probabilistic classification algorithm

We estimated the accuracy performance of our algorithm as the mean accuracy obtained in each k-fold of the test data. We got an accuracy of 88 ± 8% when discriminating AD patients from CTR, and 87 ± 4% when determining FTD patients from CTR. When we tried classifying AD vs. FTD patients, the accuracy was 82 ± 6%. Finally, we obtained an accuracy of 77 ± 6% when discriminating between the three groups (AD vs. FTD vs. CTR) (Table 2).

Table 2
Classification performance of the different approaches and the percentage of participants with a higher probability of 80% in the diagnosis grouped by diagnosis.

	AD vs CTR	FTD vs CTR	AD vs FTD
MRI all data	Accuracy: 87.7%	Accuracy: 86.9%	Accuracy: 81.8%
(N = 491)	AD: 73.4%	FTD: 74.2%	AD: 73.3%
	CTR: 64.5%	CTR: 73.3%	FTD: 74.2%
MRI reduced data	Accuracy: 88.5%	Accuracy: 85.6%	Accuracy: 84.6%
(N = 178)	AD: 67.2%	FTD: 68.1%	AD: 53.2%
	CTR: 55.3%	CTR: 54.3%	FTD: 54.4%
CSF data	Accuracy: 93.0%	Accuracy: 86.6%	Accuracy: 83.8%
(N = 178)	AD: 72.1%	FTD: 72.1%	AD: 40.6%
	CTR: 71.7%	CTR: 23.5%	FTD: 45.9%
MRI and CSF data	Accuracy: 90.3%	Accuracy: 86.5%	Accuracy: 88.5%
(N = 178)	AD: 68.1%	FTD: 70.8%	AD: 60.7%
	CTR: 64.4%	CTR: 53.2%	FTD: 55.1%

As can be seen in Fig. 1, the resulting algorithms were well-calibrated, which allowed us to create confidence ranges in the algorithm classification. The comparison of AD vs. CTR showed that 73% of AD participants and 65% of CTR participants presented a probability higher than 0.8. In the FTD vs. CTR comparison, we found 74% FTD participants and 73% CTR participants with a probability ≥ 0.8. Finally, when discriminating AD vs. FTD, we found 73% AD participants and 74% FTD participants with probabilities above 0.8 for being classified as AD or FTD, respectively. Figure 2 shows the density of the individual probabilities and how the distribution between the clinical and the algorithm diagnosis is distributed within the group with an individual probability ≥ 0.8. Notably, the algorithm diagnosis and the clinical algorithm did not always coincide, also inside the higher probability of 0.8 (high confidence).

Then, we aimed to study the FTD clinical subtypes separately. Due to limitations in sample size, we merged svPPA and nfvPPA in the same group-PPA. We obtained 91 ± 2% accuracy for classifying bvFTD patients vs. CTR and 93 ± 4% when discriminating PPA patients from CTR. In both cases, the accuracy increased compared to the accuracy reported for all FTD together (87 ± 4%). Compared with AD, we obtained 85 ± 3% for the bvFTD vs. AD comparison and 91 ± 3% for the PPA vs AD. Finally, we obtained an accuracy of 68 ± 6% discriminating bvFTD from PPA.

3.3. Important MRI regions for classification

Figure 3 shows the region weights associated with each comparison. In summary, when comparing AD versus CTR, the GM volume of the hippocampus, putamen, and amygdala played the most crucial role. For FTD vs. CTR, we found that occipital, parietal, and frontal regions emerged as the top regions for the classification. Finally, when discriminating both dementias (AD vs. FTD), we found a widespread pattern in which the CTh measures were generally more important than subcortical GM volumes, especially those in the frontal lobe.

The results of the most crucial regions in the classifications considering bvFTD and PPA participants are shown in Fig. 4. When discriminating bvFTD from CTR, the frontal and temporal lobes and the GM volume of the ventricles were the most important areas. In contrast, when differentiating between CTR and PPA participants, the top regions were GM volumes of the hippocampus, amygdala, and temporal lobe. When discriminating AD from bvFTD, the most important areas were the temporal, parietal, and occipital lobes. The frontal, parietal, and occipital lobes emerged in the PPA vs. AD discrimination. Finally, when discriminating bvFTD vs. PPA, the regions which contributed the most were the frontal, temporal, and occipital lobes.

3.4. Individual probabilities using MRI and CSF data

The group classification performance of the algorithm and the percentage of participants with an individual probability ≥ 80% using MRI-only, CSF-only, and combined MRI and CSF data are presented in Table 2. Adding NfL and 14-3-3 data to the MRI data did not improve the results. However, some conclusions could be derived from the results.

For comparing AD vs. CTR, adding CSF data to the MRI did not increase the classification rate (Table 2, Fig. 5). However, CSF data (NfL and 14-3-3) alone was enough to discriminate between AD and CTR participants.

Contrarily, in the comparison between FTD and CTR having MRI and CSF, more participants were classified with high confidence, with a probability ≥ 0.8 (see Table 2 and Fig. 5). The accuracy with only CSF data and that of MRI + CSF data was very similar, 87 ± 8% and 86 ± 9%, respectively; however, when adding CSF data, CTR subjects with a CTR-probability higher than 0.8 increased from 23–53%. Thus, in this case, combining the MRI and CSF data reduced the participants in the gray zone of the diagnosis.

Finally, when we compared AD and FTD participants, combining MRI and CSF data increased both the accuracy and the number of subjects with a probability ≥ 0.8 (see Table 2 and Fig. 5).

4. DISCUSSION

In this study, we implemented a machine learning algorithm that discriminates FTD and AD subjects using data from structural MRI. In addition, our algorithm was able to differentiate subtypes of FTD with good accuracy. Clinical diagnosis requires decisions at the individual level, and the degree of confidence in the diagnosis is key in managing the patient. We approach the clinical question of diagnosis

confidence using individual probabilities. Among our key results, we found that 74% FTD and 73% AD participants showed an individual probability ≥ 0.8 of being well-classified by the algorithm in the FTD vs. AD comparison. Adding CSF neurodegeneration markers (NfL and 14-3-3) levels did not significantly improve the diagnosis classification or the number of patients with high individual probability for the diagnosis.

Previous ML algorithms using structural MRI data have reported accuracies between 76 and 97% for AD vs. CTR, 72–88% for FTD versus CTR, 51–90% for AD versus FTD, and 54–70% in discriminating between AD, FTD, and CTR (7, 11, 18, 21, 22, 40–47). These studies used different algorithms, with the SVM being the most common.

We obtained accuracies that are in accordance with, or even outperformed, previously reported algorithms, especially for AD vs. FTD (16, 17, 48–50). We have differentiated FTD expressions (bvFTD and PPA) against AD or CTR, outperforming previously published works (11, 21, 46). First, regarding the comparisons with CTR, for bvFTD, we obtained a 91% accuracy, and in the case of the PPA participants, an accuracy of 93%. When classifying bvFTD and PPA separately against AD, we obtained accuracies up to 90% for both cases. However, when we tried to classify bvFTD vs. PPA, we obtained an accuracy of 68%, which is lower than the accuracy reported by Kim et al. (41), probably due to differences in the algorithm.

We depicted the patterns that drive accuracy for each classification setting, to obtain a comprehensive explanation of structural changes in both dementias. The GM volume of the hippocampus, putamen, and amygdala were essential in differentiating AD from CTR. By contrast, when differentiating FTD from CTR, the cortical regions were the most important, especially the CTh of occipital, parietal, and frontal regions. According to this, GM volumes of subcortical areas could help to identify AD patients, and the thickness of the cortex could be the key to identify FTD participants. This is in agreement with findings obtained with more classical analysis methods (12, 51–57). Finally, regarding the FTD variants, frontal brain regions emerged for the bvFTD, and hippocampus and temporal regions were the most important in PPA, as reported before (51).

Besides reaching good accuracies, one of the main novelties of our work is that we obtained the individual probabilities for each diagnosis in all comparisons. Notably, as we built our first set of algorithms uniquely with MRI data, these probabilities might reflect each individual's brain atrophy severity. Using these values, we could identify the participants with high diagnosis confidence (with a probability upper to 80%) and those who do not have that high confidence that could be a candidate for further evaluations. Notably, more than 70% of AD and FTD participants were classified with high diagnosis confidence in the FTD vs. AD comparison.

Other studies using multimodal information also reported high classification accuracy combining data from different imaging modalities or other biological and clinical measures (24, 40, 49, 58). Even so, in some cases, our scores with only structural MRI data showed better accuracy (18, 46, 59, 60). Here, we evaluated if adding CSF data to the MRI could improve the accuracy or the number of participants with a

high diagnosis confidence. This could help in reducing the number of clinical tests that usually have to take these patients. This approach could detect which participants need extra clinical tests and which ones; another clinical examination would not provide new evidence. In our cohort, adding NfL and 14-3-3 CSF data to the MRI data provided was beneficial for the accuracy of the group classification or the number of participants with high individual diagnosis probability, especially in the comparison where FTD participants were involved. The fact that both, NfL and 14-3-3 levels and MRI data reflect neurodegeneration and no other aspects of the pathophysiological processes in these diseases could explain the reduced improvement by adding CSF data. We did not use Aβ42, t-tau, and p-tau to avoid circularity, as they were used for the AD clinical diagnosis.

Overall, our study has several strengths. First, its good performance makes it suitable for potential implementation in a clinical setting, especially in doubtful cases or locations with limited access to expert opinion or additional biomarkers. The key to individual probabilities, thanks to the calibrated algorithm, also might represent a step toward personalized medicine. For example, it could be used to identify the best patients to receive the new drugs or for clinical trials. In addition, at the level of explainable ML, we identified the most critical regions for classification, contributing to the definition of structural atrophy patterns, and may be used for identifying target regions in further studies.

Our study also presents several limitations. First, it is unicentric. It has the advantage that all the participants had the same MRI scanner protocol and clinical criteria for the diagnosis. In the case of using the algorithm in other centers, the increased heterogeneity of the data could worsen the algorithm's performance. Another limitation regarding the FTD participants is that, when looking at the different clinical expressions, we reduced the sample size to approximately 50 participants for each group, and svPPA and nfvPPA had to be studied together. This means that the results are subject to large sampling variability. Future studies could further explore the subanalyses with the FTD phenotype subtypes in more detail. Finally, only some participants had NfL and 14-3-3 data available, and the smaller sample size might have impacted the results.

In conclusion, the proposed diagnosis algorithm has shown high accuracy classification scores with structural MRI data to discriminate AD, FTD, and CTR. This approach also provided individual MRI-based classification probability scores as an ancillary tool for studying the overlapping results between FTD and AD and a surrogate estimation for the confidence in the ML diagnosis.

Abbreviations

AD=Alzheimer’s disease

Aβ42=Amyloid-beta protein 42

bvFTD=Behavioral variant frontotemporal dementia

CSF=Cerebrospinal fluid

CTh=Cortical thickness

CTR=healthy subjects

FTD=Frontotemporal dementia

HCB=Hospital Clínic de Barcelona

ML=Machine Learning

MMSE=Mini-Mental State Examination

MRI=Magnetic Resonance Imaging

NfL=neurofilament light chain

nfvPPA=Nonfluent Variant Primary Progressive Aphasia

PPA=Primary progressive aphasia

SVM=Support Vector Machine

svPPA=Semantic Variant Primary Progressive Aphasia

Declarations

Ethical Approval and Consent to participate

The Hospital Clínic de Barcelona Ethics Committee approved the study (HCB 2019/0105), and all participants gave written informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

A. Pérez-Millan is a recipient of the French Embassy in Spain / Institut français de España fellowship and a travel fellowship from María de Maeztu Unit of Excellence (Institute of Neurosciences, University of Barcelona) MDM-2017-0729. This study was partially funded by Instituto de Salud Carlos III, Spain (grant no. PI20/0448 to Dr. R. Sanchez-Valle, PI19/00449 to Dr. A. Lladó and project PI19/00198 to Dr. M. Balasa and co-funded by the European Union, “Una manera de hacer Europa” AGAUR, Generalitat de Catalunya (SGR 2021- 01126 and 2021 SGR 00523) and by Spanish Ministry of Science and Innovation (PID2020-118386RA-I00/ AEI/10.13039/501100011033 to Dr. R. Sala-Llonch). Dr. B. Thirion is supported the KARAIB AI chair (ANR-20-CHIA-0025-01) and the H2020 Research Infrastructures Grant EBRAIN-Health 101058516. Dr. N. Falgàs is a recipient of the Joan Rodes fellowship from the Instituto de Salud Carlos III, Spain. Dr. S. Borrego-Écijais a recipient of the Joan Rodés Josep Baselga grant from FBBVA.

Authors' contributions

APM, BT, RSV, and RSL contributed to the study's design. APM, NF, SB, BB, JJP, ATM, JS, JMA, AA, NB, MB, AL, and RSV contributed to the data acquisition. APM and BT contributed to the data analyses. APM, BT, NF, AL, RSV, and RSL contributed to interpreting the data. APM, BT, RSV, and RSL contributed to the draft of the article. NF, SB BB, JJP, ATM, JS, JMA, AA, NB, MB, and AL revised the manuscript critically for important intellectual content and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors thank patients, their relatives, and healthy controls for participating in the research.

References

1. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimer’s and Dementia. 2011;7(3):270–9.
2. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. 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. Alzheimer’s and Dementia. 2011;7(3):263–9.
3. Alcolea D, Vilaplana E, Suárez-Calvet M, Illán-Gala I, Blesa R, Clarimón J, et al. CSF sAPPβ, YKL-40, and neurofilament light in frontotemporal lobar degeneration. Neurology. 2017 Jul 11;89(2):178–88.
4. Antonell A, Tort-Merino A, Ríos J, Balasa M, Borrego-Écija S, Auge JM, et al. Synaptic, axonal damage and inflammatory cerebrospinal fluid biomarkers in neurodegenerative dementias. Alzheimer’s & Dementia [Internet]. 2019 Oct 24 [cited 2022 Dec 7]; Available from: https://www.sciencedirect.com/science/article/pii/S1552526019353695

5. McFerrin MB, Chi X, Cutter G, Yacoubian TA. Dysregulation of 14-3-3 proteins in neurodegenerative diseases with Lewy body or Alzheimer pathology. Ann Clin Transl Neurol. 2017 Jul;4(7):466–77.
6. Rohrer JD, Woollacott IOC, Dick KM, Brotherhood E, Gordon E, Fellows A, et al. Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. Neurology. 2016 Sep 27;87(13):1329–36.
7. Davatzikos C, Resnick SM, Wu X, Parmpi P, Clark CM. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. NeuroImage. 2008 Jul 15;41(4):1220–7.
8. Du AT, Schuff N, Kramer JH, Rosen HJ, Gorno-Tempini ML, Rankin K, et al. Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain. 2007 Apr 1;130(4):1159–66.
9. Bocchetta M, Todd EG, Peakman G, Cash DM, Convery RS, Russell LL, et al. Differential early subcortical involvement in genetic FTD within the GENFI cohort. NeuroImage: Clinical. 2021;30:102646.
10. Borrego-Écija S, Sala-Llonch R, van Swieten J, Borroni B, Moreno F, Masellis M, et al. Disease-related cortical thinning in presymptomatic granulin mutation carriers. NeuroImage: Clinical. 2021 Jan 1;29.
11. Canu E, Agosta F, Mandic-Stojmenovic G, Stojković T, Stefanova E, Inuggi A, et al. Multiparametric MRI to distinguish early onset Alzheimer's disease and behavioural variant of frontotemporal dementia. NeuroImage: Clinical. 2017;15:428–38.
12. Contador J, Pérez-Millán A, Tort-Merino A, Balasa M, Falgàs N, Olives J, et al. Longitudinal brain atrophy and CSF biomarkers in early-onset Alzheimer's disease. NeuroImage: Clinical. 2021 Jan 1;32:102804.
13. Möller C, Hafkemeijer A, Pijnenburg YAL, Rombouts SARB, van der Grond J, Dopper E, et al. Joint assessment of white matter integrity, cortical and subcortical atrophy to distinguish AD from behavioral variant FTD: A two-center study. NeuroImage: Clinical. 2015 Jan 1;9:418–29.
14. Möller C, Vrenken H, Jiskoot L, Versteeg A, Barkhof F, Scheltens P, et al. Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. Neurobiology of Aging. 2013 Aug 1;34(8):2014–22.
15. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics [Internet]. 2014 [cited 2023 Mar 13];8. Available from: https://www.frontiersin.org/articles/10.3389/fninf.2014.00014
16. Frizzell TO, Glashutter M, Liu CC, Zeng A, Pan D, Hajra SG, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: A systematic review. Ageing Research Reviews. 2022 May 1;77:101614.
17. Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: A review of machine learning applications. NeuroImage: Clinical. 2018 Jan 1;20:506–22.
18. Bron EE, Smits M, Papma JM, Steketee RME, Meijboom R, de Groot M, et al. Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using

structural and advanced MRI. European Radiology. 2017 Aug 1;27(8):3372–82.

19. Chagué P, Marro B, Fadili S, Houot M, Morin A, Samper-González J, et al. Radiological classification of dementia from anatomical MRI assisted by machine learning-derived maps. Journal of Neuroradiology. 2021 Nov 1;48(6):412–8.
20. Klöppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, et al. Automatic classification of MR scans in Alzheimer's disease. Brain. 2008 Mar;131(3):681–9.
21. Möller C, Pijnenburg YAL, Van Der Flier WM, Versteeg A, Tijms B, De Munck JC, et al. Alzheimer disease and behavioral variant frontotemporal dementia: Automatic classification based on cortical atrophy for single-subject diagnosis. Radiology. 2016 Jun 1;279(3):838–48.
22. Pérez-Millan A, Contador J, Juncà-Parella J, Bosch B, Borrell L, Tort-Merino A, et al. Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. Human Brain Mapping [Internet]. 2023 [cited 2023 Jan 24]; Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.26205
23. Bisenius S, Mueller K, Diehl-Schmid J, Fassbender K, Grimmer T, Jessen F, et al. Predicting primary progressive aphasias with support vector machine approaches in structural MRI data. NeuroImage: Clinical. 2017 Jan 1;14:334–43.
24. Bron EE, Klein S, Papma JM, Jiskoot LC, Venkatraghavan V, Linders J, et al. Cross-cohort generalizability of deep and conventional machine learning for MRI-based diagnosis and prediction of Alzheimer's disease. NeuroImage: Clinical. 2021 Jan 1;31:102712.
25. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert MO, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage. 2011 May 15;56(2):766–81.
26. Magnin B, Mesrob L, Kinkingnéhun S, Pélégrini-Issac M, Colliot O, Sarazin M, et al. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology. 2009 Feb 1;51(2):73–83.
27. 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: Clinical. 2017 Jan 1;14:656–62.
28. Mendez MF. The accurate diagnosis of early-onset dementia. International Journal of Psychiatry in Medicine. 2006;36(4):401–12.
29. Wojtas A, Heggeli KA, Finch N, Baker M, DeJesus-Hernandez M, Younkin SG, et al. C9ORF72 repeat expansions and other FTD gene mutations in a clinical AD patient series from Mayo Clinic. Am J Neurodegener Dis. 2012 May 16;1(1):107–18.
30. Zee J van der, Sleegers K, Broeckhoven CV. Invited Article: The Alzheimer disease–frontotemporal lobar degeneration spectrum. Neurology. 2008 Oct 7;71(15):1191–7.
31. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's and Dementia. 2018 Apr 1;14(4):535–62.

32. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011 Mar;76:1006–14.
33. 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 Sep;134:2456–77.
34. Fischl B, Van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, et al. Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex. 2004 Jan;14(1):11–22.
35. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of the United States of America. 2000 Sep 26;97(20):11050–5.
36. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006 Jul 1;31(3):968–80.
37. Seidman LJ, Faraone SV, Goldstein JM, Goodman JM, Kremen WS, Matsuda G, et al. Reduced subcortical brain volumes in nonpsychotic siblings of schizophrenic patients: A pilot magnetic resonance imaging study. American Journal of Medical Genetics - Neuropsychiatric Genetics. 1997;74(5):507–14.
38. Breiman L. Random Forests. Machine Learning. 2001;45(1):5–32.
39. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(Oct):2825–30.
40. 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 Mar 23;6(3):e18111.
41. Kim JP, Kim J, Park YH, Park SB, Lee JS, Yoo S, et al. Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. NeuroImage: Clinical. 2019 Jan 1;23:101811.
42. Li B, Jang I, Riphagen J, Almaktoum R, Yochim KM, Ances BM, et al. Identifying individuals with Alzheimer's disease-like brains based on structural imaging in the Human Connectome Project Aging cohort. Human Brain Mapping. 2021 Dec 1;42(17):5535–46.
43. Lin W, Tong T, Gao Q, Guo D, Du X, Yang Y, et al. Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease Prediction From Mild Cognitive Impairment. Frontiers in Neuroscience [Internet]. 2018 [cited 2023 Apr 9];12. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2018.00777
44. Moore PJ, Lyons TJ, Gallacher J, Initiative for the ADN. Random forest prediction of Alzheimer's disease using pairwise selection from time series data. PLOS ONE. 2019 Feb 14;14(2):e0211558.
45. Salvatore C, Cerasa A, Battista P, Gilardi MC, Quattrone A, Castiglioni I. Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. Front Neurosci. 2015 Sep 1;9:307.

46. Wang J, Redmond SJ, Bertoux M, Hodges JR, Hornberger M. A Comparison of Magnetic Resonance Imaging and Neuropsychological Examination in the Diagnostic Distinction of Alzheimer's Disease and Behavioral Variant Frontotemporal Dementia. Frontiers in Aging Neuroscience [Internet]. 2016 [cited 2023 Apr 9];8. Available from: https://www.frontiersin.org/articles/10.3389/fnagi.2016.00119
47. Wang SH, Phillips P, Sui Y, Liu B, Yang M, Cheng H. Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. J Med Syst. 2018 Mar 26;42(5):85.
48. Basheera S, Sai Ram MS. Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation. Alzheimer's & Dementia: Translational Research & Clinical Interventions. 2019 Jan 1;5:974–86.
49. Dashtipour K, Taylor W, Ansari S, Zahid A, Gogate M, Ahmad J, et al. Detecting Alzheimer's Disease Using Machine Learning Methods. In: Ur Rehman M, Zoha A, editors. Body Area Networks Smart IoT and Big Data for Intelligent Health Management. Cham: Springer International Publishing; 2022. p. 89–100. (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering).
50. McCarthy J, Collins DL, Ducharme S. Morphometric MRI as a diagnostic biomarker of frontotemporal dementia: A systematic review to determine clinical applicability. NeuroImage: Clinical. 2018 Jan 1;20:685–96.
51. Avants BB, Cook PA, Ungar L, Gee JC, Grossman M. Dementia induces correlated reductions in white matter integrity and cortical thickness: A multivariate neuroimaging study with sparse canonical correlation analysis. NeuroImage. 2010 Apr 15;50(3):1004–16.
52. Dickerson BC, Goncharova I, Sullivan MP, Forchetti C, Wilson RS, Bennett DA, et al. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiology of Aging. 2001;22(5):747–54.
53. Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010 Feb;6(2):67–77.
54. Gil-Navarro S, Lladó A, Rami L, Castellví M, Bosch B, Bargalló N, et al. Neuroimaging and Biochemical Markers in the Three Variants of Primary Progressive Aphasia. DEM. 2013;35(1–2):106–17.
55. Gordon E, Rohrer JD, Fox NC. Advances in neuroimaging in frontotemporal dementia. Journal of Neurochemistry. 2016;138(S1):193–210.
56. Hodges JR, Patterson K. Semantic dementia: a unique clinicopathological syndrome. The Lancet Neurology. 2007 Nov 1;6(11):1004–14.
57. Jack CR, Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology. 2000;55(4):484–9.
58. Chen Y, Sha M, Zhao X, Ma J, Ni H, Gao W, et al. Automated detection of pathologic white matter alterations in Alzheimer's disease using combined diffusivity and kurtosis method. Psychiatry Research: Neuroimaging. 2017 Jun 30;264:35–45.

59. Bouts MJRJ, Möller C, Hafkemeijer A, van Swieten JC, Dopper E, van der Flier WM, et al. Single Subject Classification of Alzheimer's Disease and Behavioral Variant Frontotemporal Dementia Using Anatomical, Diffusion Tensor, and Resting-State Functional Magnetic Resonance Imaging. Journal of Alzheimer's Disease. 2018 Jan 1;62(4):1827–39.
60. Dyrba M, Grothe M, Kirste T, Teipel SJ. Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. Human Brain Mapping. 2015;36(6):2118–31.

Figures

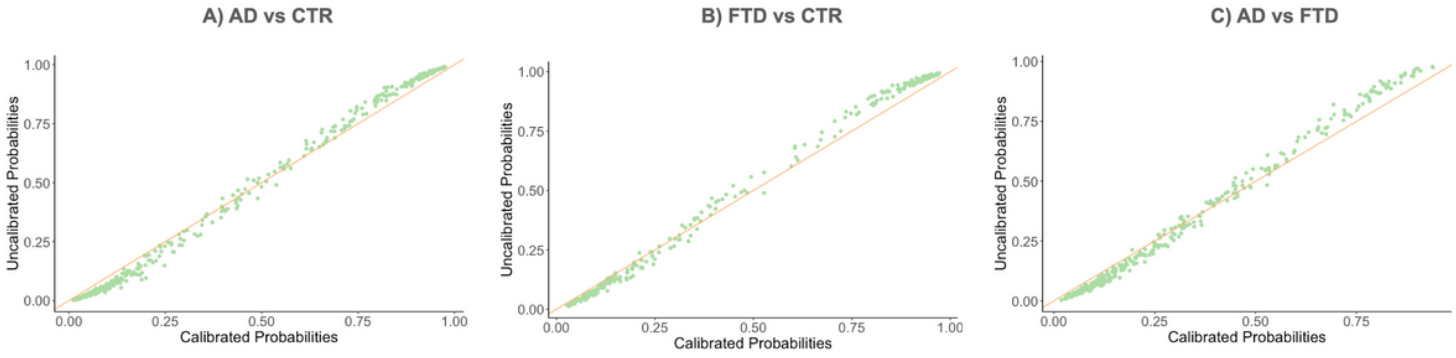


Figure 1

We plot the calibrated probabilities versus the uncalibrated probabilities. Each point represents the individual probability for each classification obtained with calibration and without calibration step. We represent the test data together.

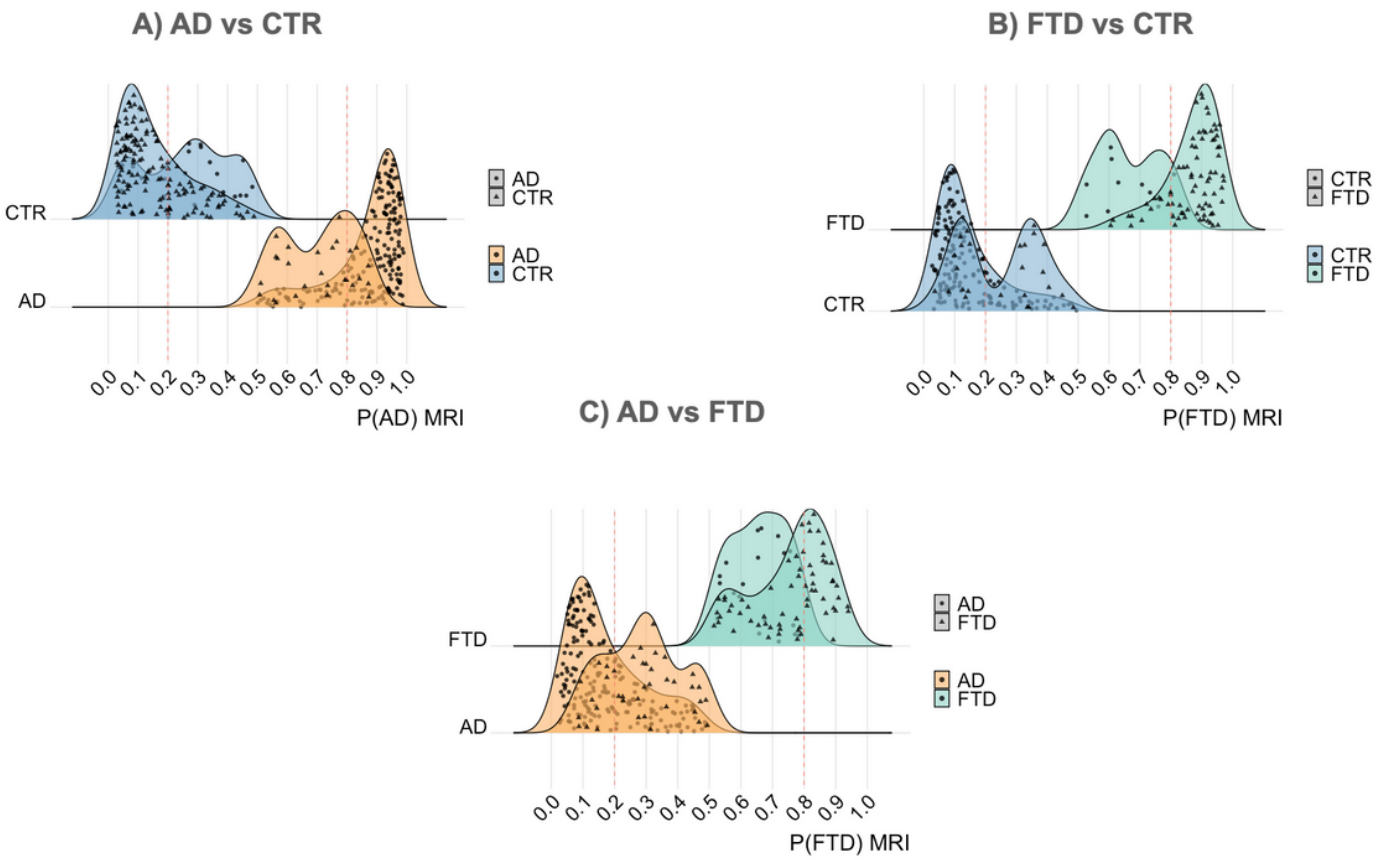


Figure 2

Density plot to study the obtained individual probabilities with the MRI-based algorithm. The clinical diagnosis is identified with triangles or circles, and the algorithm's diagnoses are plotted with different colors. CTR: healthy subjects, AD: Alzheimer's disease, FTD: frontotemporal dementia.

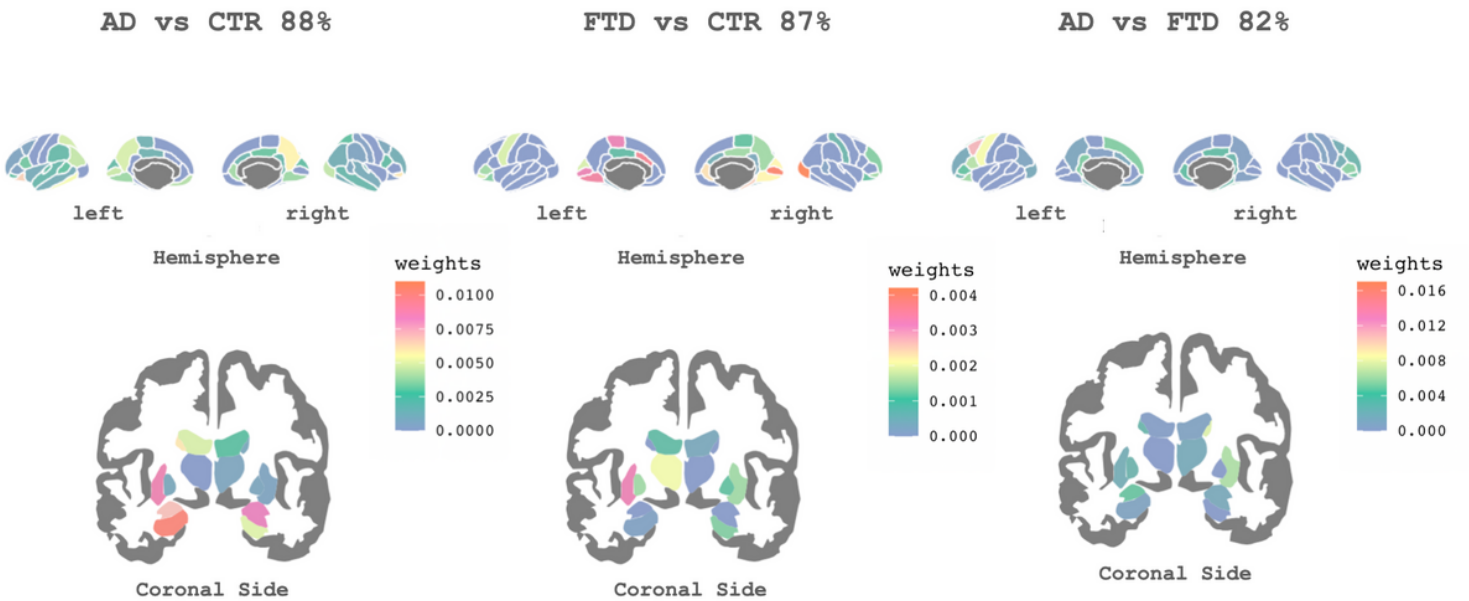


Figure 3

Cortical (top) and subcortical (bottom) patterns of the feature importance of each region associated with AD and FTD. At a higher value major importance of that region for the classification.

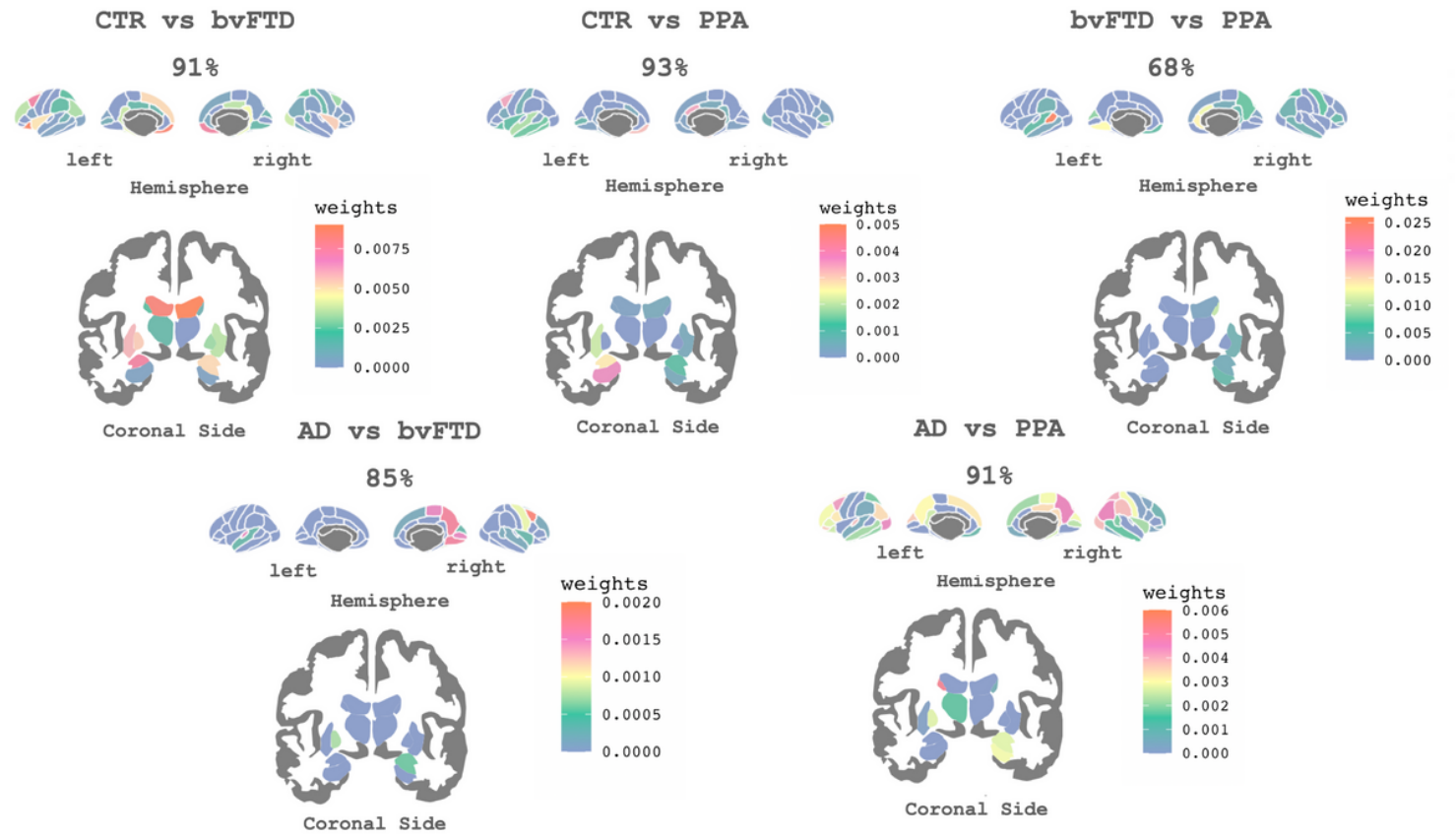


Figure 4

Cortical (top) and subcortical (bottom) patterns of the feature importance of each region associated with bvFTD and PPA. At a higher value major importance of that region for the classification.

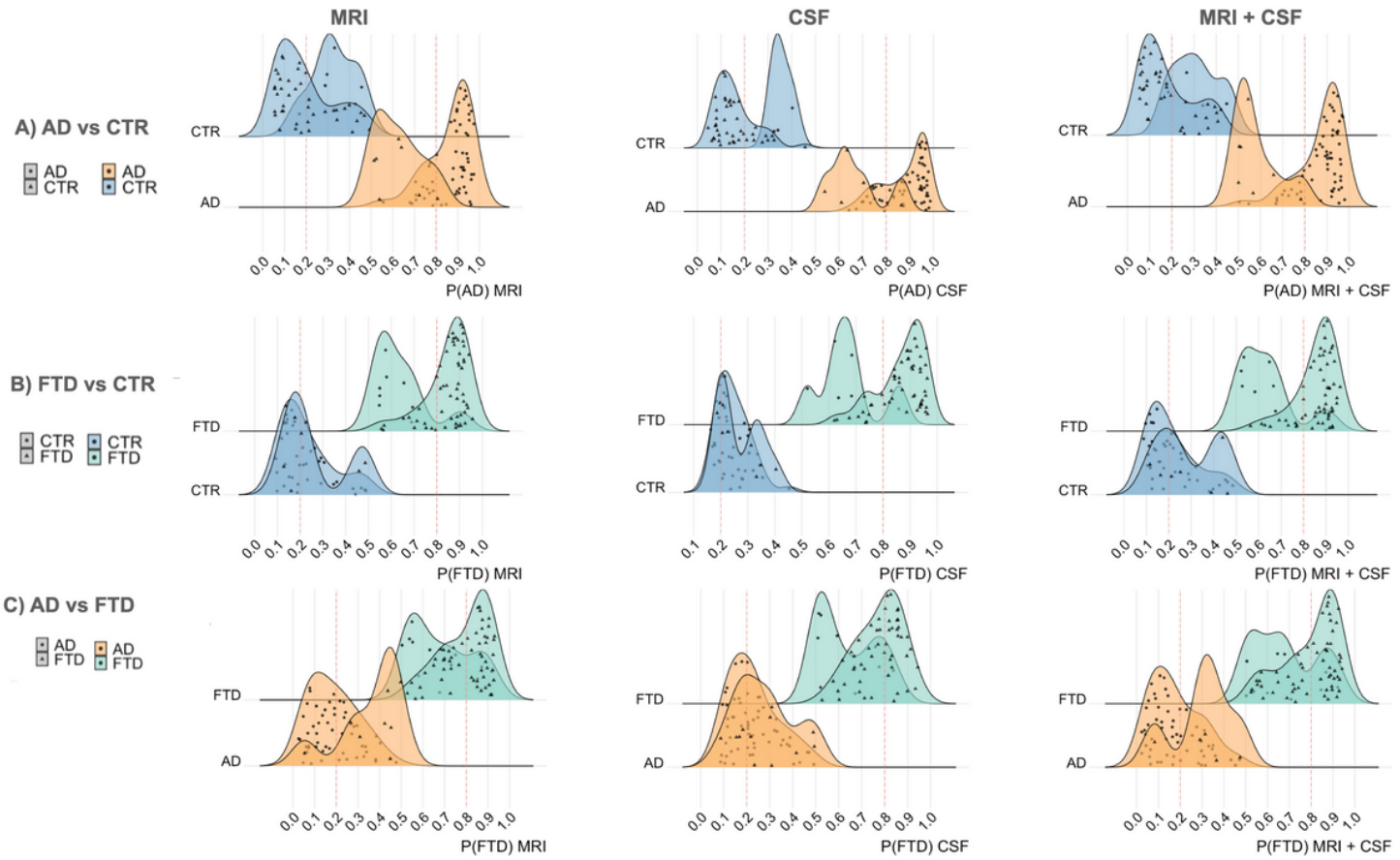


Figure 5

Density plot to study the obtained individual probabilities with the MRI-based, CSF-based, and MRI- and CSF-based algorithms. It can be seen in the clinical diagnosis with triangles or circles and the algorithm's diagnosis, plotted with different colors. CTR: healthy subjects,

Supplementary Files

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

- SupplemntaryMaterial.docx