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

Agnès Pérez-Millan, Bertrand Thirion, Neus Falgàs, Sergi Borrego-Écija, Beatriz Bosch, Jordi Juncà-Parella, Adrià Tort-Merino, Jordi Sarto, Josep Maria Augé, Anna Antonell, et al.

▶ **To cite this version:**

Agnès Pérez-Millan, Bertrand Thirion, Neus Falgàs, Sergi Borrego-Écija, Beatriz Bosch, et al.. Beyond group classification: Probabilistic differential diagnosis of frontotemporal dementia and Alzheimer's disease with MRI and CSF biomarkers. Neurobiology of Aging, 2024, 144, pp.1-11. 10.1016/j.neurobiolaging.2024.08.008 . hal-04851068

HAL Id: hal-04851068
https://inria.hal.science/hal-04851068v1

Submitted on 20 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

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

Agnès Pérez-Millan[1,2,3,4], Bertrand Thirion[4], Neus Falgàs[1], Sergi Borrego-Écija[1], Beatriz Bosch[1], Jordi Juncà-Parella[1], Adrià Tort-Merino[1], Jordi Sarto[1], Josep Maria Augé[5], Anna Antonell[1], Nuria Bargalló[6], Mircea Balasa[1], Albert Lladó[1,2], Raquel Sánchez-Valle[1,2*], Roser Sala-Llonch [2,3,7, 8*†]

[1]Alzheimer's disease and other cognitive disorders unit. Service of Neurology, Hospital Clínic de Barcelona. Fundació Recerca Clínic Barcelona-IDIBAPS, Barcelona, Spain.

[2]Institut de Neurociències, University of Barcelona, Barcelona, Spain.

[3]Department of Biomedicine, Faculty of Medicine, University of Barcelona, Barcelona, Spain

[4]Inria, CEA, Université Paris-Saclay, Paris, France

[5]Biochemistry and Molecular Genetics Department, Hospital Clínic de Barcelona, Barcelona, Spain

[6]Image Diagnostic Centre, Hospital Clínic de Barcelona, CIBER de Salud Mental, Instituto de Salud Carlos III.Magnetic Resonance Image Core Facility, IDIBAPS, Barcelona, Spain

[7]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, Spain

[8]Fundació de Recerca Clínic Barcelona-Institut d'Investigacions Biomèdiques August Pi I Sunyer (FRCB-IDIBAPS), Barcelona, Spain.

† Corresponding author at Roser Sala-Llonch, PhD. Biophysics and Bioengineering Unit. Faculty of Medicine and Health Sciences, University of Barcelona. Casanova 143, 08036 Barcelona (Spain). Tel: +34 934024516

Email address: roser.sala@ub.edu

* These authors contributed equally to this work.

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

Abbreviations

AD=Alzheimer’s disease
Aβ42=Amyloid-beta protein 42
bvFTD=Behavioral variant frontotemporal dementia
CSF=Cerebrospinal fluid
CTh=Cortical thickness
CTR=healthy controls
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

ABSTRACT

Neuroimaging and fluid biomarkers are used to differentiate frontotemporal dementia (FTD) from Alzheimer's disease (AD). We implemented a machine learning algorithm that provides individual probabilistic scores based on magnetic resonance imaging (MRI) and cerebrospinal fluid (CSF) data. We investigated whether combining MRI and CSF levels could improve the diagnosis confidence. 215 AD patients, 103 FTD patients, and 173 healthy controls (CTR) were studied. With MRI data, we obtained an accuracy of 82% for AD vs. FTD. A total of 74% of FTD and 73% of AD participants have a high probability of accurate diagnosis. Adding CSF-NfL and 14-3-3 levels improved the accuracy and the number of patients in the confidence group for differentiating FTD from AD. We obtain individual diagnostic probabilities with high precision to address the problem of confidence in the diagnosis. We suggest when MRI, CSF, or the combination are necessary to improve the FTD and AD diagnosis. This algorithm holds promise towards clinical applications as support to clinical findings or in settings with limited access to expert diagnoses.

1. INTRODUCTION

Frontotemporal dementia (FTD) is a clinically, pathologically, and genetically heterogeneous neurodegenerative disorder, which tends to be misdiagnosed with Alzheimer's Disease (AD) (Harris et al., 2015; Koedam et al., 2010; Mendez et al., 2013). Clinically, AD and FTD are inherently different, with the usual overlap occurring primarily between certain rare subtypes of these conditions (Bozeat et al., 2000). However, AD is the most frequent dementia and sometimes is the first choice for many non-specialist clinicians. The overlapping symptoms, especially in the early stages, and the lack of specific, accepted, and available diagnostic biomarkers for FTD subtypes make its diagnosis challenging (Swift et al., 2021). Although prior literature quantifying misdiagnosis between AD and FTD variants is scarce, a study by Falgàs et al. (Falgàs et al., 2019) conducted in a memory clinic showed that up to 19% of clinically suspected early-onset AD cases were ultimately diagnosed as FTD after performing CSF or neuroimaging biomarkers (FDG-PET, amyloid-PET, MRI). These misdiagnoses included both bvFTD, lvPPA, and svPPA cases. Therefore, there is a need to identify tools to help accurately diagnose dementia's underlying etiologies and their subtypes.

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) (Albert et al., 2011; McKhann et al., 2011). 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 (Alcolea et al., 2017; Antonell et al., 2019; McFerrin et al., 2017; Rohrer et al., 2016).

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 (Davatzikos et al., 2008; Du et al., 2007). MRI markers such

as atrophy measures have a neurobiological implication as could play an important role in disease diagnosis and tracking of pathologic progression in AD or FTD and are used as outcome measures in trials of potentially disease-modifying therapies (Frisoni et al., 2010; Prados et al., 2015). 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 (Bocchetta et al., 2021; Borrego-Écija et al., 2021; Canu et al., 2017; Contador et al., 2021; Möller et al., 2015, 2013). The conclusions drawn by prior studies regarding a specific MRI pattern of atrophy for AD were that the regions of the temporal, parietal, and occipital lobes were the ones more affected (Blanc et al., 2015; Möller et al., 2013; Whitwell et al., 2011). On the other hand, atrophy in the temporal and frontal lobes constitutes the specific FTD atrophy pattern (Couto et al., 2013; Möller et al., 2015; Rabinovici et al., 2008). 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 (Abraham et al., 2014; Frizzell et al., 2022; Mateos-Pérez et al., 2018) to support the clinical diagnosis of these two dementias (Bron et al., 2017; Chagué et al., 2021; Klöppel et al., 2008; Möller et al., 2016; Pérez-Millan et al., 2023). Many studies have shown that a support vector machine (SVM) with neuroimaging data differentiates AD or FTD patients from healthy controls (Bisenius et al., 2017; Bron et al., 2021; Cuingnet et al., 2011; Lampe et al., 2023; Magnin et al., 2009; Meyer et al., 2017; Pérez-Millan et al., 2023). In the context of the differential diagnosis of these two dementias, it is known that the clinical symptoms of FTD and AD can display a substantial overlap between them (Mendez, 2006; Wojtas et al., 2012; Zee et al., 2008), suggesting that additional markers may help in the differentiation. In this sense, in recent years, several studies have appeared to assess the use of artificial intelligence, including ML and deep learning, to evaluate, predict, and classify the differential diagnosis of FTD and AD with multiple data, including MRI, neuropsychological test, or biological data (Garcia-Gutierrez et al., 2021; Javeed et al., 2023; Kim et al., 2019; Maito et al., 2023; Moguilner et al., 2022; Nguyen et al., 2023; Pérez-Millan et al., 2023). These studies have presented different approaches and multiple datasets, leading to different predictive capabilities and a wide range of applicability settings. It is still not clear which is the best data to use and

whether ML or deep learning achieves the best accuracy. Thus, there is still a need to improve these algorithms, for example, by obtaining high accuracies for the different subtypes, by finding the best combination of data or by being able to obtain probability disease scores at the subject level. In addition, these previous studies use uniquely MRI data – either unimodal or multimodal MRI – or the combination of MRI data and neuropsychological test data (De Francesco et al., 2023; Moguilner et al., 2022). To our knowledge, no previous studies have combined MRI and CSF data in an ML algorithm for the differential diagnosis of FTD and AD. In previous works, we used statistical approaches to study the contribution of the biochemical markers to the structural changes (Falgàs et al., 2020) and the association between the CTh variability and CSF levels (Pérez-Millan et al., 2024). Here, we explore for the first time the combination of MRI and CSF in a ML algorithm to differentiate FTD and AD.

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, a lumbar puncture for AD markers following the hospital's standard clinical care practice, 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 (Albert et al., 2011; McKhann et al., 2011) supported by the CSF biomarkers profile suggesting underlying AD neuropathology according to National Institute on Aging/Alzheimer's Association

Research Framework 2018 (Jack et al., 2018). 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) (Gorno-Tempini et al., 2011; Rascovsky et al., 2011). Participants with clinical features of AD but with unsupportive CSF profile were excluded from the study, as well as patients with FTD who showed an AD CSF profile. Healthy controls (CTR) had cognitive performance within the normative range (cutoff 1.5 SD from the normative mean, available normative data (Peña-Casanova et al., 2009)) and normal AD CSF biomarkers. We included healthy voluntaries and individuals with subjective memory complaints; all of them performed within the normal range in all the tests of the comprehensive neuropsychological assessment reported in previous studies of the group (Contador et al., 2022; Tort-Merino et al., 2022).

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 (Fischl et al., 2004; Fischl and Dale, 2000). 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 (Desikan et al., 2006; Seidman et al., 1997). 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. Age was added as a feature to ensure that the ML algorithm captures age's true importance or influence on the model. 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 (subcortical volumes normalized by intracranial volume and CTh measures) to z-scores with the training data (all participants) and then applied this conversion to obtain the z-scores for the test dataset. We implemented a calibrated classifier with an SVM as a base estimator to predict the diagnosis of the participants, where the features of the ML algorithm were the previously described values transformed to z-scores. The calibration was set using the *CalibratedClassifierCV* from the Scikit-learn library with a cross-validation of 5 with the train set. As our proposal involves an inner loop in the calibration, to maximise the number of samples included in this procedure, the SVM base estimator hyperparameters were fixed to kernel=rbf, C=1 and gamma=1/(number of features). For each classifier, we fitted a logistic regression model that distributes the classifier's output of the decision function, which predicts a “soft” score for each sample in relation to each group. The model assumes a sigmoid shape and calibrates 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

(we merged svPPA and nfvPPA due to the sample size and named the group as 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 with stratified folds preserving the percentage of samples of each diagnosis to evaluate the performance of the classification. Then, we analysed the importance of each region for the decision of the classification through a permutation feature importance estimation (Breiman, 2001) using the test data of each run with the calibrated SVM algorithm. Feature importance is evaluated as the difference between the baseline metric and the metric obtained after permuting a feature and re-evaluating it. We used the *permutation_importance* from the Scikit-learn library. The higher the weight, the larger the importance of the feature in the classification. Notably, the units of the weights are rather arbitrary. Thus, even if they can be compared across features for a given classification problem, they should not be compared across different scenarios.

We obtained individual probabilities associated with group correspondences as output values for each test data point given by the calibrated SVM. 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 participants 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 (Pedregosa et al., 2011).

2.6 Statistics

We compared the demographic and clinical data among groups using ANOVA tests for continuous variables and Fisher test for discrete variables. Continuous variables were expressed as mean ± standard deviation (SD). The pairwise differences were evaluated using a Benjamini-Hochberg correction. The significance level was set in all the analyses at a corrected p-value < 0.05. Statistical analyses were performed using R version 4.2.1.

3. RESULTS

3.1. Sample demographics

The prospective study includes 491 participants: 215 AD, 103 FTD (56 bvFTD, 24 svPPA, 21 nfvPPA, and 2 PPA), and 173 CTR participants (74% were healthy voluntaries, and 26% were individuals with subjective memory complaints but completely normal cognition and normal AD CSF biomarkers). 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 (p-values corrected by Benjamini-Hochberg), 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, CTR were younger than AD and FTD participants.

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. The train and test were 80% and 20% of the total sample, respectively, in all the analyses. Including sex in the algorithm led to similar results. We got an accuracy of 88 ± 8% (AUC=0.87) when discriminating AD patients from CTR, and 87 ± 4% (AUC=0.85) when determining FTD patients from CTR. When we tried classifying AD vs. FTD patients, the accuracy was 82 ± 6% (AUC=0.77). All AUC and

the F1-score for all comparisons are shown in Supplementary Table 2. Notably, given a rather unbalanced scenario, the results of these metrics indicate that the overall accuracy is not likely to be driven by a single-group classification. Finally, we obtained an accuracy of 77 ± 6% when discriminating between the three groups (AD vs. FTD vs. CTR) (Table 2).

As shown in Figure 1 and in Supplementary Figure 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 (probability of being correctly classified). 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. Figure 2 shows that a higher probability reduced the misprediction. However, a high probability does not ensure a correct classification.

Then, we aimed to study the FTD clinical subtypes separately. Due to limitations in sample size, we merged svPPA and nfvPPA in the same named PPA group. 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 (Tables 3, 4 and 5 in Supplementary Material show the feature importance of all the features). It should be noted that these weights should not be compared between classification settings, as they reflect the absolute change in accuracy associated with each feature. However, the different feature orderings within each comparison offer important insights. 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. In this comparison, age emerged as a medium important feature, being in the sixth place (Table 5 Supplementary Material).

The results of the most crucial regions in the classifications considering bvFTD and PPA participants are shown in Figure 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 improved the results in some of the comparisons.

For comparing AD vs. CTR, CSF data (NfL and 14-3-3) alone was enough and better than MRI data to discriminate between AD and CTR participants (Table 2, Figure 5). Thus, in this comparison the MRI data could not be needed. In the comparison between FTD and CTR, having MRI or CSF data led us to similar results in terms of accuracy (see Table 2 and Figure 5). However, the MRI data is needed for having a high number of CTR with a CTR-probability higher than 0.8. 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 participants with a probability≥ 0.8 (see Table 2 and Figure 5). Thus, in the differential diagnosis of FTD and AD, having MRI and CSF data could bring more reliability to the diagnosis.

4. DISCUSSION

In this study, we implemented a machine learning algorithm that discriminates FTD and AD patients 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 improved the diagnosis classification or the number of patients with high individual probability for the diagnosis in some cases, especially for differentiating FTD from AD.

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 (Bron et al., 2017; Canu et al., 2017; Davatzikos et al., 2008; Dukart et al., 2011; Kim et al., 2019; Li et al., 2021; Lin et al., 2018; Möller et al., 2016; Moore et al., 2019; Pérez-Millan et al., 2023; Salvatore et al., 2015; Wang et al., 2016, 2018). 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 (Basheera and Sai Ram, 2019; Dashtipour et al., 2022; Frizzell et al., 2022; Mateos-Pérez et al., 2018; McCarthy et al., 2018). We have differentiated FTD expressions (bvFTD and PPA) against AD or CTR, outperforming previously published works (Canu et al., 2017; Lampe et al., 2023; Möller et al., 2016; Wang et al., 2016). 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. (Kim et al., 2019), probably due to differences in the algorithm. They used hierarchical classification with cortical atrophy, but in the CTh measures, they applied noise removal. Then, the hierarchical classification algorithm was constructed by applying linear discriminant analysis (LDA) in combination with principal component analysis (PCA) to the cortical thickness data. The results of their proposed algorithm using the entire hierarchical tree showed an overall 75.8% accuracy, but looking in detail at step 3 of the algorithm, which focused on the FTD group (bvFTD versus PPA), the algorithm presented 86.9% accuracy, which is higher than the obtained in our algorithm. Thus, our algorithm accurately distinguishes AD, FTD, and CTR using MRI data. However, the classification accuracy between bvFTD and PPA is lower than in some previous works.

Other studies using multimodal information also reported high classification accuracy combining data from different imaging modalities or other biological and clinical measures (Bron et al., 2021; Chen et al., 2017; Dashtipour et al., 2022; Dukart et al., 2011). Even so, in some cases, our scores with only structural MRI data showed better accuracy (Bouts et al., 2018; Bron et al., 2017; Dyrba et al., 2015; Wang et al., 2016). 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 reduce the number of clinical tests usually required to take these patients. This approach identifies participants who need additional clinical tests while sparing others from unnecessary examinations that wouldn't yield 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 in some cases, especially in the differential diagnosis of FTD and AD. In this case, the combination of the data of these two modalities increased the accuracy and the number of participants with a reliable diagnostic. In the case of discriminating AD and CTR, the CSF data presented the best performance, so the MRI data may be unnecessary or redundant. Finally, in the case of classifying FTD versus CTR, MRI and CSF data showed similar results, though the MRI data contributes to the reliability of the CTR diagnosis, so in this case, we should recommend using MRI data. The comparisons between CTR and dementia patients (AD or FTD) are a gold standard for studying the algorithm and for research proposes. For clinical utility, it could be helpful the differential diagnosis of FTD vs AD or their subtypes. We excluded Aβ42, t-tau, and p-tau from our algorithm to

avoid circularity. These biomarkers were used to confirm AD when positive or FTD when negative. In contrast, NfL and 14-3-3 were not used to determine the label feature. Although NfL and 14-3-3 correlate with Aβ42, t-tau, and p-tau, so does hippocampus volume. Therefore, this correlation with current diagnostic biomarkers of AD does not bias the algorithm, as these biomarkers were not used to create the feature label (diagnosis), which is what the ML algorithm aims to predict. The participants with available CSF NfL and 14-3-3 biomarkers were not those that the clinician was already in doubt based on the MRI, as the diagnostic criteria for these patients are supported by a CSF profile.

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. It is important to note that our results are merely an example of interpreting these probabilities in a context where participants are distributed along a spectrum between two groups, assuming that clinical conditions (or labels) are known. However, the interpretation may differ in various clinical contexts, particularly when there are more erroneous or unknown diagnoses. We suggest evaluating these algorithms with different datasets to explore various applications and scenarios.

Furthermore, 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. While the mathematical approaches used by ML algorithms differ from traditional statistical group comparisons—mainly because they account for interactions and hidden relationships between features—several key conclusions can be drawn from these patterns. Notably, our results reflect the typical AD neuroimaging pattern, found mainly in the early stages of the disease, where most patients of our sample are. The involvement of the hippocampus and subcortical areas aligns with the neuropathological continuum demonstrated by postmortem studies, such as Braak et al., 2011, showing that

tau pathology affects these regions in the earliest stages of the disease. Hippocampal atrophy is considered a hallmark feature of AD, directly related to memory impairment (Braak et al., 2011). The amygdala, involved in emotional processing and memory, also exhibits early atrophy, contributing to behavioral symptoms (Johnson et al., 2024; Punzi et al., 2024). The putamen, part of the basal ganglia, may be affected by tau pathology spreading from the hippocampus and other subcortical areas in early disease stages, impacting cognitive functions (Yang et al., 2024). By contrast, when differentiating FTD from CTR, the cortical regions were the most important, especially the CTh of occipital, parietal, and frontal regions. Accordingly, GM volumes of subcortical areas could help to identify AD patients, while CTh could be the key to identifying FTD participants. This aligns with findings obtained using more classical analysis methods (Avants et al., 2010; Contador et al., 2021; Dickerson et al., 2001; Frisoni et al., 2010; Gil-Navarro et al., 2013; Gordon et al., 2016; Hodges and Patterson, 2007; Jack et al., 2000). Regarding FTD variants, the frontal brain regions emerged for the bvFTD, while the hippocampus and temporal regions were the most important in PPA, as previously reported (Avants et al., 2010). These results align with previously described brain changes in bvFTD and PPA. bvFTD is characterized by primary degeneration affecting the frontal lobes, causing behavioral symptoms such as disinhibition, apathy, and impulsivity (Rascovsky et al., 2011). In contrast, PPA is marked by degeneration in the language areas, particularly the left temporal lobe, manifesting as impaired language skills (Gorno-Tempini et al., 2011).

Overall, our study has several strengths. First, we provide examples of scenarios where MRI or CSF data or their combination contributes to a more reliable AD or FTD diagnosis. In addition, the overall good performance of our algorithm offers opportunities for future clinical applications after validating this type of ML algorithm in other datasets, including AD and FTD or in other neurodegenerative dementias. We suggest its potential implementation could be particularly impactful in locations with limited access to expert opinion or incomplete biomarker profiles. The individual probabilities generated could significantly advance personalized medicine. The use of probabilistic algorithms like the one proposed here may be a first step towards developing new methods that consider disease severity or therapy response, helping to identify potential candidates for new drugs or additional diagnostic tests, as we could address their current diagnostic

confidence. Further studies may aim to validate these outcomes regarding individual clinical outcomes at a longitudinal level, such as response to therapy or disease trajectory.

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. Also, it is focused on only AD or FTD, so applying it to different clinical contexts could affect the results, particularly when there are more erroneous, unknown diagnoses or other neurodegenerative dementias. 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. Also, it is known that svPPA and nfvPPA present different characteristic MRI patterns, and CSF levels (NfL and 14-3-3) are both increased (Abu-Rumeileh et al., 2018; Ljubenkov et al., 2018; Seelaar et al., 2011). Thus, the feature importance of the MRI features may be influenced by the fact that we study PPA participants together. 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. Future studies may replicate our CSF results in large, multicentric cohorts to support the findings. As is common with many ML algorithms using different data sources, the availability of NfL and 14-3-3 could introduce a bias towards a subset of participants. In our study, as these biomarkers are not obtained routinely, the availability of such data depended on whether we had leftover CSF or if the levels had been determined previously. We did not perform additional testing on the new participants, but this criterion was the same for both groups, minimizing the risk of a bias.

5. CONCLUSION

In conclusion, the proposed diagnosis algorithm has shown high accuracy classification scores with structural MRI data to discriminate AD, FTD, and CTR. Furthermore, we propose guidelines suggesting when MRI, CSF, or the combination are necessary or improve the diagnosis of FTD and AD patients. This approach also provided individual

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.

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.

Availability of data

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

Conflict of interest statement

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-Écija is 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

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics 8.

Abu-Rumeileh, S., Capellari, S., Stanzani-Maserati, M., Polischi, B., Martinelli, P., Caroppo, P., Ladogana, A., Parchi, P., 2018. The CSF neurofilament light signature in rapidly progressive neurodegenerative dementias. Alz Res Therapy 10, 3. https://doi.org/10.1186/s13195-017-0331-1

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 2011. 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 7, 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

Alcolea, D., Vilaplana, E., Suárez-Calvet, M., Illán-Gala, I., Blesa, R., Clarimón, J., Lladó, A., Sánchez-Valle, R., Molinuevo, J.L., García-Ribas, G., Compta, Y., Martí, M.J., Piñol-Ripoll, G., Amer-Ferrer, G., Noguera, A., García-Martín, A., Fortea, J., Lleó, A., 2017. CSF sAPPβ, YKL-40, and neurofilament light in frontotemporal lobar degeneration. Neurology 89, 178–188. https://doi.org/10.1212/WNL.0000000000004088

Antonell, A., Tort-Merino, A., Ríos, J., Balasa, M., Borrego-Écija, S., Auge, J.M., Muñoz-García, C., Bosch, B., Falgàs, N., Rami, L., Ramos-Campoy, O., Blennow, K., Zetterberg, H., Molinuevo, J.L., Lladó, A., Sánchez-Valle, R., 2019. Synaptic, axonal damage and inflammatory cerebrospinal fluid biomarkers in neurodegenerative dementias. Alzheimer's & Dementia. https://doi.org/10.1016/j.jalz.2019.09.001

Avants, B.B., Cook, P.A., Ungar, L., Gee, J.C., Grossman, M., 2010. Dementia induces correlated reductions in white matter integrity and cortical thickness: A multivariate neuroimaging study with sparse canonical correlation analysis. NeuroImage 50, 1004–1016. https://doi.org/10.1016/j.neuroimage.2010.01.041

Basheera, S., Sai Ram, M.S., 2019. 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 5, 974–986. https://doi.org/10.1016/j.trci.2019.10.001

Bisenius, S., Mueller, K., Diehl-Schmid, J., Fassbender, K., Grimmer, T., Jessen, F., Kassubek, J., Kornhuber, J., Landwehrmeyer, B., Ludolph, A., Schneider, A., Anderl-Straub, S., Stuke, K., Danek, A., Otto, M., Schroeter, M.L., 2017. Predicting primary progressive aphasias with support vector machine approaches in structural MRI data. NeuroImage: Clinical 14, 334–343. https://doi.org/10.1016/j.nicl.2017.02.003

Blanc, F., Colloby, S.J., Philippi, N., Pétigny, X. de, Jung, B., Demuynck, C., Phillipps, C., Anthony, P., Thomas, A., Bing, F., Lamy, J., Martin-Hunyadi, C., O'Brien, J.T., Cretin, B., McKeith, I., Armspach, J.-P., Taylor, J.-P., 2015. Cortical Thickness in Dementia with Lewy Bodies and Alzheimer's Disease: A Comparison of Prodromal and Dementia Stages. PLOS ONE 10, e0127396. https://doi.org/10.1371/journal.pone.0127396

Bocchetta, M., Todd, E.G., Peakman, G., Cash, D.M., Convery, R.S., Russell, L.L., Thomas, D.L., Iglesias, J.E., Van Swieten, J.C., Jiskoot, L.C., Seelaar, H., Borroni, B., Galimberti, D., Sanchez-Valle, R., Laforce, R., Moreno, F., Synofzik, M., Graff, C., Masellis, M., Tartaglia, M.C., Rowe, J.B., Vandenberghe, R., Finger, E., Tagliavini, F., De Mendonça, A., Santana, I., Butler, C.R., Ducharme, S., Gerhard, A., Danek, A., Levin, J., Otto, M., Sorbi, S., Le Ber, I., Pasquier, F., Rohrer, J.D., 2021. Differential early subcortical involvement in genetic FTD within the GENFI cohort. NeuroImage: Clinical 30, 102646. https://doi.org/10.1016/j.nicl.2021.102646

Borrego-Écija, S., Sala-Llonch, R., van Swieten, J., Borroni, B., Moreno, F., Masellis, M., Tartaglia, C., Graff, C., Galimberti, D., Laforce, R., Rowe, J.B., Finger, E., Vandenberghe, R., Tagliavini, F., de Mendonça, A., Santana, I., Synofzik, M., Ducharme, S., Levin, J., Danek, A., Gerhard, A., Otto, M., Butler, C., Frisoni, G., Sorbi, S., Heller, C., Bocchetta, M., Cash, D.M., Convery, R.S., Moore, K.M., Rohrer, J.D., Sanchez-Valle, R., Rossor, M.N., Fox, N.C., Woollacott, I.O.C., Shafei, R., Greaves, C., Neason, M., Guerreiro, R., Bras, J., Thomas, D.L., Nicholas, J., Mead, S., Meeter, L., Panman, J., Papma, J., van Minkelen, R., Pijnenburg, Y., Indakoetxea, B., Gabilondo, A., TaintaMD, M., de Arriba, M., Gorostidi, A., Zulaica, M., Villanua, J., Diaz, Z., Olives, J., Lladó, A., Balasa, M., Antonell, A., Bargallo, N., Premi, E., Cosseddu, M., Gazzina, S., Padovani, A., Gasparotti, R., Archetti, S., Black, S., Mitchell, S., Rogaeva, E., Freedman, M., Keren, R., Tang-Wai, D., Öijerstedt, L., Andersson, C., Jelic, V., Thonberg, H., Arighi, A., Fenoglio, C., Scarpini MD, E., Fumagalli, G., Cope, T., Timberlake, C., Rittman, T., Shoesmith, C., Bartha, R., Rademakers, R., Wilke, C., Bender, B., Bruffaerts, R., Vandamme, P., Vandenbulcke, M., Maruta, C., Ferreira, C.B., Miltenberger, G., Verdelho, A., Afonso, S., Taipa, R., Caroppo, P., Di Fede, G., Giaccone, G., Prioni, S., Redaelli, V., Rossi, G., Tiraboschi, P., Duro, D., Rosario Almeida, M., Castelo-Branco, M., João Leitão, M., Tabuas-Pereira, M., Santiago, B., Gauthier, S., Rosa-Neto, P., Veldsman, M., Flanagan, T., Prix, C., Hoegen, T., Wlasich, E., Loosli, S., Schonecker, S., Semler, E., Anderl-Straub, S., 2021. Disease-related cortical thinning in presymptomatic granulin mutation carriers. NeuroImage: Clinical 29. https://doi.org/10.1016/j.nicl.2020.102540

Bouts, M.J.R.J., Möller, C., Hafkemeijer, A., van Swieten, J.C., Dopper, E., van der Flier, W.M., Vrenken, H., Wink, A.M., Pijnenburg, Y.A.L., Scheltens, P.,

Barkhof, F., Schouten, T.M., de Vos, F., Feis, R.A., van der Grond, J., de Rooij, M., Rombouts, S.A.R.B., 2018. 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 62, 1827–1839. https://doi.org/10.3233/JAD-170893

Bozeat, S., Gregory, C.A., Ralph, M.A., Hodges, J.R., 2000. Which neuropsychiatric and behavioural features distinguish frontal and temporal variants of frontotemporal dementia from Alzheimer's disease? J Neurol Neurosurg Psychiatry 69, 178–186. https://doi.org/10.1136/jnnp.69.2.178

Braak, H., Thal, D.R., Ghebremedhin, E., Del Tredici, K., 2011. Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. J Neuropathol Exp Neurol 70, 960–969. https://doi.org/10.1097/NEN.0b013e318232a379

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Bron, E.E., Klein, S., Papma, J.M., Jiskoot, L.C., Venkatraghavan, V., Linders, J., Aalten, P., De Deyn, P.P., Biessels, G.J., Claassen, J.A.H.R., Middelkoop, H.A.M., Smits, M., Niessen, W.J., van Swieten, J.C., van der Flier, W.M., Ramakers, I.H.G.B., van der Lugt, A., 2021. Cross-cohort generalizability of deep and conventional machine learning for MRI-based diagnosis and prediction of Alzheimer's disease. NeuroImage: Clinical 31, 102712. https://doi.org/10.1016/j.nicl.2021.102712

Bron, E.E., Smits, M., Papma, J.M., Steketee, R.M.E., Meijboom, R., de Groot, M., van Swieten, J.C., Niessen, W.J., Klein, S., 2017. Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using structural and advanced MRI. European Radiology 27, 3372–3382. https://doi.org/10.1007/s00330-016-4691-x

Canu, E., Agosta, F., Mandic-Stojmenovic, G., Stojković, T., Stefanova, E., Inuggi, A., Imperiale, F., Copetti, M., Kostic, V.S., Filippi, M., 2017. Multiparametric MRI to distinguish early onset Alzheimer's disease and behavioural variant of frontotemporal dementia. NeuroImage: Clinical 15, 428–438. https://doi.org/10.1016/j.nicl.2017.05.018

Chagué, P., Marro, B., Fadili, S., Houot, M., Morin, A., Samper-González, J., Beunon, P., Arrivé, L., Dormont, D., Dubois, B., Teichmann, M., Epelbaum, S., Colliot, O., 2021. Radiological classification of dementia from anatomical MRI assisted by machine learning-derived maps. Journal of Neuroradiology 48, 412–418. https://doi.org/10.1016/j.neurad.2020.04.004

Chen, Y., Sha, M., Zhao, X., Ma, J., Ni, H., Gao, W., Ming, D., 2017. Automated detection of pathologic white matter alterations in Alzheimer's disease using combined diffusivity and kurtosis method. Psychiatry Research: Neuroimaging 264, 35–45. https://doi.org/10.1016/j.pscychresns.2017.04.004

Contador, J., Pérez-Millan, A., Guillén, N., Sarto, J., Tort-Merino, A., Balasa, M., Falgàs, N., Castellví, M., Borrego-Écija, S., Juncà-Parella, J., Bosch, B., Fernández-Villullas, G., Ramos-Campoy, O., Antonell, A., Bargalló, N., Sanchez-Valle, R., Sala-Llonch, R., Lladó, A., 2022. Sex differences in early-onset Alzheimer's disease. European Journal of Neurology 29, 3623–3632. https://doi.org/10.1111/ene.15531

Contador, J., Pérez-Millán, A., Tort-Merino, A., Balasa, M., Falgàs, N., Olives, J., Castellví, M., Borrego-Écija, S., Bosch, B., Fernández-Villullas, G., Ramos-

715 Campoy, O., Antonell, A., Bargalló, N., Sanchez-Valle, R., Sala-Llonch, R.,
716 Lladó, A., 2021. Longitudinal brain atrophy and CSF biomarkers in early-onset
717 Alzheimer's disease. NeuroImage: Clinical 32, 102804.
718 https://doi.org/10.1016/j.nicl.2021.102804
719 Couto, B., Manes, F., Montañés, P., Matallana, D., Reyes, P., Velasquez, M., Yoris, A.,
720 Baez, S., Ibáñez, A., 2013. Structural neuroimaging of social cognition in
721 progressive non-fluent aphasia and behavioral variant of frontotemporal
722 dementia. Front Hum Neurosci 7, 467.
723 https://doi.org/10.3389/fnhum.2013.00467
724 Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O.,
725 Chupin, M., Benali, H., Colliot, O., 2011. Automatic classification of patients
726 with Alzheimer's disease from structural MRI: A comparison of ten methods
727 using the ADNI database. NeuroImage, Multivariate Decoding and Brain
728 Reading 56, 766–781. https://doi.org/10.1016/j.neuroimage.2010.06.013
729 Dashtipour, K., Taylor, W., Ansari, S., Zahid, A., Gogate, M., Ahmad, J., Assaleh, K.,
730 Arshad, K., Imran, M.A., Abbasi, Q., 2022. Detecting Alzheimer's Disease
731 Using Machine Learning Methods, in: Ur Rehman, M., Zoha, A. (Eds.), Body
732 Area Networks. Smart IoT and Big Data for Intelligent Health Management,
733 Lecture Notes of the Institute for Computer Sciences, Social Informatics and
734 Telecommunications Engineering. Springer International Publishing, Cham, pp.
735 89–100. https://doi.org/10.1007/978-3-030-95593-9_8
736 Davatzikos, C., Resnick, S.M., Wu, X., Parmpi, P., Clark, C.M., 2008. Individual
737 patient diagnosis of AD and FTD via high-dimensional pattern classification of
738 MRI. NeuroImage 41, 1220–1227.
739 https://doi.org/10.1016/j.neuroimage.2008.03.050
740 De Francesco, S., Crema, C., Archetti, D., Muscio, C., Reid, R.I., Nigri, A., Bruzzone,
741 M.G., Tagliavini, F., Lodi, R., D'Angelo, E., Boeve, B., Kantarci, K., Firbank,
742 M., Taylor, J.-P., Tiraboschi, P., Redolfi, A., 2023. Differential diagnosis of
743 neurodegenerative dementias with the explainable MRI based machine learning
744 algorithm MUQUBIA. Sci Rep 13, 17355. https://doi.org/10.1038/s41598-023-
745 43706-6
746 Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D.,
747 Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany,
748 R.J., 2006. An automated labeling system for subdividing the human cerebral
749 cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–
750 980. https://doi.org/10.1016/j.neuroimage.2006.01.021
751 Dickerson, B.C., Goncharova, I., Sullivan, M.P., Forchetti, C., Wilson, R.S., Bennett,
752 D.A., Beckett, L.A., DeToledo-Morrell, L., 2001. MRI-derived entorhinal and
753 hippocampal atrophy in incipient and very mild Alzheimer's disease.
754 Neurobiology of Aging 22, 747–754. https://doi.org/10.1016/S0197-
755 4580(01)00271-8
756 Du, A.T., Schuff, N., Kramer, J.H., Rosen, H.J., Gorno-Tempini, M.L., Rankin, K.,
757 Miller, B.L., Weiner, M.W., 2007. Different regional patterns of cortical
758 thinning in Alzheimer's disease and frontotemporal dementia. Brain 130, 1159–
759 1166. https://doi.org/10.1093/brain/awm016
760 Dukart, J., Mueller, K., Horstmann, A., Barthel, H., Möller, H.E., Villringer, A., Sabri,
761 O., Schroeter, M.L., 2011. Combined Evaluation of FDG-PET and MRI
762 Improves Detection and Differentiation of Dementia. PLOS ONE 6, e18111.
763 https://doi.org/10.1371/journal.pone.0018111
764 Dyrba, M., Grothe, M., Kirste, T., Teipel, S.J., 2015. Multimodal analysis of functional

765 and structural disconnection in Alzheimer's disease using multiple kernel SVM.
766 Human Brain Mapping 36, 2118–2131. https://doi.org/10.1002/hbm.22759
767 Falgàs, N., Ruiz-Peris, M., Pérez-Millan, A., Sala-Llonch, R., Antonell, A., Balasa, M.,
768 Borrego-Écija, S., Ramos-Campoy, O., Augé, J.M., Castellví, M., Tort-Merino,
769 A., Olives, J., Fernández-Villullas, G., Blennow, K., Zetterberg, H., Bargalló,
770 N., Lladó, A., Sánchez-Valle, R., 2020. Contribution of CSF biomarkers to
771 early-onset Alzheimer's disease and frontotemporal dementia neuroimaging
772 signatures. Human Brain Mapping hbm.24925.
773 https://doi.org/10.1002/hbm.24925
774 Falgàs, N., Tort-Merino, A., Balasa, M., Borrego-Écija, S., Castellví, M., Olives, J.,
775 Bosch, B., Férnandez-Villullas, G., Antonell, A., Augé, J.M., Lomeña, F.,
776 Perissinotti, A., Bargalló, N., Sánchez-Valle, R., Lladó, A., 2019. Clinical
777 applicability of diagnostic biomarkers in early-onset cognitive impairment.
778 European Journal of Neurology 26, 1098–1104.
779 https://doi.org/10.1111/ene.13945
780 Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex
781 from magnetic resonance images. Proceedings of the National Academy of
782 Sciences of the United States of America 97, 11050–11055.
783 https://doi.org/10.1073/pnas.200033797
784 Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H.,
785 Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N.,
786 Rosen, B., Dale, A.M., 2004. Automatically Parcellating the Human Cerebral
787 Cortex. Cerebral Cortex 14, 11–22. https://doi.org/10.1093/cercor/bhg087
788 Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., Thompson, P.M., 2010. The clinical
789 use of structural MRI in Alzheimer disease. Nat Rev Neurol 6, 67–77.
790 https://doi.org/10.1038/nrneurol.2009.215
791 Frizzell, T.O., Glashutter, M., Liu, C.C., Zeng, A., Pan, D., Hajra, S.G., D'Arcy,
792 R.C.N., Song, X., 2022. Artificial intelligence in brain MRI analysis of
793 Alzheimer's disease over the past 12 years: A systematic review. Ageing
794 Research Reviews 77, 101614. https://doi.org/10.1016/j.arr.2022.101614
795 Garcia-Gutierrez, F., Delgado-Alvarez, A., Delgado-Alonso, C., Díaz-Álvarez, J., Pytel,
796 V., Valles-Salgado, M., Gil, M.J., Hernández-Lorenzo, L., Matías-Guiu, J.,
797 Ayala, J.L., Matias-Guiu, J.A., 2021. Diagnosis of Alzheimer's disease and
798 behavioural variant frontotemporal dementia with machine learning-aided
799 neuropsychological assessment using feature engineering and genetic
800 algorithms. Int J Geriatr Psychiatry 37. https://doi.org/10.1002/gps.5667
801 Gil-Navarro, S., Lladó, A., Rami, L., Castellví, M., Bosch, B., Bargalló, N., Lomeña, F.,
802 Reñé, R., Montagut, N., Antonell, A., Molinuevo, J.L., Sánchez-Valle, R., 2013.
803 Neuroimaging and Biochemical Markers in the Three Variants of Primary
804 Progressive Aphasia. DEM 35, 106–117. https://doi.org/10.1159/000346289
805 Gordon, E., Rohrer, J.D., Fox, N.C., 2016. Advances in neuroimaging in frontotemporal
806 dementia. Journal of Neurochemistry 138, 193–210.
807 https://doi.org/10.1111/jnc.13656
808 Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa,
809 S.F., Ogar, J.M., Rohrer, J.D., Black, S., Boeve, B.F., Manes, F., Dronkers, N.F.,
810 Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S.,
811 Hodges, J.R., Mesulam, M.M., Grossman, M., Rascovsky, P.K., Patterson, K.,
812 Miller, B.L., Knopman, D.S., Hodges, J.R., Mesulam, M.M., Grossman, M.,
813 Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R.,
814 Mesulam, M.M., Grossman, M., 2011. Classification of primary progressive

aphasia and its variants. Neurology 76, 1006–14. https://doi.org/10.1212/WNL.0b013e31821103e6
Harris, J.M., Thompson, J.C., Gall, C., Richardson, A.M.T., Neary, D., du Plessis, D., Pal, P., Mann, D.M.A., Snowden, J.S., Jones, M., 2015. Do NIA-AA criteria distinguish Alzheimer's disease from frontotemporal dementia? Alzheimer's & Dementia 11, 207–215. https://doi.org/10.1016/j.jalz.2014.04.516
Hodges, J.R., Patterson, K., 2007. Semantic dementia: a unique clinicopathological syndrome. The Lancet Neurology 6, 1004–1014. https://doi.org/10.1016/S1474-4422(07)70266-1
Jack, C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B., Holtzman, D.M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J.L., Montine, T., Phelps, C., Rankin, K.P., Rowe, C.C., Scheltens, P., Siemers, E., Snyder, H.M., Sperling, R., Elliott, C., Masliah, E., Ryan, L., Silverberg, N., 2018. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's and Dementia 14, 535–562. https://doi.org/10.1016/j.jalz.2018.02.018
Jack, C.R., Petersen, R.C., Xu, Y., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Tangalos, E.G., Kokmen, E., 2000. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology 55, 484–489. https://doi.org/10.1212/wnl.55.4.484
Javeed, A., Dallora, A.L., Berglund, J.S., Ali, A., Ali, L., Anderberg, P., 2023. Machine Learning for Dementia Prediction: A Systematic Review and Future Research Directions. J Med Syst 47, 17. https://doi.org/10.1007/s10916-023-01906-7
Johnson, A.S., Ziaggi, G., Smith, A.C., Houlihan, H., Heuer, L.B., Guzmán, D.S., Okafor, A., Huey, E.D., Talmasov, D., Provenzano, F., Kreisl, W.C., Lao, P.J., Alzheimer's Disease Neuroimaging Initiative, 2024. Psychotic symptoms are associated with elevated tau PET signal in the amygdala independent of Alzheimer's disease clinical severity and amyloid burden. medRxiv 2024.01.12.24301221. https://doi.org/10.1101/2024.01.12.24301221
Kim, J.P., Kim, J., Park, Y.H., Park, S.B., Lee, J.S., Yoo, S., Kim, E.-J., Kim, H.J., Na, D.L., Brown, J.A., Lockhart, S.N., Seo, S.W., Seong, J.-K., 2019. Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. NeuroImage: Clinical 23, 101811. https://doi.org/10.1016/j.nicl.2019.101811
Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Automatic classification of MR scans in Alzheimer's disease. Brain 131, 681–689. https://doi.org/10.1093/brain/awm319
Koedam, E.L.G.E., Lauffer, V., van der Vlies, A.E., van der Flier, W.M., Scheltens, P., Pijnenburg, Y.A.L., 2010. Early-Versus Late-Onset Alzheimer's Disease: More than Age Alone. Journal of Alzheimer's Disease 19, 1401–1408. https://doi.org/10.3233/JAD-2010-1337
Lampe, L., Huppertz, H.-J., Anderl-Straub, S., Albrecht, F., Ballarini, T., Bisenius, S., Mueller, K., Niehaus, S., Fassbender, K., Fliessbach, K., Jahn, H., Kornhuber, J., Lauer, M., Prudlo, J., Schneider, A., Synofzik, M., Kassubek, J., Danek, A., Villringer, A., Diehl-Schmid, J., Otto, M., Schroeter, M.L., FTLD Consortium Germany, 2023. Multiclass prediction of different dementia syndromes based on multi-centric volumetric MRI imaging. Neuroimage Clin 37, 103320. https://doi.org/10.1016/j.nicl.2023.103320
Li, B., Jang, I., Riphagen, J., Almaktoum, R., Yochim, K.M., Ances, B.M.,

Bookheimer, S.Y., Salat, D.H., 2021. Identifying individuals with Alzheimer's disease-like brains based on structural imaging in the Human Connectome Project Aging cohort. Human Brain Mapping 42, 5535–5546. https://doi.org/10.1002/hbm.25626

Lin, W., Tong, T., Gao, Q., Guo, D., Du, X., Yang, Y., Guo, G., Xiao, M., Du, M., Qu, X., The Alzheimer's Disease Neuroimaging Initiative, 2018. Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease Prediction From Mild Cognitive Impairment. Frontiers in Neuroscience 12.

Ljubenkov, P.A., Staffaroni, A.M., Rojas, J.C., Allen, I.E., Wang, P., Heuer, H., Karydas, A., Kornak, J., Cobigo, Y., Seeley, W.W., Grinberg, L.T., Spina, S., Fagan, A.M., Jerome, G., Knopman, D., Boeve, B.F., Dickerson, B.C., Kramer, J., Miller, B., Boxer, A.L., Rosen, H.J., 2018. Cerebrospinal fluid biomarkers predict frontotemporal dementia trajectory. Annals of Clinical and Translational Neurology 5, 1250–1263. https://doi.org/10.1002/acn3.643

Magnin, B., Mesrob, L., Kinkingnéhun, S., Pélégrini-Issac, M., Colliot, O., Sarazin, M., Dubois, B., Lehéricy, S., Benali, H., 2009. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology 51, 73–83. https://doi.org/10.1007/s00234-008-0463-x

Maito, M.A., Santamaría-García, H., Moguilner, S., Possin, K.L., Godoy, M.E., Avila-Funes, J.A., Behrens, M.I., Brusco, I.L., Bruno, M.A., Cardona, J.F., Custodio, N., García, A.M., Javandel, S., Lopera, F., Matallana, D.L., Miller, B., Oliveira, M.O. de, Pina-Escudero, S.D., Slachevsky, A., Ortiz, A.L.S., Takada, L.T., Tagliazuchi, E., Valcour, V., Yokoyama, J.S., Ibañez, A., 2023. Classification of Alzheimer's disease and frontotemporal dementia using routine clinical and cognitive measures across multicentric underrepresented samples: a cross sectional observational study. The Lancet Regional Health – Americas 17. https://doi.org/10.1016/j.lana.2022.100387

Mateos-Pérez, J.M., Dadar, M., Lacalle-Aurioles, M., Iturria-Medina, Y., Zeighami, Y., Evans, A.C., 2018. Structural neuroimaging as clinical predictor: A review of machine learning applications. NeuroImage: Clinical 20, 506–522. https://doi.org/10.1016/j.nicl.2018.08.019

McCarthy, J., Collins, D.L., Ducharme, S., 2018. Morphometric MRI as a diagnostic biomarker of frontotemporal dementia: A systematic review to determine clinical applicability. NeuroImage: Clinical 20, 685–696. https://doi.org/10.1016/j.nicl.2018.08.028

McFerrin, M.B., Chi, X., Cutter, G., Yacoubian, T.A., 2017. Dysregulation of 14-3-3 proteins in neurodegenerative diseases with Lewy body or Alzheimer pathology. Ann Clin Transl Neurol 4, 466–477. https://doi.org/10.1002/acn3.421

McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., Phelps, C.H., 2011. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's and Dementia 7, 263–269. https://doi.org/10.1016/j.jalz.2011.03.005

Mendez, M.F., 2006. The accurate diagnosis of early-onset dementia. International Journal of Psychiatry in Medicine 36, 401–412. https://doi.org/10.2190/Q6J4-R143-P630-KW41

Mendez, M.F., Joshi, A., Tassniyom, K., Teng, E., Shapira, J.S., 2013.

915 Clinicopathologic differences among patients with behavioral variant
916 frontotemporal dementia. Neurology 80, 561–568.
917 https://doi.org/10.1212/WNL.0b013e3182815547
918 Meyer, S., Mueller, K., Stuke, K., Bisenius, S., Diehl-Schmid, J., Jessen, F., Kassubek,
919 J., Kornhuber, J., Ludolph, A.C., Prudlo, J., Schneider, A., Schuemberg, K.,
920 Yakushev, I., Otto, M., Schroeter, M.L., 2017. Predicting behavioral variant
921 frontotemporal dementia with pattern classification in multi-center structural
922 MRI data. NeuroImage: Clinical 14, 656–662.
923 https://doi.org/10.1016/j.nicl.2017.02.001
924 Moguilner, S., Birba, A., Fittipaldi, S., Gonzalez-Campo, C., Tagliazucchi, E., Reyes,
925 P., Matallana, D., Parra, M.A., Slachevsky, A., Farías, G., Cruzat, J., García, A.,
926 Eyre, H.A., Joie, R.L., Rabinovici, G., Whelan, R., Ibáñez, A., 2022. Multi-
927 feature computational framework for combined signatures of dementia in
928 underrepresented settings. J. Neural Eng. 19, 046048.
929 https://doi.org/10.1088/1741-2552/ac87d0
930 Möller, C., Hafkemeijer, A., Pijnenburg, Y.A.L., Rombouts, S.A.R.B., van der Grond,
931 J., Dopper, E., van Swieten, J., Versteeg, A., Pouwels, P.J.W., Barkhof, F.,
932 Scheltens, P., Vrenken, H., van der Flier, W.M., 2015. Joint assessment of white
933 matter integrity, cortical and subcortical atrophy to distinguish AD from
934 behavioral variant FTD: A two-center study. NeuroImage: Clinical 9, 418–429.
935 https://doi.org/10.1016/j.nicl.2015.08.022
936 Möller, C., Pijnenburg, Y.A.L., Van Der Flier, W.M., Versteeg, A., Tijms, B., De
937 Munck, J.C., Hafkemeijer, A., Rombouts, S.A.R.B., Van Der Grond, J., Van
938 Swieten, J., Dopper, E., Scheltens, P., Barkhof, F., Vrenken, H., Wink, A.M.,
939 2016. Alzheimer disease and behavioral variant frontotemporal dementia:
940 Automatic classification based on cortical atrophy for single-subject diagnosis.
941 Radiology 279, 838–848. https://doi.org/10.1148/radiol.2015150220
942 Möller, C., Vrenken, H., Jiskoot, L., Versteeg, A., Barkhof, F., Scheltens, P., van der
943 Flier, W.M., 2013. Different patterns of gray matter atrophy in early- and late-
944 onset Alzheimer's disease. Neurobiology of Aging 34, 2014–2022.
945 https://doi.org/10.1016/j.neurobiolaging.2013.02.013
946 Moore, P.J., Lyons, T.J., Gallacher, J., Initiative, for the A.D.N., 2019. Random forest
947 prediction of Alzheimer's disease using pairwise selection from time series data.
948 PLOS ONE 14, e0211558. https://doi.org/10.1371/journal.pone.0211558
949 Nguyen, H.-D., Clément, M., Planche, V., Mansencal, B., Coupé, P., 2023. Deep
950 grading for MRI-based differential diagnosis of Alzheimer's disease and
951 Frontotemporal dementia. Artificial Intelligence in Medicine 144, 102636.
952 https://doi.org/10.1016/j.artmed.2023.102636
953 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.,
954 Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos,
955 A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn:
956 Machine Learning in Python. Journal of Machine Learning Research 12, 2825–
957 2830.
958 Peña-Casanova, J., Quiñones-Ubeda, S., Gramunt-Fombuena, N., Aguilar, M., Casas,
959 L., Molinuevo, J.L., Robles, A., Rodríguez, D., Barquero, M.S., Antúnez, C.,
960 Martínez-Parra, C., Frank-García, A., Fernández, M., Molano, A., Alfonso, V.,
961 Sol, J.M., Blesa, R., NEURONORMA Study Team, 2009. Spanish Multicenter
962 Normative Studies (NEURONORMA Project): norms for Boston naming test
963 and token test. Arch Clin Neuropsychol 24, 343–354.
964 https://doi.org/10.1093/arclin/acp039

965 Pérez-Millan, A., Borrego-Écija, S., Falgàs, N., Juncà-Parella, J., Bosch, B., Tort-
966 Merino, A., Antonell, A., Bargalló, N., Rami, L., Balasa, M., Lladó, A., Sala-
967 Llonch, R., Sánchez-Valle, R., 2024. Cortical thickness modeling and variability
968 in Alzheimer's disease and frontotemporal dementia. J Neurol 271, 1428–1438.
969 https://doi.org/10.1007/s00415-023-12087-1
970 Pérez-Millan, A., Contador, J., Juncà-Parella, J., Bosch, B., Borrell, L., Tort-Merino, A.,
971 Falgàs, N., Borrego-Écija, S., Bargalló, N., Rami, L., Balasa, M., Lladó, A.,
972 Sánchez-Valle, R., Sala-Llonch, R., 2023. Classifying Alzheimer's disease and
973 frontotemporal dementia using machine learning with cross-sectional and
974 longitudinal magnetic resonance imaging data. Human Brain Mapping.
975 https://doi.org/10.1002/hbm.26205
976 Prados, F., Cardoso, M.J., Leung, K.K., Cash, D.M., Modat, M., Fox, N.C., Wheeler-
977 Kingshott, C.A.M., Ourselin, S., 2015. Measuring brain atrophy with a
978 generalized formulation of the boundary shift integral. Neurobiol Aging 36,
979 S81–S90. https://doi.org/10.1016/j.neurobiolaging.2014.04.035
980 Punzi, M., Sestieri, C., Picerni, E., Chiarelli, A.M., Padulo, C., Delli Pizzi, A., Tullo,
981 M.G., Tosoni, A., Granzotto, A., Della Penna, S., Onofrj, M., Ferretti, A., Delli
982 Pizzi, S., Sensi, S.L., Alzheimer's Disease Neuroimaging Initiative, 2024.
983 Atrophy of hippocampal subfields and amygdala nuclei in subjects with mild
984 cognitive impairment progressing to Alzheimer's disease. Heliyon 10, e27429.
985 https://doi.org/10.1016/j.heliyon.2024.e27429
986 Rabinovici, G.D., Seeley, W.W., Kim, E.J., Gorno-Tempini, M.L., Rascovsky, K.,
987 Pagliaro, T.A., Allison, S.C., Halabi, C., Kramer, J.H., Johnson, J.K., Weiner,
988 M.W., Forman, M.S., Trojanowski, J.Q., Dearmond, S.J., Miller, B.L., Rosen,
989 H.J., 2008. Distinct MRI atrophy patterns in autopsy-proven Alzheimer's
990 disease and frontotemporal lobar degeneration. American Journal of
991 Alzheimer's Disease and other Dementias 22, 474–488.
992 https://doi.org/10.1177/1533317507308779
993 Rascovsky, K., Hodges, J.R., Knopman, D., Mendez, M.F., Kramer, J.H., Neuhaus, J.,
994 van Swieten, J.C., Seelaar, H., Dopper, E.G.P., Onyike, C.U., Hillis, A.E.,
995 Josephs, K. a., Boeve, B.F., Kertesz, A., Seeley, W.W., Rankin, K.P., Johnson,
996 J.K., Gorno-Tempini, M.-L.L., Rosen, H., Prioleau-Latham, C.E., Lee, A.,
997 Kipps, C.M., Lillo, P., Piguet, O., Rohrer, J.D., Rossor, M.N., Warren, J.D., Fox,
998 N.C., Galasko, D., Salmon, D.P., Black, S.E., Mesulam, M., Weintraub, S.,
999 Dickerson, B.C., Diehl-Schmid, J., Pasquier, F., Deramecourt, V., Lebert, F.,
1000 Pijnenburg, Y., Chow, T.W., Manes, F., Grafman, J., Cappa, S.F., Freedman,
1001 M., Grossman, M., Miller, B.L., 2011. Sensitivity of revised diagnostic criteria
1002 for the behavioural variant of frontotemporal dementia. Brain 134, 2456–2477.
1003 https://doi.org/10.1093/brain/awr179
1004 Rohrer, J.D., Woollacott, I.O.C., Dick, K.M., Brotherhood, E., Gordon, E., Fellows, A.,
1005 Toombs, J., Druyeh, R., Cardoso, M.J., Ourselin, S., Nicholas, J.M., Norgren,
1006 N., Mead, S., Andreasson, U., Blennow, K., Schott, J.M., Fox, N.C., Warren,
1007 J.D., Zetterberg, H., 2016. Serum neurofilament light chain protein is a measure
1008 of disease intensity in frontotemporal dementia. Neurology 87, 1329–1336.
1009 https://doi.org/10.1212/WNL.0000000000003154
1010 Salvatore, C., Cerasa, A., Battista, P., Gilardi, M.C., Quattrone, A., Castiglioni, I., 2015.
1011 Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's
1012 disease: a machine learning approach. Front Neurosci 9, 307.
1013 https://doi.org/10.3389/fnins.2015.00307
1014 Seelaar, H., Rohrer, J.D., Pijnenburg, Y.A.L., Fox, N.C., Swieten, J.C. van, 2011.

Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. Journal of Neurology, Neurosurgery & Psychiatry 82, 476–486. https://doi.org/10.1136/jnnp.2010.212225

Seidman, L.J., Faraone, S.V., Goldstein, J.M., Goodman, J.M., Kremen, W.S., Matsuda, G., Hoge, E.A., Kennedy, D., Makris, N., Caviness, V.S., Tsuang, M.T., 1997. Reduced subcortical brain volumes in nonpsychotic siblings of schizophrenic patients: A pilot magnetic resonance imaging study. American Journal of Medical Genetics - Neuropsychiatric Genetics 74, 507–514. https://doi.org/10.1002/(SICI)1096-8628(19970919)74:5<507::AID-AJMG11>3.0.CO;2-G

Swift, I.J., Sogorb-Esteve, A., Heller, C., Synofzik, M., Otto, M., Graff, C., Galimberti, D., Todd, E., Heslegrave, A.J., van der Ende, E.L., Van Swieten, J.C., Zetterberg, H., Rohrer, J.D., 2021. Fluid biomarkers in frontotemporal dementia: past, present and future. J Neurol Neurosurg Psychiatry 92, 204–215. https://doi.org/10.1136/jnnp-2020-323520

Tort-Merino, A., Falgàs, N., Allen, I.E., Balasa, M., Olives, J., Contador, J., Castellví, M., Juncà-Parella, J., Guillén, N., Borrego-Écija, S., Bosch, B., Fernández-Villullas, G., Ramos-Campoy, O., Antonell, A., Rami, L., Sánchez-Valle, R., Lladó, A., 2022. Early-onset Alzheimer's disease shows a distinct neuropsychological profile and more aggressive trajectories of cognitive decline than late-onset. Ann Clin Transl Neurol. https://doi.org/10.1002/acn3.51689

Wang, J., Redmond, S.J., Bertoux, M., Hodges, J.R., Hornberger, M., 2016. 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 8.

Wang, S.-H., Phillips, P., Sui, Y., Liu, B., Yang, M., Cheng, H., 2018. Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. J Med Syst 42, 85. https://doi.org/10.1007/s10916-018-0932-7

Whitwell, J.L., Clifford, R.J., Przybelski, S.A., Parisi, J.E., Senjem, M.L., Boeve, B.F., Knopman, D.S., Petersen, R.C., Dickson, D.W., Josephs, K.A., 2011. Temporoparietal atrophy: a marker of AD pathology independent of clinical diagnosis. Neurobiol Aging 32, 1531–1541. https://doi.org/10.1016/j.neurobiolaging.2009.10.012

Wojtas, A., Heggeli, K.A., Finch, N., Baker, M., DeJesus-Hernandez, M., Younkin, S.G., Dickson, D.W., Graff-Radford, N.R., Rademakers, R., 2012. C9ORF72 repeat expansions and other FTD gene mutations in a clinical AD patient series from Mayo Clinic. Am J Neurodegener Dis 1, 107–118.

Yang, Z., Kinney, J.W., Cordes, D., The Alzheimer's Disease Neuroimaging Initiative, null, 2024. Uptake of 18F-AV45 in the Putamen Provides Additional Insights into Alzheimer's Disease beyond the Cortex. Biomolecules 14, 157. https://doi.org/10.3390/biom14020157

Zee, J. van der, Sleegers, K., Broeckhoven, C.V., 2008. Invited Article: The Alzheimer disease–frontotemporal lobar degeneration spectrum. Neurology 71, 1191–1197. https://doi.org/10.1212/01.wnl.0000327523.52537.86

FIGURES & TABLES

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 participants, 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**

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 (N=491)	Accuracy: 87.7% AD: 73.4% CTR: 64.5%	Accuracy: 86.9% FTD: 74.2% CTR: 73.3%	Accuracy: 81.8% AD: 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%

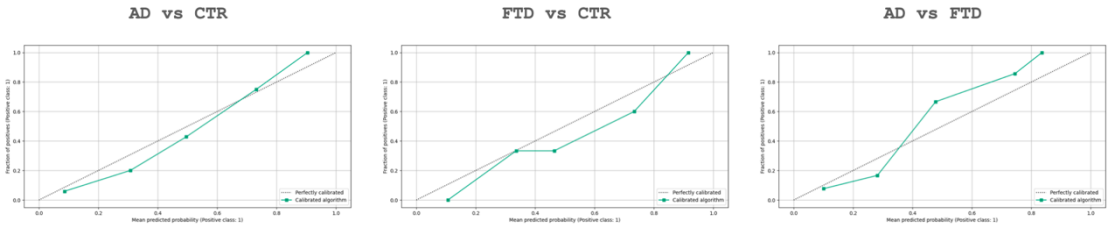

Figure 1: Calibration plots with the test dataset to evaluate the calibration quality for each classification scenario. We observe that the line plot generated by mean predicted probabilities vs fraction of positiveness (represented with a green plain line) approaches the identity line (here represented as a gray-dashed line), which would indicate a perfect calibration AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls.

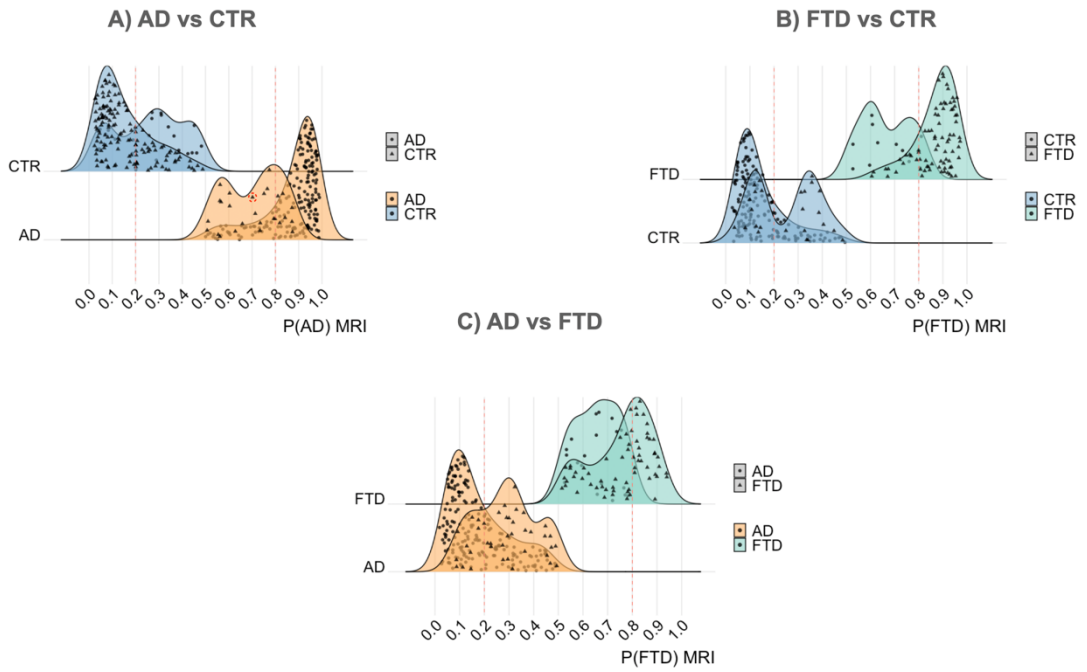

Figure 2: Density plot to study the obtained individual probabilities with the MRI-based algorithm (all the study participants included). The clinical diagnosis is identified with triangles or circles, and the algorithm's diagnoses are plotted with different colors. The vertical red dashed lines indicate the thresholds for the grey zone. We highlight a random participant (point) with a dashed red circle in the comparison of AD vs CTR for explanatory reasons. This point would have a probability of 0.7 of being AD (x-axis) and consequently a probability of 0.2 of being CTR (inverse probability). AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls, P: probability.

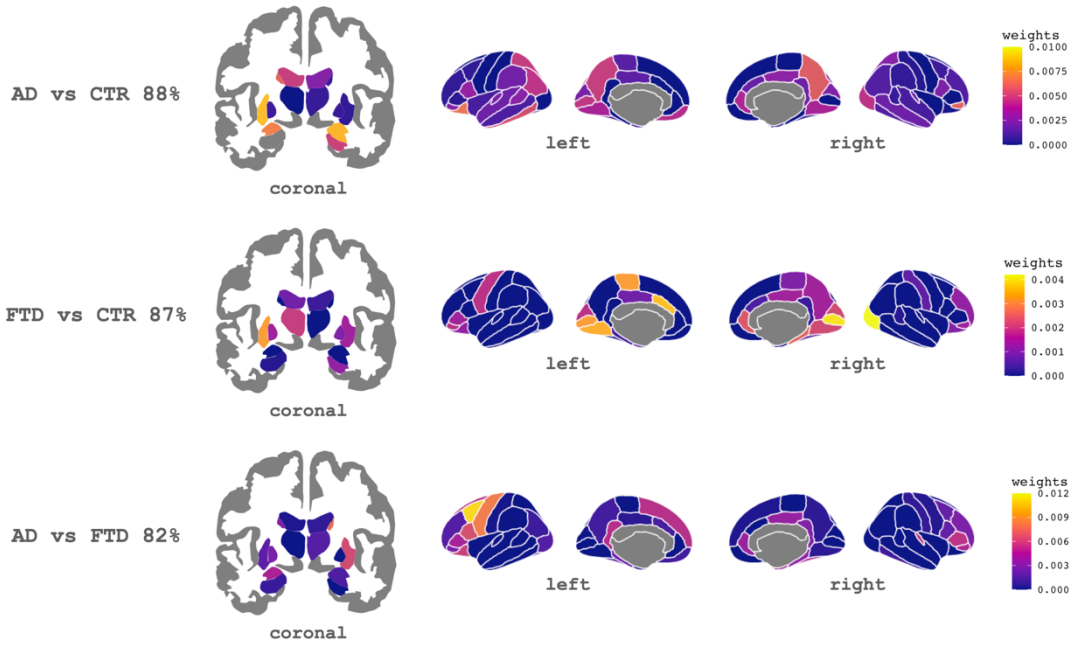

Figure 3: Cortical and subcortical patterns of the feature importance of each region associated with the different AD and FTD comparison (comparison per row). At a higher value major importance of that region for the classification. AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls.

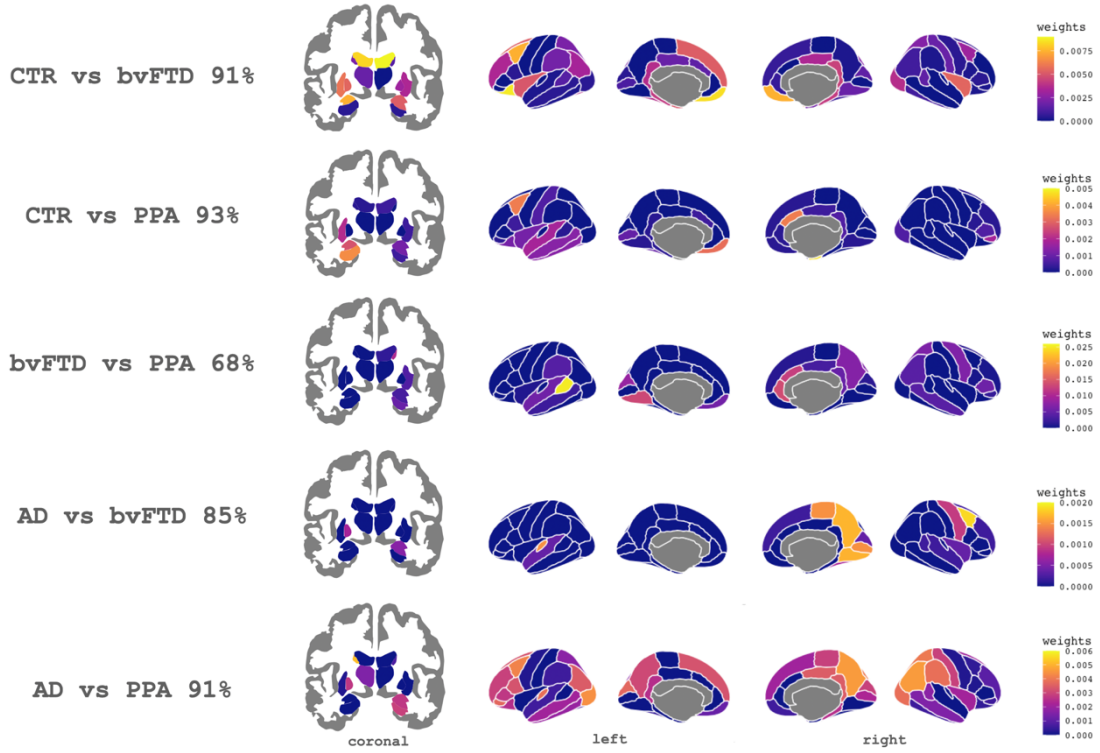

Figure 4: Cortical and subcortical patterns of the feature importance of each region associated with the different bvFTD and PPA comparison (comparison per row). At a higher value major importance of that region for the classification. AD: Alzheimer's disease, bvFTD: behavior frontotemporal dementia, PPA: primary progressive aphasia, CTR: healthy controls.

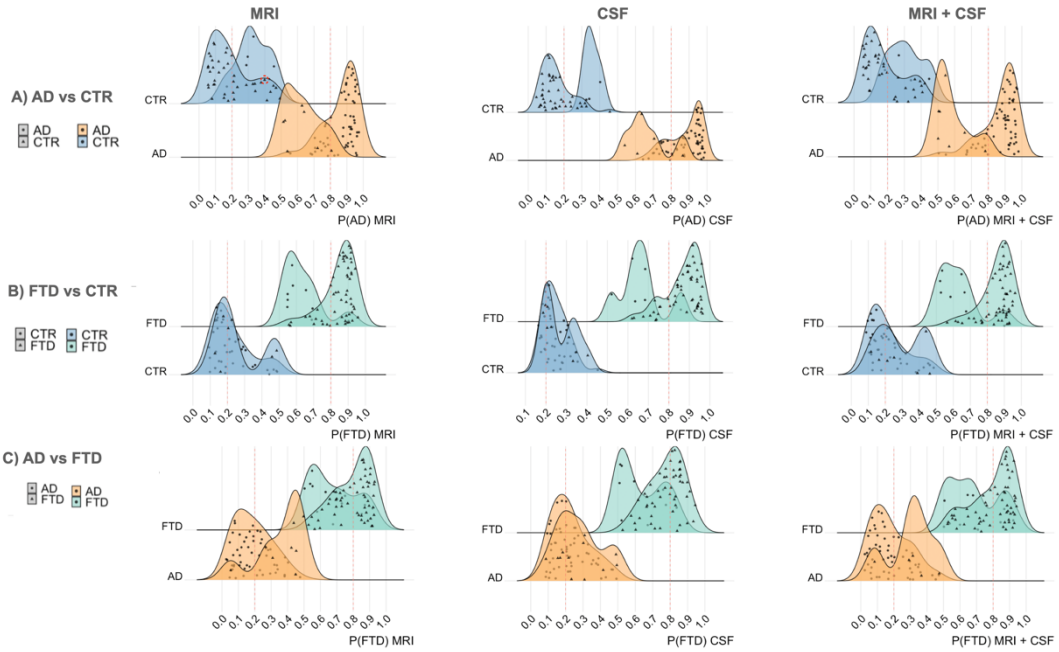

Figure 5: Density plot to study the obtained individual probabilities with the MRI-based, CSF-based, and MRI- and CSF-based algorithms (participants included in all the analyses were those with MRI and CSF data). It can be seen in the clinical diagnosis with triangles or circles and the algorithm's diagnosis, plotted with different colors. The vertical red dashed lines indicate the thresholds for the grey zone. We highlight a random participant (point) with a dashed red circle in the comparison of AD vs CTR for explanatory reasons. This point would have a probability of 0.4 of being AD (x-axis) and consequently a probability of 0.6 of being CTR (inverse probability). AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls, P: probability.