



**HAL**  
open science

## 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

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

4

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

9

10 <sup>1</sup>Alzheimer's disease and other cognitive disorders unit. Service of Neurology, Hospital Clínic de  
11 Barcelona. Fundació Recerca Clínic Barcelona-IDIBAPS, Barcelona, Spain.

12 <sup>2</sup>Institut de Neurociències, University of Barcelona, Barcelona, Spain.

13 <sup>3</sup>Department of Biomedicine, Faculty of Medicine, University of Barcelona, Barcelona, Spain

14 <sup>4</sup>Inria, CEA, Université Paris-Saclay, Paris, France

15 <sup>5</sup>Biochemistry and Molecular Genetics Department, Hospital Clínic de Barcelona, Barcelona,  
16 Spain

17 <sup>6</sup>Image Diagnostic Centre, Hospital Clínic de Barcelona, CIBER de Salud Mental, Instituto de  
18 Salud Carlos III.Magnetic Resonance Image Core Facility, IDIBAPS, Barcelona, Spain

19 <sup>7</sup>Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina  
20 (CIBER-BBN), Barcelona, Spain

21

22 <sup>8</sup>Fundació de Recerca Clínic Barcelona-Institut d'Investigacions Biomèdiques August Pi I Sunyer  
23 (FRCB-IDIBAPS), Barcelona, Spain.

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

27 Email address: [rosar.sala@ub.edu](mailto:rosar.sala@ub.edu)

28 \* These authors contributed equally to this work.

29

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

32

33

#### 34 **Abbreviations**

35 AD=Alzheimer's disease

36 A $\beta$ 42=Amyloid-beta protein 42

37 bvFTD=Behavioral variant frontotemporal dementia

38 CSF=Cerebrospinal fluid

39 CTh=Cortical thickness

40 CTR=healthy controls

41 FTD=Frontotemporal dementia

42 HCB=Hospital Clínic de Barcelona

43 ML=Machine Learning

44 MMSE=Mini-Mental State Examination

45 MRI=Magnetic Resonance Imaging

46 NfL=neurofilament light chain

47 nfvPPA=Nonfluent Variant Primary Progressive Aphasia

48 PPA=Primary progressive aphasia

49 SVM=Support Vector Machine

50 svPPA=Semantic Variant Primary Progressive Aphasia

51

52

53

54

55

56

57

58

59

60

61 **ABSTRACT**

62

63 Neuroimaging and fluid biomarkers are used to differentiate frontotemporal dementia  
64 (FTD) from Alzheimer’s disease (AD). We implemented a machine learning algorithm  
65 that provides individual probabilistic scores based on magnetic resonance imaging (MRI)  
66 and cerebrospinal fluid (CSF) data. We investigated whether combining MRI and CSF  
67 levels could improve the diagnosis confidence. 215 AD patients, 103 FTD patients, and  
68 173 healthy controls (CTR) were studied. With MRI data, we obtained an accuracy of  
69 82% for AD vs. FTD. A total of 74% of FTD and 73% of AD participants have a high  
70 probability of accurate diagnosis. Adding CSF-NfL and 14-3-3 levels improved the  
71 accuracy and the number of patients in the confidence group for differentiating FTD from  
72 AD. We obtain individual diagnostic probabilities with high precision to address the  
73 problem of confidence in the diagnosis. We suggest when MRI, CSF, or the combination  
74 are necessary to improve the FTD and AD diagnosis. This algorithm holds promise  
75 towards clinical applications as support to clinical findings or in settings with limited  
76 access to expert diagnoses.

77

78

79

80

81

82

83

84

85

86

87

88

## 89 1. INTRODUCTION

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

105

106 During the last two decades, fluid biomarker studies have substantially improved the  
107 diagnosis of neurodegenerative dementias. The current clinical criteria for AD diagnosis  
108 include cerebrospinal fluid (CSF) biomarkers, such as the amyloid-beta protein 42  
109 (A $\beta$ 42), the total tau (t-tau), and phosphorylated tau (p-tau) (Albert et al., 2011; McKhann  
110 et al., 2011). However, currently, FTD criteria do not include biochemical markers.  
111 Neurofilament light chain (NfL) levels, a marker of neuroaxonal damage, and 14-3-3  
112 protein levels, a marker of synaptic-neuronal loss, have been both proposed as nonspecific  
113 neurodegeneration markers that could support the diagnosis of FTD, although their levels  
114 are also increased in AD compared to controls (Alcolea et al., 2017; Antonell et al., 2019;  
115 McFerrin et al., 2017; Rohrer et al., 2016).

116

117 Magnetic Resonance Imaging (MRI) is broadly used in the study of AD and FTD, both  
118 at the research and the clinical levels. Visual evaluation of the atrophy pattern is mainly  
119 used in the clinical setting (Davatzikos et al., 2008; Du et al., 2007). MRI markers such

120 as atrophy measures have a neurobiological implication as could play an important role  
121 in disease diagnosis and tracking of pathologic progression in AD or FTD and are used  
122 as outcome measures in trials of potentially disease-modifying therapies (Frisoni et al.,  
123 2010; Prados et al., 2015). Quantitative MRI studies have described patterns of cortical  
124 thickness and gray matter (GM) volume loss in AD and FTD at the group level when  
125 compared separately with healthy populations (Bocchetta et al., 2021; Borrego-Écija et  
126 al., 2021; Canu et al., 2017; Contador et al., 2021; Möller et al., 2015, 2013). The  
127 conclusions drawn by prior studies regarding a specific MRI pattern of atrophy for AD  
128 were that the regions of the temporal, parietal, and occipital lobes were the ones more  
129 affected (Blanc et al., 2015; Möller et al., 2013; Whitwell et al., 2011). On the other hand,  
130 atrophy in the temporal and frontal lobes constitutes the specific FTD atrophy pattern  
131 (Couto et al., 2013; Möller et al., 2015; Rabinovici et al., 2008). However, quantitative  
132 MRI studies are only scarcely used in clinics due to technical difficulties and limited  
133 accuracy in performing the diagnosis at the individual level.

134

135 A growing body of evidence supports the role of machine learning (ML) techniques using  
136 brain MRI (Abraham et al., 2014; Frizzell et al., 2022; Mateos-Pérez et al., 2018) to  
137 support the clinical diagnosis of these two dementias (Bron et al., 2017; Chagué et al.,  
138 2021; Klöppel et al., 2008; Möller et al., 2016; Pérez-Millan et al., 2023). Many studies  
139 have shown that a support vector machine (SVM) with neuroimaging data differentiates  
140 AD or FTD patients from healthy controls (Bisenius et al., 2017; Bron et al., 2021;  
141 Cuingnet et al., 2011; Lampe et al., 2023; Magnin et al., 2009; Meyer et al., 2017; Pérez-  
142 Millan et al., 2023). In the context of the differential diagnosis of these two dementias, it  
143 is known that the clinical symptoms of FTD and AD can display a substantial overlap  
144 between them (Mendez, 2006; Wojtas et al., 2012; Zee et al., 2008), suggesting that  
145 additional markers may help in the differentiation. In this sense, in recent years, several  
146 studies have appeared to assess the use of artificial intelligence, including ML and deep  
147 learning, to evaluate, predict, and classify the differential diagnosis of FTD and AD with  
148 multiple data, including MRI, neuropsychological test, or biological data (Garcia-  
149 Gutierrez et al., 2021; Javeed et al., 2023; Kim et al., 2019; Maito et al., 2023; Moguilner  
150 et al., 2022; Nguyen et al., 2023; Pérez-Millan et al., 2023). These studies have presented  
151 different approaches and multiple datasets, leading to different predictive capabilities and  
152 a wide range of applicability settings. It is still not clear which is the best data to use and

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

165

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

172

## 173 **2. MATERIALS AND METHODS**

### 174 **2.1. Participants**

175 We recruited the participants from the Alzheimer's disease and other cognitive disorders  
176 unit of the Hospital Clínic de Barcelona (HCB), Barcelona, Spain. All participants  
177 underwent a complete clinical and cognitive evaluation, a lumbar puncture for AD  
178 markers following the hospital's standard clinical care practice, and a 3T high-resolution  
179 structural MRI scan. Participants with a history of stroke, traumatic brain injury, major  
180 psychiatric disorder, or alcohol abuse were excluded.

181 All AD participants fulfilled the criteria for mild dementia due to AD (Albert et al., 2011;  
182 McKhann et al., 2011) supported by the CSF biomarkers profile suggesting underlying  
183 AD neuropathology according to National Institute on Aging/Alzheimer's Association

184 Research Framework 2018 (Jack et al., 2018). The FTD participants fulfilled the  
185 diagnostic criteria for either behavioral variant frontotemporal dementia (bvFTD) or  
186 FTD-related primary progressive aphasia (PPA) phenotypes, including Semantic Variant  
187 Primary Progressive Aphasia (svPPA) and Nonfluent Variant Primary Progressive  
188 Aphasia (nfvPPA) (Gorno-Tempini et al., 2011; Rascovsky et al., 2011). Participants with  
189 clinical features of AD but with unsupportive CSF profile were excluded from the study,  
190 as well as patients with FTD who showed an AD CSF profile. Healthy controls (CTR)  
191 had cognitive performance within the normative range (cutoff 1.5 SD from the normative  
192 mean, available normative data (Peña-Casanova et al., 2009)) and normal AD CSF  
193 biomarkers. We included healthy voluntaries and individuals with subjective memory  
194 complaints; all of them performed within the normal range in all the tests of the  
195 comprehensive neuropsychological assessment reported in previous studies of the group  
196 (Contador et al., 2022; Tort-Merino et al., 2022).

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

199

## 200 **2.2 Biochemical markers**

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

205

## 206 **2.3 MRI acquisition**

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

212

## 213 **2.4 MRI processing**



214 We used the processing stream available in FreeSurfer version 6.0  
215 (<http://surfer.nmr.mgh.harvard.edu.sire.ub.edu/>) to perform cortical reconstruction and  
216 volumetric segmentation of the T1-weighted acquisitions. FreeSurfer allowed us to obtain  
217 cortical thickness (CTh) maps and segment the subcortical structures (Fischl et al., 2004;  
218 Fischl and Dale, 2000). From reconstructed data, we got summary measures of mean CTh  
219 and GM volumes across the left and right hemispheres and summary measures of mean  
220 CTh in 68 cortical regions and GM volumes of 16 subcortical structures, all derived from  
221 atlases available in FreeSurfer (Desikan et al., 2006; Seidman et al., 1997). The estimated  
222 intracranial volume estimated with FreeSurfer was used to normalize volume measures.  
223 All images and individual segmentations were visually inspected and manually corrected  
224 if needed.

## 225 **2.5 MRI-based individual probabilistic classification algorithm**

226 We used all CTh values, GM subcortical volumes, and the age of the participants to create  
227 our ML algorithm. Age was added as a feature to ensure that the ML algorithm captures  
228 age's true importance or influence on the model. We introduced the regional measures of  
229 both hemispheres separately, leading to a total of 84 features per subject (see  
230 Supplementary Material).

231

232 We first converted MRI data (subcortical volumes normalized by intracranial volume and  
233 CTh measures) to z-scores with the training data (all participants) and then applied this  
234 conversion to obtain the z-scores for the test dataset. We implemented a calibrated  
235 classifier with an SVM as a base estimator to predict the diagnosis of the participants,  
236 where the features of the ML algorithm were the previously described values transformed  
237 to z-scores. The calibration was set using the *CalibratedClassifierCV* from the Scikit-  
238 learn library with a cross-validation of 5 with the train set. As our proposal involves an  
239 inner loop in the calibration, to maximise the number of samples included in this  
240 procedure, the SVM base estimator hyperparameters were fixed to kernel=rbf, C=1 and  
241 gamma=1/(number of features). For each classifier, we fitted a logistic regression model  
242 that distributes the classifier's output of the decision function, which predicts a “soft”  
243 score for each sample in relation to each group. The model assumes a sigmoid shape and  
244 calibrates the probability between 0 and 1. We created classifiers for each pair of  
245 diagnostic groups (AD vs. CTR, FTD vs. CTR, and AD vs. FTD) and across the three  
246 groups (AD vs. FTD vs. CTR). Then, we subdivided the FTD group into bvFTD and PPA

247 (we merged svPPA and nfvPPA due to the sample size and named the group as PPA), and  
248 we used them as independent groups in a new set (AD vs. bvFTD, AD vs. PPA, CTR vs.  
249 bvFTD, CTR vs. PPA, and bvFTD vs. PPA). All the comparisons were performed with a  
250 5-fold cross-validation with stratified folds preserving the percentage of samples of each  
251 diagnosis to evaluate the performance of the classification. Then, we analysed the  
252 importance of each region for the decision of the classification through a permutation  
253 feature importance estimation (Breiman, 2001) using the test data of each run with the  
254 calibrated SVM algorithm. Feature importance is evaluated as the difference between the  
255 baseline metric and the metric obtained after permuting a feature and re-evaluating it. We  
256 used the *permutation\_importance* from the Scikit-learn library. The higher the weight,  
257 the larger the importance of the feature in the classification. Notably, the units of the  
258 weights are rather arbitrary. Thus, even if they can be compared across features for a  
259 given classification problem, they should not be compared across different scenarios.

260

261 We obtained individual probabilities associated with group correspondences as output  
262 values for each test data point given by the calibrated SVM. They had complementary  
263 values (i.e., the probability of one group is equal to 1 minus the probability of the other  
264 in the classification between two diagnostics), and they were directly associated with the  
265 output category (i.e., the final classification was the one with probability  $>0.5$ ). We  
266 conventionally set two levels of diagnosis confidence: an individual probability  $\geq 0.8$  (or  
267  $\leq 0.2$ ) was considered to provide high diagnosis confidence, while probabilities between  
268 0.2 and 0.8 were considered a “gray zone”, with lower or insufficient diagnosis  
269 confidence for the clinical decision. Thus, we estimated the accuracy and the number of  
270 individuals with a high probability of being from the group for each classification.

271

272 Finally, we aimed to explore if NfL and 14-3-3 levels could help diagnose the individuals  
273 of the gray zone of the MRI diagnosis for the following comparisons due to the available  
274 data: AD vs. CTR, FTD vs. CTR, and AD vs. FTD. Thus, we created a reduced dataset  
275 with participants having MRI data, NfL, and 14-3-3 levels. We trained and tested the  
276 proposed algorithm in 3 situations: MRI-based algorithm, CSF-based algorithm, and MRI  
277 and CSF-based algorithm to study if the individual probabilities towards the actual class  
278 increased. We did not include A $\beta$ 42, t-tau, and p-tau levels to avoid circularity, as these  
279 markers were used in the clinical diagnosis according to current criteria.

280

281 We implemented the ML algorithm in Python version 3.10.6 (www.python.org) with the  
282 Scikit-learn library (Pedregosa et al., 2011).

283

## 284 **2.6 Statistics**

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

290

## 291 **3. RESULTS**

### 292 **3.1. Sample demographics**

293 The prospective study includes 491 participants: 215 AD, 103 FTD (56 bvFTD, 24  
294 svPPA, 21 nfvPPA, and 2 PPA), and 173 CTR participants (74% were healthy  
295 voluntaries, and 26% were individuals with subjective memory complaints but  
296 completely normal cognition and normal AD CSF biomarkers). A subset of the study  
297 participants had CSF measures available: NfL (N=365) and 14-3-3 (N=182). Table 1  
298 shows demographic information, group statistics (p-values corrected by Benjamini-  
299 Hochberg), and biomarker levels. As expected, CSF biomarkers levels showed significant  
300 differences between groups (corrected p-value $<0.05$ ). There were differences in age and  
301 sex. As expected, based on previous studies, AD and CTR groups had more women than  
302 men; meanwhile, the FTD groups were more harmonized. Regarding age, CTR were  
303 younger than AD and FTD participants.

304

### 305 **3.2. MRI-based probabilistic classification algorithm**

306 We estimated the accuracy performance of our algorithm as the mean accuracy obtained  
307 in each k-fold of the test data. The train and test were 80% and 20% of the total sample,  
308 respectively, in all the analyses. Including sex in the algorithm led to similar results. We  
309 got an accuracy of  $88 \pm 8\%$  (AUC=0.87) when discriminating AD patients from CTR,  
310 and  $87 \pm 4\%$  (AUC=0.85) when determining FTD patients from CTR. When we tried  
311 classifying AD vs. FTD patients, the accuracy was  $82 \pm 6\%$  (AUC=0.77). All AUC and

312 the F1-score for all comparisons are shown in Supplementary Table 2. Notably, given a  
313 rather unbalanced scenario, the results of these metrics indicate that the overall accuracy  
314 is not likely to be driven by a single-group classification. Finally, we obtained an accuracy  
315 of  $77 \pm 6\%$  when discriminating between the three groups (AD vs. FTD vs. CTR) (Table  
316 2).

317

318 As shown in Figure 1 and in Supplementary Figure 1, the resulting algorithms were well-  
319 calibrated, which allowed us to create confidence ranges in the algorithm classification.  
320 The comparison of AD vs. CTR showed that 73% of AD participants and 65% of CTR  
321 participants presented a probability higher than 0.8 (probability of being correctly  
322 classified). In the FTD vs. CTR comparison, we found 74% FTD participants and 73%  
323 CTR participants with a probability  $\geq 0.8$ . Finally, when discriminating AD vs. FTD, we  
324 found 73% AD participants and 74% FTD participants with probabilities above 0.8 for  
325 being classified as AD or FTD, respectively. Figure 2 shows the density of the individual  
326 probabilities and how the distribution between the clinical and the algorithm diagnosis is  
327 distributed within the group with an individual probability  $\geq 0.8$ . Figure 2 shows that a  
328 higher probability reduced the misprediction. However, a high probability does not ensure  
329 a correct classification.

330

331 Then, we aimed to study the FTD clinical subtypes separately. Due to limitations in  
332 sample size, we merged svPPA and nfvPPA in the same named PPA group. We obtained  
333  $91 \pm 2\%$  accuracy for classifying bvFTD patients vs. CTR and  $93 \pm 4\%$  when  
334 discriminating PPA patients from CTR. In both cases, the accuracy increased compared  
335 to the accuracy reported for all FTD together ( $87 \pm 4\%$ ). Compared with AD, we obtained  
336  $85 \pm 3\%$  for the bvFTD vs. AD comparison and  $91 \pm 3\%$  for the PPA vs AD. Finally, we  
337 obtained an accuracy of  $68 \pm 6\%$  discriminating bvFTD from PPA.

338

### 339 **3.3. Important MRI regions for classification**

340 Figure 3 shows the region weights associated with each comparison (Tables 3, 4 and 5 in  
341 Supplementary Material show the feature importance of all the features). It should be  
342 noted that these weights should not be compared between classification settings, as they  
343 reflect the absolute change in accuracy associated with each feature. However, the  
344 different feature orderings within each comparison offer important insights. In summary,

345 when comparing AD versus CTR, the GM volume of the hippocampus, putamen, and  
346 amygdala played the most crucial role. For FTD vs. CTR, we found that occipital, parietal,  
347 and frontal regions emerged as the top regions for the classification. Finally, when  
348 discriminating both dementias (AD vs. FTD), we found a widespread pattern in which  
349 the CTh measures were generally more important than subcortical GM volumes,  
350 especially those in the frontal lobe. In this comparison, age emerged as a medium  
351 important feature, being in the sixth place (Table 5 Supplementary Material).

352

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

362

### 363 **3.4. Individual probabilities using MRI and CSF data**

364 The group classification performance of the algorithm and the percentage of participants  
365 with an individual probability  $\geq 80\%$  using MRI-only, CSF-only, and combined MRI and  
366 CSF data are presented in Table 2. Adding NfL and 14-3-3 data to the MRI data improved  
367 the results in some of the comparisons.

368

369 For comparing AD vs. CTR, CSF data (NfL and 14-3-3) alone was enough and better  
370 than MRI data to discriminate between AD and CTR participants (Table 2, Figure 5).  
371 Thus, in this comparison the MRI data could not be needed. In the comparison between  
372 FTD and CTR, having MRI or CSF data led us to similar results in terms of accuracy (see  
373 Table 2 and Figure 5). However, the MRI data is needed for having a high number of  
374 CTR with a CTR-probability higher than 0.8. Thus, in this case, combining the MRI and  
375 CSF data reduced the participants in the gray zone of the diagnosis. Finally, when we  
376 compared AD and FTD participants, combining MRI and CSF data increased both the

377 accuracy and the number of participants with a probability  $\geq 0.8$  (see Table 2 and Figure  
378 5). Thus, in the differential diagnosis of FTD and AD, having MRI and CSF data could  
379 bring more reliability to the diagnosis.

380

#### 381 **4. DISCUSSION**

382 In this study, we implemented a machine learning algorithm that discriminates FTD and  
383 AD patients using data from structural MRI. In addition, our algorithm was able to  
384 differentiate subtypes of FTD with good accuracy. Clinical diagnosis requires decisions  
385 at the individual level, and the degree of confidence in the diagnosis is key in managing  
386 the patient. We approach the clinical question of diagnosis confidence using individual  
387 probabilities. Among our key results, we found that 74% FTD and 73% AD participants  
388 showed an individual probability  $\geq 0.8$  of being well-classified by the algorithm in the FTD  
389 vs. AD comparison. Adding CSF neurodegeneration markers (NfL and 14-3-3) levels  
390 improved the diagnosis classification or the number of patients with high individual  
391 probability for the diagnosis in some cases, especially for differentiating FTD from AD.

392

393 Previous ML algorithms using structural MRI data have reported accuracies between 76  
394 and 97% for AD vs. CTR, 72-88% for FTD versus CTR, 51-90% for AD versus FTD,  
395 and 54-70% in discriminating between AD, FTD, and CTR (Bron et al., 2017; Canu et  
396 al., 2017; Davatzikos et al., 2008; Dukart et al., 2011; Kim et al., 2019; Li et al., 2021;  
397 Lin et al., 2018; Möller et al., 2016; Moore et al., 2019; Pérez-Millan et al., 2023;  
398 Salvatore et al., 2015; Wang et al., 2016, 2018). These studies used different algorithms,  
399 with the SVM being the most common. We obtained accuracies that are in accordance  
400 with, or even outperformed, previously reported algorithms, especially for AD vs. FTD  
401 (Basheera and Sai Ram, 2019; Dashtipour et al., 2022; Frizzell et al., 2022; Mateos-Pérez  
402 et al., 2018; McCarthy et al., 2018). We have differentiated FTD expressions (bvFTD and  
403 PPA) against AD or CTR, outperforming previously published works (Canu et al., 2017;  
404 Lampe et al., 2023; Möller et al., 2016; Wang et al., 2016). First, regarding the  
405 comparisons with CTR, for bvFTD, we obtained a 91% accuracy, and in the case of the  
406 PPA participants, an accuracy of 93%. When classifying bvFTD and PPA separately  
407 against AD, we obtained accuracies up to 90% for both cases. However, when we tried  
408 to classify bvFTD vs. PPA, we obtained an accuracy of 68%, which is lower than the

409 accuracy reported by Kim et al. (Kim et al., 2019), probably due to differences in the  
410 algorithm. They used hierarchical classification with cortical atrophy, but in the CTh  
411 measures, they applied noise removal. Then, the hierarchical classification algorithm was  
412 constructed by applying linear discriminant analysis (LDA) in combination with principal  
413 component analysis (PCA) to the cortical thickness data. The results of their proposed  
414 algorithm using the entire hierarchical tree showed an overall 75.8% accuracy, but  
415 looking in detail at step 3 of the algorithm, which focused on the FTD group (bvFTD  
416 versus PPA), the algorithm presented 86.9% accuracy, which is higher than the obtained  
417 in our algorithm. Thus, our algorithm accurately distinguishes AD, FTD, and CTR using  
418 MRI data. However, the classification accuracy between bvFTD and PPA is lower than  
419 in some previous works.

420

421 Other studies using multimodal information also reported high classification accuracy  
422 combining data from different imaging modalities or other biological and clinical  
423 measures (Bron et al., 2021; Chen et al., 2017; Dashtipour et al., 2022; Dukart et al.,  
424 2011). Even so, in some cases, our scores with only structural MRI data showed better  
425 accuracy (Bouts et al., 2018; Bron et al., 2017; Dyrba et al., 2015; Wang et al., 2016).  
426 Here, we evaluated if adding CSF data to the MRI could improve the accuracy or the  
427 number of participants with a high diagnosis confidence. This could help reduce the  
428 number of clinical tests usually required to take these patients. This approach identifies  
429 participants who need additional clinical tests while sparing others from unnecessary  
430 examinations that wouldn't yield new evidence. In our cohort, adding NfL and 14-3-3  
431 CSF data to the MRI data provided was beneficial for the accuracy of the group  
432 classification or the number of participants with high individual diagnosis probability in  
433 some cases, especially in the differential diagnosis of FTD and AD. In this case, the  
434 combination of the data of these two modalities increased the accuracy and the number  
435 of participants with a reliable diagnostic. In the case of discriminating AD and CTR, the  
436 CSF data presented the best performance, so the MRI data may be unnecessary or  
437 redundant. Finally, in the case of classifying FTD versus CTR, MRI and CSF data showed  
438 similar results, though the MRI data contributes to the reliability of the CTR diagnosis,  
439 so in this case, we should recommend using MRI data. The comparisons between CTR  
440 and dementia patients (AD or FTD) are a gold standard for studying the algorithm and  
441 for research proposes. For clinical utility, it could be helpful the differential diagnosis of  
442 FTD vs AD or their subtypes. We excluded A $\beta$ 42, t-tau, and p-tau from our algorithm to

443 avoid circularity. These biomarkers were used to confirm AD when positive or FTD when  
444 negative. In contrast, NfL and 14-3-3 were not used to determine the label feature.  
445 Although NfL and 14-3-3 correlate with A $\beta$ 42, t-tau, and p-tau, so does hippocampus  
446 volume. Therefore, this correlation with current diagnostic biomarkers of AD does not  
447 bias the algorithm, as these biomarkers were not used to create the feature label  
448 (diagnosis), which is what the ML algorithm aims to predict. The participants with  
449 available CSF NfL and 14-3-3 biomarkers were not those that the clinician was already  
450 in doubt based on the MRI, as the diagnostic criteria for these patients are supported by a  
451 CSF profile.

452

453 Besides reaching good accuracies, one of the main novelties of our work is that we  
454 obtained the individual probabilities for each diagnosis in all comparisons. Notably, as  
455 we built our first set of algorithms uniquely with MRI data, these probabilities might  
456 reflect each individual's brain atrophy severity. Using these values, we could identify the  
457 participants with high diagnosis confidence (with a probability upper to 80%) and those  
458 who do not have that high confidence that could be a candidate for further evaluations.  
459 Notably, more than 70% of AD and FTD participants were classified with high diagnosis  
460 confidence in the FTD vs. AD comparison. It is important to note that our results are  
461 merely an example of interpreting these probabilities in a context where participants are  
462 distributed along a spectrum between two groups, assuming that clinical conditions (or  
463 labels) are known. However, the interpretation may differ in various clinical contexts,  
464 particularly when there are more erroneous or unknown diagnoses. We suggest evaluating  
465 these algorithms with different datasets to explore various applications and scenarios.

466

467 Furthermore, we depicted the patterns that drive accuracy for each classification setting  
468 to obtain a comprehensive explanation of structural changes in both dementias. The GM  
469 volume of the hippocampus, putamen, and amygdala were essential in differentiating AD  
470 from CTR. While the mathematical approaches used by ML algorithms differ from  
471 traditional statistical group comparisons—mainly because they account for interactions  
472 and hidden relationships between features—several key conclusions can be drawn from  
473 these patterns. Notably, our results reflect the typical AD neuroimaging pattern, found  
474 mainly in the early stages of the disease, where most patients of our sample are. The  
475 involvement of the hippocampus and subcortical areas aligns with the neuropathological  
476 continuum demonstrated by postmortem studies, such as Braak et al., 2011, showing that



477 tau pathology affects these regions in the earliest stages of the disease. Hippocampal  
478 atrophy is considered a hallmark feature of AD, directly related to memory impairment  
479 (Braak et al., 2011). The amygdala, involved in emotional processing and memory, also  
480 exhibits early atrophy, contributing to behavioral symptoms (Johnson et al., 2024; Punzi  
481 et al., 2024). The putamen, part of the basal ganglia, may be affected by tau pathology  
482 spreading from the hippocampus and other subcortical areas in early disease stages,  
483 impacting cognitive functions (Yang et al., 2024). By contrast, when differentiating FTD  
484 from CTR, the cortical regions were the most important, especially the CTh of occipital,  
485 parietal, and frontal regions. Accordingly, GM volumes of subcortical areas could help to  
486 identify AD patients, while CTh could be the key to identifying FTD participants. This  
487 aligns with findings obtained using more classical analysis methods (Avants et al., 2010;  
488 Contador et al., 2021; Dickerson et al., 2001; Frisoni et al., 2010; Gil-Navarro et al., 2013;  
489 Gordon et al., 2016; Hodges and Patterson, 2007; Jack et al., 2000). Regarding FTD  
490 variants, the frontal brain regions emerged for the bvFTD, while the hippocampus and  
491 temporal regions were the most important in PPA, as previously reported (Avants et al.,  
492 2010). These results align with previously described brain changes in bvFTD and PPA.  
493 bvFTD is characterized by primary degeneration affecting the frontal lobes, causing  
494 behavioral symptoms such as disinhibition, apathy, and impulsivity (Rascovsky et al.,  
495 2011). In contrast, PPA is marked by degeneration in the language areas, particularly the  
496 left temporal lobe, manifesting as impaired language skills (Gorno-Tempini et al., 2011).  
497

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

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

511

512 Our study also presents several limitations. First, it is unicentric. It has the advantage that  
513 all the participants had the same MRI scanner protocol and clinical criteria for the  
514 diagnosis. In the case of using the algorithm in other centers, the increased heterogeneity  
515 of the data could worsen the algorithm's performance. Also, it is focused on only AD or  
516 FTD, so applying it to different clinical contexts could affect the results, particularly when  
517 there are more erroneous, unknown diagnoses or other neurodegenerative dementias.  
518 Another limitation regarding the FTD participants is that, when looking at the different  
519 clinical expressions, we reduced the sample size to approximately 50 participants for each  
520 group, and svPPA and nfvPPA had to be studied together. This means that the results are  
521 subject to large sampling variability. Also, it is known that svPPA and nfvPPA present  
522 different characteristic MRI patterns, and CSF levels (NfL and 14-3-3) are both increased  
523 (Abu-Rumeileh et al., 2018; Ljubenkov et al., 2018; Seelaar et al., 2011). Thus, the feature  
524 importance of the MRI features may be influenced by the fact that we study PPA  
525 participants together. Future studies could further explore the subanalyses with the FTD  
526 phenotype subtypes in more detail. Finally, only some participants had NfL and 14-3-3  
527 data available, and the smaller sample size might have impacted the results. Future studies  
528 may replicate our CSF results in large, multicentric cohorts to support the findings. As is  
529 common with many ML algorithms using different data sources, the availability of NfL  
530 and 14-3-3 could introduce a bias towards a subset of participants. In our study, as these  
531 biomarkers are not obtained routinely, the availability of such data depended on whether  
532 we had leftover CSF or if the levels had been determined previously. We did not perform  
533 additional testing on the new participants, but this criterion was the same for both groups,  
534 minimizing the risk of a bias.

## 535 **5. CONCLUSION**

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

540 classification probability scores as an ancillary tool for studying the overlapping results  
541 between FTD and AD and a surrogate estimation for the confidence in the ML diagnosis.

542

#### 543 **Ethical Approval and Consent to participate**

544 The Hospital Clínic de Barcelona Ethics Committee approved the study (HCB  
545 2019/0105), and all participants gave written informed consent.

546

#### 547 **Availability of data**

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

550

#### 551 **Conflict of interest statement**

552 The authors declare that they have no competing interests.

553

#### 554 **Funding**

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

567

#### 568 **Authors' contributions**

569 APM, BT, RSV, and RSL contributed to the study's design. APM, NF, SB, BB, JJP,  
570 ATM, JS, JMA, AA, NB, MB, AL, and RSV contributed to the data acquisition. APM  
571 and BT contributed to the data analyses. APM, BT, NF, AL, RSV, and RSL contributed  
572 to interpreting the data. APM, BT, RSV, and RSL contributed to the draft of the article.

573 NF, SB BB, JJP, ATM, JS, JMA, AA, NB, MB, and AL revised the manuscript critically  
574 for important intellectual content and approved the final version of the manuscript. All  
575 authors contributed to the article and approved the submitted version.

576

## 577 **Acknowledgments**

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

580

## 581 **REFERENCES**

582

- 583 Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J.,  
584 Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for  
585 neuroimaging with scikit-learn. *Frontiers in Neuroinformatics* 8.
- 586 Abu-Rumeileh, S., Capellari, S., Stanzani-Maserati, M., Polischi, B., Martinelli, P.,  
587 Caroppo, P., Ladogana, A., Parchi, P., 2018. The CSF neurofilament light  
588 signature in rapidly progressive neurodegenerative dementias. *Alz Res Therapy*  
589 10, 3. <https://doi.org/10.1186/s13195-017-0331-1>
- 590 Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C.,  
591 Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo,  
592 M.C., Thies, B., Phelps, C.H., 2011. The diagnosis of mild cognitive impairment  
593 due to Alzheimer's disease: Recommendations from the National Institute on  
594 Aging-Alzheimer's Association workgroups on diagnostic guidelines for  
595 Alzheimer's disease. *Alzheimer's and Dementia* 7, 270–279.  
596 <https://doi.org/10.1016/j.jalz.2011.03.008>
- 597 Alcolea, D., Vilaplana, E., Suárez-Calvet, M., Illán-Gala, I., Blesa, R., Clarimón, J.,  
598 Lladó, A., Sánchez-Valle, R., Molinuevo, J.L., García-Ribas, G., Compta, Y.,  
599 Martí, M.J., Piñol-Ripoll, G., Amer-Ferrer, G., Noguera, A., García-Martín, A.,  
600 Fortea, J., Lleó, A., 2017. CSF sAPP $\beta$ , YKL-40, and neurofilament light in  
601 frontotemporal lobar degeneration. *Neurology* 89, 178–188.  
602 <https://doi.org/10.1212/WNL.0000000000004088>
- 603 Antonell, A., Tort-Merino, A., Ríos, J., Balasa, M., Borrego-Écija, S., Auge, J.M.,  
604 Muñoz-García, C., Bosch, B., Falgàs, N., Rami, L., Ramos-Campoy, O.,  
605 Blennow, K., Zetterberg, H., Molinuevo, J.L., Lladó, A., Sánchez-Valle, R.,  
606 2019. Synaptic, axonal damage and inflammatory cerebrospinal fluid  
607 biomarkers in neurodegenerative dementias. *Alzheimer's & Dementia*.  
608 <https://doi.org/10.1016/j.jalz.2019.09.001>
- 609 Avants, B.B., Cook, P.A., Ungar, L., Gee, J.C., Grossman, M., 2010. Dementia induces  
610 correlated reductions in white matter integrity and cortical thickness: A  
611 multivariate neuroimaging study with sparse canonical correlation analysis.  
612 *NeuroImage* 50, 1004–1016. <https://doi.org/10.1016/j.neuroimage.2010.01.041>
- 613 Basheera, S., Sai Ram, M.S., 2019. Convolution neural network-based Alzheimer's  
614 disease classification using hybrid enhanced independent component analysis

615 based segmented gray matter of T2 weighted magnetic resonance imaging with  
616 clinical valuation. *Alzheimer's & Dementia: Translational Research & Clinical*  
617 *Interventions* 5, 974–986. <https://doi.org/10.1016/j.trci.2019.10.001>

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

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

630 Bocchetta, M., Todd, E.G., Peakman, G., Cash, D.M., Convery, R.S., Russell, L.L.,  
631 Thomas, D.L., Iglesias, J.E., Van Swieten, J.C., Jiskoot, L.C., Seelaar, H.,  
632 Borroni, B., Galimberti, D., Sanchez-Valle, R., Laforce, R., Moreno, F.,  
633 Synofzik, M., Graff, C., Masellis, M., Tartaglia, M.C., Rowe, J.B.,  
634 Vandenberghe, R., Finger, E., Tagliavini, F., De Mendonça, A., Santana, I.,  
635 Butler, C.R., Ducharme, S., Gerhard, A., Danek, A., Levin, J., Otto, M., Sorbi,  
636 S., Le Ber, I., Pasquier, F., Rohrer, J.D., 2021. Differential early subcortical  
637 involvement in genetic FTD within the GENFI cohort. *NeuroImage: Clinical* 30,  
638 102646. <https://doi.org/10.1016/j.nicl.2021.102646>

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

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

- 665 Barkhof, F., Schouten, T.M., de Vos, F., Feis, R.A., van der Grond, J., de Rooij,  
666 M., Rombouts, S.A.R.B., 2018. Single Subject Classification of Alzheimer's  
667 Disease and Behavioral Variant Frontotemporal Dementia Using Anatomical,  
668 Diffusion Tensor, and Resting-State Functional Magnetic Resonance Imaging.  
669 *Journal of Alzheimer's Disease* 62, 1827–1839. [https://doi.org/10.3233/JAD-](https://doi.org/10.3233/JAD-170893)  
670 170893
- 671 Bozeat, S., Gregory, C.A., Ralph, M.A., Hodges, J.R., 2000. Which neuropsychiatric  
672 and behavioural features distinguish frontal and temporal variants of  
673 frontotemporal dementia from Alzheimer's disease? *J Neurol Neurosurg*  
674 *Psychiatry* 69, 178–186. <https://doi.org/10.1136/jnnp.69.2.178>
- 675 Braak, H., Thal, D.R., Ghebremedhin, E., Del Tredici, K., 2011. Stages of the  
676 pathologic process in Alzheimer disease: age categories from 1 to 100 years. *J*  
677 *Neuropathol Exp Neurol* 70, 960–969.  
678 <https://doi.org/10.1097/NEN.0b013e318232a379>
- 679 Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32.  
680 <https://doi.org/10.1023/A:1010933404324>
- 681 Bron, E.E., Klein, S., Papma, J.M., Jiskoot, L.C., Venkatraghavan, V., Linders, J.,  
682 Aalten, P., De Deyn, P.P., Biessels, G.J., Claassen, J.A.H.R., Middelkoop,  
683 H.A.M., Smits, M., Niessen, W.J., van Swieten, J.C., van der Flier, W.M.,  
684 Ramakers, I.H.G.B., van der Lugt, A., 2021. Cross-cohort generalizability of  
685 deep and conventional machine learning for MRI-based diagnosis and prediction  
686 of Alzheimer's disease. *NeuroImage: Clinical* 31, 102712.  
687 <https://doi.org/10.1016/j.nicl.2021.102712>
- 688 Bron, E.E., Smits, M., Papma, J.M., Steketee, R.M.E., Meijboom, R., de Groot, M., van  
689 Swieten, J.C., Niessen, W.J., Klein, S., 2017. Multiparametric computer-aided  
690 differential diagnosis of Alzheimer's disease and frontotemporal dementia using  
691 structural and advanced MRI. *European Radiology* 27, 3372–3382.  
692 <https://doi.org/10.1007/s00330-016-4691-x>
- 693 Canu, E., Agosta, F., Mandic-Stojmenovic, G., Stojković, T., Stefanova, E., Inuggi, A.,  
694 Imperiale, F., Copetti, M., Kostic, V.S., Filippi, M., 2017. Multiparametric MRI  
695 to distinguish early onset Alzheimer's disease and behavioural variant of  
696 frontotemporal dementia. *NeuroImage: Clinical* 15, 428–438.  
697 <https://doi.org/10.1016/j.nicl.2017.05.018>
- 698 Chagué, P., Marro, B., Fadili, S., Houot, M., Morin, A., Samper-González, J., Beunon,  
699 P., Arrivé, L., Dormont, D., Dubois, B., Teichmann, M., Epelbaum, S., Colliot,  
700 O., 2021. Radiological classification of dementia from anatomical MRI assisted  
701 by machine learning-derived maps. *Journal of Neuroradiology* 48, 412–418.  
702 <https://doi.org/10.1016/j.neurad.2020.04.004>
- 703 Chen, Y., Sha, M., Zhao, X., Ma, J., Ni, H., Gao, W., Ming, D., 2017. Automated  
704 detection of pathologic white matter alterations in Alzheimer's disease using  
705 combined diffusivity and kurtosis method. *Psychiatry Research: Neuroimaging*  
706 264, 35–45. <https://doi.org/10.1016/j.psychresns.2017.04.004>
- 707 Contador, J., Pérez-Millan, A., Guillén, N., Sarto, J., Tort-Merino, A., Balasa, M.,  
708 Falgàs, N., Castellví, M., Borrego-Écija, S., Juncà-Parella, J., Bosch, B.,  
709 Fernández-Villullas, G., Ramos-Campoy, O., Antonell, A., Bargalló, N.,  
710 Sanchez-Valle, R., Sala-Llonch, R., Lladó, A., 2022. Sex differences in early-  
711 onset Alzheimer's disease. *European Journal of Neurology* 29, 3623–3632.  
712 <https://doi.org/10.1111/ene.15531>
- 713 Contador, J., Pérez-Millán, A., Tort-Merino, A., Balasa, M., Falgàs, N., Olives, J.,  
714 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., Lehericy, 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](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-](https://doi.org/10.1038/s41598-023-43706-6)  
745 [43706-6](https://doi.org/10.1038/s41598-023-43706-6)
- 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-](https://doi.org/10.1016/S0197-4580(01)00271-8)  
755 [4580\(01\)00271-8](https://doi.org/10.1016/S0197-4580(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., Fernández-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



815 aphasia and its variants. *Neurology* 76, 1006–14.  
816 <https://doi.org/10.1212/WNL.0b013e31821103e6>

817 Harris, J.M., Thompson, J.C., Gall, C., Richardson, A.M.T., Neary, D., du Plessis, D.,  
818 Pal, P., Mann, D.M.A., Snowden, J.S., Jones, M., 2015. Do NIA-AA criteria  
819 distinguish Alzheimer’s disease from frontotemporal dementia? *Alzheimer’s &*  
820 *Dementia* 11, 207–215. <https://doi.org/10.1016/j.jalz.2014.04.516>

821 Hodges, J.R., Patterson, K., 2007. Semantic dementia: a unique clinicopathological  
822 syndrome. *The Lancet Neurology* 6, 1004–1014. [https://doi.org/10.1016/S1474-](https://doi.org/10.1016/S1474-4422(07)70266-1)  
823 [4422\(07\)70266-1](https://doi.org/10.1016/S1474-4422(07)70266-1)

824 Jack, C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B.,  
825 Holtzman, D.M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J.L.,  
826 Montine, T., Phelps, C., Rankin, K.P., Rowe, C.C., Scheltens, P., Siemers, E.,  
827 Snyder, H.M., Sperling, R., Elliott, C., Masliah, E., Ryan, L., Silverberg, N.,  
828 2018. NIA-AA Research Framework: Toward a biological definition of  
829 Alzheimer’s disease. *Alzheimer’s and Dementia* 14, 535–562.  
830 <https://doi.org/10.1016/j.jalz.2018.02.018>

831 Jack, C.R., Petersen, R.C., Xu, Y., O’Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F.,  
832 Tangalos, E.G., Kokmen, E., 2000. Rates of hippocampal atrophy correlate with  
833 change in clinical status in aging and AD. *Neurology* 55, 484–489.  
834 <https://doi.org/10.1212/wnl.55.4.484>

835 Javeed, A., Dallora, A.L., Berglund, J.S., Ali, A., Ali, L., Anderberg, P., 2023. Machine  
836 Learning for Dementia Prediction: A Systematic Review and Future Research  
837 Directions. *J Med Syst* 47, 17. <https://doi.org/10.1007/s10916-023-01906-7>

838 Johnson, A.S., Ziaggi, G., Smith, A.C., Houlihan, H., Heuer, L.B., Guzmán, D.S.,  
839 Okafor, A., Huey, E.D., Talmasov, D., Provenzano, F., Kreisl, W.C., Lao, P.J.,  
840 Alzheimer’s Disease Neuroimaging Initiative, 2024. Psychotic symptoms are  
841 associated with elevated tau PET signal in the amygdala independent of  
842 Alzheimer’s disease clinical severity and amyloid burden. *medRxiv*  
843 2024.01.12.24301221. <https://doi.org/10.1101/2024.01.12.24301221>

844 Kim, J.P., Kim, J., Park, Y.H., Park, S.B., Lee, J.S., Yoo, S., Kim, E.-J., Kim, H.J., Na,  
845 D.L., Brown, J.A., Lockhart, S.N., Seo, S.W., Seong, J.-K., 2019. Machine  
846 learning based hierarchical classification of frontotemporal dementia and  
847 Alzheimer’s disease. *NeuroImage: Clinical* 23, 101811.  
848 <https://doi.org/10.1016/j.nicl.2019.101811>

849 Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox,  
850 N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Automatic  
851 classification of MR scans in Alzheimer’s disease. *Brain* 131, 681–689.  
852 <https://doi.org/10.1093/brain/awm319>

853 Koedam, E.L.G.E., Lauffer, V., van der Vlies, A.E., van der Flier, W.M., Scheltens, P.,  
854 Pijnenburg, Y.A.L., 2010. Early-Versus Late-Onset Alzheimer’s Disease: More  
855 than Age Alone. *Journal of Alzheimer’s Disease* 19, 1401–1408.  
856 <https://doi.org/10.3233/JAD-2010-1337>

857 Lampe, L., Huppertz, H.-J., Anderl-Straub, S., Albrecht, F., Ballarini, T., Bisenius, S.,  
858 Mueller, K., Niehaus, S., Fassbender, K., Fliessbach, K., Jahn, H., Kornhuber, J.,  
859 Lauer, M., Prudlo, J., Schneider, A., Synofzik, M., Kassubek, J., Danek, A.,  
860 Villringer, A., Diehl-Schmid, J., Otto, M., Schroeter, M.L., FTL D Consortium  
861 Germany, 2023. Multiclass prediction of different dementia syndromes based on  
862 multi-centric volumetric MRI imaging. *Neuroimage Clin* 37, 103320.  
863 <https://doi.org/10.1016/j.nicl.2023.103320>

864 Li, B., Jang, I., Riphagen, J., Almaktoum, R., Yochim, K.M., Ances, B.M.,

- 865 Bookheimer, S.Y., Salat, D.H., 2021. Identifying individuals with Alzheimer's  
866 disease-like brains based on structural imaging in the Human Connectome  
867 Project Aging cohort. *Human Brain Mapping* 42, 5535–5546.  
868 <https://doi.org/10.1002/hbm.25626>
- 869 Lin, W., Tong, T., Gao, Q., Guo, D., Du, X., Yang, Y., Guo, G., Xiao, M., Du, M., Qu,  
870 X., The Alzheimer's Disease Neuroimaging Initiative, 2018. Convolutional  
871 Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease  
872 Prediction From Mild Cognitive Impairment. *Frontiers in Neuroscience* 12.
- 873 Ljubenkov, P.A., Staffaroni, A.M., Rojas, J.C., Allen, I.E., Wang, P., Heuer, H.,  
874 Karydas, A., Kornak, J., Cobigo, Y., Seeley, W.W., Grinberg, L.T., Spina, S.,  
875 Fagan, A.M., Jerome, G., Knopman, D., Boeve, B.F., Dickerson, B.C., Kramer,  
876 J., Miller, B., Boxer, A.L., Rosen, H.J., 2018. Cerebrospinal fluid biomarkers  
877 predict frontotemporal dementia trajectory. *Annals of Clinical and Translational*  
878 *Neurology* 5, 1250–1263. <https://doi.org/10.1002/acn3.643>
- 879 Magnin, B., Mesrob, L., Kinkingnéhun, S., Péligrini-Issac, M., Colliot, O., Sarazin, M.,  
880 Dubois, B., Lehericy, S., Benali, H., 2009. Support vector machine-based  
881 classification of Alzheimer's disease from whole-brain anatomical MRI.  
882 *Neuroradiology* 51, 73–83. <https://doi.org/10.1007/s00234-008-0463-x>
- 883 Maito, M.A., Santamaría-García, H., Moguilner, S., Possin, K.L., Godoy, M.E., Avila-  
884 Funes, J.A., Behrens, M.I., Brusco, I.L., Bruno, M.A., Cardona, J.F., Custodio,  
885 N., García, A.M., Javandel, S., Lopera, F., Matallana, D.L., Miller, B., Oliveira,  
886 M.O. de, Pina-Escudero, S.D., Slachevsky, A., Ortiz, A.L.S., Takada, L.T.,  
887 Tagliazuchi, E., Valcour, V., Yokoyama, J.S., Ibañez, A., 2023. Classification of  
888 Alzheimer's disease and frontotemporal dementia using routine clinical and  
889 cognitive measures across multicentric underrepresented samples: a cross  
890 sectional observational study. *The Lancet Regional Health – Americas* 17.  
891 <https://doi.org/10.1016/j.lana.2022.100387>
- 892 Mateos-Pérez, J.M., Dadar, M., Lacalle-Aurióles, M., Iturria-Medina, Y., Zeighami, Y.,  
893 Evans, A.C., 2018. Structural neuroimaging as clinical predictor: A review of  
894 machine learning applications. *NeuroImage: Clinical* 20, 506–522.  
895 <https://doi.org/10.1016/j.nicl.2018.08.019>
- 896 McCarthy, J., Collins, D.L., Ducharme, S., 2018. Morphometric MRI as a diagnostic  
897 biomarker of frontotemporal dementia: A systematic review to determine  
898 clinical applicability. *NeuroImage: Clinical* 20, 685–696.  
899 <https://doi.org/10.1016/j.nicl.2018.08.028>
- 900 McFerrin, M.B., Chi, X., Cutter, G., Yacoubian, T.A., 2017. Dysregulation of 14-3-3  
901 proteins in neurodegenerative diseases with Lewy body or Alzheimer pathology.  
902 *Ann Clin Transl Neurol* 4, 466–477. <https://doi.org/10.1002/acn3.421>
- 903 McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas,  
904 C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C.,  
905 Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub,  
906 S., Phelps, C.H., 2011. The diagnosis of dementia due to Alzheimer's disease:  
907 Recommendations from the National Institute on Aging-Alzheimer's  
908 Association workgroups on diagnostic guidelines for Alzheimer's disease.  
909 *Alzheimer's and Dementia* 7, 263–269.  
910 <https://doi.org/10.1016/j.jalz.2011.03.005>
- 911 Mendez, M.F., 2006. The accurate diagnosis of early-onset dementia. *International*  
912 *Journal of Psychiatry in Medicine* 36, 401–412. [https://doi.org/10.2190/Q6J4-](https://doi.org/10.2190/Q6J4-R143-P630-KW41)  
913 [R143-P630-KW41](https://doi.org/10.2190/Q6J4-R143-P630-KW41)
- 914 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., Dopfer, 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., Dopfer, 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., Onofri, 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., Drueh, 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.

- 1015 Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a  
 1016 review. *Journal of Neurology, Neurosurgery & Psychiatry* 82, 476–486.  
 1017 <https://doi.org/10.1136/jnnp.2010.212225>
- 1018 Seidman, L.J., Faraone, S.V., Goldstein, J.M., Goodman, J.M., Kremen, W.S., Matsuda,  
 1019 G., Hoge, E.A., Kennedy, D., Makris, N., Caviness, V.S., Tsuang, M.T., 1997.  
 1020 Reduced subcortical brain volumes in nonpsychotic siblings of schizophrenic  
 1021 patients: A pilot magnetic resonance imaging study. *American Journal of*  
 1022 *Medical Genetics - Neuropsychiatric Genetics* 74, 507–514.  
 1023 [https://doi.org/10.1002/\(SICI\)1096-8628\(19970919\)74:5<507::AID-](https://doi.org/10.1002/(SICI)1096-8628(19970919)74:5<507::AID-AJMG11>3.0.CO;2-G)  
 1024 [AJMG11>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1096-8628(19970919)74:5<507::AID-AJMG11>3.0.CO;2-G)
- 1025 Swift, I.J., Sogorb-Esteve, A., Heller, C., Synofzik, M., Otto, M., Graff, C., Galimberti,  
 1026 D., Todd, E., Heslegrave, A.J., van der Ende, E.L., Van Swieten, J.C.,  
 1027 Zetterberg, H., Rohrer, J.D., 2021. Fluid biomarkers in frontotemporal dementia:  
 1028 past, present and future. *J Neurol Neurosurg Psychiatry* 92, 204–215.  
 1029 <https://doi.org/10.1136/jnnp-2020-323520>
- 1030 Tort-Merino, A., Falgàs, N., Allen, I.E., Balasa, M., Olives, J., Contador, J., Castellví,  
 1031 M., Juncà-Parella, J., Guillén, N., Borrego-Écija, S., Bosch, B., Fernández-  
 1032 Villullas, G., Ramos-Campoy, O., Antonell, A., Rami, L., Sánchez-Valle, R.,  
 1033 Lladó, A., 2022. Early-onset Alzheimer’s disease shows a distinct  
 1034 neuropsychological profile and more aggressive trajectories of cognitive decline  
 1035 than late-onset. *Ann Clin Transl Neurol.* <https://doi.org/10.1002/acn3.51689>
- 1036 Wang, J., Redmond, S.J., Bertoux, M., Hodges, J.R., Hornberger, M., 2016. A  
 1037 Comparison of Magnetic Resonance Imaging and Neuropsychological  
 1038 Examination in the Diagnostic Distinction of Alzheimer’s Disease and  
 1039 Behavioral Variant Frontotemporal Dementia. *Frontiers in Aging Neuroscience*  
 1040 8.
- 1041 Wang, S.-H., Phillips, P., Sui, Y., Liu, B., Yang, M., Cheng, H., 2018. Classification of  
 1042 Alzheimer’s Disease Based on Eight-Layer Convolutional Neural Network with  
 1043 Leaky Rectified Linear Unit and Max Pooling. *J Med Syst* 42, 85.  
 1044 <https://doi.org/10.1007/s10916-018-0932-7>
- 1045 Whitwell, J.L., Clifford, R.J., Przybelski, S.A., Parisi, J.E., Senjem, M.L., Boeve, B.F.,  
 1046 Knopman, D.S., Petersen, R.C., Dickson, D.W., Josephs, K.A., 2011.  
 1047 Temporoparietal atrophy: a marker of AD pathology independent of clinical  
 1048 diagnosis. *Neurobiol Aging* 32, 1531–1541.  
 1049 <https://doi.org/10.1016/j.neurobiolaging.2009.10.012>
- 1050 Wojtas, A., Heggeli, K.A., Finch, N., Baker, M., DeJesus-Hernandez, M., Younkin,  
 1051 S.G., Dickson, D.W., Graff-Radford, N.R., Rademakers, R., 2012. C9ORF72  
 1052 repeat expansions and other FTD gene mutations in a clinical AD patient series  
 1053 from Mayo Clinic. *Am J Neurodegener Dis* 1, 107–118.
- 1054 Yang, Z., Kinney, J.W., Cordes, D., The Alzheimer’s Disease Neuroimaging Initiative,  
 1055 null, 2024. Uptake of 18F-AV45 in the Putamen Provides Additional Insights  
 1056 into Alzheimer’s Disease beyond the Cortex. *Biomolecules* 14, 157.  
 1057 <https://doi.org/10.3390/biom14020157>
- 1058 Zee, J. van der, Slegers, K., Broeckhoven, C.V., 2008. Invited Article: The Alzheimer  
 1059 disease–frontotemporal lobar degeneration spectrum. *Neurology* 71, 1191–1197.  
 1060 <https://doi.org/10.1212/01.wnl.0000327523.52537.86>
- 1061

## 1062 FIGURES & TABLES

1063 Table 1: Group summaries written as each measure's mean and standard deviation. We calculated differences  
 1064 between groups using Fisher Test for sex or the Anova Test for the rest of the variables. We highlighted the  
 1065 significant group differences in bold. We measured pairwise differences with a Benjamini-Hochberg correction p-  
 1066 value. CTR: healthy participants, AD: Alzheimer's disease, FTD: frontotemporal dementia, NfL: neurofilament light  
 1067 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	<b>0.049</b>	<b>0.022</b>
Age at MRI, years (SD)	59.4 (15.0)	65.0 (9.9)	63.7 (8.3)	<b>1.3e-5</b>	<b>0.0045</b>	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)	<b>1.2e-07</b>	<b>&lt; 2e-16</b>	<b>5.9e-06</b>
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)	<b>&lt; 2e-16</b>	<b>3.0e-06</b>	<b>5.9e-06</b>

1068

1069

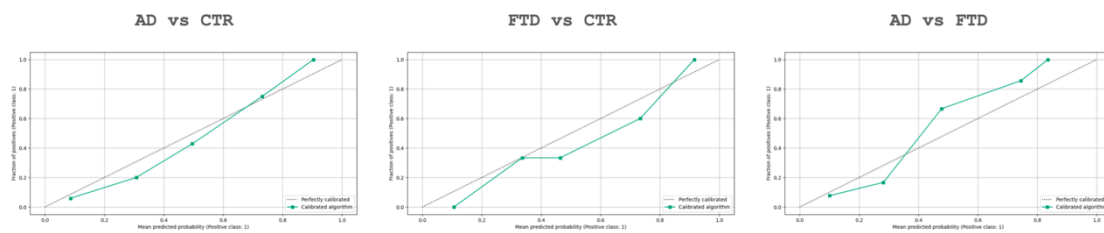
1070 Table 2: Classification performance of the different approaches and the percentage of participants with a higher  
 1071 probability of 80% in the diagnosis grouped by diagnosis.

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

<b>MRI reduced data</b> (N=178)	Accuracy: 88.5%	Accuracy: 85.6%	Accuracy: 84.6%
	AD: 67.2%	FTD: 68.1%	AD: 53.2%
	CTR: 55.3%	CTR: 54.3%	FTD: 54.4%
<b>CSF data</b> (N=178)	Accuracy: 93.0%	Accuracy: 86.6%	Accuracy: 83.8%
	AD: 72.1%	FTD: 72.1%	AD: 40.6%
	CTR: 71.7%	CTR: 23.5%	FTD: 45.9%
<b>MRI and CSF data</b> (N=178)	Accuracy: 90.3%	Accuracy: 86.5%	Accuracy: 88.5%
	AD: 68.1%	FTD: 70.8%	AD: 60.7%
	CTR: 64.4%	CTR: 53.2%	FTD: 55.1%

1072

1073



1074

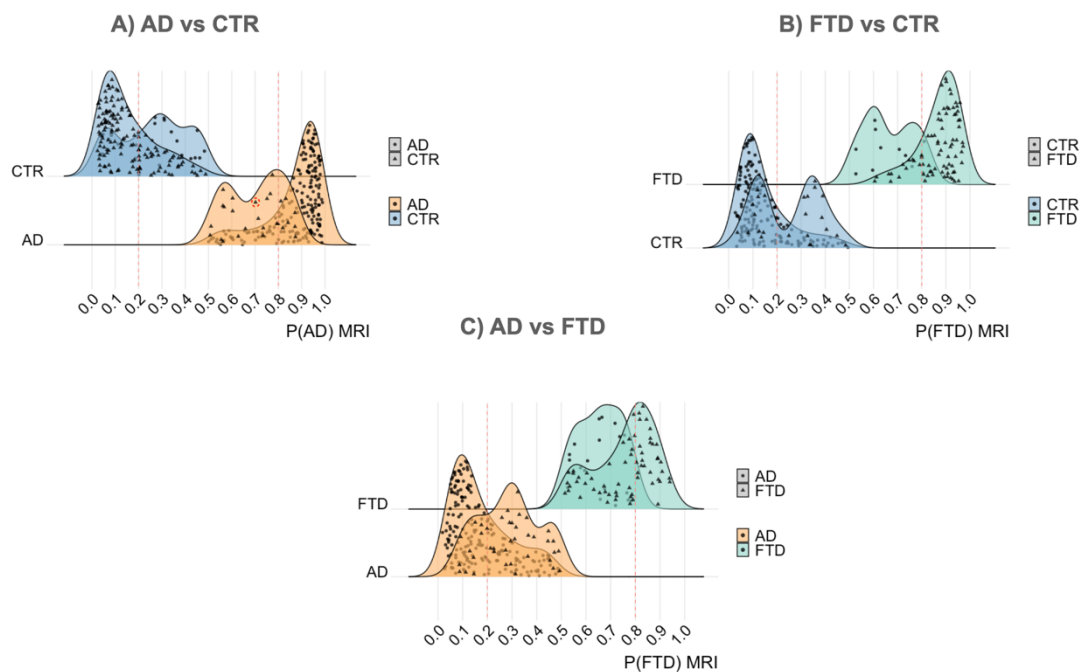
1075 *Figure 1: Calibration plots with the test dataset to evaluate the calibration quality for each classification scenario.*

1076 *We observe that the line plot generated by mean predicted probabilities vs fraction of positiveness (represented with*

1077 *a green plain line) approaches the identity line (here represented as a gray-dashed line), which would indicate a*

1078 *perfect calibration AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls.*

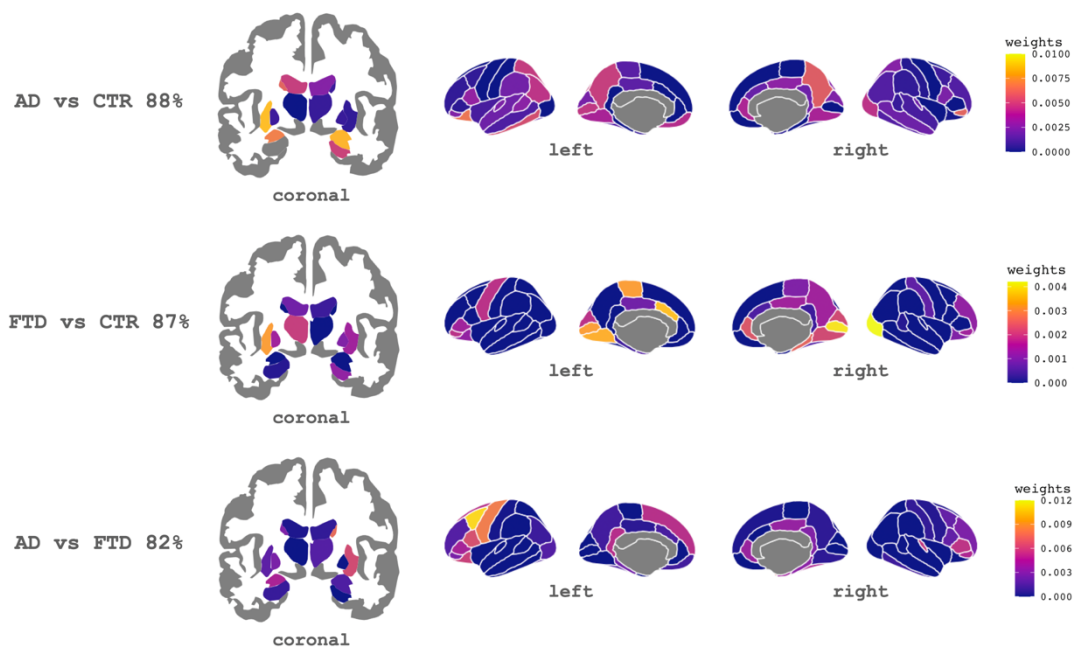
1079



1080

1081 *Figure 2: Density plot to study the obtained individual probabilities with the MRI-based algorithm (all the study*  
 1082 *participants included). The clinical diagnosis is identified with triangles or circles, and the algorithm's diagnoses are*  
 1083 *plotted with different colors. The vertical red dashed lines indicate the thresholds for the grey zone. We highlight a*  
 1084 *random participant (point) with a dashed red circle in the comparison of AD vs CTR for explanatory reasons. This*  
 1085 *point would have a probability of 0.7 of being AD (x-axis) and consequently a probability of 0.2 of being CTR*  
 1086 *(inverse probability). AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls, P:*  
 1087 *probability.*

1088

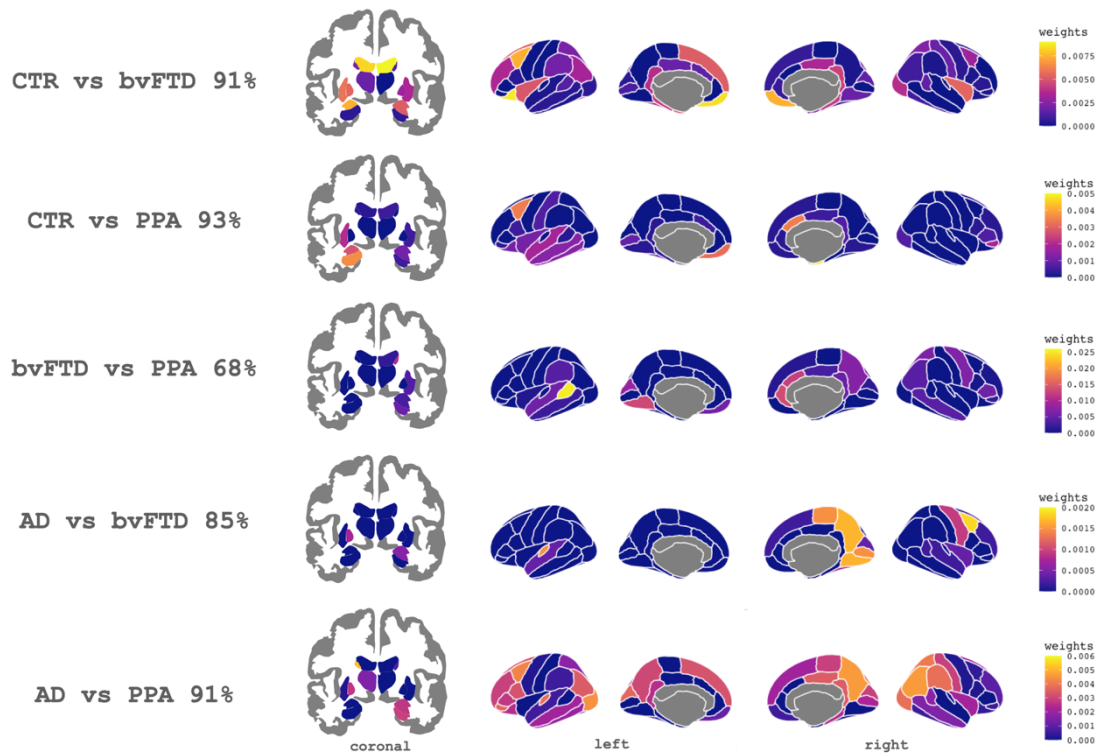


1089



1090 *Figure 3: Cortical and subcortical patterns of the feature importance of each region associated with the different AD*  
 1091 *and FTD comparison (comparison per row). At a higher value major importance of that region for the classification.*  
 1092 *AD: Alzheimer's disease, FTD: frontotemporal dementia, CTR: healthy controls.*

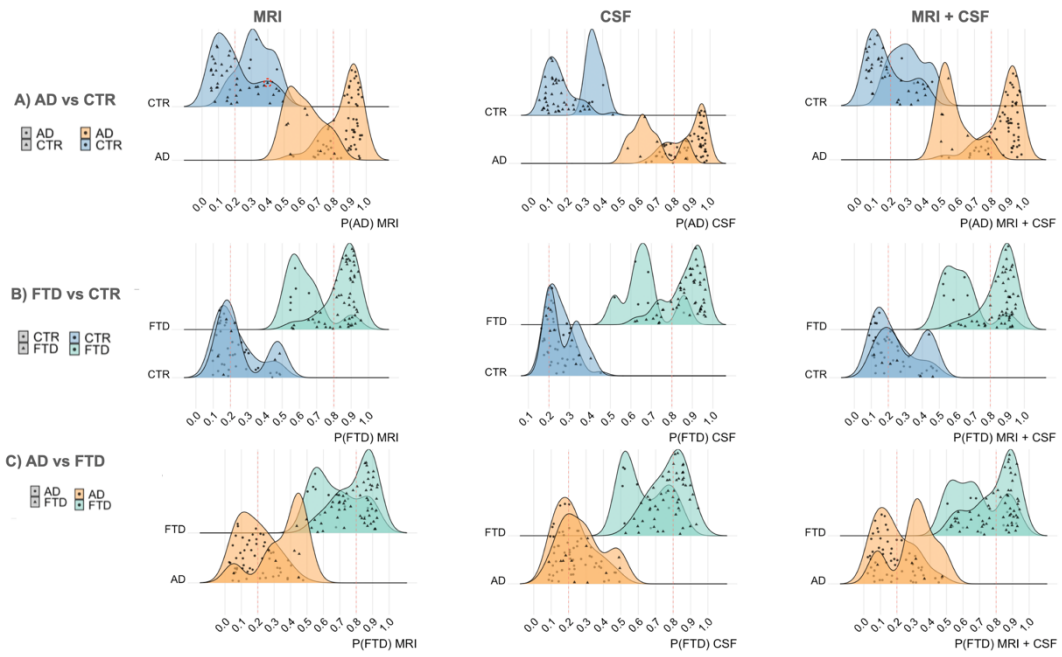
1093



1094

1095 *Figure 4: Cortical and subcortical patterns of the feature importance of each region associated with the different*  
 1096 *bvFTD and PPA comparison (comparison per row). At a higher value major importance of that region for the*  
 1097 *classification. AD: Alzheimer's disease, bvFTD: behavior frontotemporal dementia, PPA: primary*  
 1098 *progressive aphasia, CTR: healthy controls.*

1099



1100

1101 *Figure 5: Density plot to study the obtained individual probabilities with the MRI-based, CSF-based, and MRI- and*  
 1102 *CSF-based algorithms (participants included in all the analyses were those with MRI and CSF data). It can be seen*  
 1103 *in the clinical diagnosis with triangles or circles and the algorithm's diagnosis, plotted with different colors. The*  
 1104 *vertical red dashed lines indicate the thresholds for the grey zone. We highlight a random participant (point) with a*  
 1105 *dashed red circle in the comparison of AD vs CTR for explanatory reasons. This point would have a probability of*  
 1106 *0.4 of being AD (x-axis) and consequently a probability of 0.6 of being CTR (inverse probability). AD: Alzheimer's*  
 1107 *disease, FTD: frontotemporal dementia, CTR: healthy controls, P: probability.*