UNIVERSITAT DE
BARCELONA

Utilitat dels models matemàtics amb neuroimatge pel diagnòstic i l'estudi de la variabilitat en la malaltia d'Alzheimer i la demència frontotemporal

Agnès Pérez Millan

Aquesta tesi doctoral està subjecta a la llicència ***Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.***

Esta tesis doctoral está sujeta a la licencia ***Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative* Commons.**

This doctoral thesis is licensed under the ***Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.***

Utilitat dels models matemàtics amb neuroimatge pel diagnòstic i l'estudi de la variabilitat en la malaltia d'Alzheimer i la demència frontotemporal

Tesi presentada per:

Agnès Pérez Millan

Per optar al grau de doctora per la Universitat de Barcelona

Programa de Biomedicina

Facultat de Medicina i Ciències de la Salut

Dirigida per:

Roser Sala Llonch i Raquel Sánchez del Valle Díaz

Barcelona, 2023

A la meva àvia Concepció

Barcelona, 16 d'octubre de 2023

La Dr. Roser Sala Llonch, doctora en Medicina per la Universitat de Barcelona i la Dra. Raquel Sánchez del Valle Díaz, doctora en Medicina per la Universitat de Barcelona,

CERTIFIQUEN

Que la memòria titulada *"Utilitat dels models matemàtics amb neuroimatge pel diagnòstic i l'estudi de la variabilitat en la malaltia d'Alzheimer i la demència frontotemporal",* presentada per Agnès Pérez Millan, ha estat realitzada sota la nostra direcció i considerem que reuneix les condicions necessàries per ser defensada davant el Tribunal corresponent per optar al Grau de Doctor.

Signat,

Dra. Roser Sala Llonch	Dra. Raquel Sánchez-Valle
Unitat de Biofísica i Bioenginyeria,	*Unitat d'Alzheimer i altres*
Departament de Biomedicina,	*trastorns cognitius,*
Universitat de Barcelona	Servei de Neurologia
	Hospital Clínic de Barcelona

AGRAÏMENTS

Un cop arriba el moment d'escriure els agraïments és quan t'adones que una etapa més s'ha acabat, una etapa que es podria descriure com una muntanya russa molt enriquidora, però a la vegada molt dura. Voldria començar agraint a totes les persones que han col·laborat en aquesta tesi o han fet possible que jo em decidís per fer-la, des dels meus mestres de l'escola fins a les meves directores de tesi passant per la meva família i amics, perquè sense tots ells no hauria estat possible. El procés de la tesi en si ha estat realitzada gràcies a l'ajuda i el suport de moltes persones, però en especial voldria transmetre el meu agraïment:

M'agradaria començar per les meves directores de tesi, la Roser i la Raquel, ja que són les persones que han lluitat tant com jo perquè tires endavant aquesta tesi. Gràcies per l'ajuda, la guia, l'esforç, les hores de discussió, els consells i pels coneixements que m'emporto tant científicament com en l'àmbit personal. Gràcies per la confiança dipositada en mi, gràcies per la llibertat donada durant aquests anys i gràcies per enriquir-me com a persona. Resumint, aquesta tesi no hauria estat possible sense cap de les dues.

Bea, Jordi i José sense vosaltres la tesi no hauria estat possible, la vostra ajuda i disposició ha estat sempre present. M'hauria estat impossible recopilar tantes imatges, comprovar diagnòstics i consultar dubtes a mi sola. En especial a la Bea, gràcies per deixar-me robar el teu ordinador per descarregar imatges per ser el més ràpid i per les estones asseguda al meu costat mirant que tots els pacients tinguessin diagnòstic i el nombre de visites correcte. Parlant d'imatges, no em puc oblidar de la meva directora de tesi, la Roser que em va animar i ajudar a fer un doctorat sobre imatges de ressonància magnètica.

Gràcies a tots els membres de la Unitat d'Alzheimer i altres trastorns cognitius (alguns ja mencionats), inclosa la meva altra directora, que almenys si s'han escandalitzat en alguna pregunta molt bàsica meva, han dissimulat i han respost molt amablement i han tingut una paciència de sant explicant conceptes molt bàsics a una física que no tenia ni idea. Raquel gracias por hacerme entender que por muy maravilloso que parezca numéricamente, no necesariamente tiene que tener sentido clínico. Albert, gràcies per tots els comentaris i discussions que hem tingut al llarg de la tesi i que l'han enriquida. Sergi, gràcies per explicar-me més d'una vegada i dues, els diagnòstics de la DFT. Neus i Núria, gràcies per estar disponibles sempre per poder resoldre dubtes i donar suport en tot. Diana i Jordi, gràcies pel companyarisme i per ser del grup menú infantil. Adrià gràcies pels consells, i explicar-me una mica més sobre les proves neuropsicològiques. Anna gràcies pels conceptes bàsics de genètica que he pogut aprendre al teu costat. Mircea, Guada, Andrea, Yolanda, Marta, Lorena, i a tots gràcies per col·laborar en la tesi, el suport, els dinars i els sopars de joves.

A l'anomenat team "celulitas" (i algun infiltrat) que no compartim recerca, però sí espai. Natalia, Paula, Marta, Berni, Ale, Ignasi, Hector, Marc, Mar, Patri, Viki, Africa, Alicia, Rafa, Esther que us haig de dir. Gràcies per fer aquesta tesi suportable, gràcies pels esmorzar de xafardejos i pels dinars de més xafardeig, pels consells, pels plans i sobretot per ser la via d'escapament quan no em surt res del codi. Acabem la tesi i no heu aconseguit que agafi una pipeta, ara bé heu aconseguit que entri el lab sense por i que entengui alguna cosa del que feu amb les vostres cèl·lules. També gràcies pels consells de com gestionar el doctorat. Carol no ets del team "celulitas" perquè ets team peixera, per estar d'infiltrada, gràcies pels consells i per les converses de divendres tarda quan estem soles.

Many happy memories and people come to my mind when I remember the time spent in Paris. Bertrand, thanks for welcoming me to the group, for your time, and your guidance. Thanks for making everything very easy. Finally, thanks for all the knowledge I learned and brought home.

Mar i Elba, crec que sou les dues persones juntament amb la mama que heu patit molt amb aquesta tesi. Crec que heu estat un pilar per acabar-la i per superar aquest últim any de tesi. Us podria dir moltes coses, però ja estan dites i les sabeu perfectament. Simplement continueu estant presents per molt de temps!

Als meus amics, en especial el Bang Club: Laura's, Pol's, Sílvia, Francesc, Miguel, Ferran i Jöan gràcies per l'amistat i el suport per aconseguir acabar la carrera i continuar amb la tesi. Gràcies per estar sempre presents. Mariona's, Uma, Quim, Víctor's, Marc, Pau, Pol, Nil, Natàlia, gràcies per tots els moments de desconnexió tan necessaris.

Finalment, a la meva família gràcies per ser-hi sempre donant suport i estima de forma incondicional. Per fer companyia fins altes hores de la nit, per animar-me i no deixar-me caure, però sobretot per l'educació i els valors en la perseverança, la lluita, la constància, i el tracte a les persones que m'heu ensenyat des de ben petita. Tot i que hi ha hagut moments de lluita constant.

No puc acabar sense donar cert reconeixement a la xocolata que ha fet una gran contribució en aquesta tesi sobretot en l'escriptura, sense ella no hauria acabat aquest document.

ÍNDEX

Pròleg 13

Treball acadèmic relacionat 17

Llistat d'abreviatures 21

CAPÍTOL 1. Introducció 23

- 1.1 Impacte global de les demències 23
- 1.2 Malaltia d'Alzheimer 25
- 1.3 Demència frontotemporal 31
- 1.4 Diagnòstic diferencial entre la demència frontotemporal i la malaltia d'Alzheimer 34
- 1.5 Ressonància magnètica en l'estudi de la malaltia d'Alzheimer i la demència frontotemporal 36
- 1.6 Models predictius 40
- 1.7 Aprenentatge automàtic 45
- 1.8 Utilitat de l'aprenentatge automàtic en la clínica 50

CAPÍTOL 2. Hipòtesis i Objectius 53

- 2.1 Hipòtesis 53
- 2.2 Objectius 54

CAPÍTOL 3. Mètodes 57

- 3.1 Participants 57
- 3.2 Adquisició de la ressonància magnètica 60
- 3.3 Processament de les imatges de ressonància magnètica 61

3.4 Biomarcadors de líquid cefalorraquidi i sang 61

3.5 Implementació dels models matemàtics 62

CAPÍTOL 4. Resultats 65

Treball 1 67

Treball 2 85

Treball 3 115

Treball 4 147

Treball 5 175

Treball 6 201

4.1. Informe de les directores 217

4.2. Resum global dels resultats 219

CAPÍTOL 5. Discussió 227

CAPÍTOL 6. Conclusions 239

CAPÍTOL 7. Bibliografia 243

PRÒLEG

Aquesta tesi, presentada per a l'obtenció del títol de Doctor per la Universitat de Barcelona, és el resultat de diferents estudis. Aquesta tesi segueix el format del compendi de publicacions, i inclou quatre articles publicats o acceptats revisats per parells, en els quals la candidata era primera autora o primera autora amb la mateixa contribució (indicat amb *), i dos manuscrits enviats per a la seva publicació com a primera autora. Els sis treballs es presenten en el següent ordre:

1. **Agnès Pérez-Millan,** José Contador, Raúl Tudela, Aida Niñerola-Baizán, Xavier Setoain, Albert Lladó, Raquel Sánchez-Valle, Roser Sala-Llonch. Evaluating the performance of Bayesian and frequentist approaches for longitudinal modeling: application to Alzheimer's disease. Scientific Reports 2022; 2022 Agost 24. DOI:10.1038/s41598-022-18129-4. IF (2022): 4,6; Q2 Multidisciplinary Sciences.
2. **Agnès Pérez-Millan,** Sergi Borrego-Écija, Neus Falgàs, Jordi Juncà-Parella, Beatriz Bosch, Adrià Tort-Merino, Anna Antonell, Nuria Bargalló, Lorena Rami, Mircea Balasa, Albert Lladó, Roser Sala-Llonch, Raquel Sánchez-Valle. Cortical thickness modeling and variability in Alzheimer's disease and frontotemporal dementia. Acceptat en Journal of Neurology. IF (2022): 6,0; Q1 Clinical Neurology.
3. **Agnès Pérez-Millan,** José Contador, Jordi Juncà-Parella, Beatriz Bosch, Laia Borrell, Adrià Tort-Merino, Neus Falgàs, Sergi Borrego-Écija,

Nuria Bargalló, Lorena Rami, Mircea Balasa, Albert Lladó, Raquel Sánchez-Valle, Roser Sala-Llonch. Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. Human Brain Mapping 2023; 2023 Gener 20. DOI: 10.1002/hbm.26205. IF (2022): 4,8; Q1 Neuroimaging.

4. **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, Nuria Bargalló, Mircea Balasa, Albert Lladó, Raquel Sánchez-Valle, Roser Sala-Llonch. Beyond group classification: probabilistic differential diagnosis of frontotemporal dementia and Alzheimer's disease with MRI and CSF biomarkers. En revisió.
5. **Agnès Pérez-Millan,** Uma Maria Lal-Trehan Estrada, Neus Falgàs, Núria Guillén, Sergi Borrego-Écija, Beatriz Bosch, Jordi Juncà-Parella, Adrià Tort-Merino, Jordi Sarto, Josep Maria Augé, Anna Antonell, Nuria Bargalló, Raquel Ruiz-García, Laura Naranjo, Mircea Balasa, Albert Lladó, Roser Sala-Llonch, Raquel Sánchez-Valle. The Cortical Asymmetry Index (CAI) for subtyping dementia patients. En revisió.
6. **Agnès Pérez-Millan**[*], Sergi Borrego-Écija[*], John C. van Swieten, Lize Jiskoot, Fermin Moreno, Robert Laforce, Caroline Graff, Mario Masellis, Maria Carmela Tartaglia, James B. Rowe, Barbara Borroni, Elizabeth Finger, Matthis Synofzik, Daniela Galimberti, Rik Vandenberghe, Alexandre de Mendonça, Chris R. Butler, Alexander Gerhard, Simon Ducharme, Isabelle Le Ber, Isabel Santana, Florence Pasquier, Johannes

Levin, Markus Otto, Sandro Sorbi, Pietro Tiraboschi, Harro Seelaar, Tobias Langheinrich, Jonathan D. Rohrer, Roser Sala-Llonch, Raquel Sánchez-Valle, on behalf of the Genetic FTD Initiative, GENFI. Loss of brainstem white matter predicts onset and motor neuron symptoms in *C9orf72* expansion carriers: a GENFI study. Journal of Neurology 2023; 2023 Març 1. DOI: 10.1007/s00415-022-11435-x. IF (2022): 6,0; Q1 Clinical Neurology.

TREBALL ACADÈMIC RELACIONAT

Llistat de publicacions addicionals de la candidata que no estan incloses en la tesi. Aquests treballs són el resultat del treball col·laboratiu amb altres projectes durant el temps de la tesi:

1. Sergi Borrego-Écija*, **Agnès Pérez-Millan***, Anna Antonell, Laura Fort, Elif Kaya-Tilki, Alberto León-Halcón, Albert Lladó, Laura Molina-Porcel, Mircea Balasa, Jordi Juncà-Parella, Javier Vitorica, Jose Luis Venero, Tomas Deierborg, Antonio Boza-Serrano, Raquel Sánchez-Valle. Galectin-3 is upregulated in Frontotemporal dementia patients with subtype specificity. Acceptat en Alzheimer's & Dementia. IF (2022): 14,0; Q1.
2. Gonzalo Forno, José Contador, **Agnès Pérez-Millan**, Nuria Guillen, Neus Falgas, Jordi Sarto, Adrià Tort- Merino, Magdalena Castellví, Beatriz Bosch, Guadalupe Fernández-Villullas, Mircea Balasa, Anna Antonell, Roser Sala Llonch, Raquel Sanchez-Valle, Roser Sala-Llonch, Michael Hornberger, Albert Lladó. The APOE4 effect: Structural brain differences in Alzheimer's disease according to the age at symptom onset. European Journal of Neurology, 2023; 2023 Febrer 8. DOI:10.1111/ene.15657. IF (2022): 5,1; Q1.
3. Lorena Rami, María León, Natalia Valech, Nina Coll-Padrós, Beatriz Bosch, Jaume Olives, Ana Salinero, **Agnès Pérez-Millan**, José Luis Molinuevo, Raquel Sánchez-Valle, Adrià Tort-Merino. Design and

validation of the 1-week memory battery for assessing episodic memory and accelerated long-term forgetting in cognitively unimpaired subjects. Neuropsychology, 2023; 2023 Febrer 2. DOI:10.1037/neu0000879. IF (2022): 2,1; Q2, Psychology.

4. José Contador*, **Agnès Pérez-Millan***, Nuria Guillen, Jordi Sarto, Adrià Tort-Merino, Mircea Balasa, Neus Falgàs, Magdalena Castellví, Sergi Borrego-Écija, Jordi Juncà-Parella, Beatriz Bosch, Guadalupe Fernández-Villullas, Oscar Ramos-Campoy, Anna Antonell, Nuria Bargalló, Raquel Sanchez-Valle, Roser Sala-Llonch, Albert Lladó. Sex differences in early-onset Alzheimer's disease. European Journal of Neurology, 2022; 2022 Novembre 2. DOI:10.1111/ene.1553. IF (2022): 5,1; Q1.
5. Oscar Ramos-Campoy, Albert Lladó, Beatriz Bosch, Mireia Ferrer, **Agnès Pérez-Millan**, Miguel Vergara, Laura Molina-Porcel, Laura Fort-Aznar, Ricardo Gonzalo, Fermín Moreno-Izco, Guadalupe Fernández-Villullas, Mircea Balasa, Raquel Sanchez-Valle, Anna Antonell. Differential Gene Expression in Sporadic and Genetic Forms of Alzheimer's Disease and Frontotemporal Dementia in Brain Tissue and Lymphoblastoid Cell Lines. Molecular Neurobiology. 2022; 2022 Octubre 1. DOI: 10.1007/s12035-022-02969-2. IF (2022): 5,1; Q1.
6. José Contador, **Agnès Pérez-Millan**, Nuria Guillen, Adrià Tort-Merino, Mircea Balasa, Neus Falgàs, Jaume Olives, Magdalena Castellví, Sergi Borrego-Écija, Beatriz Bosch, Guadalupe Fernández-Villullas, Oscar Ramos- Campoy, Anna Antonell, Nuria Bargalló, Raquel Sanchez-Valle, Roser Sala-Llonch, Albert Lladó. Baseline MRI atrophy predicts 2-year cognitive outcomes in early-onset Alzheimer's disease. Journal of

neurology 2022; 2022 Maig 1. DOI:10.1007/s00415-021-10851-9. IF (2022): 6,0; Q1.

7. José Contador, **Agnès Pérez-Millan**, Adrià Tort-Merino, Mircea Balasa, Neus Falgàs, Jaume Olives, Magdalena Castellví, Sergi Borrego-Écija, Beatriz Bosch, Guadalupe Fernández-Villullas, Oscar Ramos-Campoy, Anna Antonell, Nuria Bargalló, Raquel Sanchez-Valle, Roser Sala-Llonch, Albert Lladó, Alzheimer's Disease Neuroimaging Initiative. Longitudinal brain atrophy and CSF biomarkers in early-onset Alzheimer's disease. NeuroImage: Clinical 2021; 2021 Agost 30. DOI:10.1016/j.nicl.2021.102804. IF (2021). 4,9; Q2 Neuroimaging.
8. Neus Falgàs, Mariona Ruiz-Peris, **Agnès Pérez-Millan**, Roser Sala-Llonch, Anna Antonell, Mircea Balasa, Sergi Borrego-Écija, Oscar Ramos-Campoy, Josep Maria Augé, Magdalena Castellví, Adrià Tort-Merino, Jaume Olives, Guadalupe Fernández-Villullas, Kaj Blennow, Henrik Zetterberg, Núria Bargalló, Albert Lladó and Raquel Sánchez-Valle. CSF biomarkers profile for the AD and FTLD signatures. Human Brain Mapping 2020; 2020 Gener 16. DOI:10.1002/hbm.24925. IF (2020): 5,0; Q1.

LLISTAT D'ABREVIATURES

Aβ42	Isofroma de 42 aminoàcids de la proteïna beta-amiloide
ACP	Anàlisi de components principals
ADNI	Alzheimer's Disease Neuroimaging Initiative
AMF	Anàlisi de múltiples factors
ANOVA	Anàlisis de la variància
APP	Afàsies progressives primàries
CAI	Índex d'Asimetria Cortical
CDR	Clinical dementia rating
CTR	Persones sanes
C9orf72	Chromosome 9 open reading frame 72
DCL	Deteriorament cognitiu lleu
DFT	Demència frontotemporal
ELISA	Assaig immunosorbent lligat a enzims
EML	Models d'efectes mixtes lineals
GENFI	Genetic Frontotemporal dementia Initiative
GFAP	Proteïna acídica fibrilar glial
GRN	Gen de la granulina
LCR	Líquid cefalorraquidi
MA	Malaltia d'Alzheimer
MAPT	Gen de la proteïna associada a microtúbuls tau
MMSE	Mini-Mental State Examination

MVS	Màquina vectorial de suport
NfL	Neurofilaments de cadena lleugera
PET	Tomografia per emissió de positrons
PSEN	Gen de la presenilina
p-tau	Tau fosforilada
RM	Ressonància magnètica
RSV	Regressió de suport vectorial
t-tau	Tau total
UCH-L1	Ubiquitin carboxy-terminal hydrolase L1
vcDFT	Variant conductual de DFT
vlAPP	Variant logopènica de APP
vnfAPP	Variant no fluent de APP
vsAPP	Variant semàntica de APP
YKL-40	Proteïna quitinasa-3-similar-1
14-3-3	Proteïna 14-3-3

CAPÍTOL 1

Introducció

1.1 IMPACTE GLOBAL DE LES DEMÈNCIES

El deteriorament cognitiu i la demència són un problema de salut freqüent poblacionalment i generen un alt grau de discapacitat i mortalitat. Actualment, en tot el món hi ha aproximadament 50 milions de persones diagnosticades amb demència i s'espera que aquesta xifra augmenti fins als 131 milions el 2050. El cost de les demències pels sistemes de salut és elevat, en 2015 eren 818 bilions de dòlars a escala mundial, i cada cop ho serà més, s'esperà que en 2050 sigui 1 bilió de dòlars (1).

Clínicament, les demències es caracteritzen per un deteriorament cognitiu i conductual prou greu per afectar l'activitat habitual de les persones que la pateixen. La demència no forma part de l'envelliment normal de les persones, però l'edat és el factor de risc més important per patir una demència, aquest fet fa que la demència sigui un repte global del nostre segle. A mes de l'edat, ser del sexe femení és també un factor no modificable de risc per patir una demència. Les demències neurodegeneratives es consideren malalties altament heretables. Tot i això, només una minoria dels casos es poden considerar que presenten una causa genètica. En aquests casos, la presència d'una mutació és un factor determinant i la malaltia s'hereta amb un patró de transmissió autosòmic dominant.

La paraula demència, no defineix una única malaltia, és una síndrome que engloba diverses malalties, tot i que els símptomes i les comorbiditats se superposen amb freqüència (2). Les malalties neurodegeneratives més comunes que causen demència són la malaltia d'Alzheimer (MA), la demència amb cossos de Lewy, i la demència frontotemporal (DFT) entre d'altres. En molts casos, les característiques clíniques d'aquestes malalties se superposen dificultant el seu diagnòstic, podent conduir a un diagnòstic erroni (3–5). En aquest sentit, es necessiten biomarcadors diagnòstics, és a dir, característiques bioquímiques, de neuroimatge o d'un altre tipus que siguin mesurables i capaces d'identificar el procés patològic.

Biològicament, les demències neurodegeneratives es desenvolupen lentament durant molts anys abans que siguin diagnosticades i fins i tot abans que apareguin els primers símptomes (6), durant el que s'anomena la fase preclínica. Poder detectar que les desencadena ens pot ajudar a desenvolupar intervencions per prevenir-les o endarrerir l'aparició dels símptomes. Durant molts anys es va considerar que la demència no era tractable ni prevenible, però un cop s'ha deixat enrere aquesta creença, s'estan posant molts esforços per trobar tractaments efectius (1). Per aquest motiu, és cada vegada més necessari, identificar biomarcadors que siguin capaços de traçar la seva evolució i la resposta als fàrmacs.

1.2 MALALTIA D'ALZHEIMER

La MA és la demència més freqüent, englobant el 60-70% dels casos de demència al món (6,7). Es tracta d'una malaltia molt llarga, on els canvis histopatològics comencen dècades abans de l'aparició dels primers símptomes (8). La MA es caracteritza clínicament per un deteriorament cognitiu progressiu, on la pèrdua de la memòria és un dels trets principals, però també afecta el llenguatge, les funcions executives i les capacitats visuoespacials (9,10). Aquestes alteracions cognitives s'acompanyen de trastorns conductuals com l'apatia o la irritabilitat entre d'altres. Les alteracions cognitives i conductuals condueixin a una pèrdua de funcionalitat i autonomia progressives i irreversibles actualment.

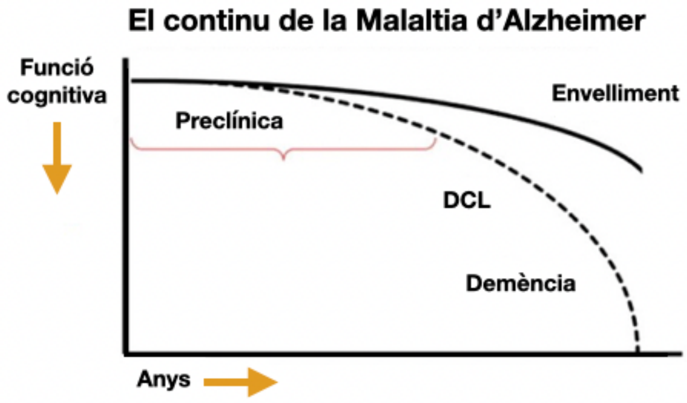

Figura 1: Mostra les fases clíniques que presenta la malaltia d'Alzheimer, s'observa una primera fase preclínica, procedida per una fase amb deteriorament cognitiu lleu (DCL) i finalment la demència. Figura adaptada de (11).

El nom de MA honora el nom del Dr. Alois Alzheimer que en 1906 va descriure el cas, d'Auguste D, una dona de 51 anys, que presentava pèrdua de memòria progressiva, desorientació i al·lucinacions, que la portarà a morir amb 55 anys (12). En aquell moment va ser descrita com una malaltia específica de l'escorça cerebral, ja que era més prima del normal (presentava el que es coneix avui en dia com atròfia). També mostrava una pèrdua neuronal i uns cúmuls anòmals intraneuronals que es van nomenar cabdells neurofibril·lars. Actualment, la MA es defineix pels dipòsits de plaques extracel·lulars d'amiloide i dipòsits intracel·lulars de la proteïna tau hiperfosforilada. L'anomenada "hipòtesi amiloide", encara que discutida per alguns autors, és la teoria més acceptada per explicar el procés fisiopatològic de la malaltia. Aquesta teoria posa com fet clau la diferència entre les formes amiloidogèniques i no amiloidogèniques de la proteïna amiloide que desencadenaria una sèrie de processos que conduirien a la disfunció sinàptica i mort neuronal, entre altres efectes. Seria aquesta disfunció sinàptica i especialment la mort neuronal la que acabarà provocant el deteriorament cognitiu i conductual característic de les persones que pateixen MA.

Els dipòsits de proteïna amiloide i tau hiperfosforilada podrien començar més de 20 anys abans dels primers símptomes, aquesta etapa de la malaltia es coneix com a fase asimptomàtica o preclínica de la malaltia. L'inici dels símptomes, o fase prodròmica de la malaltia es correspon amb un deteriorament cognitiu lleu (DCL), que evolucionarà amb el pas dels anys, a la fase de demència de la malaltia, com es pot observar en la Figura 1 (11,13).

La majoria dels casos de MA tenen una etiologia multifactorial, i no són determinats per alteracions genètiques concretes, i es consideren formes esporàdiques. Tanmateix, hi ha un petit grup de casos, menys de l'1%, que tenen un origen genètic, i per tant es consideren malalties minoritàries. En aquests casos, els pacients amb MA d'origen genètic presenten una mutació genètica que és el factor determinant per desenvolupar la malaltia, aquesta mutació s'hereta amb transmissió autosòmica dominant. Les mutacions més freqüents tenen lloc en el gen presenilina 1 (PSEN1), el gen presenilina 2 (PSEN2) i el gen de la proteïna precursora d'amiloide. L'estudi d'aquestes formes genèticament determinades suposa un model biològic excepcional per estudiar en l'humà les fases més inicials de la malaltia, especialment les fases preclíniques.

Els primers criteris pel diagnòstic clínic de la MA van ser publicats en 1984 (14). El diagnòstic definitiu de la MA restava a l'estudi neuropatològic i el diagnòstic clínic en vida se sostenia amb un quadre clínic compatible i l'exclusió d'altres potencials causes dels símptomes. L'ús d'aquests criteris diagnòstics clínics presentava fins a un 30% d'errors diagnòstics quan es comparaven amb dades neuropatològiques, posant de manifest la necessitat de biomarcadors per donar suport al diagnòstic clínic. Els avenços en el coneixement de la malaltia i el desenvolupament tecnològic en les últimes dècades han canviat la manera de fer el diagnòstic clínic, especialment en les fases primerenques de la malaltia per la incorporació de biomarcadors, ja no d'exclusió d'altres processos sinó amb biomarcadors específics per la MA. En el context del descobriment i identificació de nous biomarcadors, cal tenir en compte que el 1998 es va establir per consens que els biomarcadors específics de la MA ha de tenir una sensibilitat major al 80% per distingir la MA de les persones sanes, i una especificitat superior al 80% també per diferenciar-la d'altres (15). De la mateixa manera, la conceptualització de la MA va canviar des d'una entitat clínica a una entitat clinicobiològica. El 2011, 100 anys aproximadament després de la descripció del primer cas de MA, es van publicar els nous criteris de la MA que inclouen la diferenciació entre la fase DCL i la demència (16,17) i inclouen també per primera vegada la fase preclínica com una part de la MA, encara en absència de símptomes (11). Als últims anys s'han publicat diversos sets de criteris que tenen en comú la inclusió de biomarcadors d'amiloïdosi, de taupatia i de neurodegeneració. Els marcadors actuals acceptats d'amiloïdosi són: els nivells baixos d'Aβ42 o ràtio Aβ42/Aβ40 en líquid cefalorraquidi (LCR), la detecció de plaques d'amiloide amb les imatges de tomografia per emissió de positrons (PET) cerebrals amb traçadors d'amiloide, i més recentment la ràtio Aβ42/Aβ40 o els nivells de ptau181 en plasma. Per altra banda, com marcadors de neurodegeneració s'accepten els nivells elevats de proteïna tau total en LCR, un patró d'atròfia característic en RM o un patró d'hipometabolisme característic en PET con 18F-fluorodexoxiglucosa (8). A més d'aquests biomarcadors ben establerts, altres biomarcadors s'estan estudiant com complementaris per estudiar aspectes com la degeneració de la substància blanca (ex. nivells de cadenes lleugeres de neurofilaments (NfL)), la neuroinflamació i activació astròglial (ex. nivells de la proteïna quitinasa-3-similar-1 (YKL-40)), la pèrdua sinàptica (ex. nivells de proteïna 14-3-3 (14-3-3)), l'activació i proliferació astrocítica (ex. nivells proteïna acídica fibrilar glial (GFAP)), la degradació de proteïnes al proteasoma (ex. nivells d'Ubiquitin carboxy-terminal hydrolase L1 (UCH-L1)), etc. Com es pot veure en la Figura 2, cadascú dels biomarcadors segueix un curs diferent al llarg de la MA, amb diferent utilitat en cadascuna de les fases de la malaltia.

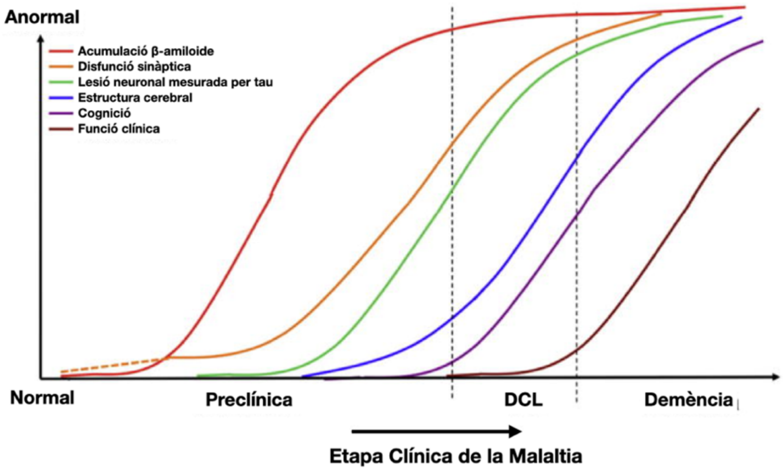

Figura 2: Mostra la trajectòria temporal dels biomarcadors claus de la malaltia d'Alzheimer, per cada una de les fases d'evolució de la malaltia. Figura adaptada de (11,18).

En aquesta tesi, de les diferents tècniques de diagnòstic de la MA, ens centrarem especialment en la RM estructural, que és una tècnica per detectar i avaluar la presència d'atròfia en diferents estructures cerebrals. L'hipocamp és una de les estructures més afectades a conseqüència de la MA, tant precoçment com en estadis més avançats, actualment l'atròfia de l'hipocamp està inclosa en els criteris de diagnòstic de la MA acceptats i prèviament mencionats. Més enllà de l'hipocamp, hi ha uns patrons d'atròfia de substància grisa de les regions corticals ben definits per la MA. Els pacients amb MA presenten una atròfia majoritàriament temporal i parietal posterior en la RM (19–21). Les imatges de RM permeten la detecció dels canvis en la substància grisa en les àrees corticals i subcorticals associades a la MA, fins i tot en etapes prodròmiques (22). A més a més, cal destacar que les imatges de RM permeten fer un seguiment de la progressió de la malaltia, per exemple, les regions com l'hipocamp, l'escorça entorrinal i el precúneus presenten una atròfia que permet fer aquest seguiment amb precisió (23,24). Aquests patrons d'atròfia característics de la MA mostren la utilitat d'ajudar en el seguiment de la malaltia i podrien ser de gran utilitat en el diagnòstic clínic dels diferents subtipus de MA. Tot i això, la seva utilitat en detectar els pacients amb MA en fase preclínica està menys establerta (22). En la Figura 3, es pot observar l'atròfia característica per les persones amb MA i s'observa que és diferent les persones sanes.

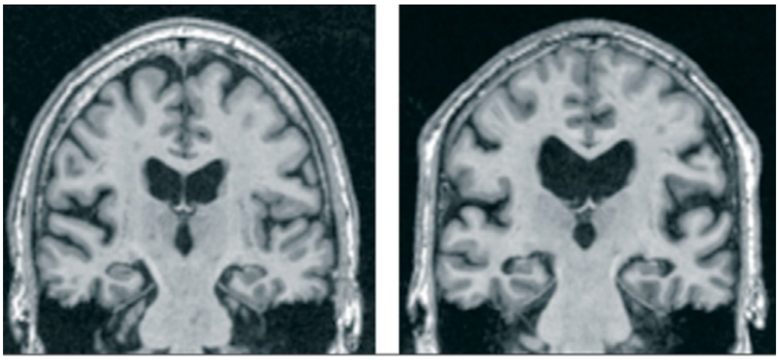

Figura 3: Mostra la ressonància magnètica de cervell de dues persones adultes per observar l'atrofia cerebral causada per la malaltia d'Alzheimer. D'esquerra a dreta, primer trobem un adult sa sense atrofia cerebral, en canvi el segon adult està diagnosticat amb malaltia d'Alzheimer, i mostra atrofia cerebral. Figura adaptada de (18).

1.3 DEMÈNCIA FRONTOTEMPORAL

La DFT és la segona demència neurodegenerativa més freqüent en menors de 65 anys, darrere de la MA, englobant aproximadament el 10% en aquest rang de població dels casos de demència (28–30). Actualment, es considera una malaltia amb una prevalença aproximada entre 15-22 per 100.000 persones, tot i que es creu que podria estar infradiagnosticada i els números podrien ser més elevats (29,31). La DFT és un terme paraigües que engloba un conjunt d'entitats neurodegeneratives que provoquen la pèrdua progressiva de neurones que té lloc principalment en els lòbuls frontal i temporal. La DFT, que ha tingut diversos noms en el transcurs de la història, va ser descrita per primer cop per Arnold Pick en 1892 (32), per això també és coneguda com a Malaltia de Pick. Arnold Pick va descriure diversos pacients que presentaven trastorns de llenguatge i alteracions conductuals amb una marcada atrofia frontal i temporal en el cervell en les seves autòpsies.

La DFT és un trastorn progressiu que es caracteritza a clínicament per canvis conductuals i disfunció cognitiva en què predomina la disfunció executiva i els problemes de llenguatge. És un grup de trastorns clínicament, neuropatològicament i genèticament heterogeni. La DFT d'origen genètic suposa aproximadament un 15% dels casos (5-30% segons les sèries) (33,34). La majoria d'aquests casos són deguts a l'herència autosòmica dominant d'expansions patològiques en el chromosome 9 open reading frame 72 (C9orf72), mutacions en el gen de granulina (GRN) o de la proteïna tau associada a microtúbuls (MAPT) (30). Les formes genètiques de DFT permeten estudiar els canvis que tenen lloc al cervell dels portadors de mutacions patogèniques anys abans de l'inici dels símptomes. En el cas de la DFT d'origen genètic el diagnòstic es pot fer a través d'un test genètic per detectar si la persona és portadora d'algunes de les mutacions que origina la DFT. En el cas de la DFT esporàdica el diagnòstic definitiu requereix l'estudi neuropatològic, que habitualment només es fa post mortem. Pel diagnòstic en vida s'estableixen dues categories diagnòstiques: probable o possible (36). El diagnòstic possible es basa en elements clínics. Un perfil d'atròfia característic amb afectació predominant dels lòbuls frontal i/o temporal en la neuroimatge estructural o bé un hipometabolisme o hipofunció d'aquestes àrees en neuroimatge nuclear dona suport al diagnòstic de DFT probable. Actualment, no hi ha biomarcadors validats del procés neuropatològic subjacent. Bioquímicament, els nivells de NfL en LCR i sang, marcadors de lesió axonal podrien ser útils en la DFT per diferenciar-la de processos psiquiàtrics i per avaluar la seva progressió (37–39). De manera similar a la MA, en la DFT s'ha definit un ordre temporal teòric segons el qual els diferents biomarcadors s'alterarien durant el transcurs de la malaltia, incloent-hi també la fase preclínica, es poden observar en la Figura 4.

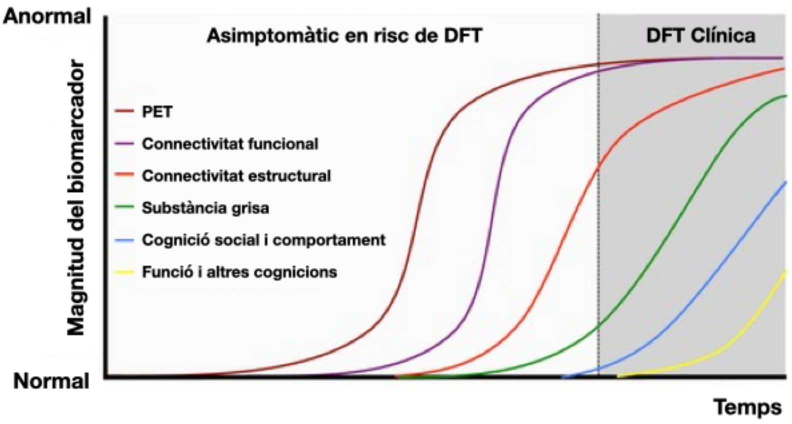

Figura 4: Mostra la trajectòria temporal dels biomarcadors claus de la demència frontotemporal, per cada una de les fases d'evolució de la malaltia. Figura adaptada de (40)

Clínicament, existeixen dos grans subtipus de DFT: la variant conductual (vcDFT), que principalment afecta la conducta i la personalitat de les persones que la pateixen, i les afàsies progressives primàries (APP), en aquesta la principal afectació acostuma a ser una alteració del llenguatge. Les APP es poden dividir en 4 subtips: la variant semàntica (vsAPP), on les persones que la pateixen presenten un llenguatge fluent, una anòmia severa i una pèrdua de la capacitat per identificar el significat de les paraules; la variant no fluent (vnfAPP), en aquest cas es caracteritza per presentar dificultats en l'emissió del llenguatge, agramatisme i apràxia de la parla; la variant logopèdica (vlAPP), que presenta com a característica principal les pauses constant en la parla, la dificultat per trobar les paraules i dificultats per repetir frases, habitualment amb una MA subjacent. Finalment, hi ha alguns pacients amb APP que no es poden classificar en algun dels grups mencionats anteriorment i s'anomenen APP inclassificables o mixtes. En aquesta tesi, a causa de les característiques de la mostra, s'estudien les variants vcDFT, vsAPP i vnfAPP. En la Figura 5, es pot observar l'atròfia característica de cada una de les variants mencionades (30,41).

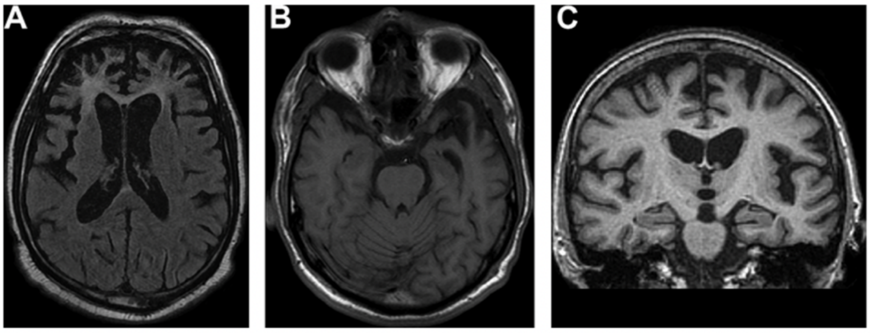

Figura 5: Mostra la ressonància magnètica de cervell de tres persones adultes per observar l'atròfia cerebral causada per la demència frontotemporal i com es diferencia en cada una de les variants clíniques. A) Pacient amb demència frontotemporal amb la variant conductual que presenta atròfia frontal B) Pacient amb demència frontotemporal amb la variant semàntica que presenta atròfia al temporal a l'hemisferi esquerre C) Pacient amb demència frontotemporal amb la variant no fluent que presenta atròfia frontal inferior de l'hemisferi esquerre i temporal superior. Figura adaptada de (30).

1.4 DIAGNÒSTIC DIFERENCIAL ENTRE LA DEMÈNCIA FRONTOTEMPORAL I LA MALALTIA D'ALZHEIMER

La superposició de símptomes amb altres malalties, especialment psiquiàtriques i la manca de marcadors de procés fisiopatològic, com s'ha mencionat abans, fa que amb freqüència pacients amb DFT pateixen un retard diagnòstic notable.

També persones amb MA sobretot menors de 65 anys, arriben a les consultes especialitzades aproximadament 3 anys després de l'inici dels símptomes (3). Així mateix, és freqüent la confusió entre la DFT i la MA, especialment en formes de presentació atípiques de la MA (42–44). És per això que hi ha una necessitat d'identificar biomarcadors específics de la DFT, per poder facilitar el seu diagnòstic. Ja s'ha vist que en el diagnòstic diferencial entre la DFT i la MA, els biomarcadors químics poden ser de gran ajuda (45), però no sempre estan disponibles o la presència de patologia concomitant podria fer-los resultar no concloents.

La combinació de neuroimatge, proves cognitives, i biomarcadors bioquímics podria millorar el diagnòstic de la DFT i de les seves diferents expressions clíniques. Els nivells de NfL en LCR i sang, són possibles biomarcadors de DFT molt prometedors, ja que indiquen dany neuroaxonal. Tot i això, podrien ser marcadors de neurodegeneració inespecífics, que també poden estar alterats amb pacients amb MA o en altres malalties neurodegeneratives (46–50). Així, tot i presentar resultats prometedors amb diferenciar la DFT de persones sanes, a l'hora de fer un diagnòstic diferencial amb la MA podrien no ser els més adequats. Els pacients amb DFT presenten nivells més alts de NfL respecte als pacients amb MA, tot i que hi ha un grau de superposició destacable (51,52). Per aquest motiu, biomarcadors de neuroimatge podrien ser d'ajuda a l'hora de realitzar un diagnòstic diferencial entre la DFT i la MA (53–55).

Els estudis de neuroimatge, especialment de RM estructural, poden millorar la precisió diagnòstica diferencial de la DFT i la MA quan els biomarcadors bioquímics no donen resultats concloents, ja que permeten identificar patrons anatòmics característics de cada una de les demències (41) on la MA presenta una atròfia majoritàriament temporal i parietal posterior, i la DFT es presenta essencialment alteracions temporals i frontals (30,50). Tot i això, és necessari estudiar la RM a escala individual, ja que aquests patrons característics de cada una de les demències, poden ser patrons en l'àmbit grupal, però no sempre són certs per a tots els pacients, pel fet que sobretot la DFT és molt heterogènia. Per una altra part, l'estudi de l'heterogeneïtat que s'observa clínicament podria donar pistes sobre els condicionants tant del fenotip, com de la progressió de la malaltia i podria ser clau per fer una avaluació personalitzada del cas. Finalment, actualment tant per la MA com per la DFT no disposem de marcadors de progressió o pronòstic a la DFT, cosa que limita la informació personalitzada al pacient i el desenvolupament d'eines terapèutiques.

1.5 RESSONÀNCIA MAGNÈTICA EN L'ESTUDI DE LA MALALTIA D'ALZHEIMER I LA DEMÈNCIA FRONTOTEMPORAL

La RM és una tècnica no invasiva que utilitza el magnetisme per crear imatges de les estructures del cos. És una tècnica molt habitual per obtenir imatges del cervell. Aquesta tècnica és no ionitzant, ja que no fa servir radiació com podria ser en el TC o el PET, i la majoria dels hospitals tenen accés a un sistema d'adquisició de RM. Dins d'un escàner de RM, s'emeten polsos de radiofreqüència en presència d'un camp magnètic molt fort (d'entre 1,5 i 7 Tesla) i aquests provoquen moviments d'excitació i relaxació en els nuclis d'hidrogen que es poden mesurar com a corbes de relaxació en el temps a les antenes receptores, anomenades T1 i T2. Els diferents teixits tenen diferents propietats magnètiques i això ens permet generar imatges de l'anatomia i la funcionalitat de l'estructura o l'òrgan estudiat. Hi ha diverses modalitats de ressonància magnètica cerebral, que depenen de la programació de la seqüència de polsos i de les mesures de T1 i/o T2. Les més importants són: la RM estructural que proporciona una imatge d'alta qualitat que permet observar detalls de l'anatomia del cervell i que és la que s'estudia en aquesta tesi; la RM funcional, que permet estudiar l'activitat cerebral; i finalment, la RM de difusió que mesura l'anisotropia de les molècules d'aigua, això ens permet estimar els tractes de les fibres cerebrals.

La física que hi ha darrere és que la RM estructural cerebral és que la magnetització dels teixits fa que varin de direcció, amb el camp magnètic que genera l'imant que envolta el tub de l'escàner. Quan es deixa de tenir camp magnètic es produeix la relaxació magnètica, és a dir, retornar al seu estat inicial abans d'haver estat alterat. Aquesta fase és la que ens permet obtenir realment les imatges que ens permeten estudiar el cervell. Cada teixit torna al seu estat en una velocitat diferent i aquesta informació és recollida, ja que es genera una ona magnètica que ens permet aconseguir aquesta informació que un cop es reconstrueixi ens permetrà identificar els diferents teixits en la imatge. En les imatges T1 (Figures 3 i 5), la substància grisa es correspon a la tonalitat de gris fort, mentre que la substància blanca al gris fluix. Finalment, el negre correspon al LCR i el blanc al crani. Depèn de la seqüència en què estigui potenciada la imatge la correspondència de colors i teixits serà diferent. Actualment, s'estan utilitzant cada cop més els escàners de RM de 3T (abans s'utilitzaven més el 1,5T que tenen una resolució menor) que permeten diferenciar la substància grisa, la blanca i el LCR sense dificultats, permeten obtenir valors d'aquestes imatges cerebrals.

Per poder avaluar l'atròfia cerebral (és una prova de la neurodegeneració), i ferla servir com una eina complementària pel diagnòstic diferencial de la MA i la DFT, és necessari tenir uns valors numèrics que permetin quantificar els aspectes visuals que s'observen. L'atròfia es pot mesurar visualment, amb escales que permeten categoritzar segons el rang d'atròfia que es presenta, o amb procediments de processament d'imatges semiautomàtics. Aquests procediments semiautomàtics que són els que fem servir per a la present tesi, es poden realitzar amb diversos softwares (FreeSurfer (https://surfer.nmr.mgh.harvard.edu), FSL (https://fsl.fmrib.ox.ac.uk), SPM (www.fil.ion.ucl.ac.uk/spm)) i s'ha demostrat que les mesures quantitatives que proporcionen tenen una correlació alta amb les mesures histològiques (56). Per poder estudiar l'atròfia en detall, aquests programes fan el processament semiautomàtic de les imatges de RM estructural per estimar el volum de substància grisa i el gruix cortical i obtenir valors globals i regionals, mitjançant atles que ens permetin identificar cada una de les regions. Per regions corticals, és habitual dividir l'escorça cerebral segons l'atles Desikan (57), que ens proporciona un mapa de 68 regions corticals de cervell (Figura 6). Mentre que per estudiar les regions subcorticals l'atles Aseg (58), és el més comú i ens permet tenir un mapa de 37 regions subcorticals. Hi ha altres atles que ens permeten segmentar el cervell en diferents granularitats, utilitzar un atles o un altre dependrà de l'objectiu de l'estudi.

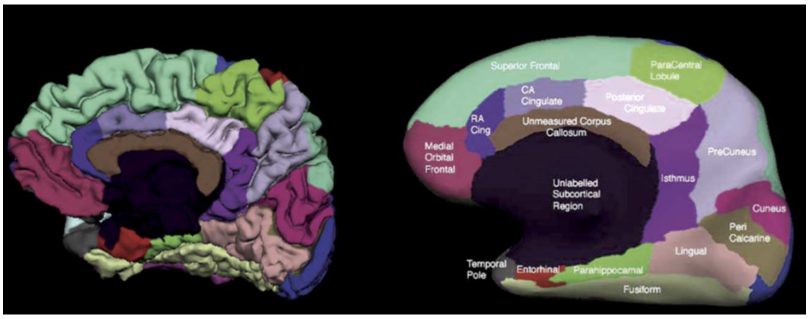
Superior Frontal
ParaCentral Lobule
CA Cingulate
Posterior Cingulate
RA Cing
Unmeasured Corpus Callosum
Medial Orbital Frontal
PreCuneus
Isthmus
Unlabelled Subcortical Region
Cuneus
Peri Calcarine
Lingual
Temporal Pole
Entorhinal
Parahippocamal
Fusiform

Figura 6: Representacions de les regions corticals segons de l'atles Desikan amb vista medial. Esquerra mostra la representació pial i la dreta mostra el cervell inflat conjuntament amb els noms de les regions que es veuen en aquesta vista. Figura adaptada de (57)

L'atròfia de l'hipocamp és una característica diagnòstica de la MA acceptada, que es pot avaluar amb mètodes semiautomàtics que ens permetin calcular el volum de l'hipocamp per poder estudiar si hi ha una disminució del seu volum, que ens estaria indicant la presència d'atròfia en aquesta regió. Tot i ser una mesura acceptada en el diagnòstic de la MA, pot ser que serveixi per diferència persones amb MA respecte persones sanes, però no suficient per fer un diagnòstic diferencial amb DFT o un diagnòstic de la MA amb un estadi precoç (59). Per aquest motiu necessitem estudiar totes les regions del cervell, ja que ens aportarà la informació necessària que podria fer realitzar el diagnòstic diferencial. Quan es vol estudiar totes les regions del cervell com un conjunt i estudiar les seves relacions i com interaccionen entre elles per poder explicar les dues demències o estudiar la progressió que presenten al llarg del temps conjuntament la seva variabilitat es necessiten eines matemàtiques avançades. Per poder aconseguir aquest diagnòstic diferencial de la DFT i la MA, hi ha un creixement d'eines estadístiques avançades i models d'intel·ligència artificial que permeten l'estudi detallat de tota la informació que aporten les imatges o la combinació d'aquestes amb altres proves accessibles (30,60–62).

1.6MODELS PREDICTIUS

El 1993 Geisser defineix els models predictius o models de predicció com “el procés pel qual un model es crea o es tria per intentar predir millor la probabilitat d’un resultat” (63). Aquesta definició es pot interpretar com que els models predictius són processos de desenvolupament d’una eina o model matemàtic que genera una predicció precisa per englobar els més actuals (64). En els fons els models predictius són un conjunt de tècniques matemàtiques i estadístiques que permeten predir el comportament de certa variable en funció d’un o uns esdeveniments coneguts. Analitzant les dades històriques i les actuals i generant un model matemàtic es poden predir resultats del futur. Fet que fa que estiguin molt presents en el nostre dia a dia i siguin de rellevància per la ciència. Són d’utilitat pel cercador Google, les xarxes socials, i en context clínic ens podrien ajudar a respondre la pregunta de si una persona patirà una malaltia concreta. Poden ser de gran utilitat per estudiar malalties i ens poden guiar cap a una medicina personalitzada en permetre definir quins tractaments mèdics poden ser més adients segons les característiques del pacient, estudiar la variabilitat, o predir quina evolució pot tenir un pacient en funció de les seves característiques. Ara bé, aquests models també poden generar prediccions inexactes o donar respostes incorrectes. Quan els models predictius fallen acostuma a ser per algun d’aquests motius (1) un preprocessament incorrecte de les dades, (2) una validació inadequada de les dades, (3) una extrapolació injustificada, per exemple, portar les dades en un espai que no és el correcte, o (4) el sobreajustament del model (64). Per tant, moltes de les vegades que fallen els models recau en les dades, és per això, que és molt important conèixer les dades i saber quines són les limitacions que comporten si volem obtenir models predictius fiables. Pel context d'aquesta tesi, ens centrarem en els models predictius aplicats en casos clínics.

Els models predictius engloben des de la regressió lineal fins a la intel·ligència artificial. Segons el disseny de l'estudi, es poden crear models amb dades transversals, és a dir, quan es disposa d'una única visita per participant, o amb dades longitudinals, quan per cada un dels participants tenim una visita inicial i d'altres successives que ens permeten estudiar la seva evolució al llarg del temps. Depèn de l'estudi, serà més adient fer estudis transversals o longitudinals segons la hipòtesi que es tingui. Tot i això, els estudis longitudinals recullen dades en diversos moments de múltiples participants, per tant, són difícils i costosos, però són àmpliament reconeguts com a recursos necessaris per entendre les malalties neurodegeneratives (65). En comparació amb els enfocaments transversals, els estudis longitudinals proporcionen una potència estadística més gran reduint l'efecte de confusió de la variabilitat entre subjectes (66). Per estudiar dades longitudinals la tècnica més establerta és el model d'efectes mixtos lineals (EML), aquests models ofereixen un marc potent i versàtil per analitzar les dades longitudinals. És més adequada que altres tècniques més clàssiques com podrien ser l'anàlisi de mesures repetides de la variància (ANOVA) o l'anàlisi transversal dels canvis percentuals (60,67,68).

Un cop tenim les dades modelitzades, se segueix amb un estudi de la interferència dels resultats, que permeten avaluar les hipòtesis que es tenen per analitzar si es compleixen o no. Estadísticament, es pot estudiar la mateixa hipòtesi des de dos enfocaments diferents: (1) l'enfocament freqüentista, el més utilitzat històricament que estableix el llindar del p valor de 0,05 per saber si és estadísticament significatiu i (2) l'enfocament bayesià que en aquest cas estudia les distribucions de probabilitat de les mostres amb probabilitats condicionals de valors coneguts. L'enfocament freqüentista en termes generals es basa en les distribucions de mostreig i en el Teorema Central del Límit (69,70). En aquest cas, els paràmetres poblacionals d'interès són tractats com valors fixos. Contràriament, l'enfocament bayesià estima els paràmetres a partir de la distribució de la població un cop es dona l'evidència amb les dades observades (69,71). Per tant, es tracten els paràmetres d'interès com variables aleatòries que es poden descriure amb distribucions de probabilitat. L'enfocament bayesià es basa en el Teorema de Bayes i es basa en l'estudi de la distribució de les dades ja existents, nomenades "prior", per poder predir el comportament futur de dades noves (72). Cada opció té els seus punts forts i febles, però les diferències en aplicacions en el camp de les malalties neurodegeneratives no ha estat molt estudiat. En termes generals, la virtut més gran de l'enfocament freqüentista és el rigor i precisió extrema en la hipòtesi concreta que s'està estudiant. I la virtut més gran de l'enfocament bayesià és que permet una interpretació realment probabilística, ja que proporciona respostes directes a preguntes sobre la magnitud probable dels efectes d'interès i, per tant, permet comparar hipòtesis d'una manera senzilla i intuïtiva (72,73).

Els models predictius no solament ens permeten predir com evolucionaran els pacients amb els anys o en funció d'alguna altra variable. Tot i que aquesta és una de les principals utilitats d'aquests models. Ens permeten portar un seguiment de l'evolució de les malalties, amb tots els beneficis que comporta. Aquest seguiment, avui en dia, és més necessari que mai, ja que permet avaluar l'eficàcia de fàrmacs o proporcionar un pronòstic precís. Per exemple, ens podria ajudar a identificar el moment idoni per iniciar el tractament i controlar la seva resposta (74). Els models predictius, que permeten realitzar estimacions en funció del temps són de gran importància per entendre la progressió de la malaltia i comparar l'estat del cervell (o altres variables, però en aquesta tesi ens centrem en les RM) entre pacients en diferents estadis de la malaltia (75–77). Tanmateix, aquests models poden generar prediccions imprecises, ja que suposen que tots els grups clínics (per aquesta tesi pacients amb MA, pacients amb DFT i persones sanes) estan ben definits i són iguals, quan biològicament sovint són molt heterogenis. Aquest enfocament ha permès detectar efectes de grups clínics, sobretot entre pacients i persones sanes, però quan es volen estudiar diferències entre grups, aquesta filosofia pot suposar tot un repte i un problema, ja que no té en compte que hi pot haver superposició de símptomes o patrons (78–80). La Figura 7 representa la base teòrica d'aquesta filosofia (Figura 7A) i alhora el que es troba en la pràctica clínica (Figura 7D), on hi ha superposicions de símptomes i patrons, co-patologies, etc. El que sembla un impediment per estimar aquests models amb precisió, actualment s'ha convertit en un gran avantatge per avançar cap a la medicina personalitzada. Podem assumir que els grups clínics no són homogenis, i estudiar aquesta heterogeneïtat que pot ser deguda a: (1) heterogeneïtat clínica, heterogeneïtat a conseqüència de diferents perfils de símptomes que es classifiquen sota el mateix trastorn. Un clar exemple és la DFT, ja sigui per les seves expressions clíniques (vcDFT, vsAPP, vnfAPP, vlAPP) com per les seves diferents causes genètiques (C9orf72, GRN, MAPT). (2) Heterogeneïtat biològica, heterogeneïtat induïda per diferents predisposicions biològiques que convergeixen en els mateixos símptomes, per exemple podrien ser les diferents mutacions que causen MA o DFT. Finalment, (3) heterogeneïtat ambiental, diferents esdeveniments ambientals que causen (o eviten) els mateixos símptomes. Estudiar les desviacions que presenten els pacients de la norma, ens pot permetre avaluar la gravetat de la malaltia, la variabilitat cognitiva i obtenir un estudi personalitzat de l'evolució de cada pacient. En el cas de la MA i la DFT, aquestes mostren una heterogeneïtat individual relativament alta en la presentació i la taxa de progressió que és rellevant per entendre millor la patogènesi d'aquestes malalties, i predir-ne la progressió i, potencialment, l'efecte dels tractaments (81).

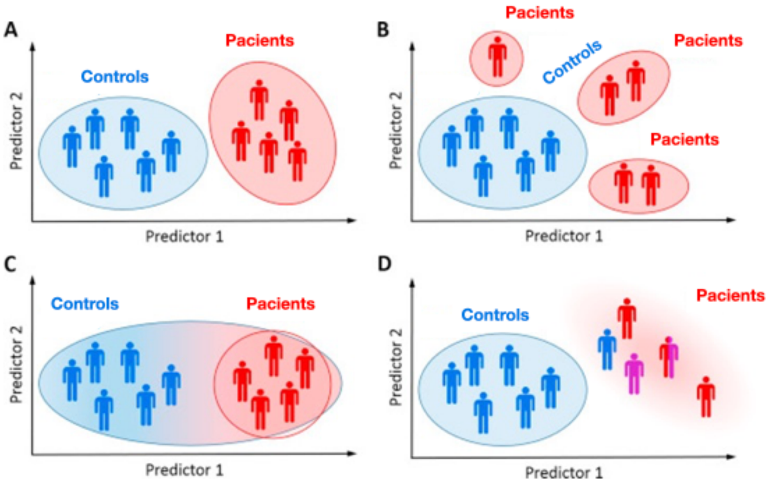
A
Controls
Pacients
Predictor 2
Predictor 1
B
Pacients
Controls
Pacients
Pacients
Predictor 2
Predictor 1
C
Controls
Pacients
Predictor 2
Predictor 1
D
Controls
Pacients
Predictor 2
Predictor 1

Figura 7: Representació dels diferents enfocaments dels grups clínics, en aquest cas entre persones sanes i pacients, però és extrapolable a més grups. A) Enfocament clàssic de casos i persones sanes on se suposa que cada grup clínic està ben definit. B) La població clínica pot estar formada per múltiples grups, cadascun amb una patologia diferent. C) La variació relacionada amb la malaltia es pot relacionar amb variació de les persones sanes. D) El grup clínic pot ser difús i heterogeni a conseqüència d'un diagnòstic erroni, de comorbiditats o de co-patologies. Figura adaptada de (78).

1.7 APRENENTATGE AUTOMÀTIC

Els inicis de la ciència computacional daten de 1950 quan Alan Turing va fer pública la següent pregunta "Les màquines poden pensar?" en un seminari on presentava la seva investigació (82). En aquell moment encara no s'havia desenvolupat la intel·ligència artificial, però sí que es considera un punt d'inflexió o, per molts, el seu naixement. Malgrat els anys que han passat, la pregunta continua sent vigent. Actualment, la intel·ligència artificial es podria definir com una ciència que permet imitar les capacitats humanes per resoldre problemes i la presa de decisions. És una definició volàtil, ja que en les últimes dècades han sorgit diferents definicions o han anat evolucionant. El terme intel·ligència artificial és molt ampli, engloba diferents branques dins d'aquesta disciplina, com l'aprenentatge automàtic –conegut també en anglès com *machine learning*–, i l'aprenentatge profund –conegut com a *deep learning* en anglès. Aquestes dues disciplines estan formades per algorismes d'intel·ligència artificial que permeten realitzar prediccions o classificacions basades en un conjunt de dades. En aquesta tesi ens centrem en l'aprenentatge automàtic.

L'aprenentatge automàtic va ser definit l'any 1959 pel científic Arthur Samuel com "el subcamp de la informàtica que dona als ordinadors la capacitat d'aprendre sense ser programats explícitament" (83). L'aprenentatge automàtic es basa en aquest concepte, ja que és una disciplina científica que desenvolupa models estadístics i algorismes que utilitzen els sistemes computacionals per dur a terme de manera eficaç una tasca específica basant-se en patrons i inferència de les dades, i millorant progressivament la seva precisió. L'objectiu principal és permetre que els ordinadors aprenguin automàticament, sense fer ús d’instruccions humanes explícites. I per aconseguir-ho se segueixen normalment 3 passos estàndards: (1) Procés de decisió, les dades que s'introdueixen a l'algorisme permeten dur a terme una estimació de quin patró segueixen, amb aquesta informació l’algorisme és capaç de fer una classificació o predicció. (2) Funció d'error, s'avalua l'error de la predicció que ha dut a terme l'algorisme fent servir les dades introduïdes. Finalment, (3) optimització del model, s'analitza la qualitat de les prediccions, és a dir, si el model és capaç d'ajustar-se a les dades introduïdes. El que es pretén en aquest pas és reduir la discrepància entre l'exemple conegut (les dades introduïdes inicialment) i l'estimació feta. Aquest pas s’acostuma a repetir fins que s’obté la millor precisió possible.

Si es vol avaluar el model d'intel·ligència artificial, es necessita realitzar dues fases, una primera fase que és el que s’anomena entrenament i la segona fase que és el testatge. Això ens permetrà avaluar el model amb unes dades diferents de les que s’han fet servir per establir els paràmetres del model. Així, les dades es divideixen en dues parts. Una part de les dades es farà servir per entrenar el model i l’altra per avaluar-ne la fiabilitat o precisió. En cas de no procedir a fer la divisió dels dos conjunts de dades, ens podem trobar amb el problema de sobre ajust, que és un problema metodològic greu en els algoritmes d’intel·ligència artificial. Si s’avalua el model amb les mateixes dades que s’han utilitzat per calcular els paràmetres, s’obté una puntuació sobre optimista, ja que repeteix les etiquetes ja vistes, però no es podria avaluar si el model és útil per predir futures dades (84,85). Tanmateix, en molts casos, la divisió de les dades en dos grups no és suficient, ja que molts models necessiten avaluar quins són els paràmetres de les funcions internes del model pel conjunt de dades estudiades. És per això que en molts casos es recomana realitzar una tercera divisió de les dades, el que s'anomena conjunt de dades de validació per resoldre el problema. El conjunt de dades de validació avalua els paràmetres òptims que ha de tenir la funció i les dades del testatge avaluen el model. Naturalment, en dividir les dades disponibles en tres conjunts, reduïm dràsticament el nombre de mostres que es poden utilitzar per entrenar el model, i els resultats poden dependre d'una elecció aleatòria particular pel conjunt de les dades estudiades. Per solucionar aquest problema, es du a terme l'estratègia del procediment de validació creuada (Figura 8), aquesta estratègia consisteix a eliminar el conjunt de dades de validació, però no la seva funció. L'enfocament més bàsic (Figura 8), és l'anomenat "k-fold" validació creuada, on el conjunt d'entrenament es divideix en k conjunts més petits (on k depèn de les dimensions dels conjunts). Aquest procediment consisteix a entrenar el model emprant els k-1 conjunts de les dades, i l'avaluació del model es realitza en les dades del conjunt restant. Per tant, les dades que entren el model en cap cas són les que es fan servir per a la validació. Aquesta divisió de les dades es fa repetides vegades, fent que totes les dades formin part de les dades d'entrenament i del conjunt de testatge, com s'observa en la Figura 8, i la mesura del rendiment que s'informa del model és la mitjana de tots els valors calculats. Aquest enfocament pot ser computacionalment costós, però permet fer l'entrenament sense reduir el conjunt de dades disponibles i fa que els models siguin precisos, extrapolables i vàlids (84,86,87).

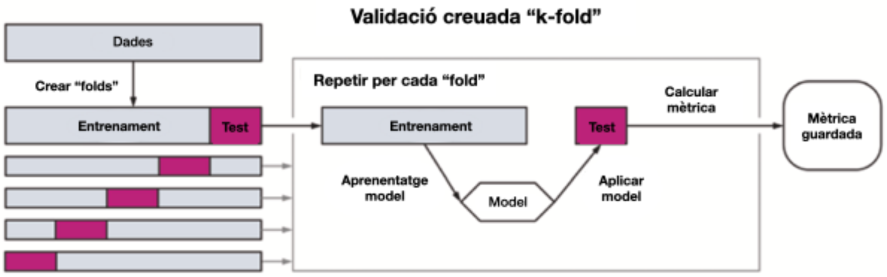

Figura 8: Esquema del procediment de la validació creuada "k-fold". S'observa que el conjunt de dades es divideix en k conjunts que es distribueixen segons conjunt entrenament i testatge aleatòriament. Figura adaptada de (85).

Els algorismes d'aprenentatge automàtic es poden classificar en 3 grups: aprenentatge supervisat, no supervisat, i semi supervisat, que és la combinació dels dos (88,89). En tots els casos es disposa d'una sèrie d'exemples o mostres, cada un amb unes determinades característiques. En l'aprenentatge supervisat, el conjunt de dades de l'entrenament estan etiquetades, per tant, l'algoritme aprèn d'exemples etiquetats. A nivell pràctic, el conjunt de dades s'estructura amb una matriu X (que no té definides les dimensions) on s'introdueixen les dades numèriques de les característiques que es necessiten per la predicció i el vector de Y on s'introdueixen les etiquetes. Els algoritmes de classificació supervisats engloben alhora dues categories principals: la regressió i la classificació. En el cas dels algoritmes de regressió, les dades que retorna l'algorisme poden tenir qualsevol valor numèric dins d'un interval, mentre que, en els algorismes de classificació, les prediccions es restringeixen a un conjunt limitat de valors (88). Posem un exemple pels dos casos, si suposem que en tots dos la matriu X són característiques mèdiques de MA o DFT, en un algorisme de classificació, podem intentar predir si aquests pacients tenen la malaltia o no. D'altra banda, en el cas de la regressió, el model avalua els valors continus per realitzar una predicció segons aquests valors continus, per exemple el valor individual de substància grisa en funció dels anys de malaltia. Els algoritmes de classificació automàtica supervisats més coneguts i utilitzats en conjunts de dades clíniques són la màquina vectorial de suport (MVS) – coneguda com *support vector machine* en anglès –, la regressió lineal i logística o els arbres de decisió, entre d'altres. En l'aprenentatge no supervisat, s'utilitzen dades no etiquetades i la finalitat és trobar similituds o diferències entre les característiques que ens permetin identificar patrons o grups de dades sense cap mena d'interferència humana. Seguint la notació algebraica de l'exemple anterior, el vector Y en aquest cas no existeix i les úniques dades són les característiques que formen la matriu X, no hi ha cap mena de variable que pugui etiquetar les dades en grups predefinits (84). Un exemple és agrupar els pacients en subgrups segons les seves característiques biològiques, cognitives o clíniques. Posteriorment, podem intentar interpretar si aquests subgrups es poden relacionar, per exemple, en si respondran a un medicament o no, en aquest cas tindríem la finalitat de predir si necessiten el medicament o no. Els algoritmes de classificació automàtica no supervisats més coneguts i utilitzats per dades clíniques són anàlisi de clústers o l'anàlisi de components principals. Finalment, l'aprenentatge semi supervisat és la combinació de les dues metodologies anteriorment explicades. En aquesta metodologia, en la fase d'entrenament part de les dades estan etiquetades mentre que les altres no. Així que les dades etiquetades serveixen per guiar la classificació de les no etiquetades. Aquest mètode acostuma a ser l'opció escollida si etiquetar totes les dades (agrupar-les en grups) té un cost molt elevat.

1.8 UTILITAT DE L’APRENENTATGE AUTOMÀTIC EN LA CLÍNICA

L'aprenentatge automàtic ha revolucionat molts camps i poden ser de gran utilitat en el desenvolupament de la cura de la MA, la DFT i altres afeccions cròniques. Aquestes tècniques es beneficien de l’accés a grans quantitats d’informació sobre els pacients, i són capaces d’aprendre i construir models robustos i d’alt rendiment, que poden ajudar a identificar marcadors de risc, identificar patrons amb utilitat diagnòstica, modelar la progressió de les malalties i fins i tot suggerir intervencions de medicina de precisió per a pacients a escala individual de manera ràpida (50). Tots i els prometedors avantatges, la metodologia està en un estat inicial, i s’han de crear algoritmes interpretables perquè siguin transferibles a la clínica i que millorin les eines disponibles.

Els algoritmes d’aprenentatge automàtic presenten molts avantatges per les malalties neurodegeneratives, en conseqüència per la MA i la DFT (90). Tot i això, els treballs publicats en aquesta àrea no es distribueixen uniformement, s’ha investigat molt en mètodes per la classificació entre pacients de MA i persones sanes, però no hi ha gaires estudis que se centrin a aprofundir en els avantatges d’aquests mètodes per estudiar el diagnòstic diferencial en les diferents demències, que és un problema clínicament rellevant (91). Els estudis recents proporcionen una evidència que sosté el paper de les tècniques d'aprenentatge automàtic que utilitzen la RM estructural (62,92,93) per donar suport al diagnòstic clínic de demències neurodegeneratives (61,94–96).

Les tècniques d'aprenentatge automàtic també ens poden ajudar a entendre la variabilitat fenotípica que presenten aquestes malalties tant en variabilitat de símptomes com d'edat d'inici o progressió i possiblement també de variabilitat en resposta terapèutica (90). En aquest sentit, l'aprenentatge automàtic no supervisat podria ser d'utilitat analitzant dades clíniques i de neuroimatge, podria identificar subtipus de MA, de DFT o caracteritzar la variabilitat individual. Així, els models d'aprenentatge automàtic ofereixen una oportunitat única per disseccionar aquestes malalties amb detall i des d'un enfocament objectiu, sense premisses clíniques preestablertes, ja que ens basem en dades recopilades sense intervenció humana (91).

En aquest context, el diagnòstic, l'anàlisi de la variabilitat i la predicció de l'evolució i la resposta terapèutica de les malalties neurodegeneratives a través dels models matemàtics és una branca d'investigació creixent, que requereix tant innovació metodològica com de models d'avaluació per l'aplicació clínica, amb una gran varietat d'aplicacions potencials. En aquest sentit, la recerca en el camp de l'aprenentatge automàtic ha d'adaptar-se a investigar els reptes clínics reals actuals; per l'altre costat la investigació clínica ha d'estar oberta als potencials beneficis de l'ús de models matemàtics, i especialment de l'aprenentatge automàtic.

CAPÍTOL 2

Hipòtesis i Objectius

2.1 HIPÒTESIS

1. La consistència dels resultats de l'anàlisi estadística segons l'aproximació freqüentista o bayesiana en bases de dades clíniques i paraclíniques de malaltia d'Alzheimer, especialment en dades longitudinals, dependrà de l'homogeneïtat de les dades.
2. Es podrà avaluar la variabilitat individual de les persones amb malaltia d'Alzheimer i demència frontotemporal amb models de predicció grupal del gruix cortical. Aquesta variabilitat es podrà relacionar amb variables clíniques i biomarcadors.

3. Les dades de ressonància magnètica estructural ens permetran classificar de forma automàtica els pacients amb malaltia d'Alzheimer i demència frontotemporal, i alhora proporcionaran mapes cerebrals que permetin la interpretació de les decisions de l'algoritme. Els estudis longitudinals milloraran el rendiment diagnòstic diferencial de la malaltia d'Alzheimer i la demència frontotemporal amb algoritmes de classificació automàtica.
4. Serà possible obtenir les probabilitats individuals de diagnòstic de la malaltia d'Alzheimer o la demència frontotemporal, i aquestes permetran identificar subjectes en la zona de baixa confiança del diagnòstic i planificar quins pacients es beneficiarien de proves addicionals.
5. Una mesura precisa de l'asimetria cerebral ens permetrà diferenciar els pacients amb demència frontotemporal respecte a pacients amb malaltia d'Alzheimer i diferenciar-ne les variants clíniques.
6. Estudis quantitatius de ressonància magnètica en algunes variants genètiques de demència frontotemporal, podrien identificar canvis d'utilitat clínica. Portadors de la mutació C9orf72, podrien presentar una pèrdua de substància blanca al tronc encefàlic per l'afectació de vies motores.

2.2 OBJECTIUS

L'objectiu general d'aquesta tesi és avaluar l'aportació dels models matemàtics avançats en les eines de diagnòstic, d'estudi de variabilitat i de progressió actuals en la malaltia d'Alzheimer i la demència frontotemporal. Els objectius específics són:

1. Avaluar l'impacte de l'ús dels diferents enfocaments estadístics en estudis longitudinals de dades clíniques i paraclíniques de ressonància magnètica de la malaltia d'Alzheimer.
2. Caracteritzar models de gruix cortical en termes que a) permetin estudiar la variabilitat individual dels pacients amb malaltia d'Alzheimer i demència frontotemporal, b) permetin predir l'evolució dels pacients amb malaltia d'Alzheimer i demència frontotemporal amb relació al temps d'evolució.
3. Explorar i comparar la utilitat diagnòstica dels algoritmes de classificació automàtica en estudis transversals i estudis longitudinals de pacients amb malaltia d'Alzheimer i demència frontotemporal.
4. Establir la probabilitat individual per la classificació entre pacients amb malaltia d'Alzheimer, demència frontotemporal, i persones sanes per estudiar la zona de baixa confiança diagnòstica de persones amb malaltia d'Alzheimer i demència frontotemporal.
5. Definir un índex d'asimetria cerebral a partir dels valors del gruix cortical que permeti identificar persones amb demència frontotemporal i les seves variants clíniques.
6. Avaluar la pèrdua de substància blanca del tronc encefàlic en portadors simptomàtics i asimptomàtics de la mutació C9orf72 comparat amb no portadors i la seva relació amb la presència de malaltia de motoneurona clínica.

CAPÍTOL 3

Mètodes

El mètode de cadascun dels estudis es detalla en el manuscrit corresponent. Tot seguit es descriuen breument de manera general les dades utilitzades, els criteris de diagnòstic dels participants i les principals tècniques utilitzades, ja que són factors comuns en els articles que formen aquesta tesi.

3.1 PARTICIPANTS

Pel primer treball es fa servir la base de dades pública Alzheimer's Disease Neuroimaging Initiative (ADNI) (https://adni.loni.usc.edu), que és una base de dades pública multicèntrica dissenyada per estudiar la MA. Les dades dels escàners de ressonància magnètica estructural de pacients amb MA, deteriorament cognitiu lleu (DCL) i persones sanes (CTR) de diverses visites són descarregades. Els criteris de diagnòstic (97) per cada un dels pacients és el següent:

a) MA: Persones adultes amb queixes cognitives amb Mini-Examen d'estat mental (MMSE) entre 20 i 26 i Escala de classificació clínica de demència (CDR) entre 0,5 i 1.
b) DCL: Persones adultes amb queixes cognitives amb MMSE entre 24 i 30 i CDR 0,5.
c) CTR: Persones adultes sanes sense queixes cognitives amb MMSE entre 24 i 30 i CDR 0.

Pel segon, tercer, quart i cinquè treball es fa servir dades de pacients MA, DFT, i CTR, visitats a la Unitat d’Alzheimer i altres trastorns cognitius del Servei de Neurologia de l’Hospital Clínic de Barcelona. El Comitè ètic de l’Hospital Clínic de Barcelona ha aprovat els estudis (HCB 2019/0105) i tots els pacients han signat el consentiment informat per participar. Per aquests estudis s’han fet servir les imatges de RM estructural, en els estudis segon, tercer i cinquè s’ha fet servir més d’una visita dels pacients, per tant, un estudi longitudinal, mentre que en el quart estudi solament es va fer servir la primera visita d’aquests pacients. En els estudis segon, quart i cinquè també s’han fet servir dades de biomarcadors de LCR i sang i proves cognitives d’aquests pacients. Els criteris de diagnòstic per aquests pacients ha estat:

a) MA: Pacients que complien els criteris clínics de demència lleu a causa de MA (16,17) confirmats per un perfil de biomarcadors de LCR que suggereix la neuropatologia subjacent de la MA segons National Institute on Aging/Alzheimer's Association Research Framework 2018 (13).
b) DFT: Pacients que complien els criteris clínics pel diagnòstic de diferents variants de DFT, exactament la variant conductual o els fenotips de l'afàsia primària progressiva semàntica o no fluent (36,98).
c) CTR: Persones adultes sanes amb rendiment cognitiu dins del rang normatiu.

Finalment, el sisè article solament ha estat centrat en DFT de causa genètica, per això s'ha fet servir les dades de RM i de proves cognitives del consorci Genetic Frontotemporal dementia Initiative (GENFI) (https://www.genfi.org/) que és un consorci que engloba diferents grups de centres de recerca d'Europa i el Canadà. L'Hospital Clínic de Barcelona en forma part a través d'una de les directores d'aquesta tesi la Dra. Raquel Sánchez-Valle com IP en el centre. Els comitès ètics de cada centre han aprovat els procediments i tots els pacients han firmat el consentiment corresponent. En aquest cas solament han format part de l'estudi els pacients amb la mutació del gen C9orf72, els criteris de diagnòstic dels pacients són els següents:

a) DFT: Portadors de l'expansió patogènica a C9orf72, asimptomàtics o simptomàtics amb diagnòstic clínic de DFT o esclerosi lateral amiotròfica.
b) CTR: Participants familiars de primer grau de portadors de la mutació C9orf72, no portadors de l'alteració genètica.

3.2 ADQUISISCIÓ DE LA RESSONÀNCIA MAGNÈTICA

Les dades de RM estructural del primer treball es van adquirir en diferents escàners 1,5T i 3T, dels fabricants Philips Healthcare (Koninklijke Philips NV, Amsterdam, Països Baixos), GE Healthcare Life Sciences (General Electric, Boston, MA, EUA) i Siemens Healthcare Diagnostics. (Siemens, Erlangen, Alemanya).

Pels treballs segon, tercer, quart i cinquè les imatges de ressonància magnètica es van adquirir en els escàners 3T Magnetom Trio Tim (Siemens Medical Systems, Alemanya) i la seva actualització el Prisma 3T (Siemens Medical Systems, Alemanya) de la Plataforma d'imatge per Ressonància magnètica de FRCB-IDIBAPS. Es va utilitzar les imatges de RM estructurals 3D d'alta resolució (ponderat en T1, MP-RAGE, temps de repetició = 2,30ms, temps d'eco = 2,98ms, 240 llesques, camp de visió = 256 mm, mida del vòxel = 1 ×1 x 1mm).

Finalment per l'últim treball les dades de RM estructural es van adquirir en 19 escàners de tres fabricants diferents: Philips Healthcare (Koninklijke Philips NV, Amsterdam, Països Baixos), GE Healthcare Life Sciences (General Electric, Boston, MA, EUA) i Siemens Healthcare Diagnostics. (Siemens, Erlangen, Alemanya). Els protocols d'adquisició van ser sotmesos a controls de qualitat per la seva harmonització.

3.3 PROCESSAMENT DE LES IMATGES DE RESSONÀNCIA MAGNÈTICA

El primer treball és l'únic que les dades de ressonància magnètica van estar ja processades externament, pel consorci ADNI (99).

Per la resta de treballs que formen aquesta tesi el processament de les imatges ha estat realitzat íntegrament per l'autora d'aquesta tesi. Es va utilitzar el processament disponible amb el software FreeSurfer (http://surfer.nmr.mgh.harvard.edu.sire.ub.edu/) versió 6.0 per dur a terme la reconstrucció cortical i la segmentació volumètrica de les adquisicions ponderades en T1. FreeSurfer ens permet obtenir mapes de gruix cortical i segmentar les estructures subcorticals (100,101). Amb les imatges reconstruïdes s'aconsegueixen les mesures resumides de volums mitjans del gruix cortical i la substància grisa dels dos hemisferis. Els valors per regions corticals i subcorticals que fem servir per a aquesta tesi són els dels atles disponibles a FreeSurfer (57,102). Totes les imatges i les segmentacions individuals es van inspeccionar visualment i es van corregir manualment en cas de ser necessari.

3.4 BIOMARCDORES DE LÍQUID CEFALORRAQUIDI i SANG

Els estudis segon, quart i cinquè inclouen biomarcadors de LCR i sang, els nivells d'aquests biomarcadors es van obtenir al laboratori del grup d'Alzheimer i altres trastorns cognitius de FRCB-IDIBAPS (Cellex) o al Centre de diagnòstic biomèdic de l'Hospital Clínic de Barcelona.

Pels biomarcadors de LCR, la seva extracció va ser amb una punció lumbar, posteriorment l'anàlisi dels diferents biomarcadors va ser utilitzat kits d'assaig immunosorbent lligat a enzims (ELISA) disponibles comercialment per determinar els nivells d'Aβ42, t-tau i p-tau (INNOTEST, Fujirebio Europe N.V., Gent, Bèlgica), NfL (IBL International, Hamburg, Alemanya) i 14-3-3 (CircuLex, MBL International Corporation, Woburn, MA).

En el cas dels biomarcadors de sang, la seva anàlisi va ser mitjançant Quanterix Simoa Neurology 4-Plex A (incloent-hi t-tau, GFAP, NfL i UCH-L1) i Quanterix Simoa p-tau181 Advantage V2 i V2.1 seguint el protocol del fabricant (Quanterix, Billerica, MA) i s'harmonitza els valors dels dos kits.

3.5 IMPLEMENTACIÓ DELS MODELS MATEMÀTICS

Els models matemàtics, els models d'intel·ligència artificial i les anàlisis estadístiques de la present tesi han estat dissenyats per l'autora d'aquesta tesi i implementats en els llenguatges de programació Pyhton (www.python.org) o R (https://www.r-project.org) principalment. Les versions dels programaris s'indiquen en els diferents treballs, ja que hi ha hagut actualitzacions al llarg de la tesi.

El primer i sisè treball tenen una base principalment estadística, i per això s'han dut a terme únicament amb el llenguatge de programació R, i en el cas del primer treballa una part també amb el llenguatge Stan (https://mc-stan.org/). Pel primer treball hi ha el codi per reproduir l'estudi en obert a Github (https://github.com/Agnes2/LME-with-a-Bayesian-and-Frequentist-Approaches.git), ja que és amb dades públiques.

El segon, tercer i quart article els models d'intel·ligència artificials han estat implementats en Python amb la llibreria Scikit-learn (103), una llibreria àmpliament utilitzada i acceptada per realitzar models d'intel·ligència artificial en diversos àmbits i contextos. Es caracteritza per tenir eines simples i eficients per l'anàlisi predictiva de les dades. Per altra banda, la part més estadística i de gràfics dels treballs ha estat desenvolupada en R.

Finalment, en el treball cinquè també s'ha fet servir la combinació dels dos llenguatges de programació (Python i R). Per desenvolupar l'índex que es presenta s'ha fet amb el llenguatge Python sense fer ús de llibreries on ja havia estat implementat. Finalment, per l'estudi estadístic i de visualització s'ha fet servir funcions de llibreries ja implementades de R.

CAPÍTOL 4

Resultats

TREBALL 1

Evaluating the performance of Bayesian and frequentist approaches for longitudinal modeling: application to Alzheimer's disease.

Agnès Pérez-Millan, José Contador, Raúl Tudela, Aida Niñerola-Baizán, Xavier Setoain, Albert Lladó, Raquel Sánchez-Valle, Roser Sala-Llonch.

Scientific Reports 2022; 2022 Agost 24.

DOI: 10.1038/s41598-022-18129-4

IF (2022): 4,6; Q2 Multidisciplinary Sciences

www.nature.com/scientificreports

scientific reports

OPEN

Evaluating the performance of Bayesian and frequentist approaches for longitudinal modeling: application to Alzheimer's disease

Agnès Pérez-Millan[1,2], José Contador[1], Raúl Tudela[3], Aida Niñerola-Baizán[3,4], Xavier Setoain[3,4], Albert Lladó[1,5], Raquel Sánchez-Valle[1] & Roser Sala-Llonch[2,3]✉

Linear mixed effects (LME) modelling under both frequentist and Bayesian frameworks can be used to study longitudinal trajectories. We studied the performance of both frameworks on different dataset configurations using hippocampal volumes from longitudinal MRI data across groups—healthy controls (HC), mild cognitive impairment (MCI) and Alzheimer's disease (AD) patients, including subjects that converted from MCI to AD. We started from a big database of 1250 subjects from the Alzheimer's disease neuroimaging initiative (ADNI), and we created different reduced datasets simulating real-life situations using a random-removal permutation-based approach. The number of subjects needed to differentiate groups and to detect conversion to AD was 147 and 115 respectively. The Bayesian approach allowed estimating the LME model even with very sparse databases, with high number of missing points, which was not possible with the frequentist approach. Our results indicate that the frequentist approach is computationally simpler, but it fails in modelling data with high number of missing values.

The availability of longitudinal data-repeated measures of the same subjects over time-provides the opportunity to study trajectories of disease biomarkers. This offers an unquestionable value, as measures of change and evolution can complement cross-sectional analyses—mainly based on group differences at a specific time point—into the understanding of neurological diseases and the evaluation of disease-modifying treatments. However, real-life longitudinal databases are often characterized by high levels of noise, high variability, and missing points that lead to unbalanced data. All these factors represent a challenge when creating the models and often limit their interpretability. In this context, the use of linear mixed effects (LME) models offers a powerful and versatile framework for analysing longitudinal data, being more adequate than classical approaches such as repeated measures analysis of variance (ANOVA) or cross-sectional analysis of percent changes[1–3].

In addition, when these biomarkers are obtained from neuroimaging data, there are additional challenges, as there are strong dependencies within subjects and timepoints. In this sense, besides the clear dependencies between the different measures of one subject, there are also dependencies between subjects that need to be modelled. LME models attempt to reconcile these schemes by combining fixed and random effects, where fixed effects are assumed to represent those parameters that are the same for the whole population, while random effects are group dependent variables assumed to consider the variance in the data explained over time and subject. In our case, the random effects will take into account the variability of the non-independent measures from different subjects[4–6].

[1]Alzheimer's Disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació Clínic per a la Recerca Biomèdica, Universitat de Barcelona, 08036 Barcelona, Spain. [2]Institute of Neurosciences. Department of Biomedicine, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Faculty of Medicine, University of Barcelona, 08036 Barcelona, Spain. [3]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, Spain. [4]Nuclear Medicine Department, Hospital Clínic Barcelona, Barcelona, Spain. [5]Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas, CIBERNED, Madrid, Spain. ✉email: roser.sala@ub.edu

After data modelling, LME models are usually followed by statistical inference procedures, which allow the researcher to generate questions about the model and to further evaluate their statistical significance and clinical relevance. In this sense, while statistical significance is well established to p-values < 0.05, or equivalent, the assessment of clinical relevance has not yet a standard analysis[7]. It has been suggested that clinical relevance can never be determined form p-values alone[8], and complementary statistics should emerge to overcome this limitation in interpretability.

Historically, the dominant approach for performing the full procedure of LME modelling + statistical inference has been the Frequentist LME (FLME) approach. However, different methods using a Bayesian LME (BLME) approach have been suggested[3,9]. As suggested in the editorial of Anna G.M. Temp et al.[10], Bayesian statistics can be used jointly with frequentist approaches to draw clinically relevant conclusions that can complement classical studies based uniquely on statistical significance.

In general words, the FLME approach is based on sampling distributions and on the Central Limit Theorem[11,12], and it treats the population parameters of interest as fixed values[11]. While in BLME, parameters are estimated from the population distribution, given the evidence provided by the observed data[11]. BLME is considered a more natural approach to answer a question, since it estimates the parameters of interest directly from the population distribution instead of estimating them from the sampling distribution[13]. The Bayesian approach treats the parameters of interest as random variables that can be described with probability distributions[11]. These posterior distributions can be compared directly without referring to statistical results of multiple tests. Overall, the differences in comparing frequentist vs Bayesian approaches in different fields have opened a debate in several fields[14–17].

Alzheimer's disease (AD) is clearly one of the research fields that will benefit from the development of longitudinal statistical methods. It is believed that AD is a slowly evolving process that likely begins years before the clinical symptoms are manifested[18,19]. Therefore, there is a strong interest in identifying subjects at high risk before the full clinical criteria for AD dementia are met[20,21], as well as in giving reliable prognosis at the subject's level. The existence of public available databases, such as the Alzheimer's disease neuroimaging initiative (ADNI) has facilitated the definition and validation of neuroimaging biomarkers for AD[22]. In this sense, the hippocampal volume (HV), derived from structural Magnetic Resonance Imaging (MRI) data, has become one of the most widely used biomarkers. Compared with healthy aging, HV is progressively affected in AD, being already reduced in patients with Mild Cognitive Impairment (MCI) due to AD and more strongly affected in advanced AD stages[23–25].

In the recent years, the longitudinal trajectories of some AD biomarkers using frequentist approaches have been widely described[1,9,26]. On the other hand, the attempts to incorporate Bayesian statistics have shown promising results[2,3,9]. Even if frequentist and Bayesian schools represent two different schools of thinking, they often complement to each other. In the present work we analysed longitudinal MRI data from the ADNI dataset, using both FLME and BLME approaches. We performed simulations of real-life datasets derived from a public big database to explore the robustness of the methods with limited sample sizes and missing data using both approaches. Our goal was to evaluate the pros and cons of these approaches in real-life scenarios. For this, we create (simulate) datasets that incorporate the common handicaps found in clinical studies, e.g., low number of participants, missing data points or unbalanced sets with the aim to provide recommendations for further studies as regards the use of frequentist and Bayesian approaches, whilst illustrating the limitations of both approaches and bringing attention to statistical significance and clinical relevance.

Materials and methods

Data.

We used longitudinal brain MRI data (T1-weighted scans, combining 1.5 and 3.0 Tesla) from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. Including participants from ADNI-1, ADNI-GO, ADNI-2 and ADNI-3. Scans had been previously preprocessed with the FreeSurfer Longitudinal stream[27], as explained elsewhere[28]. We focus our analyses on the HV, as it is a common AD biomarker and we include the total intracranial volume (ICV), as a known confound in neuroimaging studies. Therefore, we downloaded HV, ICV and demographics from the data server.

We included AD dementia and MCI patients, as well as Healthy control (HC) participants, as labelled by the ADNI consortium[21]. According to their clinical evolution, we further created the following groups:

1. *Stable HC (sHC)* subjects who were diagnosed as HC throughout the follow-up period.
2. *Converter HC (cHC)* subjects who were diagnosed as HC at baseline and progressed to MCI or AD dementia.
3. *Stable MCI (sMCI)* subjects who were diagnosed as MCI throughout the follow-up period.
4. *Converter MCI (cMCI)* subjects who were diagnosed as MCI at baseline and progressed to AD dementia.
5. *AD* subjects who were diagnosed as AD at baseline.

We initially selected subjects having at least two acquisitions and we created several datasets as starting points. Tables 1 and 2 provide descriptive statistics of our initially selected longitudinal samples.

The datasets used for the different analyses were:

1. *Dataset 1* consisted of all the available data from the 4 timepoints, as described in Tables 1 and 2 (N = 1250 subjects).
2. *Dataset 2* was a reduced version of *dataset 1* containing only sMCI and cMCI subjects (N = 680 subjects).

Variable	sHC	cHC	sMCI	cMCI	AD	p-value
N	273	78	361	319	219	
Baseline age (years)	74.3 ± 5.7	76.2 ± 5.1	72.9 ± 7.4	72.4 ± 7.5	74.7 ± 7.9	0.19
Sex (M/F)	142/131	40/38	212/149	184/135	123/96	0.41
APOE-e4 (nc/c)	207/66	52/26	203/158	121/198	64/155	< 0.0005

Table 1. Characteristics of the longitudinal ADNI sample used. Baseline age values are in mean ± standard deviation. M = male, F = female, nc = non-carriers, c = carriers. p-values indicate differences between group. We used ANOVA for baseline age, and Fisher's exact test for the other data.

Time point	sHC (N)	cHC (N)	sMCI (N)	cMCI (N)	AD (N)	Time from baseline (years)
Baseline	273	78	361	319	219	0.00
Year 0.5	243	71	326	274	187	0.51 ± 0.05
Year 1	234	70	294	275	173	1.01 ± 0.06
Year 2	206	61	233	240	97	2.02 ± 0.08

Table 2. Number of scans per time point by clinical group and time between scans. Time from baseline values are in mean ± standard deviation.

Variable	sHC	cHC	sMCI	cMCI	AD	p-value
N	172	53	187	186	72	
Baseline age (years)	74.2 ± 5.6	76.1 ± 5.4	72.0 ± 6.9	71.6 ± 7.6	74.2 ± 7.9	0.004
Sex (M/F)	96/76	24/29	104/83	107/79	40/32	0.62
APOE-ε4 (nc/c)	133/39	33/20	116/71	72/114	21/51	< 0.0005

Table 3. Characteristics of the balanced longitudinal ADNI sample used. Baseline age values are in mean ± standard deviation. M = male, F = female, nc = non-carriers, c=carriers. p-values indicate differences between group. We used ANOVA for baseline age, and Fisher's exact test for the other data.

3. *Dataset 3* was a homogeneous balanced database. We selected from *dataset 1*, subjects with 4 timepoints available. Demographics for this database are summarized in Table 3 (N = 670 subjects).
4. *Dataset 4* is a reduced version of *dataset 3* containing only sMCI and cMCI subjects (N = 373 subjects).

Implementation of LME models. As there is not a fixed rule for choosing the number of random effects in LME, we evaluated two different models. Both models included the Intercept term, or group-mean, as a random effect. For the first LME model, the fixed effects were: time from baseline, group, group-by-time interaction, baseline age, sex, APOE status, APOE-by-time interaction and ICV. For the second LME model, the slope (measured as time from baseline) was also included as a random effect and the rest of variables were left as fixed effects (see Supplementary Material for details). The selection of the variables to be included in the models was done mimicking the analysis performed by Bernal-Rusiel et al. and according to previous AD literature[29,30]. HV (the outcome variable of our model) and ICV (a fixed effect variable of the model) variables were standardized to zero mean and standard deviation of one, using Fisher's Z norm, to ensure that the estimated coefficients are all on the same scale and therefore the corresponding effect sizes are comparable.

Statistical inference. We first studied which of the two proposed LME models were more appropriate for our sample using the frequentist approach with ANOVA and the Akaike Information Criteria (AIC). We used an ANOVA with χ^2 test on the model parameters and coefficients estimated for both models and we assessed the significance with the likelihood ratio test[31].

We then used frequentist statistical inference to test some of the well-known research questions in the AD field. For that, we created a set of contrasts using F-tests and using Satterthwaite's method[32] to compute the degrees of freedom. The contrasts studied were:

1. Are there differences across the 5 groups? (i.e., ANOVA main effect).
2. Are there differences between sMCI and cMCI?
3. Are there differences between cMCI and AD?

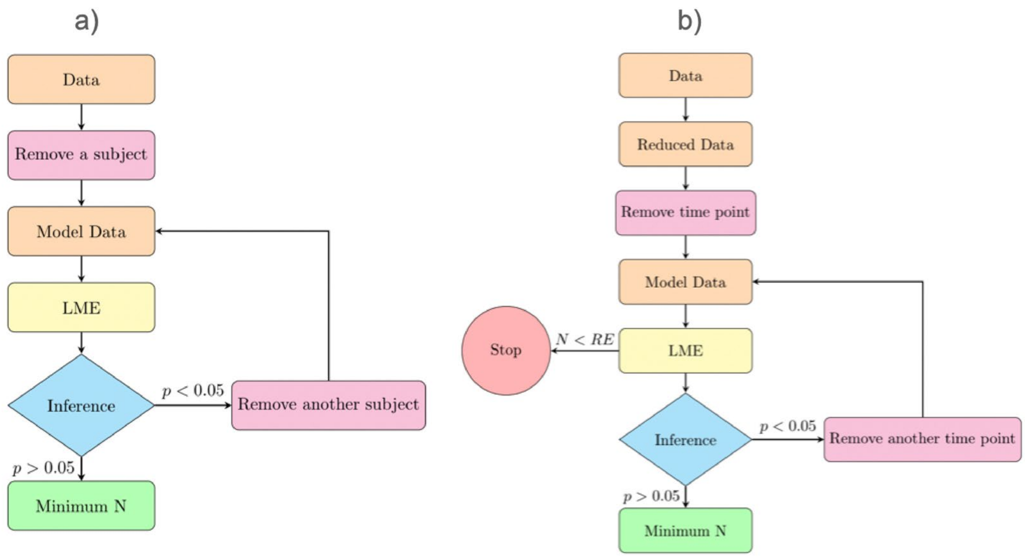

Figure 1. Simulations' scheme (**a**) Strategy for minimum N simulations. The initial data is dataset 1 or dataset 2. (**b**) Strategy of the simulation of missing time points. *RE*= random effects.

4. Are there differences between sHC and sMCI?
5. Are there differences between sHC and AD?
6. Are there differences between sHC and cMCI?
7. Are there differences between sHC and cHC?

We evaluated the LME model and tested these 7 contrasts in the datasets described previously (note that with dataset 2 or 4 we could only test contrast 1).

For the BLME approach, we also used the LME model with two random factors (the intercept and the slope). Posterior distribution measures regression parameters *ß* and contains all the information for statistical inference. We used the Credible Intervals (CrI) of this posterior distribution to study group differences. The CrI differ from the well-known Confidence Intervals (CI) in the fact that they are based in the uses of prior information and allow direct inferences about plausibility. Thus, CrI need the use of prior information to be estimated and can be interpreted as the probability in terms of plausibility[33]. We considered the four datasets and the same 7 contrasts described above.

All analyses were implemented in software R (https://www.r-project.org), version 3.6.2. For the LME model we used the *lme4* package[34] and the *rstan* package[35], so we combined R and Stan (https://mc-stan.org/) languages. The code for these analyses is available at https://github.com/Agnes2/LME-with-a-Bayesian-and-Frequentist-Approaches.git.

Simulation of real-life databases. Firstly, with the aim to provide a recommendation of the minimum N needed in these studies, we performed sequential simulations on the databases. We followed the scheme shown in Fig. 1a. We started from either dataset 1 (all groups) or dataset 2 (only MCI). We randomly selected one subject and we removed it (all its time points) from the dataset. Then we re-estimated the LME model, and we calculated the contrast of interest. This was repeated until the stopping criterion was met. At this point, we stored the last significant database, as a borderline significant dataset. Here, the stopping criterion was set at p-value > 0.05. This procedure was performed with dataset 1 (i.e., minimum N to differentiate across the 5 groups) and dataset *2* (i.e., minimum N to differentiate between sMCI and cMCI).

Further, with the aim to evaluate another typical situation in these studies, we also tested the effect of *missing timepoints*. We started from *dataset 3* (full balanced data) or *dataset 4* (MCI balanced data), and we proceeded as shown in Fig. 1b. First, we randomly selected one subject's time point of the sample and we removed it. We then estimated the FLME model, and we performed the corresponding statistical inference. We progressively removed time points from different subjects until the stopping criterion was met, and, as above, the last database was kept as a borderline significant database. The stopping criterion was set at p-value > 0.05. However, it should be mentioned that the FLME model cannot handle having more subjects' samples than random effects. Therefore, this restriction was added as an additional stopping criterion. The random effects were measured as $N \times 2$ subjects' samples, as we had a FLME model with randomly varying intercept and slope.

All the simulations were repeated over 500 iterations to account for the random selection of the subjects/timepoints to be removed at each step, leading to 500 *borderline significant databases*.

We applied BLME to the most compromised datasets found with FLME and we studied its behavior. Here, we performed a descriptive analysis in the borderline situations found with FLME. For that, we studied the obtained borderline datasets with the BLME model to estimate if they remained significant in the Bayesian framework and to evaluate the potential clinical interpretations that could be derived from them in terms of relevance.

Contrast	Dataset 1 F, p	Dataset 2 F, p	Dataset 3 F, p	Dataset 4 F, p
sMCI vs cMCI	39.2 5.6×10^{-10} *	36.1 3.2×10^{-9}*	31.8 2.5×10^{-8}*	24.2 1.3×10^{-6}*
All groups	22.8 4.1×10^{-18} *	–	15.1 8.5×10^{-12}*	–
AD vs cMCI	2.0 0.2	–	0.2 0.6	–
sHC vs sMCI	2.3 0.1	–	1.0 0.3	–
sHC vs AD	53.8 4.1×10^{-13}*	–	27.7 1.9×10^{-7}*	–
sHC vs cMCI	53.4 5.7×10^{-13}*	–	40.4 3.8×10^{-10}*	–
sHC vs cHC	2.3 0.1	–	2.7 0.1	–

Table 4. Summary of the null hypotheses tested and results of the statistical inference. *Indicates p-value < 0.05 (Bonferroni corrected).

Parameter	Interpretation	Estimate	95% CrI	
ß$_{11}$	cHC × time	− 0.03	− 0.06	0.01
ß$_{12}$	sMCI × time	− 0.02	− 0.04	0.01
ß$_{13}$	cMCI × time	− 0.08	− 0.11	− 0.06*
ß$_{14}$	AD × time	− 0.11	− 0.13	− 0.08*

Table 5. Estimation and 95% Credible Intervals (CrI) of the ßs of interest LME model fitted with a Bayesian approach. CrI borders are expressed as the 2.5% and 97.5% percentiles. *Indicates that the effect is significant (i.e., CrI does not contain zero).

Results

Statistics on ADNI longitudinal databases.

Of the two possible LME models to fit our data—one with the intercept as a random effect and another with intercept and slope as random effects—we found that the second one performed better for explaining our data. This was verified by the results of the ANOVA (p-value <<< 0.001) and by comparing their AIC values (1546.2 vs 1411.7). Therefore, all further analyses were performed with this model. To obtain comparable results, we also used intercept and slope as random effects in the BLME model.

We applied the FLME model followed by a set of F-tests to evaluate the contrasts of interest in the four databases described above. Results are shown in Table 4, and they reproduce previous reports on the field (as those presented in Ref.[1]). Mainly, we found significant differences in HV (*p*-value < 0.05) between all the five clinical groups, between sMCI and cMCI, between sHC and AD and between sHC and cMCI for the four initial datasets configurations. All these differences remained significant after correcting for multiple comparisons using Bonferroni (n = 7 tests, *p*-value < 0.05/7).

After fitting the BLME model, we obtained the joint posterior probability of the parameters. Here, we were interested in the posterior probability distribution for the *ß*s, and we used the interval from 2.5th to 97.5th percentiles to obtain the 95% CrI[36]. We focused on the *ß*s that represented change over time for the different groups (with sHC being the reference group). Results for dataset 1 are shown in Table 5. We found that the effect of time was significant (i.e., the 95% CrIs did not contain zero) for cMCI and for AD, while it was not significant for cHC and sMCI. When comparing groups, which not contain the reference group, we considered that there were differences when the corresponding CrI did not overlap. The contrasts with significant differences are the same as those depicted by the FLME approach.

Finding compromised datasets with FLME.

Minimum N simulations. By evaluating the 500 databases obtained from the procedure described in Fig. 1a and starting from dataset 1, we found that the minimum N needed to differentiate the five clinical groups (with p-value < 0.05) using the HV measure was N = 147 ± 73 overall. As the removal process followed a random order, the number of subjects within each group was not fixed by the algorithm. The group distribution resulting from the 500 databases is shown in Fig. 2a.

Similarly, with the same procedure and starting from database 2, we found that the minimum N needed to differentiate cMCI and sMCI using HV measures was N = 115 ± 64 overall. The distribution of the groups within the 500 obtained databases is shown in Fig. 2b.

Missing points simulations. For these simulations, in both cases (starting from *database 3* and *database 4*), we rapidly encountered that the limitation of number of samples < number of random effects. Thus, evidencing the low robustness of FLME approaches with highly unbalanced data.

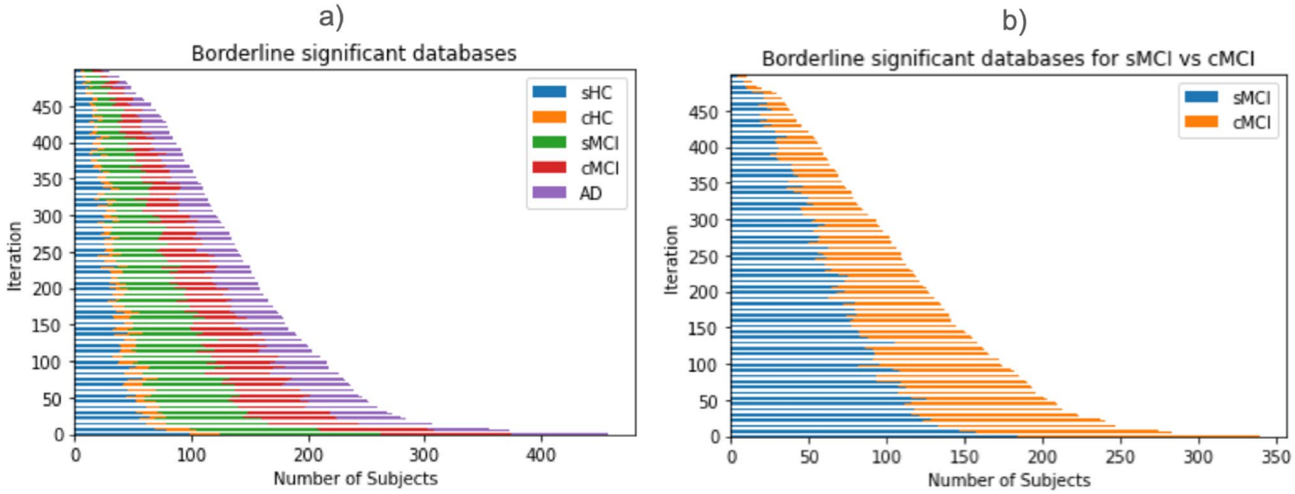

Figure 2. Distribution of subjects within each group for all the obtained databases (**a**) minimum N simulation across five clinical groups (**b**) minimum N simulation across MCI group.

By analysing the *database 3* (initial N = 670 with 4 time points per subject), with the process described in Fig. 1b, the simulations stopped at N = 612 ± 9 subjects (with different number of time points per subject), except from 3 iterations that did not converge into a failing database considered as outliers. At the moment that it was impossible to estimate the FLME model we had a mean of 2 missing time points per subject.

Similarly, with the same procedure and starting from *database 4* (initial N = 373 with 4 time points per subject) we found that the simulations stopped at N = 341 ± 7 except from 45 databases that did not stop. At the point that it was impossible to estimate the FLME model we had again a mean of 2 missing time point per subject.

Evaluating compromised datasets with BLME. *Minimum N simulations.* We studied the behaviour of BLME approach on different datasets obtained from the frequentist simulations of the minimum N. We first picked 10 different databases depicting differences across the 5 clinical groups, but that were at the limit for significance. These were randomly selected from the 500 iterations of the FLME experiments (the full characteristics of these databases are described in Supplementary Material). When studied with a BLME approach, 9 of them showed differences across the 5 groups. Figure 3a represents two of the datasets obtained in the simulation of the minimum N. Then, we selected 10 databases obtained from the simulations with dataset 2 (i.e., minimum N to find differences between sMCI and cMC). In this case, only 2 datasets remained significant when studied with the BLME approach. Figure 3b shows an example of the datasets obtained with the simulation of the minimum N.

Missing points simulations. To estimate the FLME model with a frequentist approach we encountered a practical limitation inherent to the model: the need to have more samples than random effects. Here, we selected 10 databases, from the FLME simulations, at the point that they no longer met the requirement. Thus, with these databases it was impossible to estimate a FLME model. The full characteristics of these databases are described in Supplementary Material. We studied with a BLME model these 10 databases. We found that the model can be estimated, and that all the databases depicted differences across the 5 groups (see Supplementary Material). For the 10 stopping databases created from the simulations with *dataset 4* we found similar results (see Supplementary Material), we could estimate the model and present differences for sMCI vs cMCI. Figure 3c,d represents Dataset 3 and 4 with one example of the datasets obtained after the missing points simulations for each situation.

Discussion

In this study, we explored large longitudinal neuroimage datasets obtained from ADNI to study trajectories of hippocampal volume change in AD. For that, we created LME models under the frequentist and the Bayesian frameworks. We found that both approaches have similar behavior in finding differences with the entire database. In the minimum N simulations, the Bayesian approach was slightly stricter to significance when reducing data size. In addition, our results indicated that the Bayesian approach is more robust to unbalanced and sparse databases with different number of measurements across subjects.

Firstly, our investigation supports the use of LME approaches to model longitudinal data. The results of our null hypotheses testing agreed with those reported previously in AD for the hippocampus[1,2]. Additionally, we provide evidence of the utility of these apparently more complex analyses to study compromised datasets with different time points for across subjects.

The frequentist approach allowed us to implement a method for testing the relationship between the sample characteristics (size and missing points) and the expected group differences. Even considering that the statistical threshold (here $p < 0.05$) may be rather arbitrary (see[37,38]), it is important to note that this was chosen as a controlled systematic approach to study the behavior of the databases when removing subjects, with the ultimate

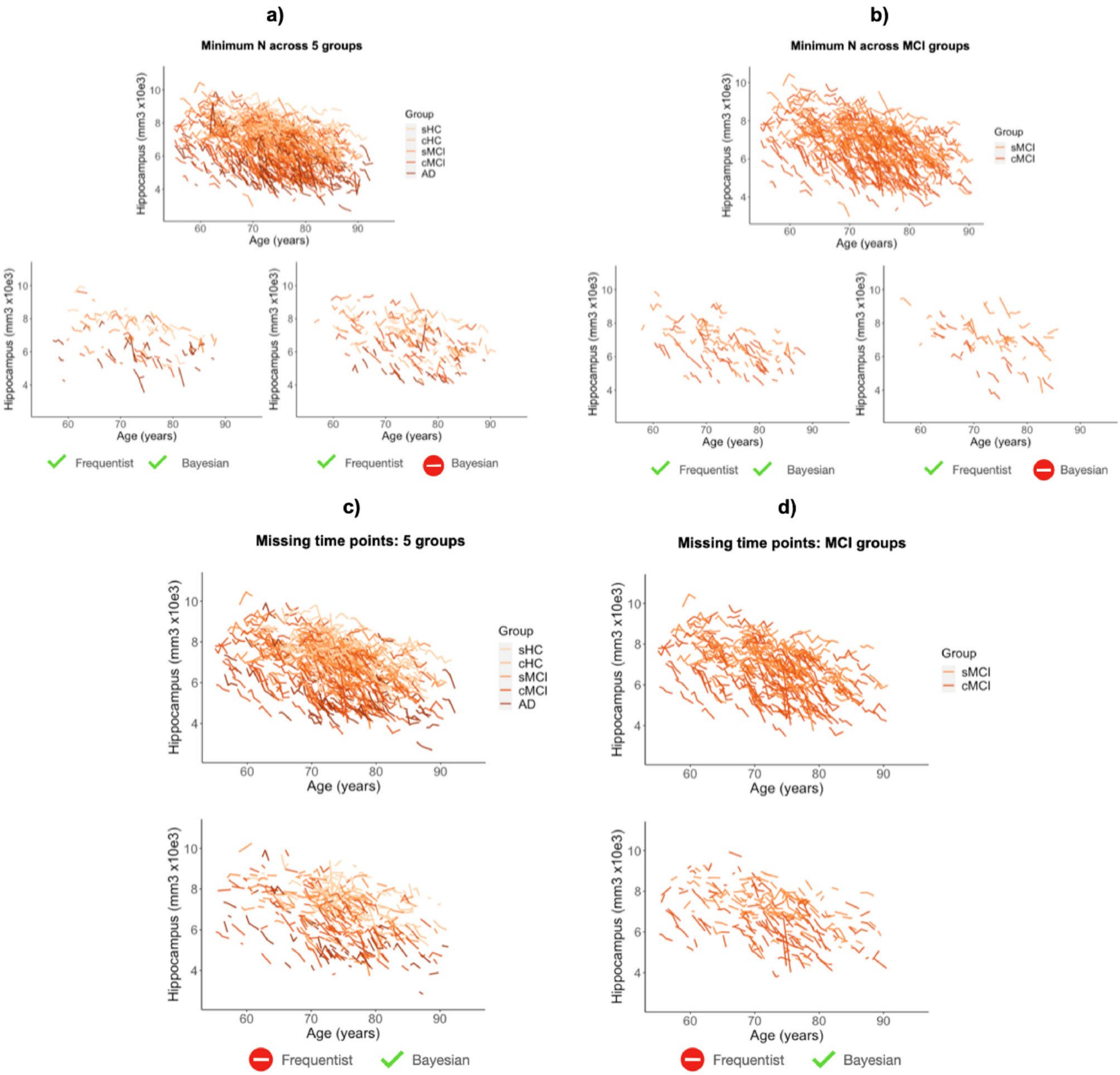

Figure 3. Hippocampus volume versus age. Top plots (**a,b**) represent dataset 1 (**a**) and 2 (**b**) with initial N. Bottom plots (**a,b**) represent four different databases obtained after the simulation of minimum N, resulting significant for frequentist and Bayesian approaches and only for frequentist approach. Top plots (**c,d**) represent dataset 3 (**c**) and 4 (**d**). Bottom plots (**c,d**) represent different databases obtained after the simulation of missing time points, being only estimable for Bayesian approach.

goal to evaluate the behavior of both approaches in different scenarios. To our knowledge, there are no previous studies addressing similar questions with neuroimaging data.

In a further step, we aimed to explore the utility of Bayesian statistics combined with LME modelling. It has been suggested that Bayesian approaches could complement the findings obtained with frequentist analyses, as they provide a more interpretable framework. Bayesian models are based on the direct estimation from the population distribution represented by the posterior distribution, instead of estimating from the hypothetical sampling distribution as it happens in the frequentist approach[13]. In this context, our BLME model can be interpreted in a probabilistic way and may offer a more direct interpretation in clinical settings than FLME[13,14]. Contrarily, the FLME approach does not accept probabilistic interpretation although many researchers use them to interpret their results[39]. As we observed when testing large databases, the two approximations of LME model often led to similar conclusions. Indeed, our results in terms of statistical significance support previous research on the role of HV as a biomarker for AD, being highly significant across groups and between converters and non-converters. However, when repeating all comparisons with BLME, we aimed to add clinical relevance to the above significance statement. This was more evidenced when studying compromised datasets. That is, by using posterior distributions, longitudinal analyses can be better adapted to real-life datasets with clinical relevance. Overall, we emphasize the need of knowing the characteristics of the sample to be able to infer the correct interpretation of the results. For example, it is known that with a non-informative prior, Bayesian approaches tend to mimic frequentist results from the numerical point of view[11]. Little by little more studies use Bayesian approaches in the context of neuroimaging and dementia. In a similar context, Cespedes et al. demonstrated that the use of BLME can be useful for estimating atrophy rates in The Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing cohort (AIBL) (https://aibl.csiro.au). Bayesian statistics were used in the ADNI database in a latent time joint mixed-effects model to provide a continuous alternative to clinical diagnosis[40]. And, in a more complex approach, the authors implemented a multi-task Bayesian learning algorithm on the ADNI database to model trajectories of biomarkers at the individual level[9]. Although these studies clearly differ from ours, they support the use of Bayesian approaches in clinical contexts. In addition, Bayesian statistics appear as a good framework to solve clinically relevant questions that cannot be addressed with frequentist approaches. For example, the absence of effects[10].

We calculated the minimum sample size that led to significant group differences with FLME, and we obtained values (overall values for the studied groups) of 147 and 115 for all groups and for MCI conversion to dementia respectively. It should be mentioned that the main goal of this study was not sample size estimation and thus, these values are rather indicative, in the sense that they are restricted to the research questions and the measure used in this study. However, we believe that they can be of interest in the context of clinical trials. In a study with frontotemporal dementia, Staffaroni et al. calculated the estimated sample size using bootstrapping techniques for different cognitive and imaging measures and they obtained values from < 100 to > 500 depending on the measure or combination of measures chosen[41].

By studying highly compromised datasets (those at the border classical frequentist significance at $p\text{-value} < 0.05$), we were able to compare the two approaches. Notably, not all the borderline databases identified with FLME (i.e., p-value nearly 0.05) remained significant with the BLME approach. This may be due to the fact that accurate modelling of the variances in the Bayesian framework led to more restrictive statistics. It should be mentioned that the ADNI database is quite heterogeneous as do not have the same time point for each subject and that our analyses did not control for some external covariates such as different centers and scanners that might add variability.

In addition, it should be noted that our strategy for comparing approaches was based on selecting the datasets with FLME followed by the evaluation of significance with BLME. This strategy allowed us to obtain important insights as regards significance and interpretability for longitudinal modelling. However, the above conclusions are restricted to this and should not be generalized to any dataset.

The other group of simulations that we implemented was related to databases with missing points. In this sense, an important drawback for FLME modelling is the need of having more subjects' samples than random effects for the model to be estimated. In practice, this was the main reason for stopping in these simulations. Instead, the Bayesian approach allowed estimating the LME model even with high number of missing points in the database. More specifically, our results show that the BLME model is feasible in a 4-timepoint database that has approximately 2 missing values for each subject, suggesting that the Bayesian framework should be chosen for longitudinal modelling in sparse databases. Other studies have demonstrated that Bayesian statistics overcome some of the limitations of classical statistical inference in non-homogeneous databases[3].

Our study has several limitations. First, one difficulty of using the Bayesian approach is its complexity in computing posterior distributions used to estimate the CrI. This has historically imposed an important barrier[13]. Although software solutions have improved in the last years, the frequentist approach is still computationally easier. Further studies should explore the utilization of Markov Chain Monte Carlo approaches to overcome this limitation. Due to this high computational cost of the BLME, the implementation of a method for testing sequential data removal with BLME was out of reach of this study. Second, the current study is based on the HV measure, and the conclusions are specific for this. To be able to generalize our conclusions to broader contexts, other MRI biomarkers for AD, and eventually other databases, should be studied. Third, in the ADNI dataset, there are acquisition differences (i.e., different scanners) which were not included in the analyses. This could have an impact on the HV measurements. Finally, the clinical diagnosis available in the ADNI dataset does not systematically include CSF validation, which, according to the latest MCI diagnostic criteria[42,43], may result in some subjects wrongly labelled as HC or MCI. Other sources of misclassification (or confusing diagnosis) refer to the fact that other pathologies may coexist in subjects diagnosed with AD and that different AD subtypes show different biomarker trajectories. Our results did not account for misdiagnosis nor subgrouping, as ground truth labels were not available. We believe that further studies, possibly using unsupervised machine learning, could account for these factors.

Data availability

Publicly available datasets were analyzed in this study. This data can be found at: adni.loni.usc.edu.

Code availability

Analyses code is available at https://github.com/Agnes2/LME-with-a-Bayesian-and-Frequentist-Approaches.git.

Received: 26 May 2022; Accepted: 5 August 2022
Published online: 24 August 2022

References

1. Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B. & Sabuncu, M. R. Statistical analysis of longitudinal neuroimage data with linear mixed effects models. *Neuroimage* **66**, 249–260 (2013).

2. Cespedes, M. I. *et al.* Comparisons of neurodegeneration over time between healthy ageing and Alzheimer's disease cohorts via Bayesian inference. *BMJ Open* **7**, e012174 (2017).
3. Ziegler, G., Penny, W. D., Ridgway, G. R., Ourselin, S. & Friston, K. J. Estimating anatomical trajectories with Bayesian mixed-effects modeling. *Neuroimage* **121**, 51–68 (2015).
4. Bliese, P. D. *Within-Group Agreement, Non-independence, and Reliability: Implications for Data Aggregation and Analysis* (ScienceOpen, 2000).
5. Guerrero, R. *et al.* Instantiated mixed effects modeling of Alzheimer's disease markers. *Neuroimage* **142**, 113–125 (2016).
6. Rivera-Lares, K., Logie, R., Baddeley, A. & Della Sala, S. Rate of forgetting is independent of initial degree of learning. *Mem. Cognit.* https://doi.org/10.3758/s13421-021-01271-1 (2022).
7. Kieser, M., Friede, T. & Gondan, M. Assessment of statistical significance and clinical relevance. *Stat. Med.* **32**, 1707–1719 (2013).
8. Van Rijn, M. H. C., Bech, A., Bouyer, J. & Van Den Brand, J. A. J. G. Statistical significance versus clinical relevance. *Nephrol. Dial. Transplant.* **32**, 6–12 (2017).
9. Aksman, L. M. *et al.* Modeling longitudinal imaging biomarkers with parametric Bayesian multi-task learning. *Hum. Brain Mapp.* **40**, 3982–4000 (2019).
10. Temp, A. G. M. *et al.* How Bayesian statistics may help answer some of the controversial questions in clinical research on Alzheimer's disease. *Alzheimer's Dement.* **17**, 917–919 (2021).
11. Lesaffre, E. & Lawson, A. B. *Bayesian Biostatistics* (Wiley, 2012).
12. Wilkinson, M. Distinguishing between statistical significance and practical/clinical meaningfulness using statistical inference. *Sports Med.* **44**, 295–301 (2014).
13. Hespanhol, L., Vallio, C. S., Costa, L. M. & Saragiotto, B. T. Understanding and interpreting confidence and credible intervals around effect estimates. *Braz. J. Phys. Ther.* **23**, 290–301 (2019).
14. Berry, D. A. Bayesian clinical trials. *Nat. Rev. Drug Discov.* **5**, 27–36 (2006).
15. Gurrin, L. C., Kurinczuk, J. J. & Burton, P. R. Bayesian statistics in medical research: An intuitive alternative to conventional data analysis. *J. Eval. Clin. Pract.* **6**, 193–204 (2000).
16. van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M. & Depaoli, S. A systematic review of Bayesian articles in psychology: The last 25 years. *Psychol. Methods* **22**, 217–239 (2017).
17. Wagenmakers, E. J. *et al.* Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychon. Bull. Rev.* **25**, 35–57 (2018).
18. Jack, C. R. *et al.* Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* **12**, 207–216 (2013).
19. Hardy, J. Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci.* **20**, 154–159 (1997).
20. Jack, C. R. *et al.* NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's Dement.* **14**, 535–562 (2018).
21. Petersen, R. C. *et al.* Alzheimer's disease neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* **74**, 201–209 (2010).
22. Caroli, A. & Frisoni, G. B. The dynamics of Alzheimer's disease biomarkers in the Alzheimer's Disease Neuroimaging Initiative cohort. *Neurobiol. Aging* **31**, 1263–1274 (2010).
23. Dickerson, B. C. *et al.* MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. *Neurobiol. Aging* **22**, 747–754 (2001).
24. Jack, C. R. *et al.* Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology* **55**, 484–489 (2000).
25. Schuff, N. *et al.* MRI of hippocampal volume loss in early Alzheimers disease in relation to ApoE genotype and biomarkers. *Brain* **132**, 1067–1077 (2009).
26. Zhang, L. *et al.* Longitudinal trajectory of Amyloid-related hippocampal subfield atrophy in nondemented elderly. *Hum. Brain Mapp.* **41**, 2037–2047 (2020).
27. Reuter, M., Schmansky, N. J., Rosas, H. D. & Fischl, B. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* **61**, 1402–1418 (2012).
28. Landau, S. & Jagust, W. *Florbetapir Processing Methods* (2015).
29. Buckner, R. L. *et al.* A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: Reliability and validation against manual measurement of total intracranial volume. *Neuroimage* **23**, 724–738 (2004).
30. Jack, C. R. *et al.* Atrophy rates accelerate in amnestic mild cognitive impairment. *Neurology* **70**, 1740–1752 (2008).
31. Fitzmaurice, G. M., Laird, N. M. & Ware, J. H. *Applied Longitudinal Analysis* (Wiley, 2011).
32. Satterthwaite, F. E. An approximate distribution of estimates of variance components. *Biometr. Bull.* **2**, 110 (1946).
33. Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D. & Wagenmakers, E. J. The fallacy of placing confidence in confidence intervals. *Psychon. Bull. Rev.* **23**, 103–123 (2016).
34. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 201–210 (2015).
35. Team, S. D. *RStan: The R interface to Sta* (2020).
36. Sorensen, T., Hohenstein, S. & Vasishth, S. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quant. Methods Psychol.* **12**, 175–200 (2015).
37. Hubbard, R. *Corrupt Research: The Case for Reconceptualizing Empirical Management and Social Science* (SAGE, 2016). https://doi.org/10.4135/9781506305332.
38. Ziliak, S. & McCloskey, D. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives* (University of Michigan Press, 2008).
39. Pocock, S. J. & Hughes, M. D. Estimation issues in clinical trials and overviews. *Stat. Med.* **9**, 657–671 (1990).
40. Li, D. *et al.* Bayesian latent time joint mixed-effects model of progression in the Alzheimer's disease neuroimaging initiative. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* **10**, 657–668 (2018).
41. Staffaroni, A. M. *et al.* Longitudinal multimodal imaging and clinical endpoints for frontotemporal dementia clinical trials. *Brain* **142**, 443–459 (2019).
42. McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7**, 263–269 (2011).
43. Albert, M. S. *et al.* The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7**, 270–279 (2011).

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Data used in the preparation of this article were obtained from the ADNI database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators is available at: www.loni.ucla.edu/ADNI/Collaboration/ADNI_Authorship_list.pdf.

Author contributions

A.P.M. and R.S.L. contributed to the design of the study. A.P.M., R.T. and R.S. contributed to the analyses of the data. A.P.M., J.C., A.N.B., R.T. and R.S.L. contributed to the interpretation of the data. A.P.M., J.C., R.T. and R.S.L. contributed to the draft of the article. A.N.B., X.S., A.L. and R.S.V. 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.

Funding

This work was supported by the Spanish Ministry of Science and Innovation [Grant Number PID2020-118386RA-I00 to RSL].

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-18129-4.

Correspondence and requests for materials should be addressed to R.S.-L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

Supplementary Material

1 Supplementary Figures and Tables

1.1 Supplementary Tables

Supplementary Table 1: Number of subjects and scans per time point for databases studied with the BLME. The databases were obtained from the frequentist simulations of the minimum N for all the groups.

	N	*sHC (N)*	*cHC (N)*	*sMCI (N)*	*cMCI (N)*	*AD (N)*	*Year 0 (N)*	*Year 0.5 (N)*	*Year 1 (N)*	*Year 2 (N)*
Database 1	92	19	7	25	14	27	92	65	66	36
Database 2	189	31	12	55	51	40	189	154	128	96
Database 3	177	39	10	53	46	29	177	150	126	90
Database 4	161	35	8	43	40	35	161	128	115	87
Database 5	127	26	7	33	37	24	127	101	84	58
Database 6	182	41	9	51	45	36	182	146	130	89
Database 7	175	34	10	53	37	41	175	138	120	88
Database 8	97	22	6	24	23	22	97	73	71	44
Database 9	91	15	5	28	22	21	91	73	55	43
Database 10	156	29	5	47	43	32	156	127	110	64

Supplementary Table 2: Number of subjects and scans per time point for databases studied with the BLME. The databases were obtained from the frequentist simulations of the minimum N for MCI group.

	N	*sMCI (N)*	*cMCI (N)*	*Year 0 (N)*	*Year 0.5 (N)*	*Year 1 (N)*	*Year 2 (N)*
Database 1	60	39	21	60	45	38	33
Database 2	123	62	61	123	102	92	65
Database 3	152	86	66	152	124	108	85
Database 4	128	68	60	128	107	92	77
Database 5	65	33	32	65	50	48	29
Database 6	96	53	43	96	81	70	47
Database 7	135	66	69	135	109	89	76
Database 8	74	38	36	74	54	43	38
Database 9	166	93	71	166	143	124	89
Database 10	133	78	55	133	107	92	71

Supplementary Table 3: Number of subjects and scans per time point for the stopping databases. 95% CrI of the ßs of interest (ß$_{11}$: cHC x time, ß$_{12}$: sMCI x time, ß$_{13}$: cMCI x time and ß$_{14}$: AD x time) LME model fitted with a Bayesian approach. CrI borders are expressed as the 2.5% and 97.5% percentiles. * Indicates that the effect is significant (i.e., CrI does not contain zero).

Stopping Databases	***sHC (N)***	***cHC (N)***	***sMCI (N)***	***cMCI (N)***	***AD (N)***	***Year 0 (N)***	***Year 0.5 (N)***	***Year 1 (N)***	***Year 2 (N)***	***95% CrI* ß$_{11}$**	***95% CrI* ß$_{12}$**	***95% CrI* ß$_{13}$**	***95% CrI* ß$_{14}$**
Database 1	158	45	173	178	67	309	298	314	321	-0.10 0.03	-0.04 0.04	-0.09 - 0.01 *	-0.15 - 0.04 *
Database 2	161	44	173	167	66	305	297	302	318	-0.15 - 0.03 *	-0.08 0.00	-0.18 - 0.10 *	-0.20 - 0.08 *
Database 3	152	42	167	166	64	304	307	286	285	-0.10 0.04	-0.07 0.03	-0.13 - 0.04 *	-0.13 0.01
Database 4	161	50	170	175	65	318	314	299	311	-0.10 0.01	-0.08 0.00	-0.12 - 0.04 *	-0.15 - 0.04 *
Database 5	166	46	169	170	67	338	302	295	301	-0.11 0.02	-0.06 0.02	-0.14 - 0.06 *	-0.14 - 0.06 *
Database 6	158	46	170	169	70	318	321	294	293	-0.14 - 0.02 *	-0.08 0.00	-0.13 - 0.04 *	-0.17 - 0.06 *
Database 7	165	47	169	173	65	323	302	299	314	-0.09 0.03	-0.05 0.03	-0.12 - 0.04 *	-0.15 - 0.05 *
Database 8	160	49	174	169	68	314	306	329	291	-0.15 - 0.02 *	-0.06 0.03	-0.15 - 0.06 *	-0.19 - 0.06 *
Database 9	157	47	174	172	67	323	303	306	302	-0.11 0.02	-0.04 0.04	-0.12 - 0.03 *	-0.12 - 0.01 *
Database 10	150	49	171	168	71	298	299	302	319	-0.06 0.07	-0.04 0.04	-0.09 0.00	-0.12 - 0.01 *

Supplementary Table 4: Number of MCI subjects and scans per time point for the stopping databases. 95% CrI of the ß of interest (cMCI x time) fitted with a Bayesian approach. CrI borders are expressed as the 2.5% and 97.5% percentiles. * Indicates that the effect is significant if the CrI does not contain zero.

Stopping Databases	*sMCI (N)*	*cMCI (N)*	*Year 0 (N)*	*Year 0.5 (N)*	*Year 1 (N)*	*Year 2 (N)*	*95% CrI*
Database 1	177	167	172	162	182	172	-0.12 - 0.03 *
Database 2	169	175	174	163	173	178	-0.13 - 0.02 *
Database 3	178	173	166	188	171	177	-0.15 - 0.05 *
Database 4	172	172	181	164	171	172	-0.14 - 0.04 *
Database 5	169	169	169	157	170	180	-0.14 - 0.05 *
Database 6	176	172	170	156	188	182	-0.13 - 0.03 *
Database 7	174	174	163	172	191	170	-0.10 - 0.01 *
Database 8	179	173	161	194	185	164	-0.10 - 0.02 *
Database 9	170	178	176	175	156	189	-0.09 - 0.01 *
Database 10	172	168	157	167	179	177	-0.14 - 0.03 *

2 Supplementary Methods

Exact expression for the models of the LME models of the Section 2.2

The model with only intercept as random effect can be written as follows, where *i* indicates the subject and *j* indicates the time point:

$$Y_{ij} = \beta_1 + \beta_2 cHC_i + \beta_3 time_{ij} + \beta_4 sMCI_i + \beta_5 cMCI_i + \beta_6 AD_i + \beta_7 APOE4_i + \beta_8 SEX_i$$

$$+\beta_9 AGE_i + \beta_{10} ICV_i + \beta_{11} cHC_i time_{ij} + \beta_{12} sMCI_i time_{ij} + \beta_{13} cMCI_i time_{ij} + \beta_{14} AD_i time_{ij}$$

$$+\beta_{15} APOE4_i time_{ij} + b_{1i} + e_{ij}$$

While the model with intercept and slope as random effects can be written as:

$$Y_{ij} = \beta_1 + \beta_2 cHC_i + \beta_3 time_{ij} + \beta_4 sMCI_i + \beta_5 cMCI_i + \beta_6 AD_i + \beta_7 APOE4_i + \beta_8 SEX_i$$

$$+\beta_9 AGE_i + \beta_{10} ICV_i + \beta_{11} cHC_i time_{ij} + \beta_{12} sMCI_i time_{ij} + \beta_{13} cMCI_i time_{ij} + \beta_{14} AD_i time_{ij}$$

$$+\beta_{15} APOE4_i time_{ij} + b_{1i} + b_{2i} time_{ij} + e_{ij}$$

With the following variables:

Y_{ij}: The variable to predicted. Here, HV.

$time_{ij}$: time from baseline (in years).

cHC_i=1: if subject is cHC and 0 otherwise.

$sMCI_i$=1: if subject is sMCI and 0 otherwise.

$cMCI_i$=1: if subject is cMCI and 0 otherwise.

AD_i=1: if subject is AD and 0 otherwise.

$APOE4_i$=1: if subject has at least one e4 allele and 0 otherwise.

SEX_i=1: if subject female and 0 otherwise.

AGE_i: age at baseline (in years).

ICV_i: total intracranial volume (in litres).

b_{1i}: random effect for the intercept term.

b_{2i}: random effect for the slope term.

e_{ij}: error term

TREBALL 2

Cortical thickness modeling and variability in Alzheimer's disease and frontotemporal dementia.

Agnès Pérez-Millan, Sergi Borrego-Écija, Neus Falgàs, Jordi Juncà-Parella, Beatriz Bosch, Adrià Tort-Merino, Anna Antonell, Nuria Bargalló, Lorena Rami, Mircea Balasa, Albert Lladó, Roser Sala-Llonch, Raquel Sánchez-Valle.

-Acceptat-

Journal of Neurology

IF (2022): 6,0; Q1

Cortical thickness modeling and variability in Alzheimer's disease and frontotemporal dementia

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

[1]Alzheimer's disease and other cognitive disorders Unit. Service of Neurology, Hospital Clínic de Barcelona. Fundació Recerca Clínic Barcelona-IDIBAPS, Barcelona, 08036, Spain.

[2]Institut de Neurociències, University of Barcelona, Barcelona, 08036, Spain.

[3]Department of Biomedicine, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Faculty of Medicine, University of Barcelona, Barcelona, 08036, Spain

[4]Atlantic Fellow for Equity in Brain Health, Global Brain Health Institute, University of California San Francisco, San Francisco, 94143, United States of America.

[5]Image Diagnostic Centre, Hospital Clínic de Barcelona, Barcelona, Spain. CIBER de Salud Mental, Instituto de Salud Carlos III. Magnetic Resonance Image Core Facility, IDIBAPS Barcelona, 08036, Spain.

[6]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, 08036, Spain.

† Corresponding author at Raquel Sanchez-Valle, MD, PhD. Alzheimer's disease and other cognitive disorders unit. Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), University of Barcelona Villarroel, 170 08036 Barcelona (Spain). Tel: +34 932275785
Email address: rsanchez@clinic.cat

* These authors contributed equally to this work.

ORCID LIST:

Agnès Pérez-Millan: https://orcid.org/0000-0002-3006-9792

Sergi Borrego-Écija: https://orcid.org/0000-0003-2557-0010

Neus Falgàs: https://orcid.org/0000-0002-3404-2765

Jordi Juncà-Parella: https://orcid.org/0000-0002-4772-2647

Beatriz Bosch: https://orcid.org/0000-0002-6094-0024

Adrià Tort-Merino: https://orcid.org/0000-0002-5646-0482

Anna Antonell: https://orcid.org/0000-0002-3286-8459

Nuria Bargalló: https://orcid.org/0000-0001-6284-5402

Lorena Rami: https://orcid.org/0000-0002-7411-1921

Mircea Balasa: https://orcid.org/0000-0002-1795-4228

Albert Lladó: https://orcid.org/0000-0002-5066-4150

Roser Sala-Llonch: https://orcid.org/0000-0003-3576-0475

Raquel Sánchez-Valle: https://orcid.org/0000-0001-7750-896X

Acknowledgments

The authors thank patients, their relatives, and healthy controls for participating in the research. This work was supported 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, AGAUR, Generalitat de Catalunya (SGR 2021- 01126) and by Spanish Ministry of Science and Innovation (PID2020-118386RA-I00/AEI/10.13039/501100011033 to Dr. R. Sala-Llonch) and María de Maeztu Unit of Excellence (Institute of Neurosciences, University of Barcelona) MDM-2017-0729. Dr. S. Borrego-Écija is a recipient of the Joan Rodés Josep Baselga grant from FBBVA. Dr. N. Falgàs is a recipient of the Joan Rodes fellowship from the Instituto de Salud Carlos III, Spain.

ABSTRACT

Background and objectives: Alzheimer’s disease (AD) and frontotemporal dementia (FTD) show different patterns of cortical thickness (CTh) loss compared with healthy controls (HC), even though there is relevant heterogeneity between individuals suffering from each of these diseases. Thus, we developed CTh models to study individual variability in AD, FTD, and HC.

Methods: We used the baseline CTh measures of 374 participants obtained from the structural MRI processed with FreeSurfer. A total of 169 AD patients (63 ± 9 years, 65 men), 88 FTD patients (64 ± 9 years, 43 men), and 122 HC (62 ± 10 years, 47 men) were studied. We fitted region-wise temporal models of CTh using Support Vector Regression. Then, we studied associations of individual deviations from the model with cerebrospinal fluid levels of neurofilament light chain (NfL) and 14-3-3 protein and Mini-Mental State Examination (MMSE). Furthermore, we used real longitudinal data from 144 participants to test model predictivity.

Results: We defined CTh spatiotemporal models for each group with a reliable fit. Individual deviation correlated with MMSE for AD and with NfL for FTD. AD patients with higher deviations from the trend presented higher MMSE values. In FTD, lower NfL levels were associated with higher deviations from the CTh prediction. For AD and HC, we could predict longitudinal visits with the presented model trained with baseline data. For FTD, the longitudinal visits had more variability.

Conclusion: We highlight the value of CTh models for studying AD and FTD longitudinal changes and variability and their relationships with cognitive features and biomarkers.

Keywords

Alzheimer's disease, frontotemporal dementia, magnetic resonance imaging, cortical thickness, predictive modeling

1. INTRODUCTION

During the last two decades, the study of several biomarkers, including neuroimaging, has substantially improved the diagnosis of neurodegenerative dementias. However, there’s still a need for reliable biomarkers to track disease, evaluate the effect of experimental drugs, or provide an accurate prognosis. Both Alzheimer’s disease (AD) and frontotemporal dementia (FTD) are characterized by prototypical clinical features and patterns of progressive brain atrophy that constitute the disease's fingerprint or signature. However, both diseases show relatively high individual heterogeneity in presentation and progression rate. The study of this variability is relevant to better understanding these diseases' pathogenesis and predicting disease progression and potentially the effect of treatments [1]. Previous studies also suggest that the age of onset might influence the longitudinal evolution of AD patients [2, 3], emphasizing the need to model age and time from symptoms onset.

Quantitative neuroimaging studies with Magnetic Resonance Imaging (MRI) have been widely used to detect brain changes across these neurodegenerative disorders, using measures such as the cortical thickness (CTh) [4–6], but mostly in research studies. Cerebrospinal fluid (CSF) biomarkers, such as the amyloid-beta protein 42 (Aβ42), the total tau (t-tau), and phosphorylated tau (p-tau) have been included in the current criteria for AD diagnosis [7, 8]. Other CSF-biomarkers, such as 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 [9–12].

Modeling approaches that account for time are of paramount importance to understanding disease progression and comparing brain status across subjects at different disease stages [13–16]. Using structural MRI, some authors described CTh loss with time, providing valuable information on the characterization of disease trajectories and validation of prognostic biomarkers [17, 18].

We hypothesized that dementia is characterized by high variability in atrophy patterns reflecting clinical and biological differences across subjects. In this prospective study with 379 subjects, we aimed to develop CTh models for each diagnosis group (AD, FTD, and healthy controls (HC)), considering time from disease onset. We further aimed to study individual variability with respect to the model and evaluate the effect of the time from disease onset, cognition, and biochemical markers in the individual deviations from the model. Finally, we aimed to test these pseudo-longitudinal models in predicting the real longitudinal CTh evolution of a subsample of subjects.

2. MATERIALS AND METHODS

2.1. Participants

The prospective study includes 379 participants recruited from the Alzheimer's disease and other cognitive disorders group of the Hospital Clínic de Barcelona (HCB), Barcelona, Spain. The study was approved by the HCB Ethics Committee (HCB 2019/0105), complied with the Declaration of Helsinki, and all participants gave written informed consent. All participants underwent a complete clinical and cognitive evaluation [3, 19] and at least one 3T high-resolution structural MRI scan. A subset of these participants had novel CSF measures. A total of 309 participants had available CSF-NfL levels, and 160 participants had available CSF-14-3-3 measures. Additionally, a subset of 144 subjects underwent a second MRI scan 2 years after the baseline and 58 subjects with 4 years after the baseline scan. Participants with a history of stroke, traumatic brain injury, major psychiatric disorder, or alcohol abuse were excluded. Participants were classified into three groups:

- AD: participants meeting criteria for MCI or mild dementia due to AD [7, 8] supported by CSF-biomarkers profile suggesting underlying AD neuropathology according to National Institute on Aging/Alzheimer's Association Research Framework 2018 [20].
- FTD: patients who met 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) [21, 22]. All FTD patients included here showed a CSF profile not suggestive of AD.
- HC: healthy adults having cognitive performance within the normative range (cutoff 1.5 SD from the normative mean).

2.2 CSF-biomarkers

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

2.3 MRI acquisition

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) was acquired 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 FreeSurfer version 6.0 (http://surfer.nmr.mgh.harvard.edu.sire.ub.edu/) to perform cortical reconstruction of the T1-weighted acquisitions [23–25]. FreeSurfer allowed us to generate automated CTh maps for the left and the right hemispheres and to obtain summary measures within regions. We used the 68 cortical parcellations derived from the Desikan atlas available in FreeSurfer [26].

2.5 CTh models

We used the CTh values at baseline to generate three CTh models over time using Support Vector Regression (SVR). For the time variable, we used the chronological age of the subjects for HC subjects and years of disease duration (YDD) for patients. YDD were calculated as the difference between the age at MRI acquisition and the age of disease onset for each patient. All groups were modeled separately. To train the three CTh models, we used the following strategy: the HC model was trained with HC participants, the AD model was trained with AD patients, and the FTD model was trained with FTD patients. We introduced regional measures of both hemispheres leading to a total of 68 CTh values per subject together with time (chronological age or YDD) to train the SVR model. The overall performance of each model was assessed using the root-mean-square error (RMSE) for each CTh region (i.e., the error between estimated and real values), averaged across individuals. During training, the following SVR hyperparameters were introduced with a Grid Search, and we used a cross-validation of 5 folds with the train set (all the cross-sectional data). Models were implemented in Python version 3.10.6 (www.python.org) with the scikit-learn library [27].

2.6. Correlation with CSF-biomarkers and cognition

We calculated the individual residuals of the model for each region and disease (Figure 1A). We obtained deviations from the HC-derived model for each patient, and we computed the correlation between these residuals and Mini-Mental State Examination (MMSE) scores and YDD. Then, we estimated the deviations from its own disease model and computed the correlation between the obtained residuals and available CSF-biomarkers levels (NfL and 14-3-3) and MMSE scores. The FTD group was first studied as a whole and further divided into subgroups (bvFTD, svPPA, and nfvPPA). A p-value<0.05 was set as the threshold for significance in the correlations after applying the correction for multiple comparisons for the number of regions with Benjamini-Hochberg for all the analyses. Correlations were implemented in R language, version 4.2.1 (https://www.r-project.org).

2.7 Prediction of longitudinal changes in CTh maps

The longitudinal data were used as unseen test samples to obtain individual predictions using the group-specific CTh models. Then, we calculated the mean absolute error (MAE) between predicted and real measures for each region (Figure 1B). We also studied the different subgroups of FTD patients (bvFTD, nfvPPA, and svPPA). These comparisons were also implemented in R version 4.2.1.

3. RESULTS

3.1. Sample demographics

Of the 379 subjects included in the analyses, 169 were AD patients (n=29, 2 years follow-up), 88 were FTD (n=27, with 2 years follow-up, and n=9, with 4 years follow-up), and 122 were HC (n=88, with 2 years follow-up, and n=49, with 4 years follow-up). The FTD group included 47 bvFTD patients (n=11, with 2 years follow-up, and n=4, with 4 years follow-up), 22 svPPA patients (n=9, with 2 years follow-up, and n=3, with 4 years follow-up), 17 nfvPPA patients (n=6, 2 years follow-up), and 2 unspecified PPA patients (n=1, 2 years follow-up). Of the 309 subjects with available NfL levels, 144 were AD patients, 63 were FTD patients, and 102 were healthy controls. For the 14-3-3 samples, we had 66 AD patients, 54 FTD patients, and 40 healthy controls with available data for our analysis. Figure 2 shows a schema of the samples available for each analysis. Demographic information and group statistics are shown in Table 1 and in Supplementary Material. In summary, as expected, CSF-biomarkers levels and MMSE scores showed significant differences between groups (corrected p-value<0.05). Notably, the time between scans was modeled inside the models to control the potential differences between individuals.

3.2. CTh models with time and correlations with disease duration and cognition

We obtained a CTh model for HC with chronological age. The mean RMSE of the model was 0.024. Then, we obtained individual deviations for HC, AD, and FTD subjects from this HC model using the residuals (Figure 3A). HC subjects showed values around zero, meaning that the observed and estimated values are closer for most subjects, indicating a good fit. We obtained high residual values (absolute values between 0.01 and 0.77) for the AD and FTD, indicating high deviations from the HC-defined model (Figure 3A). The negative residuals indicate that the real CTh values of AD and FTD subjects were lower than those estimated by the HC model. In AD, higher deviations from the HC model were only associated with higher YDD in the regions: right bankssts, right transverse temporal, right lateral orbitofrontal, and right frontal pole and with MMSE in temporal and parietal regions (corrected p-value<0.05, Figure 3B). No significant correlations were identified in FTD.

3.3. CTh models with years of disease duration and correlations with CSF-biomarkers and cognition

We estimated CTh models for AD and FTD using YDD in the temporal axis. We obtained estimations of CTh in a span from 0 to 9 YDD, and we calculated the individual variations between each individual point and the estimation. The maps of variability across regions and time are shown in Figure 4. The mean individual deviation from the model (RMSE values) for these models was: 0.08 for FTD and 0.04 for AD. We also calculated the mean values by YDD subgroups (see Supplementary Material). We studied the correlation between individual residuals from their disease model and individual CSF-NfL and CSF-14-3-3 levels and MMSE scores. In AD, the residuals of the right precentral and right entorhinal regions had a significant negative correlation with NfL levels (corrected p-value<0.05, Figure 5A). The residuals of the temporal and parietal lobes positively correlated with MMSE scores, so participants with higher residual values correspond to participants with higher MMSE scores (corrected p-value<0.05, Figure 5A). We did not find any correlation between 14-3-3 levels and the residuals for AD patients. In FTD, we found significant negative correlations (p-value<0.05, corrected) between NfL values and the residuals of the model in regions of the frontal, temporal, and parietal lobes (Figure 5B). For 14-3-3 levels and MMSE scores, we did not find significant results. When we studied the different FTD variants, we found a negative correlation between NfL and individual residuals for bvFTD in frontal, temporal, and parietal regions (corrected p-value<0.05, see Supplementary Material). For the other variants (svPPA and nfvPPA), we did not find any correlation with NfL. However, high caudal middle frontal right region deviations were associated with high MMSE scores for the nfvPPA patients (corrected p-value<0.05, see Supplementary Material). As before, 14-3-3 levels showed no significant correlation for bvFTD, svPPA, or nfvPPA. These correlations remained significant when age was included as a covariate.

3.4. CTh prediction of follow-up visits

Finally, we explored whether the predictive CTh models trained with the baseline data could be used to predict future real CTh values. For that, we calculated the mean absolute error (MAE) values contrasting the real values (Figure 6) at each longitudinal timepoint with their predicted disease model values. Also, we estimated the MAE (computed internally within the model data) for the baseline visit to use as a reference. For HC and AD (Figure 6A), the predictive model trained with baseline data allows the prediction of future visits with low error, as the mean MAE were 0.12 and 0.15, respectively, for all the timepoints. For FTD (Figure 6A), the 2 years follow-up visit could be predicted reasonably well (mean MAE 0.20 at baseline and 0.22 at 2 years), however, the model could not present an optimal prediction for the 4 years follow-up visit (mean MAE 0.25). We repeated the analysis with FTD data divided into subgroups (Figure 6B), (bvFTD, nfvPPA, and svPPA). Despite the limited sample size, we found results like the whole FTD group for bvFTD and svPPA. For the nfvPPA patients (only data at 2 years follow-up), we find the worst results in terms of predictive accuracy. In the Supplementary Material, we specify the mean MAE for each group and subgroup.

4. DISCUSSION

In this work, we implemented group-specific models of whole brain CTh according to chronological age for HC and YDD for AD and FTD patients using SVR. We modeled CTh changes with time, and we demonstrated these models' capability to identify individual variations as deviations from the norm. We studied individual deviations from the model defined for HC and from the disease-specific models. In AD, individual deviations from the HC-defined model correlated with YDD and MMSE scores. Using the disease-specific models, we found a significant inverse correlation between CSF-NfL levels and the deviation from the model in most of the brain regions for FTD patients, providing additional evidence of the relationship between imaging and fluid biomarkers in FTD. Furthermore, we found a positive correlation for AD between the MMSE scores and the deviation from the model in multiple regions. Finally, using real longitudinal data, we performed an exploratory study to evaluate the ability of the models described above to obtain individual predictions of future CTh values.

In healthy subjects, our model estimated whole-brain CTh using the chronological age of the subjects with residual values near zero, indicating a good fit. Several studies and multi-centric initiatives have recently focused on studying normative models with age in healthy samples [15]. Even if our healthy sample is limited compared to these studies, our results align with what has been described. In addition, we offer the applicability of such models with longitudinal data with satisfactory results. Furthermore, as deviations from normality have been widely used to assess disease severity and cognitive variability in youth [28] and in psychiatric disorders [29], we study if AD and FTD patients could be described with the previous HC model. As expected, individual AD and FTD CTh data diverged from the model defined with HC subjects. In AD, this divergence was significantly correlated with YDD and MMSE but not in FTD, suggesting that both time and general cognition impact more consistently in CTh changes compared to HC in AD. In AD, the regions of the temporal, parietal, and occipital lobes were the ones with the highest deviations from the HC model. These regions have been reported to be the most affected in patients with AD [30–33]. Therefore, our study can identify meaningful atrophy patterns using estimation models. On the other hand, the temporal and frontal lobes were the ones with the highest deviations for the FTD patients compared to the HC trend. In addition, temporal regions presented higher deviation than the AD group, depicting higher structural alterations in FTD [11, 32, 34, 35].

Given that models derived from the HC group that used chronological age showed high variability for these groups, we expected that modeling disease duration could better capture these variations. In this sense, we described CTh models in AD and FTD using YDD instead of the chronological age. Even though both AD and FTD individual values evidence the existence of variability within these diseases. These results complement other research highlighting these models to assess heterogeneity in neurodegenerative disorders [1, 36, 37].

We found that the individual AD deviations from the AD CTh model in the right precentral and right entorhinal inversely correlated with the NfL levels. At the regional level, the correlations between MMSE scores and the deviations from the model in AD were in the temporal and parietal lobes. In AD, MMSE has been associated with CTh [4, 33, 38]. In our case, the AD patients who deviated more from the model trend had high MMSE scores. Previous studies such as Valtteri et al. 2010 [38] showed that less CTh in frontal and temporal lobes is associated with lower MMSE scores in the AD group, which coincides with our regions and the direction of the correlation. Even if we study variability and not direct CTh measures, the regions that appear significant in the correlation analysis have high coincidence with the AD cortical pattern, which includes the temporal and parietal lobes appear [30, 31, 33, 34].

The deviations of the FTD patients from their own model presented correlations with CSF-NfL levels in the frontal, temporal, and parietal lobes. The frontal and temporal areas are known to be the most affected in patients with FTD compared to HC [32, 34, 39]. In FTD, NfL levels have been suggested as a maker of the disease severity [40, 41]. We found that higher levels of NfL corresponded to lower residual values. Thus, patients with CTh patterns that adjusted well to the model had higher CSF-NfL levels, suggesting that the model captures the trend of neurodegeneration. Instead, lower NfL levels could correspond to patients that move further away from the FTD model, as has been described for slowly progressive FTD or non-progressors FTD [42]. Indeed, when studying the different FTD variants, we found that the main pattern was mainly driven by bvFTD, as other variants, such as PPA, showed much higher deviation. In addition, the fact that the parietal lobule emerged, which has not been reported previously, could be due to the high presence of bvFTD patients in the FTD sample. At the same time, the individual deviations in PPA patients did not present any significant correlation with NfL, suggesting that the heterogeneity in these patients' CTh pattern is not related to CSF-NfL levels. These results might provide additional support to the use of NfL levels as a marker of neurodegeneration and disease severity in FTD, paving the way for its future use as an outcome measure for clinical practice. In addition, the inexistent correlation with MMSE scores in FTD patients could be due to the fact that MMSE is not a sensitive cognitive measure in these patients, contrary to AD patients.

At the longitudinal level, we explored the application of these predictive models, trained at a cross-sectional level, to predict CTh patterns for future visits. We used the longitudinal data to know if the pseudo-longitudinal models developed with baseline data could predict the CTh values of future visits of these patients. We compared the predicted values of the longitudinal data with the real values of those visits. For healthy subjects and AD, we found that the model could predict well the CTh values at future visits. For FTD patients, the variability observed at baseline was reproduced for the longitudinal data and increased at each visit. Then, we focus on the FTD variants; the bvFTD patients fit best in the future prediction model, and the nfvPPA patients were the worst. Our results align with previous studies showing that FTD is phenotypically heterogeneous [21, 22]. and support that an independent predictive model should be created for each FTD phenotypic variant. Similarly, in a recent study of genetic FTD, Poos et al. [43], identified differences between the three major FTD gene mutation carriers' trajectories, suggesting that each genetic FTD subtype should also be modeled separately. The fit showed more FTD heterogeneity than AD at baseline and longitudinal levels. Overall, the proposed models could be the starting point to be able to differentiate different subtypes or dementias through the estimation of their CTh values or variabilities. Thus, further studies with a bigger cohort should study this point in detail.

Our study has some limitations. One of the major limitations of the study is the sample size. The sample size limited the generation of different CTh models for each FTD variant, especially all models at the third visit. However, the unicentric nature of the study, even if limited in the sample size, also provided homogeneity to the data acquisition. We only tested the effect of two selected neurodegeneration biomarkers in the individual residuals from their disease model. The selection of these two biomarkers was, at a certain point, arbitrary. Future studies could evaluate whether other markers could better explain the individual variability.

In conclusion, SVR provides the opportunity to generate CTh disease models to predict longitudinal changes and to study individual variability in AD, FTD, and healthy individuals and their relationships with cognitive features and biomarkers.

Conflict of Interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Verdi S, Marquand AF, Schott JM, Cole JH (2021) Beyond the average patient: how neuroimaging models can address heterogeneity in dementia. Brain 144:2946–2953. https://doi.org/10.1093/brain/awab165

2. Mendez MF (2012) Early-onset Alzheimer's Disease: Nonamnestic Subtypes and Type 2 AD. Archives of Medical Research 43:677–685. https://doi.org/10.1016/j.arcmed.2012.11.009

3. Tort-Merino A, Falgàs N, Allen IE, et al (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

4. Contador J, Pérez-Millan A, Guillen N, et al (2021) Baseline MRI atrophy predicts 2-year cognitive outcomes in early-onset Alzheimer's disease. Journal of Neurology. https://doi.org/10.1007/s00415-021-10851-9

5. Dickerson BC, Wolk DA (2012) MRI cortical thickness biomarker predicts AD-like CSF and cognitive decline in normal adults. Neurology 78:84–90. https://doi.org/10.1212/WNL.0b013e31823efc6c

6. Du AT, Schuff N, Kramer JH, et al (2007) Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain 130:1159–1166. https://doi.org/10.1093/brain/awm016

7. Albert MS, DeKosky ST, Dickson D, et al (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

8. McKhann GM, Knopman DS, Chertkow H, et al (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's and Dementia 7:263–269. https://doi.org/10.1016/j.jalz.2011.03.005

9. Alcolea D, Vilaplana E, Suárez-Calvet M, et al (2017) CSF sAPPβ, YKL-40, and neurofilament light in frontotemporal lobar degeneration. Neurology 89:178–188. https://doi.org/10.1212/WNL.0000000000004088

10. Antonell A, Tort-Merino A, Ríos J, et al (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

11. Falgàs N, Ruiz-Peris M, Pérez-Millan A, et al (2020) Contribution of CSF biomarkers to early-onset Alzheimer's disease and frontotemporal dementia neuroimaging signatures. Human Brain Mapping hbm.24925. https://doi.org/10.1002/hbm.24925

12. McFerrin MB, Chi X, Cutter G, Yacoubian TA (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

13. Coupé P, Manjón JV, Lanuza E, Catheline G (2019) Lifespan Changes of the Human Brain In Alzheimer's Disease. Sci Rep 9:3998. https://doi.org/10.1038/s41598-019-39809-8

14. Pérez-Millan A, Contador J, Juncà-Parella J, et al (2023) Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. Human Brain Mapping. https://doi.org/10.1002/hbm.26205

15. Pomponio R, Erus G, Habes M, et al (2020) Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. NeuroImage 208:116450. https://doi.org/10.1016/j.neuroimage.2019.116450

16. Wang T, Qiu RG, Yu M (2018) Predictive Modeling of the Progression of Alzheimer's Disease with Recurrent Neural Networks. Sci Rep 8:9161. https://doi.org/10.1038/s41598-018-27337-w

17. Storsve AB, Fjell AM, Tamnes CK, et al (2014) Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: Regions of accelerating and decelerating change. Journal of Neuroscience 34:8488–8498. https://doi.org/10.1523/JNEUROSCI.0391-14.2014

18. Risacher SL, Shen L, West JD, et al (2010) Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort. Neurobiology of Aging 31:1401–1418. https://doi.org/10.1016/j.neurobiolaging.2010.04.029

19. Contador J, Pérez-Millan A, Guillén N, et al Sex differences in early-onset Alzheimer's disease. European Journal of Neurology n/a: https://doi.org/10.1111/ene.15531

20. Jack CR, Bennett DA, Blennow K, et al (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

21. Gorno-Tempini ML, Hillis AE, Weintraub S, et al (2011) Classification of primary progressive aphasia and its variants. Neurology 76:1006–14. https://doi.org/10.1212/WNL.0b013e31821103e6

22. Rascovsky K, Hodges JR, Knopman D, et al (2011) Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 134:2456–2477. https://doi.org/10.1093/brain/awr179

23. Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of the United States of America 97:11050–11055. https://doi.org/10.1073/pnas.200033797

24. Fischl B, Van Der Kouwe A, Destrieux C, et al (2004) Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex 14:11–22. https://doi.org/10.1093/cercor/bhg087

25. Reuter M, Schmansky NJ, Rosas HD, Fischl B (2012) Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 61:1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084

26. Desikan RS, Ségonne F, Fischl B, et al (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31:968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

27. Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12:2825–2830

28. Kjelkenes R, Wolfers T, Alnæs D, et al (2022) Deviations from normative brain white and gray matter structure are associated with psychopathology in youth. Dev Cogn Neurosci 58:101173. https://doi.org/10.1016/j.dcn.2022.101173

29. Rutherford S, Kia SM, Wolfers T, et al (2022) The normative modeling framework for computational psychiatry. Nat Protoc 17:1711–1734. https://doi.org/10.1038/s41596-022-00696-5

30. Whitwell JL, Clifford RJ, Przybelski SA, et al (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

31. Möller C, Vrenken H, Jiskoot L, et al (2013) Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. Neurobiology of Aging 34:2014–2022. https://doi.org/10.1016/j.neurobiolaging.2013.02.013

32. Rabinovici GD, Seeley WW, Kim EJ, et al (2008) Distinct MRI atrophy patterns in autopsy-proven Alzheimer's disease and frontotemporal lobar degeneration. American Journal of Alzheimer's Disease and other Dementias 22:474–488. https://doi.org/10.1177/1533317507308779

33. Blanc F, Colloby SJ, Philippi N, et al (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

34. Möller C, Hafkemeijer A, Pijnenburg YAL, et al (2015) Joint assessment of white matter integrity, cortical and subcortical atrophy to distinguish AD from behavioral variant FTD: A two-center study. NeuroImage: Clinical 9:418–429. https://doi.org/10.1016/j.nicl.2015.08.022

35. Couto B, Manes F, Montañés P, et al (2013) Structural neuroimaging of social cognition in progressive non-fluent aphasia and behavioral variant of frontotemporal dementia. Front Hum Neurosci 7:467. https://doi.org/10.3389/fnhum.2013.00467

36. Wolfers T, Beckmann CF, Hoogman M, et al (2020) Individual differences v. the average patient: mapping the heterogeneity in ADHD using normative models. Psychological Medicine 50:314–323. https://doi.org/10.1017/S0033291719000084

37. Zabihi M, Oldehinkel M, Wolfers T, et al (2019) Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 4:567–578. https://doi.org/10.1016/j.bpsc.2018.11.013

38. Julkunen V, Niskanen E, Koikkalainen J, et al (2010) Differences in Cortical Thickness in Healthy Controls, Subjects with Mild Cognitive Impairment, and Alzheimer's Disease Patients: A Longitudinal Study. Journal of Alzheimer's Disease 21:1141–1151. https://doi.org/10.3233/JAD-2010-100114

39. Seelaar H, Rohrer JD, Pijnenburg YAL, et al (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

40. Rohrer JD, Woollacott IOC, Dick KM, et al (2016) Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. Neurology 87:1329–1336. https://doi.org/10.1212/WNL.0000000000003154

41. Scherling CS, Hall T, Berisha F, et al (2014) Cerebrospinal fluid neurofilament concentration reflects disease severity in frontotemporal degeneration. Annals of Neurology 75:116–126. https://doi.org/10.1002/ana.24052

42. Eratne D, Keem M, Lewis C, et al (2022) Cerebrospinal fluid neurofilament light chain differentiates behavioural variant frontotemporal dementia progressors from non-progressors. J Neurol Sci 442:120439. https://doi.org/10.1016/j.jns.2022.120439

43. Poos JM, Grandpierre LDM, Ende EL van der, et al (2022) Longitudinal Brain Atrophy Rates in Presymptomatic Carriers of Genetic Frontotemporal Dementia. Neurology 99:e2661–e2671. https://doi.org/10.1212/WNL.0000000000201292

Figures & Tables

Table 1: Group summaries are given as each measure's mean and standard deviation. Differences between groups are calculated using Fisher's Test for sex or the ANOVA Test for the rest of the variables. Significant group differences are highlighted in bold, and pairwise differences were measured with a Benjamini-Hochberg correction p-value. HC: healthy controls, AD: Alzheimer's disease, FTD: frontotemporal dementia, MMSE: Mini-Mental State Examination, NfL: neurofilament light chain.

	HC	**FTD**	**AD**	**HC-FTD p-value**	**HC-AD p-value**	**FTD-AD p-values**
N at first MRI	122	88	169	------	------	------
N at second MRI	88	27	29	------	------	------
N at third MRI	49	9	------	------	------	------
Sex at first MRI, Men/Women	47/75	43/45	65/104	0.24	1.0	0.24
Sex at second MRI, Men/Women	29/59	14/13	12/17	0.33	0.59	0.59
Sex at third MRI, Men/Women	15/34	5/4	------	0.25	------	------
Age at first MRI, years (SD)	61.9 (9.8)	63.6 (8.6)	63.3 (9.4)	0.29	0.29	0.79
Years of Disease at first MRI, years (SD)	------	3.3 (2.1)	2.9 (1.8)	------	------	0.10
N MMSE	107	67	140	------	------	------
Mean MMSE (SD)	28.7 (1.2)	24.5 (4.5)	22.3 (4.5)	**5.4e-12**	**<2.0e-16**	**7.9e-5**
N NfL	102	63	144	------	------	------
Mean Nfl, pg/mL (SD)	564.3 (312.7)	2358.9 (1765.9)	1122.8 (616.9)	**<2.0e-16**	**3.9e-6**	**<2.0e-16**
N 14-3-3	40	54	66	------	------	------
Mean 14-3-3, pg/mL (SD)	2637.3 (734.7)	4381.4 (1931.6)	6458.4 (591.3)	**0.037**	**1.2e-5**	**0.0076**

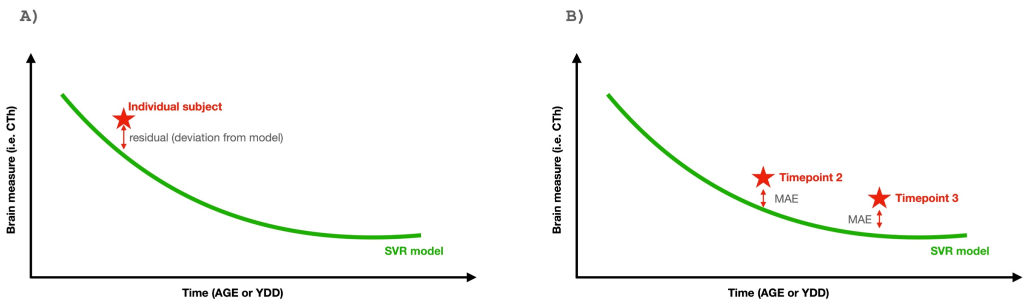
A)
Brain measure (i.e. CTh)
Individual subject
residual (deviation from model)
SVR model
Time (AGE or YDD)
B)
Brain measure (i.e. CTh)
Timepoint 2
MAE
Timepoint 3
MAE
SVR model
Time (AGE or YDD)

Figure 1: Cartoon representation of the methodology used. A) CTh model (green line) and estimation of a sample individual's deviation from the model. B) Methodology used to test the model with real follow-up visits. YDD: years of disease duration, SVR: Support Vector Regression, MAE: mean absolute error.

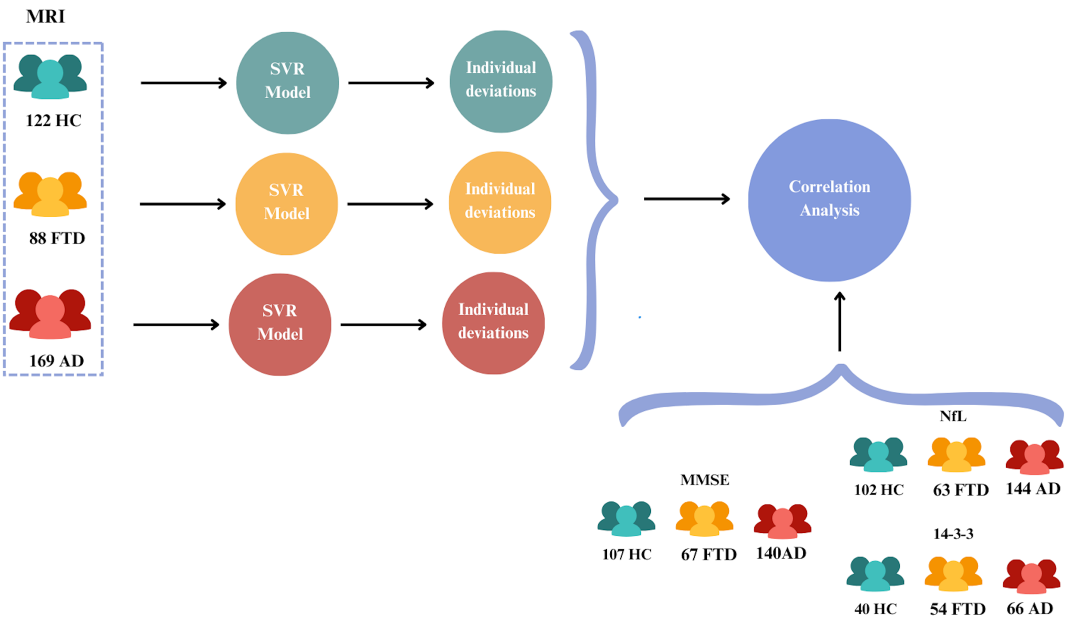
MRI
122 HC
88 FTD
169 AD
SVR Model
SVR Model
SVR Model
Individual deviations
Individual deviations
Individual deviations
Correlation Analysis
MMSE
107 HC
67 FTD
140AD
NfL
102 HC
63 FTD
144 AD
14-3-3
40 HC
54 FTD
66 AD

Figure 2. Study workflow, including the number of subjects for each group and biomarker.

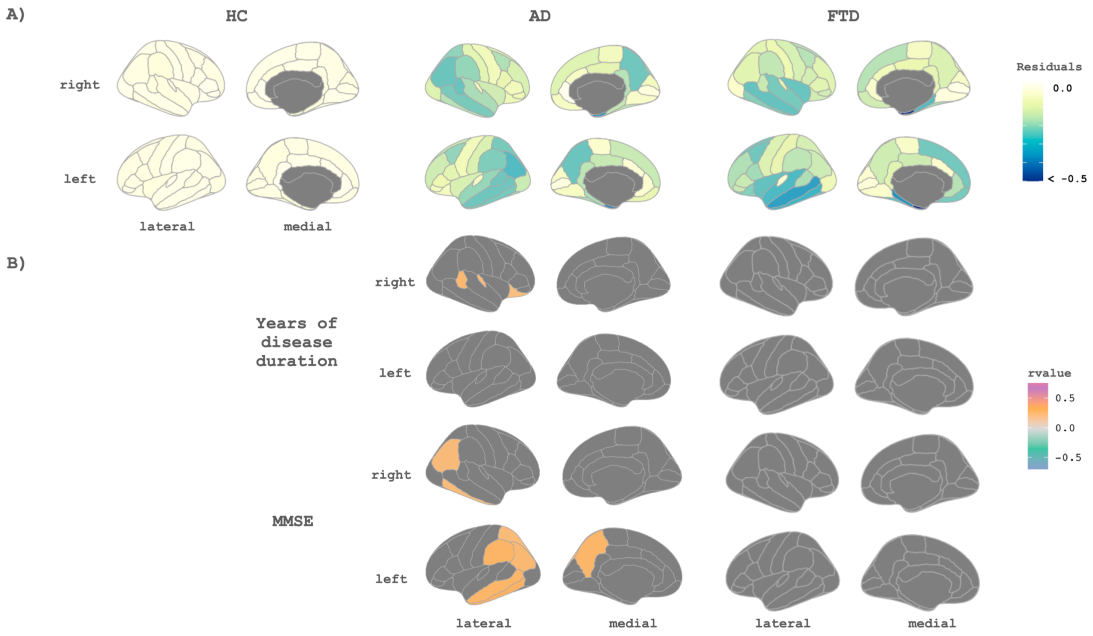
A)
HC
AD
FTD
right
left
lateral
medial
Residuals
0.0
< -0.5
B)
Years of disease duration
MMSE
right
left
right
left
rvalue
0.5
0.0
-0.5
lateral
medial
lateral
medial

Figure 3: A) *Maps of the residual values for each group against the proposed healthy subjects' model for cortical thickness. B) Maps of correlations between residuals against the HC model of cortical thickness and clinical biomarkers in HC, AD, and FTD subjects. Only significant regions are shown, the threshold was set at 0.05 with a p-value adjusted with multiple comparisons of all the CTh regions. HC: healthy controls, AD: Alzheimer's disease, FTD: frontotemporal dementia, MMSE: Mini-Mental State Examination.*

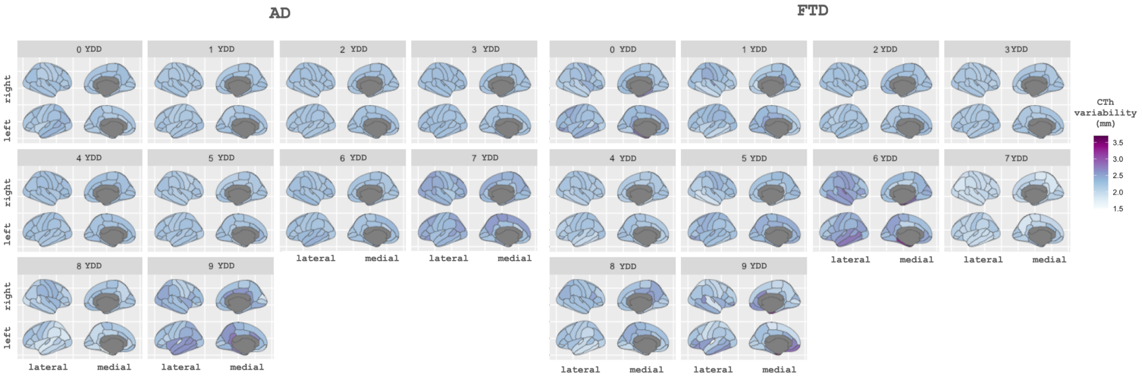
AD
FTD
0 YDD
1 YDD
2 YDD
3 YDD
4 YDD
5 YDD
6 YDD
7 YDD
8 YDD
9 YDD
right
left
lateral
medial
CTh variability (mm)
3.5
3.0
2.5
2.0
1.5

Figure 4: Pattern of variability of the cortical thickness estimation for each brain region according to years of disease (0 to 9 years) for each disease group. AD: Alzheimer's disease, FTD: frontotemporal dementia, CTh: cortical thickness, YDD: years of disease duration.

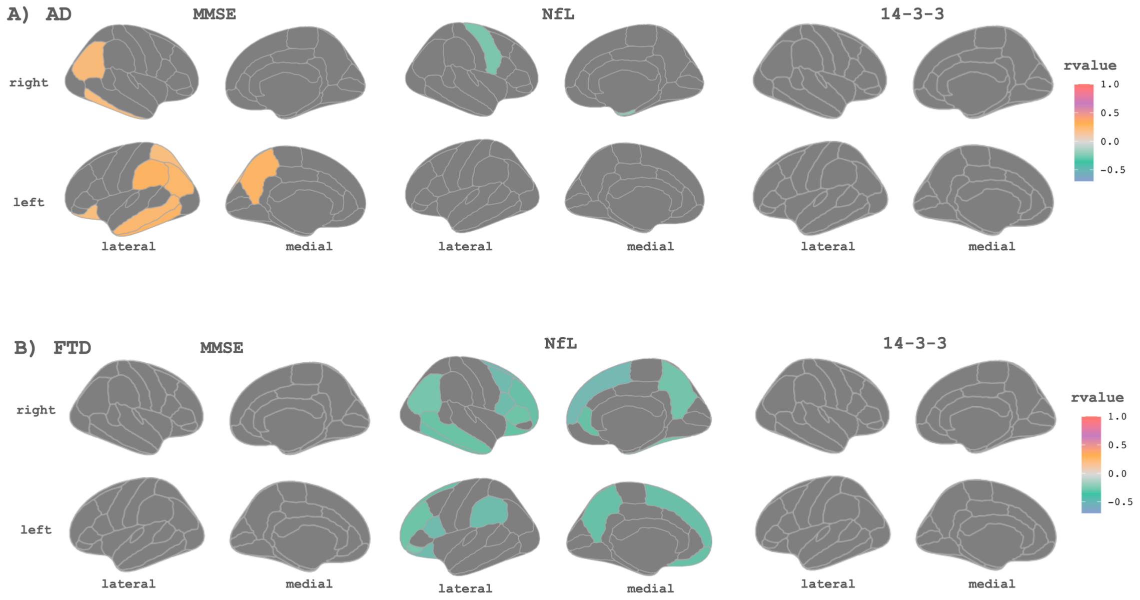
A) AD
MMSE
NfL
14-3-3
right
left
lateral
medial
rvalue
1.0
0.5
0.0
-0.5
B) FTD
MMSE
NfL
14-3-3
right
left
lateral
medial
rvalue
1.0
0.5
0.0
-0.5

Figure 5: Maps of correlations between individual residuals for the group-specific cortical thickness model and cerebrospinal fluid-biomarkers and MMSE scores for AD and FTD participants. Only significant regions are shown. The threshold was set at 0.05 with a p-value adjusted with multiple comparisons of all the CTh regions. *A) Correlations within the AD group. B) Correlations within the FTD group. AD: Alzheimer's disease, FTD: frontotemporal dementia, MMSE: Mini-Mental State Examination, NfL: neurofilament light chain.*

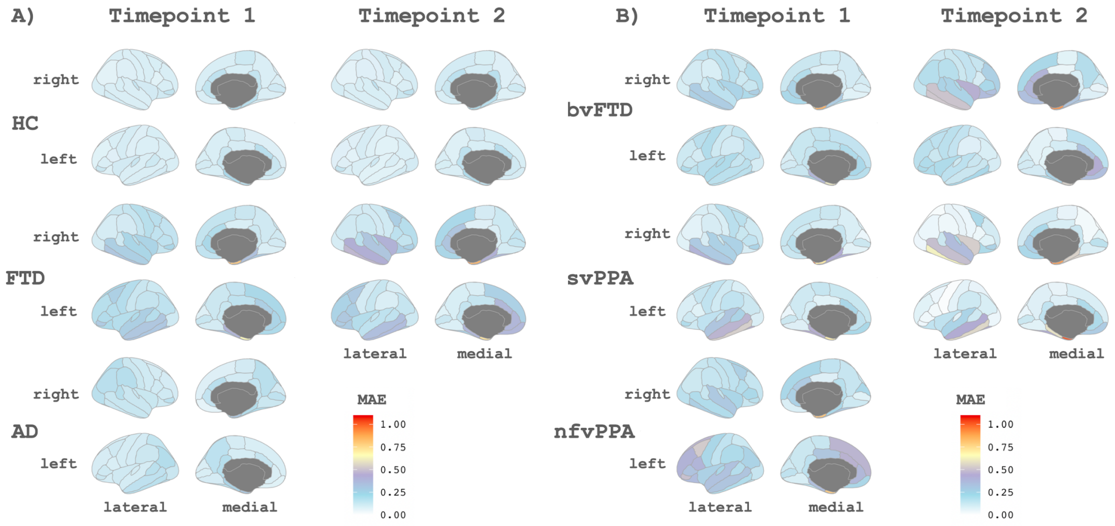
A)
Timepoint 1
Timepoint 2
HC
FTD
AD
right
left
lateral
medial
MAE
1.00
0.75
0.50
0.25
0.00
B)
Timepoint 1
Timepoint 2
bvFTD
svPPA
nfvPPA
right
left
lateral
medial
MAE
1.00
0.75
0.50
0.25
0.00

Figure 6: A) Maps of the mean absolute error values of the predictions for the longitudinal data for each region for the HC, AD, and FTD groups. B) Maps of the mean absolute error values of the predictions for the longitudinal data for the different FTD variants predicted with the FTD model. HC: healthy controls, AD: Alzheimer's disease, FTD: frontotemporal dementia, bvFTD: behavior frontotemporal dementia, svPPA: semantic variant primary progressive aphasia, nfvPPA: nonfluent variant primary progressive aphasia, MAE: mean absolute error.

Supplementary Material

Sample demographics

We present the sample demographics for the different variants of frontotemporal dementia (FTD) participants in Table 1.

Table 1: Group summaries are given as each measure's mean and standard deviation. Differences between groups are calculated using Fisher Test for sex or the ANOVA Test for the rest of the variables. Significant group differences are highlighted in bold, pairwise differences were measured with a Benjamini-Hochberg correction p-value). bvFTD: behavior frontotemporal dementia, svPPA: semantic variant primary progressive aphasia, nfvPPA: nonfluent variant primary progressive aphasia, MMSE: Mini-Mental State Examination, NfL: neurofilament light chain.

	bvFTD	**svPPA**	**nfvPPA**	**bvFTD-svPPA p-value**	**bvFTD-nfvPPA p-value**	**svPPA-nfvPPA p-value**
N at first MRI	47	22	17	------	------	------
N at second MRI	11	9	6	------	------	------
N at third MRI	4	3	0	------	------	------

Sex at first MRI, Men/Women	23/24	13/9	7/10	0.68	0.78	0.68
Sex at second MRI, Men/Women	6/5	6/3	2/4	0.67	0.67	0.67
Sex at third MRI, Men/Women	2/2	------	2/1	1.0	------	------
Age at first MRI, years (SD)	63.6 (9.1)	64.0 (8.0)	64.0 (8.8)	0.99	0.99	0.99
Years of Disease at first MRI, years (SD)	3.4 (2.1)	3.0 (2.4)	3.6 (1.8)	0.77	0.75	0.75
N MMSE	37	17	11	------	------	------
Mean MMSE (SD)	24.3 (5.0)	24.4 (3.3)	25.9 (4.4)	0.93	0.60	0.60
N NfL	32	18	12	------	------	------
Mean Nfl, pg/mL (SD)	2215.2 (1988.6)	2167.9 (1146.1)	2976.9 (1957.1)	0.93	0.34	0.34
N 14-3-3	29	15	10	------	------	------

Mean 14-3-3, pg/mL (SD)	4534.2 (2143.3)	4585.7 (1710.2)	3632.1 (1536.7)	0.93	0.35	0.35

CTh models with years of disease duration and correlations with CSF-biomarkers and cognition

We present the results of the mean residuals of the model grouped by baseline YDD in Table 2.

Table 2: Mean of the CTh residuals across all regions for each model and separated by years of disease duration. AD: Alzheimer's disease, FTD: frontotemporal dementia, and YDD: Years of disease duration.

	FTD	**AD**
0 YDD, mean residual (SD)	0.04 (0.08)	0.02 (0.10)
1 YDD, mean residual (SD)	0.01 (0.14)	0.02 (0.13)
2 YDD, mean residual (SD)	0.01 (0.13)	0.01 (0.12)
3 YDD, mean residual (SD)	0.04 (0.14)	0.01 (0.08)
4 YDD, mean residual (SD)	0.05 (0.16)	0.00 (0.10)
5 YDD, mean residual (SD)	0.02 (0.21)	0.03 (0.11)
6 YDD, mean residual (SD)	0.13 (0.07)	0.00 (0.11)
7 YDD, mean residual (SD)	0.12 (0.26)	0.06 (0.07)
8 YDD, mean residual (SD)	0.04 (0.10)	0.08 (0.02)
9 YDD, mean residual (SD)	0.03 (0.25)	0.06 (0.02)

Correlation with CSF-biomarkers and cognition

We studied the correlation between individual residuals from the FTD disease model and individual CSF-NfL and CSF-14-3-3 levels and MMSE scores for the FTD variants. (Figure 1).

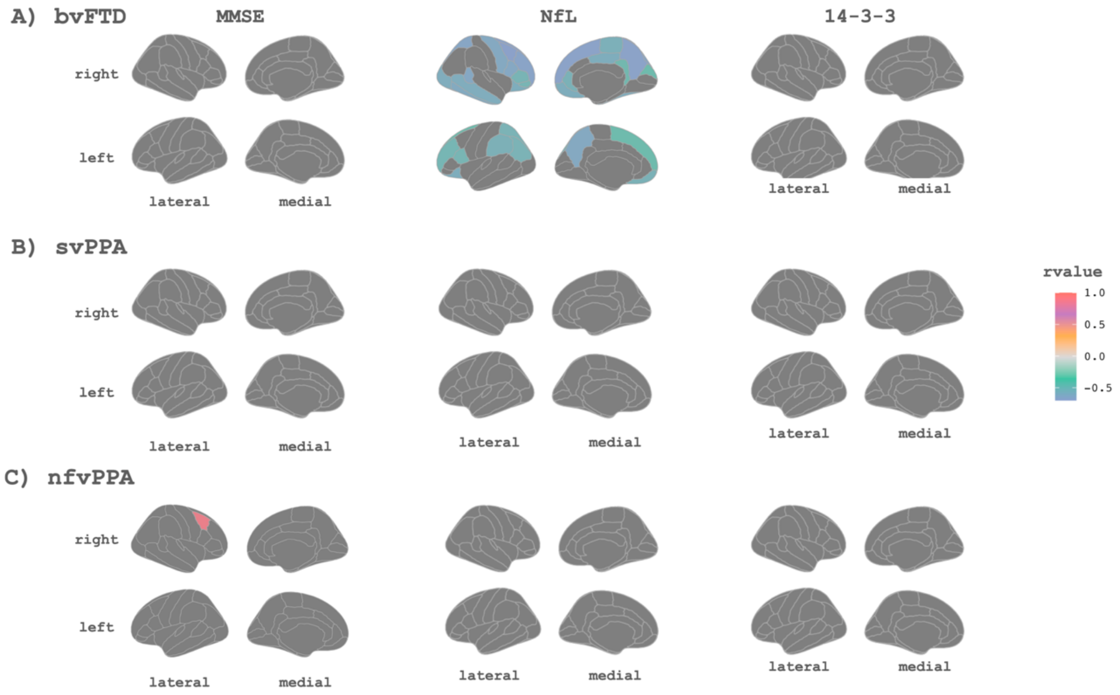

Figure 1: Maps of correlations between residuals for the cortical thickness model and cerebrospinal fluid-biomarkers and MMSE scores for the FTD variants. Only significant regions are shown. The threshold was set at 0.05 with a p-value adjusted with multiple comparisons of all the CTh regions. bvFTD: behavior frontotemporal dementia, svPPA: semantic variant primary progressive aphasia, nfvPPA: nonfluent variant primary progressive aphasia, MMSE: Mini-Mental State Examination, NfL: neurofilament light chain.

Prediction of Cortical Thickness for the follow-up visits

We estimated the mean absolute error (MAE) for all the visits (Table 3).

Table 3: The mean absolute error for each group. We estimated the mean for all the regions. HC: healthy controls, AD: Alzheimer's disease, FTD: frontotemporal dementia, bvFTD: behavior frontotemporal dementia, svPPA: semantic variant primary progressive aphasia, nfvPPA: nonfluent variant primary progressive aphasia.

	Baseline	Timepoint 2	Timepoint 3
HC	0.12	0.12	0.12
AD	0.15	0.15	------
FTD	0.20	0.22	0.25
bvFTD	0.20	0.20	0.23
svPPA	0.20	0.22	0.23
nfvPPA	0.20	0.26	------

TREBALL 3

Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data.

Agnès Pérez-Millan, José Contador, Jordi Juncà-Parella, Beatriz Bosch, Laia Borrell, Adrià Tort-Merino, Neus Falgàs, Sergi Borrego-Écija, Nuria Bargalló, Lorena Rami, Mircea Balasa, Albert Lladó, Raquel Sánchez-Valle, Roser Sala-Llonch

Human Brain Mapping 2023; 2023 Gener 20

DOI: 10.1002/hbm.26205

IF (2022): 4,8; Q1 Neuroimaging

Received: 22 July 2022 | Revised: 1 December 2022 | Accepted: 4 January 2023
DOI: 10.1002/hbm.26205

RESEARCH ARTICLE

WILEY

Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data

Agnès Pérez-Millan[1,2] | José Contador[1] | Jordi Juncà-Parella[1] | Beatriz Bosch[1] | Laia Borrell[2] | Adrià Tort-Merino[1] | Neus Falgàs[1,3] | Sergi Borrego-Écija[1] | Nuria Bargalló[4] | Lorena Rami[1] | Mircea Balasa[1,3] | Albert Lladó[1] | Raquel Sánchez-Valle[1] | Roser Sala-Llonch[2,5]

[1]Alzheimer's Disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació Clínic per a la Recerca Biomèdica, Universitat de Barcelona, Barcelona, Spain

[2]Department of Biomedicine, Faculty of Medicine, Institute of Neurosciences, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona, Spain

[3]Atlantic Fellow for Equity in Brain Health, Global Brain Heath Institute, University of California San Francisco, Trinity College Dublin, San Francisco, California, USA

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

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

Correspondence
Roser Sala-Llonch, PhD, Biophysics and Bioengineering Unit, Department of Biomedicine, Faculty of Medicine, University of Barcelona, Casanova 143, Barcelona 08036, Spain.
Email: roser.sala@ub.edu

Funding information
Department de Salut – Generalitat de Catalunya, Grant/Award Numbers: PERIS 2016-2020, SLT008/18/00061; Instituto de Salud Carlos III (ISCIII), Grant/Award Numbers: 20143810, PI20/0448, PI19/00198, PI19/00449; Spanish Ministry of Science and Innovation, Grant/Award Number: AEI/10.13039/501100011033; María de Maeztu Unit of Excellence (Institute of Neurosciences, University of Barcelona, Grant/Award Number: MDM-2017-0729; Joan Rodés Josep Baselga grant FBBVA; Rio Hortega, Grant/Award Number: CM21/00024

Abstract

Alzheimer's disease (AD) and frontotemporal dementia (FTD) are common causes of dementia with partly overlapping, symptoms and brain signatures. There is a need to establish an accurate diagnosis and to obtain markers for disease tracking. We combined unsupervised and supervised machine learning to discriminate between AD and FTD using brain magnetic resonance imaging (MRI). We included baseline 3T-T1 MRI data from 339 subjects: 99 healthy controls (CTR), 153 AD and 87 FTD patients; and 2-year follow-up data from 114 subjects. We obtained subcortical gray matter volumes and cortical thickness measures using FreeSurfer. We used dimensionality reduction to obtain a single feature that was later used in a support vector machine for classification. Discrimination patterns were obtained with the contribution of each region to the single feature. Our algorithm differentiated CTR versus AD and CTR versus FTD at the cross-sectional level with 83.3% and 82.1% of accuracy. These increased up to 90.0% and 88.0% with longitudinal data. When we studied the classification between AD versus FTD we obtained an accuracy of 63.3% at the cross-sectional level and 75.0% for longitudinal data. The AD versus FTD versus CTR classification has reached an accuracy of 60.7%, and 71.3% for cross-sectional and longitudinal data respectively. Disease discrimination brain maps are in concordance

Raquel Sánchez-Valle and Roser Sala-Llonch contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2023 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

with previous results obtained with classical approaches. By using a single feature, we were capable to classify CTR, AD, and FTD with good accuracy, considering the inherent overlap between diseases. Importantly, the algorithm can be used with cross-sectional and longitudinal data.

KEYWORDS
Alzheimer's disease, frontotemporal dementia, machine learning, magnetic resonance imaging, neuroimaging markers

1 | INTRODUCTION

Alzheimer's disease (AD) and frontotemporal dementia (FTD) are common forms of dementia, with different, but partly overlapping, symptoms, and brain signatures. This complicates the differential diagnosis and might lead to misdiagnosis (Balasa et al., 2011; Mendez, 2006). On the other hand, accurate characterization of longitudinal trajectories is needed for prognosis and disease tracking purposes. In recent years, clinical studies have been complemented with automated or semiautomated algorithms, which hold promise toward computer-aided diagnosis. In this line, machine learning (ML) techniques use clinical and biomarker data to learn patterns that can be used as an aid for differential disease diagnosis and tracking.

Magnetic resonance imaging (MRI) has been widely used to detect disease-specific brain changes across different neurodegenerative disorders. Concretely, using structural MRI, studies have described patterns of cortical thickness (CTh) and gray matter (GM) volume loss both in AD and FTD when compared separately with healthy populations (Bocchetta et al., 2021; Canu et al., 2017; Contador et al., 2021; Möller et al., 2013). In addition, in a previous study including both diseases, we showed that distinct brain atrophy patterns could potentially help in differentiating AD and FTD (Falgàs et al., 2020). More recently, measures derived from MRI have been used within ML algorithms to differentiate these diseases (Bron et al., 2017; Li et al., 2021; Möller et al., 2016; Penny et al., 2007). These approaches have shown good results. However, the large number of features needed can result in computationally expensive methods that are difficult to implement in a clinical setting. In addition, obtaining spatial patterns of the features driving classification can be challenging in some cases. These two issues limit the applicability and interpretability of the algorithms.

Besides the description of atrophy patterns at a specific time point, longitudinal neuroimaging studies have gained popularity in AD and FTD (Bejanin et al., 2020; Irish et al., 2018; Sintini et al., 2020). Using structural MRI, studies have described the trajectories of GM volume and CTh loss with time, providing valuable information on the characterization of disease trajectories and validation of prognostic biomarkers (Risacher et al., 2010; Storsve et al., 2014). When compared with cross-sectional studies, longitudinal designs can explore the large heterogeneity of the effect of between-subject brain changes which requires repeated measures and longitudinal designs (Bernal-Rusiel et al., 2013; Thompson et al., 2011). Thus, while cross-sectional studies depict patterns of differential or overlapping brain atrophy, longitudinal studies are needed to understand the differences in disease trajectories. In the context of ML, although longitudinal data are sometimes used for diagnostic confirmation and/or for exploring the predictive values of baseline acquisitions, few approaches include longitudinal data in the training to improve the models.

In this study, we developed a classification algorithm using cross-sectional and longitudinal MRI data from 399 subjects, including AD and FTD patients and healthy controls (CTR). We implemented a feature reduction algorithm using unsupervised techniques followed by a widely used classifier, namely the support vector machine (SVM). Besides maximizing classification performance, by studying the weights of the features from the unsupervised part, we aimed to investigate the brain patterns that collected a higher amount of variance for the different classification settings, thus providing interpretability to our results.

2 | MATERIALS AND METHODS

2.1 | Participants

We selected 339 subjects prospectively recruited from the Alzheimer's disease and other cognitive disorders unit of the Hospital Clínic de Barcelona (HCB), all having a complete clinical work-up and at least one 3T high-resolution structural MRI scan. Additionally, a subset of subjects underwent a second acquisition after ~2 years. Participants were classified into three groups:

- AD: patients who presented AD biomarker profiles suggesting underlying AD neuropathology (*abnormal amyloid and tau*, A + T+) with neurodegeneration (N+) according to National Institute on Aging/Alzheimer Association Research Framework 2018 (Jack et al., 2018) and Mini-Mental State Examination (MMSE) ≥ 18. They also fulfilled diagnostic criteria for mild cognitive impairment due to AD or AD mild dementia (Albert et al., 2011; McKhann et al., 2011).
- FTD: patients who met diagnostic criteria for either behavioral variant frontotemporal dementia (bvFTD) or primary progressive aphasia (PPA), including semantic variant PPA (svPPA), and nonfluent variant PPA (nfvPPA) (Gorno-Tempini et al., 2011; Rascovsky

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2011). FTD showed normal values for AD cerebrospinal fluid (CSF) biomarkers.

- CTR: healthy adults having cognitive performance within the normative range (cutoff 1.5 SD from the normative mean) and normal levels of AD CSF biomarkers.

Subjects were collected as part of other ongoing and past studies within the HCB. All were approved by the HCB Ethics Committee and all participants gave written informed consent.

2.2 | MRI acquisition

A high-resolution 3D structural data set (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) was acquired for each individual at each time point in a 3T Magnetom Trio Tim scanner (Siemens Medical Systems, Germany) or in a 3T Prisma scanner (Siemens Medical Systems, Germany) at HCB using equivalent acquisition protocols. Before these analyses, we performed tests to evaluate the interscanner variability (results can be found in Supplementary Material S1), and we found low variability induced by the scanner.

2.3 | 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. For longitudinal data, we used the longitudinal stream in FreeSurfer. All FreeSurfer preprocessing steps are reported in detail elsewhere (Fischl & Dale, 2000; Fischl et al., 2004; Reuter et al., 2012). Briefly, FreeSurfer allowed us to generate automated CTh maps and segmentation of the subcortical structures. From reconstructed data, we obtained global measures of mean CTh and GM volumes of the left and right hemispheres (lh and rh). In addition, we used the summary measures of mean CTh in 68 cortical parcellations and GM volumes of 17 subcortical structures, all derived from atlases available in FreeSurfer (Desikan et al., 2006; Seidman et al., 1997). Volume measures were normalized by the estimated intracranial volume. All images and individual segmentations were visually inspected and manually corrected if needed.

2.4 | Cross-sectional study: Brain signatures and classification performance

We used all the CTh values and subcortical volumes at baseline obtained with FreeSurfer to create our ML pipeline. We introduced the global and regional measures of both hemispheres (rh and lh) separately leading to a total of 103 values per subject. Subcortical regions were normalized using the intracranial volume and then all values (subcortical volumes and CTh measures) were converted to z-score. Therefore, the features of our ML algorithm were the subcortical volumes and the CTh measures transformed to z-scores obtained after processing the T1-weighed MRI images.

Firstly, we performed a principal component analysis (PCA) to reduce the number of regional measures (all CTh measures and all volumetric values) to a single feature, by keeping only the first principal component (PC). The first PC is the one with the highest value of explained variance, so it is the best choice to study the weights used in the transformation while keeping good classification accuracy. The weights to obtain this first PC will be used to provide a regional interpretation of the classification results. Second, this feature was introduced into a SVM algorithm to perform classifications between the three groups and between pairs of groups. Eventually, the weights of the PCA were used to create disease-specific patterns. These patterns, allow having an interpretable ML algorithm that returns the accuracy and the associated patterns of the classification between groups. So, the proposed pipeline is not only focused on obtaining the best accuracy, but it also focuses on giving explainability to the ML algorithm.

To address circularity in the PCA and to avoid overfitting, we implemented all the steps in a cross-validated setting, in which train data was used in the entire PCA + SVM, and a group of subjects was hold-out to be used as test data. For testing, original data were projected into the PCA space to obtain the unique feature and this was used in the SVM classification. Overall performance was assessed using k-fold cross-validation with 20 iterations of the procedure explained above. Figure 1 shows a schematic representation of the algorithm. Additionally, the SVM hyperparameters: kernel (options: linear, rbf, and polynomial), *C* (values = [0.1,1, 10, 100, and 1000]), and gamma (values = [1, 0.1, 0.01, and 0.001]) were introduced with a Grid Search using an additional cross-validation of 10, with the train set. The algorithm was implemented in Python version 3.8 (www.python.org), and we used the library scikit-learn (Pedregosa et al., 2011) for PCA, Grid Search, and SVM.

To assess the effect of age in the classification, we performed two complementary analyses: (1) we evaluated if age itself was able to classify the different groups; (2) We repeated all analyses by adding age as a feature together with all the regions. To assess the effect of sex in the classification we repeated all analyses by adding sex as a feature together with all the regions.

2.5 | Longitudinal study: Brain signatures and classification

We used all global and regional volumes and CTh measures derived from the FreeSurfer longitudinal stream for the longitudinal classification analysis. As in the cross-sectional analysis, we introduced the measures of both hemispheres (rh and lh) separately, subcortical regions were divided by intracranial volume, and all values were converted to z-score.

Here, the overall pipeline also consisted of an unsupervised algorithm followed by an SVM classification. For the unsupervised part,

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

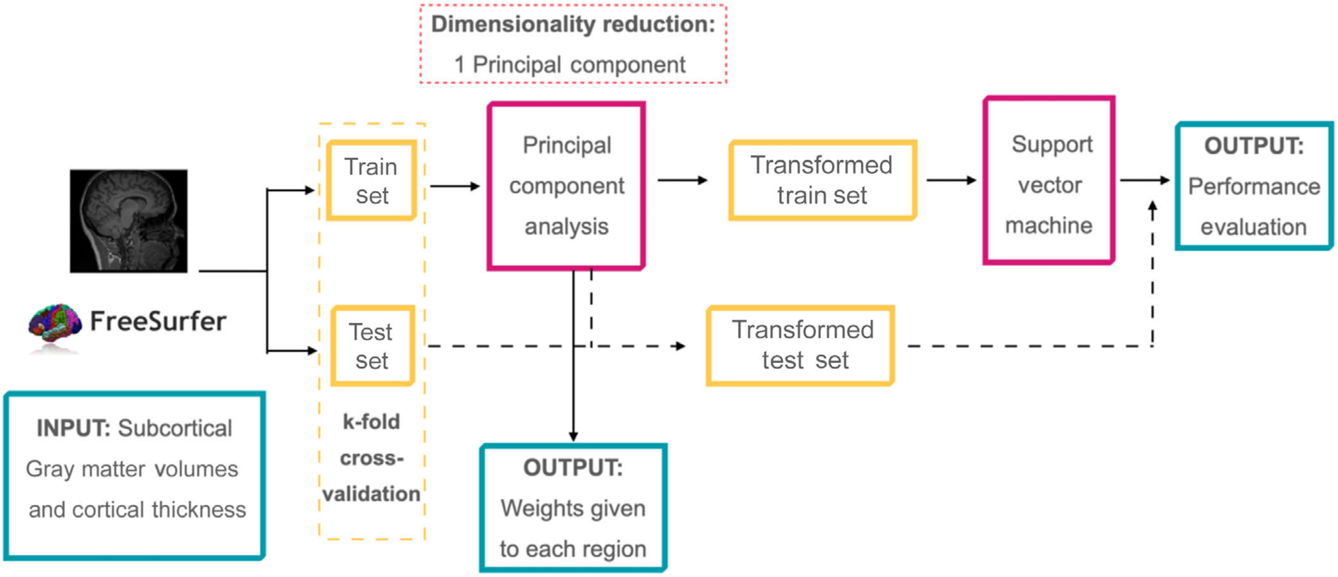
Dimensionality reduction:
1 Principal component
Train set
Principal component analysis
Transformed train set
Support vector machine
OUTPUT: Performance evaluation
FreeSurfer
Test set
Transformed test set
INPUT: Subcortical Gray matter volumes and cortical thickness
k-fold cross-validation
OUTPUT: Weights given to each region

FIGURE 1 Machine Learning Algorithm scheme

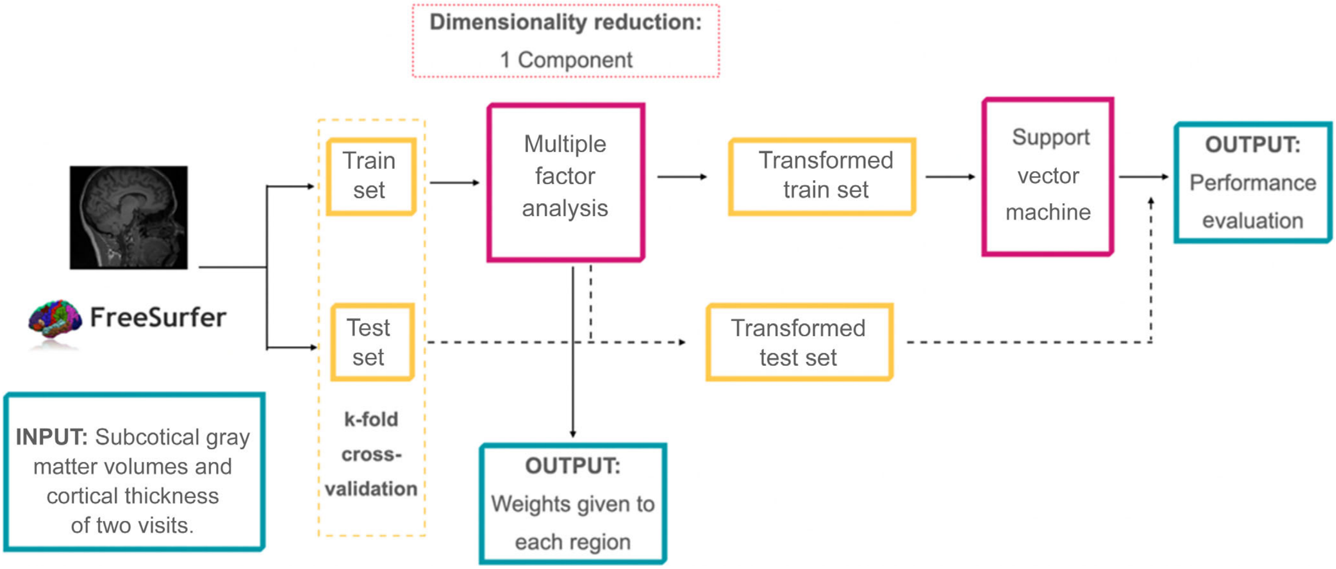
Dimensionality reduction:
1 Component
Train set
Multiple factor analysis
Transformed train set
Support vector machine
OUTPUT: Performance evaluation
FreeSurfer
Test set
Transformed test set
INPUT: Subcotical gray matter volumes and cortical thickness of two visits.
k-fold cross-validation
OUTPUT: Weights given to each region

FIGURE 2 Longitudinal Machine Learning Algorithm scheme

we performed a multiple factor analysis (MFA), which is a generalization of PCA. MFA allows implementing a factor analysis with repeated samples (i.e., the same measures for different time points for each subject; Abdi et al., 2013). As before, we only kept the first component. We kept the first component with the same objective as we did with the PCA in the cross-sectional analyses. The weights to obtain this first component will give us the interpretability of the classification results of the algorithm. Then, this first component was introduced to the SVM algorithm. Here, the use of MFA facilitates the creation of a longitudinal ML in the SVM context, while keeping the interpretability of results. As before, overall performance was assessed using k-fold cross-validation with $k = 20$, in which test data was not included in any of the procedures within the pipeline (MFA and SVM). As an internal quality control, we revised the number of subjects included in the train/test set for each iteration, and we observed that they were equally distributed across iterations. Figure 2 shows a scheme of the algorithm.

With the objective to study how the longitudinal data can help to improve the classification in ML algorithms and how should be introduced to the algorithm, different settings were considered for the longitudinal analysis: (1) including all regional measures from time point 1 and time point 2 as different observations for each subject in the MFA decomposition. (2) including all measures from each time point together with a *change* variable (computed as time point2–time point1) for each region in the MFA decomposition. (3) using only the change variable in a standard PCA analysis. So, this setting will be the equivalent of a study with longitudinal data but analyzed at a cross-sectional level. In addition, to allow comparison with the cross-

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sectional analysis we performed a PCA decomposition of the data from time point 1 with the subsample having a longitudinal acquisition.

Finally, for each of the four decomposition settings mentioned above, we used the weights of the obtained component from the MFA/PCA to create disease-specific patterns. The SVM optimal parameters were obtained with a Grid Search with cross-validation of 10, with the train set.

The algorithm was implemented in Python version 3.8 we used the libraries prince (https://pypi.org/project/prince/) for MFA and scikit-learn (Pedregosa et al., 2011) for PCA, Grid Search, and SVM.

3 | RESULTS

3.1 | Sample demographics

Of the 339 subjects included in the analyses, 153 were AD patients ($n = 20$ with longitudinal data); 87 were FTD ($n = 26$ with longitudinal data), and 99 were CTR ($n = 68$ longitudinal). The FTD group included subjects with 49 bvFTD patients ($n = 14$ longitudinal data), 20 svPPA patients ($n = 7$ longitudinal data), 17 nfvPPA patients ($n = 4$ longitudinal data), and 1 unspecified PPA patient ($n = 1$ longitudinal data). Demographic information and group statistics are shown in Table 1. In summary, there was a significantly greater proportion of men in the FTD group compared with the AD and CTR groups. In addition, CTR subjects were slightly younger than AD and FTD groups ($p < 0.05$) at the first visit. However, there were no differences in age for the second visit. The time between scanners was lower in the FTD group compared with AD and CTR ($p < .0001$) but did not differ between AD and CTR.

3.2 | Cross-sectional analyses: Classification

Our algorithm had an accuracy of 83.3% ± 12.7% in the CTR versus AD classification, 82.1% ± 14.7% for CTR versus FTD, 63.3% ± 9.1% for AD versus FTD, and 60.7% ± 12.7% when discriminating the three groups. Including age or sex in the algorithm led to similar results (see Supplementary Material S1)

3.3 | Cross-sectional analyses: Brain patterns and relevant features

The weights of each of the regional measures (CTh cortical regions and GM volumes for subcortical regions) were obtained as their contribution to the main feature (i.e., the first component). Figure 3 shows these PCA weights for the separate analysis of CTR versus AD and CTR versus FTD. As all algorithms were cross-validated using k-fold, these patterns show the mean weights obtained with all the train sets across iterations.

As can be seen in Figure 3, for AD, the three most important regions were rh supramarginal, lh supramarginal, and rh precuneus. For the FTD group, these were rh supramarginal, rh superior frontal, and rh inferior parietal. The complete list of regions with their associated weights for each setting is shown in Supplementary Material S1.

3.4 | Longitudinal analyses: Classification

First, when using the two time points separately for each subject in the MFA, we obtained an accuracy of 90.0% ± 14.7% in the CTR versus AD classification, 88.0% ± 16.4% for CTR versus FTD, 75.0% ± 36.9% for AD versus FTD, and 71.3% ± 13.1% for the three-group classification (Table 2)

Including change as an additional measure for each subject and region led to higher classification scores for CTR versus AD and the three-group classification. The accuracies were 94.5% ± 11.8% in the CTR versus AD classification, 87.8% ± 17.8% for CTR versus FTD, 60.8% ± 33.4% for AD versus FTD, and 77.7% ± 19.0% when discriminating the three groups. Interestingly, including change in the algorithm increased the accuracy for CTR versus AD and decreased the accuracy for CTR versus FTD. Including age or sex in all the longitudinal analyses led to similar results (see Supplementary Material S1)

TABLE 1 Group summaries are given as the mean and the SD of each measure.

				CTR–AD	CTR–FTD	AD–FTD
	CTR	AD	FTD	*p*-value	*p*-value	*p*-value
N at first MRI	99	153	87	—	—	—
N at second MRI	68	20	26	—	—	—
Sex at first MRI, men/women	30/69	59/94	47/40	.22	**.005**	**.03**
Sex at second MRI, men/women	18/50	10/10	15/11	.09	**.02**	.77
Age at first MRI, years (SD)	60.2 (10.5)	64.3 (9.7)	63.6 (8.3)	**.003**	**.03**	.5
Age at second MRI, years (SD)	65.0 (7.2)	62.1 (4.5)	63.8 (5.9)	.3	.5	.5
Time between MRIs, years (SD)	2.1 (0.4)	1.9 (0.3)	1.4 (0.5)	.3	**2.5 e-9**	**4.7 e-5**

Note: Differences between groups are calculated using Fisher test for sex and the ANOVA test for the rest of the variables. Significant group differences are highlighted in bold, and pairwise differences were measured with Benjamini–Hochberg correction (*p*-values threshold .05).
Abbreviations: AD, Alzheimer's disease; CTR, controls; FTD, frontotemporal dementia; MRI, magnetic resonance imaging.

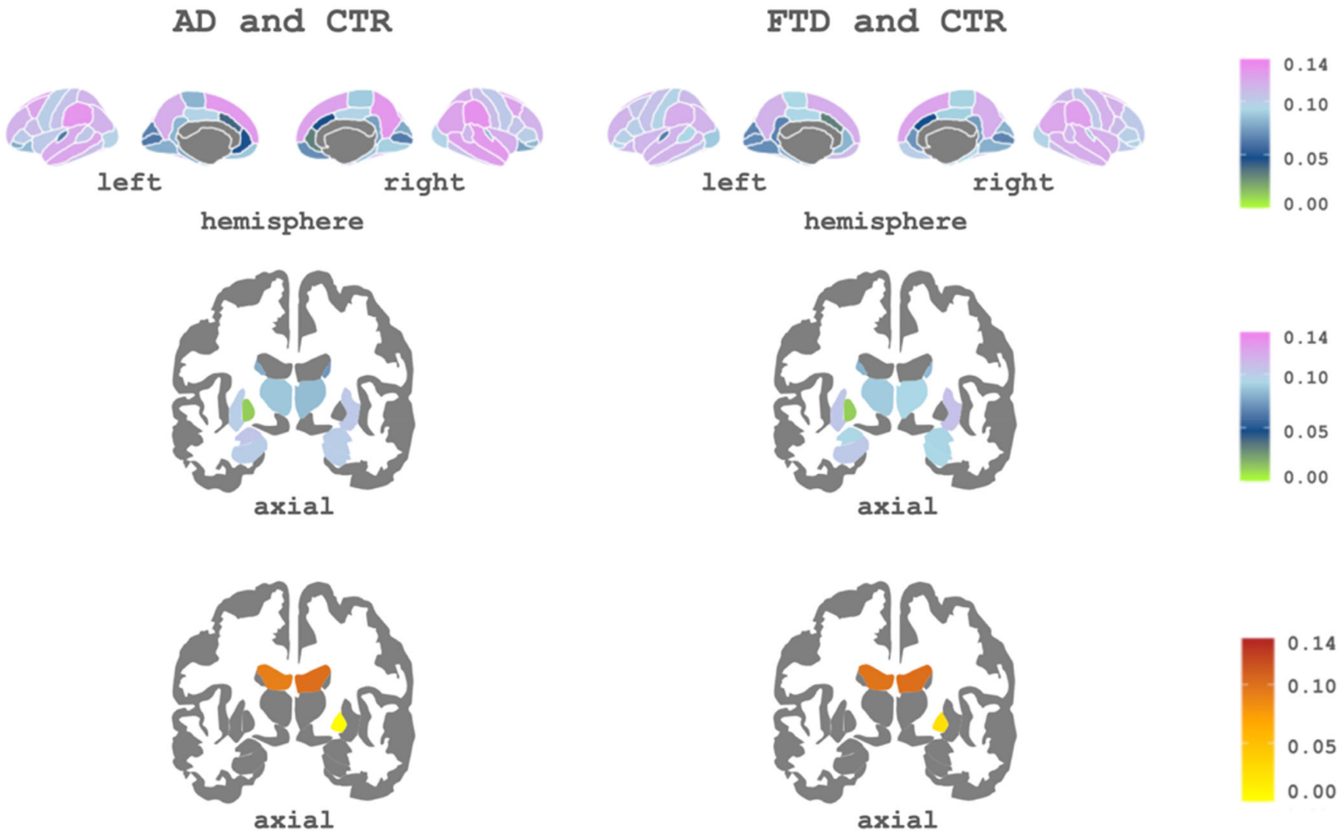

FIGURE 3 Subcortical and cortical patterns of the first principal component's weights associated with Alzheimer's disease (AD) and frontotemporal dementia (FTD). Top: Cortical regions of interest (ROIs) included in the component. Bottom: subcortical ROIs of the component. The cool color scale represents negative weights, and the warm scale represents positive weights within the component

TABLE 2 Classification performance of the different approaches.

		Longitudinal reduced sample			
	Cross-sectional all subjects (%)	Tp1, Tp2 (%)	Tp1, Tp2, change (%)	Only change (%)	Baseline of the longitudinal sample
CTR vs. AD	83.3 ± 12.7	90.0 ± 14.7	**94.5 ± 11.8**	86.5 ± 13.2	79.5 ± 9.8
CTR vs. FTD	82.1 ± 14.7	**88.0 ± 16.4**	87.8 ± 17.8	80.0 ± 15.4	81.0 ± 16.0
AD vs. FTD	63.3 ± 9.1	**75.0 ± 36.9**	60.8 ± 33.4	58.3 ± 32.2	67.5 ± 32.2
CTR vs. AD vs. FTD	60.7 ± 12.7	71.3 ± 13.1	**77.7 ± 19.0**	46.7 ± 21.2	59.7 ± 10.4

Note: The best accuracy for each group is highlighted in bold.
Abbreviations: AD, Alzheimer's disease; CTR, controls; FTD, frontotemporal dementia, Tp1, timepoint 1; Tp2, timepoint 2.

When including only change as a feature for each region (i.e., without the individual time point data), we obtained lower classification scores compared with the full-longitudinal setting. The classification scores were: 86.5% ± 13.2% in the CTR versus AD classification, 80.0% ± 15.4% for CTR versus FTD, 58.3% ± 32.2% for AD versus FTD, and 46.7% ± 21.2% for the three-class classification. This result indicates that this variable is not the unique source of discrimination power across groups. However, we observed that including it in the longitudinal study together with the measures at each time point was beneficial for classification

Finally, to be able to directly compare the two approaches, we repeated the cross-sectional analyses for the baseline data of the subjects that had longitudinal data available. We obtained a cross-sectional accuracy of 79.5% ± 9.8% in the CTR versus AD classification, 81.0% ± 16.0% for CTR versus FTD, 67.5% ± 32.2% for AD versus FTD, and 59.7 ± 10.4% for the three-class classification. These results highlight the increase in performance for the longitudinal approach when compared with a cross-sectional approach with an equivalent sample. Table 2 shows a summary of all the accuracies for each analysis and comparison

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

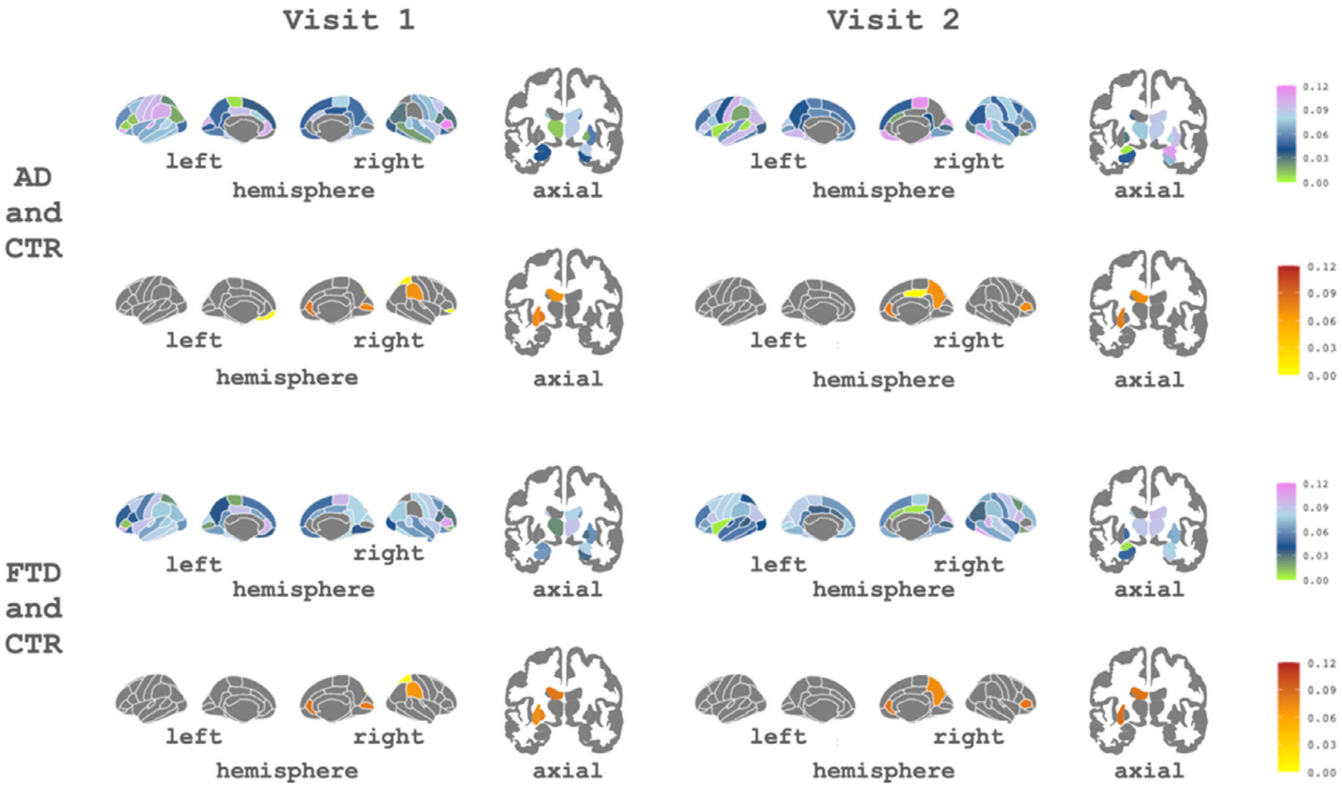

FIGURE 4 Subcortical and cortical patterns of the contributions of the first component associated with Alzheimer's disease (AD) and frontotemporal dementia (FTD) for both visits. The cool color scale represents negative weights, and the warm scale represents positive weights within the component

3.5 | Longitudinal analyses: Brain patterns and relevant structures

The contributions of each brain region (CTh cortical regions and GM volumes for subcortical regions) to the first MFA component are shown in Figure 4. Using these weights, we created a pattern for each comparison and visit: CTR versus AD and CTR versus FTD. As before, patterns were obtained with the mean weights obtained with all train sets across iterations.

As can be seen in Figure 4, for the longitudinal analysis the most important regions for differentiating AD were rh pars triangularis, rh entorhinal, and lh transverse temporal from Visit 1, and rh paracentral, rh inferior temporal, and rh amygdala from Visit 2. For FTD, these were rh pars triangularis, lh pars orbitalis, and lh rostral anterior cingulate from Visit 1 and rh temporal pole, rh inferior temporal, and rh bankssts from Visit 2. The complete list of features and associated weights for each comparison can be found in Supplementary Material S1.

4 | DISCUSSION

In this study, we show the utility of machine learning approaches for the differential diagnosis of AD and FTD. We used a combination of unsupervised and supervised techniques that allowed (1) reducing all subcortical GM volumes and CTh measures from MRI into a single feature and (2) obtaining blind classifications from MRI data while providing meaningful atrophy patterns at the group level using the weights of the unsupervised part of the algorithm. We implemented two equivalent algorithms for cross-sectional and longitudinal data, and we observed that longitudinal approaches outperformed cross-sectional analyses in differentiating AD and FTD.

At the cross-sectional level, we demonstrated that ML can be used to support clinical diagnosis using a method that is not computationally expensive. Previous ML neuroimaging studies with cross-sectional structural MRI data have shown accuracies between 80% and 95% for AD versus CTR, 72%–88% for FTD versus CTR, and 69%–89% for AD versus FTD (Bron et al., 2017; Klöppel et al., 2008; Möller et al., 2016). In the case of the three-class classification, Bron et al. (2017) reported an accuracy of 70%. In our cohort, we obtained accuracy scores that are in agreement with these studies. Our study has a higher sample size than those mentioned above and differs in two important methodological aspects: first, we used a feature reduction algorithm that results in a single component being fed into the classification stage. This component encompasses information from all CTh values and GM volume measures. Even if there is an obvious risk of losing information in this step, we demonstrated that it is still useful for discrimination, while reducing the computational cost. Second, our FTD group included patients with both PPA and bvFTD variants, not only bvFTD patients as in some of the abovementioned

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

studies. Even if further studies with similar algorithms should attempt to differentiate disease subtypes, we believe that our approach can be very useful to support differential diagnosis across diseases. Noteworthy, the wide range of misclassified subjects both in ours in previous studies suggests that some heterogeneities and overlaps need to be further explored (Habes et al., 2020).

We developed an algorithm that used longitudinal data for training the classification models, which resulted in a noticeable increase in classification accuracy for all comparisons, increasing up to 15% of the overall accuracy score. Only a few studies have used longitudinal data in ML algorithms. For example, previous studies reported accuracies between 80% and 94% for AD versus CTR (Gavidia-Bovadilla et al., 2017; Guo et al., 2020; Zhang et al., 2017). In our cohort, we obtain accuracy scores of 90% and 95% for this specific comparison, outperforming previous approaches. To our knowledge, no previous studies have used longitudinal data to differentiate FTD versus CTR or AD versus FTD. Our approach uses the MFA decomposition, which captures the variance of the data while modeling intra-subject variability.

The accuracy obtained in all the longitudinal experiments suggests that the proposed methodology is useful to support the diagnosis problem of distinguishing AD from FTD. Longitudinal analyses allowed studying disease change and disease trajectories. Here, we first applied the ML algorithm combining all longitudinal data in a repeated measures fashion (Visit 1 plus Visit 2). We then tested the algorithm using the longitudinal data (Visit 1 and Visit 2) in combination with the variable change that was the difference between both visits. We observed that including the change variable in the algorithm boosted the accuracy for CTR versus AD and worsened the accuracy for CTR versus FTD. This could indicate that in AD patients the change variable has higher discriminative power. On the contrary, in FTD the change variable did not add discriminative power (indeed, there is a slight decrease in accuracy), indicating that the patterns themselves provide more discriminability than the change. This finding could be associated with the fact that FTD patients could show floor effect in atrophy rates due to an advanced stage of disease, while AD patients included in this study might be in a less advanced stage (Pankov et al., 2016). Finally, as regards the abovementioned results, we acknowledge that the change variable, computed as a difference, might add collinearity to the algorithm. However, by including it, we were able to make guesses about the different spatial patterns being important for classification.

Besides maximizing performance, we were interested in depicting the spatial patterns that drive accuracy for each classification setting. In ML, this represents a crucial step for developing algorithms that are interpretable at the biological and pathological levels (Stiglic et al., 2020). Here, we used the weights of the PCA/MFA components to identify the regions that contributed the most to the group variances. To add explainability to our algorithm, we studied the most important regions providing classification between pathologies. In general, we found widespread brain patterns of variance, with common AD/FTD regions appearing within the ones on the top of the lists, depicting pathological patterns in concordance with the literature (Möller et al., 2016; Rabinovici et al., 2008). However, our classification feature was obtained from the PCA, and therefore it includes contributions across all brain regions, accounting for both overlapping and differential patterns across disorders. ML techniques search for robust interactions between features (in our case the brain regions), so it is plausible that AD and FTD patterns present both differential and overlapping regions. Previous works (Davatzikos et al., 2008; Falgàs et al., 2020; Laakso et al., 2000) have also found some overlapping patterns of degeneration across disorders. Overall, we believe that the joint analysis of overlapping patterns (i.e., indicating common neurodegeneration) and specific regional alterations will be crucial in future works investigating differential diagnosis.

We developed an ML algorithm for differential diagnosis of FTD and AD with good to excellent accuracy. Our algorithm combines unsupervised and supervised methodologies. In addition, we adapted our algorithms to longitudinal data, which was not included in previous works in the field (Bron et al., 2017; Chagué et al., 2021; Klöppel et al., 2008; Möller et al., 2016). Therefore, the main novelty is the combination of advanced methods (PCA/MFA and SVM) together with the use of longitudinal data, a key point to understanding neurodegenerative diseases (Clifford et al., 1999). Moreover, the PCA was not only used for the feature reduction step, as in previous studies (Bachli et al., 2020; Davatzikos et al., 2008; Kim et al., 2019), but also for obtaining the weights of the features for interpreting the findings.

Our study has some limitations. Regarding the sample size: First, it is important to consider that our sample size at baseline, is lower compared with the sample size which can be obtained in a multicentric study. However, all the data has been acquired in the same center, allowing us to have the same MRI scanner protocol and the same clinical criteria for the diagnosis. Another limitation is the relatively small sample size at the follow-up visit, in some analyses, especially for the comparison between AD and FTD, the number of subjects was low, due to the difficulty of obtaining longitudinal samples in these dementias. Regarding the MRI data, it is known that the combination of different modalities of MRI, such as diffusion tensor imaging (DTI), resting-state, amyloid positron emission tomography, or arterial spin labeling could improve the accuracy scores (Agostinho et al., 2022; Bron et al., 2014, 2017), while in our case we only used structural MRI. Thus, as future work, we could include some of these image modalities in our algorithm. Finally, the different clinical variants in the FTD group (bvFTD and PPA) were studied as a single group and there is the possibility that our results are biased by the different FTD atrophy patterns. Due to sample size limitations, it was impossible to subdivide the FTD group to study in detail the different variants of the pathology. Future studies taking into consideration this point should be considered.

In our study, we focused on differentiating AD and FTD using ML on MRI data. Other studies have reported differences between these pathologies using a wide range of methods and MRI sequences. For example, the work of Du et al. (2007) used structural MRI to compare regional CTh between AD and FTD and their relationship with neuropsychological scores at a vertex-wise level. They found that FTD patients, compared with AD patients, had a thinner cortex in parts of

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the bilateral parietal and precuneus regions. Similarly, Steketee et al. (2016) used perfusion and structural MRI to differentiate AD and FTD. They studied the sensitivity, specificity, and diagnostic performance of the different regions and found that AD patients, compared with FTD patients, showed hypoperfusion in the posterior cingulate cortex. However, they also found that regional atrophy did not differ between AD and FTD. In this sense, we could differentiate AD and FTD with atrophy values, suggesting that ML techniques, that incorporate information across the whole brain, might be more sensitive. Finally, Avants et al. (2010) used sparse canonical correlation analysis with DTI and T1-weighed MRI data to identify patterns of reduced white matter (WM) integrity for AD and FTD. They found that, in FTD, frontal and temporal degeneration is correlated across modalities. In AD, they reported a significant association between CTh and WM in parietal and temporal regions. Considering all these studies, in future work, we aim to add different MRI modalities and other clinical variables to our algorithm.

5 | CONCLUSION

In conclusion, our study leads to three important points: First, the combination of PCA or MFA and SVM successfully separates patients with AD or FTD from CTR subjects. Second, they perform well in the differential diagnosis of two pathologies (AD vs FTD). Thirdly, the follow-up visits are beneficial in ML algorithms to distinguish these pathologies, especially AD. All these points suggest an important breakthrough in computer-aided diagnostic image analysis for clinical research and practice.

AUTHOR CONTRIBUTIONS

Agnès Pérez-Millan, Laia Borrell, Raquel Sánchez-Valle, and Roser Sala-Llonch contributed to the design of the study. Agnès Pérez-Millan, José Contador, Jordi Juncà-Parella, and Beatriz Bosch contributed to the analyses of the data. Agnès Pérez-Millan, José Contador, Albert Lladó, Raquel Sánchez-Valle, and Roser Sala-Llonch contributed to the interpretation of the data. Agnès Pérez-Millan, José Contador, Raquel Sánchez-Valle, and Roser Sala-Llonch contributed to the draft of the article. Jordi Juncà-Parella, Beatriz Bosch, Laia Borrell, Adrià Tort-Merino, Neus Falgàs, Sergi Borrego-Écija, Nuria Bargalló, Lorena Rami, Mircea Balasa, and Albert Lladó revised the article critically for important intellectual content and approved the final version of the article. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank patients, their relatives, and healthy controls for their participation in the research. This work was supported by Instituto de Salud Carlos III, Spain (grant no. 20143810 and PI20/0448 to Dr. Raquel Sánchez-Valle), Spanish Ministry of Science and by Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union (project PI19/00449 to Dr. Albert Lladó and project PI19/00198 to Dr. Mircea Balasa), by Department de Salut – Generalitat de Catalunya (PERIS 2016–2020 SLT008/18/00061 to Dr. A. Lladó) and by Spanish Ministry of Science and Innovation (AEI/10.13039/501100011033 to Dr. R. Sala-Llonch) and María de Maeztu Unit of Excellence (Institute of Neurosciences, University of Barcelona) MDM-2017-0729. Dr. S. Borrego-Écija is a recipient of the Joan Rodés Josep Baselga grant from FBBVA. Dr. Neus Falgàs recipient of Rio Hortega grant (CM21/00024).

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Agnès Pérez-Millan https://orcid.org/0000-0002-3006-9792
José Contador https://orcid.org/0000-0002-9692-1570
Jordi Juncà-Parella https://orcid.org/0000-0002-4772-2647
Beatriz Bosch https://orcid.org/0000-0002-6094-0024
Laia Borrell https://orcid.org/0000-0002-2519-2518
Adrià Tort-Merino https://orcid.org/0000-0002-5646-0482
Neus Falgàs https://orcid.org/0000-0002-3404-2765
Sergi Borrego-Écija https://orcid.org/0000-0003-2557-0010
Nuria Bargalló https://orcid.org/0000-0001-6284-5402
Lorena Rami https://orcid.org/0000-0002-7411-1921
Mircea Balasa https://orcid.org/0000-0002-1795-4228
Albert Lladó https://orcid.org/0000-0002-5066-4150
Roser Sala-Llonch https://orcid.org/0000-0003-3576-0475

REFERENCES

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*, 149–179. https://doi.org/10.1002/wics.1246

Agostinho, D., Caramelo, F., Moreira, A. P., Santana, I., Abrunhosa, A., & Castelo-Branco, M. (2022). Combined structural MR and diffusion tensor imaging classify the presence of Alzheimer's disease with the same performance as MR combined with amyloid positron emission tomography: A data integration approach. *Frontiers in Neuroscience*, *15*, 638175. https://doi.org/10.3389/fnins.2021.638175

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 Dement*, *7*, 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

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

Bachli, M. B., Sedeño, L., Ochab, J. K., Piguet, O., Kumfor, F., Reyes, P., Torralva, T., Roca, M., Cardona, J. F., Campo, C. G., Herrera, E., Slachevsky, A., Matallana, D., Manes, F., García, A. M., Ibáñez, A., & Chialvo, D. R. (2020). Evaluating the reliability of neurocognitive biomarkers of neurodegenerative diseases across countries: A machine learning approach. *NeuroImage*, *208*, 116456. https://doi.org/10.1016/j.neuroimage.2019.116456

Balasa, M., Gelpi, E., Antonell, A., Rey, M. J., Sánchez-Valle, R., Molinuevo, J. L., & Lladó, A. (2011). Clinical features and APOE genotype of pathologically proven early-onset Alzheimer disease. *Neurology*, *76*, 1720–1725. https://doi.org/10.1212/WNL.0b013e31821a44dd

Bejanin, A., Tammewar, G., Marx, G., Cobigo, Y., Iaccarino, L., Kornak, J., Staffaroni, A. M., Dickerson, B. C., Boeve, B. F., Knopman, D. S., Gorno-Tempini, M., Miller, B. L., Jagust, W. J., Boxer, A. L., Rosen, H. J., & Rabinovici, G. D. (2020). Longitudinal structural and metabolic changes in frontotemporal dementia. *Neurology*, *95*, E140–E154. https://doi.org/10.1212/WNL.0000000000009760

Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B., & Sabuncu, M. R. (2013). Statistical analysis of longitudinal neuroimage data with linear mixed effects models. *NeuroImage*, *66*, 249–260. https://doi.org/10.1016/j.neuroimage.2012.10.065

Bocchetta, M., Todd, E. G., Peakman, G., Cash, D. M., Convery, R. S., Russell, L. L., Thomas, D. L., Eugenio Iglesias, J., 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., … Zulaica, M. (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

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

Bron, E. E., Steketee, R. M. E., Houston, G. C., Oliver, R. A., Achterberg, H. C., Loog, M., van Swieten, J. C., Hammers, A., Niessen, W. J., Smits, M., Klein, S., & for the Alzheimer's Disease Neuroimaging Initiative. (2014). Diagnostic classification of arterial spin labeling and structural MRI in presenile early stage dementia. *Human Brain Mapping*, *35*, 4916–4931. https://doi.org/10.1002/hbm.22522

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

Clifford, J. R., Petersen, R. C., Xu, Y. C., O'Brien, P. C., Smith, G. E., Ivnik, R. J., et al. (1999). Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology*, *52*, 1397–1403. https://doi.org/10.1212/wnl.52.7.1397

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-Campoy, O., Antonell, A., Bargalló, N., Sanchez-Valle, R., Sala-Llonch, R., Lladó, A., & Alzheimer's Disease Neuroimaging Initiative. (2021). Longitudinal brain atrophy and CSF biomarkers in early-onset Alzheimer's disease. *NeuroImage Clinical*, *32*, 102804. https://doi.org/10.1016/j.nicl.2021.102804

Davatzikos, C., Resnick, S. M., Wu, X., Parmpi, P., & Clark, C. M. (2008). Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. *NeuroImage*, *41*, 1220–1227. https://doi.org/10.1016/j.neuroimage.2008.03.050

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*, 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Du, A. T., Schuff, N., Kramer, J. H., Rosen, H. J., Gorno-Tempini, M. L., Rankin, K., Miller, B. L., & Weiner, M. W. (2007). Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. *Brain*, *130*, 1159–1166. https://doi.org/10.1093/brain/awm016

Falgàs, N., Ruiz-Peris, M., Pérez-Millan, A., Sala-Llonch, R., Antonell, A., Balasa, M., Borrego-Écija, S., Ramos-Campoy, O., Augé, J. M., Castellví, M., Tort-Merino, A., Olives, J., Fernández-Villullas, G., Blennow, K., Zetterberg, H., Bargalló, N., Lladó, A., & Sánchez-Valle, R. (2020). Contribution of CSF biomarkers to early-onset Alzheimer's disease and frontotemporal dementia neuroimaging signatures. *Human Brain Mapping*, *41*, 2004–2013. https://doi.org/10.1002/hbm.24925

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11050–11055. https://doi.org/10.1073/pnas.200033797

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., et al. (2004). Automatically Parcellating the human cerebral cortex. *Cerebral Cortex*, *14*, 11–22. https://doi.org/10.1093/cercor/bhg087

Gavidia-Bovadilla, G., Kanaan-Izquierdo, S., Mataroa-Serrat, M., & Perera-Lluna, A. (2017). Early prediction of Alzheimer's disease using null longitudinal model-based classifiers. *PLoS One*, *12*, e0168011. https://doi.org/10.1371/journal.pone.0168011

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*, 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Guo, M., Li, Y., Zheng, W., Huang, K., Zhou, L., Hu, X., Yao, Z., & Hu, B. (2020). A novel conversion prediction method of MCI to AD based on longitudinal dynamic morphological features using ADNI structural MRIs. *Journal of Neurology*, *267*, 2983–2997. https://doi.org/10.1007/s00415-020-09890-5

Habes, M., Grothe, M. J., Tunc, B., McMillan, C., Wolk, D. A., & Davatzikos, C. (2020). Disentangling heterogeneity in Alzheimer's disease and related dementias using data-driven methods. *Biological Psychiatry*, *88*, 70–82. https://doi.org/10.1016/j.biopsych.2020.01.016

Irish, M., Landin-Romero, R., Mothakunnel, A., Ramanan, S., Hsieh, S., Hodges, J. R., & Piguet, O. (2018). Evolution of autobiographical memory impairments in Alzheimer's disease and frontotemporal dementia – A longitudinal neuroimaging study. *Neuropsychologia*, *110*, 14–25. https://doi.org/10.1016/j.neuropsychologia.2017.03.014

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's Dementia*, *14*, 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

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., Jr., Ashburner, J., & Frackowiak, R. S. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain*, *131*, 681–689. https://doi.org/10.1093/brain/awm319

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Laakso, M. P., Frisoni, G. B., Könönen, M., Mikkonen, M., Beltramello, A., Geroldi, C., Bianchetti, A., Trabucchi, M., Soininen, H., & Aronen, H. J. (2000). Hippocampus and entorhinal cortex in frontotemporal dementia and Alzheimer's disease: A morphometric MRI study. *Biological Psychiatry*, *47*, 1056–1063. https://doi.org/10.1016/S0006-3223(99)00306-6

Li, B., Jang, I., Riphagen, J., Almaktoum, R., Yochim, K. M., Ances, B. M., Bookheimer, S. Y., Salat, D. H., & For the Alzheimer's Disease Neuroimaging Initiative. (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

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 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

Möller, C., Pijnenburg, Y. A. L., van der Flier, W. M., Versteeg, A., Tijms, B., de Munck, J. C., Hafkemeijer, A., Rombouts, S. A., van der Grond, J., van Swieten, J., Dopper, E., Scheltens, P., Barkhof, F., Vrenken, H., & Wink, A. M. (2016). Alzheimer disease and behavioral variant frontotemporal dementia: Automatic classification based on cortical atrophy for single-subject diagnosis. *Radiology*, *279*, 838–848. https://doi.org/10.1148/radiol.2015150220

Möller, C., Vrenken, H., Jiskoot, L., Versteeg, A., Barkhof, F., Scheltens, P., & van der Flier, W. M. (2013). Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. *Neurobiology of Aging*, *34*, 2014–2022. https://doi.org/10.1016/j.neurobiolaging.2013.02.013

Pankov, A., Binney, R. J., Staffaroni, A. M., Kornak, J., Attygalle, S., Schuff, N., Weiner, M. W., Kramer, J. H., Dickerson, B. C., Miller, B. L., & Rosen, H. J. (2016). Data-driven regions of interest for longitudinal change in frontotemporal lobar degeneration. *NeuroImage Clinical*, *12*, 332–340. https://doi.org/10.1016/j.nicl.2015.08.002

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830 Retrieved from October 22, 2019 http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html

Penny, W., Friston, K., Ashburner, J., Kiebel, S., & Nichols, T. (2007). *Statistical parametric mapping: The analysis of functional brain images*. Elsevier Ltd. https://doi.org/10.1016/B978-0-12-372560-8.X5000-1

Rabinovici, G. D., Seeley, W. W., Kim, E. J., Gorno-Tempini, M. L., Rascovsky, K., Pagliaro, T. A., Allison, S. C., Halabi, C., Kramer, J. H., Johnson, J. K., Weiner, M. W., Forman, M. S., Trojanowski, J. Q., DeArmond, S. J., Miller, B. L., & Rosen, H. (2008). Distinct MRI atrophy patterns in autopsy-proven Alzheimer's disease and frontotemporal lobar degeneration. *American Journal of Alzheimer's Disease and Other Dementias*, *22*, 474–488. https://doi.org/10.1177/1533317507308779

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M. L., Rosen, H., … Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*, 2456–2477. https://doi.org/10.1093/brain/awr179

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*, 1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084

Risacher, S. L., Shen, L., West, J. D., Kim, S., McDonald, B. C., Beckett, L. A., et al. (2010). Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort. *Neurobiology of Aging*, *31*, 1401–1418. https://doi.org/10.1016/j.neurobiolaging.2010.04.029

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 Part B. Neuropsychiatric Genetics*, *74*, 507–514. https://doi.org/10.1002/(SICI)1096-8628(19970919)74:5<507::AID-AJMG11>3.0.CO;2-G

Sintini, I., Graff-Radford, J., Senjem, M. L., Schwarz, C. G., Machulda, M. M., Martin, P. R., Jones, D. T., Boeve, B. F., Knopman, D. S., Kantarci, K., Petersen, R. C., Jack, C. R., Jr., Lowe, V. J., Josephs, K. A., & Whitwell, J. L. (2020). Longitudinal neuroimaging biomarkers differ across Alzheimer's disease phenotypes. *Brain*, *143*, 2281–2294. https://doi.org/10.1093/brain/awaa155

Steketee, R. M. E., Bron, E. E., Meijboom, R., Houston, G. C., Klein, S., Mutsaerts, H. J. M. M., Mendez Orellana, C. P., de Jong, F. J., van Swieten, J. C., van der Lugt, A., & Smits, M. (2016). Early-stage differentiation between presenile Alzheimer's disease and frontotemporal dementia using arterial spin labeling MRI. *European Radiology*, *26*, 244–253. https://doi.org/10.1007/s00330-015-3789-x

Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *WIREs Data Mining and Knowledge Discovery*, *10*, e1379. https://doi.org/10.1002/widm.1379

Storsve, A. B., Fjell, A. M., Tamnes, C. K., Westlye, L. T., Overbye, K., Aasland, H. W., & Walhovd, K. B. (2014). Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: Regions of accelerating and decelerating change. *The Journal of Neuroscience*, *34*, 8488–8498. https://doi.org/10.1523/JNEUROSCI.0391-14.2014

Thompson, W. K., Hallmayer, J., & O'Hara, R. (2011). Design considerations for characterizing psychiatric trajectories across the life span: Application to effects of APOE-ε4 on cerebral cortical thickness in Alzheimer's disease. *The American Journal of Psychiatry*, *168*, 894–903. https://doi.org/10.1176/appi.ajp.2011.10111690

Zhang, J., Liu, M., An, L., Gao, Y., & Shen, D. (2017). Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images. *IEEE Journal of Biomedical and Health Informatics*, *21*, 1607–1616. https://doi.org/10.1109/JBHI.2017.2704614

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pérez-Millan, A., Contador, J., Juncà-Parella, J., Bosch, B., Borrell, L., Tort-Merino, A., Falgàs, N., Borrego-Écija, S., Bargalló, N., Rami, L., Balasa, M., Lladó, A., Sánchez-Valle, R., & Sala-Llonch, R. (2023). Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. *Human Brain Mapping*, 1–11. https://doi.org/10.1002/hbm.26205

10970193, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hbm.26205 by Readcube (Labtiva Inc.), Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Supplementary Material

Intra-Scanner Variability

We studied 3 MRI-longitudinal datasets with 2 time-points, separated 2 years, all having available T1-weighted scans, with equal or equivalent acquisition protocols before and after a scanner upgrade. Being ‘scanner 1’ a 3T Siemens TrioTim MRI scanner and ‘scanner 2’ a 3T Siemens Prisma scanner, the groups were as follows: (1) A group of 14 healthy controls (median and IQR age: 58.4 [54.7, 61.4]) scanned with scanner 1 at both time points (CTR); (2) a group of 17 healthy controls (median and IQR age: 59.7 [49.0, 60.9]) scanned at baseline with scanner 1 and after 2 years with scanner 2 (CTRc), and (3) a group of 12 AD patients (median and IQR age: 60.4 [58.0, 65.8]), scanned with scanner 1 at both time points. We used the longitudinal processing stream of FreeSurfer to obtain summary measures, including subcortical GM volume and cortical thickness (CTh) using the same atlases as the main manuscript. We studied the intra-class correlation coefficient (ICC) to assess the reliability of the metrics for CTR and CTRc groups. We calculated the percent difference (PcD) to obtain the relative difference of GM measures between time points, for all groups. With the PcD measure, we aimed to study the magnitude of the changes due to the scanner in comparison with the effect of AD in atrophy. We further focused on hippocampal atrophy to provide a practical example, as it is considered a biomarker for AD.

We found that the ROI-CTh measures were the most affected by the scanner upgrade, with a mean ICC across regions of 0.8 and 0.92 for the CTRc and CTR groups respectively (36 out of 68 regions were significantly different). However, mean ICC values were close to 1 for both datasets for these metrics, indicating good reliability. As regards PcD, we consistently observed that longitudinal differences were stronger in AD than in CTR or CTRc groups (Figure 1). Mean PcD values for CTh (across parcellations) were: -0.20% for CTR, -1.53% for CTRc and -4.57% for AD. And mean PcD for subcortical GM volumes were (across atlas regions): -3.11% for CTR, -2.76% for CTRc and -10% for AD. Hippocampal atrophy was similar between CTR and CTRc with values close to zero, with no significant differences between them. In AD, hippocampal volume loss was stronger, showing significant differences between CTR and CTRc (Figure 2).

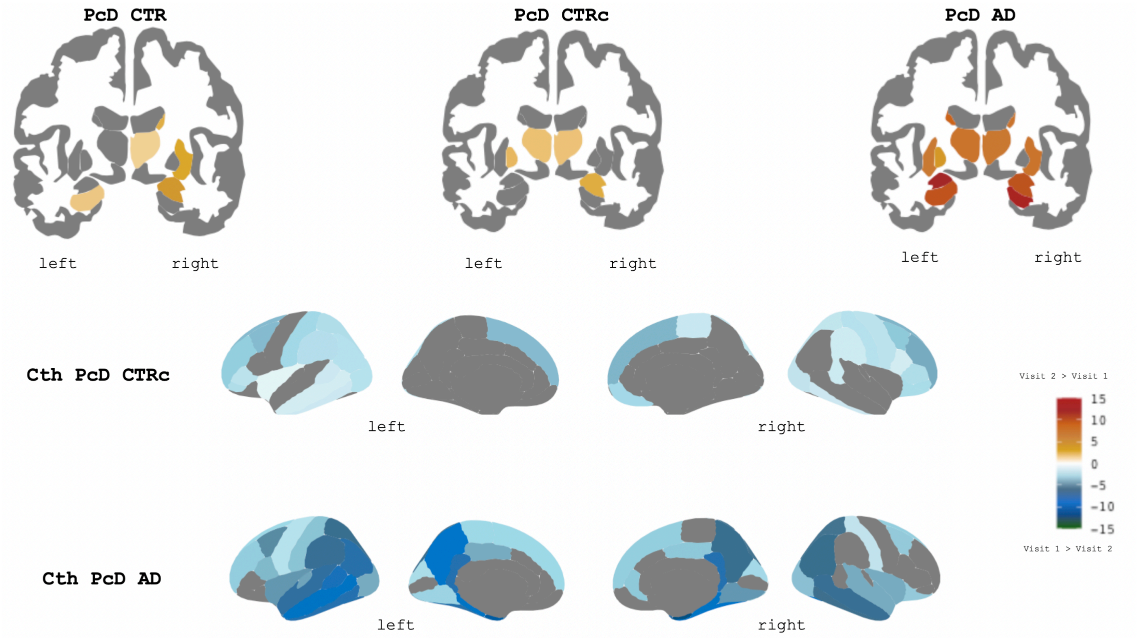
PcD CTR
PcD CTRc
PcD AD
left
right
Cth PcD CTRc
Cth PcD AD
Visit 2 > Visit 1
15
10
5
0
-5
-10
-15
Visit 1 > Visit 2

Figure 1: Mean PcD values for CTR, CTRc, and AD. Only significant regions are shown. The cool color scale represents Visit 1 > Visit 2 and the warm scale represents Visit 2 > Visit 1 within the PcD

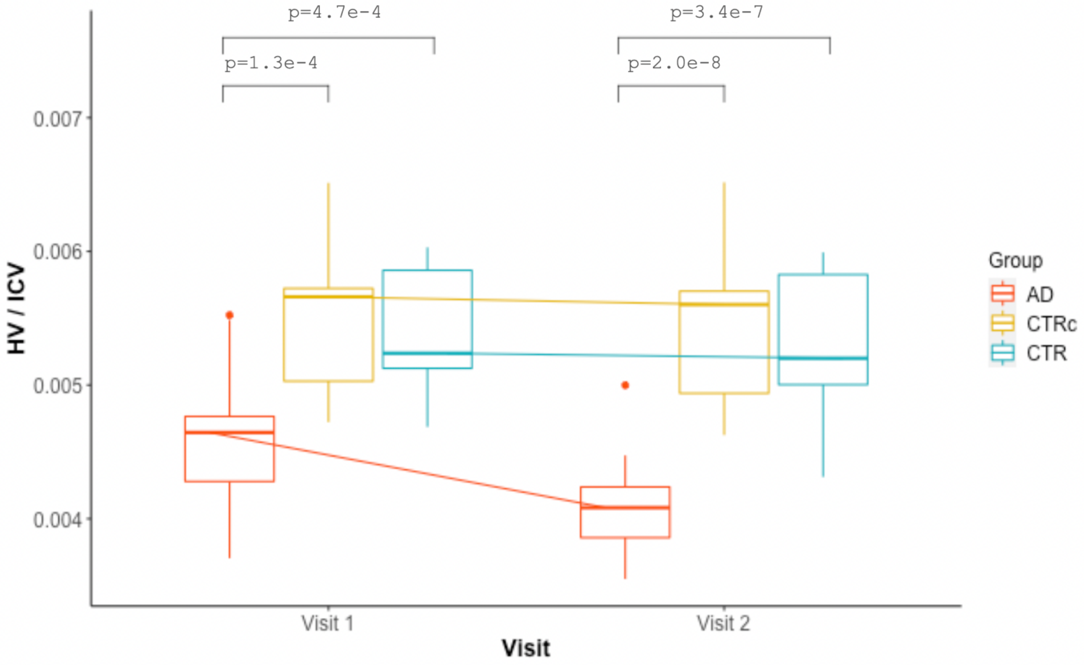

Figure 2: Boxplot of the HV normalized by the intracranial volume (ICV) for each visit. Pairwise comparisons show significant differences in some comparisons.

Disease Classification with Age: Cross-sectional study

The ML algorithm, which combines all cortical CTh values and the subcortical GM volumes (obtained with FreeSurfer) including age, had the following accuracy: 81.8 ± 13.2 % in the CTR vs AD classification, 82.6 ± 14.8 % for CTR vs FTD, 62.1 ± 7.8 % for AD vs FTD and 61.9 ± 10.5 % when discriminating the 3 groups.

Disease Classification with Sex: Cross-sectional study

The ML algorithm, which combines all cortical CTh values and the subcortical GM volumes (obtained with FreeSurfer) including the sex variable, had the following accuracy: 81.8 ± 14.2 % in the CTR vs AD classification, 82.6 ± 14.8 % for CTR vs FTD, 61.3 ± 11.6 % for AD vs FTD and 61.1 ± 8.5 % when discriminating the 3 groups.

Cross-sectional analyses: Brain patterns and relevant regions

The complete list of regions ordered from most to least important for AD and CTR is shown in Table 1.

Table 1: Relevant regions for the AD and CTR classification.

Region	Weight (absolute value)
right supramarginal	0.132398024
left supramarginal	0.131250765
right precuneus	0.129766071
right superior temporal	0.129413757
left superior frontal	0.128659062
right middle temporal	0.125845316
right superior frontal	0.124020465
right caudal middle frontal	0.123852731
right inferior parietal	0.123621522
left caudal middle frontal	0.123231611
left middle temporal	0.122663053
right fusiform	0.120213889
left superior temporal	0.120048274
right superior parietal	0.118839236
right insula	0.118765178
left precuneus	0.118004610
left inferior parietal	0.115784908

left rostral middle frontal	0.115510007
right rostral middle frontal	0.114876770
left superior parietal	0.114797351
left fusiform	0.114594705
right bankssts	0.113284782
right postcentral	0.109881772
left pars opercularis	0.109683926
left inferior temporal	0.108852163
right inferior temporal	0.107537313
left postcentral	0.107362484
right pars opercularis	0.106304893
left amygdala	0.104860639
left bankssts	0.104428045
right entorhinal	0.103358363
right precentral	0.102994215
left insula	0.102988526
right putamen	0.102658807
left pars triangularis	0.101768935
right lateral ventricle	0.101735621
right hippocampus	0.100563834
left precentral	0.100517210
right pars triangularis	0.100163562
right amygdala	0.100078512
left hippocampus	0.099910649
left putamen	0.099615513
left lateral occipital	0.097343502
left posterior cingulate	0.096700975
right posterior cingulate	0.096259173
right lateral occipital	0.095844236
left entorhinal	0.094493914
left lateral ventricle	0.092100142
right lingual	0.090868299
right paracentral	0.089423868
right temporal pole	0.088178150
left thalamus proper	0.087852884
left medial orbitofrontal	0.087723867
right parahippocampal	0.087467138
left temporal pole	0.085780580
right thalamus proper	0.084828698
left lateral orbitofrontal	0.083730125
right pars orbitalis	0.083725304
left isthmus cingulate	0.082190575

left paracentral	0.082091817
left lingual	0.080314217
left caudate	0.079843286
left pars orbitalis	0.079370561
right transverse temporal	0.078883607
right cuneus	0.078698332
right lateral orbitofrontal	0.076470462
left parahippocampal	0.076362321
right isthmus cingulate	0.075751401
right caudate	0.072829449
right medial orbitofrontal	0.069739841
left frontal pole	0.068055185
left cuneus	0.065380244
right pericalcarine	0.063486021
left transverse temporal	0.063066450
left pericalcarine	0.061865656
right frontal pole	0.053554998
left rostral anterior cingulate	0.048428281
right caudal anterior cingulate	0.046019643
right cerebellum cortex	0.043014358
left caudal anterior cingulate	0.039211475
left cerebellum cortex	0.038275172
right rostral anterior cingulate	0.032349503
left pallidum	0.011922263
right pallidum	0.003399224

The complete list of regions ordered from most to least important for FTD and CTR is shown in Table 2.

Table 2: Relevant regions for the FTD and CTR classification.

Region	**Weight (absolute value)**
right supramarginal	0.12619778
right superior frontal	0.12332762
right inferior parietal	0.12161832
right superior temporal	0.12083992
left superior frontal	0.11887638
right middle temporal	0.11872332
right precuneus	0.11845395
left supramarginal	0.11843901
right insula	0.11704590

right precentral	0.11701397
right caudal middle frontal	0.11673290
left precuneus	0.11561300
left middle temporal	0.11526148
left superior temporal	0.11500142
left rostral middle frontal	0.11475845
left inferior parietal	0.11467779
left insula	0.11401137
left lateral orbitofrontal	0.11376168
right bankssts	0.11361550
right fusiform	0.11165565
right superior parietal	0.11154816
left medial orbitofrontal	0.11091987
right pars opercularis	0.11069583
left superior parietal	0.11067223
right inferior temporal	0.10994370
left precentral	0.10761547
right putamen	0.10754690
left fusiform	0.10590622
left caudal middle frontal	0.10558708
left inferior temporal	0.10462069
Left putamen	0.10385312
right lateral orbitofrontal	0.10299590
right rostral middle frontal	0.10284533
Left hippocampus	0.10214976
right postcentral	0.10201190
right lateral ventricle	0.10122240
left temporal pole	0.10118039
right entorhinal	0.10117219
right temporal pole	0.10074654
right medial orbitofrontal	0.09983957
left pars opercularis	0.09925066
left pars triangularis	0.09922243
left lateral ventricle	0.09883158
right pars triangularis	0.09867895
left postcentral	0.09709643
left entorhinal	0.09616336
left pars orbitalis	0.09606334
right lateral occipital	0.09568428
right pars orbitalis	0.09518599
right parahippocampal	0.09420009
left posterior cingulate	0.09401839

left amygdala	0.09372896
left bankssts	0.09369281
left paracentral	0.09330762
right hippocampus	0.09330148
right thalamus proper	0.09304978
right amygdala	0.09242262
right paracentral	0.09135317
right posterior cingulate	0.09069424
left thalamus proper	0.08921889
left lateral occipital	0.08871286
right transverse temporal	0.08668020
right caudate	0.08429756
left rostral anterior cingulate	0.08214760
Left caudate	0.08184923
right cuneus	0.08122028
right lingual	0.08062650
left parahippocampal	0.07943174
left frontal pole	0.07608782
right isthmus cingulate	0.07415441
right rostral anterior cingulate	0.07143054
left lingual	0.06932710
left pericalcarine	0.06840119
left cuneus	0.06686781
right pericalcarine	0.06636366
left isthmus cingulate	0.06569378
left transverse temporal	0.06448232
right frontal pole	0.05999339
right caudal anterior cingulate	0.04743249
right cerebellum cortex	0.03503372
left cerebellum cortex	0.03419993
left caudal anterior cingulate	0.03183794
right pallidum	0.02479197
left pallidum	0.01180515

Longitudinal Classification with Age

The ML longitudinal algorithm, combining all cortical CTh values and the subcortical GM volumes (obtained with FreeSurfer) including age, had the following accuracy: 90.0 ± 14.7 % in the CTR vs AD classification, 88.0 ± 16.4 % for CTR vs FTD, 72.5 ± 40.2 % for AD vs FTD and 70.5 ± 13.9 % when discriminating the 3 groups.

The ML longitudinal algorithm combining all cortical CTh values and the subcortical GM volumes at each visit, the difference between both visits for each measure, and age, had the following accuracy: 93.5 ± 12.2 % in the CTR vs AD classification, 85.5 ± 18.7 % for CTR vs FTD, 60.8 ± 33.4 % for AD vs FTD and 79.3 ± 19.4 % when discriminating the 3 groups.

Disease Classification with Sex: Longitudinal

The ML longitudinal algorithm, which combines all cortical CTh values and the subcortical GM volumes (obtained with FreeSurfer) including the sex variable, had the following accuracy: 90.0 ± 14.7 % in the CTR vs AD classification, 88.0 ± 16.4 % for CTR vs FTD, 75.0 ± 36.9 % for AD vs FTD and 70.5 ± 13.9 % when discriminating the 3 groups.

Longitudinal analyses: Brain patterns and relevant regions

The complete list of regions ordered from most to least important for AD and CTR for visit 1 is shown in Table 3.

Table 3: Relevant regions for the AD and CTR classification for visit 1.

Hemisphere	Region	Weight (absolute value)
right	pars triangularis	0.118172
right	entorhinal	0.110376
left	transverse temporal	0.108773
left	pars orbitalis	0.105024
left	supramarginal	0.102022
left	rostral anterior cingulate	0.101265
left	cuneus	0.100044
left	temporal pole	0.099950
left	postcentral	0.097928
left	precentral	0.093428
left	posterior cingulate	0.092619
right	caudal middle frontal	0.091058
right	caudal anterior cingulate	0.090806
right	parahippocampal	0.090057
right	frontal pole	0.089258

left	parahippocampal	0.088595
right	insula	0.088363
right	thalamus proper	0.087210
right	amygdala	0.086191
right	lateral ventricle	0.083690
right	rostral anterior cingulate	0.083240
left	bankssts	0.081812
left	pallidum	0.081781
right	paracentral	0.081355
right	transverse temporal	0.081115
left	caudal anterior cingulate	0.077248
right	pericalcarine	0.076638
right	temporal pole	0.076573
left	frontal pole	0.076310
right	bankssts	0.076241
left	caudate	0.074540
right	precentral	0.073781
left	caudal middle frontal	0.073628
left	middle temporal	0.073518
left	putamen	0.072987
left	inferior temporal	0.071299
right	supramarginal	0.070180
right	lingual	0.069422
left	superior temporal	0.068992
left	lateral ventricle	0.068276
right	superior temporal	0.067320
right	postcentral	0.065620
right	cuneus	0.065499
left	insula	0.065208
left	lateral orbitofrontal	0.064892
left	rostral middle frontal	0.059575
right	inferior temporal	0.055750
left	entorhinal	0.055531
left	lingual	0.055197
right	isthmus cingulate	0.054670
right	caudate	0.054669
right	putamen	0.054669
right	precuneus	0.050482
left	lateral occipital	0.050000
right	fusiform	0.049660
right	medial orbitofrontal	0.049584
left	precuneus	0.048811

right	lateral orbitofrontal	0.046896
right	superior frontal	0.046674
left	isthmus cingulate	0.046614
left	pericalcarine	0.045486
right	posterior cingulate	0.044347
left	amygdala	0.043514
right	hippocampus	0.042535
left	hippocampus	0.042131
left	superior frontal	0.035765
right	lateral occipital	0.035728
left	fusiform	0.031348
right	rostral middle frontal	0.030046
right	inferior parietal	0.029368
right	pars opercularis	0.028268
left	superior parietal	0.024416
right	middle temporal	0.020374
left	inferior parietal	0.017754
right	pallidum	0.017304
left	pars opercularis	0.012286
left	thalamus proper	0.011450
left	medial orbitofrontal	0.010741
left	pars triangularis	0.009616
right	superior parietal	0.008360
right	pars orbitalis	0.007793
left	paracentral	0.006384

The complete list of regions ordered from most to least important for AD and CTR for visit 2 is shown in Table 4.

Table 4: Relevant regions for the AD and CTR classification for visit 2.

Hemisphere	Region	Weight (absolute value)
right	paracentral	0.113792
right	inferior temporal	0.108894
right	amygdala	0.108818
right	parahippocampal	0.107927
right	bankssts	0.107347
right	medial orbitofrontal	0.106943
right	cuneus	0.106709
right	temporal pole	0.103368

right	fusiform	0.099968
left	postcentral	0.099282
left	lingual	0.097540
right	entorhinal	0.096164
left	middle temporal	0.094651
left	inferior parietal	0.094532
left	parahippocampal	0.094078
left	temporal pole	0.093919
left	pars triangularis	0.093051
right	putamen	0.092860
left	pars orbitalis	0.089565
right	lateral ventricle	0.089344
right	frontal pole	0.089172
left	caudal middle frontal	0.089151
right	transverse temporal	0.087917
right	thalamus proper	0.087723
left	cuneus	0.086478
left	frontal pole	0.084532
right	supramarginal	0.083918
right	caudate	0.083690
left	superior parietal	0.083148
right	pallidum	0.081741
right	rostral anterior cingulate	0.081133
left	entorhinal	0.079680
left	putamen	0.078981
right	precentral	0.078617
left	thalamus proper	0.075900
right	superior parietal	0.074704
left	pars opercularis	0.074513
right	pars triangularis	0.073674
right	middle temporal	0.073556
right	hippocampus	0.072210
left	lateral ventricle	0.071315
right	superior temporal	0.071017
right	precuneus	0.069748
right	pars opercularis	0.065132
right	lingual	0.064493
left	pericalcarine	0.063289
right	rostral middle frontal	0.061189
left	superior temporal	0.060877
left	posterior cingulate	0.059678
left	rostral anterior cingulate	0.058585

right	pars orbitalis	0.058152
left	medial orbitofrontal	0.057767
left	fusiform	0.056820
left	paracentral	0.055976
left	inferior temporal	0.055475
left	lateral orbitofrontal	0.054304
left	caudal anterior cingulate	0.052900
left	superior frontal	0.052121
left	rostral middle frontal	0.052115
right	postcentral	0.048481
right	insula	0.047640
right	caudal middle frontal	0.047179
right	superior frontal	0.045793
left	caudate	0.045378
right	inferior parietal	0.045032
left	isthmus cingulate	0.045026
left	hippocampus	0.044391
left	precuneus	0.042913
left	precentral	0.042860
right	isthmus cingulate	0.040560
left	lateral occipital	0.034668
right	pericalcarine	0.032705
left	pallidum	0.028913
right	lateral occipital	0.028649
right	lateral orbitofrontal	0.023552
left	transverse temporal	0.019127
right	caudal anterior cingulate	0.019027
left	supramarginal	0.018742
left	bankssts	0.006251
left	insula	0.005267
left	amygdala	0.004162
right	posterior cingulate	0.001359

The complete list of regions ordered from most to least important for FTD and CTR for visit 1 is shown in Table 5.

Table 5: Relevant regions for the FTD and CTR classification for visit 1.

Hemisphere	**Region**	**Weight (absolute value)**
right	pars triangularis	0.105423

left	pars orbitalis	0.102475
left	rostral anterior cingulate	0.097304
right	paracentral	0.095766
left	insula	0.091756
left	posterior cingulate	0.091293
right	thalamus proper	0.090531
left	postcentral	0.088563
right	frontal pole	0.088454
right	temporal pole	0.088353
right	insula	0.087168
right	precentral	0.086933
left	parahippocampal	0.086874
left	caudate	0.086819
left	inferior temporal	0.086419
left	caudal middle frontal	0.085461
right	bankssts	0.085295
right	rostral anterior cingulate	0.085151
left	transverse temporal	0.084872
left	middle temporal	0.084822
right	amygdala	0.083502
left	lateral ventricle	0.082818
right	lateral ventricle	0.082746
right	cuneus	0.082366
left	cuneus	0.081553
right	pericalcarine	0.081429
right	isthmus cingulate	0.081139
right	precuneus	0.080997
right	postcentral	0.080438
right	inferior temporal	0.079328
left	bankssts	0.077580
right	entorhinal	0.077305
right	parahippocampal	0.077143
left	supramarginal	0.077142
right	inferior parietal	0.077078
left	pallidum	0.076968
left	inferior parietal	0.073785
left	putamen	0.071318
left	frontal pole	0.070684
right	transverse temporal	0.070083
left	fusiform	0.068111
right	supramarginal	0.067212
right	caudal anterior cingulate	0.066970

left	superior temporal	0.066159
left	caudal anterior cingulate	0.065228
right	posterior cingulate	0.065010
left	temporal pole	0.063964
left	lingual	0.063639
right	middle temporal	0.063611
left	hippocampus	0.062493
left	lateral occipital	0.061538
right	caudate	0.060309
right	putamen	0.060309
right	caudal middle frontal	0.060267
left	amygdala	0.059806
right	superior temporal	0.059646
left	isthmus cingulate	0.058942
right	lateral orbitofrontal	0.058269
left	precentral	0.054454
right	medial orbitofrontal	0.052517
left	superior frontal	0.052147
right	superior frontal	0.050172
left	lateral orbitofrontal	0.047253
left	rostral middle frontal	0.046318
left	pars opercularis	0.043936
left	entorhinal	0.042194
right	lateral occipital	0.041520
left	precuneus	0.038030
right	lingual	0.037376
left	medial orbitofrontal	0.036232
right	hippocampus	0.032376
right	pallidum	0.031284
right	pars opercularis	0.031156
right	rostral middle frontal	0.028339
left	thalamus proper	0.023958
left	superior parietal	0.023169
left	pericalcarine	0.021973
right	fusiform	0.020885
left	paracentral	0.018195
right	pars orbitalis	0.013612
left	pars triangularis	0.012474
right	superior parietal	0.000610

The complete list of regions ordered from most to least important for FTD and CTR for visit 2 is shown in Table 6.

Table 6: Relevant regions for the FTD and CTR classification for visit 2.

Hemisphere	Region	Weight (absolute value)
right	temporal pole	0.105374
right	inferior temporal	0.103468
right	bankssts	0.099943
right	cuneus	0.096823
right	entorhinal	0.096567
right	pars opercularis	0.094847
right	lateral ventricle	0.093809
right	pars orbitalis	0.091971
left	temporal pole	0.091373
right	thalamus proper	0.091193
left	inferior parietal	0.089250
left	pars opercularis	0.088880
right	precentral	0.088848
left	thalamus proper	0.088363
right	frontal pole	0.088288
left	cuneus	0.087542
right	hippocampus	0.087128
left	paracentral	0.087083
left	putamen	0.086956
left	parahippocampal	0.086598
right	transverse temporal	0.086444
right	superior temporal	0.086016
left	precuneus	0.085628
left	lateral ventricle	0.084933
left	precentral	0.083897
right	pars triangularis	0.083271
right	caudate	0.082746
left	caudal middle frontal	0.082737
right	rostral anterior cingulate	0.081898
left	postcentral	0.081469
left	frontal pole	0.080682
left	pericalcarine	0.079441
right	amygdala	0.078945
left	rostral anterior cingulate	0.078445
right	paracentral	0.077605

left	caudate	0.077212
right	supramarginal	0.074233
left	pars orbitalis	0.072481
left	supramarginal	0.072293
right	precuneus	0.071317
right	rostral middle frontal	0.070820
left	lingual	0.070494
right	putamen	0.069779
right	superior parietal	0.068416
right	medial orbitofrontal	0.067916
right	pallidum	0.067289
right	parahippocampal	0.066909
right	middle temporal	0.066576
left	pars triangularis	0.064636
right	lingual	0.064115
left	superior frontal	0.063180
left	medial orbitofrontal	0.061147
left	lateral orbitofrontal	0.059321
right	superior frontal	0.058516
right	postcentral	0.057398
left	entorhinal	0.056978
left	isthmus cingulate	0.054416
left	middle temporal	0.053484
left	rostral middle frontal	0.053454
right	isthmus cingulate	0.052028
right	insula	0.051652
left	superior temporal	0.051070
right	fusiform	0.050336
left	pallidum	0.048768
right	lateral occipital	0.048630
left	hippocampus	0.047515
left	inferior temporal	0.045986
left	superior parietal	0.042554
left	lateral occipital	0.039869
left	transverse temporal	0.038380
left	bankssts	0.035047
left	fusiform	0.035022
left	posterior cingulate	0.034307
right	inferior parietal	0.032824
right	pericalcarine	0.032559
left	caudal anterior cingulate	0.032013
right	caudal middle frontal	0.023647

right	lateral orbitofrontal	0.021188
right	caudal anterior cingulate	0.011242
left	amygdala	0.006234
right	posterior cingulate	0.004857
left	insula	0.004029

TREBALL 4

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, Nuria Bargalló, Mircea Balasa, Albert Lladó, Raquel Sánchez-Valle, Roser Sala-Llonch.

-En revisió-

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, *†]

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

[2]Institut de Neurociències, University of Barcelona, Barcelona, 08036, Spain.

[3]Department of Biomedicine, Faculty of Medicine, University of Barcelona, Barcelona, 08036, 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, Barcelona, Spain. CIBER de Salud Mental, Instituto de Salud Carlos III.Magnetic Resonance Image Core Facility, IDIBAPS

[7]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), 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, SVM, individual probability

Key points

- The classification between Alzheimer’s disease and Frontotemporal dementia should account for probabilities.
- There is a gray zone of subjects with non-confident or unclear diagnoses that should be studied deeper.
- The availability of multimodal data contributes to improving diagnosis.

ABSTRACT

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

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

Results: 215 AD patients (65 ± 10 years, 137 women), 103 FTD patients (64 ± 8 years, 49 women), and 173 CTR (59 ± 15 years, 106 women) were studied. With MRI data only, we obtained accuracies of 88% in the AD vs. healthy controls (CTR) classification, 87% for FTD vs. CTR, 82% for AD vs. FTD, and 80% when differentiating the three groups. A total of 74% of FTD and 73% of AD participants have a high (≥ 0.8) probability of accurate diagnosis in the FTD vs. AD comparison. Adding CSF-NfL and 14-3-3 levels did not improve the accuracy or the number of patients in the high diagnosis confidence group.

Conclusion: Our approach, using ML algorithms that provide individual probabilities, holds promise towards individual diagnoses, especially in doubtful cases as support to clinical findings or in settings with limited access to expert diagnoses.

1. INTRODUCTION

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

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

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

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

In this study, we aimed to develop a probabilistic computer-aided diagnosis of FTD and AD with MRI data to study the overlapping and differential brain patterns of these two neurodegenerative disorders. Then, we addressed the clinical problem of diagnosis confidence using individual prediction probabilities. We proposed investigating whether combining MRI and CSF biomarkers could help differentiate these two dementias and gain more confidence in the diagnosis.

2. MATERIALS AND METHODS

2.1. Participants

We recruited the participants from the Alzheimer's disease and other cognitive disorders unit of the Hospital Clínic de Barcelona (HCB), Barcelona, Spain. All participants underwent a complete clinical and cognitive evaluation and a 3T high-resolution structural MRI scan.

All AD participants fulfilled the criteria for mild dementia due to AD [1, 2] supported by the CSF biomarkers profile suggesting underlying AD neuropathology according to National Institute on Aging/Alzheimer's Association Research Framework 2018 [31]. The FTD participants fulfilled the diagnostic criteria for either behavioral variant frontotemporal dementia (bvFTD) or FTD-related primary progressive aphasia (PPA) phenotypes, including Semantic Variant Primary Progressive Aphasia (svPPA) and Nonfluent Variant Primary Progressive Aphasia (nfvPPA) [32, 33]. Healthy adults (CTR) had cognitive performance within the normative range.

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

2.2 Biochemical markers

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

2.3 MRI acquisition

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

2.4 MRI processing

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

2.5 MRI-based individual probabilistic classification algorithm

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

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

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

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

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

3. RESULTS

3.1. Sample demographics

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

3.2. MRI-based probabilistic classification algorithm

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

As can be seen in 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. In the FTD vs. CTR comparison, we found 74% FTD participants and 73% CTR participants with a probability ≥ 0.8. Finally, when discriminating AD vs. FTD, we found 73% AD participants and 74% FTD participants with probabilities above 0.8 for being classified as AD or FTD, respectively. Figure 2 shows the density of the individual probabilities and how the distribution between the clinical and the algorithm diagnosis is distributed within the group with an individual probability ≥ 0.8. Notably, the algorithm diagnosis and the clinical algorithm did not always coincide, also inside the higher probability of 0.8 (high confidence).

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

3.3. Important MRI regions for classification

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

The results of the most crucial regions in the classifications considering bvFTD and PPA participants are shown in 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 did not improve the results. However, some conclusions could be derived from the results.

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

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

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

4. DISCUSSION

In this study, we implemented a machine learning algorithm that discriminates FTD and AD subjects using data from structural MRI. In addition, our algorithm was able to differentiate subtypes of FTD with good accuracy. Clinical diagnosis requires decisions at the individual level, and the degree of confidence in the diagnosis is key in managing the patient. We approach the clinical question of diagnosis confidence using individual probabilities. Among our key results, we found that 74% FTD and 73% AD participants showed an individual probability≥0.8 of being well-classified by the algorithm in the FTD vs. AD comparison. Adding CSF neurodegeneration markers (NfL and 14-3-3) levels did not significantly improve the diagnosis classification or the number of patients with high individual probability for the diagnosis.

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

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

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

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

Other studies using multimodal information also reported high classification accuracy combining data from different imaging modalities or other biological and clinical measures [24, 40, 49, 58]. Even so, in some cases, our scores with only structural MRI data showed better accuracy [18, 46, 59, 60]. Here, we evaluated if adding CSF data to the MRI could improve the accuracy or the number of participants with a high diagnosis confidence. In our cohort, adding NfL and 14-3-3 CSF data to the MRI data provided low benefit for the accuracy of the group classification or the number of participants with high individual diagnosis probability. The fact that both, NfL and 14-3-3 levels and MRI data reflect neurodegeneration and no other aspects of the pathophysiological processes in these diseases could explain the lack of improvement by adding CSF data. We did not use Aβ42, t-tau, and p-tau to avoid circularity, as they were used for the AD clinical diagnosis.

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

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

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

REFERENCES

1. Albert MS, DeKosky ST, Dickson D, et al (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

2. McKhann GM, Knopman DS, Chertkow H, et al (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's and Dementia 7:263–269. https://doi.org/10.1016/j.jalz.2011.03.005

3. Alcolea D, Vilaplana E, Suárez-Calvet M, et al (2017) CSF sAPPβ, YKL-40, and neurofilament light in frontotemporal lobar degeneration. Neurology 89:178–188. https://doi.org/10.1212/WNL.0000000000004088

4. Antonell A, Tort-Merino A, Ríos J, et al (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

5. McFerrin MB, Chi X, Cutter G, Yacoubian TA (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

6. Rohrer JD, Woollacott IOC, Dick KM, et al (2016) Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. Neurology 87:1329–1336. https://doi.org/10.1212/WNL.0000000000003154

7. Davatzikos C, Resnick SM, Wu X, et al (2008) Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. NeuroImage 41:1220–1227. https://doi.org/10.1016/j.neuroimage.2008.03.050

8. Du AT, Schuff N, Kramer JH, et al (2007) Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain 130:1159–1166. https://doi.org/10.1093/brain/awm016

9. Bocchetta M, Todd EG, Peakman G, et al (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

10. Borrego-Écija S, Sala-Llonch R, van Swieten J, et al (2021) Disease-related cortical thinning in presymptomatic granulin mutation carriers. NeuroImage: Clinical 29:. https://doi.org/10.1016/j.nicl.2020.102540

11. Canu E, Agosta F, Mandic-Stojmenovic G, et al (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

12. Contador J, Pérez-Millán A, Tort-Merino A, et al (2021) Longitudinal brain atrophy and CSF biomarkers in early-onset Alzheimer's disease. NeuroImage: Clinical 32:102804. https://doi.org/10.1016/j.nicl.2021.102804

13. Möller C, Hafkemeijer A, Pijnenburg YAL, et al (2015) Joint assessment of white matter integrity, cortical and subcortical atrophy to distinguish AD from behavioral variant FTD: A two-center study. NeuroImage: Clinical 9:418–429. https://doi.org/10.1016/j.nicl.2015.08.022

14. Möller C, Vrenken H, Jiskoot L, et al (2013) Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. Neurobiology of Aging 34:2014–2022. https://doi.org/10.1016/j.neurobiolaging.2013.02.013

15. Abraham A, Pedregosa F, Eickenberg M, et al (2014) Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics 8:

16. Frizzell TO, Glashutter M, Liu CC, et al (2022) Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: A systematic review. Ageing Research Reviews 77:101614. https://doi.org/10.1016/j.arr.2022.101614

17. Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, et al (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

18. Bron EE, Smits M, Papma JM, et al (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

19. Chagué P, Marro B, Fadili S, et al (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

20. Klöppel S, Stonnington CM, Chu C, et al (2008) Automatic classification of MR scans in Alzheimer's disease. Brain 131:681–689. https://doi.org/10.1093/brain/awm319

21. Möller C, Pijnenburg YAL, Van Der Flier WM, et al (2016) Alzheimer disease and behavioral variant frontotemporal dementia: Automatic classification based on cortical atrophy for single-subject diagnosis. Radiology 279:838–848. https://doi.org/10.1148/radiol.2015150220

22. Pérez-Millan A, Contador J, Juncà-Parella J, et al (2023) Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. Human Brain Mapping. https://doi.org/10.1002/hbm.26205

23. Bisenius S, Mueller K, Diehl-Schmid J, et al (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

24. Bron EE, Klein S, Papma JM, et al (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

25. Cuingnet R, Gerardin E, Tessieras J, et al (2011) Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage 56:766–781. https://doi.org/10.1016/j.neuroimage.2010.06.013

26. Magnin B, Mesrob L, Kinkingnéhun S, et al (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

27. Meyer S, Mueller K, Stuke K, et al (2017) Predicting behavioral variant frontotemporal dementia with pattern classification in multi-center structural MRI data. NeuroImage: Clinical 14:656–662. https://doi.org/10.1016/j.nicl.2017.02.001

28. Mendez MF (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

29. Wojtas A, Heggeli KA, Finch N, et al (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

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

31. Jack CR, Bennett DA, Blennow K, et al (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

32. Gorno-Tempini ML, Hillis AE, Weintraub S, et al (2011) Classification of primary progressive aphasia and its variants. Neurology 76:1006–14. https://doi.org/10.1212/WNL.0b013e31821103e6

33. Rascovsky K, Hodges JR, Knopman D, et al (2011) Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 134:2456–2477. https://doi.org/10.1093/brain/awr179

34. Fischl B, Van Der Kouwe A, Destrieux C, et al (2004) Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex 14:11–22. https://doi.org/10.1093/cercor/bhg087

35. Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of

the United States of America 97:11050–11055. https://doi.org/10.1073/pnas.200033797

36. Desikan RS, Ségonne F, Fischl B, et al (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31:968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

37. Seidman LJ, Faraone SV, Goldstein JM, et al (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

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

39. Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12:2825–2830

40. Dukart J, Mueller K, Horstmann A, et al (2011) Combined Evaluation of FDG-PET and MRI Improves Detection and Differentiation of Dementia. PLOS ONE 6:e18111. https://doi.org/10.1371/journal.pone.0018111

41. Kim JP, Kim J, Park YH, et al (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

42. Li B, Jang I, Riphagen J, et al (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

43. Lin W, Tong T, Gao Q, et al (2018) Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer’s Disease Prediction From Mild Cognitive Impairment. Frontiers in Neuroscience 12:

44. Moore PJ, Lyons TJ, Gallacher J, Initiative for the ADN (2019) Random forest prediction of Alzheimer’s disease using pairwise selection from time series data. PLOS ONE 14:e0211558. https://doi.org/10.1371/journal.pone.0211558

45. Salvatore C, Cerasa A, Battista P, et al (2015) Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer’s disease: a machine learning approach. Front Neurosci 9:307. https://doi.org/10.3389/fnins.2015.00307

46. Wang J, Redmond SJ, Bertoux M, et al (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:

47. Wang S-H, Phillips P, Sui Y, et al (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

48. Basheera S, Sai Ram MS (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

49. Dashtipour K, Taylor W, Ansari S, et al (2022) Detecting Alzheimer's Disease Using Machine Learning Methods. In: Ur Rehman M, Zoha A (eds) Body Area Networks. Smart IoT and Big Data for Intelligent Health Management. Springer International Publishing, Cham, pp 89–100

50. McCarthy J, Collins DL, 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

51. Avants BB, Cook PA, Ungar L, et al (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

52. Dickerson BC, Goncharova I, Sullivan MP, et al (2001) MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiology of Aging 22:747–754. https://doi.org/10.1016/S0197-4580(01)00271-8

53. Frisoni GB, Fox NC, Jack CR, et al (2010) The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 6:67–77. https://doi.org/10.1038/nrneurol.2009.215

54. Gil-Navarro S, Lladó A, Rami L, et al (2013) Neuroimaging and Biochemical Markers in the Three Variants of Primary Progressive Aphasia. DEM 35:106–117. https://doi.org/10.1159/000346289

55. Gordon E, Rohrer JD, Fox NC (2016) Advances in neuroimaging in frontotemporal dementia. Journal of Neurochemistry 138:193–210. https://doi.org/10.1111/jnc.13656

56. Hodges JR, 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

57. Jack CR, Petersen RC, Xu Y, et al (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

58. Chen Y, Sha M, Zhao X, et al (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

59. Bouts MJRJ, Möller C, Hafkemeijer A, et al (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

60. Dyrba M, Grothe M, Kirste T, Teipel SJ (2015) Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. Human Brain Mapping 36:2118–2131. https://doi.org/10.1002/hbm.22759

Acknowledgments

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

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Agnès Pérez-Millan: https://orcid.org/0000-0002-3006-9792
Bertrand Thirion: https://orcid.org/0000-0001-5018-7895
Neus Falgàs: https://orcid.org/0000-0002-3404-2765
Sergi Borrego-Écija: https://orcid.org/0000-0003-2557-0010
Beatriz Bosch: https://orcid.org/0000-0002-6094-0024
Jordi Juncà-Parella: https://orcid.org/0000-0002-4772-2647
Adrià Tort-Merino: https://orcid.org/0000-0002-5646-0482
Jordi Sarto: https://orcid.org/0000-0002-6030-3375
Josep Maria Augé: https://orcid.org/0000-0001-6383-7308
Anna Antonell: https://orcid.org/0000-0002-3286-8459
Nuria Bargalló: https://orcid.org/0000-0001-6284-5402
Mircea Balasa: https://orcid.org/0000-0002-1795-4228
Albert Lladó: https://orcid.org/0000-0002-5066-4150
Raquel Sánchez-Valle: https://orcid.org/0000-0001-7750-896X
Roser Sala-Llonch: https://orcid.org/0000-0003-3576-0475

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 subjects, AD: Alzheimer's disease, FTD: frontotemporal dementia, NfL: neurofilament light chain.

	CTR	**AD**	**FTD**	**CTR-AD p-values**	**CTR-FTD p-values**	**AD-FTD p-values**
N MRI	173	215	103	---	---	---
Sex at MRI, Men/Women	67/106	78/137	54/49	0.67	**0.049**	**0.022**
Age at MRI, years (SD)	59.4 (15.0)	65.0 (9.9)	63.7 (8.3)	**1.3e-5**	**0.0045**	0.39
N CSF NfL	112	175	78	---	---	---

CSF NfL, pg/mL (SD)	536.1 (312.6)	1134.7 (587.1)	2340.6 (1736.3)	**1.2e-07**	**< 2e-16**	**5.9e-06**
N CSF 14-3-3	50	68	64	---	---	---
CSF 14-3-3, pg/mL (SD)	2531.9 (748.2)	5727.3 (2303.5)	4234.9 (1869.1)	**< 2e-16**	**3.0e-06**	**5.9e-06**

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

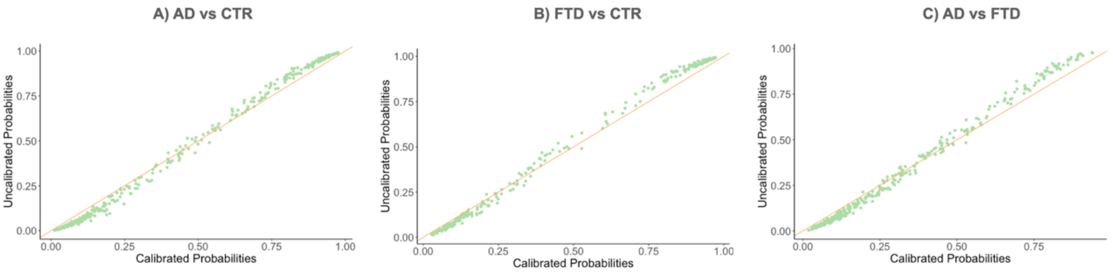

Figure 1: We plot the calibrated probabilities versus the uncalibrated probabilities. Each point represents the individual probability for each classification obtained with calibration and without calibration step. We represent the test data together.

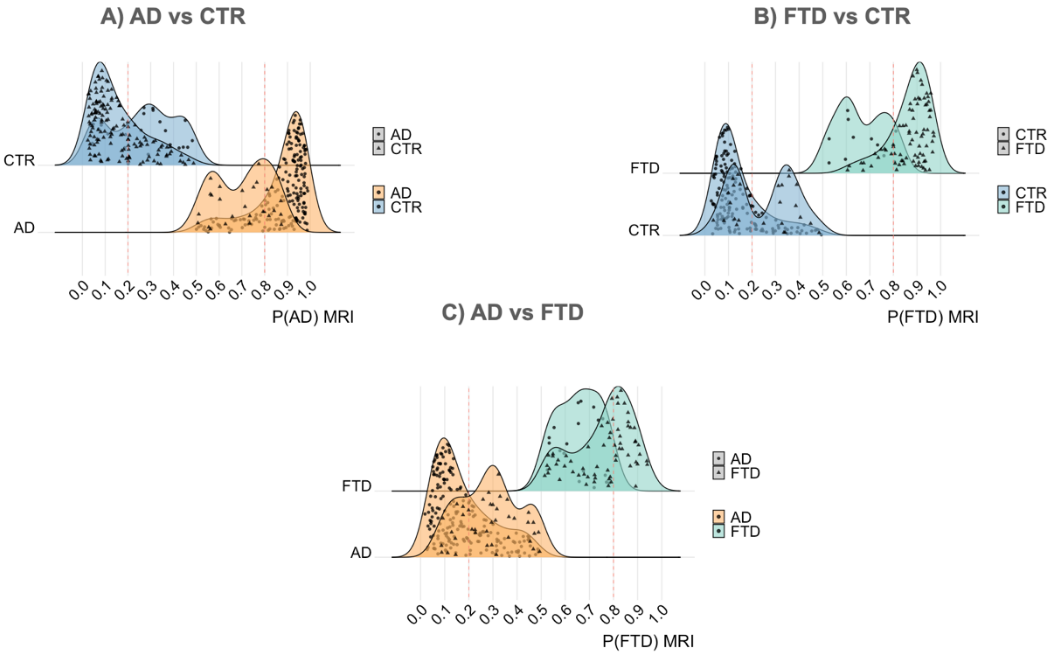

Figure 2: Density plot to study the obtained individual probabilities with the MRI-based algorithm. The clinical diagnosis is identified with triangles or circles, and the algorithm's diagnoses are plotted with different colors. CTR: healthy subjects, AD: Alzheimer's disease, FTD: frontotemporal dementia.

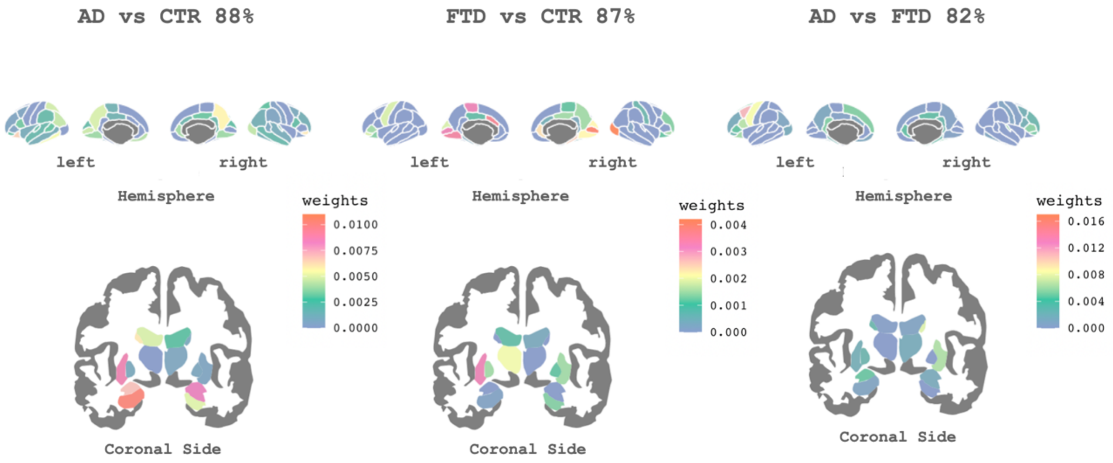
AD vs CTR 88%
FTD vs CTR 87%
AD vs FTD 82%
left
right
Hemisphere
weights
0.0100
0.0075
0.0050
0.0025
0.0000
Coronal Side
left
right
Hemisphere
weights
0.004
0.003
0.002
0.001
0.000
Coronal Side
left
right
Hemisphere
weights
0.016
0.012
0.008
0.004
0.000
Coronal Side

Figure 3: Cortical (top) and subcortical (bottom) patterns of the feature importance of each region associated with AD and FTD. At a higher value major importance of that region for the classification.

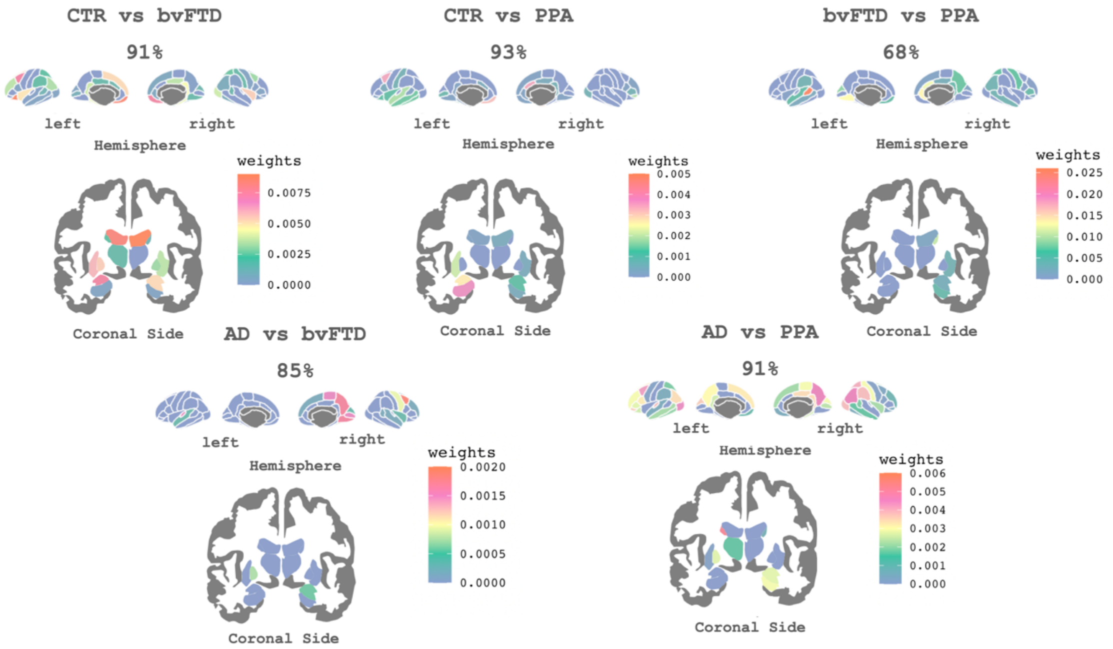
CTR vs bvFTD
91%
left
right
Hemisphere
weights
0.0075
0.0050
0.0025
0.0000
Coronal Side
CTR vs PPA
93%
left
right
Hemisphere
weights
0.005
0.004
0.003
0.002
0.001
0.000
Coronal Side
bvFTD vs PPA
68%
left
right
Hemisphere
weights
0.025
0.020
0.015
0.010
0.005
0.000
Coronal Side
AD vs bvFTD
85%
left
right
Hemisphere
weights
0.0020
0.0015
0.0010
0.0005
0.0000
Coronal Side
AD vs PPA
91%
left
right
Hemisphere
weights
0.006
0.005
0.004
0.003
0.002
0.001
0.000
Coronal Side

Figure 4: Cortical (top) and subcortical (bottom) patterns of the feature importance of each region associated with bvFTD and PPA. At a higher value major importance of that region for the classification.

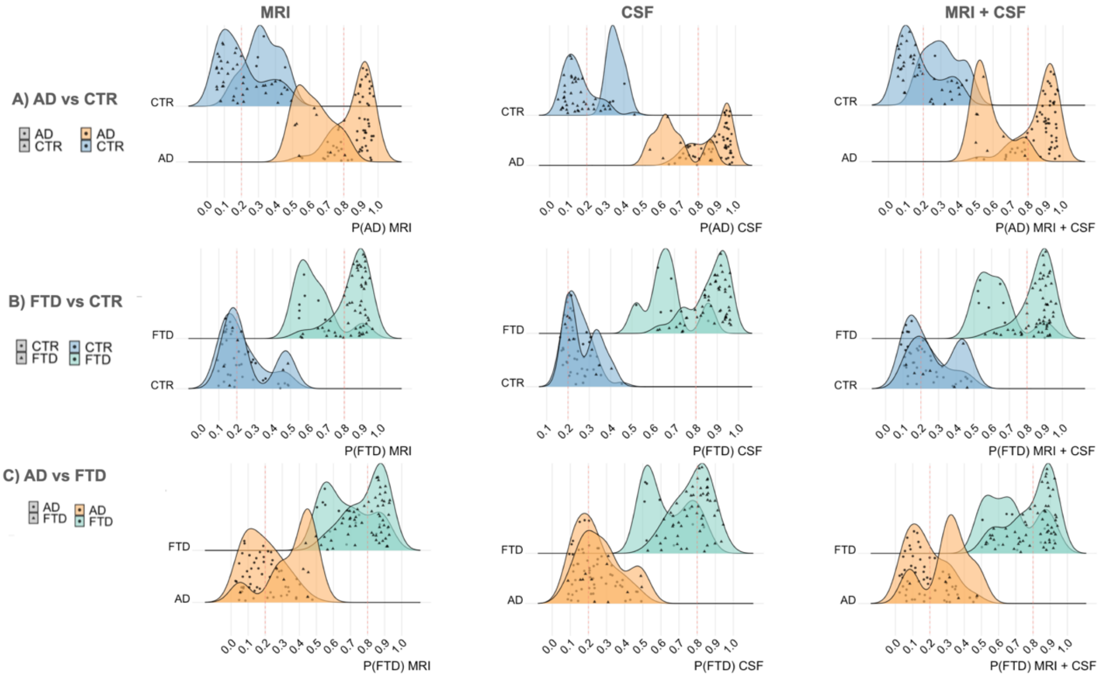

Figure 5: Density plot to study the obtained individual probabilities with the MRI-based, CSF-based, and MRI- and CSF-based algorithms. It can be seen in the clinical diagnosis with triangles or circles and the algorithm's diagnosis, plotted with different colors. CTR: healthy subjects,

Supplementary Material

MRI Features

The complete list of regions used for all the classifications.

Region
left amygdala
left bankssts
left caudal anterior cingulate
left caudal middle frontal
left caudate
left cerebellum cortex
left cuneus
left entorhinal
left frontal pole
left fusiform
left hippocampus
left inferior parietal
left inferior temporal
left insula
left isthmus cingulate
left lateral occipital
left lateral orbitofrontal
left lateral ventricle
left lingual
left medial orbitofrontal
left middle temporal
left pallidum
left paracentral
left parahippocampal
left pars opercularis
left pars orbitalis
left pars triangularis
left pericalcarine
left postcentral
left posterior cingulate
left precentral
left precuneus
left putamen

left rostral anterior cingulate
left rostral middle frontal
left superior frontal
left superior parietal
left superior temporal
left supramarginal
left temporal pole
left thalamus proper
left transverse temporal
right amygdala
right bankssts
right caudal anterior cingulate
right caudal middle frontal
right caudate
right cerebellum cortex
right cuneus
right entorhinal
right frontal pole
right fusiform
right hippocampus
right inferior parietal
right inferior temporal
right insula
right isthmus cingulate
right lateral occipital
right lateral orbitofrontal
right lateral ventricle
right lingual
right medial orbitofrontal
right middle temporal
right pallidum
right paracentral
right parahippocampal
right pars opercularis
right pars orbitalis
right pars triangularis
right pericalcarine
right postcentral
right posterior cingulate
right precentral
right precuneus
right putamen

right rostral anterior cingulate
right rostral middle frontal
right superior frontal
right superior parietal
right superior temporal
right superior temporal
right supramarginal
right thalamus proper
right transverse temporal

TREBALL 5

The Cortical Asymmetry Index (CAI) for subtyping dementia patients.

Agnès Pérez-Millan, Uma Maria Lal-Trehan Estrada, Neus Falgàs, Núria Guillén, Sergi Borrego-Écija, Beatriz Bosch, Jordi Juncà-Parella, Adrià Tort-Merino, Jordi Sarto, Josep Maria Augé, Anna Antonell, Nuria Bargalló, Raquel Ruiz-García, Laura Naranjo, Mircea Balasa, Albert Lladó, Roser Sala-Llonch, Raquel Sánchez-Valle

-En revisió-

The Cortical Asymmetry Index (CAI) for subtyping dementia patients

Agnès Pérez-Millan[1,2,3], Uma Maria Lal-Trehan Estrada[2,3], Neus Falgàs[1], Núria Guillén[1] Sergi Borrego-Écija[1], Beatriz Bosch[1], Jordi Juncà-Parella[1], Adrià Tort-Merino[1], Jordi Sarto[1], Josep Maria Augé[4], Anna Antonell[1], Nuria Bargalló[5], Raquel Ruiz-García[6], Laura Naranjo[6], Mircea Balasa[1], Albert Lladó[1,2], Roser Sala-Llonch[2,3,7*], Raquel Sánchez-Valle[1,2] *†

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

[2] Institut de Neurociències, University of Barcelona, Barcelona, 08036, Spain.

[3] Department of Biomedicine, University of Barcelona, Barcelona, 08036, Spain

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

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

[6] Immunology Service, Biomedical Diagnostic Center, Hospital Clínic de Barcelona, Barcelona, Spain.

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

† Corresponding author at Raquel Sanchez-Valle, MD, PhD. Alzheimer's disease and other cognitive disorders unit. Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), University of Barcelona Villarroel, 170 08036 Barcelona (Spain). Tel: +34 932275785

Email address: rsanchez@clinic.cat

* These authors contributed equally to this work.

Word count: Main text: 2787 words. Abstract: 224 words. Tables: 1; Figures: 5; References: 54

Keywords: asymmetry index, the Jensen-Shannon distance, MRI, Alzheimer's Disease, frontotemporal dementia, CSF, plasma, longitudinal study

ABSTRACT

Introduction: Frontotemporal dementia (FTD) patients usually show more asymmetric atrophy patterns than Alzheimer's Disease (AD). Here, we define the individual Cortical Asymmetry Index (CAI) and explore its diagnostic utility.

Methods: We collected structural T1-MRI scans from 554 participants, including FTD (all variants), AD, and healthy controls, and processed them using Freesurfer. We defined the CAI using measures based on a metric derived from information theory with the cortical thickness measures. A subset of the study participants had additional follow-up MRIs, cerebrospinal fluid (CSF), or plasma measures. We analyzed differences at cross-sectional and longitudinal levels. We then clustered FTD and AD subjects based on the CAI values and studied the fluid biomarkers characteristics within each cluster.

Results: CAI differentiated FTD, AD, and healthy controls. It also distinguished the semantic variant Primary Progressive Aphasia (svPPA) from the other FTD phenotypes. In FTD, the CAI increased over time. The cluster analysis identified two subgroups within FTD, characterized by different CSF and plasma neurofilament-light (NfL) levels, and two subgroups within AD, with different plasma Glial fibrillary acidic protein (GFAP) levels. In AD, CAI correlated with plasma-GFAP and Mini-Mental State Examination (MMSE); in FTD, the CAI was associated with NfL levels (CSF and plasma).

Conclusions: The method proposed here for the CAI at the individual level could quantify asymmetries previously described visually. The CAI could define clinically and biologically meaningful disease subgroups.

1. INTRODUCTION

Frontotemporal dementia (FTD) is a clinically, pathologically, and genetically heterogeneous neurodegenerative disorder associated with frontal and temporal atrophy. FTD patients tend to be misdiagnosed with Alzheimer's Disease (AD), especially at relatively young ages, although early behavioral and personality changes typical of FTD are not common in initial AD (Harris et al., 2015; Koedam et al., 2010; Mendez et al., 2013). However, AD is the most frequent dementia and sometimes is the first choice for many non-specialist clinicians. For this reason, there is a need to identify tools to help accurately diagnose dementia's underlying etiologies and their subtypes. Cortical asymmetry has been associated with characteristically clinical features in psychiatric and neurological conditions (Kong et al., 2018; Toga and Thompson, 2003). In this sense, compared with AD, FTD subjects show a more asymmetrical pattern at visual inspection (Gordon et al., 2016; Rabinovici and Miller, 2010; Seelaar et al., 2011; Whitwell, 2019).

Within the FTD phenotypes, the semantic variant of Primary Progressive Aphasia (svPPA) is the most asymmetric (Gordon et al., 2016; Rohrer and Rosen, 2013; Thompson et al., 2003). In this sense, the study of brain asymmetry could help in the diagnosis of the different phenotypes of FTD: behavioral variant frontotemporal dementia (bvFTD) or FTD-related primary progressive aphasia (PPA) phenotypes, including svPPA and Nonfluent Variant Primary Progressive Aphasia (nfvPPA) (Gorno-Tempini et al., 2011; Rascovsky et al., 2011). A measure of cortical asymmetry could help in the early differential diagnosis or monitoring the progress of neurodegeneration (Herzog and Magoulas, 2021; Toga and Thompson, 2003).

The differences between measures from the left and right hemispheres can be detected in the structural magnetic resonance imaging (MRI) (Guadalupe et al., 2017; Kong et al., 2018; Toga and Thompson, 2003). There is no established way to calculate the asymmetry, and different approaches have been used in multiple works on various brain diseases (Douglas et al., 2018; Herzog and Magoulas, 2021; Kalavathi et al., 2017; Kong et al., 2020; Maingault et al., 2016; Sarica et al., 2018; Schijven et al., 2023).

Then, computational algorithms can obtain measures of cortical asymmetry from the global and regional cortical thickness (CTh), surface area, or subcortical volumes. The most common method estimates the asymmetry as (L-R)/[(L+R)/2], being L for the left hemisphere measure and R for the right hemisphere measure. However, previously published works indicate that approaches using the probability distribution to obtain an asymmetric index could get better results and are currently being used in other biological areas (Estrada et al., 2022; Newton et al., 2014; Ramakrishnan and Bose, 2017).

In this study, first, we defined the Cortical Asymmetry Index (CAI) using measures derived from the information theory with CTh measures. Secondly, we aimed to study if the CAI might help in the differential diagnosis of FTD vs. AD and if it had a differential effect across the FTD clinical expressions. Then, we wanted to study cortical asymmetry changes at the longitudinal level. Finally, we explored the capability of the CAI to identify subgroups and its correlation with fluid biomarkers and cognitive measures.

2. METHODS

2.1. Participants

Participants were recruited from the Alzheimer's disease and other cognitive disorders unit of the Hospital Clínic de Barcelona (HCB), Barcelona, Spain. All participants underwent a complete clinical and cognitive evaluation and a 3T high-resolution structural MRI scan. A subset of the study participants had available follow-up visits, cerebrospinal fluid (CSF), or plasma measures.

All AD participants fulfilled the criteria for mild cognitive impairment due to AD or mild dementia due to AD (Albert et al., 2011; McKhann et al., 2011) supported by the CSF biomarkers results according to the National Institute on Aging/Alzheimer's Association Research Framework 2018 (Jack et al., 2018). We included amnesic and non-amnesic participants in the study. The FTD participants fulfilled the diagnostic criteria for bvFTD, svPPA, or nfvPPA (Gorno-Tempini et al., 2011; Rascovsky et al., 2011).

The study was approved by the HCB Ethics Committee (HCB 2019/0105), and all the participants signed the informed consent.

2.2 Biochemical markers

We used commercially available single-analyte enzyme-linked immunosorbent assay (ELISA) kits to determine levels of CSF neurofilament-light chain (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 group laboratory, Barcelona, Spain.

Plasma biomarkers concentrations were measured using single molecule array (SIMOA), Quanterix Neurology 4-Plex A including total tau (t-tau), glial fibrillary acidic protein (GFAP), NfL, and Ubiquitin C-terminal hydrolase L1 (UCH-L1) and the Quanterix p-tau181 Advantage V2 and V2.1 assays following the manufacturer's protocol (Quanterix, Billerica, MA), we harmonized the values of the two kits.

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 generate automated CTh maps; steps are reported in detail elsewhere (Fischl et al., 2004; Fischl and Dale, 2000). This study used the mean CTh measures in 68 cortical parcellations (34 per hemisphere) derived from atlases available in FreeSurfer (Desikan et al., 2006). All images and individual segmentations were visually inspected and manually corrected if needed.

2.5 Cortical Asymmetry Index (CAI)

We obtained the CAI by calculating the Jensen-Shannon distance (JSD), a methodology based on a metric derived from information theory (Lin, 1991). The JSD measures the difference between two probability distributions (P1 and P2) using the square root of the Jensen-Shannon divergence, which is based on the Kullback-Leibler divergence (KLD). The KLD is also referred to as relative entropy between two distributions. It is calculated as the negative sum of the probability of each event in P1 multiplied by the logarithm of the probability of the same event in P2 divided by the probability of the event in P1. The JSD is the symmetric and normalized version of KLD, so the JSD from one probability P1 to another probability P2 is the same as the JSD of P2 from P1. We obtained the CAI using the distribution of CTh measures of each brain hemisphere (previously obtained with the software FreeSurfer) that represents the difference in CTh distribution between brain hemispheres. Thus, the CAI allows us to quantify brain structural asymmetry at the individual level. CAI is a non-dimensional measure with higher values indicating a more asymmetric brain.

The CAI was calculated using in-house methods implemented in Python version 3.10.6 (www.python.org).

2.6 Statistical Analysis

We first compared the CAI between all groups, including healthy controls, AD, and FTD patients. Then, we performed comparisons across the different FTD phenotypes. In both cases, we used the ANOVA test, corrected by age when needed, and a post hoc analysis with the Tukey test to assess multiple comparisons. Finally, we performed longitudinal analyses with Linear Mixed Effects Models (LME) to study changes between visits; we added age at MRI, sex, and age of onset (for AD and FTD) as fixed effects.

We then analyzed the CAI using cluster analyses for FTD and AD participants separately to evaluate whether subgroups emerged within each disease. We used hierarchical clustering, precisely agglomerative nesting cluster with Ward's method and Manhattan metric. We determined the optimal number of clusters in each case with the Silhouette strategy. Then, with the cluster groups obtained, we studied differences in fluid biomarkers across subgroups with non-parametric tests such as the Kruskal Wallis or Wilcoxon tests. All biomarkers' levels were converted to z-score before these calculations.

Finally, we evaluated the association between CAI and the CSF and plasma biomarkers that emerged in the cluster analysis. In addition, we studied the correlation of CAI with age and Mini-Mental State Examination (MMSE). We used Spearman's rank correlation coefficient due to the limited samples.

We implemented the statistical analysis and the cluster analysis (with cluster and factoextra packages) in R version 4.2.1

3. RESULTS

3.1. Demographics

We included 554 participants: 230 AD (30 with follow-up MRI scans with two years approximately between scans), 101 FTD (30 with follow-up MRI scans with one year and a half approximately between scans), and 173 healthy controls (96 with follow-up MRI scans with two years approximately between scans). The 101 FTD participants were distributed as follows: 55 bvFTD (14 follow-up MRI scans), 21 nfvPPA (7 follow-up MRI scans), and 24 svPPA (9 follow-up MRI scans). AD participants were 64% women, and the mean age was 65.3 ± 9.7; FTD participants were 48% women, the mean age was 63.7 ± 8.4, and healthy control participants were 61% women, and the mean age was 59.4 ± 15.0. Healthy controls were younger than AD and FTD participants ($p<0.05$). As expected, CSF and plasma biomarkers levels showed significant differences between groups, except for plasma t-tau and UCH-L1 (Table 1).

3.2. Cross-sectional CAI differences

AD participants had a CAI of 0.40 ± 0.28, FTD participants 0.59 ± 0.52, and healthy controls 0.16 ± 0.10. We found higher CAI in AD and FTD patients than healthy controls (adjusted p-value < 0.0001), meaning a more asymmetric brain structure. FTD participants also presented a higher CAI than AD patients (adjusted p-value<0.0001) (Figure 1).

Then, we studied the CAI across FTD clinical phenotypes. The bvFTD patients had a mean CAI of 0.46 ± 0.46, nfvPPA patients 0.62 ± 0.62, and svPPA patients 0.83 ± 0.48. When studying the differences between FTD clinical phenotypes, svPPA presented higher CAI values than bvFTD (adjusted p-value<0.01) (Figure 2).

3.3. Longitudinal changes in CAI

The mean CAI for the follow-up visits was 0.42 ± 0.31 for AD patients, 0.80 ± 0.77 for FTD patients, and 0.16 ± 0.12 for healthy controls. At the longitudinal level (Figure 3), we found that CAI significantly increased in FTD, indicating a significantly more asymmetric brain over time (p-value=0.012 between baseline and follow-up, measured as the interaction with age). We did not find changes between visits in AD patients and healthy controls.

3.4. Cluster Group

The cluster analysis was used to identify subgroups based on their CAI. Then, we compared biomarker values between clusters. This method could divide the FTD group into Cluster 1 (N=41) and Cluster 2 (N=58). Cluster 1 presented a more asymmetric brain with 62.5% svPPA, 35.2% bvFTD, and 33.3% nfvPPA of the total cohort participants. In contrast, Cluster 2 showed a less asymmetric brain and grouped as 37.5% svPPA, 64.8% bvFTD, and 66.7% nfvPPA of the total cohort participants. CSF and plasma levels of NfL were increased in Cluster 1 compared to Cluster 2 (p-value=0.0059 and p-value=0.042, respectively) (Figure 4a).

Within AD participants, the cluster analysis also yielded two groups, presenting differences in plasma GFAP levels (p-value=0.035) (Figure 4b). Then, we tried to find an explanation based on the clinical profile to explain the clusters. However, we did not find any difference between clusters at a clinical level. The distribution of the participants in the two clusters was not grouped due to sex, age of onset (early-onset AD vs. late-onset AD), age of disease duration, AD phenotypes (amnesic vs. non-amnesic), MMSE, APOE status, or the combination of these variables.

3.5. Relation between CAI, fluid biomarkers, and global cognition measures

We assessed whether CAI values correlated with the levels of the fluid biomarkers that emerged in the cluster analysis, in addition to MMSE and age (Figure 5). We found a weak positive correlation in AD for plasma-GFAP ($r = 0.20$, p-value < 0.05) and a weak negative correlation with MMSE ($r = -0.26$, p-value < 0.001). For FTD patients, we found a significant positive correlation of CAI values with NfL levels; both CSF-NfL ($r = 0.27$, p-value < 0.05) and plasma-NfL ($r = 0.41$, p-value < 0.05). We found no correlation between CAI and fluid biomarkers in the different FTD clinical phenotypes. Healthy controls did not present any correlation between the CAI and age either. Age did not correlate with CAI in any of the groups.

4. DISCUSSION

In this study, we propose a methodology to estimate the brain structural asymmetries, namely the CAI, for neurodegenerative diseases, especially FTD and AD. We used an information theory method to calculate each individual's CAI. CAI mean values differentiate healthy controls from dementia participants, and FTD from AD, and between FTD phenotypes. We studied the brain asymmetries over time and found an increase in asymmetry for FTD. Cluster analysis using the CAI values for both FTD and AD enabled us to identify disease subgroups and study the relationship between brain asymmetry and the levels of fluid biomarkers.

Our first finding, the increased asymmetry for FTD, aligns with what has been described in the literature using a mix of non-standardized measures and visual inspection. The svPPA is the most asymmetrical clinical phenotype of FTD, usually showing greater left than right atrophy of the temporal lobes (Bruun et al., 2019;

Gordon et al., 2016; Gorno-Tempini et al., 2011, 2004; Mesulam et al., 2014; Rohrer and Rosen, 2013; Schroeter et al., 2007; Seelaar et al., 2011; Whitwell et al., 2015). Also, in agreement with our findings, the FTD brain asymmetry has been reported to increase over time (Rohrer and Rosen, 2013; Whitwell et al., 2015). In contrast, AD is commonly described as symmetric dementia: the atrophy presented in one hemisphere of the brain, usually, is similar in the other. However, some studies reported asymmetric patterns in AD (Derflinger et al., 2011; Low et al., 2019; Mesulam et al., 2014; Roe et al., 2021; Rombouts et al., 2000). In our study, AD patients present lower CAI when compared with FTD patients, meaning less asymmetry. However, if compared to healthy controls, they present an asymmetrical brain.

Previous works studying brain asymmetries in FTD and AD usually report these asymmetries at the visual level. Still, further studies are needed that present a methodology to evaluate and quantify this phenomenon systematically. Although some studies have been done to quantify asymmetries (Boccardi et al., 2002; Bocchetta et al., 2018; Bruun et al., 2019; Low et al., 2019; Schroeter et al., 2007), we chose a more complex methodology based on information theory to obtain the estimation of an asymmetric index. For this reason, we proposed to estimate the asymmetric index with the JSD measure. This novel measure based on entropy has recently been used to study similarities in different biological and clinical areas. For example, it has been used to help in the detection of ischemic stroke, to study long-term surgical outcomes in the brain for epilepsy patients, to study cancer cells to find link cells with near-identical gene expression, or to study the individual metabolic network in patients with type 2 diabetes (Estrada et al., 2022; La Manno et al., 2021; Li et al., 2022; Zhu et al., 2022). We demonstrated that the CAI defined here could be of help for the differentiation of the clinical expressions or dementias, to study the progression of the disease, or to identify subgroups.

Other studies have explored quantification strategies for brain symmetries to differentiate between different dementias successfully (Bruun et al., 2019; Kitagaki et al., 1998). However, there is currently no standard and accurate methodology for this, and the CAI presented here showed promising results. Also, it was helpful in identifying FTD clinical phenotypes. For example, svPPA presented different levels compared to bvFTD. We found that the svPPA is the most asymmetric clinical expression of FTD, which is in accordance with previous literature, using visual scales (Seelaar et al., 2011; Whitwell, 2019). We replicated the previous results in this study and quantified these differences using the CAI.

Our cluster analysis yielded FTD and AD participants into two subgroups according to their asymmetry indexes. As cluster analysis is a no supervised statistical approach, we aimed to find an explanation using clinical and biomarker data. Among the two subgroups for FTD patients, the more asymmetric cluster was enriched with svPPA participants and the other cluster with nfvPPA and bvFTD. The cluster analysis was in accordance with the differences studied between FTD clinical phenotypes using ANOVAs, where we found a significant difference between svPPA and bvFTD. Overall, we observed that svPPA presented the highest CAI compared to nfvPPA and bvFTD. AD participants were also subdivided into two subgroups; however, we could not find an explanation based on the clinical profile of the subjects. We analyzed its association with fluid biomarkers to study further implications of cortical brain asymmetry. We found that in the FTD Cluster 1, CAI values were associated with higher levels of NfL in both CSF and plasma with respect to the bvFTD and nfvPPA and that higher CAI predicted higher NfL (CSF and plasma) in FTD patients. This suggests that NfL (CSF and plasma) is directly associated with brain asymmetry. Previous studies reported that NfL levels in CSF and plasma were associated with brain atrophy (Illán-Gala et al., 2021; Meeter et al., 2016; Rohrer et al., 2016). However, the association between NfL and brain asymmetries was not investigated. When studying AD, we found that different CAI groups were associated with plasma GFAP levels. The AD-less asymmetric cluster presented higher GFAP levels. This biomarker has been previously associated with brain atrophy due to aging or disease severity (Ebenau et al., 2022; Traub et al., 2022). However, its association with brain asymmetry has not been studied before. Then, we examined the correlation between GFAP levels and CAI in AD patients and obtained a positive correlation. Overall, these associations between brain asymmetries and fluid biomarkers suggest that both contribute to defining AD subgroups.

Finally, FTD participants presented higher levels of brain asymmetry over time, suggesting that the CAI could indicate FTD progression. Previous studies have shown that FTD's different genetic or clinical expressions behave differently in becoming more asymmetric over follow-up visits (Mahoney et al., 2012; Meeter et al., 2017; Whitwell et al., 2012). Regarding AD, no differences in CAI over time were found for AD patients.

In conclusion, the proposed CAI measure might identify differences between AD and FTD and between FTD phenotypes. It might be a sensitive measure to determine the disease progression of the FTD participants. Cortical asymmetry drives subgrouping inside FTD and AD, associated with various fluid biomarker levels. Overall, this highlights the potential relevance of quantifying cortical asymmetry. Further studies deepening into the genetic expressions of FTD or investigating the predictive value of the CAI towards clinical progression might be of interest.

REFERENCES

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

Boccardi, M., Laakso, M.P., Bresciani, L., Geroldi, C., Beltramello, A., Frisoni, G.B., 2002. Clinical characteristics of frontotemporal patients with symmetric brain atrophy. European Archives of Psychiatry and Clinical Neurosciences 252, 235–239. https://doi.org/10.1007/s00406-002-0388-z

Bocchetta, M., Iglesias, J.E., Scelsi, M.A., Cash, D.M., Cardoso, M.J., Modat, M., Altmann, A., Ourselin, S., Warren, J.D., Rohrer, J.D., 2018. Hippocampal Subfield Volumetry: Differential Pattern of Atrophy in Different Forms of Genetic Frontotemporal Dementia. Journal of Alzheimer's Disease 64, 497–504. https://doi.org/10.3233/JAD-180195

Bruun, M., Koikkalainen, J., Rhodius-Meester, H.F.M., Baroni, M., Gjerum, L., van Gils, M., Soininen, H., Remes, A.M., Hartikainen, P., Waldemar, G., Mecocci, P., Barkhof, F., Pijnenburg, Y., van der Flier, W.M., Hasselbalch, S.G., Lötjönen, J., Frederiksen, K.S., 2019. Detecting frontotemporal dementia syndromes using MRI biomarkers. NeuroImage: Clinical 22, 101711. https://doi.org/10.1016/j.nicl.2019.101711

Derflinger, S., Sorg, C., Gaser, C., Myers, N., Arsic, M., Kurz, A., Zimmer, C., Wohlschläger, A., Mühlau, M., 2011. Grey-Matter Atrophy in Alzheimer's Disease is Asymmetric but not Lateralized. Journal of Alzheimer's Disease 25, 347–357. https://doi.org/10.3233/JAD-2011-110041

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Douglas, P.K., Gutman, B., Anderson, A., Larios, C., Lawrence, K.E., Narr, K., Sengupta, B., Cooray, G., Douglas, D.B., Thompson, P.M., McGough, J.J., Bookheimer, S.Y., 2018. Hemispheric brain asymmetry differences in youths with attention-deficit/hyperactivity disorder. NeuroImage: Clinical 18, 744–752. https://doi.org/10.1016/j.nicl.2018.02.020

Ebenau, J.L., Pelkmans, W., Verberk, I.M.W., Verfaillie, S.C.J., Bosch, K.A. van den, Leeuwenstijn, M. van, Collij, L.E., Scheltens, P., Prins, N.D., Barkhof, F., Berckel, B.N.M. van, Teunissen, C.E., Flier, W.M. van der, 2022. Association of CSF, Plasma, and Imaging Markers of Neurodegeneration With Clinical Progression in People With Subjective Cognitive Decline. Neurology 98, e1315–e1326. https://doi.org/10.1212/WNL.0000000000200035

Estrada, U.M.L.-T., Meeks, G., Salazar-Marioni, S., Scalzo, F., Farooqui, M., Vivanco-Suarez, J., Gutierrez, S.O., Sheth, S.A., Giancardo, L., 2022. Quantification of infarct core signal using CT imaging in acute ischemic stroke. NeuroImage: Clinical 34, 102998. https://doi.org/10.1016/j.nicl.2022.102998

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of the United States of America 97, 11050–11055. https://doi.org/10.1073/pnas.200033797

Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex 14, 11–22. https://doi.org/10.1093/cercor/bhg087

Gordon, E., Rohrer, J.D., Fox, N.C., 2016. Advances in neuroimaging in frontotemporal dementia. Journal of Neurochemistry 138, 193–210. https://doi.org/10.1111/jnc.13656

Gorno-Tempini, M.L., Dronkers, N.F., Rankin, K.P., Ogar, J.M., Phengrasamy, L., Rosen, H.J., Johnson, J.K., Weiner, M.W., Miller, B.L., 2004. Cognition and anatomy in three variants of primary progressive aphasia. Annals of Neurology 55, 335–346. https://doi.org/10.1002/ana.10825

Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S.F., Ogar, J.M., Rohrer, J.D., Black, S., Boeve, B.F., Manes, F., Dronkers, N.F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R., Mesulam, M.M., Grossman, M., Rascovsky, P.K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R., Mesulam, M.M., Grossman, M., Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R., 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

Guadalupe, T., Mathias, S.R., vanErp, T.G.M., Whelan, C.D., Zwiers, M.P., Abe, Y., Abramovic, L., Agartz, I., Andreassen, O.A., Arias-Vásquez, A., Aribisala, B.S., Armstrong, N.J., Arolt, V., Artiges, E., Ayesa-Arriola, R., Baboyan, V.G., Banaschewski, T., Barker, G., Bastin, M.E., Baune, B.T., Blangero, J., Bokde, A.L.W., Boedhoe, P.S.W., Bose, A., Brem, S., Brodaty, H., Bromberg, U., Brooks, S., Büchel, C., Buitelaar, J., Calhoun, V.D., Cannon, D.M., Cattrell, A., Cheng, Y., Conrod, P.J., Conzelmann, A., Corvin, A., Crespo-Facorro, B., Crivello, F., Dannlowski, U., de Zubicaray, G.I., de Zwarte, S.M.C., Deary, I.J., Desrivières, S., Doan, N.T., Donohoe, G., Dørum, E.S., Ehrlich, S., Espeseth, T., Fernández, G., Flor, H., Fouche, J.-P., Frouin, V., Fukunaga, M., Gallinat, J., Garavan, H., Gill, M., Suarez, A.G., Gowland, P., Grabe, H.J., Grotegerd, D., Gruber, O., Hagenaars, S., Hashimoto, R., Hauser, T.U., Heinz, A., Hibar, D.P., Hoekstra, P.J., Hoogman, M., Howells, F.M., Hu, H., Hulshoff Pol, H.E., Huyser, C., Ittermann, B., Jahanshad, N., Jönsson, E.G., Jurk, S., Kahn, R.S., Kelly, S., Kraemer, B., Kugel, H., Kwon, J.S., Lemaitre, H., Lesch, K.-P., Lochner, C., Luciano, M., Marquand, A.F., Martin, N.G., Martínez-Zalacaín, I., Martinot, J.-L., Mataix-Cols, D., Mather, K., McDonald, C., McMahon, K.L., Medland, S.E., Menchón, J.M., Morris, D.W., Mothersill, O., Maniega, S.M., Mwangi, B., Nakamae, T., Nakao, T., Narayanaswaamy, J.C., Nees, F., Nordvik, J.E., Onnink, A.M.H., Opel, N., Ophoff, R., Paillère Martinot, M.-L., Papadopoulos Orfanos, D., Pauli, P., Paus, T., Poustka, L., Reddy, J.YC., Renteria, M.E., Roiz-Santiáñez, R., Roos, A., Royle, N.A., Sachdev, P., Sánchez-Juan, P., Schmaal, L., Schumann, G., Shumskaya, E., Smolka, M.N., Soares, J.C., Soriano-Mas, C., Stein, D.J., Strike, L.T., Toro, R., Turner, J.A., Tzourio-Mazoyer, N., Uhlmann, A., Hernández, M.V., van den Heuvel, O.A., van der Meer, D., van Haren, N.E.M.. ., Veltman, D.J., Venkatasubramanian, G., Vetter, N.C., Vuletic, D., Walitza, S., Walter, H., Walton, E., Wang, Z., Wardlaw, J., Wen, W., Westlye, L.T., Whelan, R., Wittfeld, K., Wolfers, T., Wright, M.J., Xu, J., Xu, X., Yun, J.-Y., Zhao, J., Franke, B., Thompson, P.M., Glahn, D.C., Mazoyer, B., Fisher, S.E., Francks, C., 2017. Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. Brain Imaging and Behavior 11, 1497–1514. https://doi.org/10.1007/s11682-016-9629-z

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

Herzog, N.J., Magoulas, G.D., 2021. Brain Asymmetry Detection and Machine Learning Classification for Diagnosis of Early Dementia. Sensors (Basel) 21, 778. https://doi.org/10.3390/s21030778

Illán-Gala, I., Lleo, A., Karydas, A., Staffaroni, A.M., Zetterberg, H., Sivasankaran, R., Grinberg, L.T., Spina, S., Kramer, J.H., Ramos, E.M., Coppola, G., La Joie, R., Rabinovici, G.D., Perry, D.C., Gorno-Tempini, M.L., Seeley, W.W., Miller, B.L., Rosen, H.J., Blennow, K., Boxer, A.L., Rojas, J.C., 2021. Plasma Tau and Neurofilament Light in Frontotemporal Lobar Degeneration and Alzheimer Disease. Neurology 96, e671–e683. https://doi.org/10.1212/WNL.0000000000011226

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

Kalavathi, P., Senthamilselvi, M., Prasath, V.B.S., 2017. Review of Computational Methods on Brain Symmetric and Asymmetric Analysis from Neuroimaging Techniques. Technologies 5, 16. https://doi.org/10.3390/technologies5020016

Kitagaki, H., Mori, E., Yamaji, S., Ishii, K., Hirono, N., Kobashi, S., Hata, Y., 1998. Frontotemporal dementia and Alzheimer disease: evaluation of cortical atrophy with automated hemispheric surface display generated with MR images. Radiology 208, 431–439. https://doi.org/10.1148/radiology.208.2.9680572

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

Kong, X.-Z., Boedhoe, P.S.W., Abe, Y., Alonso, P., Ameis, S.H., Arnold, P.D., Assogna, F., Baker, J.T., Batistuzzo, M.C., Benedetti, F., Beucke, J.C., Bollettini, I., Bose, A., Brem, S., Brennan, B.P., Buitelaar, J., Calvo, R., Cheng, Y., Cho, K.I.K., Dallaspezia, S., Denys, D., Ely, B.A., Feusner, J., Fitzgerald, K.D., Fouche, J.-P., Fridgeirsson, E.A., Glahn, D.C., Gruner, P., Gürsel, D.A., Hauser, T.U., Hirano, Y., Hoexter, M.Q., Hu, H., Huyser, C., James, A., Jaspers-Fayer, F., Kathmann, N., Kaufmann, C., Koch, K., Kuno, M., Kvale, G., Kwon, J.S., Lazaro, L., Liu, Y., Lochner, C., Marques, P., Marsh, R., Martínez-Zalacaín, I., Mataix-Cols, D., Medland, S.E., Menchón, J.M., Minuzzi, L., Moreira, P.S., Morer, A., Morgado, P., Nakagawa, A., Nakamae, T., Nakao, T., Narayanaswamy, J.C., Nurmi, E.L., O'Neill, J., Pariente, J.C., Perriello, C., Piacentini, J., Piras, Fabrizio, Piras, Federica, Pittenger, C., Reddy, Y.C.J., Rus-Oswald, O.G., Sakai, Y., Sato, J.R., Schmaal, L., Simpson, H.B., Soreni, N., Soriano-Mas, C., Spalletta, G., Stern, E.R., Stevens, M.C., Stewart, S.E., Szeszko, P.R., Tolin, D.F., Tsuchiyagaito, A., van Rooij, D., van Wingen, G.A., Venkatasubramanian, G., Wang, Z., Yun, J.-Y., Kong, X.-Z., Boedhoe, P.S.W., Abe, Y., Alonso, P., Ameis, S.H., Anticevic, A., Arnold, P.D., Assogna, F., Baker, J.T., Banaj, N., Bargalló, N., Batistuzzo, M.C., Benedetti, F., Beucke, J.C., Bollettini, I., Bose, A., Brandeis, D., Brem, S., Brennan, B.P., Buitelaar, J., Busatto, G.F., Calvo, A., Calvo, R., Cheng, Y., Cho, K.I.K., Ciullo, V., Dallaspezia, S., Denys, D., de Vries, F.E., de Wit, S.J., Dickie, E., Drechsler, R., Ely, B.A., Esteves, M., Falini, A., Fang, Y., Feusner, J., Figee, M., Fitzgerald, K.D., Fontaine, M., Fouche, J.-P., Fridgeirsson, E.A., Gruner, P., Gürsel, D.A., Hall, G., Hamatani, S., Hanna, G.L., Hansen, B., Hauser, T.U., Hirano, Y., Hoexter, M.Q., Hu, H., Huyser, C., Ikari, K., Jahanshad, N., James, A., Jaspers-Fayer, F., Kathmann, N., Kaufmann, C., Koch, K., Kuno, M., Kvale, G., Kwon, J.S., Lazaro, L., Liu, Y., Lochner, C., Magalhães, R., Marques, P., Marsh, R., Martínez-Zalacaín, I., Masuda, Y., Mataix-Cols, D., Matsumoto, K., McCracken, J.T., Menchón, J.M., Miguel, E.C., Minuzzi, L., Moreira, P.S., Morer, A., Morgado, P., Nakagawa, A., Nakamae, T., Nakao, T., Narayanaswamy, J.C., Narumoto, J., Nishida, S., Nurmi, E.L., O'Neill, J., Pariente, J.C., Perriello, C., Piacentini, J., Piras, Fabrizio, Piras, Federica, Pittenger, C., Poletti, S., Reddy, Y.C.J., Reess, T., Rus-Oswald, O.G., Sakai, Y., Sato, J.R., Schmaal, L., Shimizu, E., Simpson, H.B., Soreni, N., Soriano-Mas, C., Sousa, N., Spalletta, G., Stern, E.R., Stevens, M.C., Stewart, S.E., Szeszko, P.R., Takahashi, J., Tang, J., Thorsen, A.L., Tolin, D.F., Tsuchiyagaito, A., van Rooij, D., van Wingen, G.A., van der Werf, Y.D., Veltman, D.J., Vecchio, D., Venkatasubramanian, G., Walitza, S., Wang, Z., Watanabe, A., Xu, J., Xu, X., Yamada, K., Yoshida, T., Yun, J.-Y., Zarei, M., Zhao, Q., Zhou, C., Thompson, P.M., Stein, D.J., van den Heuvel, O.A., Francks, C., Thompson, P.M., Stein, D.J., van den Heuvel, O.A., Francks, C., 2020. Mapping Cortical and Subcortical Asymmetry in Obsessive-Compulsive Disorder: Findings From the ENIGMA Consortium. Biological Psychiatry, Obsessive-Compulsive Disorder and Developmental Disorders 87, 1022–1034. https://doi.org/10.1016/j.biopsych.2019.04.022

Kong, X.-Z., Mathias, S.R., Guadalupe, T., ENIGMA Laterality Working Group, Glahn, D.C., Franke, B., Crivello, F., Tzourio-Mazoyer, N., Fisher, S.E., Thompson, P.M., Francks, C., 2018. Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. Proceedings of the National Academy of Sciences 115, E5154–E5163. https://doi.org/10.1073/pnas.1718418115

La Manno, G., Siletti, K., Furlan, A., Gyllborg, D., Vinsland, E., Mossi Albiach, A., Mattsson Langseth, C., Khven, I., Lederer, A.R., Dratva, L.M., Johnsson, A., Nilsson, M., Lönnerberg, P., Linnarsson, S., 2021. Molecular architecture of the developing mouse brain. Nature 596, 92–96. https://doi.org/10.1038/s41586-021-03775-x

Li, Y.-L., Wu, J.-J., Ma, J., Li, S.-S., Xue, X., Wei, D., Shan, C.-L., Hua, X.-Y., Zheng, M.-X., Xu, J.-G., 2022. Alteration of the Individual Metabolic Network of the Brain Based on Jensen-Shannon Divergence Similarity Estimation in Elderly Patients With Type 2 Diabetes Mellitus. Diabetes 71, 894–905. https://doi.org/10.2337/db21-0600

Lin, J., 1991. Divergence measures based on the Shannon entropy. IEEE Transactions on Information Theory 37, 145–151. https://doi.org/10.1109/18.61115

Low, A., Mak, E., Malpetti, M., Chouliaras, L., Nicastro, N., Su, L., Holland, N., Rittman, T., Rodríguez, P.V., Passamonti, L., Bevan-Jones, W.R., Jones, P.S., Rowe, J.B., O'Brien, J.T., 2019. Asymmetrical atrophy of thalamic subnuclei in Alzheimer's disease and amyloid-positive mild cognitive impairment is associated with key clinical features. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 11, 690–699. https://doi.org/10.1016/j.dadm.2019.08.001

Mahoney, C.J., Downey, L.E., Ridgway, G.R., Beck, J., Clegg, S., Blair, M., Finnegan, S., Leung, K.K., Yeatman, T., Golden, H., Mead, S., Rohrer, J.D., Fox, N.C., Warren, J.D., 2012. Longitudinal neuroimaging and neuropsychological profiles of frontotemporal dementia with C9ORF72 expansions. Alz Res Therapy 4, 41. https://doi.org/10.1186/alzrt144

Maingault, S., Tzourio-Mazoyer, N., Mazoyer, B., Crivello, F., 2016. Regional correlations between cortical thickness and surface area asymmetries: A surface-based morphometry study of 250 adults. Neuropsychologia, The Neural Bases of Hemispheric Specialisation 93, 350–364. https://doi.org/10.1016/j.neuropsychologia.2016.03.025

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

Meeter, L.H., Dopper, E.G., Jiskoot, L.C., Sanchez-Valle, R., Graff, C., Benussi, L., Ghidoni, R., Pijnenburg, Y.A., Borroni, B., Galimberti, D., Laforce, R.J., Masellis, M., Vandenberghe, R., Ber, I.L., Otto, M., van Minkelen, R., Papma, J.M., Rombouts, S.A., Balasa, M., Öijerstedt, L., Jelic, V., Dick, K.M., Cash, D.M., Harding, S.R., Jorge Cardoso, M., Ourselin, S., Rossor, M.N., Padovani, A., Scarpini, E., Fenoglio, C., Tartaglia, M.C., Lamari, F., Barro, C., Kuhle, J., Rohrer, J.D., Teunissen, C.E., van Swieten, J.C., 2016. Neurofilament light chain: a biomarker for genetic frontotemporal dementia. Annals of Clinical and Translational Neurology 3, 623–636. https://doi.org/10.1002/acn3.325

Meeter, L.H., Kaat, L.D., Rohrer, J.D., van Swieten, J.C., 2017. Imaging and fluid biomarkers in frontotemporal dementia. Nat Rev Neurol 13, 406–419. https://doi.org/10.1038/nrneurol.2017.75

Mendez, M.F., Joshi, A., Tassniyom, K., Teng, E., Shapira, J.S., 2013. Clinicopathologic differences among patients with behavioral variant frontotemporal dementia. Neurology 80, 561–568. https://doi.org/10.1212/WNL.0b013e3182815547

Mesulam, M.-M., Weintraub, S., Rogalski, E.J., Wieneke, C., Geula, C., Bigio, E.H., 2014. Asymmetry and heterogeneity of Alzheimer's and frontotemporal pathology in primary progressive aphasia. Brain 137, 1176–1192. https://doi.org/10.1093/brain/awu024

Newton, P.K., Mason, J., Hurt, B., Bethel, K., Bazhenova, L., Nieva, J., Kuhn, P., 2014. Entropy, complexity, and Markov diagrams for random walk cancer models. Sci Rep 4, 7558. https://doi.org/10.1038/srep07558

Rabinovici, G.D., Miller, B.L., 2010. Frontotemporal Lobar Degeneration. CNS Drugs 24, 375–398. https://doi.org/10.2165/11533100-000000000-00000

Ramakrishnan, N., Bose, R., 2017. Analysis of healthy and tumour DNA methylation distributions in kidney-renal-clear-cell-carcinoma using Kullback–Leibler and Jensen–Shannon distance measures. IET Syst Biol 11, 99–104. https://doi.org/10.1049/iet-syb.2016.0052

Rascovsky, K., Hodges, J.R., Knopman, D., Mendez, M.F., Kramer, J.H., Neuhaus, J., van Swieten, J.C., Seelaar, H., Dopper, E.G.P., Onyike, C.U., Hillis, A.E., Josephs, K. a., Boeve, B.F., Kertesz, A., Seeley, W.W., Rankin, K.P., Johnson, J.K., Gorno-Tempini, M.-L.L., Rosen, H., Prioleau-Latham, C.E., Lee, A., Kipps, C.M., Lillo, P., Piguet, O., Rohrer, J.D., Rossor, M.N., Warren, J.D., Fox, N.C., Galasko, D., Salmon, D.P., Black, S.E., Mesulam, M., Weintraub, S., Dickerson, B.C., Diehl-Schmid, J., Pasquier, F., Deramecourt, V., Lebert, F., Pijnenburg, Y., Chow, T.W., Manes, F., Grafman, J., Cappa, S.F., Freedman, M., Grossman, M., Miller, B.L., 2011. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 134, 2456–2477. https://doi.org/10.1093/brain/awr179

Roe, J.M., Vidal-Piñeiro, D., Sørensen, Ø., Brandmaier, A.M., Düzel, S., Gonzalez, H.A., Kievit, R.A., Knights, E., Kühn, S., Lindenberger, U., Mowinckel, A.M., Nyberg, L., Park, D.C., Pudas, S., Rundle, M.M., Walhovd, K.B., Fjell, A.M., Westerhausen, R., 2021. Asymmetric thinning of the cerebral cortex across the adult lifespan is accelerated in Alzheimer's disease. Nat Commun 12, 721. https://doi.org/10.1038/s41467-021-21057-y

Rohrer, J.D., Rosen, H.J., 2013. Neuroimaging in frontotemporal dementia. International Review of Psychiatry 25, 221–229. https://doi.org/10.3109/09540261.2013.778822

Rohrer, J.D., Woollacott, I.O.C., Dick, K.M., Brotherhood, E., Gordon, E., Fellows, A., Toombs, J., Druyeh, R., Cardoso, M.J., Ourselin, S., Nicholas, J.M., Norgren, N., Mead, S., Andreasson, U., Blennow, K., Schott, J.M., Fox, N.C., Warren, J.D., Zetterberg, H., 2016. Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. Neurology 87, 1329–1336. https://doi.org/10.1212/WNL.0000000000003154

Rombouts, S.A.R.B., Barkhof, F., Witter, M.P., Scheltens, P., 2000. Unbiased whole-brain analysis of gray matter loss in Alzheimer's disease. Neuroscience Letters 285, 231–233. https://doi.org/10.1016/S0304-3940(00)01067-3

Sarica, A., Vasta, R., Novellino, F., Vaccaro, M.G., Cerasa, A., Quattrone, A., The Alzheimer's Disease Neuroimaging Initiative, 2018. MRI Asymmetry Index of Hippocampal Subfields Increases Through the Continuum From the Mild Cognitive Impairment to the Alzheimer's Disease. Frontiers in Neuroscience 12.

Schijven, D., Postema, M.C., Fukunaga, M., Matsumoto, J., Miura, K., de Zwarte, S.M.C., van Haren, N.E.M., Cahn, W., Hulshoff Pol, H.E., Kahn, R.S., Ayesa-Arriola, R., Ortiz-García de la Foz, V., Tordesillas-Gutierrez, D., Vázquez-Bourgon, J., Crespo-Facorro, B., Alnæs, D., Dahl, A., Westlye, L.T., Agartz, I., Andreassen, O.A., Jönsson, E.G., Kochunov, P., Bruggemann, J.M., Catts, S.V., Michie, P.T., Mowry, B.J., Quidé, Y., Rasser, P.E., Schall, U., Scott, R.J., Carr, V.J., Green, M.J., Henskens, F.A., Loughland, C.M., Pantelis, C., Weickert, C.S., Weickert, T.W., de Haan, L., Brosch, K., Pfarr, J.-K., Ringwald, K.G., Stein, F., Jansen, A., Kircher, T.T.J., Nenadić, I., Krämer, B., Gruber, O., Satterthwaite, T.D., Bustillo, J., Mathalon, D.H., Preda, A., Calhoun, V.D., Ford, J.M., Potkin, S.G., Chen, J., Tan, Y., Wang, Z., Xiang, H., Fan, F., Bernardoni, F., Ehrlich, S., Fuentes-Claramonte, P., Garcia-Leon, M.A., Guerrero-Pedraza, A., Salvador, R., Sarró, S., Pomarol-Clotet, E., Ciullo, V., Piras, F., Vecchio, D., Banaj, N., Spalletta, G., Michielse, S., van Amelsvoort, T., Dickie, E.W., Voineskos, A.N., Sim, K., Ciufolini, S., Dazzan, P., Murray, R.M., Kim, W.-S., Chung, Y.-C., Andreou, C., Schmidt, A., Borgwardt, S., McIntosh, A.M., Whalley, H.C., Lawrie, S.M., du Plessis, S., Luckhoff, H.K., Scheffler, F., Emsley, R., Grotegerd, D., Lencer, R., Dannlowski, U., Edmond, J.T., Rootes-Murdy, K., Stephen, J.M., Mayer, A.R., Antonucci, L.A., Fazio, L., Pergola, G., Bertolino, A., Díaz-Caneja, C.M., Janssen, J., Lois, N.G., Arango, C., Tomyshev, A.S., Lebedeva, I., Cervenka, S., Sellgren, C.M., Georgiadis, F., Kirschner, M., Kaiser, S., Hajek, T., Skoch, A., Spaniel, F., Kim, M., Kwak, Y.B., Oh, S., Kwon, J.S., James, A., Bakker, G., Knöchel, C., Stäblein, M., Oertel, V., Uhlmann, A., Howells, F.M., Stein, D.J., Temmingh, H.S., Diaz-Zuluaga, A.M., Pineda-Zapata, J.A., López-Jaramillo, C., Homan, S., Ji, E., Surbeck, W., Homan, P., Fisher, S.E., Franke, B., Glahn, D.C., Gur, R.C., Hashimoto, R., Jahanshad, N., Luders, E., Medland, S.E., Thompson, P.M., Turner, J.A., van Erp, T.G.M., Francks, C., 2023. Large-scale analysis of structural brain asymmetries in schizophrenia via the ENIGMA consortium. Proceedings of the National Academy of Sciences 120, e2213880120. https://doi.org/10.1073/pnas.2213880120

Schroeter, M.L., Raczka, K., Neumann, J., Yves von Cramon, D., 2007. Towards a nosology for frontotemporal lobar degenerations—A meta-analysis involving 267 subjects. NeuroImage 36, 497–510. https://doi.org/10.1016/j.neuroimage.2007.03.024

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

Thompson, S.A., Patterson, K., Hodges, J.R., 2003. Left/right asymmetry of atrophy in semantic dementia: Behavioral–cognitive implications. Neurology 61, 1196–1203. https://doi.org/10.1212/01.WNL.0000091868.28557.B8

Toga, A.W., Thompson, P.M., 2003. Mapping brain asymmetry. Nat Rev Neurosci 4, 37–48. https://doi.org/10.1038/nrn1009

Traub, J., Otto, M., Sell, R., Homola, G.A., Steinacker, P., Oeckl, P., Morbach, C., Frantz, S., Pham, M., Störk, S., Stoll, G., Frey, A., 2022. Serum glial fibrillary acidic protein indicates memory impairment in patients with chronic heart failure. ESC Heart Failure 9, 2626–2634. https://doi.org/10.1002/ehf2.13986

Whitwell, J.L., 2019. Chapter 3 - FTD spectrum: Neuroimaging across the FTD spectrum, in: Becker, J.T., Cohen, A.D. (Eds.), Progress in Molecular Biology and Translational Science, Brain Imaging. Academic Press, pp. 187–223. https://doi.org/10.1016/bs.pmbts.2019.05.009

Whitwell, J.L., Boeve, B.F., Weigand, S.D., Senjem, M.L., Gunter, J.L., Baker, M.C., DeJesus-Hernandez, M., Knopman, D.S., Wszolek, Z.K., Petersen, R.C., Rademakers, R., Jack Jr, C.R., Josephs, K.A., 2015. Brain atrophy over time in genetic and sporadic frontotemporal dementia: a study of 198 serial magnetic resonance images. European Journal of Neurology 22, 745–752. https://doi.org/10.1111/ene.12675

Whitwell, J.L., Weigand, S.D., Boeve, B.F., Senjem, M.L., Gunter, J.L., DeJesus-Hernandez, M., Rutherford, N.J., Baker, M., Knopman, D.S., Wszolek, Z.K., Parisi, J.E., Dickson, D.W., Petersen, R.C., Rademakers, R., Jack, C.R., Josephs, K.A., 2012. Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. Brain 135, 794–806. https://doi.org/10.1093/brain/aws001

Zhu, Z., Zhang, Z., Gao, X., Feng, L., Chen, D., Yang, Z., Hu, S., 2022. Individual Brain Metabolic Connectome Indicator Based on Jensen-Shannon Divergence Similarity Estimation Predicts Seizure Outcomes of Temporal Lobe Epilepsy. Frontiers in Cell and Developmental Biology 9.

Acknowledgments

The authors thank patients, their relatives, and healthy controls for participating in research studies. 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, AGAUR, Generalitat de Catalunya (SGR 2021- 01126) and by Spanish Ministry of Science and Innovation (PID2020-118386RA-I00/ AEI/10.13039/501100011033 to Dr. R. Sala-Llonch). María de Maeztu Unit of Excellence (Institute of Neurosciences, University of Barcelona) MDM-2017-0729. 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.

Data Availability Statement

Codified data supporting this study's findings are available from the corresponding author upon reasonable request.

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	**AD**	**FTD**	**CTR-AD p-values**	**CTR-FTD p-values**	**AD-FTD p-values**
N MRI	173	230	101	---	---	---
Sex at MRI, Men/Women	67/106	84/146	53/48	0.68	**0.048**	**0.023**
Age at MRI, years (SD)	59.4 (15.0)	65.3 (9.9)	63.7 (8.3)	**1.6e-6**	**0.0049**	0.24
CSF NfL, pg/mL (SD)	536.1 (312.6)	1106.9 (570.4)	2340.6 (1736.3)	**1.6e-7**	**< 2e-16**	**< 2e-16**
CSF 14-3-3, pg/mL (SD)	2531.9 (748.2)	5727.3 (2303.5)	4234.9 (1869.1)	**< 2e-16**	**3.0e-6**	**5.9e-6**
CSF YKL-40, pg/mL (SD)	270.1 (123.0)	328.6 (132.9)	315.3 (127.0)	**0.0028**	0.15	0.62
Plasma t-tau, pg/mL (SD)	4.4 (4.4)	4.2 (1.7)	3.7 (1.3)	0.65	0.65	0.65
Plasma p-tau,	2.4 (2.6)	5.4 (7.7)	1.3 (1.0)	**0.0011**	0.37	**0.0019**

pg/mL (SD)						
Plasma NfL, pg/mL (SD)	9.2 (6.6)	15.0 (6.5)	21.2 (15.9)	**1.7e-6**	**2.2e-9**	**8.0e-4**
Plasma GFAP, pg/mL (SD)	96.5 (64.9)	252.6 (154.6)	154.6 (101.7)	**1.0e-15**	**0.041**	**0.00066**
Plasma UCH-L1, pg/mL (SD)	35.1 (76.7)	17.5 (17.6)	22.1 (14.4)	0.15	0.71	0.44

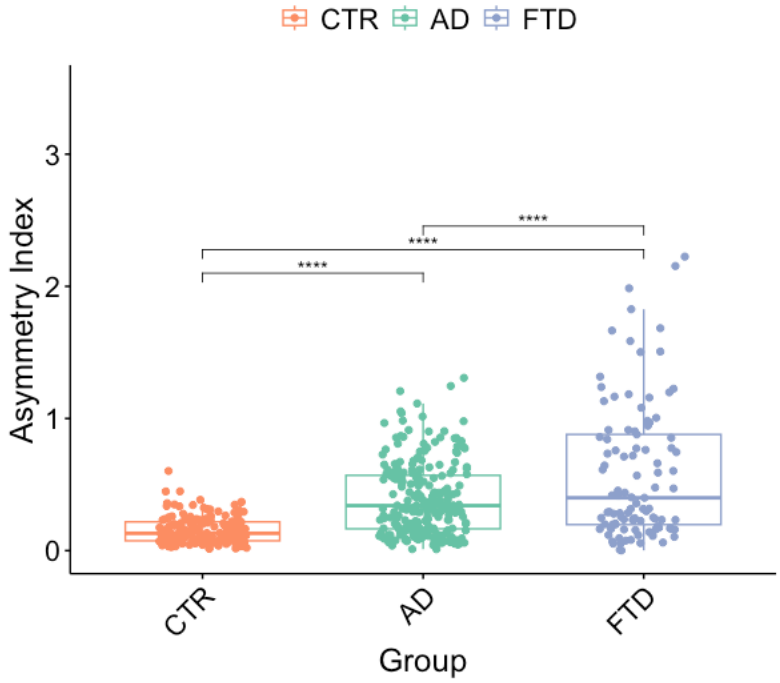

*Figure 1: Asymmetric index significant differences between AD and FTD patients and healthy controls. The symbol represents **** p<0.0001.*

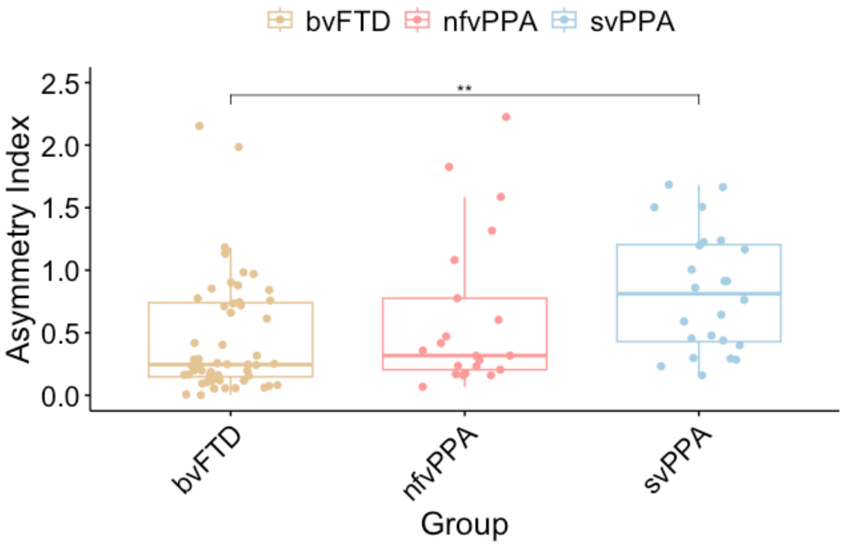

*Figure 2: Asymmetric index significant differences FTD variants. The symbol represents **p<0.01.*

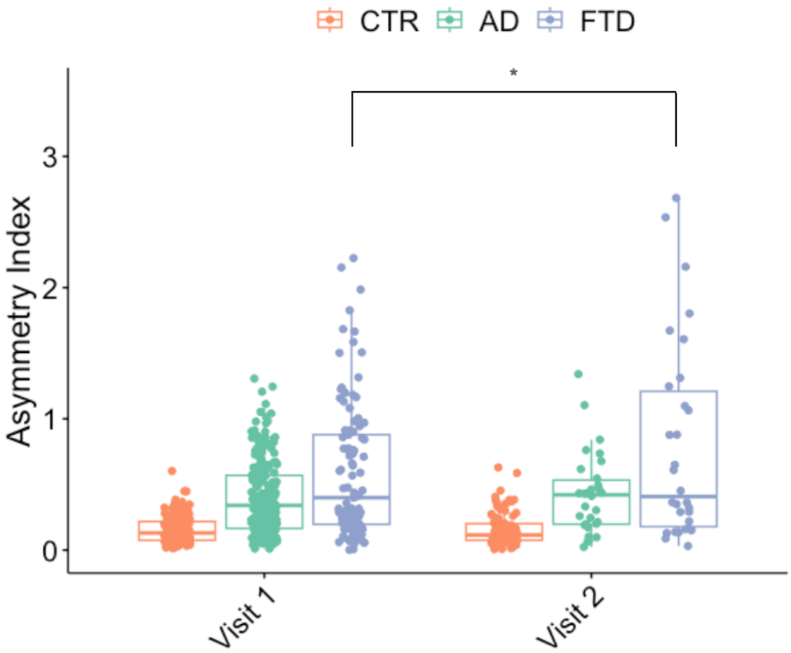

*Figure 3: Boxplot for Alzheimer's Disease patients (AD), frontotemporal dementia patients (FTD), and healthy controls (CTR) for three different visits. The symbol represents * p<0.05.*

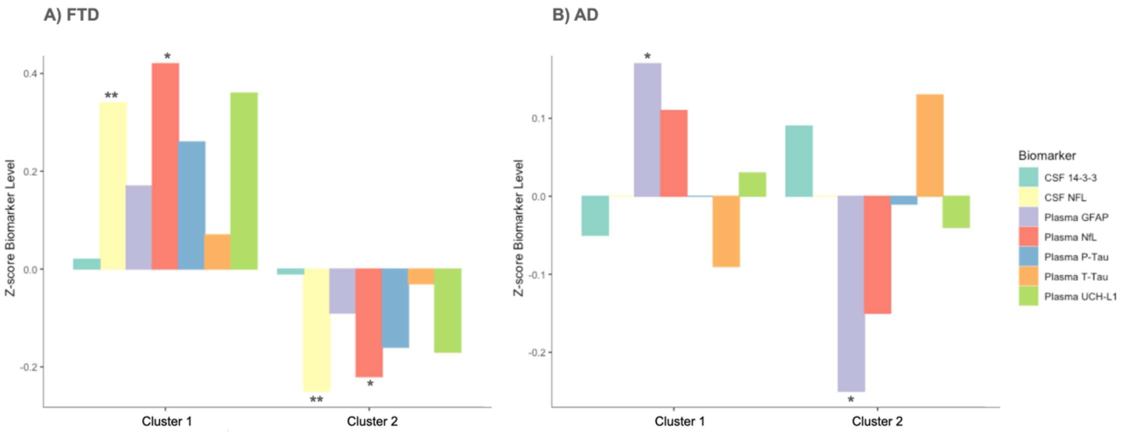

*Figure 4: Mean z-scores of each biomarker level within each group. A) Represents the FTD participants, and B) Represents the AD participants. Cluster 1, in both cases, is the most asymmetric. The symbol represents * p<0.05 and ** p<0.01.*

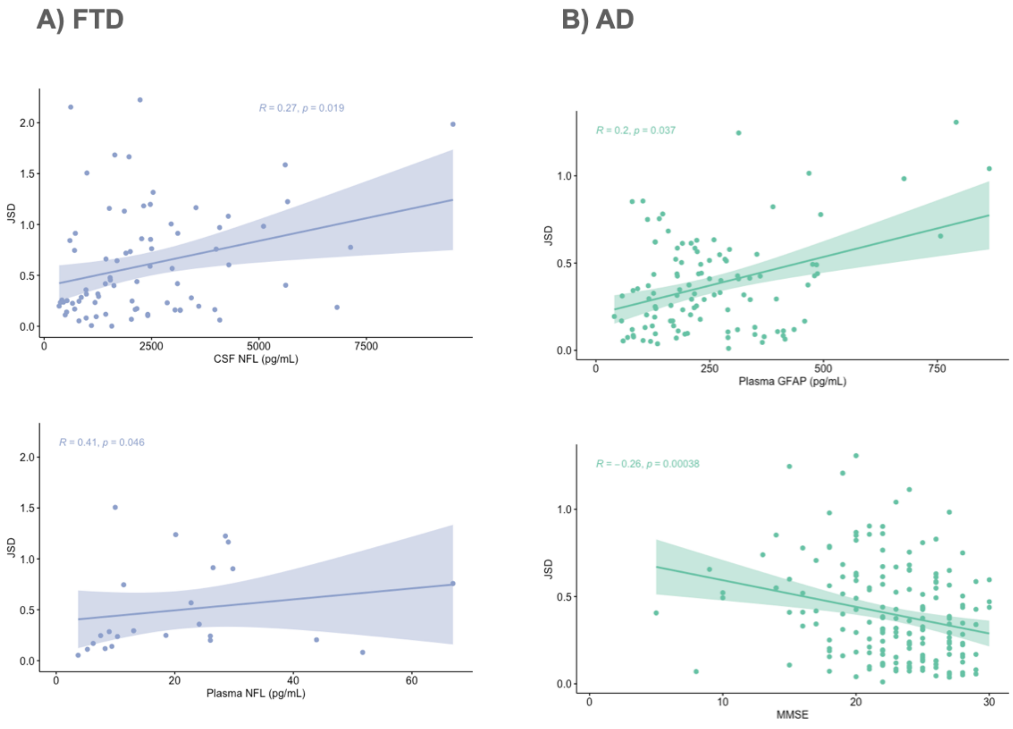

Figure 5: Correlation plots between the Cortical Asymmetric Index and biomarkers levels for the significant biomarkers (p-value<0.05). A) FTD participants and B) AD patients.

TREBALL 6

Loss of brainstem white matter predicts onset and motor neuron symptoms in *C9orf72* expansion carriers: a GENFI study.

Agnès Pérez-Millan*, Sergi Borrego-Écija*, John C. van Swieten, Lize Jiskoot, Fermin Moreno, Robert Laforce, Caroline Graff, Mario Masellis, Maria Carmela Tartaglia, James B. Rowe, Barbara Borroni, Elizabeth Finger, Matthis Synofzik, Daniela Galimberti, Rik Vandenberghe, Alexandre de Mendonça, Chris R. Butler, Alexander Gerhard, Simon Ducharme, Isabelle Le Ber, Isabel Santana, Florence Pasquier, Johannes Levin, Markus Otto, Sandro Sorbi, Pietro Tiraboschi, Harro Seelaar, Tobias Langheinrich, Jonathan D. Rohrer, Roser Sala-Llonch, Raquel Sánchez-Valle, on behalf of the Genetic FTD Initiative, GENFI

Journal of Neurology 2023; 2023 Març 1

DOI: 10.1007/s00415-022-11435-x

IF (2022): 6,0; Q1 Clinical Neurology

Journal of Neurology
https://doi.org/10.1007/s00415-022-11435-x

ORIGINAL COMMUNICATION

Loss of brainstem white matter predicts onset and motor neuron symptoms in *C9orf72* expansion carriers: a GENFI study

Agnès Pérez-Millan[1,2] · Sergi Borrego-Écija[1] · John C. van Swieten[3] · Lize Jiskoot[3,4] · Fermin Moreno[5,6] · Robert Laforce[7] · Caroline Graff[8,9] · Mario Masellis[10] · Maria Carmela Tartaglia[11] · James B. Rowe[12] · Barbara Borroni[13] · Elizabeth Finger[14] · Matthis Synofzik[15,16] · Daniela Galimberti[17,18] · Rik Vandenberghe[19,20] · Alexandre de Mendonça[21] · Chris R. Butler[22,23] · Alexander Gerhard[24,25] · Simon Ducharme[26,27] · Isabelle Le Ber[28,29] · Isabel Santana[30,31] · Florence Pasquier[32,33] · Johannes Levin[34,35] · Markus Otto[36] · Sandro Sorbi[37] · Pietro Tiraboschi[38] · Harro Seelaar[3] · Tobias Langheinrich[39] · Jonathan D. Rohrer[4] · Roser Sala-Llonch[2,40] · Raquel Sánchez-Valle[1] · The Genetic FTD Initiative, GENFI

Received: 5 August 2022 / Revised: 17 October 2022 / Accepted: 18 October 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany 2022

Abstract

Background and objectives The *C9orf72* expansion is the most common genetic cause of frontotemporal dementia (FTD) and/or motor neuron disease (MND). Corticospinal degeneration has been described in post-mortem neuropathological studies in these patients, especially in those with MND. We used MRI to analyze white matter (WM) volumes in presymptomatic and symptomatic *C9orf72* expansion carriers and investigated whether its measure may be helpful in predicting the onset of symptoms.

Methods We studied 102 presymptomatic *C9orf72* mutation carriers, 52 symptomatic carriers: 42 suffering from FTD and 11 from MND, and 75 non-carriers from the Genetic Frontotemporal dementia Initiative (GENFI). All subjects underwent T1-MRI acquisition. We used FreeSurfer to estimate the volume proportion of WM in the brainstem regions (midbrain, pons, and medulla oblongata). We calculated group differences with ANOVA tests and performed linear and non-linear regressions to assess group-by-age interactions.

Results A reduced WM ratio was found in all brainstem subregions in symptomatic carriers compared to both noncarriers and pre-symptomatic carriers. Within symptomatic carriers, MND patients presented a lower ratio in pons and medulla oblongata compared with FTD patients. No differences were found between presymptomatic carriers and non-carriers. Clinical severity was negatively associated with the WM ratio. *C9orf72* carriers presented greater age-related WM loss than non-carriers, with MND patients showing significantly more atrophy in pons and medulla oblongata.

Discussion We find consistent brainstem WM loss in *C9orf72* symptomatic carriers with differences related to the clinical phenotype supporting the use of brainstem measures as neuroimaging biomarkers for disease tracking.

Keywords Frontotemporal dementia · *C9orf72* · GENFI · Brainstem

Introduction

Frontotemporal dementia (FTD) refers to a heterogeneous group of neurodegenerative disorders that mainly affects the frontal and temporal lobes of the brain producing behavioral and language impairment [1]. Amyotrophic lateral sclerosis (ALS) is the most frequent motor neuron disease. It is caused by the neurodegeneration of motor neurons and the corticospinal and corticobulbar tracts leading to progressive weakness and muscular atrophy [2]. Due to the scientific advances in the last decades, it is now recognized that

---

Agnès Pérez-Millan and Sergi Borrego-Écija have contributed equally.

Roser Sala-Llonch and Raquel Sánchez-Valle have contributed equally.

List of GENFI consortium authors in the Appendix.

✉ Raquel Sánchez-Valle
rsanchez@clinic.cat

Extended author information available on the last page of the article

Published online: 29 November 2022

FTD and ALS are part of a clinical, neuropathological, and genetic continuum [3–6].

Although frequency varies geographically, the pathological hexanucleotide expansion in the *chromosome 9 open reading frame 72* (*C9orf72)* gene is the most common genetic cause of FTD, and ALS [7, 8]. The *C9orf72* repeat expansion is inherited with an autosomal dominant pattern with almost full penetrance leading to disease onset at a mean age of 58 years, although a wide range of age of onset (20–90 s) has been described [9]. The correlation between parental age at onset and individual age at onset for *C9orf72* expansion carriers is weak ($r = 0.32$), and thus, not useful for individual predictions [9]. In the same way, whether the symptom onset would appear in form of FTD, or ALS remains unpredictable. However, future disease-modifying drugs might be useful for both clinical phenotypes and treatments might be more useful when used in early or even presymptomatic phases of the disease. For that reason, there is a need for biomarkers that are able to provide information about the proximity of onset and track disease progression in both phenotypes. In this sense, cohorts of mutation carriers, such as the genetic frontotemporal initiative (GENFI), provide the opportunity to study the first stages of the disease and to identify markers of symptom onset and progression [10].

Previous studies have described structural changes in presymptomatic FTD subjects using brain MRI [11–15]. Concerning *C9orf72* carriers, previous studies have shown presymptomatic brain changes in the thalamus, cerebellum, hippocampus, amygdala, and hypothalamus [16, 17]. Most of these studies have focused on grey matter. In contrast, white matter (WM) degeneration has received comparatively less attention but demonstrates early and widespread WM integrity loss in *C9orf72* carriers [18].

The neuropathological examination of ALS patients reveals loss of motor neurons and the consequent degeneration of the corticospinal and corticobulbar tracts [19, 20]. This degeneration leads to lateral sclerosis of the spinal cord which gives the name to the disease. In addition to spinal cord changes, ALS patients also present relevant atrophy of the white matter areas that contain the corticospinal and corticobulbar tracts at the brainstem, especially the pyramids in the medulla oblongata. Previous work has demonstrated that changes in the spinal cord and brainstem in ALS can be detected in vivo using structural MRI [21, 22]. In a recent study, Querin et al. reported significant WM reduction in the spinal cord of presymptomatic *C9orf72* carriers using cervical cord MRI [23]. Assessing WM changes in the brainstem presents some potential benefits from cervical spinal cord evaluation, as the possibility of being measured with other brain changes in the brain MRI.

In this work, we investigate the utility of brainstem WM loss as a biomarker for *C9orf72* patients. We hypothesize that symptomatic *C9orf72* carriers would present more WM loss in the brainstem compared to non-carriers, especially in those patients with motor neuron symptoms. We also aim to study whether WM loss is identifiable in presymptomatic *C9orf72* carriers.

Materials and methods

Participants

Two hundred thirty five participants' data were obtained from the data freeze 4 (DF4) of the GENFI, an international multicenter study of known carriers of a pathogenic mutation or at risk of carrying a mutation because a first-degree relative was a known symptomatic carrier [11]. Symptomatic subjects were FTD or ALS patients carrying the *C9orf72* pathogenic expansion. Presymptomatic and noncarriers subjects were all first-degree relatives of *C9orf72* mutation carriers who consent to be tested for their genetic status.

All participants' imaging data were acquired at each time point using 3 T on scanners from three different manufacturers: Philips Healthcare (Koninklijke Philips NV, Amsterdam, Netherlands), GE Healthcare Life Sciences (General Electric, Boston, MA, USA) and Siemens Healthcare Diagnostics (Siemens, Erlangen, Germany). Protocols were designed to harmonize across scanners and sites as much as possible [11]. Subjects were classified into four groups according to their genetic status (carriers or non-carriers) and their clinical diagnosis as follows: (a) non-carriers; (b) presymptomatic *C9orf72* carriers if no diagnostic criteria were fulfilled, (c) symptomatic *C9orf72* carriers with FTD presentations in the form of behavioral variant FTD [24] or primary progressive aphasia [25] and (d) symptomatic *C9orf72* carriers with MND presentation in form of ALS or ALS-FTD [26, 27]. The disease stage of all participants was scored following the global and sum of boxes Clinical Dementia Rating adapted to FTD patients (CDR® + NACC-FTLD) rating scale [28]. The severity of motor neuron symptoms was scored with the ALS Functional Rating Scale-Revised (ALSFRS-R), a validated rating instrument for monitoring the progression of disability in ALS patients [29]. The ALSFRS-R obtains a final index of disability by scoring 12 different motor and respiratory items from 4 (no disability) to 0 (marked disability). Written informed consent was obtained from all participants. All procedures were approved by local ethics committees at each site.

MRI acquisition and processing

Participants underwent a 1.1-mm isotropic resolution volumetric T1 MRI imaging on a 3T scan using the sequences defined within the GENFI consortium. Nineteen scanners were used across different sites. MRIs of all subjects were downloaded from GENFI database and processed using

FreeSurfer version 6.0 (http://surfer.nmr.mgh.harvard.edu/) in the same center.

After the standard FreeSurfer segmentation and parcellation [30–32], we used an additional FreeSurfer pipeline to segment the brainstem region and its three main structures (midbrain, pons, and medulla oblongata) [33]. Figure 1 represents the imaging methodology to obtain the brainstem region segmentation. We assessed the WM parcel for the brainstem structures by multiplying each of the regions by the WM mask. To remove the effect of brain size, we calculated the ratio of WM for each of its structures (midbrain, pons, and medulla oblongata) using the total volume of the corresponding region (region-WM volume/region-whole volume). All images were visually inspected and manually corrected when needed.

Statistical analysis

Differences in demographic data between groups were assessed using ANOVA test for continuous variables and Fisher test for dichotomous data. Post-hoc studies were assessed for both cases to identify the pair-wise group differences, using T-tests or Fisher test accordingly. Statistical significance was set at $p < 0.05$, with corrections for multiple comparisons using the Benjamini–Hochberg procedure.

We used ANOVA test to study group differences in the WM ratio for the brainstem subregions. Age at baseline, sex and scanner were used as covariates. Then, Tukey's HSD test was used to identify pairwise differences between groups with Benjamini–Hochberg corrections for multiple comparisons. We compared the non-carriers, the presymptomatic carriers, carriers with FTD, and carriers with MND with the same procedure. Differences in the WM ratio between CDR® + NACC-FTLD global stages were assessed using Kruskal–Wallis test for all carriers, while Spearman's rank correlation coefficient was used to study the relationship between the WM ratio and the CDR® + NACC-FTLD sum of boxes and the ALSFRS-R. We evaluated multiple linear and non-linear regressions (logarithmic, polynomial to the second, third and fourth order) to test the association between the WM ratio (dependent variable) and the genetic status, age, and their interaction. For these analyses we added scanner and sex as covariates. Models were compared using R^2 and the Akaike information criterion (AIC). R (https://www.r-project.org/) version 4.0.5 was used for all analyses.

Results

Demographic and clinical characteristics of participants

After the data quality assessment the sample was reduced to 229 participants due to the segmentation problems identified. The final sample used in the analyses included: 102 presymptomatic carriers, 52 symptomatic carriers (41 FTD and 11 ALS or ALS-FTD), and 75 non-carriers (Table 1). Some of the acquisitions ($N = 43$ subjects) had a limited Field of View, so it was not possible to measure the entire medulla oblongata ROI. Thus, these images were not included in the sub-analyses of this region (21 presymptomatic, 7 symptomatic, and 15 non-carriers).

We found significant differences between the four groups (non-carriers, presymptomatic, symptomatic-FTD, symptomatic-ALS) in sex and age. Both symptomatic groups were older than the non-carriers and presymptomatic groups ($p < 0.0001$). Therefore, these variables were included as covariates in all further analyses. No significant differences were found in any demographic or clinical variables between non-carriers and presymptomatic carriers (Table 1).

Group differences in brainstem WM ratio

Non-carriers showed WM ratios very close to 1 (0.96 for the midbrain, 0.99 for the pons, and 0.97 for the medulla). No differences were found in any region between the presymptomatic and the non-carrier groups. The *C9orf72* FTD group

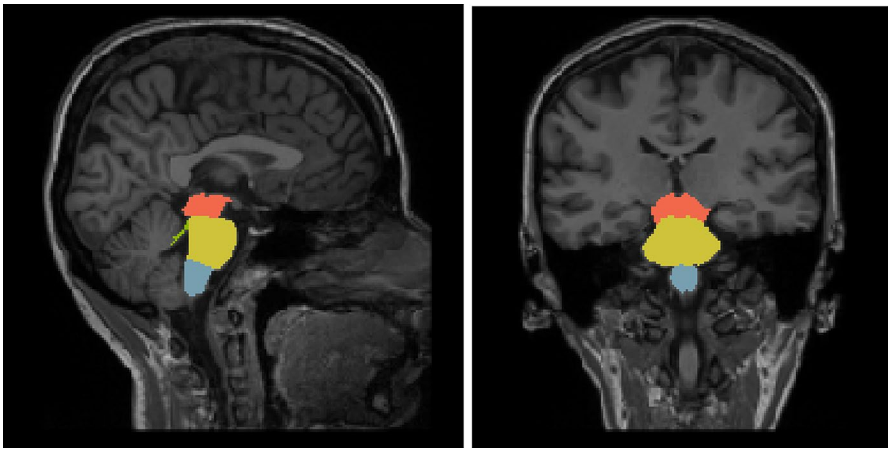

Fig. 1 The brainstem segmentation for all the matters for two different views. Orange represents the midbrain region, yellow represents the pons region and blue represents the medulla oblongata region. In this case, this subject is a healthy control

Table 1 Baseline demographics for controls, presymptomatic and both symptomatic carriers groups

	Non-carriers	*C9orf72* presymptomatic carriers	*C9orf72* FTD carriers	*C9orf72* MND carriers
Number of participants	75	102	41	11
Sex (f/m)	48/27	63/39	16/25*	4/7*
Age, years Mean (sd)	45.2 (12.6)	44.9 (11.8)	62.8 (8.4)**	62.6 (6.4)**
Age at onset, years Mean (sd)	–	–	57.2 (9.5)	59.5 (6.1)
EYO, years Mean (sd)	− 15.0 (11.6)	− 13.8 (11.9)	5.1 (6.1)**	1.4 (4.0)**
CDR® + NACC-FTLD Global Median (range)	–	–	2 (1–3)	2 (1–3)
CDR® + NACC-FTLD Sum of Boxes Median (range)	–	–	12.5 (1–22)	7.5 (1–18)

Brainstem subregions volumes and WM ratio. Show the group differences for the whole volume/WM ratio

EYO estimated years to onset, *FTD* frontotemporal dementia, *f* female, *m* male, *MND* motor neuron disease, *sd* standard deviation

*Statistical differences ($p < 0.05$) compared with non-carriers and presymptomatic carriers

**Statistical differences ($p < 0.0001$) compared with non-carriers and presymptomatic carriers

showed a lower WM ratio than non-carriers and presymptomatic carriers in all regions ($p < 0.01$ in the medulla, and $p < 0.0001$ in the midbrain and pons). The *C9orf72* MND group showed a lower WM ratio than the non-carriers and the presymptomatic carriers in all regions ($p < 0.0001$ in all comparisons). The MND group also showed a lower WM ratio than the FTD group in the medulla ($p < 0.0001$), and pons ($p < 0.0001$; (Fig. 2).

WM ratio across the severity of cognitive and motor symptoms

When studying the relationship between the WM ratio with the global CDR® + NACC-FTLD rating scale for all carriers subjects, we observed that higher clinical scores were significatively associated with lower WM ratios in all brainstem regions (Kruskal–Wallis $p < 0.001$ for all regions; Fig. 3A). Pairwise comparisons between CDR® + NACC-FTLD stages were performed for consecutive stages, depicting significant differences between the CDR = 0.5 and CDR = 1

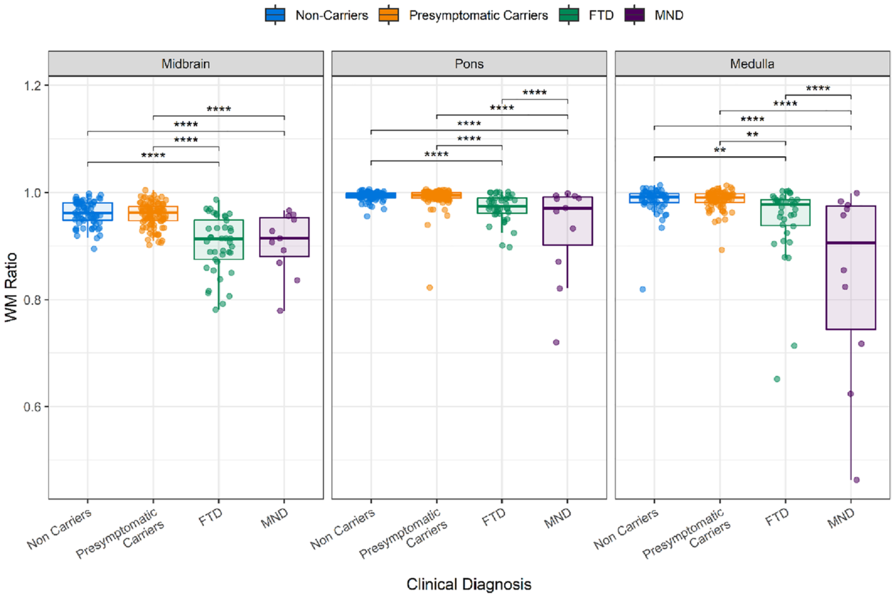

Fig. 2 Boxplot of the WM ratio volume of each brainstem region at baseline. Indicates $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$, $****p < 0.0001$

stages in the midbrain ($p < 0.05$). Additionally, moderate significant negative correlations between the WM ratio and the CDR® + NACC FTLD sum of boxes were also found for all brainstem regions (midbrain $r = -0.57$, pons $r = -0.49$ and medulla oblongata $r = -0.45$; $p < 0.0001$ all; Fig. 3B).

To assess if the WM ratio was correlated to the severity of the motor neuron symptoms, we evaluate its relationship with the ALSFRS-R score in *C9or72* carriers (Fig. 3C). We found a weak negative correlation in pons ($r = -0.37$, $p < 0.05$), but a moderate negative correlation in midbrain ($r = -0.45$; $p < 0.001$) and medulla ($r = -0.46$, $p < 0.01$).

Brainstem WM ratio and age trajectories according to the genetic status

When comparing the relationship between the WM ratio and age, we found that carriers and non-carriers showed similar trajectories until the 6th decade of life. After this

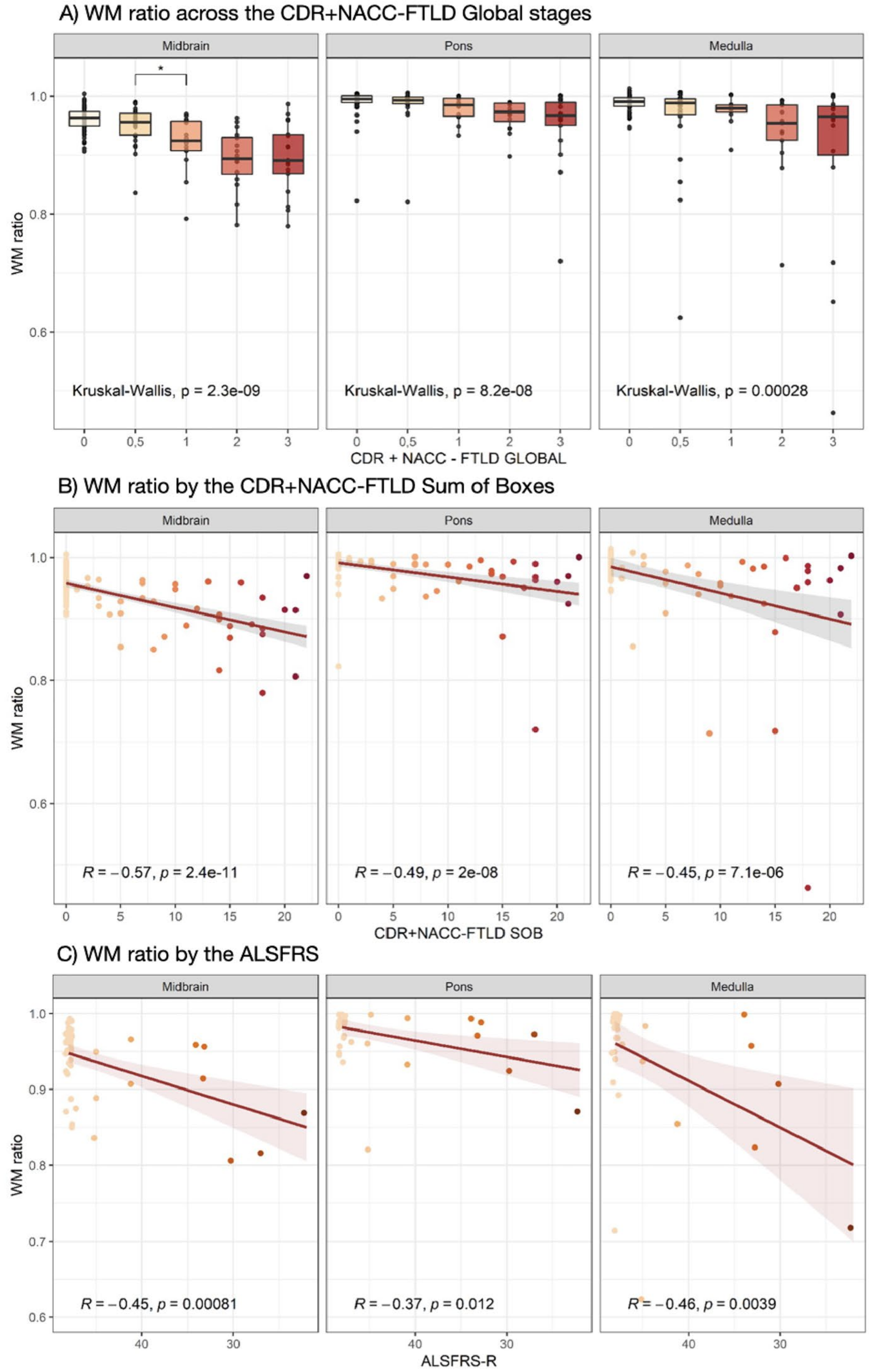

Fig. 3 **A** Boxplot of WM ratio across the *CDR® + NACC-FTLD* Global stages for the carriers' participants. Pairwise comparisons between stages were performed only for consecutive stages, finding significant differences between the 0.5 and the 1 stages in the midbrain: *$p < 0.05$, **B** Scatter plots of WM ratio by the *CDR® + NACC-FTLD* Sum Of Boxes. Red lines represent the correlation analyses, **C** Scatter plots of WM ratio by the ALSFRS-R in the different regions for the carriers' participants. Red lines represent the correlation analyses

age, carriers presented a greater loss of WM ratio than non-carriers, especially in the midbrain (Fig. 4A). The multiple linear regression comparing carriers and non-carriers showed similar results (Table 2). For both groups, age was related to lower WM ratios in the midbrain ($p < 0.001$). Carriers showed a greater loss of WM ratio by age than non-carriers in the midbrain ($p < 0.05$), suggesting a further loss of WM due to neurodegeneration. No other statistical differences were found between carriers and non-carriers. Due to the distribution of the trajectories, we also explored

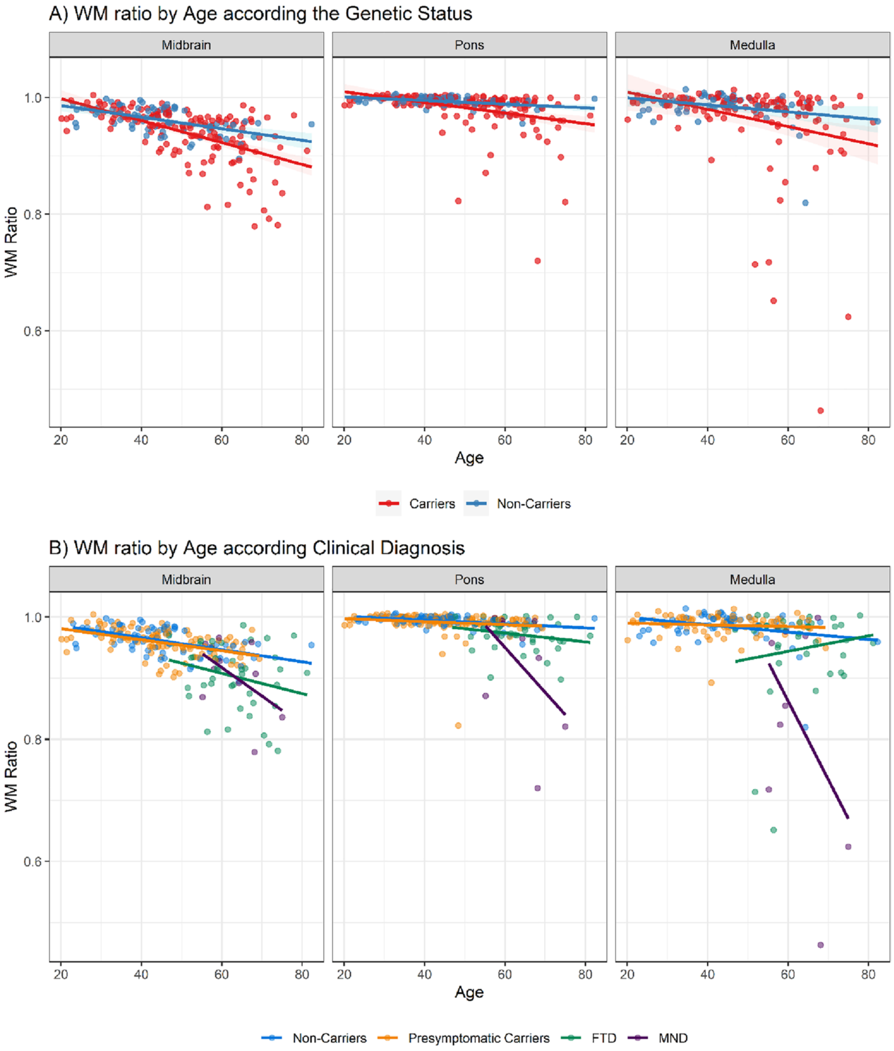

Fig. 4 Scatter plot showing the correlation between the WM ratio and age for each of the studied groups: the whole brainstem, the midbrain, the pons, and the medulla oblongata

Table 2 Multiple linear regression coefficients for comparing carriers and non-carriers

	Midbrain			Pons			Medulla		
	β	sd	*p*	*β*	sd	*p*	*β*	sd	*p*
Intercept	1.0121	0.0146	**< 0.0001**	1.0149	0.0126	**< 0.0001**	1.0284	0.0319	**< 0.0001**
Age	– 0.0010	0.0003	**< 0.001**	– 0.0003	0.0002	0.183	– 0.0007	0.0006	0.282
Scanner	– 0.0004	0.0004	0.257	– 0.0006	0.0003	0.076	– 0.0010	0.0009	0.261
Sex									
Female vs male	– 0.0056	0.0042	0.196	– 0.0050	0.0037	0.177	– 0.0159	0.0095	0.096
Genetic status									
Carriers vs noncarriers	0.02865	0.0172	0.098	0.0194	0.0148	0.192	0.0247	0.0368	0.502
Age × genetic status									
Carriers vs noncarriers	– 0.0008	0.0004	**0.019**	– 0.0005	0.0003	0.080	– 0.0007	0.0007	0.342

Significant group differences (p<0.05) are highlighted in bold non-linear regressions, but they did not improve the linear model significatively.

Brainstem WM ratio and age trajectories according to the clinical status

Finally, we assessed the brainstem WM trajectories by age according to the clinical status to evaluate if subjects with different clinical diagnoses present different trajectories of brainstem WM during the disease. In that sense, the MND group showed a greater loss of WM by age in all regions compared to FTD patients, the medulla being the region with the highest effect of age in WM loss for this group of patients (Fig. 4B; Table 3).

Discussion

In the present study, we used brain MRI scans from the GENFI consortium to investigate whether corticospinal and corticobulbar tracts neurodegeneration is measurable in the brainstem structures of *C9orf72* carriers. Symptomatic *C9orf72* expansion carriers showed consistent alterations in brainstem WM that correlated with clinical severity. Subjects with motor neuron symptoms presented more WM loss in the brainstem than those without motor symptoms.

Brainstem neuroimaging abnormalities have been investigated by means of semi-automated volumetry methods, especially in progressive supranuclear palsy [21, 34]. Concerning *C9orf72* expansion carriers, previous work found no structural volumetric gray matter (GM) impairment in the brainstem [16, 35]. However, the evaluation of brainstem WM in *C9of72* was lacking. Here, we developed a measure of WM degeneration consisting of the proportion of the brainstem volume occupied by WM. We chose the proportion of WM instead of its whole volume to avoid differences due to different brain sizes. Assessed in the non-carriers as controls, this WM ratio showed values close to 1, reflecting that, in normal conditions, the relative GM volume in the brainstem is scarce. However, these high values might reflect an overestimation of the WM volumes. Previous neuroimaging studies have shown that small brainstem pathways might be artificially enlarged due to the inclusion of crossing fibers [36, 37]. Despite this limitation, our work found differences between groups, reflecting the utility of this measure as a neuroimaging biomarker.

We found a lower brainstem WM ratio in symptomatic *C9orf72* carriers compared to non-carriers regardless of their clinical phenotype. These differences were found in the three sub-structures (midbrain, pons, and medulla oblongata), suggesting widespread neurodegeneration of the corticospinal tracts. No differences were found between presymptomatic carriers and controls. This finding would suggest that the neurodegeneration of the WM tracts appears near the onset of the symptoms, pointing to the brainstem WM ratio as a biomarker of conversion in *C9orf72* carriers. Whether the WM neurodegeneration occurs before or after the symptom's onset remains unclear. Our study did not show WM changes in the presymptomatic carriers' group. By contrast, Querin et al. recently observed spinal cord WM atrophy in presymptomatic *C9orf72* carriers who were older than 40 years [23]. This could suggest that the spinal cord would show signs of WM alterations before the brainstem, or it could be the result of including participants who were far from the estimated year of onset in our study. The observed relationship between the brainstem WM ratio and age sheds light on this point. Overall, all subjects showed a mild loss of WM over the years with both groups, carriers, and non-carriers, showing no differences until the 6th decade of life

Table 3 Multiple linear regression coefficients for assessing the brainstem WM trajectories by age according to the clinical status

	Midbrain			Pons			Medulla		
	β	sd	*p*	*β*	sd	*p*	*β*	sd	*p*
Intercept	1.0110	0.0137	**< 0.0001**	1.0110	0.0116	**< 0.0001**	1.0146	0.0271	**< 0.0001**
Age	− 0.0010	0.0003	**< 0.001**	− 0.0003	0.0002	0.151	− 0.0006	0.0005	0.274
Scanner	− 0.0005	0.0004	0.149	− 0.0003	0.0003	0.307	− 0.0002	0.0008	0.822
Sex									
Female vs male	0.0001	0.0004	0.998	− 0.0004	0.0034	0.895	− 0.0033	0.0084	0.691
Clinical group									
Presymptomatic vs control	− 0.0083	0.0176	0.638	− 0.0056	0.0143	0.694	− 0.0175	0.0341	0.608
FTD vs control	− 0.0045	0.0380	0.905	− 0.0061	0.0308	0.841	− 0.1444	0.0730	**0.049**
MND vs control	0.1992	0.0931	**0.033**	0.3871	0.0756	**< 0.0001**	0.6220	0.1735	**< 0.001**
Age × clinical group									
Presymptomatic vs control	0.0001	0.0004	0.729	0.0001	0.0003	0.773	0.0004	0.0007	0.553
FTD vs control	− 0.0005	0.0006	0.369	0.0003	0.0005	0.500	0.0019	0.0012	0.118
MND vs control	− 0.0037	0.0015	**0.013**	− 0.0070	0.0012	**< 0.0001**	− 0.0122	0.0028	**< 0.0001**

Significant group differences ($p<0.05$) are highlighted in bold when *C9orf72* carriers suffer a greater WM loss, especially in the midbrain. Of note, this decade of life coincides with the onset of symptoms reported recently by Moore et al., reinforcing the idea of the brainstem WM ratio as a possible biomarker of conversion [9].

In consonance with neuropathological studies, patients with MND showed significantly more atrophy in the pons and especially in the medulla oblongata compared to FTD. Similar results were found in the multivariate analyses, where patients presenting in form of MND suffered further loss of WM ratio than the other groups, particularly in the medulla oblongata (Fig. 4B). We hypothesize that this greater loss of WM in *C9orf72* carriers is due to the neurodegeneration of the corticospinal and corticobulbar tracts in patients who develop motor neuron symptoms. These results suggest that the brainstem WM ratio, especially in the medulla oblongata, could be an interesting biomarker to predict motor neuron symptoms in *C9orf72* carriers. This finding is particularly relevant because the form of onset in *C9orf72* carriers is highly unpredictable, and patients with motor neuron symptoms have a worse overall prognosis. Moreover, most neuroimage biomarkers studied in *C9orf72* carriers have focused on cortical atrophy, but MND patients may present only subtle cortical atrophy, especially in those with bulbar onset where, theoretically, brainstem changes were supposed to be more remarkable.

Additionally, we evaluated if the WM ratio could monitor the disease progression. For this purpose, we assessed the WM ratio across the different disease stages measured with the CDR® + NACC-FTLD scale. Here, a biological gradient was found, with patients in more advanced stages showing lower WM ratios. This loss of WM was greater in the midbrain with significant differences between the CDR = 0.5 and the CDR = 1 stages in the region. We also found a negative correlation between the CDR® + NACC-FTLD sum of boxes and the WM ratio in the brainstem. This correlation was, again, strongest in the midbrain ($r = -0.60$). We also evaluated the correlation between the WM ratio and the severity of the motor neuron symptoms in *C9orf72* carriers. A negative correlation between the WM ratio and the ALSFRS-R was found in all the brainstem regions. However, this correlation was highly influenced by subjects without motor neuron symptoms.

Our study has some limitations. First, it is important to consider that the brainstem WM visualization is challenging due to the small size of the pathways, the high density of their distributions, lower contrast, and image distortions associated with in vivo acquisitions. As mentioned before, brain volumetry could overestimate WM volumes. Despite this possible limitation, we found that our methodology is valid to find differences between groups. To support and complement our results, other MRI modalities such as Diffusion tensor imaging (DTI) may be studied in the future. Another limitation is the relatively small sample size. Even using data from a multicentric study, in some analyses, especially for the MND subgroup, the number of subjects was low, due to the low prevalence of the disease. This small number of MND patients did not allow us to study differences between subjects with bulbar or spinal onset.

In conclusion, our data suggest that WM loss in the brainstem might be a marker of clinical conversion and disease progression monitoring in C9orf72 carriers, especially in carriers presenting with motor neuron symptoms. Additional studies with extended follow-up data might be needed to confirm these findings.

Appendix

List of GENFI consortium authors

- Abbe Ullgren, Center for Alzheimer Research, Division of Neurogeriatrics, Karolinska Institutet, Stockholm, Sweden
- Adeline Rollin, CHU, CNR-MAJ, Labex Distalz, LiCEND Lille, France
- Agnès Camuzat, Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France
- Aitana Sogorb Esteve, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Alazne Gabilondo, Cognitive Disorders Unit, Department of Neurology, Donostia University Hospital, San Sebastian, Gipuzkoa, Spain; Neuroscience Area, Biodonostia Health Research Insitute, San Sebastian, Gipuzkoa, Spain
- Albert Lladó, Alzheimer's disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic, Barcelona, Spain
- Alberto Benussi, Centre for Neurodegenerative Disorders, Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy
- Alexis Brice, Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France
- Ana Gorostidi, Neuroscience Area, Biodonostia Health Research Insitute, San Sebastian, Gipuzkoa, Spain
- Ana Verdelho, Department of Neurosciences and Mental Health, Centro Hospitalar Lisboa Norte - Hospital de Santa Maria & Faculty of Medicine, University of Lisbon, Lisbon, Portugal
- Andrea Arighi, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Neurodegenerative Diseases Unit, Milan, Italy
- Anna Antonell, Alzheimer's disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic, Barcelona, Spain

- Anne Bertrand, Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France
- Annerose Engel, Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany
- Annick Vogels, Department of Human Genetics, KU Leuven, Leuven, Belgium
- Arabella Bouzigues, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Aurélie Funkiewiez, Centre de référence des démences rares ou précoces, IM2A, Département de Neurologie, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France
- Benedetta Nacmias, Department of Neuroscience, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy
- Benjamin Bender, Department of Diagnostic and Interventional Neuroradiology, University of Tübingen, Tübingen, Germany
- Camilla Ferrari, Department of Neuroscience, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy
- Carlo Wilke, Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research and Center of Neurology, University of Tübingen, Tübingen, Germany; Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
- Carolin Heller, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Carolina Maruta, Laboratory of Language Research, Centro de Estudos Egas Moniz, Faculty of Medicine, University of Lisbon, Lisbon, Portugal
- Caroline V. Greaves, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Carolyn Timberlake, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
- Catarina B. Ferreira, Laboratory of Neurosciences, Faculty of Medicine, University of Lisbon, Lisbon, Portugal
- Catharina Prix, Neurologische Klinik, Ludwig-Maximilians-Universität München, Munich, Germany
- Chiara Fenoglio, University of Milan, Centro Dino Ferrari, Milan, Italy
- Christen Shoesmith, Department of Clinical Neurological Sciences, University of Western Ontario, London, Ontario, Canada
- Cristina Polito, Department of Biomedical, Experimental and Clinical Sciences “Mario Serio”, Nuclear Medicine Unit, University of Florence, Florence, Italy
- Daisy Rinaldi, Centre de référence des démences rares ou précoces, IM2A, Département de Neurologie, AP-HP - Hôpital Pitié-Salpêtrière (DMU Neurosciences Paris 6), Paris, France
- Dario Saracino, Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière (DMU Neurosciences Paris 6), Paris, France
- David Cash, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- David L. Thomas, Neuroimaging Analysis Centre, Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, Queen Square, London, UK
- David Tang-Wai, The University Health Network, Krembil Research Institute, Toronto, Canada
- Diana Duro, Faculty of Medicine, University of Coimbra, Coimbra, Portugal
- Ekaterina Rogaeva, Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Canada
- Elio Scarpini, University of Milan, Centro Dino Ferrari, Milan, Italy
- Elisabeth Wlasich, Neurologische Klinik, Ludwig-Maximilians-Universität München, Munich, Germany
- Emanuele Buratti, Molecular Pathology Laboratory, International Centre for Genetic Engineering and Biotechnology (ICGEB), 34149 Trieste, Italy
- Emily Todd, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Enrico Premi, Stroke Unit, ASST Brescia Hospital, Brescia, Italy
- Frederico Simões do Couto, Faculdade de Medicina, Universidade Católica Portuguesa
- Gabriel Miltenberger, Faculty of Medicine, University of Lisbon, Lisbon, Portugal
- Gemma Lombardi, IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy
- Giacomina Rossi, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Giorgio Fumagalli, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Neurodegenerative Diseases Unit, Milan, Italy
- Giorgio Giaccone, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Giuseppe Di Fede, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Gregory Kuchcinski, Univ Lille, France
- Hanya Benotmane, UK Dementia Research Institute at University College London, UCL Queen Square Institute of Neurology, London, UK
- Henrik Zetterberg, UK Dementia Research Institute at University College London, UCL Queen Square Institute of Neurology, London, UK

- Imogen J. Swift, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Jackie Poos, Department of Neurology, Erasmus Medical Center, Rotterdam, Netherlands
- Janne M. Papma, Department of Neurology, Erasmus Medical Center, Rotterdam, Netherlands
- Jennifer Nicholas, Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK
- João Durães, Neurology Department, Centro Hospitalar e Universitario de Coimbra, Coimbra, Portugal
- Jolina Lombardi, Department of Neurology, University of Ulm, Ulm
- Jordi Juncà-Parella, Alzheimer's disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic, Barcelona, Spain
- Jordi Sarto, Alzheimer's disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic, Barcelona, Spain
- Jorge Villanua, OSATEK, University of Donostia, San Sebastian, Gipuzkoa, Spain
- Kiran Samra, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Koen Poesen, Laboratory for Molecular Neurobiomarker Research, KU Leuven, Leuven, Belgium
- Linn Öijerstedt, Center for Alzheimer Research, Division of Neurogeriatrics, Department of Neurobiology, Care Sciences and Society, Bioclinicum, Karolinska Institutet, Solna, Sweden
- Lisa Graf Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research and Center of Neurology, University of Tübingen, Tübingen, Germany
- Lucia Giannini, Department of Neurology, Erasmus Medical Center, Rotterdam, Netherlands
- Lucy L. Russell, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Maria João Leitão, Centre of Neurosciences and Cell Biology, Universidade de Coimbra, Coimbra, Portugal
- Maria Rosario Almeida, Faculty of Medicine, University of Coimbra, Coimbra, Portugal
- Maria Serpente, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Neurodegenerative Diseases Unit, Milan, Italy
- Marisa Lima, Neurology Department, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
- Marta Cañada, CITA Alzheimer, San Sebastian, Gipuzkoa, Spain
- Martina Bocchetta, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Maryna Polyakova, Department for Neurology, Max Planck Institute for Human Cognitive and Brain Sciences and Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany
- Mathieu Vandenbulcke, Geriatric Psychiatry Service, University Hospitals Leuven, Belgium; Neuropsychiatry, Department of Neurosciences, KU Leuven, Leuven, Belgium
- Maxime Bertoux, Inserm 1172, Lille, France
- Michele Veldsman, Nuffield Department of Clinical Neurosciences, Medical Sciences Division, University of Oxford, Oxford, UK
- Miguel Castelo-Branco, Faculty of Medicine, University of Coimbra, Coimbra, Portugal
- Miguel Tábuas-Pereira, Neurology Department, Centro Hospitalar e Universitario de Coimbra, Coimbra, Portugal
- Mikel Tainta, Neuroscience Area, Biodonostia Health Research Insitute, San Sebastian, Gipuzkoa, Spain
- Mircea Balasa, Alzheimer's disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic, Barcelona, Spain
- Miren Zulaica, Neuroscience Area, Biodonostia Health Research Insitute, San Sebastian, Gipuzkoa, Spain
- Morris Freedman, Baycrest Health Sciences, Rotman Research Institute, University of Toronto, Toronto, Canada
- Myriam Barandiaran, Cognitive Disorders Unit, Department of Neurology, Donostia University Hospital, San Sebastian, Gipuzkoa, Spain; Neuroscience Area, Biodonostia Health Research Insitute, San Sebastian, Gipuzkoa, Spain
- Nuria Bargalló, Imaging Diagnostic Center, Hospital Clínic, Barcelona, Spain
- Olivia Wagemann, Neurologische Klinik, Ludwig-Maximilians-Universität München, Munich, Germany
- Olivier Colliot, Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France
- Paola Caroppo, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Patricia Alves, Neuroscience Area, Biodonostia Health Research Insitute, San Sebastian, Gipuzkoa, Spain; Department of Educational Psychology and Psychobiology, Faculty of Education, International University of La Rioja, Logroño, Spain
- Paul Thompson, Division of Neuroscience and Experimental Psychology, Wolfson Molecular Imaging Centre, University of Manchester, Manchester, UK
- Pedro Rosa-Neto, Translational Neuroimaging Laboratory, McGill Centre for Studies in Aging, McGill University, Montreal, Québec, Canada
- Philip Van Damme, Neurology Service, University Hospitals Leuven, Belgium; Laboratory for Neurobiology, VIB-KU Leuven Centre for Brain Research, Leuven, Belgium

- Pietro Tiraboschi, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Rachelle Shafei, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Rhian S. Convery, Department of Neurodegenerative Disease, Dementia Research Centre, UCL Queen Square Institute of Neurology, London, UK
- Rick van Minkelen, Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, Netherlands
- Robart Bartha, Department of Medical Biophysics, The University of Western Ontario, London, Ontario, Canada; Centre for Functional and Metabolic Mapping, Robarts Research Institute, The University of Western Ontario, London, Ontario, Canada
- Roberto Gasparotti, Neuroradiology Unit, University of Brescia, Brescia, Italy
- Ron Keren, The University Health Network, Toronto Rehabilitation Institute, Toronto, Canada
- Rosa Rademakers, Center for Molecular Neurology, University of Antwerp
- Rose Bruffaerts, Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium
- Sabrina Sayah, Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France
- Sandra Black, Sunnybrook Health Sciences Centre, Sunnybrook Research Institute, University of Toronto, Toronto, Canada
- Sandra Loosli, Neurologische Klinik, Ludwig-Maximilians-Universität München, Munich, Germany
- Sara Mitchell, Sunnybrook Health Sciences Centre, Sunnybrook Research Institute, University of Toronto, Toronto, Canada
- Sara Prioni, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Sarah Anderl-Straub, Department of Neurology, University of Ulm, Ulm, Germany
- Serge Gauthier, Alzheimer Disease Research Unit, McGill Centre for Studies in Aging, Department of Neurology & Neurosurgery, McGill University, Montreal, Québec, Canada
- Sónia Afonso, Instituto Ciencias Nucleares Aplicadas a Saude, Universidade de Coimbra, Coimbra, Portugal
- Sonja Schönecker, Neurologische Klinik, Ludwig-Maximilians-Universität München, Munich, Germany
- Stefano Gazzina, Neurophysiology Unit, ASST Spedali Civili, Brescia, Italy
- Thibaud Lebouvier, Univ Lille, France
- Thomas Cope, Department of Clinical Neuroscience, University of Cambridge, Cambridge, UK
- Timothy Rittman, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
- Tobias Hoegen, Neurologische Klinik, Ludwig-Maximilians-Universität München, Munich, Germany
- Valentina Bessi, Department of Neuroscience, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy
- Valentina Cantoni, Centre for Neurodegenerative Disorders, Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy
- Veronica Redaelli, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano, Italy
- Vesna Jelic, Division of Clinical Geriatrics, Karolinska Institutet, Stockholm, Sweden
- Vincent DeramecourtUniv Lille, France
- Vittoria Borracci, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Neurodegenerative Diseases Unit, Milan, Italy

Acknowledgements The authors thank all the volunteers for their participation in this study. SBE is a recipient of the Joan Rodés Josep Baselga grant from the FBBVA. This study was partially funded by Fundació Marató de TV3, and Instituto de Salud Carlos III, Spain (grant no. 20143810 and PI20/0448 to RSV). The GENFI study has been supported by the Medical Research Council UK, the Italian Ministry of Health and the Canadian Institutes of Health Research as part of a Centres of Excellence in Neurodegeneration grant, as well as other individual funding to investigators. KM has received funding from an Alzheimer's Society PhD studentship. JDR acknowledges support from the National Institute for Health Research (NIHR) Queen Square Dementia Biomedical Research Unit and the University College London Hospitals Biomedical Research Centre, the Leonard Wolfson Experimental Neurology Centre, the UK Dementia Research Institute, Alzheimer's Research UK, the Brain Research Trust and the Wolfson Foundation. JCvS was supported by the Dioraphte Foundation grant 09-02-03-00, the Association for Frontotemporal Dementias Research Grant 2009, The Netherlands Organization for Scientific Research (NWO) grant HCMI 056-13-018, ZonMw Memorabel (Deltaplan Dementie, project number 733 051 042), Alzheimer Nederland and the Bluefield project. CG have received funding from JPND-Prefrontals VR Dnr 529-2014-7504, VR: 2015-02926, and 2018-02754, the Swedish FTD Initiative-Schörling Foundation, Alzheimer Foundation, Brain Foundation and Stockholm County Council ALF. DG has received support from the EU Joint Programme – Neurodegenerative Disease Research (JPND) and the Italian Ministry of Health (PreFrontALS) grant 733051042. JBR is funded by the Wellcome Trust (103838) and the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. MM has received funding from a Canadian Institutes of Health Research operating grant and the Weston Brain Institute and Ontario Brain Institute. RV has received funding from the Mady Browaeys Fund for Research into Frontotemporal Dementia. EF has received funding from a CIHR grant #327387. JDR is an MRC Clinician Scientist (MR/M008525/1) and has received funding from the NIHR Rare Diseases Translational Research Collaboration (BRC149/NS/MH), the Bluefield Project and the Association for Frontotemporal Degeneration. MS was supported by a grant 779257 "Solve-RD" from the Horizon 2020 research and innovation programme.

Funding The funding sources have no role in the design of this study, its execution, analyses, interpretation of the data, or the decision to submit results.

Data availability statement The dataset analyzed for the current study is from the research consortia GENFI. Data will be shared according to the GENFI data sharing agreement, after review by the GENFI data access committee with final approval granted by the GENFI steering committee.

Declarations

Conflicts of interest JBR reports consultancy for Asceneuron, Biogen, UCB, and SV Healthcare and research grants from Janssen, Lilly, and AZ-Medimmune. JL reports speakers fees from Bayer Vital, consulting fees from Axon Neuroscience, nonfinancial support from Abbvie, compensation for part time CMO from MODAG, author fees from Thieme medical publishers and from W. Kohlhammer GbmH medical publishers, all outside the submitted work. JDR Rohrer has served as a consultant for Biogen, Ionis, Alector, Wave Life Sciences, and Astex. RSV has served in Advisory boards Meetings for Wave Life Sciences, Ionis and Novo Nordisk and received personal fees for participating in educational activities from Janssen, Roche Diagnostics and Neuroxpharma and funding to her institution for research projects from Biogen and Sage Pharmaceuticals. The other authors report no disclosures relevant to the manuscript.

References

1. Bang J, Spina S, Miller BL (2015) Frontotemporal dementia. Lancet 386:1672–1682. https://doi.org/10.1016/S0140-6736(15)00461-4
2. Hardiman O, Al-Chalabi A, Chio A et al (2017) Amyotrophic lateral sclerosis. Nat Rev Dis Primers. https://doi.org/10.1038/nrdp.2017.71
3. Burrell JR, Halliday GM, Kril JJ et al (2016) The frontotemporal dementia-motor neuron disease continuum. Lancet 388:919–931. https://doi.org/10.1016/S0140-6736(16)00737-6
4. Borrego-Écija S, Turon-Sans J, Ximelis T et al (2021) Cognitive decline in amyotrophic lateral sclerosis: Neuropathological substrate and genetic determinants. Brain Pathol 31:e12942. https://doi.org/10.1111/BPA.12942
5. Rademakers R, Neumann M, MacKenzie IR (2012) Advances in understanding the molecular basis of frontotemporal dementia. Nat Rev Neurol 8:423–434. https://doi.org/10.1038/NRNEUROL.2012.117
6. Cerami C, Marcone A, Crespi C et al (2015) Motor neuron dysfunctions in the frontotemporal lobar degeneration spectrum: a clinical and neurophysiological study. J Neurol Sci 351:72–77. https://doi.org/10.1016/J.JNS.2015.02.039
7. DeJesus-Hernandez M, Mackenzie IR, Boeve BF et al (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72:245–256. https://doi.org/10.1016/j.neuron.2011.09.011
8. Renton AE, Majounie E, Waite A et al (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72:257–268. https://doi.org/10.1016/j.neuron.2011.09.010
9. Moore KM, Nicholas J, Grossman M et al (2020) Age at symptom onset and death and disease duration in genetic frontotemporal dementia: an international retrospective cohort study. Lancet Neurol 19:145–156. https://doi.org/10.1016/S1474-4422(19)30394-1
10. Rohrer JD, Warren JD, Fox NC, Rossor MN (2013) Presymptomatic studies in genetic frontotemporal dementia. Revue Neurologique 169:820–824. https://doi.org/10.1016/j.neurol.2013.07.010
11. Rohrer JD, Nicholas JM, Cash DM et al (2015) Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the Genetic Frontotemporal dementia Initiative (GENFI) study: a cross-sectional analysis. Lancet Neurol 14:253–262. https://doi.org/10.1016/S1474-4422(14)70324-2
12. Walhout R, Schmidt R, Westeneng H-J et al (2015) Brain morphologic changes in asymptomatic C9orf72 repeat expansion carriers. Neurology 85:1780–1788. https://doi.org/10.1212/WNL.0000000000002135
13. Cash DM, Bocchetta M, Thomas DL et al (2018) Patterns of gray matter atrophy in genetic frontotemporal dementia: results from the GENFI study. Neurobiol Aging 62:191–196. https://doi.org/10.1016/j.neurobiolaging.2017.10.008
14. Bertrand A, Wen J, Rinaldi D et al (2018) Early cognitive, structural, and microstructural changes in presymptomatic C9orf72 carriers younger than 40 years. JAMA Neurol 75:236–245. https://doi.org/10.1001/JAMANEUROL.2017.4266
15. Panman JL, Jiskoot LC, Bouts MJRJ et al (2019) Gray and white matter changes in presymptomatic genetic frontotemporal dementia: a longitudinal MRI study. Neurobiol Aging 76:115–124. https://doi.org/10.1016/j.neurobiolaging.2018.12.017
16. Bocchetta M, Todd EG, Peakman G et al (2021) Differential early subcortical involvement in genetic FTD within the GENFI cohort. NeuroImage: Clin. https://doi.org/10.1016/j.nicl.2021.102646
17. Papma JM, Jiskoot LC, Panman JL et al (2017) Cognition and gray and white matter characteristics of presymptomatic C9orf72 repeat expansion. Neurology 89:1256–1264. https://doi.org/10.1212/WNL.0000000000004393
18. Jiskoot LC, Bocchetta M, Nicholas JM et al (2018) Presymptomatic white matter integrity loss in familial frontotemporal dementia in the GENFI cohort: a cross-sectional diffusion tensor imaging study. Ann Clin Transl Neurol 5:1025–1036. https://doi.org/10.1002/ACN3.601
19. Brown RH, Al-Chalabi A (2017) Amyotrophic lateral sclerosis. N Engl J Med 377:162–172. https://doi.org/10.1056/NEJMra1603471
20. Grinberg LT, Rueb U, Heinsen H (2011) Brainstem: neglected locus in neurodegenerative diseases. Front Neurol. https://doi.org/10.3389/FNEUR.2011.00042
21. Bede P, Chipika RH, Finegan E et al (2019) Brainstem pathology in amyotrophic lateral sclerosis and primary lateral sclerosis: a longitudinal neuroimaging study. Neuroimage Clin 24:102054. https://doi.org/10.1016/j.nicl.2019.102054
22. Pioro EP, Turner MR, Bede P (2020) Neuroimaging in primary lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 21:18–27. https://doi.org/10.1080/21678421.2020.1837176
23. Querin G, Bede P, El Mendili MM et al (2019) Presymptomatic spinal cord pathology in c9orf72 mutation carriers: a longitudinal neuroimaging study. Ann Neurol 86:158–167. https://doi.org/10.1002/ANA.25520
24. Rascovsky K, Hodges JR, Knopman D et al (2011) Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 134:2456–2477. https://doi.org/10.1093/brain/awr179
25. Gorno-Tempini ML, Hillis AE, Weintraub S et al (2011) Classification of primary progressive aphasia and its variants. Neurology 76:1006–1014. https://doi.org/10.1212/WNL.0B013E31821103E6
26. Strong MJ, Abrahams S, Goldstein LH et al (2017) Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): revised diagnostic criteria. Amyotroph Lateral Scler Frontotemporal Degener 18:153–174. https://doi.org/10.1080/21678421.2016.1267768

27. Brooks BR (1994) El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. J Neurol Sci 124(Suppl):96–107. https://doi.org/10.1016/0022-510x(94)90191-0
28. Miyagawa T, Brushaber D, Syrjanen J et al (2020) Use of the CDR® plus NACC FTLD in mild FTLD: data from the ARTFL/LEFFTDS consortium. Alzheimer's Dement 16:79–90. https://doi.org/10.1016/J.JALZ.2019.05.013
29. Cedarbaum JM, Stambler N, Malta E et al (1999) The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. J Neurol Sci 169:13–21. https://doi.org/10.1016/S0022-510X(99)00210-5
30. Reuter M, Schmansky NJ, Rosas HD, Fischl B (2012) Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage 61:1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084
31. Fischl B, Salat DH, van der Kouwe AJW et al (2004) Sequence-independent segmentation of magnetic resonance images. Neuroimage 23(Suppl 1):S69-84. https://doi.org/10.1016/j.neuroimage.2004.07.016
32. Fischl B, Salat DH, Busa E et al (2002) Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33:341–355. https://doi.org/10.1016/s0896-6273(02)00569-x
33. Iglesias JE, Van Leemput K, Bhatt P et al (2015) Bayesian segmentation of brainstem structures in MRI. Neuroimage 113:184–195. https://doi.org/10.1016/j.neuroimage.2015.02.065
34. Bocchetta M, Iglesias JE, Chelban V et al (2020) Automated brainstem segmentation detects differential involvement in atypical Parkinsonian syndromes. J Mov Disord 13:39–46. https://doi.org/10.14802/jmd.19030
35. Bocchetta M, Malpetti M, Todd EG et al (2021) Looking beneath the surface: the importance of subcortical structures in frontotemporal dementia. Brain Commun 3:fcab158. https://doi.org/10.1093/braincomms/fcab158
36. Salamon N, Sicotte N, Alger J et al (2005) Analysis of the brainstem white-matter tracts with diffusion tensor imaging. Neuroradiology 47:895–902. https://doi.org/10.1007/S00234-005-1439-8
37. Ford A, Colon-Perez L, Triplett WT et al (2013) Imaging white matter in human brainstem. Front Hum Neurosci. https://doi.org/10.3389/FNHUM.2013.00400

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Agnès Pérez-Millan[1,2] · Sergi Borrego-Écija[1] · John C. van Swieten[3] · Lize Jiskoot[3,4] · Fermin Moreno[5,6] · Robert Laforce[7] · Caroline Graff[8,9] · Mario Masellis[10] · Maria Carmela Tartaglia[11] · James B. Rowe[12] · Barbara Borroni[13] · Elizabeth Finger[14] · Matthis Synofzik[15,16] · Daniela Galimberti[17,18] · Rik Vandenberghe[19,20] · Alexandre de Mendonça[21] · Chris R. Butler[22,23] · Alexander Gerhard[24,25] · Simon Ducharme[26,27] · Isabelle Le Ber[28,29] · Isabel Santana[30,31] · Florence Pasquier[32,33] · Johannes Levin[34,35] · Markus Otto[36] · Sandro Sorbi[37] · Pietro Tiraboschi[38] · Harro Seelaar[3] · Tobias Langheinrich[39] · Jonathan D. Rohrer[4] · Roser Sala-Llonch[2,40] · Raquel Sánchez-Valle[1] · The Genetic FTD Initiative, GENFI

[1] Alzheimer's Disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer, University of Barcelona, Villarroel, 170, 08036 Barcelona, Spain

[2] Department of Biomedicine, Faculty of Medicine, Institute of Neurosciences, University of Barcelona, 08036 Barcelona, Spain

[3] Department of Neurology and Alzheimer Center Erasmus MC, Erasmus MC University Medical Center, Rotterdam, The Netherlands

[4] Department of Neurodegenerative Disease, Dementia Research Centre, UCL Institute of Neurology, Queen Square, London, UK

[5] Cognitive Disorders Unit, Department of Neurology, Donostia University Hospital, San Sebastian, Gipuzkoa, Spain

[6] Neuroscience Area, Biodonostia Health Research Institute, San Sebastian, Gipuzkoa, Spain

[7] Département des Sciences Neurologiques, Clinique Interdisciplinaire de Mémoire, CHU de Québec, and Faculté de Médecine, Université Laval, Quebec City, QC, Canada

[8] Division of Neurogeriatrics, Department of Neurobiology, Care Sciences and Society, Bioclinicum, Center for Alzheimer Research, Karolinska Institutet, Solna, Sweden

[9] Unit for Hereditary Dementias, Theme Aging, Karolinska University Hospital, Solna, Sweden

[10] Sunnybrook Health Sciences Centre, Sunnybrook Research Institute, University of Toronto, Toronto, Canada

[11] Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Canada

[12] Department of Clinical Neurosciences and Cambridge University Hospitals NHS Trust and Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

[13] Department of Clinical and Experimental Sciences, Centre for Neurodegenerative Disorders, University of Brescia, Brescia, Italy

[14] Department of Clinical Neurological Sciences, University of Western Ontario, London, ON, Canada

[15] Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research and Center of Neurology, University of Tübingen, Tübingen, Germany

[16] Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

[17] Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy

[18] Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy

[19] Laboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, Leuven, Belgium

[20] Neurology Service, University Hospitals Leuven, Leuven, Belgium

[21] Faculty of Medicine, University of Lisbon, Lisbon, Portugal

[22] Nuffield Department of Clinical Neurosciences, Medical Sciences Division, University of Oxford, Oxford, UK

[23] Department of Brain Sciences, Imperial College London, London, UK

[24] Division of Neuroscience and Experimental Psychology, Wolfson Molecular Imaging Centre, University of Manchester, Manchester, UK

[25] Department of Geriatric Medicine and Nuclear Medicine, Center for Translational Neuro- and Behavioral Sciences, University Medicine Essen, Essen, Germany

[26] Department of Psychiatry, Douglas Mental Health University Institute, McGill University, Montreal, Canada

[27] Department of Neurology and Neurosurgery, McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada

[28] Sorbonne Université, Paris Brain Institute – Institut du Cerveau – ICM, Inserm U1127, CNRS UMR 7225, AP-HP - Hôpital Pitié-Salpêtrière (DMU Neurosciences Paris 6), Paris, France

[29] Département de Neurologie, AP-HP - Hôpital Pitié-Salpêtrière (DMU Neurosciences Paris 6), Paris, France

[30] Neurology Service, Faculty of Medicine, University Hospital of Coimbra (HUC), University of Coimbra, Coimbra, Portugal

[31] Center for Neuroscience and Cell Biology, Faculty of Medicine, University of Coimbra, Coimbra, Portugal

[32] Univ Lille, Lille, France

[33] CHU, CNR-MAJ, Labex Distalz, LiCEND, Lille, France

[34] Neurologische Klinik und Poliklinik, Ludwig-Maximilians-Universität, Munich, Germany

[35] German Center for Neurodegenerative Diseases (DZNE), Munich, Germany

[36] Department of Neurology, University of Ulm, Ulm, Germany

[37] Department of Neurofarba, University of Florence, Florence, Italy

[38] Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy

[39] Division of Neuroscience and Experimental Psychology, Wolfson Molecular Imaging Centre, University of Manchester, Manchester, UK

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

4. 1 INFORME DE LES DIRECTORES

Aquesta tesis doctoral per compendi d'articles ha estat realitzada sota la nostra direcció i considerem que reuneix les condicions necessàries. Inclou sis articles, dels quals 4 ja han estat publicats o acceptats i els altres estan sota revisió.

Títol	Nom de la revista; DOI	Any	IF (2022); Quartil	Contribució personal
Evaluating the performance of Bayesian and frequentist approaches for longitudinal modeling: application to Alzheimer's disease	Scientific Reports; 10.1038/s41598-022-18129-4	2022	4,6; Q2	Disseny de l'estudi, recol·lecció de les dades, anàlisi dels resultats, interpretació dels resultats, redacció i revisió del manuscrit
Cortical thickness modeling and variability in Alzheimer's disease and frontotemporal dementia.	Acceptat en Journal of Neurology	---	6,0; Q1	Disseny de l'estudi, recol·lecció de les dades, processament de les imatges de RM, anàlisi dels resultats, interpretació dels resultats, redacció i revisió del manuscrit
Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data	Human Brain Mapping; 10.1002/hbm.26205	2023	4,8; Q1	Disseny de l'estudi, recol·lecció de les dades, processament de les imatges de RM, anàlisi dels resultats, interpretació dels resultats, redacció i revisió del manuscrit

Beyond group classification: probabilistic differential diagnosis of frontotemporal dementia and Alzheimer's disease with MRI and CSF biomarkers	En revisió	---	---	Disseny de l'estudi, recol·lecció de les dades, processament de les imatges de RM, anàlisi dels resultats, interpretació dels resultats, redacció del manuscrit
The Cortical Asymmetry Index (CAI) for subtyping dementia patients	En revisió	---	---	Disseny de l'estudi, recol·lecció de les dades, processament de les imatges de RM, anàlisi dels resultats, interpretació dels resultats, redacció del manuscrit
Loss of brainstem white matter predicts onset and motor neuron symptoms in *C9orf72* expansion carriers: a GENFI study	Journal of Neurology; D10.1007/s00415-022-11435-x	2023	6,0; Q1	Disseny de l'estudi, recol·lecció de les dades, processament de les imatges de RM, anàlisi dels resultats, interpretació dels resultats, redacció i revisió del manuscrit

El sisè treball ha estat realitzat en contribució equivalent pels dos primers autors (Agnès Pérez Millan i el Dr. Sergi Borrego Écija). Cap dels treballs que formen la present tesi han estat utilitzats implícitament o explícitament en cap altra tesi.

Signat,

Dra. Roser Sala Llonch

Dra. Raquel Sánchez-Valle

4.2 RESUM GLOBAL DELS RESULTATS

Primer treball:

En aquest treball explorem el volum de l'hipocamp amb pacients de la base de dades longitudinals ADNI (adni.loni.usc.edu). Avaluem quina aproximació estadística, la freqüentista o la bayesiana, és l'òptima per estudiar models longitudinals tant en bases de dades homogènies (tots els pacients tenen el mateix nombre de vistes) com heterogènies (manquen visites d'alguns pacients). Per això, creem models d'efectes lineals mixts (ELM) en els dos marcs estadístics. En una base heterogènia en què no tots els pacients tenen el mateix nombre de visites, però amb un nombre molt elevat de pacients (1250 participants), trobem que els dos enfocaments són capaços d'identificar els mateixos grups de pacients. Així, el comportament que presenten les dues aproximacions és molt similar. Tot i això, l'enfocament bayesià necessita molt més poder computacional per poder estudiar les trajectòries d'aquests pacients.

Tanmateix, si la base de dades no és tan extensa, les dues aproximacions estadístiques ja no es comporten de la mateixa manera. Per estudiar aquest fet partim d'una base de dades homogènia extensa (N=670 participants, 4 visites cada un). Posteriorment, realitzem dues simulacions mitjançant l'eliminació progressiva de dades. En la primera, per estudiar el mínim nombre de subjectes que permeti separar els grups, trobem que l'enfocament bayesià va ser lleugerament més estricte. En la segona, per avaluar l'efecte de la manca de punts, l'aproximació bayesiana presenta un poder estadístic superior a la freqüentista, ja que a mesura que es redueixen el nombre de visites de forma aleatòria, arriba un punt que l'aproximació freqüentista no pot estimar el model. En canvi, l'aproximació bayesiana, no té aquesta limitació. Així, l'enfocament bayesià és més robust per a bases de dades desequilibrades i disperses amb un nombre diferent de mesures entre subjectes. En canvi, l'aproximació freqüentista és computacionalment més senzilla, però falla en la modelització de dades amb un gran nombre de valors que falten.

Segon treball:

En aquest treball hem implementat models d'estimació del gruix cortical segons l'edat cronològica per persones sanes, i segons els anys de durada de la malaltia per pacients amb MA i DFT. Els models han estat creats amb una regressió de suport vectorial (RSV) i amb dades de RM estructural de la Unitat d'Alzheimer i altres trastorns cognitius de l'Hospital Clínic de Barcelona (379 pacients amb una visita i 144 pacients amb 2-3 visites). Amb aquests models podem identificar les variacions individuals com a desviacions de la norma, utilitzant models específics per persones sanes, MA i DFT. Amb el model generat amb persones sanes, veiem que el model s'adapta al seu propi grup, però que els pacients amb MA i DFT presenten una variabilitat molt elevada. A més, les desviacions segons aquest model per pacients amb MA correlacionen amb les puntuacions dels anys de durada de la malaltia i MMSE, mentre que els pacients amb DFT no presenten correlacions.

Posteriorment, usem els models específics per la MA i la DFT per estudiar les desviacions del model dins de cada grup en relació amb biomarcadors bioquímics i les puntuacions del MMSE. Per als pacients amb DFT, trobem una correlació inversa significativa entre els nivells de LCR-NfL i la desviació del model a la majoria de les regions del cervell, proporcionant evidències addicionals de la relació entre la imatge i els biomarcadors bioquímics en DFT. A més, trobem una correlació positiva per als pacients amb MA entre les puntuacions MMSE i la desviació del model en diverses regions.

Finalment, utilitzem les dades longitudinals per avaluar la capacitat dels models per obtenir prediccions individuals dels valors futurs del gruix cortical. Trobem que els models de persones sanes i MA, són capaços de predir les futures vistes, en canvi, el model de DFT mostra una elevada variabilitat.

Tercer treball:

En aquest treball hem implementat un algoritme de classificació automàtica per diferenciar pacients amb MA i DFT i persones sanes amb RM estructural. La combinació d'un primer algoritme on es redueixen el nombre de les variables segons una aproximació no supervisada i posteriorment un algoritme de classificació automàtic supervisat ens permet classificar els grups amb dades transversals i longitudinals. L'algoritme ha estat entrenat i testejat amb dades de la Unitat d'Alzheimer i altres trastorns cognitius de l'Hospital Clínic de Barcelona (339 pacients amb una visita i 114 pacients amb 2 visites).

L’algorisme implementat classifica a nivell transversal els pacients sans versus MA i versus DFT amb un 83,3% i un 82,1% de precisió respectivament. Aquesta precisió augmenta fins a un 90,0% i un 88,0% respectivament amb dades longitudinals. Quan estudiem la classificació entre MA versus DFT aconseguim una precisió del 63,3% a nivell transversal i del 75,0% a nivell longitudinal. Finalment, la classificació MA versus DFT versus persones sanes assoleix una precisió del 60,7% i del 71,3% per a les dades transversals i longitudinals, respectivament. Per tant, observem que l'enfocament longitudinal proporciona un rendiment superior al transversal, assolim millors precisions en el diagnòstic automàtic per tots els casos.

Quart treball:

Aquest treball consisteix a implementar un algoritme de classificació automàtica que retorna el diagnòstic i les probabilitats individuals associades a cada grup. El treball s’ha fet amb dades transversals de RM estructural i biomarcadors de LCR de la base de dades de la Unitat d'Alzheimer i altres trastorns cognitius de l'Hospital Clínic de Barcelona que inclou 491 pacients (persones amb MA, persones amb DFT i persones sanes). L’algoritme és capaç de discriminar amb èxit els pacients amb DFT i MA amb bona precisió, i permet identificar els subtipus de variants clíniques de la DFT. Només amb dades de RM, obtenim precisions del 88% en la classificació MA versus persones sanes, 87% per DFT versus persones sanes, 82% per MA versus DFT i 80% en diferenciar els tres grups clínics. Un total del 74% dels participants amb DFT i el 73% dels participants amb

MA tenen una probabilitat alta (≥80%) de tenir un diagnòstic precís en la comparació de DFT versus MA.

Les dades dels nivells de NfL i 14-3-3 en LCR s'estudien primer de forma independent i en combinació amb dades de RM. Els resultats que s'extreuen d'aquesta informació depenen de cada una de les comparacions. Quan es comparen els pacients amb MA respecte persones sanes, els nivells de LCR són suficients per assolir un diagnòstic precís. En el cas dels pacients amb DFT versus les persones sanes, tant les dades solament de RM o LCR, com la combinació d'ambdues modalitats donen resultats similars. Finalment, per diferenciar pacients amb MA versus DFT, la combinació de les dades de RM i LCR, és la que aconsegueix el diagnòstic més precís i el nombre més nombrós de pacients amb una probabilitat major del 80% de tenir un encert en el diagnòstic.

Cinquè treball:

En aquest treball mesurem les asimetries del còrtex cerebral amb un índex que nomenem Índex d'Asimetria Cortical (CAI). L'estudi es realitza amb dades de la Unitat d'Alzheimer i altres trastorns cognitius de l'Hospital Clínic de Barcelona que inclou 554 pacients (persones amb DFT, persones amb MA i persones sanes). El CAI és una mesura derivada de la teoria de la informació que ens permet obtenir com és d'asimètric el cervell de manera individual. El CAI ens permet diferenciar les persones sanes dels participants amb demència, seguidament la DFT de la MA, i finalment els diferents fenotips de la DFT. A més a més, longitudinalment permet identificar els canvis en l'asimetria cerebral al llarg del temps pels pacients amb DFT. Finalment, duem a terme unes anàlisis de clústers amb els valors del CAI pels pacients amb DFT i per MA. Identifiquem dos clústers per cada demència, que presenten diferències significatives amb nivells de biomarcadors de líquid cefalorraquidi i de sang. Els pacients amb DFT s'agrupen en dos clústers, un estaria format principalment pels pacients amb la variant semàntica i l'altre per la variant no fluent i la conductual. Els dos clústers presenten diferències en els nivells de NfL tant en LCR com en sang. Els pacients amb MA mostren dos clústers també, que a priori no hem sigut capaços d'interpretar des de la clínica, però que es diferencien amb els nivells de GFAP en sang.

Sisè treball:

En aquest treball utilitzem la RM estructural per mesurar la substància blanca de les estructures del tronc encefàlic dels portadors de la mutació C9orf72. Les dades d'aquest estudi de DFT genètica són del consorci internacional i multicèntric GENFI (https://www.genfi.org/), i formen part de l'estudi 102 portadors asimptomàtics de la mutació C9orf72, 72 portadors simptomàtics de la mutació C9orf72 i 75 no portadors. Estudiem les diferències que puguin presentar les diferents estructures del tronc encefàlic per aquests pacients i trobem que amb aquestes estructures podem diferenciar els diferents pacients, però no les persones sanes de les asimptomàtiques, però sí que es diferencien els pacients amb DFT i els pacients amb DFT que presenten símptomes de neurones motores. També realitzem regressions lineals per avaluar les interaccions entre els diferents grups clínics i l'edat. En aquest cas trobem que els pacients amb símptomes de neurones motores presenten més pèrdua de substància blanca al tronc cerebral que els que no tenien símptomes motors. Finalment, els portadors de l'expansió C9orf72 simptomàtics mostren alteracions consistents en la substància blanca del tronc cerebral que es correlaciona amb la gravetat clínica.

CAPÍTOL 5

Discussió

Els treballs que conformen la present tesi doctoral pretenen aportar informació de les possibilitats i limitacions dels models matemàtics amb dades de RM estructural per l'estudi de la MA i la DFT. També volen ampliar el coneixement sobre els marcadors útils pel diagnòstic diferencial d'aquestes demències i les seves variants clíniques, i estudiar la complementarietat entre els biomarcadors d'imatge i els biomarcadors bioquímics. Així mateix, la present tesi doctoral pretén aprofundir en l'heterogeneïtat clínica que presenten aquestes demències estudiant la variabilitat individual dels pacients amb MA i DFT respecte de la tendència grupal en funció del temps.

En el primer treball (104) vam avaluar si l'homogeneïtat o heterogeneïtat de les dades longitudinals en bases de dades extenses en neuroimatge té alguna implicació a l'hora de realitzar l'estudi estadístic de les trajectòries. Per aquest estudi ens vam centrar en la MA, tot i que els resultats es poden extrapolar a altres demències o malalties. Això ho vam fer, ja que teníem accés a una base de dades molt extensa com és ADNI amb 1250 pacients amb 5 visites. I alhora, podíem utilitzar el volum de l'hipocamp, que és un biomarcador de MA d'imatge en RM estructural que ha estat ben descrit i és acceptat en els actuals criteris de diagnòstic de la MA (16,17,105–107), així que ens permetia centrar-nos en l'aproximació estadística. Per això, vam crear models ELM segons l'aproximació freqüentista i la bayesiana per estudiar les diferències entre els grups al llarg del temps i explorar quina aproximació és la més recomanable en situacions diverses. En primer lloc, vam trobar que els dos enfocaments coincidien en els resultats de les proves d'hipòtesis nul·les amb tot el conjunt de dades amb els prèviament descrits en la literatura amb dades ADNI (60,67,108). En aquest cas, la diferència que presentarien les dues aproximacions seria el cost computacional, ja que l'enfocament bayesià va necessitar molt més poder computacional i temps per poder estudiar les trajectòries respecte a l'aproximació freqüentista. Ara bé, s'ha de valorar la diferent interpretació que presenten cada una de les aproximacions, no solament el cost de l'estimació del model. L'aproximació bayesiana proporciona un marc més interpretable, ja que es basa en l'estimació directa a partir de la distribució de la població representada per la distribució posterior, en lloc d'estimar a partir de la distribució mostral hipotètica com passa en l'enfocament freqüentista. Així que, en entorns clínics l'aproximació bayesiana pot oferir una interpretació més directa, perquè els resultats es poden interpretar de manera probabilística, contràriament a l'aproximació freqüentista que no accepta aquesta interpretació encara que molts investigadors la utilitzen (71,73). Tanmateix, si la base de dades no és tan extensa i presenta molts pacients als quals els hi manquen visites, les dues aproximacions estadístiques ja no es comporten de la mateixa manera. En aquest sentit, un inconvenient important per al modelatge ELM segons un enfocament freqüentista és la necessitat de tenir més mostres de participants que efectes aleatoris per estimar el model. En canvi, l'enfocament bayesià permet estimar el model fins i tot amb un nombre elevat de punts que falten a la base de dades. Més concretament, els nostres resultats van mostrar que el model ELM segons un enfocament bayesià és possible en una base de dades de 4 punts de temps que té aproximadament 2 valors que falten per a cada participant, cosa que suggereix que per a la modelització longitudinal en bases de dades escasses s'hauria de triar el marc bayesià. Altres estudis han demostrat que l'estadística bayesiana supera algunes de les limitacions de la inferència estadística clàssica en bases de dades no homogènies (68,109). En conjunt, destaquem la necessitat de conèixer les característiques de la mostra de dades per poder inferir la correcta interpretació dels resultats. Així que, en l'estudi vam proposar que l'enfocament bayesià és més robust per a bases de dades desequilibrades i disperses amb un nombre diferent de mesures entre subjectes.

En el segon treball (acceptat) ens vam centrar a implementar models específics grupals per MA, DFT i persones sanes del gruix cortical segons l'edat cronològica per a persones sanes i segons els anys de malaltia per a pacients amb MA i DFT, per estudiar les desviacions que presenten respecte a les persones sanes i la tendència grupal. Donat que en el primer treball, havíem observat, igual que altres autors (110–112) que un model basat en probabilitats presenta més avantatges, en aquest segon treball vam seguir en aquesta línia i el model del gruix cortical és estimat amb una RSV, que deriva dels models d'aprenentatge automàtic d'intel·ligència artificial que té una base probabilística.

Amb aquests models vam realitzar un estudi de la variabilitat que presenta cada individu respecte de la tendència grupal i vam associar aquestes desviacions amb els símptomes en el cas dels pacients. Pels pacients amb MA, vam trobar que les desviacions individuals del model correlacionaven amb les puntuacions del MMSE, sobretot en els lòbuls temporal i parietal, on les desviacions individuals més elevades del model s'associaven als individus amb puntuacions més elevades de MMSE. La relació entre el MMSE i el gruix cortical ha estat ben descrita prèviament (113–115). En el nostre treball, vam poder identificar que els pacients amb puntuacions més elevades de MMSE són aquells en els que el gruix cortical es desvia més de l'esperat segons els anys de malaltia. En el cas dels pacients amb DFT, vam trobar una correlació inversa entre els nivells de NfL en LCR i la desviació del model, proporcionant evidències addicionals a les ja descrites per altres autors sobre la relació entre la imatge i els biomarcadors de fluids per aquests pacients. Per exemple, els nivells de NfL s'han suggerit com a factor de gravetat per la DFT (38,39). Els nivells alts de NfL van estar associats a les desviacions del model del gruix cortical més baixes. Per tant, els pacients amb patrons de gruix cortical que s'adapten bé al model tenen nivells més elevats de NfL en LCR, així que el model de gruix cortical pels pacients amb DFT va capturar la tendència de neurodegeneració. En canvi, els nivells més baixos de NfL corresponien a pacients que s'allunyaven més del model DFT, similar al que s'ha descrit com una DFT no progressiva (116). En estudiar les diferents variants clíniques de la DFT, trobem que la vcDFT és la que tindria més contribució en el model, ja que les altres variants APP mostren una desviació més alta. Aquests resultats estan en la línia del que s'ha descrit prèviament, on la DFT és una demència heterogènia (30,36,98). A més, en altres estudis d'aquesta tesi, utilitzant algoritmes de classificació automàtica, també observem que la variant vcDFT té una contribució més important a l'hora de definir les tendències o patrons dels pacients amb DFT i on els patrons d'importància no coincideixen entre les diferents variants clíniques de la DFT. Tenint en compte les dues troballes, necessitaríem crear models específics per cada una de les variants per poder predir els canvis en el gruix cortical en detall per aquests pacients.

Els resultats del segon treball, tot i que estava centrat en l'estudi de la variabilitat, també ens van permetre predir el gruix cortical en visites futures segons l'edat per persones sanes i segons els anys de malaltia per pacients amb MA. En canvi, pels pacients amb DFT, no vam poder assolir unes prediccions fiables de les seves futures vistes a causa de l'alta variabilitat que presentaven els resultats.

Els resultats del segon estudi van mostrar que els pacients amb DFT i MA mostren una variabilitat individual tant a nivell transversal com longitudinal en les imatges de RM estructural que podrien ser útils per un diagnòstic personalitzat. Per això, en el tercer treball (117) que forma aquesta tesi ens vam centrar a estudiar la utilitat de l'aprenentatge automàtic per al diagnòstic diferencial de la DFT i la MA a nivell transversal i longitudinal amb dades de RM estructural, i alhora obtenir interpretabilitat dels resultats. Per realitzar aquesta classificació entre pacients vam utilitzar el classificador MVS, que és el model d'aprenentatge automàtic supervisat àmpliament utilitzat amb dades de neuroimatge per pacients amb MA i DFT (53,94,96,118–121). Però, aquest model no permet les classificacions amb dades longitudinals, és per això, que vam aplicar tècniques no supervisades abans de realitzar la classificació amb el MVS. Aquestes tècniques van ser una anàlisi de components principals (ACP) per l'estudi transversal i el seu anàleg l'anàlisi de múltiples factors (AMF) per les dades longitudinals (122). Això, ens va permetre fer la classificació tant a nivell transversal com longitudinal i poder obtenir uns resultats interpretables, ja que solament ens vam quedar amb la primera component, així que amb les contribucions d'aquesta component vam poder estudiar quines són les variables més importants per la classificació. Aquesta combinació d'algoritmes supervisats i no supervisats ens va permetre assolir uns nivells de precisió tant a nivell transversal com longitudinal alts. L'algoritme va classificar a nivell transversal els pacients sans versus MA i versus DFT amb un 83% i un 82% de precisió respectivament. Aquesta precisió va augmentar fins a un 90% i un 88% amb dades longitudinals, per MA i DFT respectivament. Quan vam estudiar la classificació entre MA versus DFT vam aconseguir una precisió del 63% a nivell transversal i del 75% a nivell longitudinal. Finalment, la classificació MA versus DFT versus persones sanes va assolir una precisió del 61% i del 71% per a les dades transversals i longitudinals, respectivament. Aquestes precisions estan dins del rang dels estudis prèviament publicats, o són superiors (94–96,123–125). Les diferències entre aquest estudi i els prèviament publicats per altres grups són que vam poder classificar dades longitudinals amb un MVS i millorar els resultats transversals, i que vam obtenir els patrons de les regions cerebrals més importants per la classificació. En aquest estudi, ens vam centrar a assolir un algoritme d'aprenentatge automàtic amb una alta precisió que permetés diferenciar els pacients amb MA, DFT i les persones sanes, però donant molta importància a la seva interpretabilitat, no ens vam focalitzar solament en maximitzar el rendiment. L'algoritme ens va permetre aconseguir un patró d'atròfia diferent per cada una de les malalties, amb els pesos de les tècniques no supervisades, identificant les regions més rellevants per realitzar el diagnòstic automàtic. En general, vam trobar que les regions més importants per la classificació coincidien amb els patrons patològics d'acord amb la literatura de la MA i la DFT (96,126). Tanmateix, aquests algoritmes cerquen interaccions robustes entre les diferents variables que s'introdueixen (en el nostre cas les regions cerebrals), per la qual cosa és possible que apareguin regions que no coincideixen amb els patrons diferencials de MA o DFT o patrons superposats. Treballs anteriors (55,127,128) també han trobat alguns patrons de degeneració superposats entre les dues demències.

La principal limitació del tercer treball és que no vam poder estudiar la fiabilitat amb què l'algoritme dona el resultat del diagnòstic. Vam observar que, més enllà del diagnòstic diferencial entre la MA i la DFT necessitem resultats per establir si els diagnòstics són fiables o aleatoris. Per això, en el quart estudi (no publicat) d'aquesta tesi, vam decidir implementar un altre algoritme, aquest cop a nivell transversal que ens permetés obtenir aquest diagnòstic diferencial de la MA i la DFT conjuntament amb la fiabilitat dels resultats, i que fos interpretable, com ja ho era el creat anteriorment. Així, vam desenvolupar mètodes de classificació calibrada i vam considerar que els participants amb una probabilitat major al 80% per algun dels diagnòstics tindrien una classificació fiable. Els pacients amb una probabilitat individual menor al 80%, estaven dins del que anomenem zona de baixa confiança del diagnòstic. Vam identificar aquests pacients com els que necessitarien més proves clíniques per poder realitzar el seu diagnòstic diferencial. I, per això, vam anar un pas més enllà i vam decidir avaluar si la RM i els nivells de LCR (NfL i 14-3-3, biomarcadors de neurodegeneració) millorarien la classificació. En aquest estudi, com en l'anterior, vam assolir precisions de l'orde, o millors, que les prèviament descrites en la literatura per discriminar la MA i la DFT amb dades de RM estructural. En el nostre cas, tenim 88% per MA versus persones sanes, 87% per DFT versus persones sanes, 82% per MA versus DFT i 80% per discriminar els tres grups clínics (53,62,93,96,127,129–133). A més, en aquest segon estudi de classificació automàtica per pacients amb MA i DFT, vam ser capaços de millorar els nostres propis resultats, mantenint la interpretabilitat i, d'altra banda, obtenint un valor de la fiabilitat diagnòstica dels resultats que en l'estudi previ no teníem. Ara bé, en l'estudi previ, podíem classificar amb dades longitudinals i vam observar que aquestes aportaven una informació que ens permetia millorar el nostre diagnòstic. Queda per explorar, doncs, la combinació dels estudis longitudinals amb els algoritmes de classificació calibrada per assolir nivells de confiança diagnòstica en dades longitudinals.

Entre els resultats clau d'aquest quart estudi, vam trobar que el 74% dels participants amb DFT i el 73% amb MA mostraven una probabilitat individual major al 80% d'estar ben classificats per l'algoritme en la comparació DFT versus MA. També, vam explorar la classificació automàtica per les variants clíniques de la DFT (vcDFT, vsAPP i vnfAPP) amb unes precisions superiors a les descrites prèviament (53,77,96). En primer lloc, pel que fa a les comparacions amb persones sanes, per a vcDFT, vam obtenir una precisió del 91%, i en el cas dels participants APP, una precisió del 93%. Quan vam classificar vcDFT i APP per separat contra MA, vam aconseguir precisions de fins al 90% per als dos casos. Tanmateix, quan vam intentar classificar vcDFT versus APP, assolim una precisió del 68% que no seria acceptable des del punt de vista clínic.

En aquest quart treball, com vam realitzar en el tercer treball d'aquesta tesi, obtenim uns patrons que ens indiquen la importància de cada regió en la classificació, tot i que en aquest cas han estat mesurats amb un test de permutacions, ja que és un algoritme diferent del descrit prèviament. El volum de la substància grisa de l'hipocamp, el putamen i l'amígdala van ser essencials per diferenciar MA de persones sanes. Per altra banda, les regions corticals van ser les més importants per diferenciar la DFT de persones sanes, especialment les regions occipitals, parietals i frontals. Aquests resultats estan d'acord amb les troballes amb mètodes d'anàlisi més clàssics (19,40,134–137) i amb els estudis previs que s'han realitzat en aquesta tesi. Finalment, pel que fa a les variants clíniques de la DFT, les regions frontals van ser les més destacades per la vcDFT i les temporals i l'hipocamp per la variant APP, també ha estat descrit prèviament (134,136,137).

Finalment, en aquest quart treball vam observar que la combinació de la RM i els nivells de LCR, solament seria necessària per al diagnòstic diferencial de la MA i la DFT. En canvi, quan vam estudiar el diagnòstic versus una persona sana, per la MA, la RM no era necessària i per la DFT, amb una de les dues proves era suficient per obtenir la mateixa fiabilitat en el diagnòstic. Aquesta proposta d'algoritme ens permet no solament ajudar en el diagnòstic diferencial dels pacients, sinó també en avaluar quines proves serien necessàries per a cada pacient individualment.

En els treballs 3 i 4, vam crear uns algoritmes d’aprenentatge automàtic interpretables. En aquests, en estudiar les regions més importants, vam observar que no són patrons simètrics. Per tant, en el cinquè treball (no publicat), vam estudiar l’asimetria cerebral per avaluar si podia ajudar en el diagnòstic diferencial de la DFT i la MA o, fins i tot, en la diferenciació de les diferents variants de la DFT (30,98,138–141). Hi ha treballs anteriors que estudien les asimetries cerebrals en DFT i MA, però majoritàriament avaluen l’asimetria de manera visual. Tot i que s'han fet alguns estudis per quantificar les asimetries de manera sistemàtica (138,142–144), són reduïts. Nosaltres vam escollir una metodologia més complexa basada en la teoria de la informació per obtenir l'estimació d'un índex asimètric. Aquest índex d'asimetria cerebral proposat es basa en l'entropia de les imatges de RM que nomenen Índex d'Asimetria Cortical (CAI). Aquesta mesura ens va permetre diferenciar les persones sanes de les que pateixen demència (MA i DFT), les dues demències i els fenotips de la DFT, mostrant que la vsAPP és la més asimètrica. Quan vam estudiar l’asimetria al llarg del temps, vam trobar un augment d’aquesta pels pacients amb DFT, en canvi, en la MA no vam detectar cap diferència al llarg del temps en aquesta mesura.

Una anàlisi no supervisat de clústers, ens va permetre agrupar els pacients amb DFT i MA en dos subgrups segons la seva asimetria cerebral. Al ser un estudi no supervisat, vam voler donar una explicació interpretable dels resultats mitjançant dades clíniques i de biomarcadors. Entre els dos subgrups de pacients amb DFT, el clúster més asimètric agrupava els pacients amb vsAPP i l'altre clúster amb els pacients amb vcDFT i vnfAPP. Aquests dos grups alhora es diferenciaven segons els nivells de NfL en LCR i sang. Això ens va fer suggerir que els nivells de NfL (en LCR i sang) s'associen directament amb l'asimetria cerebral. Treballs previs han demostrat l'associació dels nivells de NfL amb l'atròfia cerebral (37,38,49). No obstant això, no estudien l'associació entre els nivells de NfL i l'asimetria cerebral. A l'estudiar els dos clústers de MA, no vam ser capaços de trobar una explicació basada en el perfil clínic dels participants. Tanmateix, quan vam analitzar la seva associació amb biomarcadors de fluids vam observar que els nivells de GFAP en sang mostraven diferències significatives entre els dos clústers. Aquest biomarcador s'ha associat prèviament amb l'atròfia cerebral a causa de l'envelliment o la gravetat de la malaltia (145,146). Així i tot, la seva associació amb l'asimetria cerebral no s'ha estudiat abans. Així, en aquest cinquè estudi de la tesi vam demostrar que el CAI, definit en aquest mateix estudi, podria ser d'ajuda per a la diferenciació de les expressions clíniques o demències, per estudiar la progressió de la malaltia o per identificar subgrups.

Finalment, en el sisè treball d'aquesta tesi (147), vam aplicar mesures d'avaluació de RM en una variant concreta de DFT: la DFT genètica amb la mutació C9orf72. Les imatges de RM estructurals ens permeten mesurar la neurodegeneració del tronc encefàlic i estudiar si les anomalies que es mostren poden ser un biomarcador per la DFT amb la mutació C9orf72. Per això, primer vam utilitzar mètodes de volumetria semiautomatitzats per obtenir les diferents regions del tronc encefàlic i obtenir el volum de substància blanca. En aquest treball, vam trobar una proporció de la substància blanca del tronc encefàlic menor pels portadors simptomàtics en comparació als no portadors i les persones sanes, aquests resultats són per les tres subestructures del tronc encefàlic (mesencèfal, pons i bulb raquidi), així que ens podria suggerir una neurodegeneració generalitzada en aquesta zona del cervell. Tanmateix, no vam identificar diferències entre les persones sanes i els portadors asimptomàtics de la mutació C9orf72, així que la neurodegeneració en aquesta zona, probablement apareix prop de l'inici dels símptomes de la malaltia. Els pacients amb DFT amb símptomes de neurones motores van mostrar una atròfia significativament major, sobretot en el bulb raquidi, que els pacients que solament presenten DFT. Aquests resultats proposen que la relació de la substància blanca del tronc encefàlic, especialment al bulb raquidi, podria ser un biomarcador interessant per predir els símptomes de la neurona motora en els portadors de la mutació C9orf72. Aquesta troballa és especialment rellevant perquè la forma d'inici en els portadors de la mutació C9orf72 és altament impredictible i els pacients amb símptomes de neurones motores tenen un pitjor pronòstic general.

CAPÍTOL 6

Conclusions

1. L'aproximació freqüentista i la bayesiana tenen comportaments similars, per l'anàlisi de dades longitudinals de ressonància magnètica en la malaltia d'Alzheimer, si les dades són homogènies, sent la freqüentista computacionalment més simple. En canvi, quan les dades són heterogènies i amb un elevat nombre de visites que manquen, només l'aproximació bayesiana permet modelar les dades.
2. Els models predictius basats en gruix cortical permeten a) mesurar i relacionar la variabilitat del gruix cortical que presenten a escala individual els pacients amb malaltia d'Alzheimer i demència frontotemporal amb marcadors clínics i biomarcadors bioquímics, b)

predir la morfologia cerebral en l'evolució de pacients amb malaltia d'Alzheimer i demència frontotemporal.

3. La combinació d'algoritmes d'aprenentatge automàtic no supervisat i supervisat permet diferenciar la malaltia d'Alzheimer i la demència frontotemporal amb una única característica derivada de la ressonància magnètica estructural, alhora que proporciona mapes d'interpretabilitat. Les dades longitudinals permeten una millora en el diagnòstic diferencial entre malalties respecte als estudis transversals.
4. L'algoritme proposat amb dades de ressonància magnètica estructural i els nivells de líquid cefalorraquidi de cadenes lleugeres de neurofilaments (NfL) i proteïna 14-3-3 proporciona un diagnòstic vàlid per la malaltia d'Alzheimer i la demència frontotemporal. A més, permet obtenir probabilitats individuals per estudiar zones de baixa confiança del diagnòstic. La contribució diferencial de la ressonància magnètica estructural i els biomarcadors en líquid cefalorraquidi depèn dels grups clínics a classificar, així, pel diagnòstic diferencial entre malaltia d'Alzheimer i demència frontotemporal, la combinació de les dues modalitats és la que assoleix la millor classificació.
5. L'índex d'asimetria cerebral suggerit, estimat amb les mesures regionals del gruix cortical i derivat de l'entropia, és capaç de diferenciar la demència frontotemporal de la malaltia d'Alzheimer, i la variant semàntica respecte a la variant no fluent i la variant conductual.
6. Els portadors simptomàtics de la mutació C9orf72 presenten una pèrdua de substància blanca del tronc encefàlic que podria ser un marcador pel

seguiment de la malaltia, indicant els pacients amb afectació de neurones motores.

CAPÍTOL 7

Bibliografia

1. World Alzheimer Report. World Alzheimer Report 2015 - The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International; 2015.

2. Grossman H, Bergmann C, Parker S. Dementia: a brief review. Mt Sinai J Med. 2006 Nov;73(7):985–92.

3. Balasa M, Gelpi E, Antonell A, Rey MJ, Sánchez-Valle R, Molinuevo JL, et al. Clinical features and APOE genotype of pathologically proven early-onset Alzheimer disease. Neurology. 2011 May 17;76(20):1720–5.

4. Mendez MF. The accurate diagnosis of early-onset dementia. International Journal of Psychiatry in Medicine. 2006;36(4):401–12.

5. Arvanitakis Z, Shah RC, Bennett DA. Diagnosis and Management of Dementia: A Review. JAMA. 2019 Oct 22;322(16):1589–99.

6. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. Lancet Neurol. 2016 Apr;15(5):455–532.

7. Lobo A, Launer LJ, Fratiglioni L, Andersen K, Di Carlo A, Breteler MM, et al. Prevalence of dementia and major subtypes in Europe: A collaborative study

of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Neurology. 2000;54(11 Suppl 5):S4-9.

8. Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. The Lancet Neurology. 2013 Feb;12(2):207–16.

9. Mortamais M, Ash JA, Harrison J, Kaye J, Kramer J, Randolph C, et al. Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. Alzheimer's & Dementia. 2017 Apr 1;13(4):468–92.

10. Weintraub S, Wicklund AH, Salmon DP. The Neuropsychological Profile of Alzheimer Disease. Cold Spring Harb Perspect Med. 2012 Jan 4;2(4):a006171.

11. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011 May;7(3):280–92.

12. Alzheimer A. Über eine eigenartige Erkrankung der Hirnrinde. Allgemeine Zeitschrift fur Psychiatrie und Psychisch-gerichtliche Medizin. 1907;

13. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's and Dementia. 2018 Apr 1;14(4):535–62.

14. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984 Jul;34(7):939–44.

15. Consensus report of the Working Group on: "Molecular and Biochemical Markers of Alzheimer's Disease". The Ronald and Nancy Reagan Research Institute of the Alzheimer's Association and the National Institute on Aging Working Group. Neurobiol Aging. 1998;19(2):109–16.

16. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's and Dementia. 2011;7(3):270–9.

17. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's and Dementia. 2011;7(3):263–9.

18. Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 2010 Jan;9(1):119.
19. Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010 Feb;6(2):67–77.
20. Plant C, Teipel SJ, Oswald A, Böhm C, Meindl T, Mourao-Miranda J, et al. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. NeuroImage. 2010 Mar 1;50(1):162–74.
21. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007 Aug;6(8):734–46.
22. Pini L, Pievani M, Bocchetta M, Altomare D, Bosco P, Cavedo E, et al. Brain atrophy in Alzheimer's Disease and aging. Ageing Research Reviews. 2016 Sep 1;30:25–48.
23. Jack CR, Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology. 2000;55(4):484–9.
24. Scahill RI, Schott JM, Stevens JM, Rossor MN, Fox NC. Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. Proceedings of the National Academy of Sciences. 2002 Apr 2;99(7):4703–7.
25. Hampel H, O'Bryant SE, Molinuevo JL, Zetterberg H, Masters CL, Lista S, et al. Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. Nat Rev Neurol. 2018 Nov;14(11):639–52.
26. Sarto J, Ruiz-García R, Guillén N, Ramos-Campoy Ó, Falgàs N, Esteller D, et al. Diagnostic Performance and Clinical Applicability of Blood-Based Biomarkers in a Prospective Memory Clinic Cohort. Neurology. 2022 Nov 30;10.1212/WNL.0000000000201597.
27. Suárez-Calvet M, Karikari TK, Ashton NJ, Lantero Rodríguez J, Milà-Alomà M, Gispert JD, et al. Novel tau biomarkers phosphorylated at T181, T217 or T231 rise in the initial stages of the preclinical Alzheimer's continuum when only subtle changes in Aβ pathology are detected. EMBO Molecular Medicine. 2020 Dec 7;12(12):e12921.
28. Garre-Olmo J, Genís Batlle D, Del Mar Fernández M, Marquez Daniel F, De Eugenio Huélamo R, Casadevall T, et al. Incidence and subtypes of early-onset dementia in a geographically defined general population. Neurology. 2010 Oct 5;75(14):1249–55.
29. Hogan DB, Jetté N, Fiest KM, Roberts JI, Pearson D, Smith EE, et al. The Prevalence and Incidence of Frontotemporal Dementia: a Systematic

Review. Canadian Journal of Neurological Sciences. 2016 Apr;43(S1):S96–109.

30. Seelaar H, Rohrer JD, Pijnenburg YAL, Fox NC, Swieten JC van. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. Journal of Neurology, Neurosurgery & Psychiatry. 2011 May 1;82(5):476–86.

31. Onyike CU, Diehl-Schmid J. The Epidemiology of Frontotemporal Dementia. Int Rev Psychiatry. 2013 Apr;25(2):130–7.

32. Pick A. Über die Beziehunhen der Senilen Hirnatrophie zur Aphasie. Prager medizinische Wochenschrift. 1892;17:165–7.

33. Lashley T, Rohrer JD, Mead S, Revesz T. Review: an update on clinical, genetic and pathological aspects of frontotemporal lobar degenerations. Neuropathol Appl Neurobiol. 2015 Dec;41(7):858–81.

34. Rohrer JD, Guerreiro R, Vandrovcova J, Uphill J, Reiman D, Beck J, et al. The heritability and genetics of frontotemporal lobar degeneration. Neurology. 2009 Nov 3;73(18):1451–6.

35. Rohrer JD, Nicholas JM, Cash DM, van Swieten J, Dopper E, Jiskoot L, et al. Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the Genetic Frontotemporal dementia Initiative (GENFI) study: A cross-sectional analysis. The Lancet Neurology. 2015 Mar 1;14(3):253–62.

36. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011 Sep;134:2456–77.

37. Meeter LH, Dopper EG, Jiskoot LC, Sanchez-Valle R, Graff C, Benussi L, et al. Neurofilament light chain: a biomarker for genetic frontotemporal dementia. Annals of Clinical and Translational Neurology. 2016;3(8):623–36.

38. Rohrer JD, Woollacott IOC, Dick KM, Brotherhood E, Gordon E, Fellows A, et al. Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. Neurology. 2016 Sep 27;87(13):1329–36.

39. Scherling CS, Hall T, Berisha F, Klepac K, Karydas A, Coppola G, et al. Cerebrospinal fluid neurofilament concentration reflects disease severity in frontotemporal degeneration. Annals of Neurology. 2014;75(1):116–26.

40. Gordon E, Rohrer JD, Fox NC. Advances in neuroimaging in frontotemporal dementia. Journal of Neurochemistry. 2016;138(S1):193–210.

41. Rabinovici GD, Miller BL. Frontotemporal Lobar Degeneration. CNS Drugs. 2010 May 1;24(5):375–98.

42. Harris JM, Thompson JC, Gall C, Richardson AMT, Neary D, du Plessis D, et al. Do NIA-AA criteria distinguish Alzheimer's disease from frontotemporal dementia? Alzheimer's & Dementia. 2015 Feb 1;11(2):207–15.

43. Koedam ELGE, Lauffer V, van der Vlies AE, van der Flier WM, Scheltens P, Pijnenburg YAL. Early-Versus Late-Onset Alzheimer's Disease: More than Age Alone. Journal of Alzheimer's Disease. 2010 Jan 1;19(4):1401–8.

44. Mendez MF, Joshi A, Tassniyom K, Teng E, Shapira JS. Clinicopathologic differences among patients with behavioral variant frontotemporal dementia. Neurology. 2013 Feb 5;80(6):561–8.

45. Jack CR, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, et al. A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology. 2016 Aug 2;87(5):539–47.

46. Antonell A, Tort-Merino A, Ríos J, Balasa M, Borrego-Écija S, Auge JM, et al. Synaptic, axonal damage and inflammatory cerebrospinal fluid biomarkers in neurodegenerative dementias. Alzheimer's & Dementia [Internet]. 2019 Oct 24 [cited 2022 Dec 7]; Available from: https://www.sciencedirect.com/science/article/pii/S1552526019353695

47. Forgrave LM, Ma M, Best JR, DeMarco ML. The diagnostic performance of neurofilament light chain in CSF and blood for Alzheimer's disease, frontotemporal dementia, and amyotrophic lateral sclerosis: A systematic review and meta-analysis. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2019 Dec 1;11:730–43.

48. Gaiottino J, Norgren N, Dobson R, Topping J, Nissim A, Malaspina A, et al. Increased Neurofilament Light Chain Blood Levels in Neurodegenerative Neurological Diseases. PLOS ONE. 2013 de set;8(9):e75091.

49. Illán-Gala I, Lleo A, Karydas A, Staffaroni AM, Zetterberg H, Sivasankaran R, et al. Plasma Tau and Neurofilament Light in Frontotemporal Lobar Degeneration and Alzheimer Disease. Neurology. 2021 Feb 2;96(5):e671–83.

50. World Alzheimer Report. World Alzheimer Report 2021: Journey through the diagnosis of dementia. 2021 Sep 21 [cited 2023 Jul 15]; Available from: https://www.alzint.org/resource/world-alzheimer-report-2021/

51. Landqvist Waldö M, Frizell Santillo A, Passant U, Zetterberg H, Rosengren L, Nilsson C, et al. Cerebrospinal fluid neurofilament light chain protein levels in subtypes of frontotemporal dementia. BMC Neurology. 2013 May 29;13(1):54.

52. Skillbäck T, Farahmand B, Bartlett JW, Rosén C, Mattsson N, Nägga K, et al. CSF neurofilament light differs in neurodegenerative diseases and predicts severity and survival. Neurology. 2014 Nov 18;83(21):1945–53.

53. Canu E, Agosta F, Mandic-Stojmenovic G, Stojković T, Stefanova E, Inuggi A, et al. Multiparametric MRI to distinguish early onset Alzheimer's disease and behavioural variant of frontotemporal dementia. NeuroImage: Clinical. 2017;15:428–38.

54. Du AT, Schuff N, Kramer JH, Rosen HJ, Gorno-Tempini ML, Rankin K, et al. Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain. 2007 Apr 1;130(4):1159–66.

55. Falgàs N, Ruiz-Peris M, Pérez-Millan A, Sala-Llonch R, Antonell A, Balasa M, et al. Contribution of CSF biomarkers to early-onset Alzheimer's disease and frontotemporal dementia neuroimaging signatures. Human Brain Mapping. 2020;hbm.24925.

56. Cardinale F, Chinnici G, Bramerio M, Mai R, Sartori I, Cossu M, et al. Validation of FreeSurfer-estimated brain cortical thickness: comparison with histologic measurements. Neuroinformatics. 2014 Oct;12(4):535–42.

57. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006 Jul 1;31(3):968–80.

58. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron. 2002 Jan 31;33(3):341–55.

59. Falgàs N, Sánchez-Valle R, Bargalló N, Balasa M, Fernández-Villullas G, Bosch B, et al. Hippocampal atrophy has limited usefulness as a diagnostic biomarker on the early onset Alzheimer's disease patients: A comparison between visual and quantitative assessment. NeuroImage: Clinical. 2019 Jan 1;23:101927.

60. Bernal-Rusiel JL, Greve DN, Reuter M, Fischl B, Sabuncu MR. Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. NeuroImage. 2013;66:249–60.

61. Chagué P, Marro B, Fadili S, Houot M, Morin A, Samper-González J, et al. Radiological classification of dementia from anatomical MRI assisted by machine learning-derived maps. Journal of Neuroradiology. 2021 Nov 1;48(6):412–8.

62. Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: A review of machine learning applications. NeuroImage: Clinical. 2018 Jan 1;20:506–22.

63. Geisser S. Predictive Inference: An Introduction. Chapman and Hall/CRC; 1993.

64. Kuhn M, Johnson K. Applied Predictive Modeling [Internet]. New York, NY: Springer; 2013 [cited 2023 Jul 9]. Available from: http://link.springer.com/10.1007/978-1-4614-6849-3

65. Clifford JR, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, et al. Prediction of AD with MRI-Based Hippocampal Volume in Mild Cognitive Impairment. Neurology. 1999 Apr 22;52(7):1397–403.

66. Thompson WK, Hallmayer J, O'Hara R. Design considerations for characterizing psychiatric trajectories across the life span: Application to effects of APOE-ε4 on cerebral cortical thickness in Alzheimer's disease. American Journal of Psychiatry. 2011 Sep;168(9):894–903.

67. Cespedes MI, Fripp J, McGree JM, Drovandi CC, Mengersen K, Doecke JD. Comparisons of neurodegeneration over time between healthy ageing and Alzheimer's disease cohorts via Bayesian inference. BMJ Open. 2017 Feb 1;7(2).

68. Ziegler G, Penny WD, Ridgway GR, Ourselin S, Friston KJ. Estimating anatomical trajectories with Bayesian mixed-effects modeling. NeuroImage. 2015 Nov 1;121:51–68.

69. Lesaffre, E. Lawson AB. Bayesian biostatistics. Wiley; 2012.

70. Wilkinson M. Distinguishing between statistical significance and practical/clinical meaningfulness using statistical inference. Sports Medicine. 2014 Mar;44(3):295–301.

71. Hespanhol L, Vallio CS, Costa LM, Saragiotto BT. Understanding and interpreting confidence and credible intervals around effect estimates. Brazilian Journal of Physical Therapy. 2019 Jul 1;23(4):290–301.

72. Gurrin LC, Kurinczuk JJ, Burton PR. Bayesian statistics in medical research: an intuitive alternative to conventional data analysis. Journal of Evaluation in Clinical Practice. 2000 May 1;6(2):193–204.

73. Berry DA. Bayesian clinical trials. Nature Reviews Drug Discovery. 2006 Feb;5(1):27–36.

74. Poos JM, Grandpierre LDM, Ende EL van der, Panman JL, Papma JM, Seelaar H, et al. Longitudinal Brain Atrophy Rates in Presymptomatic Carriers of Genetic Frontotemporal Dementia. Neurology. 2022 Dec 13;99(24):e2661–71.

75. Coupé P, Manjón JV, Lanuza E, Catheline G. Lifespan Changes of the Human Brain In Alzheimer's Disease. Sci Rep. 2019 Mar 8;9(1):3998.

76. Pomponio R, Erus G, Habes M, Doshi J, Srinivasan D, Mamourian E, et al. Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. NeuroImage. 2020 Mar 1;208:116450.

77. Wang T, Qiu RG, Yu M. Predictive Modeling of the Progression of Alzheimer's Disease with Recurrent Neural Networks. Sci Rep. 2018 Jun 15;8(1):9161.

78. Marquand AF, Rezek I, Buitelaar J, Beckmann CF. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. Biological Psychiatry. 2016 Oct 1;80(7):552–61.

79. Wolfers T, Beckmann CF, Hoogman M, Buitelaar JK, Franke B, Marquand AF. Individual differences v. the average patient: mapping the heterogeneity in ADHD using normative models. Psychological Medicine. 2020 Jan;50(2):314–23.

80. Wolfers T, Doan NT, Kaufmann T, Alnæs D, Moberget T, Agartz I, et al. Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. JAMA Psychiatry. 2018 Nov 1;75(11):1146–55.

81. Verdi S, Marquand AF, Schott JM, Cole JH. Beyond the average patient: how neuroimaging models can address heterogeneity in dementia. Brain. 2021 Nov 29;144(10):2946–53.

82. Turing AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. Mind. 1950 Oct 1;LIX(236):433–60.

83. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development. 1959 Jul;3(3):210–29.

84. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning [Internet]. New York, NY: Springer; 2009 [cited 2023 Jul 16]. (Springer Series in Statistics). Available from: http://link.springer.com/10.1007/978-0-387-84858-7

85. Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science [Internet]. Cham: Springer International Publishing; 2019 [cited 2023 Jul 16]. Available from: http://link.springer.com/10.1007/978-3-319-99713-1

86. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. 2001 Mar 3;14.

87. Rao RB, Fung G, Rosales R. On the Dangers of Cross-Validation. An Experimental Evaluation. In: Proceedings of the 2008 SIAM International Conference on Data Mining [Internet]. Society for Industrial and Applied Mathematics; 2008 [cited 2023 Jul 16]. p. 588–96. Available from: https://epubs.siam.org/doi/10.1137/1.9781611972788.54

88. Alpaydin E. Introduction to Machine Learning, fourth edition. MIT Press; 2010. 709 p.

89. Bishop C. Pattern Recognition and Machine Learning [Internet]. Springer; 2006 [cited 2023 Aug 1]. Available from: https://link.springer.com/book/9780387310732

90. Myszczynska MA, Ojamies PN, Lacoste AMB, Neil D, Saffari A, Mead R, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat Rev Neurol. 2020 Aug;16(8):440–56.

91. Colliot O, editor. Machine Learning for Brain Disorders [Internet]. New York, NY: Springer US; 2023 [cited 2023 Jul 26]. (Neuromethods; vol. 197). Available from: https://link.springer.com/10.1007/978-1-0716-3195-9

92. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics [Internet]. 2014 [cited 2023 Mar 13];8. Available from: https://www.frontiersin.org/articles/10.3389/fninf.2014.00014

93. Frizzell TO, Glashutter M, Liu CC, Zeng A, Pan D, Hajra SG, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: A systematic review. Ageing Research Reviews. 2022 May 1;77:101614.

94. Bron EE, Smits M, Papma JM, Steketee RME, Meijboom R, de Groot M, et al. Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using structural and advanced MRI. European Radiology. 2017 Aug 1;27(8):3372–82.

95. Klöppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, et al. Automatic classification of MR scans in Alzheimer's disease. Brain. 2008 Mar;131(3):681–9.

96. Möller C, Pijnenburg YAL, Van Der Flier WM, Versteeg A, Tijms B, De Munck JC, et al. Alzheimer disease and behavioral variant frontotemporal dementia: Automatic classification based on cortical atrophy for single-subject diagnosis. Radiology. 2016 Jun 1;279(3):838–48.

97. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. Neurology. 2010;74(3):201–9.

98. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011 Mar;76:1006–14.

99. Landau S, Jagust W. Florbetapir processing methods. 2015;

100. Fischl B, Salat DH, Van Der Kouwe AJW, Makris N, Ségonne F, Quinn BT, et al. Sequence-independent segmentation of magnetic resonance images. In: NeuroImage. Academic Press; 2004. p. S69–84.

101. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of the United States of America. 2000 Sep 26;97(20):11050–5.

102. Seidman LJ, Faraone SV, Goldstein JM, Goodman JM, Kremen WS, Matsuda G, et al. Reduced subcortical brain volumes in nonpsychotic siblings of schizophrenic patients: A pilot magnetic resonance imaging study. American Journal of Medical Genetics - Neuropsychiatric Genetics. 1997;74(5):507–14.

103. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(Oct):2825–30.

104. Pérez-Millan A, Contador J, Tudela R, Niñerola-Baizán A, Setoain X, Lladó A, et al. Evaluating the performance of Bayesian and frequentist approaches for longitudinal modeling: application to Alzheimer's disease. Sci Rep. 2022 Aug 24;12(1):14448.

105. Besson FL, Joie RL, Doeuvre L, Gaubert M, Mézenge F, Egret S, et al. Cognitive and Brain Profiles Associated with Current Neuroimaging Biomarkers of Preclinical Alzheimer's Disease. J Neurosci. 2015 Jul 22;35(29):10402–11.

106. Márquez F, Yassa MA. Neuroimaging Biomarkers for Alzheimer's Disease. Mol Neurodegeneration. 2019 Jun 7;14(1):21.

107. Sarazin M, Chauviré V, Gerardin E, Colliot O, Kinkingnéhun S, de Souza LC, et al. The Amnestic Syndrome of Hippocampal type in Alzheimer's Disease: An MRI Study. Journal of Alzheimer's Disease. 2010 Jan 1;22(1):285–94.

108. Aksman LM, Scelsi MA, Marquand AF, Alexander DC, Ourselin S, Altmann A. Modeling Longitudinal Imaging Biomarkers with Parametric Bayesian Multi-Task Learning. Human Brain Mapping. 2019;40(13):3982–4000.

109. Temp AGM, Lutz MW, Trepel D, Tang Y, Wagenmakers E, Khachaturian AS, et al. How Bayesian statistics may help answer some of the controversial questions in clinical research on Alzheimer's disease. Alzheimer's & Dementia. 2021 Jun 11;17(6):917–9.

110. Filippone M, Marquand AF, Blain CRV, Williams SCR, Mourão-Miranda J, Girolami M. PROBABILISTIC PREDICTION OF NEUROLOGICAL DISORDERS WITH A STATISTICAL ASSESSMENT OF NEUROIMAGING DATA MODALITIES. Ann Appl Stat. 2012 Dec 27;6(4):1883–905.

111. Lorenzi M, Filippone M, Frisoni GB, Alexander DC, Ourselin S. Probabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in Alzheimer's disease. NeuroImage. 2019 Apr 15;190:56–68.

112. Young J, Modat M, Cardoso MJ, Mendelson A, Cash D, Ourselin S. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in

patients with mild cognitive impairment. NeuroImage: Clinical. 2013 Jan 1;2:735–45.

113. Blanc F, Colloby SJ, Philippi N, Pétigny X de, Jung B, Demuynck C, et al. Cortical Thickness in Dementia with Lewy Bodies and Alzheimer's Disease: A Comparison of Prodromal and Dementia Stages. PLOS ONE. 2015 Jun 10;10(6):e0127396.

114. Contador J, Pérez-Millan A, Guillen N, Tort-Merino A, Balasa M, Falgàs N, et al. Baseline MRI atrophy predicts 2-year cognitive outcomes in early-onset Alzheimer's disease. Journal of Neurology [Internet]. 2021; Available from: https://doi.org/10.1007/s00415-021-10851-9

115. Julkunen V, Niskanen E, Koikkalainen J, Herukka SK, Pihlajamäki M, Hallikainen M, et al. Differences in Cortical Thickness in Healthy Controls, Subjects with Mild Cognitive Impairment, and Alzheimer's Disease Patients: A Longitudinal Study. Journal of Alzheimer's Disease. 2010 Jan 1;21(4):1141–51.

116. Eratne D, Keem M, Lewis C, Kang M, Walterfang M, Farrand S, et al. Cerebrospinal fluid neurofilament light chain differentiates behavioural variant frontotemporal dementia progressors from non-progressors. J Neurol Sci. 2022 Nov 15;442:120439.

117. Pérez-Millan A, Contador J, Juncà-Parella J, Bosch B, Borrell L, Tort-Merino A, et al. Classifying Alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. Human Brain Mapping [Internet]. 2023 [cited 2023 Jan 24]; Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.26205

118. Bisenius S, Mueller K, Diehl-Schmid J, Fassbender K, Grimmer T, Jessen F, et al. Predicting primary progressive aphasias with support vector machine approaches in structural MRI data. NeuroImage: Clinical. 2017 Jan 1;14:334–43.

119. Bron EE, Klein S, Papma JM, Jiskoot LC, Venkatraghavan V, Linders J, et al. Cross-cohort generalizability of deep and conventional machine learning for MRI-based diagnosis and prediction of Alzheimer's disease. NeuroImage: Clinical. 2021 Jan 1;31:102712.

120. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert MO, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage. 2011 May 15;56(2):766–81.

121. Magnin B, Mesrob L, Kinkingnéhun S, Pélégrini-Issac M, Colliot O, Sarazin M, et al. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology. 2009 Feb 1;51(2):73–83.

122. Abdi H, Williams LJ, Valentin D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. Wiley Interdisciplinary Reviews: Computational Statistics. 2013 Mar 1;5(2):149–79.

123. Gavidia-Bovadilla G, Kanaan-Izquierdo S, Mataroa-Serrat M, Perera-Lluna A. Early prediction of Alzheimer's disease using null longitudinal model-based classifiers. PLoS ONE [Internet]. 2017 Jan 1 [cited 2022 Apr 24];12(1). Available from: /pmc/articles/PMC5207395/

124. Guo M, Li Y, Zheng W, Huang K, Zhou L, Hu X, et al. A novel conversion prediction method of MCI to AD based on longitudinal dynamic morphological features using ADNI structural MRIs. Journal of Neurology. 2020 Oct 1;267(10):2983–97.

125. Zhang J, Liu M, An L, Gao Y, Shen D. Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images. IEEE Journal of Biomedical and Health Informatics. 2017 Nov 1;21(6):1607–16.

126. Rabinovici GD, Seeley WW, Kim EJ, Gorno-Tempini ML, Rascovsky K, Pagliaro TA, et al. Distinct MRI atrophy patterns in autopsy-proven Alzheimer's disease and frontotemporal lobar degeneration. American Journal of Alzheimer's Disease and other Dementias. 2008;22(6):474–88.

127. Davatzikos C, Resnick SM, Wu X, Parmpi P, Clark CM. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. NeuroImage. 2008 Jul 15;41(4):1220–7.

128. Laakso MP, Frisoni GB, Könönen M, Mikkonen M, Beltramello A, Geroldi C, et al. Hippocampus and entorhinal cortex in frontotemporal dementia and Alzheimer's disease: a morphometric MRI study. Biological Psychiatry. 2000 Jun 1;47(12):1056–63.

129. Dashtipour K, Taylor W, Ansari S, Zahid A, Gogate M, Ahmad J, et al. Detecting Alzheimer's Disease Using Machine Learning Methods. In: Ur Rehman M, Zoha A, editors. Body Area Networks Smart IoT and Big Data for Intelligent Health Management. Cham: Springer International Publishing; 2022. p. 89–100. (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering).

130. Dukart J, Mueller K, Horstmann A, Barthel H, Möller HE, Villringer A, et al. Combined Evaluation of FDG-PET and MRI Improves Detection and Differentiation of Dementia. PLOS ONE. 2011 Mar 23;6(3):e18111.

131. McCarthy J, Collins DL, Ducharme S. Morphometric MRI as a diagnostic biomarker of frontotemporal dementia: A systematic review to determine clinical applicability. NeuroImage: Clinical. 2018 Jan 1;20:685–96.

132. Salvatore C, Cerasa A, Battista P, Gilardi MC, Quattrone A, Castiglioni I. Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. Front Neurosci. 2015 Sep 1;9:307.

133. Wang J, Redmond SJ, Bertoux M, Hodges JR, Hornberger M. A Comparison of Magnetic Resonance Imaging and Neuropsychological Examination in the Diagnostic Distinction of Alzheimer's Disease and Behavioral Variant Frontotemporal Dementia. Frontiers in Aging Neuroscience [Internet]. 2016 [cited 2023 Apr 9];8. Available from: https://www.frontiersin.org/articles/10.3389/fnagi.2016.00119

134. Avants BB, Cook PA, Ungar L, Gee JC, Grossman M. Dementia induces correlated reductions in white matter integrity and cortical thickness: A multivariate neuroimaging study with sparse canonical correlation analysis. NeuroImage. 2010 Apr 15;50(3):1004–16.

135. Dickerson BC, Goncharova I, Sullivan MP, Forchetti C, Wilson RS, Bennett DA, et al. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiology of Aging. 2001;22(5):747–54.

136. Gil-Navarro S, Lladó A, Rami L, Castellví M, Bosch B, Bargalló N, et al. Neuroimaging and Biochemical Markers in the Three Variants of Primary Progressive Aphasia. DEM. 2013;35(1–2):106–17.

137. Hodges JR, Patterson K. Semantic dementia: a unique clinicopathological syndrome. The Lancet Neurology. 2007 Nov 1;6(11):1004–14.

138. Bruun M, Koikkalainen J, Rhodius-Meester HFM, Baroni M, Gjerum L, van Gils M, et al. Detecting frontotemporal dementia syndromes using MRI biomarkers. NeuroImage: Clinical. 2019 Jan 1;22:101711.

139. Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, et al. Cognition and anatomy in three variants of primary progressive aphasia. Annals of Neurology. 2004;55(3):335–46.

140. Rohrer JD, Rosen HJ. Neuroimaging in frontotemporal dementia. International Review of Psychiatry. 2013 Apr 1;25(2):221–9.

141. Whitwell JL, Boeve BF, Weigand SD, Senjem ML, Gunter JL, Baker MC, et al. Brain atrophy over time in genetic and sporadic frontotemporal dementia: a study of 198 serial magnetic resonance images. European Journal of Neurology. 2015;22(5):745–52.

142. Bocchetta M, Iglesias JE, Scelsi MA, Cash DM, Cardoso MJ, Modat M, et al. Hippocampal Subfield Volumetry: Differential Pattern of Atrophy in Different Forms of Genetic Frontotemporal Dementia. Journal of Alzheimer's Disease. 2018 Jan 1;64(2):497–504.

143. Low A, Mak E, Malpetti M, Chouliaras L, Nicastro N, Su L, et al. Asymmetrical atrophy of thalamic subnuclei in Alzheimer's disease and amyloid-positive mild cognitive impairment is associated with key clinical features. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2019;11(1):690–9.

144. Schroeter ML, Raczka K, Neumann J, Yves von Cramon D. Towards a nosology for frontotemporal lobar degenerations—A meta-analysis involving 267 subjects. NeuroImage. 2007 Jul 1;36(3):497–510.

145. Ebenau JL, Pelkmans W, Verberk IMW, Verfaillie SCJ, Bosch KA van den, Leeuwenstijn M van, et al. Association of CSF, Plasma, and Imaging Markers of Neurodegeneration With Clinical Progression in People With Subjective Cognitive Decline. Neurology. 2022 Mar 29;98(13):e1315–26.

146. Traub J, Otto M, Sell R, Homola GA, Steinacker P, Oeckl P, et al. Serum glial fibrillary acidic protein indicates memory impairment in patients with chronic heart failure. ESC Heart Failure. 2022;9(4):2626–34.

147. Pérez-Millan A, Borrego-Écija S, van Swieten JC, Jiskoot L, Moreno F, Laforce R, et al. Loss of brainstem white matter predicts onset and motor neuron symptoms in C9orf72 expansion carriers: a GENFI study. J Neurol. 2023 Mar 1;270(3):1573–86.