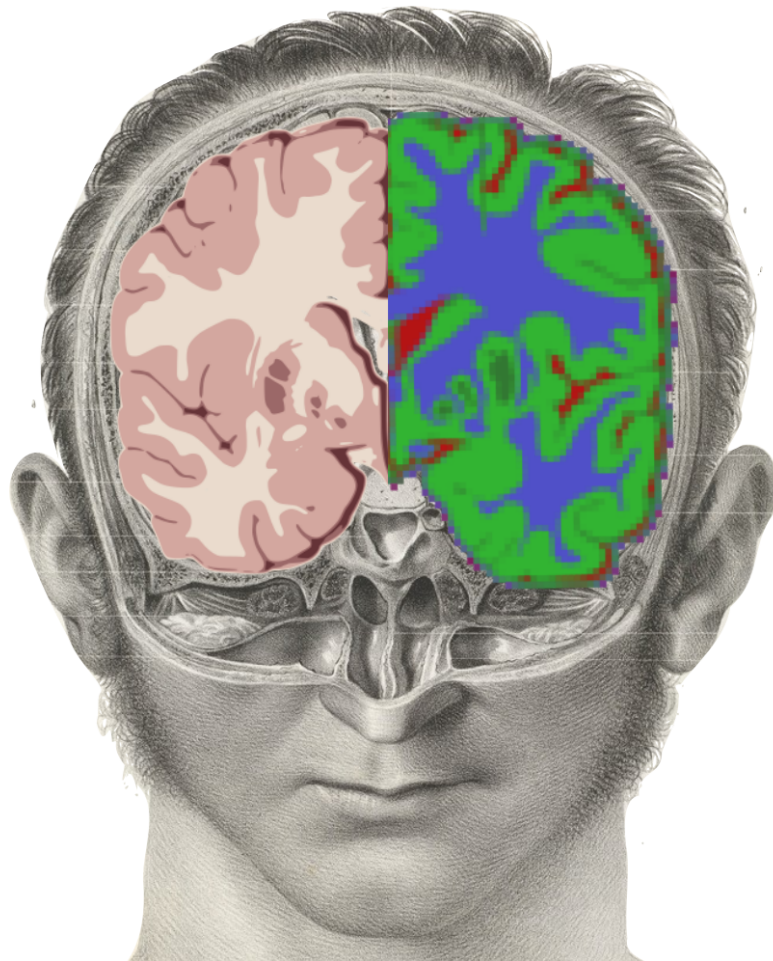# Computational anatomy strategies for characterization of brain patterns associated with Alzheimer's disease

Diana Lorena Giraldo Franco

Supervisors **prof. dr. E. Romero | prof. dr. J. Sijbers | dr. B. Jeurissen**

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Science
Faculty of Science | Antwerpen, 2022

University of Antwerp

University
of Antwerp

UNIVERSIDAD
NACIONAL
DE COLOMBIA

Faculty of Science
Doctor of science

Faculty of Engineering
PhD in engineering

Computational anatomy strategies for characterization of brain patterns associated with Alzheimer's disease

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of science
at University of Antwerp

PhD in Engineering
at Universidad Nacional de Colombia

Diana Lorena Giraldo Franco

Supervisors
prof. dr. E. Romero
prof. dr. J. Sijbers
dr. B. Jeurissen

Antwerp, 2022
Bogotá, 2022

Jury

Chairman
prof. dr. M. Verhoye, University of Antwerp, Belgium

Supervisors
prof. dr. E. Romero, Universidad Nacional de Colombia, Colombia
prof. dr. J. Sijbers, University of Antwerp, Belgium
dr. B. Jeurissen, University of Antwerp, Belgium

Members
prof. dr. I. Jelescu, Université de Lausanne, Switzerland
prof. dr. N. Malpica, Universidad Rey Juan Carlos, Spain
prof. dr. S. Van Doorslaer, University of Antwerp, Belgium
prof. dr. F. A. González Osorio, Universidad Nacional de Colombia, Colombia
prof. dr. D. A. Garzón Alvarado, Universidad Nacional de Colombia, Colombia

Contact
Diana Lorena Giraldo Franco

🏛 Universidad Nacional de Colombia, Computer Imaging and Medical Applications
Laboratory - CIM@LAB
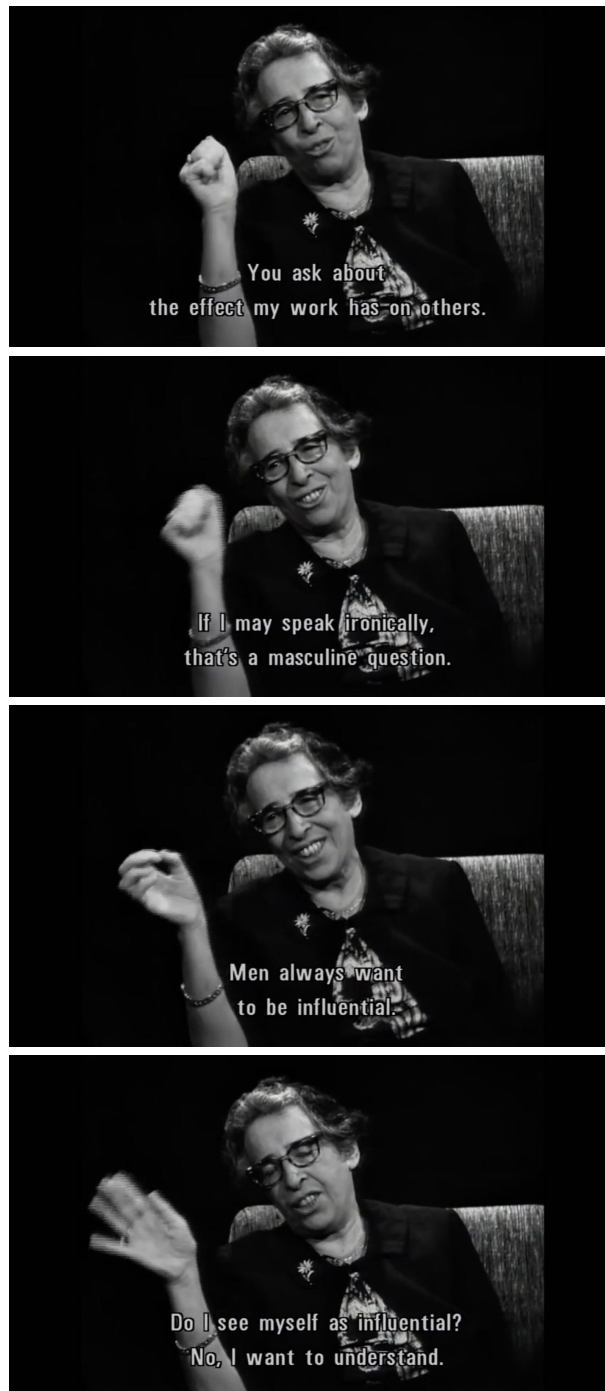🏛 Universiteit Antwerpen, imec-VisionLab
✉ dlgiraldof@unal.edu.co
✉ diana.giraldofranco@uantwerpen.be
✉ dia.giraldo@gmail.com

Hannah Arendt in "Zur Person" (1964)

Acknowledgements

This is the space where I can recognize all the people who made this work possible, the ones who gave me incredible opportunities and support, the ones who gave me the conditions to take those opportunities, and the ones who were around during these years under quite unexpected circumstances.

I want to express my sincere gratitude to Eduardo Romero, who was been a great teacher during the last 10 years. His guidance and advice have played a fundamental role in who I am as a researcher and as a person. I deeply appreciate his constant support, even in the most difficult and uncertain times, and the fact that he always believed in me way more than I did in myself. I need to highlight his patience insistence on teaching me an important lesson for life, that sometimes the best you can do is to shut up and listen.

To Jan Sijbers for giving me the incredible opportunity to work on the Visionlab, opening my eyes to a whole new environment and culture. To Ben Jeurissen for his dedication, time, help, trust, and fruitful conversations that shaped a large part of this thesis. Their kindness and willingness to support me were (and still are) remarkable. I will always appreciate the warmth, joy and consideration with which Jan and Ben received me in Antwerp for the first time during the autumn of 2015, when I was still in a state of confusion about language, weather, food, and all the things related to being in a different continent for the first time in my life.

I was given great opportunities during my career, opportunities that I could take thanks to the hard work of my parents. They encouraged me and gave me the conditions to be admitted to a public university to study Mathematics. Their effort and support on many of the decisions I make have made this road possible.

Es más que necesario escribir estos agradecimientos también en español. Esta tesis representa un punto emblemático de un camino sin final, y este camino no habría sido posible sin el esfuerzo y apoyo de mi papás ♡. A ellos, quiénes trabajaron gran parte de sus vidas para asegurarse de que tuviéramos las condiciones y oportunidades que para ellos fueron escasas, les debo poder estar escribiendo estos agradecimientos.

To the Cim@lab research group for being a great environment to exchange ideas and opinions. Special thanks Germán, Chars, and Josué, people who I can consider my friends. Thanks for your help, support, conversations and for all the unforgettable moments we have shared. The most special thanks go to Josué, my partner, for having my back for the last couple of years. I doubt I would have finished this thesis under these unexpected circumstances without his care and love ❤.

I also want to thank the Visionlab people who welcomed me with kindness, especially to Gabriel and Piet with whom I got the luck to share an office and quickly became my

vi

friends. We had a lot of fun together, and I will never forget Pearl Harbour thanks to the two of you... unless something strange happens with my memory (pun intended).

To all my friends who gave me their unconditional support and were by my side during these years: Mabe, Peri, Laura A., Laura E., Emmis, Rodrigo, Daniel, Beatriz, Cata y Federico. I consider myself a lucky person because there are many people that I can call my friends. I want to extend my gratitude to them, although some of these people still think this thesis is about robotic arms for babies with cancer.

<div align="right">

Muchas gracias
Thanks a lot
Dank je wel

</div>

Abstract

Computational anatomy strategies for characterization of brain patterns associated with Alzheimer's disease

Alzheimer's disease (AD) is one of the most complex systematic malfunctions of the nervous system that are known. The clinical symptoms of this neurodegenerative disease are alterations in cognition and behaviour that can lead to the onset of a dementia syndrome. Disease mechanisms that lead to neurodegeneration and cognitive impairment in sporadic AD are not well understood yet, making it difficult to predict the clinical progression of patients at the early stages of the AD continuum. Currently, no single biomarker or exam is sufficient to diagnose AD and existing standard instruments are not sensitive enough to detect subtle changes, predict the clinical course, and recognize heterogeneous forms of AD. This thesis presents two computational anatomy strategies aiming to identify and quantify neurodegeneration patterns associated with different clinical stages along the AD continuum using two different modalities of magnetic resonance imaging. A third contribution consists of a data-driven strategy to develop a set of domain-specific scores that result useful to estimate the risk of and predict the progression from mild cognitive impairment to dementia. Evaluation of these strategies with machine-learning and statistical inference methods demonstrate the potential of the proposed quantitative tools to help patients' clinical management and monitoring and could be used to improve the evaluation of potential disease-modifying interventions.

Keywords: Alzheimer's Disease, Neuroimaging, Medical Image Processing, Magnetic Resonance Imaging, Cognitive Impairment.

viii

Resumen

Estrategias de anatomía computacional para la caracterización de patrones cerebrales asociados a la enfermedad de Alzheimer

La enfermedad de Alzheimer (EA) es una de las fallas sistemáticas del sistema nervioso más complejas que se conocen. Los síntomas clínicos de esta enfermedad neurodegenerativa son alteraciones de la cognición y el comportamiento que pueden conducir a la aparición de un síndrome de demencia. Los mecanismos de la enfermedad que conducen a la neurodegeneración y al deterioro cognitivo en la EA aún no se conocen bien, lo que dificulta la predicción de la evolución clínica de los pacientes en las primeras fases de la EA. Actualmente, ningún biomarcador o examen es suficiente para diagnosticar la EA y los instrumentos estándar existentes no son lo suficientemente sensibles para detectar cambios sutiles, predecir el curso clínico o reconocer presentaciones atípicas de EA. Esta tesis presenta dos estrategias de anatomía computacional destinadas a identificar y cuantificar los patrones de neurodegeneración asociados a diferentes etapas clínicas a lo largo del continuo de la EA utilizando dos modalidades diferentes de imágenes de resonancia magnética. Una tercera contribución consiste en una estrategia guiada por datos para desarrollar un conjunto de puntajes específicas por dominio que resultan útiles para estimar el riesgo y predecir la progresión del deterioro cognitivo leve a la demencia. La evaluación de estas estrategias con métodos de aprendizaje automático y de inferencia estadística demuestra el potencial de las herramientas cuantitativas propuestas para ayudar al manejo y el seguimiento clínico de los pacientes y podría utilizarse para mejorar la evaluación de posibles intervenciones que puedan modificar el curso de la enfermedad.

Palabras clave: Enfermedad de Alzheimer, Neuroimágenes, Procesamiento de imágenes médicas, Imágenes de Resonancia Magnética, Deterioro Cognitivo.

Samenvatting

Computationele anatomie strategieën voor karakterisering van hersenpatronen geassocieerd met de ziekte van Alzheimer

De ziekte van Alzheimer (AD) is een van de meest complexe systemische storingen van het zenuwstelsel die bekend zijn. De klinische symptomen van deze neurodegeneratieve ziekte zijn veranderingen in cognitie en gedrag die kunnen leiden tot het ontstaan van een dementiesyndroom. De ziektemechanismen die leiden tot neurodegeneratie en cognitieve stoornissen bij sporadische AD zijn nog niet goed begrepen, waardoor het moeilijk is om de klinische progressie van patiënten in de vroege stadia van het AD continuüm te voorspellen. Momenteel is geen enkele biomarker of onderzoek voldoende om de diagnose AD te stellen en de bestaande standaardinstrumenten zijn niet gevoelig genoeg om subtiele veranderingen te detecteren, het klinische verloop te voorspellen en heterogene vormen van AD te herkennen. Dit proefschrift presenteert twee computationele anatomiestrategieën die gericht zijn op het identificeren en kwantificeren van neurodegeneratiepatronen geassocieerd met verschillende klinische stadia in het AD continuüm, gebruikmakend van twee verschillende modaliteiten van magnetische resonantie beeldvorming. Een derde bijdrage bestaat uit een data-gestuurde strategie om een reeks van domeinspecifieke scores te ontwikkelen die bruikbaar zijn om het risico in te schatten op en de progressie te voorspellen van milde cognitieve stoornissen naar dementie. Evaluatie van deze strategieën met machine-learning en statistische inferentie methoden tonen het potentieel aan van de voorgestelde kwantitatieve instrumenten om het klinisch management en de monitoring van patiënten te helpen en zouden gebruikt kunnen worden om de evaluatie van potentiële ziekte-modificerende interventies te verbeteren.

Sleutelwoorden: Ziekte van Alzheimer, Neurobeeldvorming, Medische Beeldverwerking, Magnetische Resonantie Beeldvorming, Cognitieve Stoornis.

x

Contents

List of acronyms and abbreviations

AD	Alzheimer's Disease
ADAS-Cog	Alzheimer's Disease Assessment Scale - Cognition
ADC	Apparent diffusion coefficient
ADD	Alzheimer's Disease Dementia
ADNI	Alzheimer's Disease Neuroimaging Initiative
AFD	Apparent fibre density
AUC	Area under the curve
$A\beta_{42}$	Amyloid-β peptide 42
CDR	Clinical Dementia Rating
CFA	Confirmatory factor analysis
CI	Cognitive impairment
CN	Cognitively normal
CSD	Constrained spherical deconvolution
CSF	Cerebrospinal fluid
CU	Cognitively unimpaired
DT	Diffusion Tensor
DTI	Diffusion Tensor Imaging
DW-MRI	Diffusion weighted magnetic resonance imaging
EER	Equal error rate
EMD	Earth-mover's distance
FA	Fractional anisotropy
FBA	Fixel-based analysis
FC	Fibre cross-section
fODF	Fibre orientation distribution function
FWE	Family-wise error
GLM	General linear model
GM	Grey matter
ICV	Intracranial volume
ILF	Inferior longitudinal fasciculus
ilr	Isometric log-ratios
MCI	Mild cognitive impairment
MD	Mean diffusivity

MMSE	Mini–Mental State Examination
MoCA	Montreal Cognitive Assessment
MRI	Magnetic resonance imaging
MSMT-CSD	Multi-shell multi-tissue constrained spherical deconvolution
MTL	Medial temporal lobe
P-tau	Phosphorylated tau
PET	Positron emission tomography
PVE	Partial volume effects
RF	Random forests
ROC curve	Receiver operating characteristic curve
SH	Spherical harmonics
SRB	Standardized regression based
VBA	Voxel-based analysis
VBM	Voxel-based morphometry
WM	White matter

Introduction

1.1 Alzheimer's disease dementia

Dementia is a syndrome characterized by the progressive deterioration of cognitive function as a result of a brain disease. This syndrome can affect memory, thinking, judgment and behavior up to the point people are unable to perform daily life activities and require constant assistance for the rest of their life. The biggest known risk factor for a person to develop dementia is ageing. A report combining multiple studies estimated the incidence of dementia doubles with every 6.3 year increase in age [148]. The increased life expectancy and aging of the general population have made of dementia a global public health concern. According to global estimates, in 2016 there were 43.8 million individuals living with dementia worldwide [135] and it is expected that by the year 2030 this number will reach 75 million [148]. In addition to the associated mortality, most of the social and monetary impacts of dementia stem from disability, posing an increasing burden on caregivers and healthcare systems.

The most common cause of dementia is Alzheimer's disease (AD), a neurodegenerative disorder with no effective disease-modifying treatment currently available [67, 125, 166]. This disease is pathologically defined by the presence of Amyloid-β plaques and neurofibrillary tau deposits [171, 92] associated with neuronal and synaptic loss (Figure 1.1). Although these processes might lead to cognitive impairment and dementia, Alzheimer's pathology can be present in people who did not show symptoms during their lifetime [50].

Excluding the genetic mutations that cause the early-onset hereditary AD and account for less than 5% of AD cases, the etiology of late-onset AD is complex and poorly understood [67]. Experts believe that Alzheimer's develops as a result of multiple factors such as genetic, lifestyle and environment. Besides ageing, other risk factors have been identified, including vascular diseases (*e.g.* hypertension, obesity), genetic susceptibility, and life-style factors such as diet, physical and mental activity, alcohol consumption, and education level [171].

Traditionally, AD has been recognized in terms of its typical clinical manifestation, that is the multi-domain amnestic dementia. This typical expression of AD is characterized by the progressive deterioration of episodic memory and other cognitive domains such as language, executive function, attention, and visuospatial abilities [210]. Examination
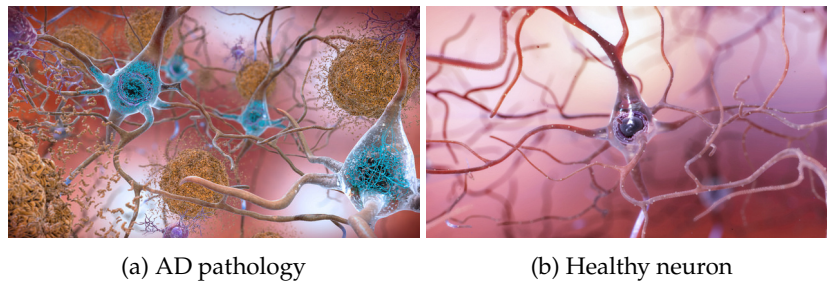
(a) AD pathology   (b) Healthy neuron

Figure 1.1: Illustration of AD pathology compared with healthy neural tissue. AD is defined by the abnormal accumulation of two proteins that form extracellular Amyloid-β plaques (in dark yellow), and intracellular tangles of tau (in blue). Image source: National Institute on Aging, National Institutes of Health (NIA-NIH).

of the brain in autopsy-confirmed cases of AD has shown a characteristic pattern for the location of AD degeneration: initially, it appears in the entorhinal cortex progressing through the hippocampus and medial temporal structures (Shown in Figure 1.2), to eventually affect association cortices [19, 132]. This neuropathological pathway correlates with the clinical picture of typical AD which starts as an amnestic syndrome of the hippocampal type accompanied by some impairment in executive functions or naming abilities [48].

Although the typical AD is the most frequent, some atypical presentations of AD have been recognized. These atypical forms vary from the amnestic syndrome presenting a predominant executive function impairment [203], aphasia and visuospatial dysfunctions [50]. The existence of such atypical forms of AD has been confirmed with neuropathological examinations of brain tissue [132] and analyses of cortical atrophy patterns [46] finding cases with spared hippocampal atrophy but posterior cortical atrophy.

All the known clinical signs and symptoms of AD are related to disturbances in cognition and behavior. Depending on the presentation and stage of the disease, one or multiple cognitive domains can be affected. Such symptoms are evaluated with the neuropsychological examination consisting of tests to assess the overall level of Cognitive Impairment (CI), specific tests to detect alterations in particular domains like memory, and interviews with the patient and relatives to grade the severity of dementia-like symptoms. The clinical onset of AD is determined by the clinical diagnosis of dementia (or major neurocognitive disorders), requiring the cognitive deficits to interfere with the ability to perform everyday activities [8]. However, cognitive impairment can be detected before it compromises the daily functioning, *i.e.* in a pre-dementia stage of the disease.

Mild cognitive impairment

When individuals show some cognitive decline in one or more cognitive domains, but this decline does not interfere with their activities or behavior, they are diagnosed with mild cognitive impairment (MCI). This broad category includes patients with similar clinical features but a variety of different causes including neurodegenerative disease such as AD, psychiatric conditions like depression, or even side effects of certain medications

(a) Hippocampus

(b) Medial Temporal Lobe (MTL)
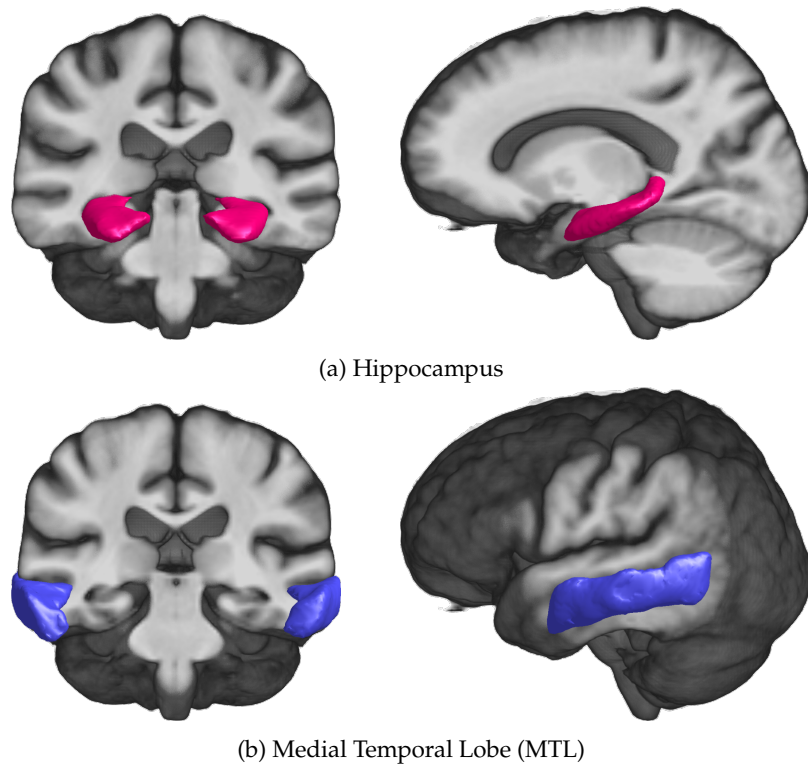
Figure 1.2: Characteristic AD lesions have been found in the hippocampus and medial temporal lobe. Atrophy of these brain areas has been recognized as a *"topographical"* biomarker of AD.

can be responsible for the perceived cognitive impairment.

Traditionally, MCI has been classified into two sub-types: amnestic and non-amnestic [143], depending on whether there is memory impairment or not. In non-amnestic MCI, cognitive domains like language, visuospatial skills, or executive function show some impairment that drives the distinction from normal ageing. Initially, amnestic MCI was thought to represent a prodromal form of AD [143] but nowadays it is known that multiple pathologies can cause amnestic MCI and not all AD cases show memory dysfunction at pre-clinical stages.

Although the group of subjects with amnestic MCI has a homogeneous clinical pheno-type, it is highly heterogeneous in terms of the underlying biology, compromise of other cognitive domains, and specific domain decline trajectories [134, 54, 27]. This hetero-geneity makes it difficult to predict clinical progression for MCI patients because they can progress to dementia, remain stable or even revert to cognitively normal. That is also the reason why identifying which individuals with MCI are more likely to develop AD dementia is a research priority [177] and an active field of study [1].

Clinical diagnosis of AD

Clinical diagnosis of Alzheimer has been guided by the NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association) criteria published in 1984 [124], it stratifies the confidence of the diagnosis in probable and definite depending on the level of certainty that dementia syndrome is caused by Alzheimer's pathology:

- **Probable AD:** based on the diagnosis of dementia with progressive impairment of memory and other cognitive functions with no presence of other diseases that can cause cognitive deficits.

- **Definite AD:** based on a diagnosis of probable AD while the patient was alive and evidence of AD pathology from tissue examination *post-mortem* [89].

According to these criteria, the amnestic syndrome was a core feature for the clinical diagnosis of probable AD, recognizing only the typical manifestation of AD. Update proposals for diagnostic criteria acknowledge the atypical forms of AD by including other clinical phenotypes different from memory impairment [50].

Diagnosis criteria give the general guidelines to diagnose AD in clinical settings but the way the guidelines are implemented is highly variable across medical facilities. The diagnosis of dementia mainly relies on clinical examination and neuropsychological testing. However, the diagnostic guidelines do not specify which tests or population standards should be applied, therefore each medical center chooses the tests and cut-offs used to ascertain whether the patient's results are normal or abnormal. For example, a systematic analysis of the global burden of dementia found 230 different diagnostic procedures across 237 studies in the dementia literature [135] and the main source of this heterogeneity is the use of different tests and cut-off scores during cognitive screening.

Given the research advances to prove Alzheimer's pathology with *in vivo* biomarkers, there have been various proposals to include the use of biomarkers for AD diagnosis in dementia [49, 50, 92] and pre-dementia stages of the disease [5, 51]. These recommendations are specially useful in research settings where early detection of Alzheimer's pathology can be part of the inclusion criteria for clinical trials.

The use of biomarkers for the "early" detection of AD pathology without symptoms of dementia in clinical practice is a matter of current debate [172, 76, 90]. Considering that there are not disease-modifying treatments available and it is not certain if the pathological signs do inevitably lead to dementia syndrome, an early diagnosis of Alzheimer's could create a psychological burden in patients that may never develop dementia [172].

Biomarkers for AD

The definite diagnosis of AD is done by the assessment of characteristic structural lesions through the pathological examination of brain tissue. These characteristic lesions are formed by abnormal accumulation of proteins, specifically: extracellular deposit of amyloid-β and neurofibrillary tangles of the protein tau. The two proteins define two groups of *in vivo* pathophysiological markers of Alzheimer's pathology [48, 92]:

- Markers of amyloidosis: low levels of amyloid-β peptide 42 ($A\beta_{42}$) or $A\beta_{42}/A\beta_{02}$ ratio in cerebrospinal fluid (CSF), and high cortical binding values for positron emission tomography (PET) with Pittsburgh compound B (PiB).

- Markers of tauopathy: high CSF total tau (t-tau) and phosphorylated tau (P-tau), and PET with tau ligands.

These CSF and PET biomarkers have demonstrated to be correlated with the pathological marks of AD: amyloid plaques and neurofibrillary tau deposits, however, they are largely static and give little information about disease stage or progression [48].

Other biomarkers assess subsequent pathological brain changes related with AD progression such as synaptic and neuronal loss. These include Fluorodeoxyglucose (FDG) - PET, which measures glucose uptake and it is sensitive to neuronal dysfunction, and structural Magnetic Resonance Imaging (MRI) to detect atrophy in certain areas such as the medial temporal lobe (MTL) and hippocampus. Although these *"topographical"* biomarkers are less specific for AD, they do correlate with disease severity [48], and can improve disease characterization [50] and prediction of cognitive decline in MCI patients [92].

Biomarkers give evidence of Alzheimer's pathology and AD-related pathological changes in any stage of the disease and MCI subjects with a combination of positive biomarkers are more likely to progress to dementia than those with negative biomarkers [144, 171]. However, none of them alone is sufficient enough to diagnose AD and the consistent finding across studies is that the combination of different biomarkers significantly improves the diagnostic accuracy and prediction of future cognitive decline [50, 92, 64].

Although the use of biomarkers is widely accepted and implemented in research settings, there are important concerns that prevent their extended use in general clinical practice. First, almost all of them are subject to methodologic variations [51], in particular, CSF biomarkers are highly variable across laboratories and techniques [144, 67], which makes it difficult to standardize cut-off points for abnormality. Secondly, validation of the clinical usefulness of biomarkers is still incomplete because most of the studies have been conducted in selected samples that might not be representative of real-world populations [64, 23]. Lastly, CSF and PET biomarkers are invasive and expensive, therefore the access to them is limited in different settings [5].

Biomarkers have opened up the possibility of detecting Alzheimer's pathology before any cognitive symptom, *i.e.* at the preclinical stage. It should be noted that recent studies have shown that most individuals with positive biomarkers for AD are not symptomatic [94] and might remain cognitively healthy during their lifetime [76]. Under biomarker-based criteria these subjects would be diagnosed with the disease without certainty they will actually develop the dementia syndrome. Although this group of individuals, the ones with "preclinical AD" [92] or "at risk of AD" [50], might be useful in clinical trials to test possible early interventions, a clinical diagnosis based only on biomarker positivity would pose an unnecessary burden on them and their relatives [51, 172].

Physiopathology: the underlying chain of events

Despite all the scientific efforts and advances during the last decades, there is little clarity about the mechanisms that lead from the disease-defining proteinopathies to neurodegeneration and cognitive impairment [67, 203]. The underlying biological processes in AD are not well understood and they are, probably, one of the most complex systematic malfunctions of the nervous system that are known [86].

A large body of research, including disease-modifying trials, has been based on the amyloid cascade hypothesis proposed almost 30 years ago [83]. According to this hypothesis, the abnormal aggregation of the amyloid-β peptide is the initial cause of a linear sequence of pathological changes in AD: formation of macroscopic plaques and tau deposits, neurodegeneration, and cognitive impairment. Cross-sectional studies suggest that the entire process from amyloid-β accumulation to the onset of dementia can take up to 20 years [194].

Aligned with the amyloid cascade hypothesis, a temporal model outlining the change of biomarkers was postulated by Jack and colleagues [93]. According to this model, the first biomarker to change is the CSF $A\beta_{42}$ followed by amyloid PET and CSF tau, then FDG-PET and structural MRI become abnormal indicating neuronal injury and atrophy, and finally, the cognitive symptoms appear. Although this biomarker model recognizes that the two proteinopathies might be initiated independently, it does incorporate the idea that amyloid-β changes can accelerate antecedent tauopathy [93]. The basic structure of this model agrees with the amyloid cascade hypothesis in the assumption that pathological changes occur in a linear sequence initiated by amyloid-β accumulation.

The amyloid cascade hypothesis is supported by the observation that the genetic mutations associated with the hereditary form of AD are also known to over-express amyloid-β [104]. Although this hypothesis might explain the physiopathology in hereditary AD, it is not sufficient to explain the development of sporadic AD, which is the most prevalent form of the disease (> 95% of cases). Growing evidence from animal models, human studies, and failed clinical trials for disease-modifying therapies suggests that the relation between amyloid-β accumulation and Alzheimer's dementia is not as direct as stated in the amyloid cascade hypothesis [86, 104]. Studies based on *post-mortem* tissue examination and *in vivo* biomarker analysis have shown that a considerable amount of cognitively intact elder people have amyloid aggregation in their brains that could be considered pathological [86, 94]. Additionally, the amount of amyloid-β plaques does not correlate with neurodegeneration and cognitive decline [104]. With the available evidence, there is no definite explanation yet for how the amyloid-β deposition could lead to neurodegeneration and cognitive impairment [203].

Nowadays it is recognized that amyloid aggregation is not sufficient to cause AD and that a complex interaction between multiple factors could be a better explanation than a simple linear sequence model [86, 92]. The long list of evidence-based alternatives to the amyloid cascade hypothesis favor other initial causes for AD (*e.g.*, failure of autophagy, mitochondrial function, cell cycle control, Ca^{2+} homeostasis) and consider multiple factors that may be responsible for neuronal damage and cognitive impairment such as inflammation, glucose metabolism and DNA damage [86, 203, 110]. More research is needed to disentangle and understand the biological mechanisms and pathological processes involved in sporadic AD physiopathology [51], closing those gaps in knowledge

are important to develop successful disease-modifying treatments [67] and make more accurate predictions about progression at the early stages of the disease.

Neuroanatomical changes

The common characteristic among most of the possible disease models is the final pathway in which neurodegeneration is the pathological feature most proximate to cognitive decline [92]. Recent studies support the observation that neurodegeneration is not the result of a linear cascade of events but the result of the interaction between multiple mechanisms involving positive and negative feedback loops [203]. Accepted neurodegeneration biomarkers capture different scales of this process: FDG-PET and CSF total tau are indicators of neuronal metabolism and damage [92, 48] while MRI biomarkers give information about macro-structural atrophy of brain tissue. These MRI biomarkers are the ones that become abnormal in the closest temporal proximity to the cognitive impairment, however, macro-structural changes could be detected with structural MRI up to ten years before the onset of clinical symptoms [196].

Early research with T_1-weighted MRI established macro-structural landmarks of the disease such as hippocampal and MTL atrophy [91, 170, 44], although these alterations are not specific for AD [188, 61], they are nowadays accepted as topographical biomarkers for disease staging and risk assessment [51, 92]. Longitudinal analyses of grey matter loss [193, 174] have resulted in defined sequential patterns of cortical atrophy starting in temporal and limbic cortices, particularly the entorhinal cortex, progressing with time to frontal and occipital brain regions matching the trajectory of brain lesions observed *post-mortem*, this pattern of atrophy is observed first in the left hemisphere and occurs faster than in its right counterpart [193, 207]. In the advanced stages of the disease, there is noticeable shrinkage in most neocortical areas accompanied by significant expansion of the ventricles.

Atrophy of the hippocampus and MTL are characteristic of typical AD that manifests predominately with memory impairment. Studies in relatively large samples of patients have shown there is heterogeneity in cortical and subcortical grey matter atrophy patterns and this heterogeneity is related to atypical manifestations or differences in compromised cognitive domains [203, 209, 46, 55]. Research with structural MRI [213, 160, 190] has confirmed the existence of atrophy patterns that are consistent with the three subtypes that were defined pathologically [132]: typical AD, hippocampal-sparing AD, and limbic-predominant AD. The observed heterogeneity supports the idea that there might be different pathways that lead to neurodegeneration [203], thus restricting the anatomical markers to independent volumetric measures of very specific regions might be an oversimplification that hampers the identification of patient subgroups.

Alzheimer's has been considered a grey matter disease because the defining brain lesions, intra-neuronal neurofibrillary tangles and extracellular senile plaques, occur mainly in the grey matter. However, there is evidence that pathological changes also occur in the white matter [24], including abnormal levels of $A\beta_{42}$ observed in *post-mortem* tissue examination [28], regional atrophy [167], presence of lesions [169], reduction of microstructural integrity [2, 9, 219, 74, 176, 128], and connectivity failures [100, 187]. White matter micro and macro-structure can be examined *in vivo* using different modalities of MRI. Studies with Diffusion Tensor Imaging (DTI) have revealed diffusion abnormalities

in white matter regions such as the splenium of the corpus callosum, superior, middle and inferior longitudinal fasciculi, corticospinal tracts, and limbic system tracts including the fornix, cingulum bundle, and parahippocampal gyrus [2, 47]. Although some works considered the white matter changes a consequence of neuronal degeneration in the grey matter explained by Wallerian degeneration, there is growing evidence suggesting that abnormalities in the white matter might occur independently and could be detected before grey matter changes [28, 100]. These findings support the idea that other pathological mechanisms like neuroinflammation and prion-like propagation might play an important role in disease physiopathology [24, 100, 110].

Some MRI-based analyses have found that neuroanatomical changes in AD are related to cognitive alterations. For instance, MTL atrophy is correlated with impairment in memory and language [170, 81, 43, 190], and thickness of the parieto-occipital cortex is associated with visuomotor speed [43], visuospatial and executive functioning [190]. Although such correlations exist, and it is known that abrupt damage in certain brain areas can affect specific cognitive functions, it would be inaccurate to attribute certain cognitive skills to a single brain region. Nowadays it is recognized that complex brain functions involve a variety of brain regions functionally connected, and that brain structures are involved in a wide variety of cognitive and functional processes. In this context, the initial neurodegeneration in particular areas might not be enough to cause the characteristic AD decline in specific cognitive functions, and that failures in the functional networks could be the ones responsible for clinical symptoms [100].

Measuring disease outcomes

Clinical stages of the disease are defined by the severity of the symptoms, *i.e.*, the level of cognitive and functional impairment. Therefore, evolution of AD is assessed with neuropsychological tests which evaluate the cognitive abilities and behavior of the patient. Some of the most used scales to determine the severity of the disease are the Clinical Dementia Rating (CDR) [130], and the Alzheimer's Disease Assessment Scale - Cognition (ADAS-Cog) [162]. The neuropsychological test battery used for diagnosis and monitoring also includes screening tests to measure the overall cognition such as the Mini–Mental State Examination (MMSE) [60] and the Montreal Cognitive Assessment (MoCA) [133]; tests to assess the compromise of specific cognitive domains like the Rey auditory verbal learning test (AVLT), the Logical Memory test [208], the Clock Drawing test [77], the Category Fluency test [131], and the Trail Making test [158]; and tests to evaluate the ability to perform everyday activities such as the Functional Assessment Questionnaire (FAQ)[145]. It is important to point out that the neuropsychological tests are the first clinical tests a patient is subject to when there are suspicions of cognitive decline or self-reported memory concern. Therefore these tests are the entry point to other assessments such as evaluation of neuroimages, biomarkers, risk factors, and longitudinal monitoring of cognition and behavior.

Level of cognitive decline determines the disease severity and marks two of the disease milestones: diagnosis of MCI and dementia onset. Monitoring the small cognitive changes consistent with disease progression is a challenging task due to the existence of a long clinically silent phase in AD [76], combined with a large variety of temporary factors that could also alter cognitive performance in individuals with or without the disease. This issue is particularly relevant in the design of clinical trial for disease-

modifying interventions where outcome measures, or end points, are defined to assess the effectiveness of the intervention. Indeed, one of the possible reasons for the long list of failed disease-modifying trials could be the poor performance of outcome measures to detect cognitive changes due to low sensitivity and high measurement variance [125, 166].

Although there are no standard measures for clinical outcomes, and different assessments from the neuropsychological test battery have been used across clinical trials [67], one of the most widely used measures is the ADAS-Cog [166]. This test evaluates multiple cognitive areas and combines the results to give a single number that should indicate the level of overall cognitive impairment. However, some studies have showed that the ADAS-Cog has: low reliability for measuring cognitive change [80], high variance due to measurement errors, ceiling effects of its sub-scores, and it is insensitive for patients in mild stages [125, 106, 166].

Improved measures of disease outcomes should be:

- Robust and sensitive enough to detect cognitive changes at early stages [4].

- Include parts that would be sensitive for heterogeneous forms of AD [64].

- Useful to make predictions about patients' progression and evaluate the risk of developing dementia.

Problem

A lot of unknowns

In sporadic AD, the causes of the disease are largely unknown, the biological mechanisms that lead from proteinopathies to cognitive impairment are unclear [203, 172], the complexity and heterogeneity of those mechanisms are not well understood yet [67], and the transition between what is considered healthy, or normal, ageing and AD is not well defined [4]. There is also no certainty about whether someone with the AD pathological signs will develop some cognitive impairment during their lifetime [172], neither about whether a patient with MCI could go back to cognitively normal, will remain as MCI or progress to dementia.

Although there are accepted biomarkers for diagnosis and monitoring, not a single one is sufficient to diagnose AD or predict disease progression [50, 92, 64]. Moreover, it remains unclear what are thresholds, anatomical distributions, or combinations of abnormal biomarkers that better predict the emergence and evolution of clinical symptoms [4].

What is needed

Existing instruments for patient evaluation and monitoring are not sensitive enough to detect subtle changes, predict progression, and recognize heterogeneous forms of AD. Although there have been several advances in this field, more research is needed to develop and validate markers that help to identify disease patterns or profiles that could predict the clinical course of the disease [177, 4].

Better markers and strategies to identify and quantify the pathological brain changes occurring with disease progression have the potential to impact the clinical management of patients, the design of clinical trials for disease-modifying treatments, improve the assessment of effectiveness for those interventions, and will open the doors to precision medicine [190, 76].

1.2 Computational anatomy in AD

Thanks to their noninvasive nature and increasing availability, MR image modalities have been a tremendous source of information to study brain anatomy abnormalities related with neurodegenerative diseases [14]. As T_1-weighted MRI provides good contrast between tissues, it has been widely used to analyse and localize macrostructural changes in AD such as volume loss or shrinkage of the cortex [65].

Regional volumetry

MR-based volumetry of the hippocampus and MTL are nowadays accepted as topographical biomarkers for AD. Measuring the volume of specific structures is completely dependent on their segmentation. In particular for the hippocampus, its manual segmentation is time consuming, requires extensive training and suffers of high inter-observer variability [12]. For this reason some studies have proposed automatic or semi-automatic segmentation methods [29, 26, 114], or indirect measures of hippocampal atrophy [12], as alternatives to manual delineation of this structure [44]. Volume measures of larger sets of regions can also give information to better characterize the disease, several studies have used automated tools to estimate the volume and cortical thickness of anatomical regions of interest, and analyse which regions help better to discriminate between patients and healthy controls [112, 39, 149, 84, 173, 182]. Most of these strategies rely on pre-defined segmentation of an anatomical template that is then registered to each brain image. Although this process could potentially affect the accuracy of the segmentation, results have shown that volumetric and cortical measures obtained with these approaches can be effectively used to distinguish between patients and healthy controls using T_1-weighted MRI acquired in realistic clinical settings [39, 173, 215].

Voxel-based morphometry

Voxel-based morphometry (VBM) [13] is one of the most common frameworks to perform statistical inference with brain images. The VBM approach can be divided in three key steps:

1. **Spatial Normalisation:** Registration of all images to a common reference space defined by a template image.

2. **Tissue segmentation:** Spatially normalised images are segmented into grey matter (GM), white matter (WM), and CSF. Tissue segmentation maps are often smoothed with a Gaussian kernel to boost the signal-to-noise ratio and alleviate the effect of registration misalignments during spatial normalisation.

3. **Statistical analysis:** Test statistic values are computed at each image voxel resulting in a "statistical parametric map", and finally the corresponding p-values for the tested hypothesis are calculated. This last step needs to take into account that multiple tests are being performed simultaneously (one test per voxel) and resulting p-values need to be corrected accordingly to control for false positives.

These steps of standard VBM are illustrated in Figure 1.3.
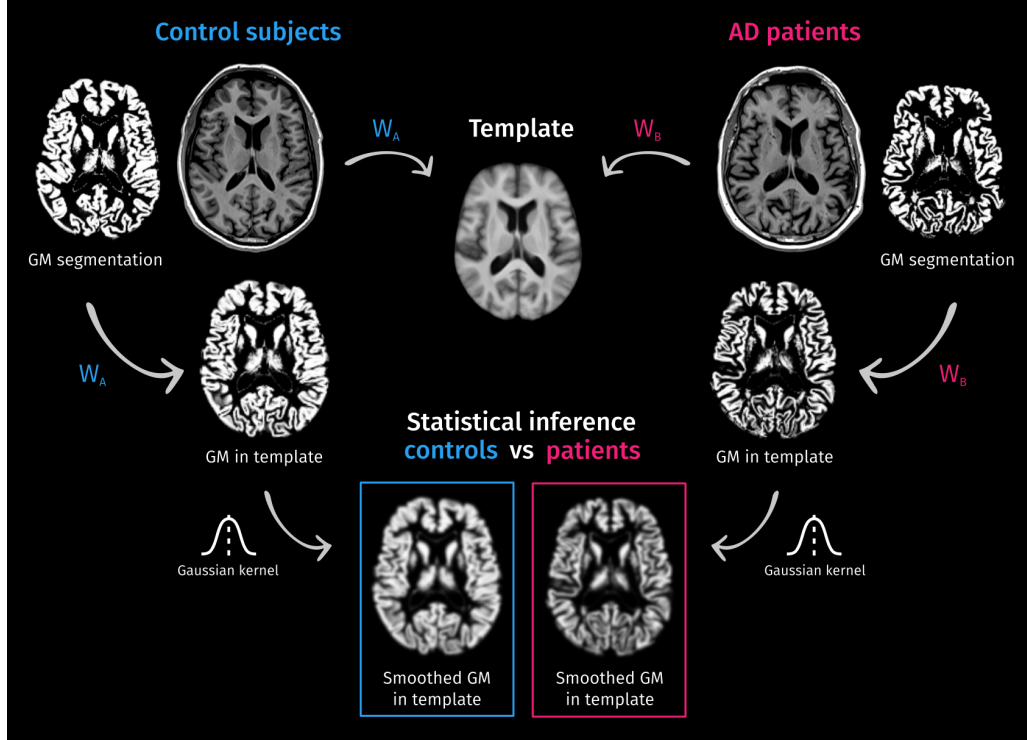
Figure 1.3: Steps involved in voxel-based morphometry (VBM), originally described by Ashburner and Friston [13] to compare the local concentrations of grey matter (GM) between groups of subjects. The VBM approach starts with the spatial normalisation of brain images by registering all of them to a template image. Having the warping from subject to template (W), spatially normalised GM map can be obtained by segmenting normalised images or transforming the GM segmentations [107]. Then, GM maps are filtered with a Gaussian kernel and finally these smoothed GM maps are the inputs for voxel-wise statistical inference.

In AD research, VBM has been applied mainly to find GM differences between AD or MCI patients and healthy controls [81, 126, 196]. Some studies have combined the VBM approach with machine learning methods to automatically classify structural images between AD, MCI and controls or predict progression from MCI to AD dementia. In these approaches, the spatially normalised tissue segmentation maps have been used directly as the inputs of a classifier [105, 32], or have served as an intermediate processing step before dimensionality reduction and feature selection for classification [35, 204, 126, 117, 34, 129, 220, 17].

Other features extracted from structural MRI

Multiple image-based markers for AD diagnosis and monitoring have been proposed and tested in the literature employing traditional image processing methods. In 1998,

Freeborough and Fox published a study relying on texture analysis of T_1-weighted MRI [62] using texture features calculated from the grey level co-occurrence matrix (GLCM) [82] of 2D image slices, this proposal showed promising classification performance when distinguishing between AD patients and controls. Similar texture analysis in the corpus callosum and thalamus also showed significant differences between patients diagnosed with mild AD, amnestic MCI and normal ageing controls [36]. In the same direction, 3D texture markers of the hippocampus have also showed good results classifying between AD, MCI and controls [218, 185].

Shape analysis has been also proposed as a potential image-based marker for AD diagnosis and prognosis. Particularly for sub-cortical structures such as the hippocampus, it was demonstrated that shape features outperform volumetric measures in AD vs controls and MCI vs controls classification tasks [68] and that shape asymmetries are better predictors of dementia onset than size asymmetries [207]. An additional group of proposals includes those inspired by pattern recognition methods such as saliency analysis to find scale-invariant descriptors [195, 165], and pattern matching techniques [147]. All of these also reported competitive classification performance between AD and controls.

Automatic classification of structural MRI

Markers for disease diagnosis and prognosis extracted from structural MRI are often evaluated by using them as inputs for one or more of the following binary classification tasks: cognitively normal (CN) elderly subjects against MCI patients, CN against AD dementia patients, and stable MCI patients against patients who progressed from MCI to AD dementia (sMCI/pMCI). A relatively recent review [157] summarised the reported classification performance of different methods using T_1-weighted MRI and other brain imaging modalities. Table 1.1 present the reported accuracy in some of the previously mentioned works.

Most of the time, direct comparison of classification performance between proposed methods is not possible due to their design and methodological differences in terms of sample size, group selection, inclusion criteria, pre-processing pipelines, cross-validation schemes, and reported evaluation metrics [168]. A couple of studies have performed a direct comparison of different methods by using the same dataset and pre-processing to test pre-defined classification or prediction tasks [32, 21, 121]. For instance, Cuingnet *et al.* [32] evaluated the classification performance of ten approaches using T_1-weighted MRI of 509 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI). For CN vs AD classification, the best performance metrics (81% sensitivity and 95% specificity) were achieved by using the voxels of the GM probabilistic segmentation map directly as features for classification [105]. For CN vs MCI, the best classification results (73% sensitivity and 85% specificity) were obtained by a method performing downsampling of the GM probability map and selection of the voxel locations which better discriminate between AD and CN to finally use that information for classification [204]. A more recent challenge [21] compared several methods for multi-class classification in three diagnostic groups (AD, MCI, and CN) with 354 previously unseen T_1-weighted MRI, the best performance (63% accuracy and 78.8% Area under the ROC-curve) was achieved by a method combining five types of features: volume of seven bilaterally joined regions (including the whole brain), cortical thickness of four lobes and the cingulate gyrus, and hippocampal volume, shape and texture scores. [184].

Authors	Classification accuracy		
	CN/AD	CN/MCI	sMCI/pMCI
Klöppel *et al.*, 2008 [105]	81.1 %	-	-
Fan *et al.*, 2008 [59]	94.3 %	-	-
McEvoy *et al.*, 2009 [123]	89.0 %	-	-
Magnin *et al.*, 2009 [117]	94.5 %	-	-
Rueda *et al.*, 2014 [165]	86.1 %	- %	-
Beheshti and Demirel, 2016 [17]	89.7 %	-	-
Davatzikos *et al.*, 2008 [35]	-	90.0 %	-
Misra *et al.*, 2009 [126]	-	-	81.5 %
Desikan *et al.*, 2009 [39]	95.0 %	95.0 %	-
Gerardin *et al.*, 2009 [68]	94.0 %	83.0 %	-
Sorensen *et al.*, 2016 [185]	91.2 %	76.4 %	74.2 %

Table 1.1: Reported accuracy for each binary classification of some works in the literature using structural MRI. Adapted from Rathore *et al.*, 2017 [157].

Examining the tissue microstructural properties

Macrostructural atrophy caused by AD is accompanied, or preceded, by microstructural changes of tissue integrity. By measuring the diffusion of water molecules in different directions, Diffusion-weighted (DW) MRI provides information about the microstructural barriers of diffusion such as myelin sheaths, axonal and cell membranes. This imaging modality has the unique potential to reveal the organization of tissue at a cellular-scale *in vivo* and non-invasively [197].

Early studies investigating disease-related abnormalities with DW-MRI used the apparent diffusion coefficient (ADC), a parameter of the free diffusion model where isotropic Gaussian diffusion is assumed. In this simple model, the diffusion signal S depends on the applied diffusion weight (b-value):

$$S(b) = S(0)e^{-b \times ADC}$$
$$ADC = -\frac{1}{b} \ln\left(\frac{S(b)}{S(0)}\right) \tag{1.1}$$

The ADC summarizes at voxel level many microscopic processes that affect the diffusion of water molecules [108], hence it captures the alterations resulting from microstructural changes. In the case of AD, one particular study reported that higher ADC in the hippocampus was related to a higher risk of progression from amnestic MCI to AD dementia [101].

The diffusion tensor (DT) model [16] is an extension of the ADC model that incorporates the dependency of the diffusion signal on the directions $u \in \mathcal{S}^2$ of the magnetic field gradient applied during image acquisition, being able to describe anisotropic diffusion:

$$S(b, u) = S(0)e^{-bu^T Du} \tag{1.2}$$

Where the diffusion tensor D is a 3×3 symmetric positive-definite matrix that has associated three orthogonal eigenvectors and three positive eigenvalues λ_1, λ_2 and λ_3. This diffusion tensor is often represented by a 3D ellipsoid as in Figure 1.4.
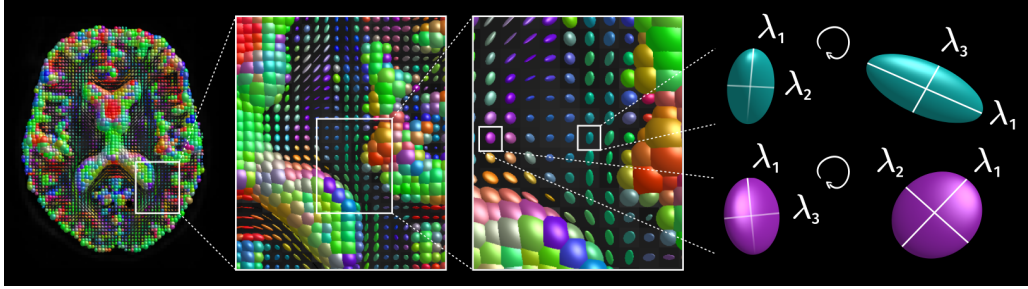

Figure 1.4: A diffusion tensor models the signal in each voxel of the Diffusion-weighted MRI, this is commonly known as Diffusion Tensor Imaging (DTI).

The eigenvalues of the diffusion tensor are used to compute some measures such as the fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD):

$$
\begin{aligned}
FA &= \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \\
MD &= \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \\
AD &= \lambda_1 \\
RD &= \frac{\lambda_2 + \lambda_3}{2}
\end{aligned}
\tag{1.3}
$$

These metrics describe diffusion behavior allowing to infer some properties of underlying brain tissue as illustrated in Figure 1.5.

The vast majority of research investigating microstructural differences between AD patients and controls have used the diffusion tensor (DT) model and its derived metrics to describe tissue diffusivity properties mostly in the WM [3, 47, 176, 219, 2, 189, 37, 109, 122], but also in the GM [212, 113]. Consistent findings across tensor-based studies analyzing WM show increased MD and reduced FA in the splenium, the cingulum bundle including the parahippocampal gyrus, and the superior, middle and inferior longitudinal fasciculus; meanwhile increased FA in crossing-fibre areas such as the corticospinal tract has also been reported in AD patients compared with control subjects [176, 2, 47]. Fewer studies have used DT metrics to study differences in GM areas focusing in certain regions
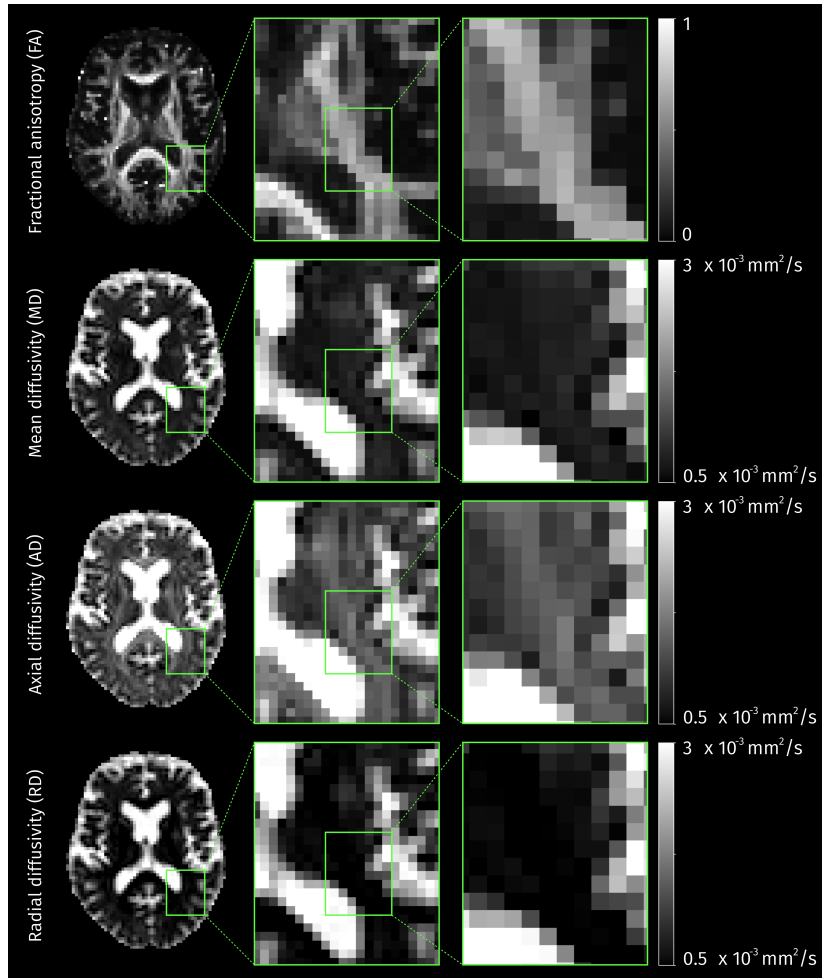
Figure 1.5: Brain maps of metrics derived from the DT model.

of interest, the common finding among them is increased MD in the hippocampus and the posterior cingulate cortex [212, 113, 85].

The DT model was the first diffusion model to be widely adopted in clinical and neuroscience research due to its simplicity [38]. However, this simplicity comes with some important limitations. For instance, it assumes the diffusion displacement probability distribution has a Gaussian form (which is not necessarily true given the complexity of diffusion barriers in the brain tissue), and it cannot represent crossing fibre configurations (which are highly prevalent in WM [98]). Another important limitation comes from what the diffusion tensor metrics are actually capturing given the limited spatial resolution of DW-MRI, in a given voxel it is highly probable that the DT model is not only representing the diffusion of one tissue type (WM or GM) but it is also accounting for the partial volume effects (PVE) with surrounding CSF, therefore DT-derived metrics could be capturing macrostructural atrophy effects rather than microstructural properties [139, 85]. Different extensions or alternatives have appeared to overcome DT model limitations, with different acquisition requirements regarding the number of gradient

orientations and b-values that are needed to estimate model parameters.

For example, diffusion kurtosis imaging (DKI) quantifies the non-gaussianity of diffusion in biologic tissues [97] by adding an excess kurtosis term to the model. Then, the signal attenuation along a certain diffusion direction u is modelled as:

$$\ln\left(\frac{S(b,u)}{S(0)}\right) = -bD_u + \frac{1}{6}b^2D_u^2K_u^2 \tag{1.4}$$

Where D_u and K_u are estimates for the diffusion coefficient and diffusional kurtosis in the direction u. Estimation of these parameters requires fitting a quadratic function of the b-value, therefore data needs to be acquired with at least two non-zero b-values, being one of them relatively high (≥ 1500 s/mm^2) to allow better appreciation of non-gaussianity. This type of DW-MRI, acquired with multiple non-zero b-values, is referred as "multi-shell" given that the acquisition gradients lie in multiple spheres.

The constrained spherical deconvolution (CSD) method to model WM is related to the notion of spheres in the space of acquisition gradients. Given a b-value, the diffusion signal can be represented as a function over the unit sphere using spherical harmonics (SH) basis functions (see Figure 1.6).

(a) $b = 700$ s/mm^2 (b) $b = 1000$ s/mm^2 (c) $b = 2800$ s/mm^2
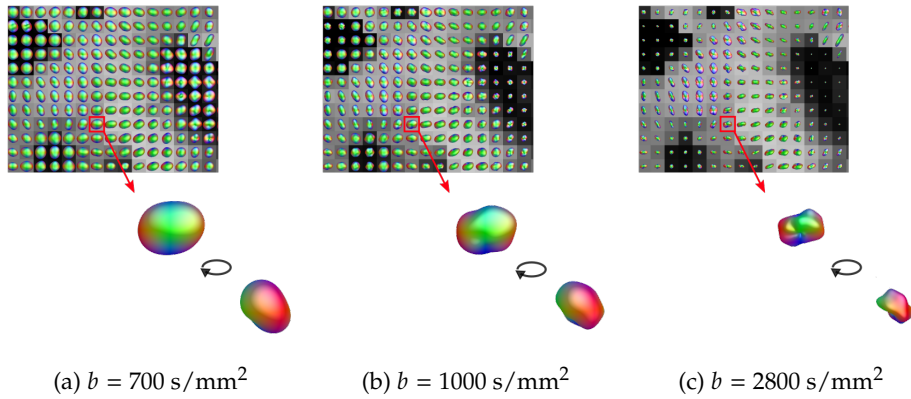
Figure 1.6: Representation of the diffusion signal in spherical harmonics for three different values of diffusion-weight (b-value).

Then, the diffusion signal observed at a constant b-value is modelled as the spherical convolution of a fibre orientation distribution function (fODF) with a single fibre response function [199], as illustrated in Figure 1.7.

The fODF can be "recovered" as the deconvolution of a single fibre response function (that needs to be estimated) from the observed signal, while enforcing non-negativity of the fODF lobes [198]. The fODF is a continuous function that could represent any underlying fibre configuration, effectively overcoming the "crossing-fibre" problem [38], an example of fODF map is shown in Figure 1.8.

These examples of more advanced models, DKI and CSD, have been recently used to investigate AD-related changes by comparing their corresponding diffusion-derived metrics between control subjects and AD patients. For instance, exploratory analysis suggested mean kurtosis could be more sensitive than FA or MD to detect initial degeneration
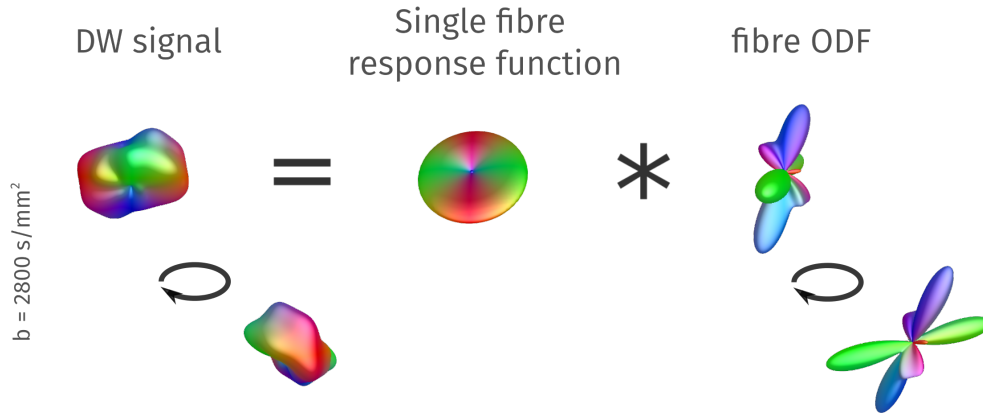
Figure 1.7: When DW-MRI is acquired with a constant non-zero b-value, the observed signal in a voxel can be modelled as the spherical convolution of a fibre orientation distribution function (ODF) with a single fibre white matter response function.
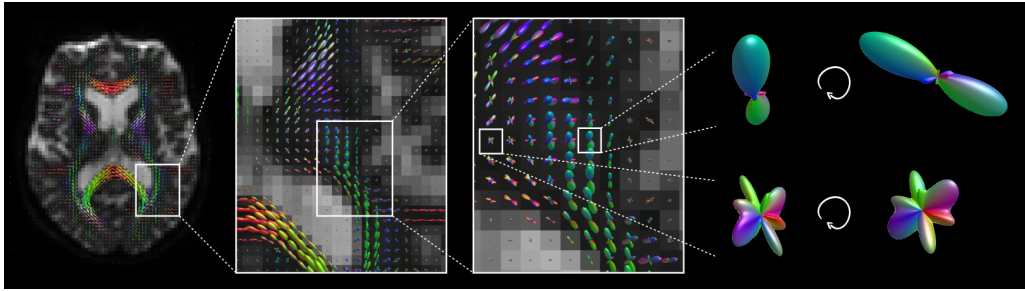
Figure 1.8: The underlying white matter is modelled with continuous fibre orientation distribution functions (fODF) represented in spherical harmonic basis.

of some WM structures such as the splenium of the corpus callosum and the corona radiata [183]. Analysis of fibre-specific measures derived from fODF showed differences of WM micro and macrostructure between AD patients and controls in specific fibre tracts including the cingulum bundle, the splenium and genu of the corpus callosum, the uncinate fasciculus, and arcuate fasciculus.

1.3 This Thesis

Exploration and evaluation of new markers to identify and quantify changes related to AD progression is a highly relevant research path that could, in the mid-term, improve the management and monitoring of patients, and help the evaluation of potential disease-modifying treatments.

It is recognized that alterations in brain anatomy are the pathological features most proximate to cognitive decline [92]; therefore, neuroanatomical markers and neuropsychological information provide direct information about disease progression.

In this context, this thesis presents a set of data-driven strategies that identify and quantify anatomical and cognitive pathological patterns associated with different clinical stages along the AD continuum. These strategies constitute the three main contributions of this thesis:

- In the first contribution, we propose a strategy that captures changes in brain anatomy by comparing the content distribution in different anatomical regions using information from T_1-weighted MRI. We demonstrate this quantitative strategy is useful for the automated classification of brain images between patients at different stages and controls. Furthermore, this characterization automatically finds out a multidimensional pattern of AD progression which is directly related to anatomical changes in specific areas. This contribution is presented in Chapter 2 and has been published in a journal article:

  - Diana L. Giraldo, Juan D. García-Arteaga, Simón Cárdenas-Robledo, Eduardo Romero. *Characterization of brain anatomical patterns by comparing region intensity distributions: Applications to the description of Alzheimer's disease*. Brain and Behavior. 2018; 8:e00942. `https://doi.org/10.1002/brb3.942` [71]

- The second contribution presents a comprehensive neuroimaging approach for the study of AD-related abnormalities in brain anatomy combining multiple interrelated measures of tissue integrity derived directly from Diffusion weighted MRI. Differences of WM properties and tissue compositions between MCI, ADD patients and age-matched cognitively normal subjects are investigated, as well as the possible correlations of diffusion-derived measures with CSF biomarkers. This part of the thesis work is presented in Chapter 3. Part of this work was presented at a conference:

  - Diana Giraldo, Hanne Struyfs, David A. Raffelt, Paul M. Parizel, Sebastiaan Engelborghs, Eduardo Romero, Jan Sijbers, Ben Jeurissen. *Fixel-Based Analysis of Alzheimer's Disease Using Multi-Tissue Constrained Spherical Deconvolution of Multi-Shell Diffusion MRI*. International Society of Magnetic Resonance in Medicine. Honolulu, USA. 2017. [70]

  A manuscript has been submitted for publication to a journal.

- The third contribution presents a data-driven method to characterize the cognitive state of MCI patients with a set of domain-specific scores obtained by learning to combine and weight sub-scores from the neuropsychological test battery. Using machine learning methods, we show the developed scores highlight subgroups

of MCI patients who exhibit different risks of progression to AD dementia and have better classification performance than standard outcomes when predicting conversion from MCI to dementia up to 5 years after neuropsychological evaluation. This contribution is presented in Chapter 4 and has been published in a journal and a conference:

– Diana L. Giraldo, Jan Sijbers, Eduardo Romero *Quantification of cognitive impairment to characterize heterogeneity of patients at risk of developing Alzheimer's disease dementia*. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2021; 13(1):e12237. `https://doi.org/10.1002/dad2.12237` [73]

– Diana L. Giraldo, Jan Sijbers, Eduardo Romero. *Quantifying cognition and behavior in normal aging, mild cognitive impairment, and Alzheimer's disease*. Proc. 13th International Conference on Medical Information Processing and Analysis. San Andrés - Colombia, 2017. `https://doi.org/10.1117/12.2287036` [72]

Finally, Chapter 5 presents some conclusions, discuss the potential impact of the contributions and suggest some possible research directions for future work.

Chapter 2

Comparing region intensity distributions

2.1   Introduction

A large number of studies have proposed automatic methods to extract features and classify T_1-weighted MRI between controls, MCI, and AD patients [32, 21, 157]. In a relatively recent review, Rathore *et al.* [157] established three automatic classification categories based on the feature extraction method from structural MRI: density maps-based, cortical surface-based, and pre-defined region-based. As pointed out by the authors, most investigations in the latter category use only hippocampus features since changes in this region are well known. These studies do not consider differences in other brain regions out of the MTL, ignoring subtle changes and possible complex patterns of the disease compromising multiple regions. On the other hand, density maps-based methods inspired by VBM classify structural MRI using whole-brain information. However, their adoption in clinical practice remains limited or almost not existing. One important reason is that, in many cases, the high-dimensional features are not easily interpretable in terms of spatial patterns of anatomical changes and cannot be related to the clinical picture, so they appear as "black boxes" to clinicians.

Despite the promising classification results of machine learning methods, their contribution to the characterization and understanding of the disease progression remains limited. Another drawback of these automatic classification approaches is that, although they do compare brains, their notion of distance has no meaning in terms of the disease progression hampering their use for exploring the pathways of the AD continuum.

In this work, we introduce a strategy that allows the quantification of brain differences by comparing the intensity distributions of several anatomical regions in the whole brain. An underlying hypothesis of this approach is that the differences between AD patients and controls are correlated to tissue constituents, a feature mirrored by the composition of gray level intensities in T_1-weighted MRI. The guiding principle incorporated in this proposal is that patients do not follow a single unique direction when transitioning from healthy ageing to AD. Instead, AD patients are assumed to drift away from a healthy state in multiple possible directions, i.e. control subjects form a relatively compact cluster whereas AD cases tend to separate towards pathological states in more than one direction.

2.2   Methods

The basis for the method is the quantitative measurement of differences between subjects in separate regions of the brain. The process can be roughly divided in two stages. First, each anatomical region is described by comparing its intensity histograms between all subjects in the sample. The second part consists on extracting regional features and performing the classification of subjects between AD/MCI patients and controls using ensemble classifiers (See Figure 2.1).
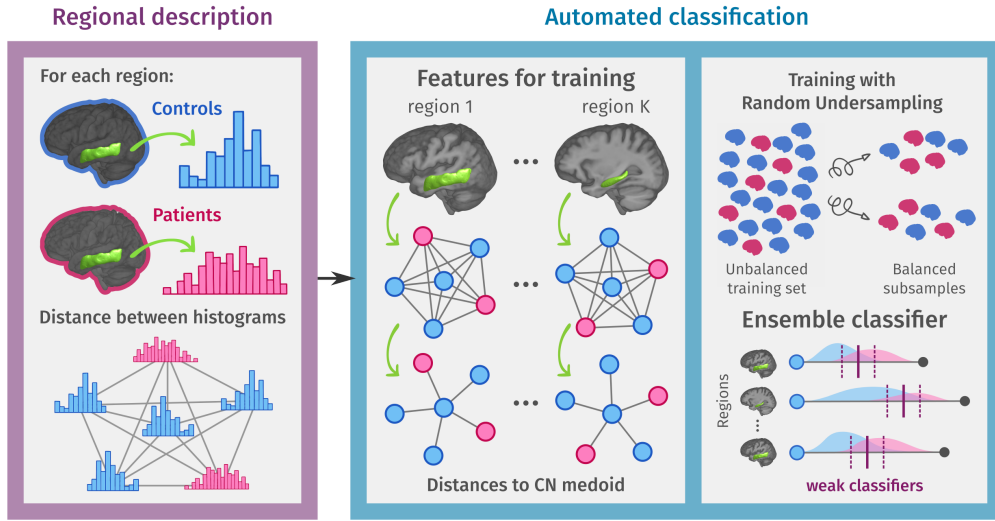
Figure 2.1: Overview of the proposed methodology. First, each one of the considered regions is described using structural MRI from all participants, including cognitively normal (CN) controls and MCI/AD dementia patients. In the second part, features for automated classification consist of the distances to the CN medoid (chosen as the reference point), and ensemble classifiers are trained with all regional features following a random undersampling boosting strategy to account for class imbalance in the sample.

2.2.1   Region description

2.2.1.1   Coarse brain parcellation

The very first step is to parcellate each brain image into a set of anatomical regions. To obtain a coarse parcellation, brain volumes were registered to the MNI152 structural template with an affine transformation calculated using the FSL (FMRIB Software Library) linear registration tool Flirt [95, 96], and then the Harvard-Oxford brain atlas (RRID:SCR_001476) was used to partition each registered brain into 96 cortical regions (48 per hemisphere) and 17 subcortical regions. The linear registration approach results in slightly displaced inter-subject anatomic regions, yet this is unlikely to affect the distribution of gray levels within these regions. Such claim may be supported by the fact that overlap between partitioned brains is at least 97%.

2.2.1.2 Region similarity across subjects

Ultimately the objective of our analysis is to be able to quantify the level of similarity or dissimilarity between subjects. Furthermore, it is expected that this distance is related to the diagnostic groups subjects belong to, i.e. two control subjects should have a smaller distance between them than a control and an AD patient. In this work, the tissue distribution of each anatomical region was described by its intensity histogram, and differences of regional tissue distribution between subjects were quantified by measuring the distances between histograms (Figure 2.2).

Region segmentation Region intensity histograms Distance between histogram
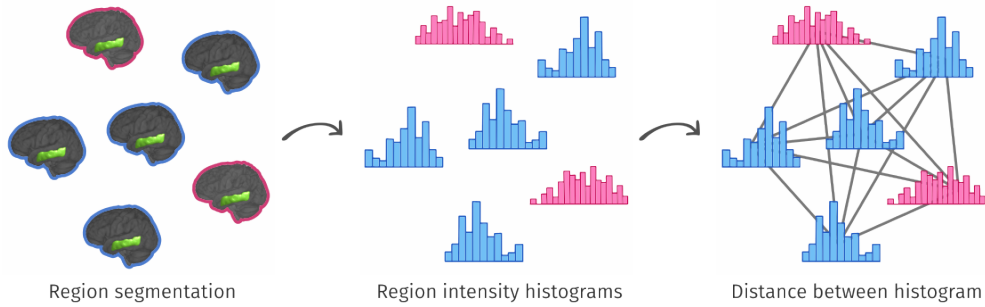
Figure 2.2: Region characterization starts with extracting equivalent anatomical regions (left). Regional information is represented the median-centered histogram of intensities (center). Histograms of the same region for different subjects are compared with the Earth-Mover's Distance (right).

All intensity histograms describing region anatomy had 64 bins, however, the bin cuts were defined individually for each case taking into account the intensity range for the whole brain image. To make histograms comparable and eliminate differences due to image intensity range, all histograms were shifted so that the center of mass was aligned to the central bin of the histogram. Then, histograms corresponding to the same region for different subjects were compared with the Earth-Mover's Distance (EMD) [164].

Earth-mover's distance formulation

The EMD calculates the minimum cost of transforming one histogram into another by solving a linear optimization problem in which certain units of the $\mathcal{S} = \{S_1, \ldots, S_n\}$ histogram, have to be moved to fill the m bins of histogram $C = \{C_1, \ldots, C_m\}$

The movement of one unit from bin $i \in \mathcal{S}$ to bin $j \in C$ has an associated cost p_{ij}. The solution consists in a set of movements $\{x_{ij}^*\}_{i,j=1}^{n,m}$ that form C and minimize the total movement cost.
The optimization problem can be written in terms of the amount of *"earth"*, in this case

units x_{ij}, that is moved from bin $i \in \mathcal{S}$ to bin $j \in \mathcal{C}$, as follows:

$$
\begin{aligned}
\underset{X}{\text{minimize}} \quad & \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{m} x_{ij} \le S_i, \text{ for } i \in \{1, \ldots, n\} \\
& \sum_{i=1}^{n} x_{ij} \ge C_j, \text{ for } j \in \{1, \ldots, m\} \\
& x_{ij} \ge 0, \text{ for } i \in \{1, \ldots, n\} \text{ and } j \in \{1, \ldots, m\}
\end{aligned}
\tag{2.1}
$$

In this case, the cost of moving one unit is set to the absolute distance between bins, *i.e.*, $p_{ij} = |i - j|$. Given the solution $\{x_{ij}^*\}_{i,j=1}^{n,m}$, the EMD between $\mathcal{S}$ and $\mathcal{C}$ is the normalized total cost:

$$
\text{EMD}(\mathcal{S}, \mathcal{C}) = \frac{1}{\sum x_{ij}^*} \sum_{i=1}^{n} \sum_{j=1}^{m} |i - j| x_{ij}^*
\tag{2.2}
$$

When the compared histograms have the same integral, as in this work, the problem is symmetric and the EMD is a metric equivalent to the Wasserstein's distance. A minimal example of the EMD between two histograms p and q is shown in Figure 2.3.
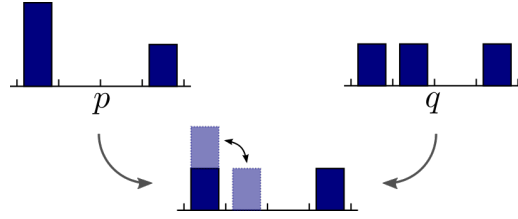
Figure 2.3: In this case the EMD between p and q is the cost of moving one unit from one bin to the next divided by the total mass: 1/3. Note that, in this case, the distance function is symmetric ($\text{EMD}(p, q) = \text{EMD}(q, p)$).

2.2.2 Automated classification of brain images

For each one of the considered anatomical regions, the result of the previous step is a matrix of pairwise distances between subjects in the data sample. Taking distance to a reference point incorporates the guiding principle that patients drift away from a healthy state, therefore we chose the medoid of the control group as such reference point (Figure 2.4). The medoid is the element of a set with the minimal mean distance to the other elements in the set, *i.e.* for a given set A and a distance function δ the medoid is defined as:

$$
\text{medoid}(A) = \underset{x \in A}{\arg\min} \sum_{y \in A} \delta(x, y)
\tag{2.3}
$$

Pairwise distances
between all cases

Pairwise distances
between controls

Distance to medoid
of controls

Figure 2.4: Given the pairwise distances between histograms of the same region across all subjects in the sample (left), a reference point was chosen by selecting the "most central" case within the control group (middle), distances to that reference point were taken as the regional features for classification (right).

Distance to the medoid of controls captures how much region anatomy is drifting away from a group representing healthy anatomy. Distance values for all regions characterize the whole-brain anatomy for each subject and constitute the features for classification. Two binary classification tasks were considered: AD patients vs controls and MCI vs controls. For this purpose, ensemble classifiers were trained with an Adaptative Boosting (ADABoost) [63] approach that iteratively updates the weights of various weak classifiers, giving more importance to samples misclassified in earlier rounds. Simple thresholds of the features were used as weak classifiers. To alliviate the class imbalance in training data, random undersampling of data was used during boosting, an strategy known as RUSBoost [175]. Once ensemble classifiers were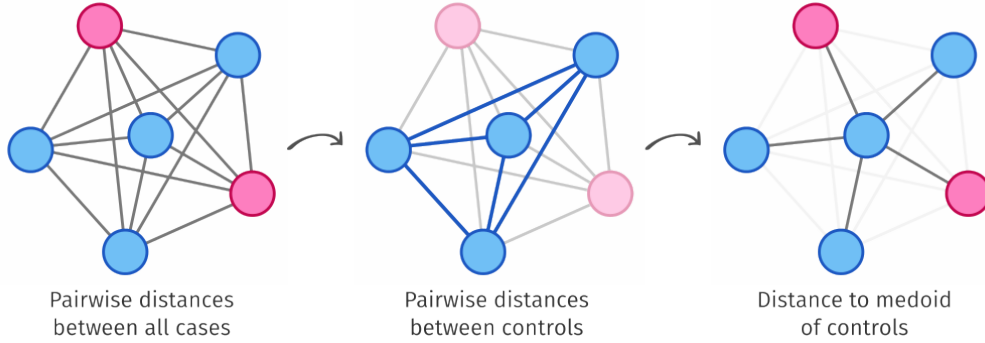 trained, each the relative importance of each feature was computed as the weighted sum of mislabeled classes for each predictor. The importance of the features for each classification task says how much a regions helps to differentiate between groups and therefore is an indicator of the degree to which each region is affected by the disease.

2.3 Evaluation

2.3.1 Data

The proposed strategy was evaluated using T_1-weighted MRI from a subset of cases in the Open Access Series of Imaging Studies (OASIS-1) database [120]. The sample consisted of 136 cases between 60 and 80 years old, from which 66 were the control group (CN), 50 corresponded to MCI patients, and 20 were patients diagnosed with mild AD. The description of each diagnostic groups in terms of age, gender and cognitive scores is shown in Table 2.1.

Structural MRI in OASIS-1 database were acquired with 1.5 T Vision scanners (Siemens, Erlangen, Germany), using magnetization prepared rapid gradient-echo (MP-RAGE) sequences. Raw T_1-weighted MR images have a voxel size of $1 \times 1 \times 1.25 mm^3$, with a

Group	N	Age	Gender (F/M)	CDR	MMSE
CN	66	70.8 ± 5.6	48/18	0	29.1 ± 1.1
MCI	50	72.8 ± 5.0	28/22	0.5	26.0 ± 3.5
AD	20	74.3 ± 4.3	13/20	1	20.8 ± 3.7

Table 2.1: Description of diagnostic groups from OASIS including their scores for the Clinical Dementia Rating (CDR) and Mini-mental state examination (MMSE).

resolution of $256 \times 256 \times 128$. Images were spatially warped into the 1988 atlas space of Talairach and Tournoux with a rigid transformation, averaged motion-corrected, skull-stripped, and finally gain-field corrected [120]. Voxel size after pre-processing is $1 \times 1 \times 1mm^3$ with image resolution of $176 \times 208 \times 176$. For more detailed information about the database see `https://www.oasis-brains.org/`.

Generalization of the presented method was tested with a different set of T_1-weighted MRI from the Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD) database [118]. This sample was composed of 23 healthy controls and 46 subjects diagnosed with probable Alzheimer's disease. The distribution of age, gender and clinical scores of this dataset is presented in Table 2.2.

Group	N	Age	Gender (F/M)	CDR	MMSE
CN	23	69.7 ± 7.1	11/12	0	29.4 ± 0.8
AD	46	69.3 ± 7.2	27/19	1 ± 0.4	19.2 ± 4.0

Table 2.2: Description of diagnostic groups from MIRIAD database.

Images in MIRIAD were acquired with a 1.5 T Signa MRI scanner (GE Medical systems, Milwaukee, WI), using a T_1-weighted Inversion Recovery Prepared Fast Spoiled Gradient Recalled (IR-FSPGR) sequence. Other imaging parameters were: matrix size of 256×256 and 124 $1.5mm$ coronal partitions. Pre-processing of these images included warping into the Talairach and Tournoux atlas and skull-stripped using FSL tools [96].

2.3.2   Cross-validation

Two different cross-validation schemes were used to test the automated classification between groups: the first evaluation was done only with data from OASIS following a leave-one-out scheme, i.e. iteratively training with the whole set but one and then using the resulting classifier to classify the case set aside. The second scheme aimed to evaluate the generalizability of the proposed characterization by training the classifier with data from OASIS database and testing it with data from MIRIAD database.

Classification performance was assessed via a receiver operating characteristic curve (ROC) calculating its area under the curve (AUC) and equal error rate (EER). The instance

of the curve with the best trade-off between false-positive rate and false-negatives negative was selected to report the sensitivity and specificity.

2.4 Results

2.4.1 Classification between patients and controls

The resulting ROC curves for classification experiments with OASIS data are shown in Figure 2.5. When classifying between controls and AD cases, the EER is 0.1 and the AUC is 0.92, as the EER indicates the best trade-off between false positives and false negatives, the sensitivity and specificity of this classification is 0.9. Classification between controls and MCI patients shows an EER of 0.3 and AUC of 0.74, implying a sensitivity and specificity of 0.7.
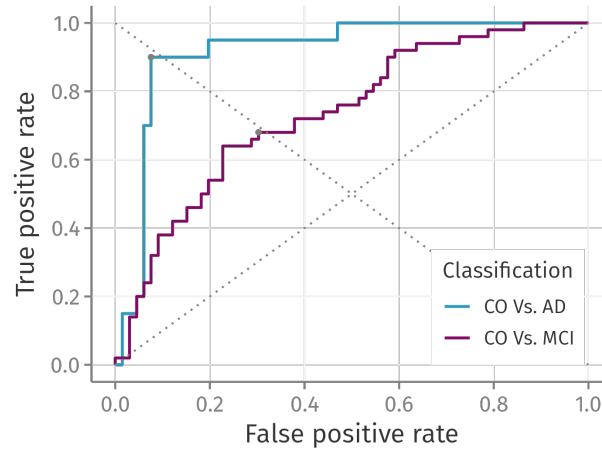
Figure 2.5: Receiver operating characteristic curves for classification experiments within OASIS database. Classification between controls and AD patients (blue line) gave an AUC of 0.91 and EER of 0.1 (False positive rate = 0.1, True positive rate = 0.9). Classification between controls and MCI (purple line) resulted in an AUC of 0.74 and EER of 0.3 (False positive rate = 0.3, True positive rate = 0.7).

Classification across databases resulted in the ROC curve shown in Figure 2.6, with an AUC of 0.92. According to the decision threshold with the best trade-off between errors, a sensitivity of 85% could be achieved with 91% of specificity. These results show a good overall performance with a high accuracy. The errors consist mostly of False Positives (7 cases) whereas the number of False Negatives remains relatively low (2 cases).

2.4.2 Region importance

For each classification task, anatomical regions were ranked according to the average of importance across iterations of the leave-one-out validation scheme with OASIS data.
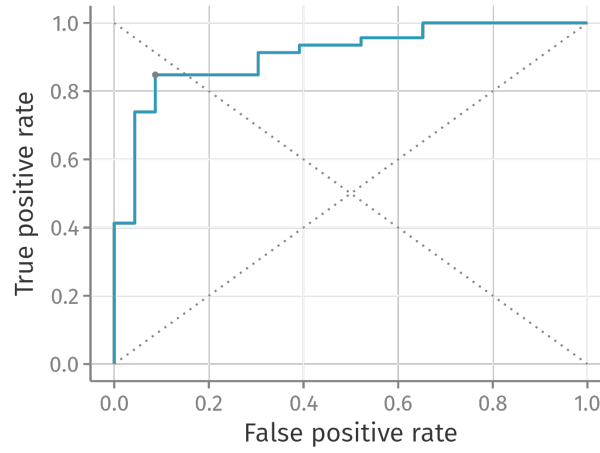
Figure 2.6: Receiver operating characteristic curve for the classification between controls and AD patiens of MIRIAD cases when the classifier was trained with OASIS cases. This classification showed an AUC of 0.92 while the best trade-off between the two types of errors is achieved with a False positive rate of 0.09 with a True positive rate of 0.85.

The ten most relevant regions to discern between controls, AD and MCI patients are shown in Tables 2.3 and 2.4, respectively.

In the case of CN vs AD classification, the importance to distinguish between groups is concentrated in a few regions (shown in Figure 2.7), the top 10 most relevant features summed more than 66.5% of the importance. It is reasonably expected that only the hippocampi (ranked first and third) account for 24% of the importance.

Figure 2.7: The ten most relevant regions for automated classification between controls and AD patients. Names of the regions and their relevance are presented in Table 2.3.

Provided that anatomical changes in MCI are not expected as evident as they might be in mild AD, differences are subtle and more regions need to be taken into account to distinguish between MCI and controls. The top ten most relevant regions to classify between these two groups account for less than 37% of the importance (shown in Figure 2.8). The observation that importance is more spread across regions hints that early

Rank		Region	Importance (%)
1	Right	Hippocampus	15.51
2	Right	Planum Temporale	13.35
3	Left	Hippocampus	8.40
4	Left	Thalamus	7.16
5	Right	Paracingulate Gyrus	4.83
6	Right	Middle Temporal Gyrus, anterior division	4.38
7	Left	Insular Cortex	4.17
8	Right	Putamen	3.59
9	Left	Frontal Orbital Cortex	2.71
10	Right	Amygdala	2.39

Table 2.3: Top ten most relevant regions for the classification between controls and AD patients.

structural changes might be more complex and not restricted to the already known anatomical areas.

Rank of relevance: ■1 ■2 ■3 ■4 ■5 ■6 ■7 ■8 ■9 ■10

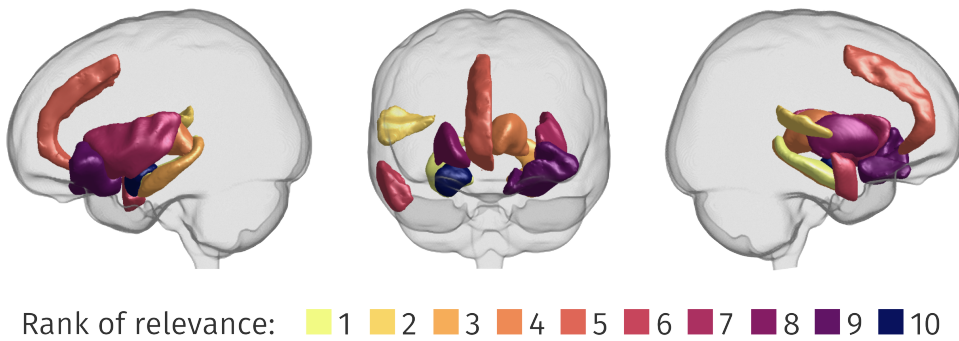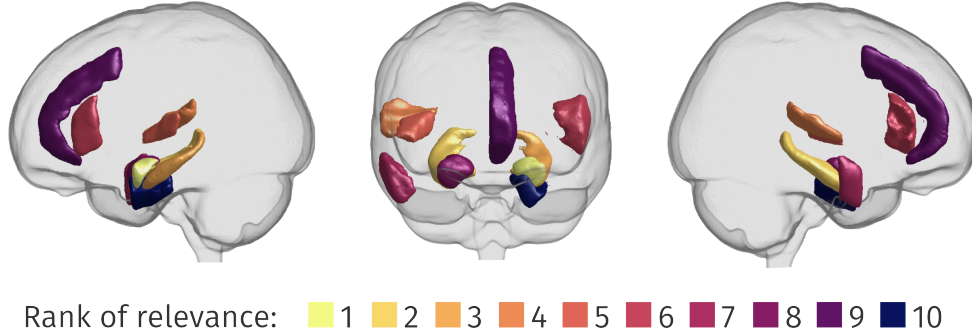Figure 2.8: The ten most relevant regions for automated classification between controls and MCI patients. Names of the regions and their relevance are presented in Table 2.4.

It is worth mentioning that the feature importance value says how much information it adds to the other features, that is to say a region with little relevancy is in any case informative, but this information might be redundant and shared by other regions. This statement is illustrated by the distributions of the two hippocampi feature values in Figure 2.9: although features for both regions show similar distributions and strong inter-class separation, the right hippocampus is more relevant than the left one, which shows almost half of the importance in Table 2.3. This difference appears during the classifier training phase: the weak classifier, based on the left hippocampus, mostly confirms the results of its right counterpart, i.e. since the same cases are discriminated by both left and right weak classifiers, the former is considered redundant because it

	Region	Importance (%)
Left	Amygdala	6.35
Right	Hippocampus	5.41
Left	Hippocampus	4.78
Right	Planum Temporale	3.56
Right	Heschl's Gyrus	3.24
Left	Inferior Frontal Gyrus, pars triangularis	3.10
Right	Middle Temporal Gyrus, anterior division	2.95
Right	Amygdala	2.55
Left	Paracingulate Gyrus	2.40
Left	Parahippocampal Gyrus, anterior division	2.18

Table 2.4: Top ten most relevant regions for the classification between control subjects and patients with mild cognitive impairment.

does not give much additional information and its weight is decreased in the ensemble of classifiers. Because of this, regions showing prevalent differences between groups are ranked higher, whereas those regions useful to classify particular cases are ranked lower.

When analyzing the distributions of the feature value for the most relevant regions and their opposite hemisphere equivalences (shown in Figure 2.9), there are strong observable differences between CN, which form relatively compact groups, and AD patients, which tend to be more scattered and diverge from CN, while the MCI group falls between them. This trend is particularly remarkable in the amygdala, hippocampus, planum temporale and thalamus, where the CN and AD feature value distributions look well separated.

2.5   Discussion

This section presents a fully automated strategy that detects characteristic structural brain patterns associated to the presence of the Alzheimer's disease. The method derives a regional descriptor that captures the changes in tissue constituency which is characteristic of any neurodegenerative disease. This regional descriptor is based on the comparison of intensity histograms between subjects, assuming gray levels in structural MRI correlate with tissue composition, an assumption that is supported by the fact that image contrast in T_1-weighted MRI is the product of relaxation differences between tissue types.

The approach herein described has an advantage over other automated classification methods since it is clinically interpretable by standing out actual patterns of the disease. Machine learning based analyses have helped to move from the classical local approaches in pre-defined regions to the exploration of more complex descriptors using artificial vision techniques. Although such descriptors are useful to separate groups of individuals, most of these features are not useful for finding out anatomo-physiological correlations
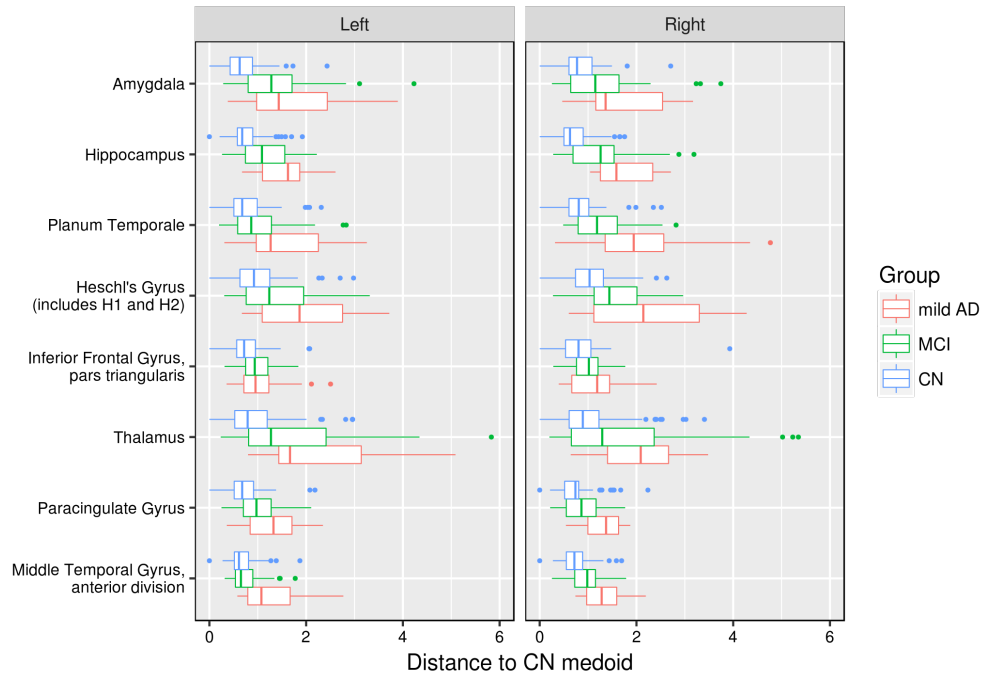
Figure 2.9: Distributions of distances to regional reference, the CN medoid, for the most relevant regions together with their contralateral equivalent.

that enhance the understanding of a particular disease.

The characterization presented here also captures disease progression patterns in multiple directions determined by anatomical changes in different brain regions. This is illustrated by Figure 2.10 which shows the median distance (per group) to the CN medoid for a group of brain regions.

This figure also suggests that equivalent regions in the two hemispheres could not show the same progression rate and then the level of discrimination between subjects is better when the left and right equivalent regions are taken separately, this claim was corroborated with additional classification experiments with the OASIS database in which left and right regions were combined. As shown in Table 2.5, performance measures (AUC and EER) are slightly worse when for both classification tasks.

	CO vs. MCI		CO vs. AD	
	AUC	EER	AUC	EER
Separating hemispheres	0.74	0.30	0.92	0.10
Combining hemispheres	0.72	0.36	0.90	0.20

Table 2.5: Comparison of classification performance when information from the same region in both hemispheres is combined

Figure 2.10: Group median distance to CN medoid for a group of anatomical regions. The directions of these polar graphics correspond to the 8 most relevant regions in the classification task: 1. amygdala, 2. hippocampus, 3. planum temporale, 4. Heschl's gyrus, 5. inferior frontal gyrus, 6. thalamus, 7. paracingulate gyrus and 8. middle temporal gyrus (anterior division).

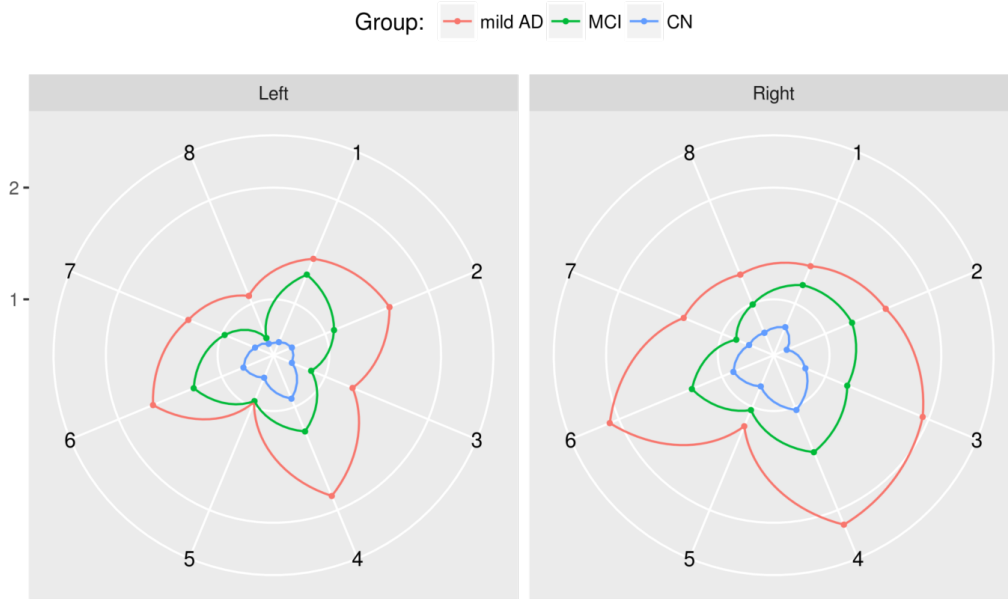The presented strategy did effectively discriminate between patients and controls. Two previous works performed an automated classification between CN and AD using exactly the same data and validation scheme but different feature extraction approaches. In the first one, the work by Toews *et. al.* [195], they propose a technique to learn local scale-invariant anatomical features by evaluating saliency in image scale-spaces and classify cases depending on the occurrence of such features. Following this feature-based morphometry approach, they achieved an EER of 0.2 when classifying between CN and AD. The second work we can directly compare with is the one by Rueda *et. al.* [165] which presents a strategy that fuses different feature-scale saliency maps and uses this information to feed the classifier, classification between CN and AD following this strategy achieves an EER of 0.14. The classification results presented in this chapter constitute an improvement over both works with an EER of 0.1 in the same experiment.

Besides outperforming previous works using the same data and validation scheme, the classification between AD patients and controls achieved 90% sensitivity and specificity while the best performing methods out of 10 compared in [32] reported up to 81% sensitivity and 95% specificity (using a different database). Similar or slightly worse classification results were reported for methods relying on voxel-based morphometry, region volumetry or different feature extraction methods [35, 217, 211, 59, 117, 147, 105, 216]. It should be noted that beyond developing a fully automatic classification pipeline, this strategy finds out a multidimensional expression of AD progression, which is directly related to anatomical changes in specific brain regions. The quantitative measures of anatomical changes proposed here can be used to describe and evaluate brain images in

terms of this multidimensional pattern.

2.6 Products

Journal paper

- **Diana L. Giraldo**, Juan D. García-Arteaga, Simón Cárdenas-Robledo, Eduardo Romero. *Characterization of brain anatomical patterns by comparing region intensity distributions: Applications to the description of Alzheimer's disease*. Brain and Behavior. 2018; 8:e00942. `https://doi.org/10.1002/brb3.942` [71]

Indirect products: conference papers

- **Diana L. Giraldo**, Juan D. García-Arteaga, Eduardo Romero. *Finding regional models of the Alzheimer disease by fusing information from neuropsychological tests and structural MR images*. Proc. SPIE Medical Imaging 2016. San Diego - USA, 2016. `https://doi.org/10.1117/12.2217021`

- Sebastian Maglioni, **Diana L. Giraldo**, Juan Duarte, Nelson Velasco, Eduardo Romero. *Description of brain volumetric changes in Alzheimer disease using region-based morphometry*. Proc. 14th International Symposium on Medical Information Processing and Analysis. Mazatlan - Mexico, 2018. `https://doi.org/10.1117/12.2511533`

- Santiago Silva, **Diana L. Giraldo**, Eduardo Romero. *Sulci characterization to predict progression from mild cognitive impairment to Alzheimer's disease*. Proc. 15th International Symposium on Medical Information Processing and Analysis. Medellín - Colombia, 2019. `https://doi.org/10.1117/12.2540437`

Chapter **3**

Investigating tissue-specific abnormalities in AD with DW-MRI

3.1 Introduction

Several studies have investigated the effect of AD on brain anatomy using MRI, most of them focused on grey matter (GM) degeneration and cortical atrophy patterns [44, 29, 105, 112, 39, 68, 147, 32, 21, 137, 33, 191, 215]. In contrast to structural MRI, diffusion-weighted MRI (DW-MRI) allows revealing microstructural effects of AD, mostly in the white matter (WM) where the diffusion of water is shaped by the architecture of axonal membranes and myelin sheaths. Most diffusion studies in AD and MCI have used the diffusion tensor model and its derived metrics, such as fractional anisotropy (FA) and mean diffusivity (MD), to detect WM degeneration induced by the disease. Consistent findings across tensor-based studies reveal a widespread increase of MD in the WM and decrease of FA in certain WM areas including the splenium, the cingulum bundle, the superior longitudinal fasciculus, the uncinate fasciculus, and the parahippocampal gyrus [176, 47, 2, 122, 45]. Research in the early stages of AD has suggested that WM microstructural degeneration is not always secondary to neuronal loss [24] and may be an early pathological feature preceding detectable hippocampal atrophy [219, 87].

Some studies have also reported a naively counter-intuitive increase of FA in crossing fibre areas such as the corticospinal tracts for AD patients compared to controls [47, 189]. The increase of FA can be explained by the partial loss or degeneration of specific fibre populations in WM regions where multiple fibre bundles with different directions meet, which are both highly prevalent in the human brain white matter [98] and cannot be faithfully represented by the diffusion tensor model. More complex models are therefore needed to infer fibre-specific information from diffusion MRI. Constrained spherical deconvolution (CSD) was introduced to overcome that limitation by modelling the WM in each voxel as a continuous fibre orientation distribution function (fODF) [198]. To estimate these fODFs, traditional CSD requires high angular resolution DW-MRI acquired with a constant non-zero diffusion weight (b-value), also referred to as a single-shell acquisition. Fibre-specific measures derived using CSD on single-shell data have been recently used to investigate WM differences between healthy controls and patients with AD, finding degeneration along specific fibre pathways such as the splenium of the

corpus callosum, the cingulum bundle in its posterior and parahippocampal aspects, the uncinate fasciculus and the arcuate fasciculus [128].

Given the limited spatial resolution of DW-MRI, the observed diffusion signal in a voxel might originate from multiple tissue types and/or the surrounding cerebrospinal fluid (CSF). These partial volume effects (PVE) can affect any diffusion measure of microstructural integrity. For instance, when the PVE due to CSF contamination is corrected in the GM, differences in MD between controls and AD patients are attenuated [85], suggesting that previously reported diffusion abnormalities in GM areas [212] were likely due to CSF contamination caused by macroscopic atrophy rather than a change in GM microstructural properties. The PVE also affects traditional "single-shell" CSD, where the diffusion-weighted signal is modeled solely as WM content and thus spurious features and biases in quantitative parameters are produced in the presence of GM or CSF [99].

The contribution of each tissue type to the signal can be quantified by exploiting their distinct diffusion signal dependency on b-value. As CSF signal decays much faster than GM and WM signals, it is possible to distinguish between CSF signal and tissue signal using only one non-zero b-value in conjunction with the corresponding $b = 0$ data. Therefore, diffusion measures can be corrected for CSF contamination at interfaces between WM/CSF and GM/CSF in studies using single-shell data [85, 128, 52]. However, to also discriminate between WM and GM signal profiles, more than one non-zero b-value is needed. When DW-MRI is acquired with a multi-shell scheme, it is possible to separate the observed signal in a voxel into the contributions from each tissue type. Multi-shell multi-tissue constrained spherical deconvolution (MSMT-CSD) [99] exploits the different tissue signal dependencies on the b-value to improve fODF estimation by quantifying the portion of the signal attributed to each macroscopic tissue type (WM, GM, and CSF). Therefore, in addition to effectively correcting for PVE in the WM modelling, this approach provides diffusion-derived measures of tissue-like content within each voxel.

In this work, AD-related abnormalities in brain tissue were studied by performing a comprehensive analysis of tissue-specific measures derived from multi-shell diffusion MRI. Using MSMT-CSD, the obtained multi-tissue model is composed of the PVE-corrected fODF along with the total contributions of three tissue types, WM, GM, and CSF, also called tissue-like fractions. Differences between control subjects, patients with MCI, and dementia due to AD were investigated by comparing fibre integrity measures and tissue composition between groups following two parallel approaches: fixel-based analysis (FBA) [155] for the WM fODF, and voxel-based analysis (VBA) [13, 180] for the tissue fractions. The comprehensive analysis we present here constitutes a holistic neuroimaging approach for the study of the AD continuum combining multiple interrelated measures of tissue integrity derived from DW-MRI.

3.2 Study data

Participants

Patients with MCI due to AD ($n = 29$) and AD dementia (ADD) ($n = 23$), as well as cognitively healthy controls ($n = 27$) were included in the study (see Table 3.1). The

diagnosis of MCI due to AD and ADD was done according to the NIA-AA research criteria [5] while taking into account clinical data, neuropsychological examination, structural MRI, and, in some cases, CSF biomarkers [183, 181]. Controls were selected from the research database of the UAntwerp Reference Center for Biological Markers of Dementia (BIODEM) [181]. They consisted of volunteers for biomarker research ($n = 43$), having a normal neuropsychological examination and no evidence of central nervous system pathology after extensive investigation [136]. The study was approved by the local ethics committee and all subjects gave written informed consent.

Image acquisition and pre-processing

Data were acquired on a Siemens 3T MRI scanner with a (32)-channel head coil using a multi-slice, single-shot EPI, spin-echo imaging sequence. Diffusion weightings of $b = 0, 700, 1000$ and $2800 s/mm^2$ were applied in $10, 25, 40$ and 75 directions, respectively. Other imaging parameters were: voxel size of $2.5 \times 2.5 \times 2.5 mm^3$, matrix size of 96×96, and 40 axial slices. During the study, the gradient set of the MRI scanner was upgraded from 40 to 80 mT/m; following this upgrade, the sequence TR and TE were changed from $6000/116 ms$ to $5900/83 ms$, with all other parameters remaining fixed. The acquisition time was approximately 16 min. A T_1-weighted MR image was additionally acquired with a voxel size of $1 \times 1 \times 1 mm^3$.

Each DW-MRI dataset was pre-processed using a state-of-the-art pipeline. Data were first denoised using random matrix theory, thereby increasing the signal-to-noise ratio (SNR) without spatially smoothing the data [205]. Then, Gibbs-ringing artefacts were suppressed [103], head motion and eddy current-induced distortions were corrected [11, 10], and inhomogeneities of the B1 field were accounted for [202]. Finally, images were up-sampled to $1.25 \times 1.25 \times 1.25 mm^3$ to improve the accuracy of subsequent spatial normalization [153, 53]. The T_1-weighted image was used to compute the intracranial volume (ICV) with SPM12 [119].

CSF biomarkers

37 individuals underwent a lumbar puncture less than 3 months before or after image acquisition. This subset included 8 controls, 19 patients with MCI due to AD, and 10 patients with AD dementia. CSF biomarker analyses were performed with single parameter ELISA kits following standard procedures [181]. CSF levels of Amyloid-β of 42 amino acids (Aβ_{1-42}), total tau (T-tau), and phosphorylated tau at threonine 181 (P-tau$_{181}$) were considered to investigate linear relations of these biomarkers with integrity measures derived from multi-shell DW-MRI.

3.3 Methods

The methodology in this work can be divided into three parts. First, CSF and GM signal contributions as well as the full WM fODF were extracted in each voxel using MSMT-CSD. Second, the tissue decomposition maps were spatially normalized across

Group	Sample size	Sex F/M	Age in years mean (sd)	Subjects with CSF biomarkers	Scanner gradient strength 40/80 mT/m
CO	27	12/15	70.3 ± 3.8	8	19/8
MCI	29	16/13	72.0 ± 3.6	19	15/14
ADD	23	10/13	71.4 ± 4.0	10	19/4
All	**79**	38/41	71.3 ± 3.8	37	53/26

Table 3.1: Description of data per group: cognitively healthy controls (CO), patients with MCI due to AD, and AD dementia (ADD). Final column indicates the number of subjects in each group for which data were acquired before vs. after scanner hardware upgrade. sd: standard deviation.

the study subjects, achieved by calculating a study-specific template and transforming all the subjects' data to this template using a multi-channel registration method. In the third step, spatially normalised information was analysed with non-parametric hypothesis tests. A schematic overview of the pipeline is presented in Figure 3.1.

3.3.1 Multi-tissue decomposition

A multi-tissue model was obtained by applying MSMT-CSD to each DW-MR dataset. To perform MSMT-CSD, a representative signal response for each of WM, GM, and CSF was estimated using an unsupervised method based on specific tissue diffusivity properties [41, 40]. Average tissue responses were obtained across subjects (separate average responses were calculated for the scans acquired before the upgrade and for those acquired after the upgrade to facilitate consistent tissue decompositions before and after the upgrade), and then, using these averaged tissue responses, MSMT-CSD was applied to each dataset. To assure WM fODF, GM, and CSF contribution maps were comparable across subjects they were normalised with a multi-tissue approach that minimizes the average difference between 1 and the sum of the three tissue-like contributions while simultaneously performing bias field correction [151, 42]. The resulting multi-tissue decomposition consists of the WM fODF along with the GM and CSF contributions to the signal, the WM contribution map is extracted from the WM fODF as l=0 term of the spherical harmonic (SH) expansion [22]. The minimum contribution of each tissue-like component was set at 1e−8.

3.3.2 Spatial normalisation

3.3.2.1 Population template

A study-specific template was built from a set of 24 cases including 12 controls, 6 patients with MCI and 6 patients with AD. These sub-groups were age-matched and balanced
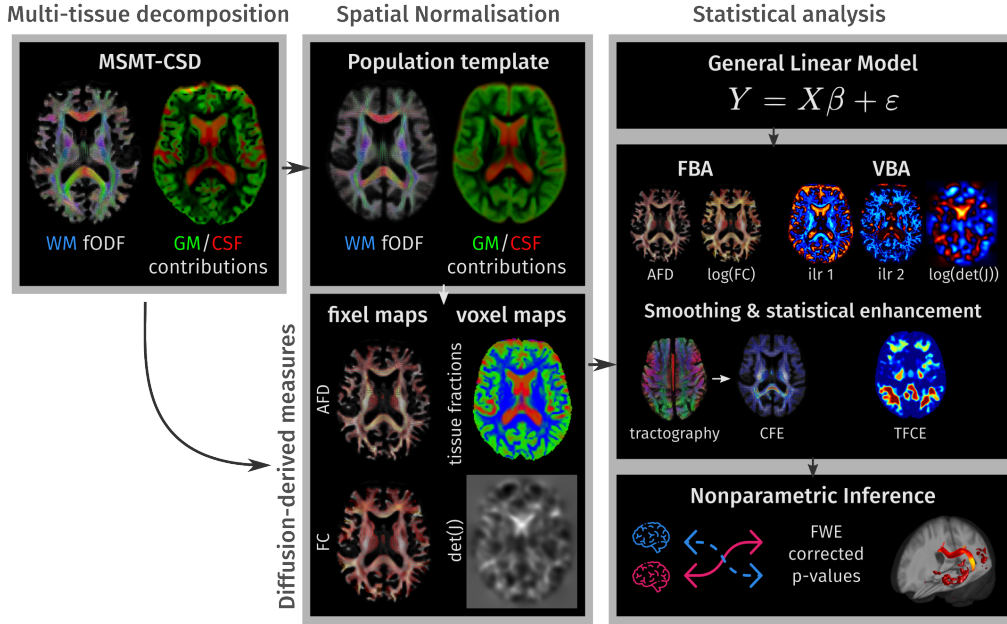
Figure 3.1: Methodology overview. In the first step, a multi-tissue decomposition is obtained from the multi-shell diffusion data. In the second part, the information is spatially normalised to a population template calculated for the study population, and each subject is represented by 6 feature maps: 2 fixel maps containing the apparent fibre density (AFD) and fibre cross-section (FC), and 4 voxel maps containing the tissue-like signal fractions for WM, GM and CSF, and 1 map with the local volumetric changes induced by the spatial deformations. Finally, differences of these measures between groups of subjects are investigated following non-parametric statistical frameworks for fixel- and voxel-based analysis.

by gender. This multi-tissue population template was constructed with an iterative atlas building framework [152] that used a multi-channel nonlinear diffeomorphic registration algorithm [146] to align the fODFs as well as the GM and CSF-like contribution maps (See Figure 3.2). The same registration algorithm was applied to align the multi-tissue decompositions for all participant scans to the population template.

3.3.2.2 Diffusion-derived measures

The fODF is a continuous function represented in the SH basis, which can represent multiple fibre populations crossing within a single voxel. To facilitate quantification and statistical analysis, these are segmented to estimate within each voxel a finite number of discrete fibre orientations [179]. The term fixel is used to refer to a specific population of fibres oriented in a specific direction within a specific voxel [154]. The integral of the fODF ascribed to each fixel is proportional to the volume of fibres aligned in the corresponding direction; this measure is known as *Apparent Fibre Density* (AFD) [153] and it has been demonstrated to effectively quantify specific fibre integrity in crossing fibre regions [161]. The AFD values extracted from the fODF can be mapped to their
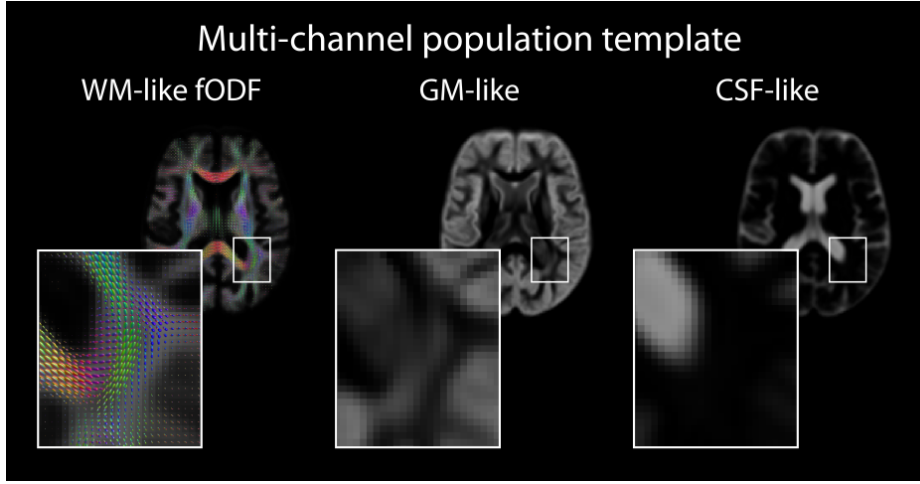
Figure 3.2: The study-specific population template is composed of a white matter fibre orientation distribution function (WM fODF) template along with the voxel templates containing the tissue-like contributions for grey matter (GM) and cerebrospinal fluid (CSF).

respective fixels as shown in Figure 3.3.

During the spatial normalisation process, the multi-tissue model is warped to match the population template. When applied to voxel maps, the warping causes expansion or contraction of regions in the spatially normalised image. In a particular voxel, this volumetric change (with respect to the population template) is captured by the determinant of the Jacobian matrix J. This concept has been extended to the fixel-based analysis framework by accounting for the effect of the Jacobian transformation along different fibre directions [153, 155]. Given the unitary vector corresponding to a fixel f, the change in scale along this direction is $\|Jf\|$, and the total volumetric change is the product between $\|Jf\|$ and the change in the area perpendicular to f; the latter of these is a measure of the variation in fibre bundle cross-section (FC) [153, 155] and is calculated as:

$$FC(f) = \frac{det(J)}{\|Jf\|} \tag{3.1}$$

The determinant of the Jacobian as well as the fibre bundle cross-section measure are not absolute measures of volume or area but rather measures of changes relative to the population template. When one of these measures is smaller than one, the corresponding features are smaller in the subject space than in the template space, and vice-versa.

In some neuroimaging analysis pipelines, a modulation step is carried out to combine the model-derived normalised measures with the macroscopic changes induced by the spatial normalisation to capture both mesoscopic and macroscopic changes. However, it has been shown that the use of these modulated measures for hypothesis testing can lead to decreased sensitivity, probably due to the introduction of multiplicative noise [150]. For this reason, in our analysis we treat model-derived measures and morphological measures separately. To ensure model-derived measures (tissue-like contributions and AFD) represent true fractions of the signal, they were divided by the sum of the three
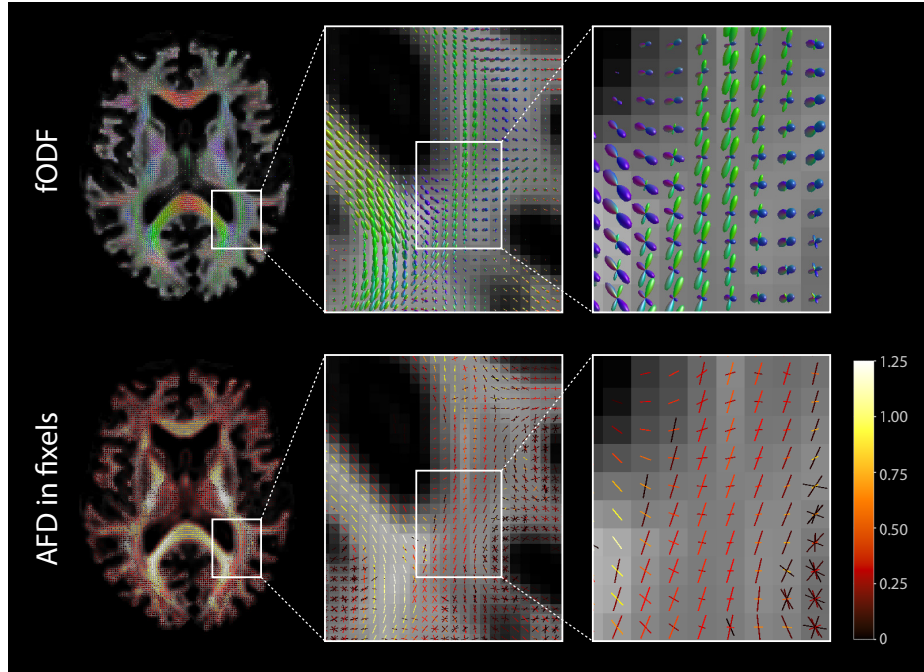
Figure 3.3: The fibre orientation distribution functions (fODF) can encode multiple fibre populations within a single voxel; each of these fibre populations is described with directional elements called *"fixels"*. Each fixel is here coloured according to the value of Apparent Fibre Density (AFD).

tissue-like contributions at each voxel. In template space, each subject is described by two fixel maps and four voxel maps. Fixel maps contain the two fibre specific measures: AFD and the fibre cross-sectional (FC) area. The set of voxel maps consist of the three tissue-like fraction maps accompanied by the determinant of the Jacobian matrix.

3.3.3 Statistical Analysis

Hypothesis testing to detect differences of measures between controls, MCI and ADD patients was done using the General Linear Model (GLM) framework including age, gender, intracranial volume (ICV), and scanner version as covariates. Non-parametric permutation tests were performed to calculate family-wise error (FWE) corrected p-values for each hypothesis by computing an empirical null distribution for the enhanced statistic [214, 6]. In this study, permutation testing was conducted with 5000 permutations and the significance level was set at $\alpha = 0.05$.

Fixel measures were compared using FBA [155, 178], while voxel-wise measures were studied using VBA [13, 180]. Both approaches are closely related and include many equivalent steps: data smoothing, statistical enhancement and calculation of p-values with correction for multiple comparisons.

3.3.3.1 Fixel-based analysis of fibre-specific measures

Smoothing and statistical enhancement of fixel-wise quantitative parameters was based on a fixel-fixel connectivity matrix, encoding fractional connectivity between fixels based on streamlines tractography. A whole-brain tractogram of 10 million streamlines was generated from the population fODF template using the iFOD2 algorithm [201]; from this a subset of 2 million streamlines was extracted using the Spherical-deconvolution Informed Filtering of Tractograms (SIFT) method [179] to reduce density biases in the reconstruction. Elements of the fixel-fixel connectivity matrix are calculated as the fraction of streamlines intersecting one fixel that also intersect another fixel [154]. These data were used both for smoothing of fixel-wise measures in conjunction with an isotropic Gaussian kernel with FWHM = 10 mm, and for performing statistical enhancement via Connectivity-based Fixel Enhancement (CFE), for which the default parameters were used ($E = 2$; $H = 3$; $C = 0.5$).

An omnibus F-test was performed first to detect any effect across the three groups and the two fixel measures: AFD and FC (first log-transformed for normality). Pairwise differences for the two measures were interrogated by simultaneously testing multiple contrasts in a GLM (3 pairs of groups × 2 measures × 2 effect directions) within the set of fixels that showed significant group effects according to the omnibus F-test. Strong FWE-corrected p-values were computed by generating a single null distribution for the 12 contrasts [6].

3.3.3.2 Voxel-based analysis of voxel-wise measures

Statistical analysis of the three tissue-like fractions $\{T_{csf}, T_{gm}, T_{wm}\}$ should take into account the compositional nature of this data: $0 < T_i < 1$ and $T_{csf} + T_{gm} + T_{wm} = 1$. The latter implies the three tissue-like fractions are not linearly independent with only two degrees of freedom, therefore projecting them to a 2-dimensional space is more appropriate for the statistical analysis than treating the three measures independently. At each voxel, the tissue-like fraction values were mapped into a 2-dimensional space using the isometric log-ratio (ilr) transformation [57], an approach that was recently adopted to study the tissue composition of lesions in AD using DWI [127]. The two independent isometric log-ratios were calculated as follows:

$$
\begin{aligned}
ilr_1 &= \frac{1}{\sqrt{6}} \ln\left[\frac{T_{csf} \times T_{gm}}{T_{wm}^2}\right] \\
ilr_2 &= \frac{1}{\sqrt{2}} \ln\left[\frac{T_{csf}}{T_{gm}}\right]
\end{aligned}
\tag{3.2}
$$

The isometric log-ratios can capture changes in the relation between the three tissue-like fractions, an example of resulting ilr is shown in Figure 3.4. Increasing ilr_1 values could reflect: decreased WM-like fraction, accompanied by increased GM or CSF-like fractions, or unchanged WM-like fraction with increased product of GM and CSF-like fractions. Increases in ilr_2 reflect an increase in the CSF-like fraction relative to the GM-like fraction.

Before statistical analysis, voxel maps containing the ilr were smoothed using a 3D Gaus-
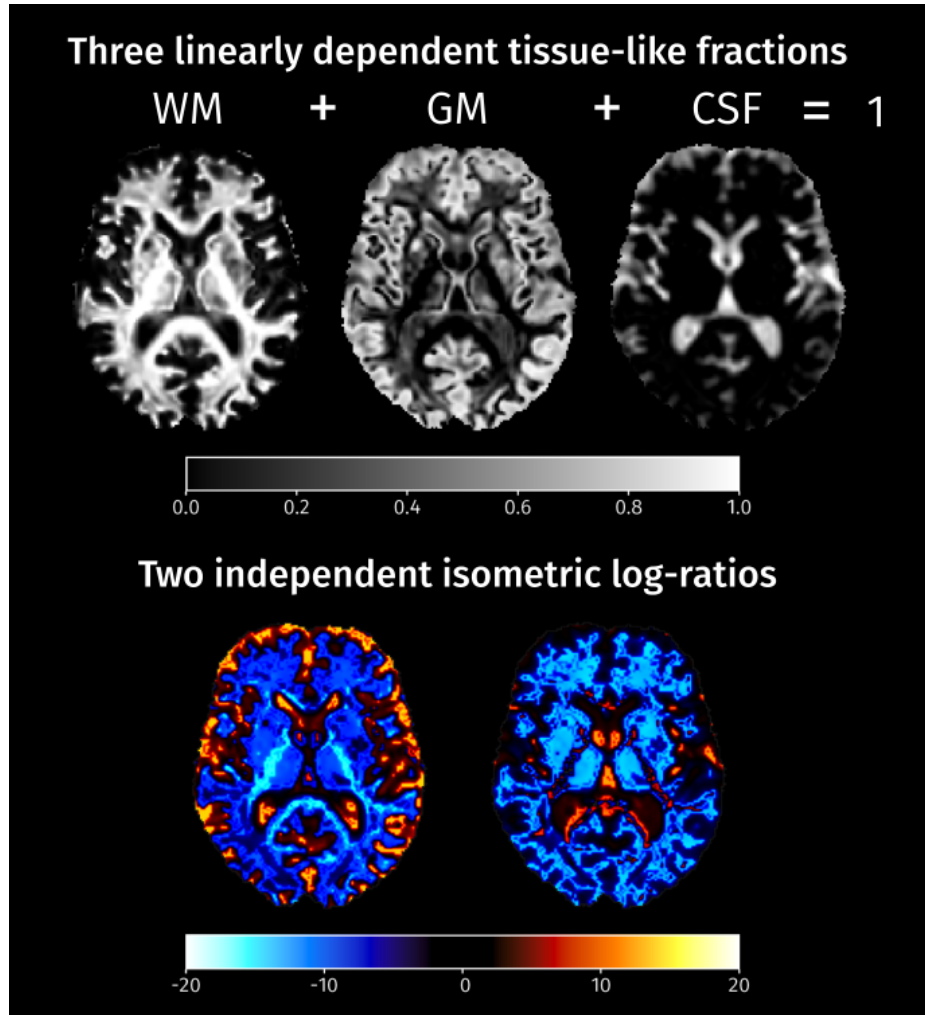
Figure 3.4: The three tissue-like fractions are bounded and linearly dependent with only two degrees of freedom, these fractions were transformed into two independent isometric log-ratios following Equation 3.2.

sian kernel with FWHM = 5 mm (voxel maps are smoothed with a narrower Gaussian filter than the spatial kernel used in FBA due to the latter being additionally constrained by fixel-fixel connectivity).

Voxel-based analysis also included an initial omnibus F-test to detect any effect across groups and voxel-wise measures. *Post hoc* testing for pairwise differences of isometric log-ratios and Jacobian determinant (log-transformed for normality) were performed simultaneously while applying strong correction of p-values over the 18 contrasts (3 pairs of groups $\times$ 3 measures $\times$ 2 effect directions). Statistical enhancement was done using the Threshold-free cluster enhancement (TFCE) method applied with default parameters (E = 0.5 and H = 2) [180].

3.3.3.3 Linear correlations with CSF biomarkers

Possible relationships between CSF biomarkers and tissue degeneration were explored by testing the linear correlations of levels of CSF A-β_{1-42}, total tau, and P-tau$_{181}$ with the extracted fixel- and voxel-wise measures. The CSF levels of each biomarker were considered as a continuous regressor in a GLM that also included age, gender, intracranial volume (ICV) and scanner as covariates. In these analyses, the relation of each biomarker with diffusion-derived measures was tested with two omnibus *F*-tests, one for the two fixel measures, and another one for the three voxel-wise measures. If any significant effects were detected, *post hoc* testing was performed while applying strong FWE correction across contrasts. Effects were considered significant when the FWE corrected *p*-values associated with the alternative hypotheses are below the significance level ($\alpha = 0.05$).

Implementation

All steps in the analyses were performed using *MRtrix3* (version 3.0.2) [200] (`https://www.mrtrix.org/`). During DW-MRI preprocessing, *MRtrix3* scripts invoke the **"eddy"** tool from *FSL* [11] and **"N4BiasFieldCorrection"** from *ANTs* [202].

3.4 Results

Specific tracts where fibre integrity measures decrease with the disease emerged from the FBA approach. From the VBA pipeline, widespread areas in the brain showed differences in tissue-like content and macroscopic volume changes.

3.4.1 Fixel-based analysis

The integrity of WM fibres was evaluated using two fixel-wise measures: apparent fibre density (AFD) and fibre bundle cross-section (FC). The initial omnibus *F*-test identified an extensive set of fixels where the fibre measures differ across disease stages (FWE-corrected $p < 0.05$), Figure 3.5 shows the streamline segments corresponding to those fixels where significant effects were detected. Effects are present in the splenium and tapetum of the corpus callosum (CC), the inferior longitudinal fasciculus (ILF), the uncinate fasciculus, the thalamo-occipital projection, the cortico-spinal tract (left), the cingulum bundle (right), the parahippocampal part of the cingulum bundle (left), and the left arcuate fasciculus.

From *post hoc* pairwise comparisons of the two fixel measures, areas with significant differences were detected for 4 out of 12 tests when setting the significance level at 0.05 after strong FWE correction. Patients with MCI and ADD show less AFD than controls in the splenium and tapetum of the CC (See Figure 3.6), while FC decreases in both groups of patients (compared to controls) are present in other white matter tracts, including the splenium, such as the left corticospinal tract, left uncinate fasciculus, and right ILF. For ADD patients, decreased FC is also detected in the right cingulum, left arcuate fasciculus, left parahippocampal gyrus, and left thalamo-occipital projections (see Figure 3.7). Many
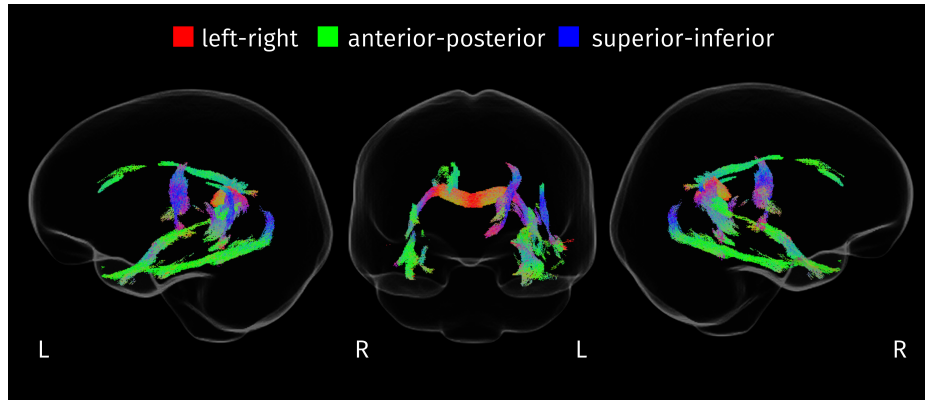
Figure 3.5: Streamline segments in the population template tractogram corresponding to fixels where the disease stage has a significant effect on any of the two fixel-specific measures (FWE-corrected $p < 0.05$). Streamlines are coloured according to their orientation.

of these fibre tracts overlap with the areas resulting from the F-test shown in Figure 3.5, explaining most of the significant effects in the omnibus test. Pairwise differences of fixel measures between ADD and MCI did not meet the level of statistical significance.

3.4.2 Voxel-based analysis

Widespread significant group effects across the three voxel-wise measures (2 *ilr* parameters and the Jacobian determinant) were identified by the omnibus F-test across much of the template analysis mask. As shown in Figure 3.8, voxel-wise measures differ across disease stages in 44.1% of the analysed brain area.

After applying strong FWE correction across the 18 *post hoc* pairwise comparisons, voxels with significant effects were detected for 9 of the tested contrasts. Increases of the first isometric log-ratio (ilr_1 in Equation 3.2) were detected in MCI and ADD patients compared with controls, and in ADD compared with MCI. The increment of this ratio indicates: the reduction of WM-like fraction accompanied by increased CSF or GM-like fractions, or the increased product of CSF and GM-like fractions while WM-like remains constant. Significant increases of the second isometric log-ratio (ilr_2 in Equation 3.2) were revealed for both MCI and ADD patients when compared to control subjects. Increased ilr_2 could be the result of: decreased GM-like fraction with increased or constant CSF-like fraction, or increased GM-like fraction with also increased CSF-like and therefore decreased WM-like fraction. Figure 3.9 shows the absolute variation of tissue-like fractions in areas where significant increases in ilr_1 or ilr_2 were detected. When MCI subjects are compared against controls, significant changes in tissue-like composition are detected in the intersection between the insular cortex and planum polare, in the cingulate cortex, the amygdala, the hippocampus, the caudate, and some WM areas such as the cingulum and inside the temporal lobe surrounding the ILF (see Figure 3.9 top row). In the case of ADD patients when compared to controls, all of these observations are recapitulated and expanded, along with changes in the temporal cortex, the temporal pole, the CC, the superior and inferior longitudinal fasciculi, the
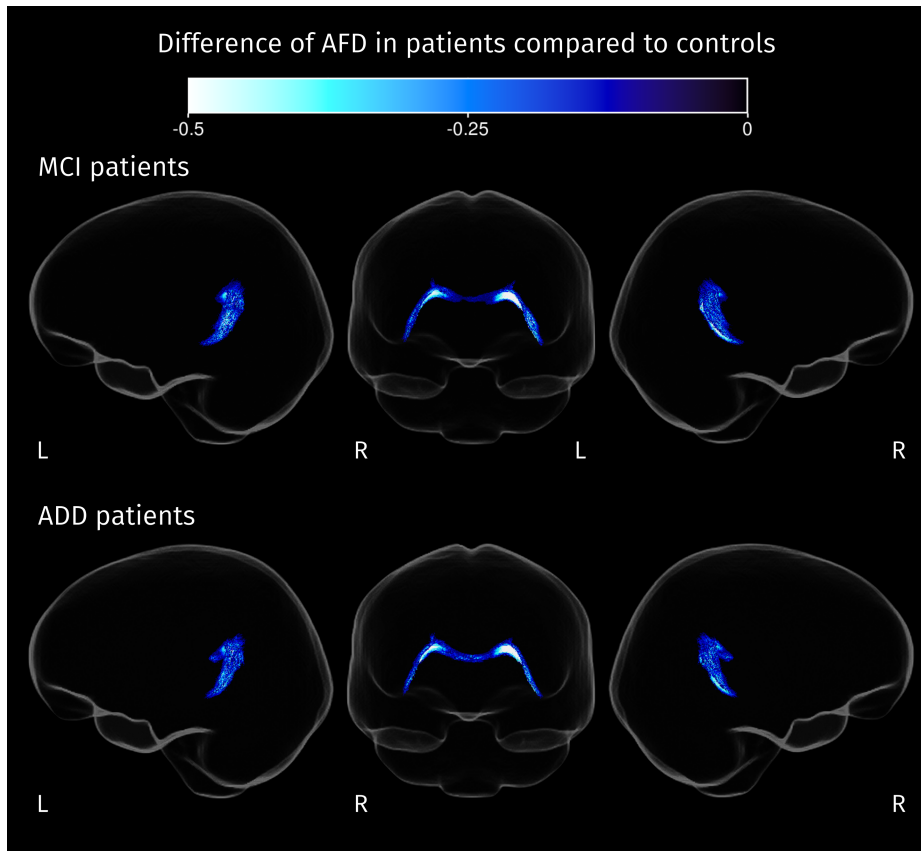
Figure 3.6: Section of the corpus callosum where AFD is significantly reduced in both groups of patients compared to control subjects (strong FWE-corrected $p < 0.05$). Colour corresponds to the value of the difference between the mean AFD for patients and the mean AFD for controls. No statistically significant differences of AFD between ADD and MCI patients were detected.

parahippocampal gyrus of the cingulum, and the thalamic radiations (see Figure 3.9 middle row). Significant differences of tissue-like composition between ADD and MCI patients were also detected in the inferior temporo-occipital region of the right temporal lobe and in the right cingulum (see Figure 3.9 bottom row).

Significant differences of local volume were detected in both MCI and ADD patients groups when compared with controls. In Figure 3.10 are shown the resulting pairwise differences in the determinant of the Jacobian matrix (which accounts for volumetric changes induced by spatial normalisation). For both groups of patients, ventricles are significantly larger and there is a significant shrinkage of the anterior part of the left temporal lobe. For ADD patients, the significant reduction of the local volume in the temporal lobe was detected in both hemispheres, and it reached the angular gyrus where the posterior parts of the middle and inferior longitudinal fasciculus are located.
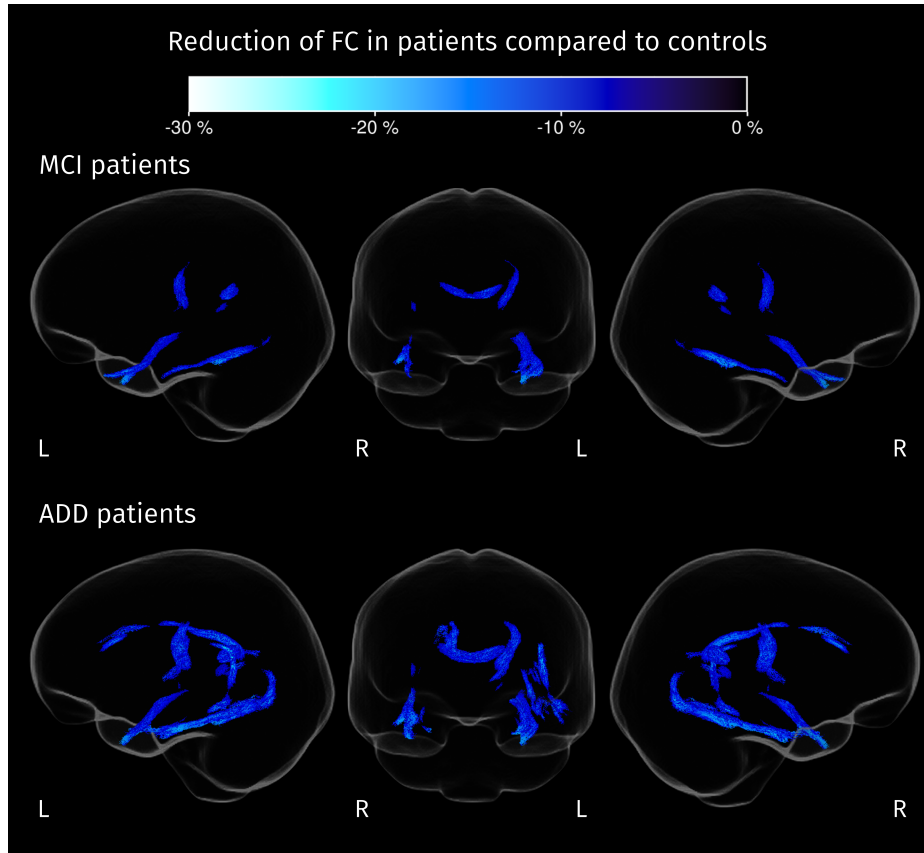
Figure 3.7: Streamline segments corresponding to fixels where FC is significantly reduced in patients compared to controls (strong FWE-corrected $p < 0.05$). Colour corresponds to the percentage of change in each group of patients compared with the control group. Given that analyses were performed with the $\log(FC)$, this value was calculated as $\exp(\beta_{AD}) - 1$ with β_{AD} representing the difference between means of $\log(FC)$ for controls and AD dementia patients. No significant differences of FC between ADD and MCI patients were detected.

3.4.3 Correlation between CSF biomarkers and diffusion-derived measures

Significant effects of $A\beta_{1-42}$ in fixel- and voxel-wise measures were detected with the omnibus F-tests. From post hoc testing negative correlations between $A\beta_{1-42}$ level and the fixel-wise / voxel-wise measures of macroscopic area/volume change relative to the template were found in the interface between the ventricles and the genu of the CC (shown in Figure 3.11). The negative regression coefficient in this case means that the lower the CSF $A\beta_{1-42}$ levels, and thus the more pathological, the greater the volumetric change caused by the registration to the population template.

Interestingly, but not surprisingly, a significant linear correlation between $A\beta_{1-42}$ level and ilr_2 was found in the left hippocampus, and also in the anterior part of the cingulate

Figure 3.8: Brain areas where the disease stage has a significant effect on any of the voxel-wise measures (FWE-corrected $p < 0.05$).

cortex, in the left dorsal anterior insula, and in the genu of the CC (see Figure 3.12). From the F-tests for the other two CSF biomarkers, total tau and P-tau$_{181}$, no significant correlations were detected between them and diffusion-derived measures.

3.5 Discussion

In this work, we presented a comprehensive analysis of AD effects in brain tissue by comparing tissue decompositions from multi-shell DWI between groups of subject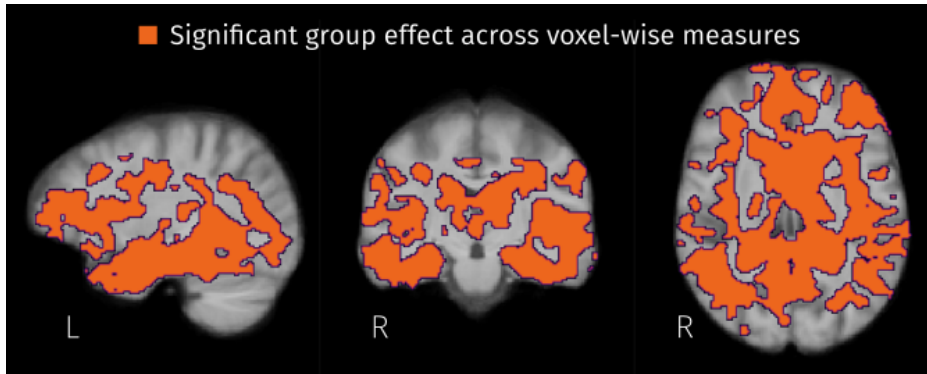s belonging to the AD continuum and cognitively healthy controls. To the best of our knowledge, this is the first study that combines multi-shell multi-tissue CSD with both fixel and voxel-based analysis approaches to detect changes of tissue diffusivity properties related to AD progression. With MSMT-CSD, we can obtain a multi-tissue model that estimates the contribution of each tissue type to the diffusion signal while modelling the WM fibre configuration taking into account the PVE.

Differences of fibre-specific measures

To study WM integrity, fibre specific measures, namely apparent fibre density (AFD) and fibre bundle cross-section (FC), were investigated following the FBA approach [155]. The integration of the MSMT-CSD within the FBA pipeline allows improving the estimation of the mentioned fibre measures because it gives a more precise fODF in voxels where WM/GM and WM/CSF signals are mixed [99]. After applying the strong FWE correction to *post hoc* pairwise one-sided comparisons in two directions, decreased AFD was detected in the splenium and tapetum of the CC for both MCI and ADD patients compared to controls (Figure 3.6), while macro-structural decreases of fibre bundle cross-section were found in several WM tracts (Figure 3.7). A previous investigation applied the FBA framework to study the WM integrity in AD and MCI patients using single-shell diffusion data [128] finding specific fibre tracts with significant decreases of the WM integrity measures in AD patients when compared to healthy controls. There are common findings such as reduced AFD in the splenium and reduced FC the right cingulum,
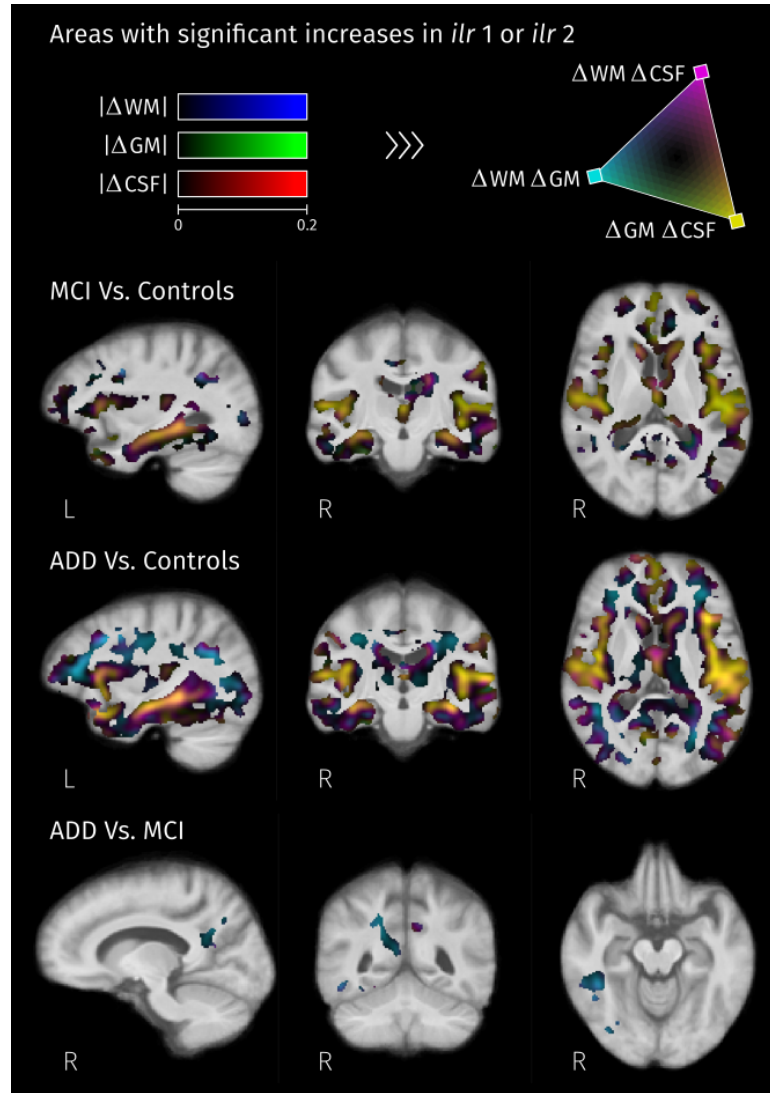
Figure 3.9: Brain areas where at least one of isometric log-ratios is significantly greater (strong FWE-corrected $p < 0.05$) in patients than in controls, and in ADD patients compared to MCI. The absolute value of the difference between mean tissue-like fraction between groups is represented in a different colour channel for each tissue type: red for CSF, green for GM, and blue for WM. The resulting colour corresponds to the combination of change for multiple tissues.

uncinate fasciculus, and ILF. Mito et al., 2018 also reported a larger set of fibres tracts with decreased AFD in ADD patients, those tracts include the parahippocampal cingulum, the inferior fronto-occipital fasciculus and the left fornix.

Some of the tracts with macrostructural differences, which manifested as a significant reduction of fibre bundle cross-section, correspond with tracts that have previously been reported to show differences in diffusion-based measures such as FA and MD. Our results
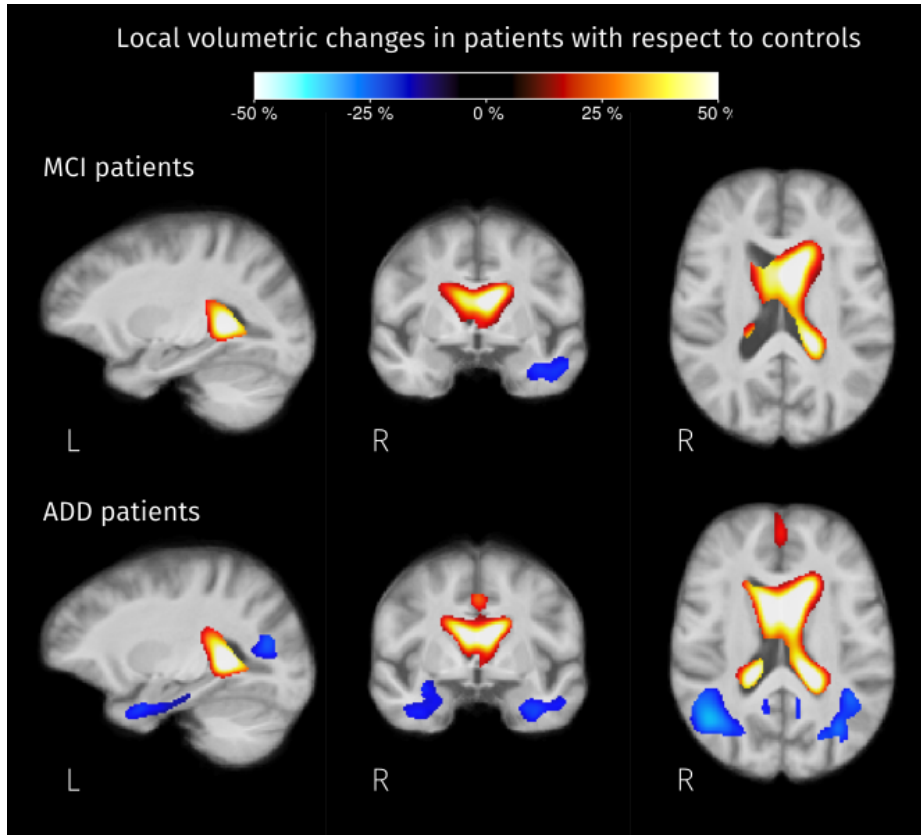
Figure 3.10: Brain areas where the Jacobian determinant is significantly different (strong FWE-corrected $p < 0.05$) in patients compared to controls. The colourmap represents the percentage of change in local volume compared to the control group. Analogous to FC, analyses for this value were performed in the log domain therefore this percentage of change was calculated as $\exp(\beta_{PT}) - 1$ where β_{PT} represents the difference between the mean value of the logarithm of the Jacobian determinant for MCI/AD patients and the mean of those values for controls.

suggest some of those changes previously attributed to microstructural properties could in fact be macroscopic effects captured by DTI metrics due to PVE [206]. The study of voxel-based metrics derived from the diffusion tensor and diffusion kurtosis models has also reported reduced WM integrity measures in the cingulum, the uncinate fasciculus, the arcuate fasciculus, and the ILF [47, 2, 183, 109]. In the corticospinal pathway, previous works have reported reduced mean kurtosis, increased free-water index, and increased FA in regions where the corticospinal tracts cross with other ones [183, 52, 47, 189], which can be observed when WM degeneration occurs in a subset of crossing fibre populations. This hypothesis is consistent with the results herein presented, which show degeneration specifically along the fibre bundles in the corticospinal tract in AD patients.
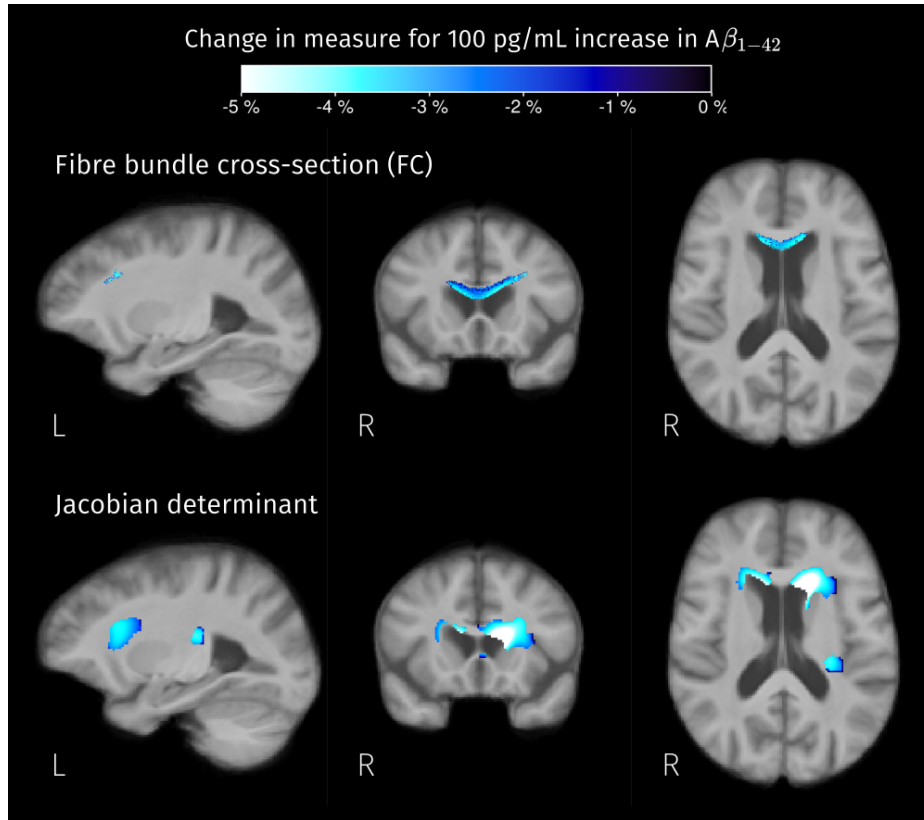
Figure 3.11: Brain areas where there is a significant linear relation between CSF biomarker for $A\beta_{1-42}$ and measures of macroscopic change relative to the population template. Colour corresponds to the percentage of change in these measures for 100 pg/mL increase in biomarker value. Given that analyses for fibre cross-section and Jacobian determinant were performed in the log scale, the colour-coded effects in significant areas were calculated as $\exp(\beta \times 100) - 1$ where β is the GLM coefficient of $A\beta_{1-42}$ for the corresponding measure.

Differences of tissue-like composition

Detected changes in tissue-like composition (Figure 3.9) in the ILF, cingulum, thalamic radiations and superior temporoparietal areas correspond to reduction of WM-like fraction (See Figure 3.13), these changes are concordant with previously reported decreased FA and increased MD in such areas [122, 2, 47] suggesting a widespread degeneration of diffusion barriers in WM. It is worth mentioning that WM-like reduction in temporal and parietal structures coincides with a significant reduction of the local volume (See Figure 3.10), meaning that the WM degeneration is also detectable at the macroscopic level and might be more advanced than the degeneration observed in the frontal areas where no volumetric differences were detected.

When looking at the GM-like fraction variation between groups, shown in Figure 3.14, the decreases detected in the cortical and subcortical areas are consistent with the widely
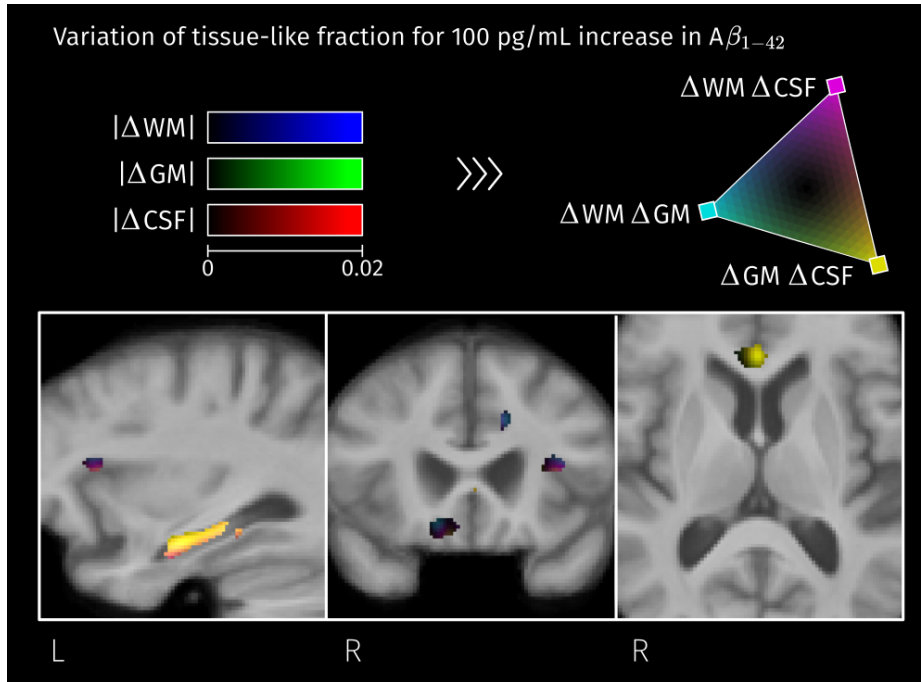
Figure 3.12: Significant linear relation between CSF biomarker for $A\beta_{1-42}$ and the second isometric log-ratio was found in the left hippocampus. Colourmap represents the estimated GLM coefficient of $A\beta_{1-42}$ (multiplied by 100) for the GM-like and CSF-like fractions.

reported landmarks of the disease: hippocampal atrophy and cortical atrophy in the temporal lobe. Therefore, these differences most likely correspond to actual GM degeneration. Some of the observed changes in tissue-like composition, manifested as increases in ilr, are the result of increased GM-like fraction (Figure 3.14), from the calculation of ilr we know that there must be an increase of the CSF-like fraction too, and therefore a reduction of the WM-like fraction; as this effect is mostly observed in WM areas such as the longitudinal fasciculi (cyan areas in Figure 3.9), this is a change consistent with degeneration of diffusion barriers in WM. Increased CSF-like fraction accompanying these changes (See Figure 3.15) is compatible with recent research reporting differences of the free-water index between AD and MCI compared to controls [52]. Although their analyses use a different diffusion model, the CSF-like signal fraction obtained in this work is conceptually close to the free-water index, both being related to the part of the signal produced by isotropic unrestricted diffusion.

Correlations with CSF biomarkers

We included an exploration of the relation between CSF biomarkers and the different diffusion-derived measures, resulting in significant effects of $A\beta_{1-42}$ levels in macroscopic measures and tissue composition in certain areas. Negative correlations between measures of volumetric change (with respect to the template) and CSF $A\beta_{1-42}$ values
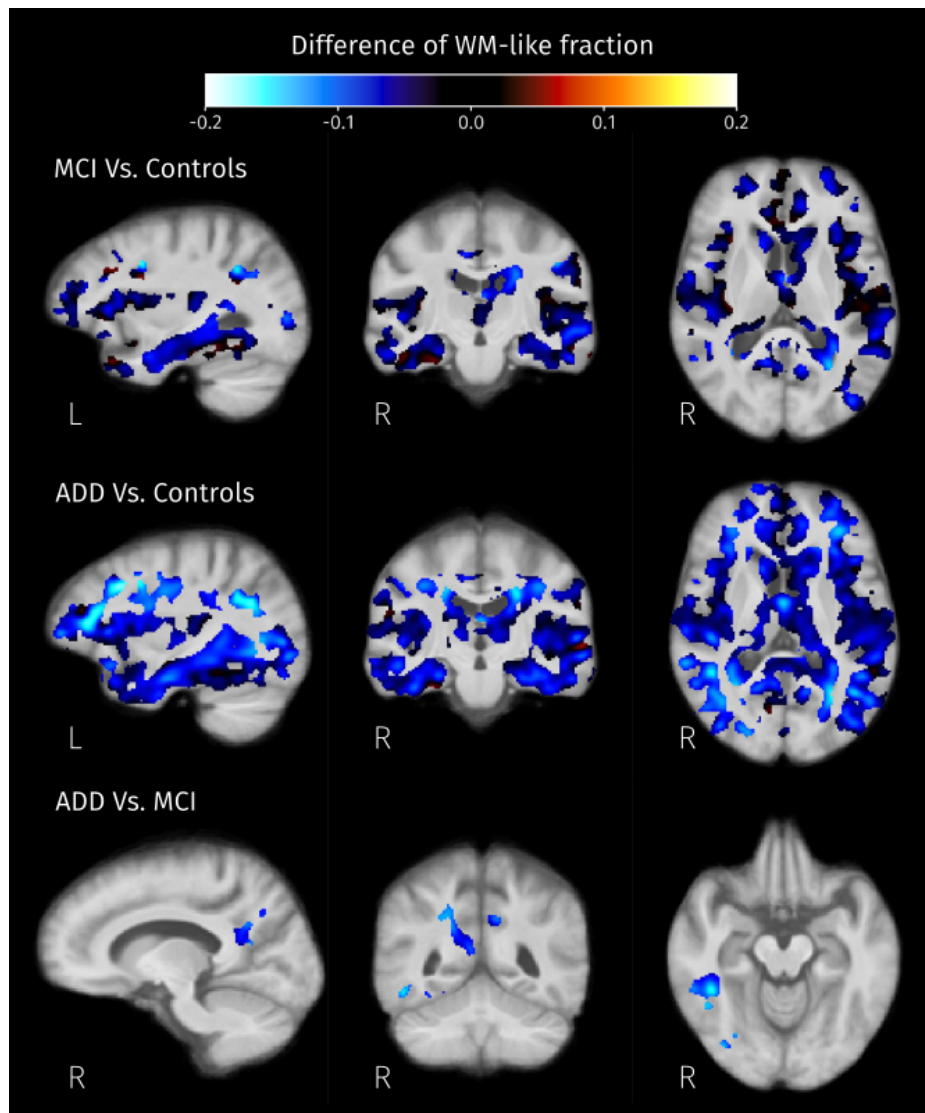
Figure 3.13: Difference of mean WM-like fraction between groups in brain areas where at least one of isometric log-ratios is significantly different between pairs of groups (strong FWE-corrected $p < 0.05$).

were found in the interface between the ventricles and the genu of the CC (Figure 3.11), indicating expansion of the ventricles in the presence of decreased (i.e. pathological) values of this biomarker. All effects of CSF $A\beta_{1-42}$ levels in fibre-specific measures detected with the omnibus F-test were confirmed as macrostructural effects driven by the volumetric differences with respect to the population template, and no significant correlations with AFD were detected at the significance threshold level after applying the strong correction for multiple comparisons during *post hoc* testing; this is an important consideration because some of the previous findings of significant correlations between CSF biomarkers and DTI measures of WM integrity have been reported without proper
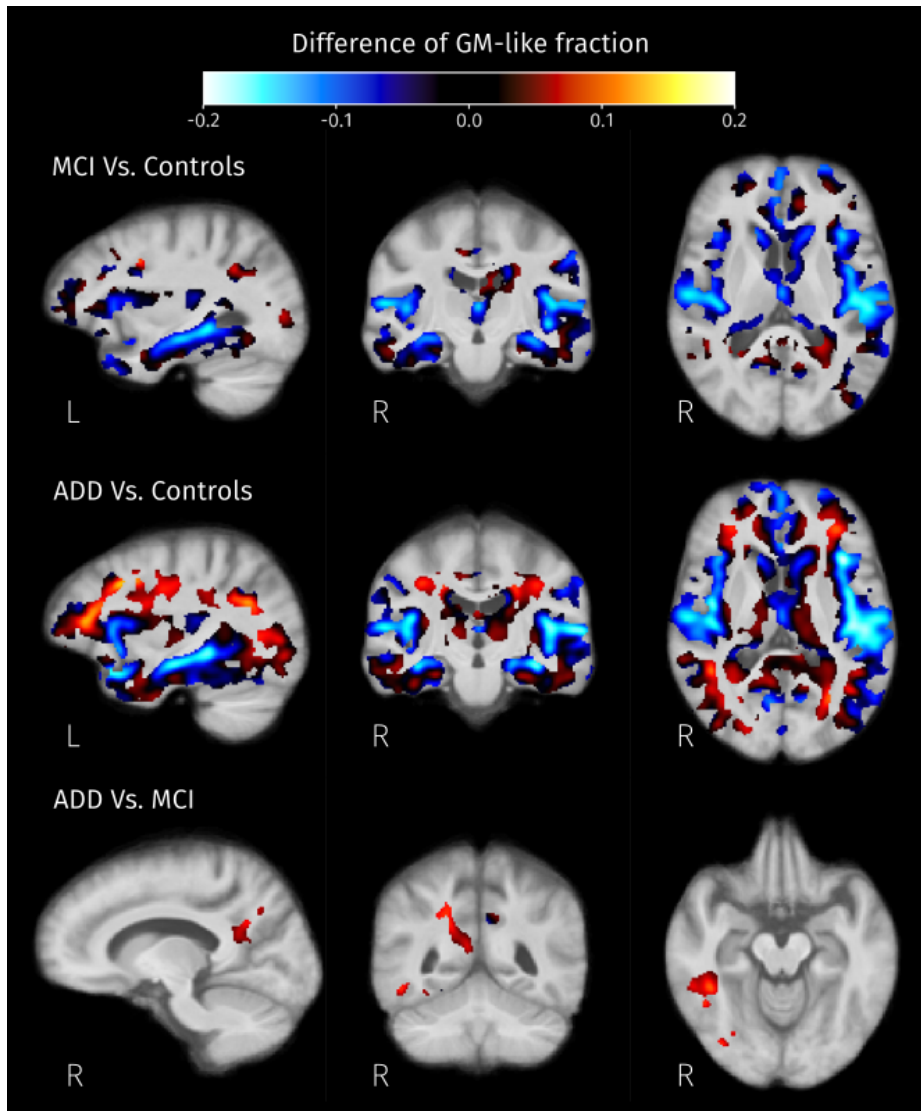
Figure 3.14: Difference of mean GM-like fraction between groups in brain areas where at least one of isometric log-ratios is significantly different between pairs of groups (strong FWE-corrected $p < 0.05$).

adjustments for multiple comparisons [7].

The significant correlations between CSF $A\beta_{1-42}$ levels and tissue-like composition come from different combinations of tissue-like content variation depending on their location (seen as different colors in Figure 3.12). As CSF $A\beta_{1-42}$ levels are more pathological, GM-like fraction decreases in the left hippocampus and genu of the CC, WM-like fraction decreases in the cingulate cortex and insula, and CSF-like fraction increases in all these areas.
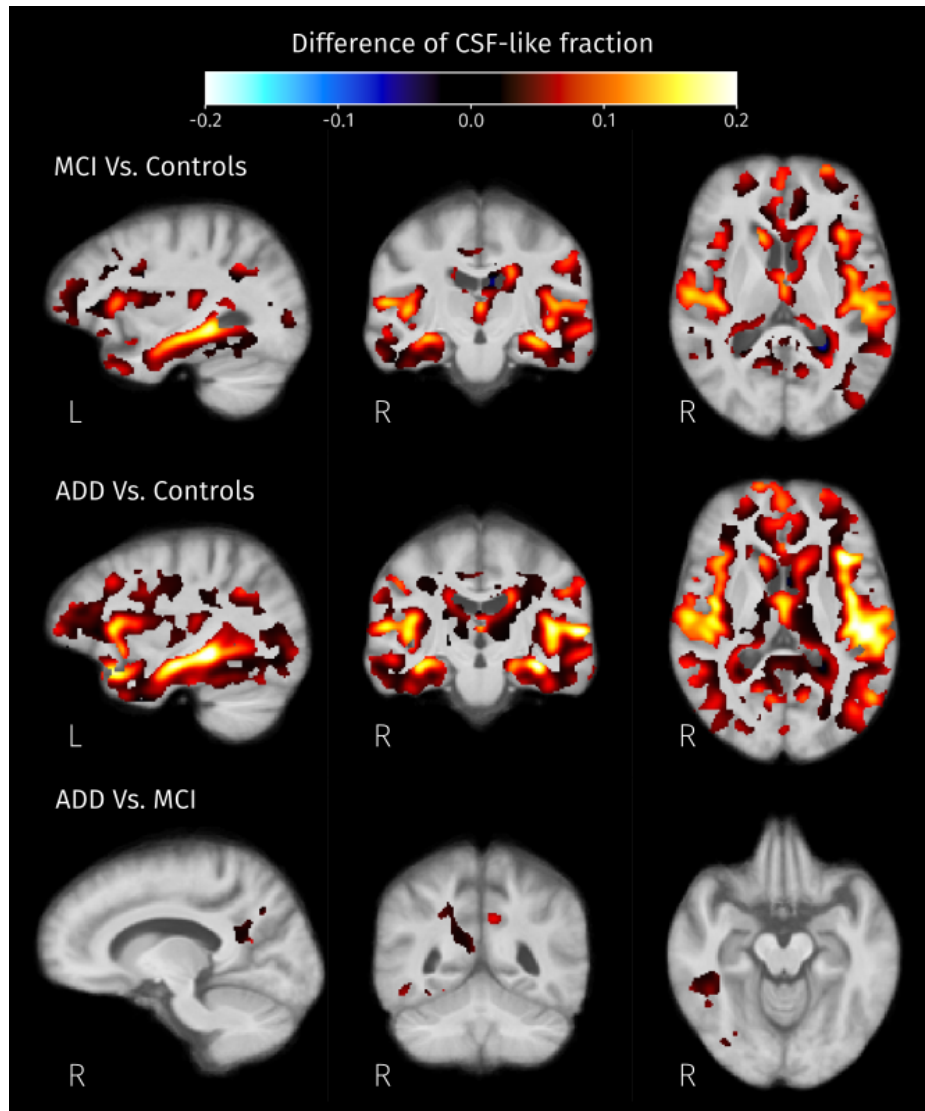
Figure 3.15: Difference of mean CSF-like fraction between groups in brain areas where at least one of isometric log-ratios is significantly different between pairs of groups (strong FWE-corrected $p < 0.05$).

Limitations

One limitation of this work is the restricted field of view of the acquired DW-MR images. As a result, the analysed area did not include the superior slices of the brain, where GM changes might be ubiquitous. Geometric distortions due to field inhomogeneity were not corrected, as no explicit image data tailored for this purpose were acquired; therefore high variability in high susceptibility areas could limit the power to detect significant differences.

Conclusions

This study demonstrates that there are widespread significant differences between the brains of patients with AD dementia and MCI due to AD, and those of age-matched healthy controls. The comprehensive analysis framework presented here facilitates simultaneous macro- and microscopic assessment of CSF, GM, and WM, all from a single DW-MRI data set. This study includes a strong control for false positives supporting the robustness of reported findings. Abnormalities related to AD symptomatic stages were detected in specific WM fibre pathways, cortical and subcortical GM, as well as macroscopic patterns such as temporal lobe atrophy and ventricle expansion.

3.6 Products

Conference abstract and oral presentation

- **Diana L. Giraldo**, Hanne Struyfs, David A. Raffelt, Paul M. Parizel, Sebastiaan Engelborghs, Eduardo Romero, Jan Sijbers, Ben Jeurissen. *Fixel-Based Analysis of Alzheimer's Disease Using Multi-Tissue Constrained Spherical Deconvolution of Multi-Shell Diffusion MRI*. International Society of Magnetic Resonance in Medicine. Honolulu, USA. 2017.

Indirect product: Journal paper

- Helena Verhelst, **Diana Giraldo**, Catharine Vander Linden, Guy Vingerhoets, Ben Jeurissen, Karen Caeyenberghs *Cognitive Training in Young Patients With Traumatic Brain Injury: A Fixel-Based Analysis*. Neurorehabilitation and Neural Repair 33. 2019; https://doi.org/10.1177/1545968319868720

Manuscript to be submitted

- **Diana L. Giraldo**, Robert Smith, Hanne Struyfs, Ellis Niemantsverdriet, Ellen De Roeck, Maria Bjerke, Sebastiaan Engelborghs, Eduardo Romero, Jan Sijbers, Ben Jeurissen. *Investigating tissue-specific abnormalities in Alzheimer's disease with multi-shell diffusion MRI*.

Chapter 4

Improving the quantitative characterization of cognitive impairment

4.1 Introduction

Currently, there is no cure for AD and the vast majority of clinical trials for disease-modifying drugs, designed to slow down AD progression from MCI to dementia, have so far failed [125]. Aside the questioned efficacy of the tested treatments, other possible reasons for failures may come up from two sources. First, heterogeneity of recruited participants, including advanced AD and variable MCI manifestations, or participants without any underlying pathology [125]. Second, standard cognitive outcomes, set as endpoints, might be highly variable and not sensitive enough to detect subtle cognitive performance changes [125, 166]. This is the case of the widely used Alzheimer's disease Assessment Scale - Cognitive (ADAS-Cog), that has shown high variability and poor sensitivity, likely by measurement errors, patient heterogeneity, and ceiling effects of its sub-scores, making some sub-scores uninformative in patients at early stages [166, 156, 80].

Composite outcomes computed with informative sub-scores from one or multiple tests have demonstrated to be more robust and sensible measures to detect cognitive and functional changes in MCI [166, 156]. However, single composite scores may mask the heterogeneity of cognitive impairment.

Patients diagnosed with MCI show varying levels of impairment in different cognitive domains beyond memory, including language, visuospatial skills, attention and executive function [116, 210, 66]. This heterogeneity is likely linked to differences in the clinical evolution [186, 78]. Therefore, evaluation of domain specific changes could help to identify individuals at greater risk of progressing to dementia. Composite scores for measuring specific domain impairment have been proposed for memory [30] and executive function [69]. These scores mitigate the effect of measurement errors for individual items while combining informative sub-scores from multiple tests. Evaluation of these two previously proposed scores demonstrated they show better performance than individual test scores in detecting domain changes over time and predicting conversion from MCI to dementia [30, 69].

This chapter presents a data-driven framework which learns to combine and weight sub-scores from the neuropsychological test battery to calculate a set of domain-specific composite scores that quantify impairment in 6 domains: memory, language, visuospatial abilities, executive functioning, orientation and attention. The weighting scheme was obtained by estimating the parameters of a multi-factor model with Confirmatory Factor Analysis (CFA). The usefulness of the developed composite scores in MCI was evaluated in two different tasks using machine learning methods. First, the set of composite scores was taken as input for unsupervised cluster analysis, aiming to identify different sub-groups of individuals in the MCI sample. Second, we tested the ability of composite scores to predict progression from MCI to dementia within specific time windows, ranging from 1 to 5 years, and compared the performance against standard outcomes.

4.2   Methods

The data-driven methodology presented here is divided in two parts (Figure 4.1). The first part consists in learning the parameters for sub-score standardization and domain scores calculation. The second part evaluates the composite scores in two automated tasks: clustering of patients diagnosed with MCI, and predicting progression to dementia.
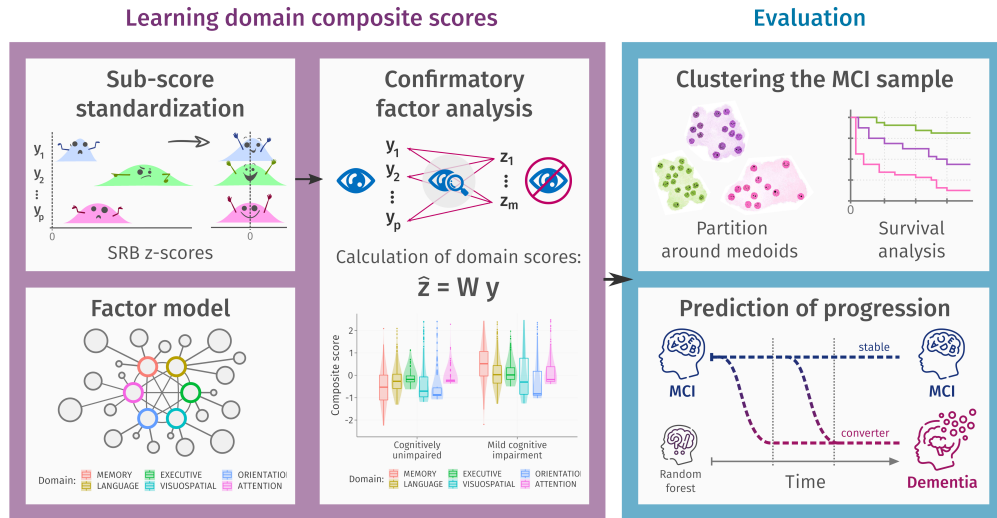
Figure 4.1: The proposed data-driven methodology can be divided in two blocks: learning and evaluation. During the learning phase, the parameters for sub-score standardization and domain composite calculation are estimated using a data sample including cognitively unimpaired participants and MCI patients. In the second part, the calculated domain scores for a separate sample of MCI are evaluated in terms of two tasks: unsupervised clustering of patients and prediction of future progression to dementia within different time windows.

4.2.1 Participants data

Data was provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The ADNI is a public-private partnership with the primary goal of testing whether magnetic resonance imaging (MRI), positron emission tomography (PET), biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For additional and up-to-date information, see `www.adni-info.org`. The dataset herein used comprised 680 patients with MCI and 668 cognitively unimpaired (CU) participants. The demographics and characteristics of these groups are presented in Table 4.1, corresponding to the first visit with the available information.

4.2.1.1 Data partition

The ADNI sample was split following the two methodological parts: learning and evaluation. For the learning set, 60% of the CU sample ($n = 400$) was taken as normative data for sub-score standardization while the remaining 40% ($n = 268$) and 40% of the MCI sample ($n = 272$) were used to learn the parameters for calculating the composite scores with CFA. The evaluation set corresponded to the remaining 60% of MCI participants ($n = 408$), for which composite scores were calculated using the parameters from the learning set.

	Learning set			Evaluation set
	CU ($n = 668$)		MCI ($n = 680$)	
	Normative data	CFA	CFA	Evaluation
	($n = 400$)	($n = 268$)	($n = 272$)	($n = 408$)
Sex (% female)	54.5	59.3	44.5	40.7
Age (mean ± sd)	73.4 ± 6.9	72.6 ± 8.0	72.6 ± 8.2	72.8 ± 7.8
APOE-ε4 (% carriers)	31.0	28.9	43.8	46.7
CDR-SOB (mean ± sd)	0.1 ± 0.2	0.1 ± 0.2	1.5 ± 1.0	1.5 ± 1.0
MMSE (mean ± sd)	29.2 ± 1.1	28.9 ± 1.2	27.9 ± 1.8	28.0 ± 1.7
ADAS-Cog (mean ± sd)	10.0 ± 4.7	11.1 ± 4.5	16.4 ± 6.8	15.0 ± 6.8

Table 4.1: Description of sets used in each step of the methodology, including the percentage of carriers of the ε4 allele of the apolipoprotein E (APOE) gene, and distributions of total scores for the Mini-Mental State Examination (MMSE), Clinical Dementia Rating - Sum of Boxes (CDR-SOB), and the Alzheimer's Disease Assessment Scale - Cognition (ADAS-Cog). *sd: standard deviation.

4.2.1.2 Neuropsychological data

Sub-scores from nine different tests were used in the present study, namely: the Alzheimer's Disease Assessment Scale - Cognition (ADAS-Cog) [162], Mini-Mental State Examination (MMSE) [60], Montreal Cognitive Assessment (MoCA) [133], Rey auditory verbal learning test (AVLT), Logical Memory test immediate and delayed [208], Clock Drawing test [77], Category Fluency test [131], Trail Making A and B [158], and one of the naming tests depending on its availability: Boston Naming test [102] or Multilingual Naming test [75]. The initial list of 50 sub-scores is presented in Table 4.2.

Sub-score code	Test	Description
Q1SCORE	ADAS-Cog	Word Recall
Q2SCORE	ADAS-Cog	Commands
Q3SCORE	ADAS-Cog	Constructional Praxis
Q4SCORE	ADAS-Cog	Delayed Word Recall
Q5SCORE	ADAS-Cog	Naming
Q6SCORE	ADAS-Cog	Ideational Praxis
Q7SCORE	ADAS-Cog	Orientation
Q8SCORE	ADAS-Cog	Word Recognition
Q9SCORE	ADAS-Cog	Remembering Test Instructions
Q10SCORE	ADAS-Cog	Comprehension
Q11SCORE	ADAS-Cog	Word-finding Difficulty
Q12SCORE	ADAS-Cog	Language
Q13SCORE	ADAS-Cog	Number cancellation
MMORITIME	MMSE	Orientation to time
MMORISPACE	MMSE	Orientation to space
MMREGI	MMSE	Three word registration
MMRECALL	MMSE	Three word recall
MMSPELLBKW	MMSE	Spelling a 5 letters word backwards
MMNAM	MMSE	Naming 2 objects
MMCOMMAND	MMSE	Following a verbal command
MMREPEAT	MMSE	Repeating a short sentence
MMREAD	MMSE	Reading a sentence wit an instruction
MMWRITE	MMSE	Writing a sentence about anything
MMDRAW	MMSE	Copying a drawing
TRAILS	MoCA	Trails
CUBE	MoCA	Copying a cube drawing
MOCACLOCK	MoCA	Drawing a clock

MOCANAM	MoCA	Naming 3 animals
MOCADIG	MoCA	Repeating digits forward and backwards
MOCALET	MoCA	Tapping with the hand when a letter is read from a list
MOCASERIAL	MoCA	Serial subtraction starting at 100
MOCAREP	MoCA	Repeating 2 sentences
MOCAFLUEN	MoCA	Naming words that beginwith the letter F
MOCAABS	MoCA	Abstraction of similarities between words
MOCADLREC	MoCA	Five word recall
MOCAORI	MoCA	Orientation to time and space
CLOCKSCOR	Clock Drawing test	Drawing a clock with details
COPYSCOR	Clock Drawing test	Copying the drawing of a clock
TRAASCOR	Trail Making test	Time to complete Part A
TRABSCOR	Trail Making test	Time to complete Part B
LIMMTOTAL	Logical Memory test	Immediate recall of a story read by the examiner
LDELTOTAL	Logical Memory test	Delayed recall of a story read by the examiner
CATANIMSC	Category Fluency test	Naming animals
RAVLT.IMMED	Rey AVLT	Repeating a list of 15 words 5 times
AVTOT6	Rey AVLT	Recall of the first list of words after a second list was read
AVTOTB	Rey AVLT	Repeating words from the second list
AVDEL30MIN	Rey AVLT	Recall of words from the first list after 30 minutes
AVDELTOT	Rey AVLT	Delayed recognition of written words from the first list
BMNOCUE	Boston / Multilingual Naming test	Naming objects in pictures
BMCUED	Boston / Multilingual Naming test	Naming objects in pictures after semantic cues

Table 4.2: The set of 50 sub-scores from 9 neuropsychological tests initially considered for the analysis.

4.2.2 Learning domain composite scores

4.2.2.1 Sub-score standardization

Given the heterogeneous scales of neuropsychological tests, some of the scales were inverted to ensure that increasing values correspond to poorer performance. The initial set of 50 sub-scores were transformed into standardized regression based (SRB) z-scores using the parameters learnt from a normative sample. Specifically, each sub-score x was modelled as linear function of age and years of education:

$$x = \beta_0 + \beta_{ed} \times \text{Education} + \beta_{age} \times \text{Age} + \varepsilon \tag{4.1}$$

Linear regression parameters $\beta_0, \beta_{ed}, \beta_{age}$ and $\sigma^2 = \text{var}(\varepsilon)$ were estimated with data from 400 CU participants. Then, the corresponding SRB z-score for each participant i (denoted y_i for consistency with upcoming formulations) was calculated as:

$$y_i = \frac{x_i - \widehat{x_i}}{\sigma} = \frac{x_i - \left(\beta_0 + \beta_{ed} \times \text{Education}_i + \beta_{age} \times \text{Age}_i\right)}{\sigma} \tag{4.2}$$

The sub-score from the naming test after a semantic clue (BMCUED) was dropped from further analysis because higher values, after scale inversion, can be associated with poor performance or perfect performance without the cue.

4.2.2.2 Derivation of domain scores

The estimation of composite measures for 6 different domains was done by proposing and testing a factor model which links a set of sub-scores from multiple tests with six domains: memory, language, visuospatial abilities, executive functioning, orientation and attention (See Figure 4.2). Before establishing a factor model, variability of sub-scores and pairwise correlations were examined in the data partition used for CFA. Sub-scores whose variance was inflated by a few outliers were not included in the model, neither were the sub-scores showing no significant correlation (greater than 0.25) with any other one and were not evaluating a similar task. The factor model was proposed taking into account what sub-scores evaluate, but also the number of previous works that performed Factor Analysis on similar neuropsychological test batteries [30, 69, 138, 72].

Factor analysis formulation

Factor analysis methods exploit the correlations between observed measures $\mathbf{y}$ to quantify the influence of unobserved factors $\mathbf{z}$ [31]. Let the vector $\mathbf{y_i} \in \mathbb{R}^{m \times 1}$ be the i-th observation of m variables, and $\mathbf{z_i} \in \mathbb{R}^{p \times 1}$ the unobserved measures of p factors. The *common factor model* states:

$$\mathbf{y_i} = \mathbf{Fz_i} + \mathbf{e_i} \tag{4.3}$$

Or component-wise:

$$y_{ij} = \sum_{k=1}^{m} f_{jk} z_{ik} + e_{ij} \tag{4.4}$$
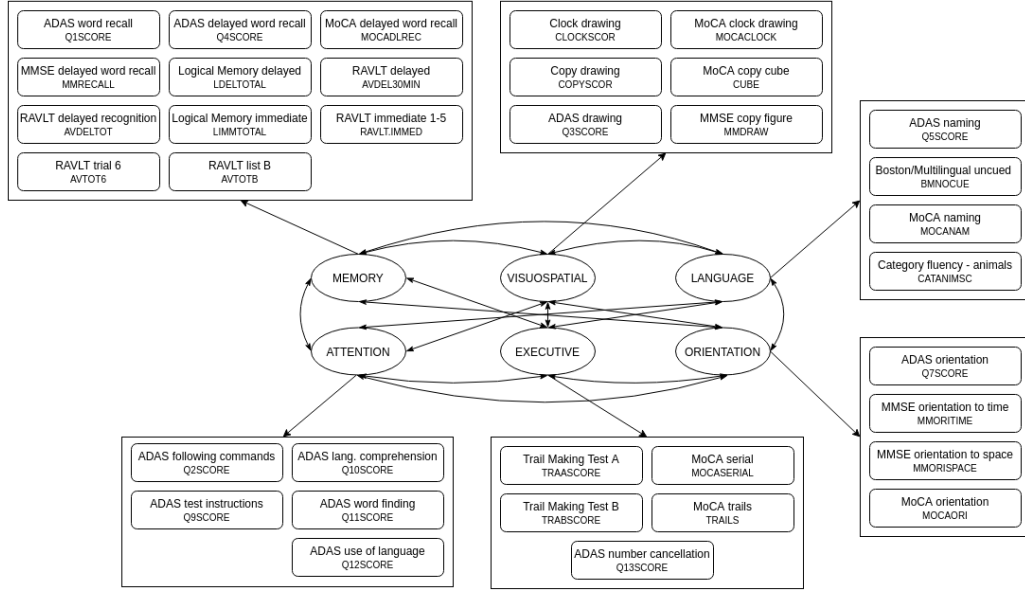
Figure 4.2: Proposed factor model connecting 6 cognitive domains with 35 sub-scores from 9 different neuropsychological tests. Sub-scores code and description is presented in Table 4.2.

Matrix $\mathbf{F} \in \mathbb{R}^{m \times p}$ contains the *factor loadings*, also known as the factor structure. Residuals $e_{\cdot j} \in \mathbb{R}^{1 \times m}$ contain the portion of the j-th variable that is not defined by the factors and matrix of residuals correlations $\mathbf{C_e} \in \mathbb{R}^{m \times m}$ is assumed diagonal. Therefore, components of the correlation matrix between observed variables $\mathbf{C_y} \in \mathbb{R}^{m \times m}$ are given by:

$$\left[\mathbf{C_y}\right]_{jl} = \operatorname{corr}(y_{\cdot j}, y_{\cdot l})$$
$$= \sum_{s=1}^{m} \sum_{t=1}^{m} f_{js} f_{lt} \left[\mathbf{C_z}\right]_{st} \quad \text{for } j \neq l \tag{4.5}$$

In Exploratory Factor Analysis (EFA) the matrices $\mathbf{F}$, $\mathbf{C_z}$, and $\mathbf{C_e}$ are estimated without any assumptions about the underlying factor structure. On the contrary, Confirmatory Factor Analysis (CFA) estimates those matrices for an hypothesized factor model from Equation 4.5. CFA was performed with the *lavaan* package [163] in R (v.3.6.3) using the unweighted least squares estimator. Model fit was evaluated by the Root Mean Square Error of the Approximation (RMSEA) and the Tucker-Lewis index (TLI), these measures evaluate a model in relation to a baseline model which assumes all variables to be independent.

Domain scores calculation

The factor structure matrix $\mathbf{F} \in \mathbb{R}^{m \times p}$ quantifies the influence of the p unobserved factors over the m observed variables. The set of unobserved factors $\mathbf{z_i}$ for a particular case i can be calculated as linear combinations of its observations $\mathbf{y_i}$ [15, 79]. That is, the estimated

vector of factors, $\widehat{\mathbf{z_i}}$, is given by:

$$\widehat{\mathbf{z_i}} = \mathbf{Wy_i} \qquad (4.6)$$

With $\mathbf{W} \in \mathbb{R}^{p \times m}$ a weight matrix that needs to be estimated. A solution that minimizes the sum of squares of the uniqueness [15], *i.e.*, the portion of the observations variance that is not explained by the factors, is given by:

$$\mathbf{W} = \left(\mathbf{F}^T \mathbf{C_e}^{-1} \mathbf{F}\right)^{-1} \mathbf{F}^T \mathbf{C_e}^{-1} \qquad (4.7)$$

The resulting estimated factor values quantify dysfunction of the different domains included in the model. The learnt set of weights can be used to calculate the domain specific scores of new observations once they have been transformed into SRB z-scores.

4.2.3 Evaluation

4.2.3.1 Clustering the MCI sample

By exploring the existence of MCI subgroups with an unsupervised clustering method, the six composite scores expose different cognitive profiles in the MCI sample. Specifically, the Partition Around Medoids (PAM) method, also known as k-medoids, iteratively splits the data set in k clusters, being the k representative points the most central points (medoid) in each cluster, and the remaining points assigned to the cluster with the nearest representative point [111, 159].

Here we incorporated the inherent relations between domains by including the covariance matrix $\mathbf{C_z}$ in the calculation of distance between subjects, the distance between a pair of subjects i and j, described by their domain scores $\mathbf{z_i}$ and $\mathbf{z_j}$, was defined as:

$$d(\mathbf{z_i}, \mathbf{z_j})^2 = (\mathbf{z_i} - \mathbf{z_j})\mathbf{C_z}(\mathbf{z_i} - \mathbf{z_j})^T \qquad (4.8)$$

Unlike the Mahalanobis distance, the distance between two subjects is weighted by the covariance between factors, thereby ensuring that the largest variance dimensions contribute more to the differences between subjects. The matrix used for this step was the estimated covariance matrix between domains $\mathbf{C_z}$ resulting from the CFA.

The number of clusters was set by revising a collection of 30 indices [25] for multiple options of k from 2 to 10. Cluster stability for these possible partitions ($2 \leq k \leq 10$) was also evaluated following a bootstrap approach. For a given partition in k subgroups, this process consists in partitioning a sub-sample of the data (80%), calculating the subset of observations that remains in their initial cluster and repeating this process multiple times (1000 iterations). The overlap between the initial clusters and bootstrap clusters was assessed via the Jaccard coefficient and the mean value of this index over the total of repetitions is reported for all clusters.

Differences between MCI subgroups

Resulting subgroups of MCI participants were compared in terms of their composite scores per domain and their risk of progression to dementia. Pairwise domain score differences between sub-groups were examined with Wilcoxon-Mann-Whitney U tests while applying the Bonferroni correction for multiple comparisons. A multivariate Cox proportional hazard regression model tested the sub-group effect in the progression from MCI to AD dementia while controlling for age, gender and years of education. Kaplan-Meier survival curves illustrated progression to dementia of the different MCI sub-groups, and curves were compared using omnibus and pairwise log-rank tests. A multivariate Cox proportional hazard regression model was used to evaluate the effect of the cognitive profile on the progression from MCI to AD dementia while controlling for age, gender and years of education. The resulting hazard ratios (HR) account for the risk difference of each MCI sub-group with respect to a reference group.

4.2.3.2 Prediction of progression to dementia

Domain specific scores were also evaluated in the automated prediction of progression from MCI to AD dementia. This evaluation consisted in classifying MCI patients as either stable or converters following the time window approach[141] fixing five different time periods: 12, 24, 36, 48, and 60 months. The 6 composite scores along with age, gender and years of education were used to train random forest classifiers [20]. A Random forest (RF) is an ensemble of decision trees constructed using a bootstrap aggregating approach. To create each decision tree, a new training set is generated by sampling, uniformly and with replacement, the original training set. This procedure ensures the collection of trees comes from independent identically distributed samples. The prediction is given by the majority voting of the decision trees in the ensemble, effectively improving the prediction accuracy [20].

Classification performance was assessed by constructing the Receiver Operating Characteristic (ROC) curve and calculating its Area Under the Curve (AUC). Depending on the time window, data for training the classifier might be highly unbalanced. This was taken into account when designing the cross validation scheme: at each iteration, a random forest classifier was trained with a balanced subset by randomly selecting the 70% of the underrepresented class with an equal number of samples from the other class. The classifier was tested with the remaining observations, in some cases reaching a larger number of samples. This process was repeated 1000 times per time window.

4.3 Results

4.3.1 Parameters for composite scores calculation

4.3.1.1 Sub-score standardization

Parameters to calculate the SRB z-scores (Equation 4.2), obtained from the linear regression with a sample of 400 cognitively unimpaired individuals are presented in Table

4.3.

Sub-score	Intercept (β_0)	Education (β_{ed})	Age (β_{age})	σ
Q1SCORE	1.044	-0.045	0.034	1.294
Q2SCORE	0.072	-0.001	0	0.353
Q3SCORE	0.178	-0.015	0.006	0.55
Q4SCORE	-0.549	-0.065	0.06	1.811
Q5SCORE	-0.094	-0.001	0.002	0.265
Q6SCORE	0.145	-0.001	-0.001	0.231
Q7SCORE	0.258	-0.009	0	0.321
Q8SCORE	5.095	0.035	-0.037	2.375
Q9SCORE	-0.042	-0.006	0.002	0.148
Q10SCORE	-0.203	0.001	0.003	0.148
Q11SCORE	-0.266	0	0.004	0.257
Q12SCORE	-0.138	0.001	0.002	0.171
Q13SCORE	-0.405	-0.001	0.015	0.819
MMORITIME	-0.024	0.001	0.001	0.301
MMORISPACE	0.432	-0.009	-0.002	0.351
MMREGI	-0.02	-0.001	0.001	0.1
MMRECALL	-0.556	-0.016	0.015	0.621
MMSPELLBKW	0.529	-0.032	0.002	0.492
MMNAM	-0.05	0.001	0.001	0.05
MMCOMMAND	0.103	-0.007	0.001	0.244
MMREPEAT	0.252	-0.005	-0.001	0.264
MMREAD	0.028	0	0	0.05
MMWRITE	-0.02	0.001	0	0.05
MMDRAW	0.183	-0.01	0.001	0.241
TRAILS	-0.078	-0.008	0.004	0.27
CUBE	0.451	-0.03	0.004	0.447
MOCACLOCK	-0.058	-0.003	0.005	0.483
MOCANAM	-0.102	0	0.003	0.276
MOCADIG	-0.001	-0.008	0.003	0.26
MOCALET	-0.113	-0.002	0.003	0.177
MOCASERIAL	0.153	-0.016	0.003	0.385
MOCAREP	0.14	-0.018	0.005	0.454
MOCAFLUEN	0.198	-0.016	0.003	0.385

MOCAABS	0.457	-0.029	0.002	0.393
MOCADLREC	-2.029	-0.049	0.074	1.691
MOCAORI	0.271	-0.005	-0.002	0.254
CLOCKSCOR	0.092	-0.012	0.005	0.536
COPYSCOR	0.167	0.003	-0.001	0.42
TRAASCOR	2.772	-0.099	0.42	9.601
TRABSCOR	9.705	-1.968	1.386	39.867
LIMMTOTAL	15.084	-0.242	-0.012	3.086
LDELTOTAL	15.238	-0.234	-0.002	3.27
CATANIMSC	36.053	-0.62	0.174	5.35
RAVLT.IMMED	2.848	-0.473	0.456	9.81
AVTOT6	-0.858	-0.15	0.123	3.248
AVTOTB	4.328	-0.072	0.089	1.988
AVDEL30MIN	-1.107	-0.075	0.129	3.9
AVDELTOT	-1.478	-0.042	0.058	2.267
BMNOCUE	2.035	-0.134	0.026	1.995
BMCUED	29.761	0.032	-0.007	0.592

Table 4.3: Parameters for sub-score standardization estimated from normative data. Complete description of each sub-score code is presented in Table 4.2.

4.3.1.2 Factor analysis

The proposed factor model was constructed with 35 sub-scores linked to 6 cognitive domains: memory, language, executive function, visuo-spatial, orientation and attention. Fit statistics given by Confirmatory Factor Analysis indicate a good model fit (RMSEA = 0.09, TLI = 0.95). Once the factor model parameters are estimated, dysfunction measures for each domain are obtained as linear combinations sub-scores, the resulting set of weights for domain score calculation are presented in Table 4.4.

	Domain					
Sub-score	Memory	Language	Executive	Visuospatial	Orientation	Attention
Q1SCORE	0.126	0	0	0	0	0
Q4SCORE	0.125	0	0	0	0	0
MOCADLREC	0.061	0	0	0	0	0
RAVLT.IMMED	0.15	0	0	0	0	0
AVTOT6	0.128	0	0	0	0	0
AVTOTB	0.052	0	0	0	0	0

AVDEL30MIN	0.068	0	0	0	0	0
AVDELTOT	0.052	0	0	0	0	0
LIMMTOTAL	0.085	0	0	0	0	0
LDELTOTAL	0.113	0	0	0	0	0
MMRECALL	0.039	0	0	0	0	0
Q5SCORE	0	0.097	0	0	0	0
MOCANAM	0	0.093	0	0	0	0
BMNOCUE	0	0.291	0	0	0	0
CATANIMSC	0	0.306	0	0	0	0
Q13SCORE	0	0	0.026	0	0	0
TRAASCOR	0	0	0.065	0	0	0
TRABSCOR	0	0	0.224	0	0	0
MOCASERIAL	0	0	0.036	0	0	0
TRAILS	0	0	0.025	0	0	0
CLOCKSCOR	0	0	0	0.395	0	0
COPYSCOR	0	0	0	0.229	0	0
MOCACLOCK	0	0	0	0.448	0	0
Q3SCORE	0	0	0	0.157	0	0
CUBE	0	0	0	0.164	0	0
MMDRAW	0	0	0	0.048	0	0
Q7SCORE	0	0	0	0	0.331	0
MMORITIME	0	0	0	0	0.313	0
MMORISPACE	0	0	0	0	0.114	0
MOCAORI	0	0	0	0	0.418	0
Q9SCORE	0	0	0	0	0	0.074
Q10SCORE	0	0	0	0	0	0.092
Q11SCORE	0	0	0	0	0	0.204
Q12SCORE	0	0	0	0	0	0.148
Q2SCORE	0	0	0	0	0	0.085

Table 4.4: Weight of each sub-score in the calculation of the 6 domain scores (Matrix $\mathbf{W}^T$). Complete description of each sub-score code is presented in Table 4.2.

Domain dysfunction scores of all subjects were calculated using the learnt parameters while differences between cognitively normal subjects and MCI patients were tested using Mann-Whitney U tests. As one test was performed per domain, p-values were adjusted using the Bonferroni correction for the 6 tests. The 6 domains differ significantly

between these groups: memory ($r = 0.59, p < 0.00005$), language ($r = 0.34, p < 0.00005$), executive functioning ($r = 0.33, p < 0.00005$), visuospatial abilities ($r = 0.20, p < 0.00005$), orientation ($r = 0.25, p < 0.00005$), and attention ($r = 0.21, p < 0.00005$).

4.3.2 Sub-groups of MCI patients

The cognitive state of MCI participants was characterized by the six domain scores and different impairment profiles were found in the MCI patient sample by cluster analysis. Once the distance between subjects is estimated, there are multiple criteria to choose the number of clusters (k) in which data could be divided. After examining 30 different indices [25], data partition in 4 clusters was suggested by 13 of these indices. Additionally, the mean cluster stability index was checked for multiple values of k resulting in values above 0.85 for $2 \leq k \leq 4$. Partition around medoids (PAM) was applied to divide the sample of 408 MCI patients in 4 different subgroups. The description of these subgroups is presented in Table 4.5 along with the description of the entire group of cognitively unimpaired participants as a reference. Figure 4.3 shows the distributions of domain dysfunction scores for each one of the MCI subgroups. A total of 60 pairwise tests were performed to compare domain composite scores between MCI subgroups and against the CU group, effect size r was computed for each test and p-values were adjusted for multiple comparisons using the Bonferroni correction. Two profiles were observed at the extremes of the dysfunction spectrum: group 1 exhibits the lowest impairment in all domains, with all score distributions being comparable with the CU group, and group 4 has the highest average dysfunction scores in 5 out of 6 domains.

In the control-like subgroup 49 out of 159 individuals progressed to dementia on the course of the follow-up, those participants converted on average 44.5 months after evaluation. This particular sub-group supports previous findings which suggest a considerable number of false positives in the diagnosis of MCI in ADNI database [54, 58].

Characterization of MCI participants with the 6 proposed domain dysfunction scores revealed 4 different cognitive profiles in the sample of ADNI participants diagnosed with MCI:

- The first subgroup (MCI 1) with the lowest mean dysfunction scores for all 6 domains compared to the other MCI subgroups. When compared to controls, this group shows significantly higher memory dysfunction ($r = 0.31, p < 0.00005$) and lower visuospatial dysfunction score ($r = 0.17, p = 0.00008$). Indeed, these participants should have exhibited some memory impairment during the neuropsychological evaluation to be diagnosed with MCI according to the ADNI criteria.

- Subjects in MCI 2 show higher impairment in memory than MCI 1 ($r = 0.39, p < 0.00005$), language ($r = 0.44, p < 0.00005$), executive function ($r = 0.25, p = 0.0012$), and visuospatial abilities ($r = 0.75, p < 0.00005$). Although the attention dysfunction does not differ from MCI 1, the difference of this domain with respect to the CU group is significant but small ($r = 0.18, p = 0.00001$).

- The third subgroup (MCI 3) differs from MCI 2 only in memory ($r = 0.25, p = 0.015$) and orientation ($r = 0.81, p < 0.00005$).

	Cognitively unimpaired	MCI subgroup			
		MCI 1	MCI 2	MCI 3	MCI 4
N	668	159	129	88	32
Age (years)	73.1±7.4	72.4±7.8	72.2±7.8	74.6±7.5	72.9±7.7
Sex (% female)	56.4	43.4	34.9	43.2	43.8
APOE-ε4 (% carriers)	30.2	35.9	47.3	55.2	75.0
Memory	-0.59±0.70	-0.04±0.63	0.57±0.72	0.95±0.73	1.51±0.67
Language	-0.26±0.56	-0.18±0.49	0.39±0.72	0.50±1.01	0.66±1.00
Executive	-0.14±0.26	-0.09±0.26	0.13±0.47	0.19±0.44	0.31±0.62
Visuospatial	-0.27±0.96	-0.71±0.43	0.83±1.09	0.49±1.40	1.22 ± 1.35
Orientation	-0.51±0.82	-0.70±0.41	-0.59±0.45	1.60±0.98	5.06±1.84
Attention	-0.15±0.35	-0.08±0.38	0.11±0.72	0.28±1.00	0.22±0.89
Mean CSI	-	0.97	0.94	0.90	0.85
Cox proportional HR	-	ref.	2.57	3.84	7.68
95% CI	-	ref.	1.59 - 4.20	2.33 - 6.30	4.32 - 13.70

Table 4.5: Description of MCI subgroups, along with the CU sample for reference. Demographic information, mean and standard deviation (sd) of domain composite scores, mean cluster stability index (CSI), and proportional hazard ratios (HR) with their 95% confidence intervals (CI).

- The last subgroup MCI 4 differs from MCI 3 in memory ($r = 0.32, p = 0.026$) and orientation ($r = 0.73, p < 0.00005$).

Kaplan-Meier survival curves for the 4 subgroups of MCI are illustrated in Figure 4.4, according to the omnibus log-rank test, survival curves for the 4 subgroups differ significantly ($\chi^2_3 = 64.2, p \leq 0.001$). According to the pairwise comparison between curves, MCI subgroup 1 exhibits significantly lower progression probability than subgroup 2 ($\chi^2_1 = 15.61, p = 0.0001$), and subgroup 4 has significantly higher progression probability than subgroup 3 ($\chi^2_1 = 5.74, p = 0.02$). Although the difference between subgroups 2 and 3 does not reach the significance level of 0.05 after false discovery rate correction, the adjusted p-value is still relatively low ($\chi^2_1 = 3.65, p = 0.056$). The resulting MCI subgroups show with distinctive survival curves confirming that the different cognitive profiles are related with different progression risk.

Differences of progression risk across MCI sub-groups were quantified using multivariate Cox models taking the control-like subgroup (MCI 1) as reference and including gender, age and years of education as covariates. The resulting proportional HR are presented in Table 4.5, HR estimates for MCI subgroups 2 and 3 compared with the control-like subgroup are 2.57 (95% CI [$1.59 - 4.20$]) and 3.84 (95% CI [$2.33 - 6.30$]), re-
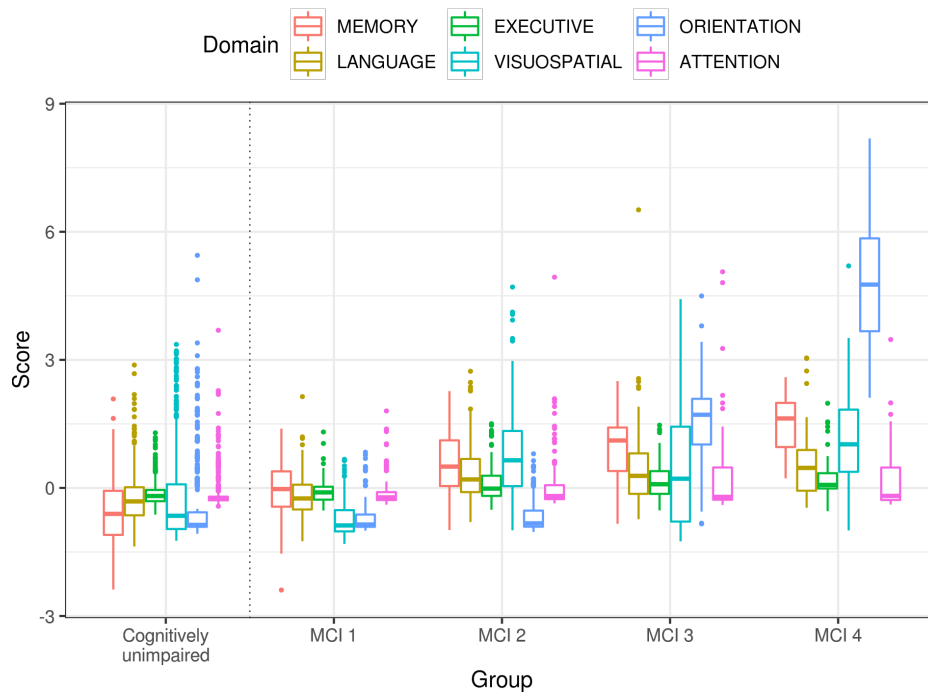
Figure 4.3: Distribution of domain dysfunction scores per MCI subgroup including the complete group of cognitively unimpaired (CU) participants as reference.

spectively. Significantly higher hazard ratio results for MCI subgroup 4 which have a risk of progression to AD dementia around 7.7 (95% CI [4.32 − 13.70]) times higher than the risk for the control-like subgroup. From the Cox model, age, years of education and gender had no effect.

4.3.3 Automated prediction of progression to AD dementia

Random Forest classifiers were trained to classify between MCI patients who remained stable (sMCI) and the ones who converted to dementia (cMCI) using data from the evaluation set. The number of cases in each one of these two groups depended on the time window being considered, Table 4.6 presents the number of cases used for training and testing the classifiers for the 5 time windows considered 4.6.

Included features for classification were the six domain scores and years of education, age, and gender. To compare with standard outcome measures, at each iteration of the validation scheme, two additional classifiers were trained while including the same covariates. The first one was trained with the scores of commonly used neuropsychological tests, namely the ADAS-Cog, MMSE, MoCA, and AVLT while the second was trained only with the ADAS-Cog. The number of trees for all RF was set at 200. The distribution of AUC values per time period across the 1000 iterations for the three classifiers is shown in Figure 4.5, mean AUC for classification with domain scores are 0.68, 075, 0.74, 0.74, and 0.76 for prediction within 12, 24, 36, 48, and 60 months, respectively.
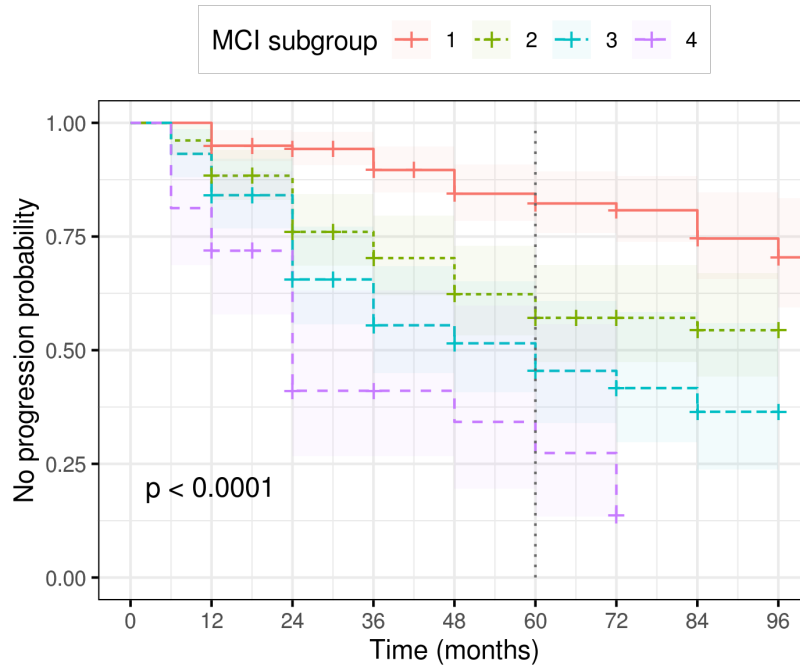
Figure 4.4: Kaplan-Meier curves for the found MCI sub-groups

Classifier performance is significantly higher when trained with domain scores rather than with the set of test totals, including the ADAS-Cog. When predicting MCI conversion within 12 months, resulting mean AUCs are 0.68 and 0.63 (Cohen's $d = 0.73$, $p \leq 0.00001$) for classifiers trained with dysfunction scores and total tests, respectively. When the conversion prediction is done within 60 months, these mean AUC values are 0.76 and 0.69 (Cohen's $d = 1.59$, $p \leq 0.00001$), respectively.

Although it might be counter-intuitive that prediction performance is better for the long term than for the short term, this is likely due to the varying number of cases used for training and testing at each time window. With longer time windows, the number of stable MCI subjects decreases while the number of MCI who converted to dementia increases. Although all RF classifiers were trained with balanced sets of cases, classifiers within 1 year were trained with fewer samples and tested with larger and more unbalanced sets, making this experiment more challenging than the classification within longer time windows (See Table 4.6).

4.3.3.1 Direct comparison with state-of-the-art predictors

To compare the prediction of MCI progression to dementia with domain scores against other composite scores and predictors in the literature, nine different sets of features were used to train the RF classifiers following a random sampling cross-validation scheme with 200 iterations. The nine sets of predictors are:

Time window	Total		Training RF		Testing RF	
	sMCI	cMCI	sMCI	cMCI	sMCI	cMCI
12 months	356	46	32	32	324	14
24 months	263	82	57	57	206	25
36 months	206	99	69	69	137	30
48 months	159	114	80	80	79	34
60 months	109	122	76	76	33	46

Table 4.6: Number of MCI subjects that remained stable (sMCI) and converted to dementia (cMCI) within each time window, along with the number of subjects per class that were used to train and test the Random Forest (RF) classifier at each iteration of the cross-validation scheme.

1. PROPOSED domain-specific composite scores.

2. PROPOSED domain-specific composite scores, with the Clinical Dementia Rating (CDR) - Sum of Boxes, and the Functional Assessment Questionnaire (FAQ).

3. ADAS Tree [115] = 1.05*Q1SCORE + 0.38*Q2SCORE + 0*Q3SCORE + 1.17*Q4SCORE + 0.61*Q5SCORE + 0.13*Q6SCORE + 1.13*Q7SCORE + 0.41*Q8SCORE + 0.54*Q9SCORE + 0.49*Q10SCORE + 0.69*Q11SCORE + 0.39*Q12SCORE + 0.68*Q13SCORE.

4. Composite [88] = Q1SCORE + Q4SCORE + Q7SCORE + CDRSB + FAQTOTAL.

5. Cognitive composite 1 [156]: CC1 = ADAS3 + (75-RAVLT.IMMED) + (30 - MMTOTAL).

6. Cognitive composite 2 [156]: CC2 = ADAS3 + CDMEMORY.

7. Cognitive–functional composite 1 [156]: CFC1 = CC1 + FAQTOTAL.

8. Cognitive–functional composite 2 [156]: CFC2 = CC2 + FAQTOTAL.

9. Selected features [140]: TRABSCOR, Forget.index, RAVLT.IMMED, TOTAL13, TRAASCOR, AVTOT6, LIMMTOTAL, CATANIMSC, AVDEL30MIN, FAQTOTAL, LDELTOTAL, MOCADLREC, AVDELTOT, BNTTOTAL, Q4SCORE, Q8SCORE, MMTOTAL, Q1SCORE, MOCAFLUEN, CDORIENT, CDHOME, AVTOTB.

The distribution of the AUC values for the 9 classification experiments along the 5 time windows is presented in Figure 4.6. The proposed domain composite scores outperform the other predictors that rely only on cognitive measures. When functional measures such as the FAQ and CDR are included in the set of predictors, AUC values improve for all the time windows. In particular, MCI conversion prediction with the domain composite scores and functional measures is slightly better than the prediction with 22 selected features from the battery of cognitive and functional assessments.
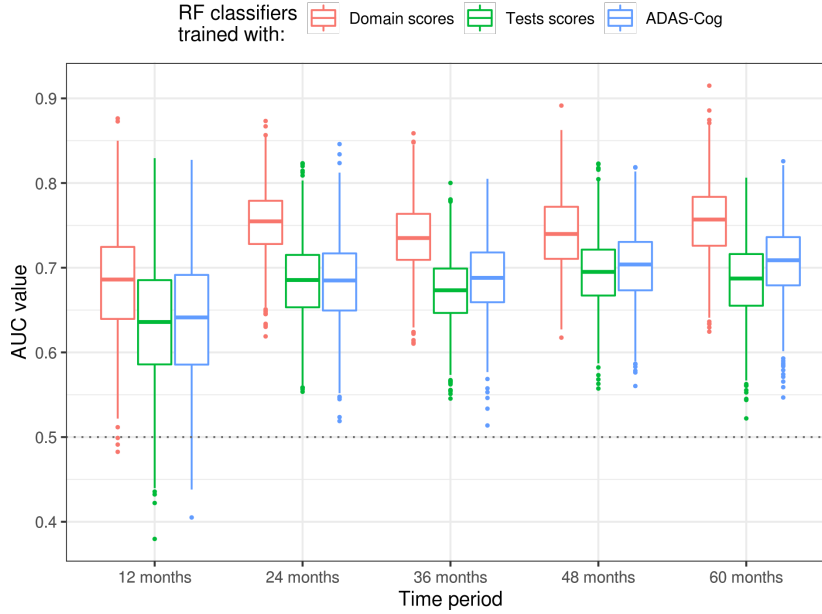
Figure 4.5: Distribution of AUC values for prediction of progression from MCI to dementia within 12, 24, 36, 48 and 60 months. Classifiers trained with composite domain scores consistently outperform classifiers trained with the ADAS-Cog, and with the set of total tests scores from ADAS-Cog, Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), and Rey auditory verbal learning test (AVLT).

4.4 Discussion

This work has introduced a data-driven methodology to characterize the cognitive state of patients diagnosed with MCI by developing specific domain scores using sub-scores from the neuropsychological tests battery applied to the ADNI participants. These domain scores highlight sub-groups of MCI patients who exhibit different risks of progression to AD dementia, and show better performance than standard outcomes when predicting conversion from MCI to dementia up to 5 years.

Factor model and composite scores

A 6 factor model estimates simultaneously composite scores for all the domains. By learning the weights for domain score calculation from a sample containing both CU and MCI in similar proportions, we can capture a more general statistical structure of the cognitive evaluation than if we had used a narrower sample within the spectrum of impairment. This is an extension of previous works that establish single factor models to obtain a composite measure for particular domains such as memory [30] and executive functioning [69]. Memory composite score in this work strongly agrees with the one hypothesized for ADNI-Mem [30], resulting therefore in highly correlated memory measures ($r = -0.943$, $p < 0.00005$). Executive function score proposed here is also correlated with ADNI-EF [69] ($r = -0.818$, $p < 0.00005$), even though sub-scores from ADAS-Cog
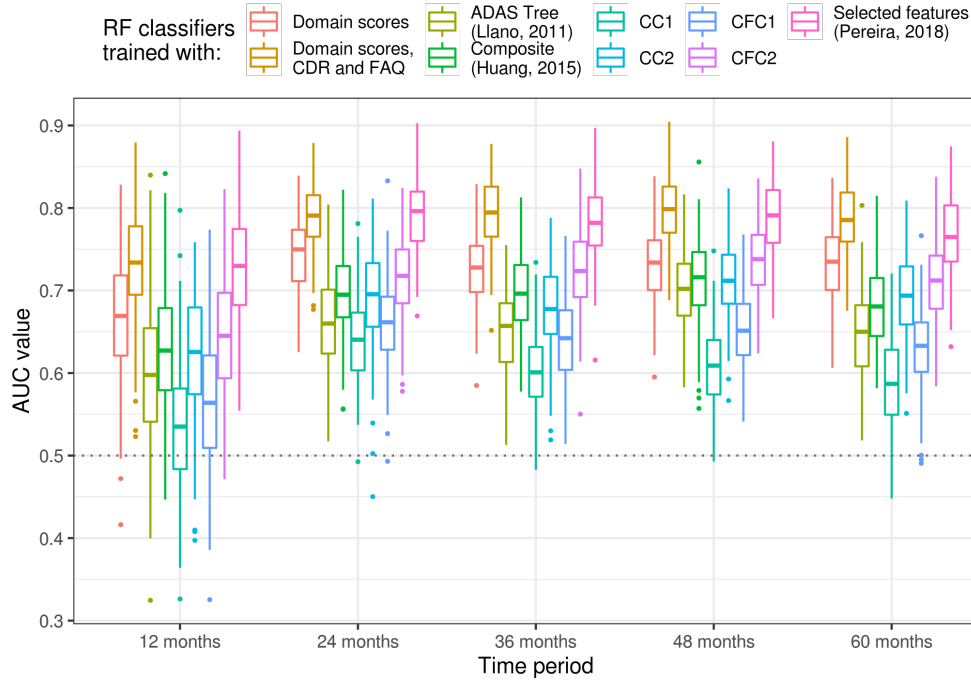
Figure 4.6: Distribution of AUC values for MCI conversion prediction within 12, 24, 36, 48 and 60 months using different sets of features.

and MoCA, not considered in ADNI-EF, were herein included.

MCI heterogeneity

As weights for domain score calculation were obtained as a solution that minimizes the portion of the variance that is not explained by the factors [15], the obtained composite scores do mitigate the effect of individual measurement errors, leading to more robust measures of impairment for each domain. This is a methodological advantage over previous works that studied MCI heterogeneity with separate neuropsychological scores per domain [18, 54, 58]. Another methodological advantage consists in adapting the notion of distance between subjects by including the domain covariance in the metrics. Most of the state-of-the-art research performs the cluster analysis [142, 18, 54, 56] using the euclidean distance to compare sets of cognitive variables between individuals. However, this distance relies on the assumption of orthogonality between dimensions and therefore each measure is considered independent from the other ones, an assumption hard to hold and far from the given nature of the data.

The cognitive characterization presented here produced a partition of the MCI group into 4 different sub-groups. Beyond the methodological differences, the obtained division is, to some extent, consistent with previous works investigating cognitive heterogeneity in MCI with ADNI data [18, 54, 58]. All these works also identified a sub-group of control-like individuals in the group of participants diagnosed with MCI according to ADNI criteria, and 2 or 3 MCI sub-groups which vary in the level of impairment

of memory, executive functions[18], and language[54]. In this work, the separation between the remaining three MCI sub-groups is guided by two domains that covariate closely, memory and orientation, while showing relatively similar levels of impairment in language, executive functioning, visuospatial abilities, and attention to the CU and the control-like MCI sub-group. Examination of future progression to dementia for the different MCI sub-groups in this study resulted in well differentiated survival curves, providing evidence for the usefulness of the proposed characterization to stratify the risk of progression to dementia during the upcoming 5 years. Therefore, the progressive risk of progression from MCI 1 to MCI 4 seems to be driven by memory and orientation. Although the important role of orientation might be unexpected, it is coherent with previous works that have identified orientation sub-scores among the most sensitive measures of cognitive change [156, 88]. The four MCI subgroups are similar in terms of age and sex distribution, but they exhibit differences in terms of the percentages of APOE-e4 carriers. Although the relation between APOE status and risk of AD dementia is widely known, the fact that this known pattern was exposed, in an unsupervised way, by orientation impairment might be worthy of further analysis in future work.

Predicting progression from MCI to dementia

The domain scores were also evaluated at automatically predicting future progression from MCI to AD dementia. Cross-validated results demonstrate that classifiers trained with our composite scores consistently outperform classifiers trained with the ADAS-Cog and multiple standard cognitive measures in addition to the ADAS-Cog, such as the MMSE, MoCA, and the AVLT.Prediction with domain scores also outperforms prediction with other cognitive composite scores in the literature [115, 156, 88]. When the domain scores are accompanied by the Clinical Dementia Rating (CDR) and the Functional Activities Questionnaire (FAQ), prediction performance is slightly better than the prediction with a set of 22 selected neuropsychological features [140].

Considering that psychiatric conditions may play an important role in the development of cognitive impairment, we tested if the addition of psychiatric information improved the performance of progression prediction. Classification experiments adding the Geriatric Depression Scale (GDS) and the abbreviated version of the Neuropsychiatric Inventory (NPI-Q) to the composite domain scores result in a very modest improvement of AUC values (Comparative results shown in Figure 4.7). It suggests psychiatric symptoms give little additional information that could be used to distinguish between MCI patients that will or will not progress to dementia.

Limitations

One important limitation of this study is that only data from ADNI was used, so generalization to other samples of population was not tested. The main reason for this is that the proposed methodology needs the sub-scores from neuropsychological tests and information with this level of detail is not available in other public databases. Survival analysis and progression prediction were based on data labels provided by ADNI, however recent studies have highlighted some flaws of the MCI diagnosis in ADNI database. First, it relies on a single test to evaluate memory leading to a high number of false
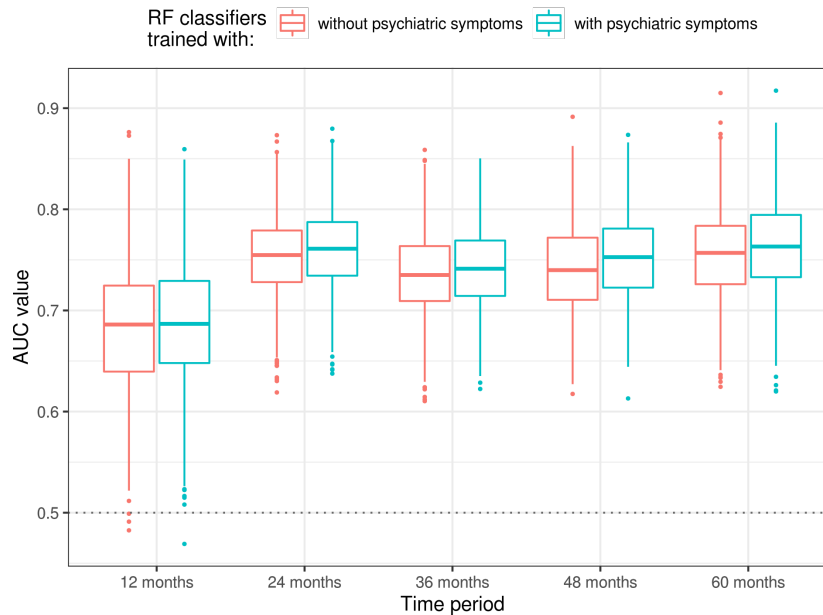
Figure 4.7: Distribution of AUC values for MCI conversion prediction within 12, 24, 36, 48 and 60 months. Classifiers were trained with domain scores, age, sex, years of education, with and without assessments of psychiatric symptoms.

positives [55]. Secondly, the MCI diagnostic criteria was not applied consistently after the first visit [192] and around 35% of subjects considered as stable MCI after a year did not meet all criteria so the continuation of MCI diagnosis appeared to be driven only by the CDR score.

Conclusions

The presented set of composite scores leads to a quantitative characterization of cognitive state for MCI patients. The presented results demonstrate that, relying only in the neuropsychological assessment, these composite domain scores are useful to stratify MCI patients and predict their future progression to dementia. Therefore, those scores could be easily included for patient monitoring or clinical trials. Future work should include longitudinal evaluation of domain dysfunction, along with AD biomarkers, that could improve understanding of the continuum between MCI and AD dementia.

4.5 Products

Journal paper

- **Diana L. Giraldo**, Jan Sijbers, Eduardo Romero. *Quantification of cognitive impairment to characterize heterogeneity of patients at risk of developing Alzheimer's disease*

dementia. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2021; 13(1):e12237. `https://doi.org/10.1002/dad2.12237` [73]

Conference papers

- **Diana L. Giraldo**, Jan Sijbers, Eduardo Romero. *Quantifying cognition and behavior in normal aging, mild cognitive impairment, and Alzheimer's disease.* Proc. 13th International Conference on Medical Information Processing and Analysis. San Andrés - Colombia, 2017. `https://doi.org/10.1117/12.2287036`

- German A. Pabón, **Diana L. Giraldo**, Eduardo Romero. *Mining relations between neuropsychological data to characterize Alzheimer's disease.* Accepted to the joint conference: 17th International Symposium on Medical Information Processing and Analysis (SIPAIM) - 10th Symposium on Medical Instrumentation and Imaging (SIIM). To be held in November 2021.

All methods and analysis in this section were implemented in R (version 3.6.3), code for processing ADNI data, reproducing the reported results, and calculate composite scores in new data is available in `https://github.com/diagiraldo/neuropsycho_adni` .

Chapter 5

Conclusions

This thesis has presented three strategies that address relevant needs in AD research. In Chapters 2 and 3 we present two contributions in the field of computational anatomy using different modalities of magnetic resonance imaging:

- We introduced a method to quantitatively describe regional anatomy extracting grey scale intensity information from T_1 weighted MRI and used this description in the automatic classification of whole-brain images. Previous works have included information from a predefined set of anatomical regions or have used whole-brain information leading to high-dimensional features that have no direct interpretation in terms of disease progression. The approach we presented falls in between these two kinds of analysis by quantifying changes in a set of anatomical regions covering the whole brain cortex and subcortical structures. The proposed metric quantifies how much regional tissue constituency is drifting away from what is considered normal, resembling the way clinicians evaluate anatomy with structural MRI but expanding this evaluation to several brain regions. This quantitative description of multiple brain areas exposes multi-dimensional patterns of AD progression that could be used to describe or evaluate anatomical changes along the AD continuum in clinical scenarios.

  The use of ensemble classifiers allows the assessment of how much additional information each region gives for the classifier to decide whether a case is a control or a patient. Although the presented methodology explored a set of anatomical regions covering the whole brain without giving preference to certain areas, the resulting set of most informative regions agrees with the widely reported changes in the temporal lobe.

- We presented a comprehensive analysis of micro- and macrostructural differences between groups using multi-shell DW-MRI data. Although several works have investigated the microstructural differences between AD or MCI patients and controls, most of them have used the diffusion tensor model to capture the underlying tissue properties, suffering from the known limitations of this model. The analysis pipeline we presented integrates: i) advanced models for the diffusion signal that represent crossing fibre configurations in WM and effectively separate tissue diffusivity properties for different tissue types, ii) the fixel-based analysis framework

to investigate changes in specific fibre pathways along with the voxel-based analysis to investigate tissue composition and volumetric changes, and iii) appropriate statistical inference methods that support the robustness of the results.

Results of the analyses revealed that patients with MCI and dementia due to AD exhibit degeneration of microstructural diffusion barriers in both white and grey matter in several brain areas such as the splenium of the corpus callosum, the cingulum and cingulate cortex, the insular cortex, and in the temporal lobe including its cortex, withe matter connections and subcortical structures. Volumetric changes, indicating macrostructural atrophy, were detected in temporal and parietal areas suggesting tissue degeneration might be more advanced than the one observed in superior and frontal areas. In addition to the investigation of group differences, we also applied the analysis pipeline to the exploration of linear relations between CSF biomarkers and diffusion-derived measures of micro- and macrostructural changes, finding significant correlations between CSF $A\beta_{1-42}$ levels and GM degeneration in the left hippocampus and expansion of the frontal horn of the lateral ventricles.

The presented methodology is a holistic neuroimaging approach that can be used to test linear hypotheses about tissue constituency and morphology. Beyond the study of group differences, it can be employed to interrogate correlations with disease quantitative markers, being also a methodological contribution to the investigation of AD-related neurodegeneration processes. The proposed approach relies on an advanced imaging acquisition technique (high angular resolution, multi-shell DW-MRI), which is not widely available out of research contexts, thus the potential implementation of presented diffusion-derived measures of tissue integrity in clinical scenarios is very unlikely. However, they could be used to evaluate the effect of potential disease-modifying treatments in preventing, slowing down, or even reversing microstructural degeneration of diffusion barriers.

Additionally to the contributions in computational anatomy, in Chapter 4 we presented a data-driven strategy aiming to improve the quantitative assessment of cognitive abilities in MCI patients:

- We developed a methodology to calculate a set of composite scores that quantify the level of impairment in six different cognitive domains: memory, language, visuospatial abilities, executive functioning, orientation and attention. These composite scores were obtained by combining and weighting sub-scores extracted from commonly used neuropsychological tests. This strategy incorporates the advantages of composite scores, *e.g.* robustness and sensitivity, with a domain specificity that facilitates the study of cognitive impairment heterogeneity. The proposed composite scores demonstrated to be useful for finding subgroups of MCI patients with different risks of progression to dementia and were able to better predict progression than standard outcomes. These results supports the idea that assessing domain-specific impairment could help to delineate cognitive profiles linked with differences in the clinical evolution.

  Domain-specific composite scores calculation could be easily included in the routine neuropsychological evaluation, giving useful information about the cognitive progression pattern and risk of progression to dementia within certain time. Furthermore, these scores could give more precise measures of the effects of therapeutic interventions designed to alleviate the cognitive consequences of AD.

In summary, this thesis presented a set of computational strategies with a common aim, the identification and quantification of pathological changes associated with AD. These contributions use information from neuroimages and cognitive evaluation to characterize patterns of disease progression. Assessment of pathological brain patterns with quantitative tools, like the ones herein developed, help patients' clinical management and monitoring and could improve the evaluation of potential, and urgently needed, disease-modifying treatments.

Perspectives

The contributions of this thesis have great potential for the study of AD-related pathological processes. There is still some work that could be done to validate each one of those strategies and evaluate how they compare to traditional methods in AD research in terms of providing better insights about AD progression patterns and patients' evolution along the AD continuum.

The first methodological contribution, the description of regional changes with distances between image intensity histograms, could be compared with traditional descriptions of regional anatomy: volume and cortical thickness. Such comparison could evaluate if the proposed description discriminates better between groups of subjects or if it is more sensitive to subtle longitudinal changes in brain anatomy. Regarding the second contribution, the comparison of tissue diffusivity properties between groups, the reported findings could be compared with results of investigating traditional diffusion tensor metrics such as FA and MD. Moreover, it would be interesting to examine how the diffusion-derived maps of tissue-like content relate to the tissue "concentration" maps used for VBM with structural images. The domain composite scores, presented in the third contribution, could be applied to longitudinal data to test how sensitive are these scores to longitudinal changes and explore the progression trajectories of different profiles of cognitive impairment.

A straightforward next step would be to integrate the computational anatomy descriptors with the domain-specific scores to explore the relationship between profiles of cognitive impairment and different neurodegeneration pathways. Investigation of the relationship between regional anatomical differences and levels of compromise per domain is possible with the available data in the Alzheimer's Disease Neuroimaging Initiative (ADNI). However, investigation of the relations between cognitive impairment and degeneration of diffusion barriers with the proposed strategies would require multi-shell DW-MRI paired with information from neuropsychological tests with an adequate level of granularity. To the best of our knowledge, only very few cases from ADNI satisfy those conditions.

Bibliography

[1] 2020 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 16(3):391–460, 2020.

[2] Julio Acosta-Cabronero and Peter J. Nestor. Diffusion tensor imaging in Alzheimer's disease: insights into the limbic-diencephalic network and methodological considerations. *Frontiers in Aging Neuroscience*, 6:266, 2014.

[3] Julio Acosta-Cabronero, Guy B. Williams, George Pengas, and Peter J. Nestor. Absolute diffusivities define the landscape of white matter degeneration in alzheimer's disease. *Brain*, 133(2):529–539, 2010.

[4] Paul S. Aisen, Jeffrey Cummings, Clifford R. Jack, John C. Morris, Reisa Sperling, Lutz Frölich, Roy W. Jones, Sherie A. Dowsett, Brandy R. Matthews, Joel Raskin, Philip Scheltens, and Bruno Dubois. On the path to 2025: understanding the Alzheimer's disease continuum. *Alzheimer's Research & Therapy*, 9(1), August 2017.

[5] Marilyn S. Albert, Steven T. DeKosky, Dennis Dickson, Bruno Dubois, Howard H. Feldman, Nick C. Fox, Anthony Gamst, David M. Holtzman, William J. Jagust, Ronald C. Petersen, Peter J. Snyder, Maria C. Carrillo, Bill Thies, and Creighton H. Phelps. 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 & Dementia*, 7(3):270 – 279, 2011.

[6] Bianca A.V. Alberton, Thomas E. Nichols, Humberto R. Gamba, and Anderson M. Winkler. Multiple testing correction over contrasts for brain imaging. *NeuroImage*, 216:116760, 2020.

[7] Kylie H. Alm and Arnold Bakker. Relationships between diffusion tensor imaging and cerebrospinal fluid metrics in early stages of the Alzheimer's disease continuum. *Journal of Alzheimer's Disease*, 70(4):965–981, August 2019.

[8] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Association, Washington, DC, 5th ed. edition, 2013.

[9] I.K. Amlien and A.M. Fjell. Diffusion tensor imaging of white matter degeneration in Alzheimer's disease and mild cognitive impairment. *Neuroscience*, 276(Supplement C):206 – 215, 2014. Secrets of the CNS White Matter.

[10] Jesper L. R Andersson, Mark S. Graham, Enikő Zsoldos, and Stamatios N. Sotiropoulos. Incorporating outlier detection and replacement into a nonparametric framework for movement and distortion correction of diffusion MR images. *NeuroImage*, 141:556 – 572, 2016.

[11] Jesper L.R. Andersson and Stamatios N. Sotiropoulos. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, 125:1063–1078, 2016.

[12] Babak A. Ardekani, Antonio Convit, and Alvin H. Bachman. Analysis of the miriad data shows sex differences in hippocampal atrophy progression. *Journal of Alzheimer's Disease*, 50(3):847–857, Feb 2016.

[13] John Ashburner and Karl J. Friston. Voxel-based morphometry—The methods. *NeuroImage*, 11(6):805 – 821, 2000.

[14] Ravi Bansal, Lawrence H. Staib, Andrew F. Laine, Xuejun Hao, Dongrong Xu, Jun Liu, Myrna Weissman, and Bradley S. Peterson. Anatomical brain images alone can accurately diagnose chronic neuropsychiatric illnesses. *PLoS ONE*, 7(12):1–21, 12 2012.

[15] M. S. Bartlett. The statistical conception of mental factors. *British Journal of Psychology. General Section*, 28(1):97–104, jul 1937.

[16] Peter J. Basser and Carlo Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance, Series B*, 111(3):209–219, 1996.

[17] Iman Beheshti and Hasan Demirel. Feature-ranking-based Alzheimer's disease classification from structural MRI. *Magnetic Resonance Imaging*, 34(3):252 – 263, 2016.

[18] Mark W. Bondi, Emily C. Edmonds, Amy J. Jak, Lindsay R. Clark, Lisa Delano-Wood, Carrie R. McDonald, Daniel A. Nation, David J. Libon, Rhoda Au, Douglas Galasko, and David P. Salmon. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *Journal of Alzheimer's Disease*, 42(1):275–289, 2014.

[19] H. Braak and E. Braak. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, 82(4):239–259, September 1991.

[20] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001.

[21] Esther E. Bron, Marion Smits, Wiesje M. van der Flier, Hugo Vrenken, Frederik Barkhof, Philip Scheltens, Janne M. Papma, Rebecca M.E. Steketee, Carolina Méndez Orellana, Rozanna Meijboom, Madalena Pinto, Joana R. Meireles, Carolina Garrett, António J. Bastos-Leite, Ahmed Abdulkadir, Olaf Ronneberger, Nicola Amoroso, Roberto Bellotti, David Cárdenas-Peña, Andrés M. Álvarez Meza, Chester V. Dolph, Khan M. Iftekharuddin, Simon F. Eskildsen, Pierrick Coupé, Vladimir S. Fonov, Katja Franke, Christian Gaser, Christian Ledig, Ricardo Guerrero, Tong Tong, Katherine R. Gray, Elaheh Moradi, Jussi Tohka, Alexandre Routier, Stanley Durrleman, Alessia Sarica, Giuseppe Di Fatta, Francesco Sensi, Andrea Chincarini, Garry M. Smith, Zhivko V. Stoyanov, Lauge Sørensen, Mads Nielsen, Sabina Tangaro, Paolo Inglese, Christian Wachinger, Martin Reuter, John C. van Swieten, Wiro J. Niessen, and Stefan Klein. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *NeuroImage*, 111:562 – 579, 2015.

[22] Fernando Calamante, Robert E. Smith, Jacques-Donald Tournier, David Raffelt, and Alan Connelly. Quantification of voxel-wise total fibre density: Investigating the problems associated with track-count mapping. *NeuroImage*, 117:284–293, 2015.

[23] Marco Canevelli, Ilaria Bacigalupo, Giuseppe Gervasi, Eleonora Lacorte, Marco Massari, Flavia Mayer, Nicola Vanacore, and Matteo Cesari. Methodological issues in the clinical validation of biomarkers for Alzheimer's disease: The paradigmatic example of CSF. *Frontiers in Aging Neuroscience,*, 11:282, 2019.

[24] Francesca Caso, Federica Agosta, and Massimo Filippi. Insights into white matter damage in Alzheimer's disease: From postmortem to in vivo diffusion tensor MRI studies. *Neurodegenerative Diseases*, 16(1-2):26–33, dec 2015.

[25] Malika Charrad, Nadia Ghazzali, Véronique Boiteau, and Azam Niknafs. NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, Articles*, 61(6):1–36, 2014.

[26] Marie Chupin, Emilie Gérardin, Rémi Cuingnet, Claire Boutet, Louis Lemieux, Stéphane Lehéricy, Habib Benali, Line Garnero, and Olivier Colliot. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus*, 19(6):579–587, 2009.

[27] Simon Cloutier, Howard Chertkow, Marie-Jeanne Kergoat, Serge Gauthier, and Sylvie Belleville. Patterns of cognitive decline prior to dementia in persons with mild cognitive impairment. *Journal of Alzheimer's Disease*, 47(4):901–913, August 2015.

[28] Lyndsey E Collins-Praino, Yitshak I Francis, Erica Y Griffith, Anne F Wiegman, Jonathan Urbach, Arlene Lawton, Lawrence S Honig, Etty Cortes, Jean Paul G Vonsattel, Peter D Canoll, James E Goldman, and Adam M Brickman. Soluble amyloid beta levels are elevated in the white matter of Alzheimer's patients, independent of cortical plaque severity. *Acta Neuropathologica Communications*, 2(1), August 2014.

[29] Olivier Colliot, Gaël Chételat, Marie Chupin, Béatrice Desgranges, Benoît Magnin, Habib Benali, Bruno Dubois, Line Garnero, Francis Eustache, and Stéphane Lehéricy. Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. *Radiology*, 248, 07 2008.

[30] Paul K. Crane, Adam Carle, Laura E. Gibbons, Philip Insel, R. Scott Mackin, Alden Gross, Richard N. Jones, Shubhabrata Mukherjee, S. McKay Curtis, Danielle Harvey, Michael Weiner, Dan Mungas, and for the Alzheimer's Disease Neuroimaging Initiative. Development and assessment of a composite score for memory in the Alzheimer's disease neuroimaging initiative (ADNI). *Brain Imaging and Behavior*, 6(4):502–516, Dec 2012.

[31] Robert Cudeck. Exploratory factor analysis. In Howard E.A. Tinsley and Steven D. Brown, editors, *Handbook of Applied Multivariate Statistics and Mathematical Modeling*, pages 265 – 296. Academic Press, San Diego, 2000.

[32] Rémi Cuingnet, Emilie Gerardin, Jérôme Tessieras, Guillaume Auzias, Stéphane Lehéricy, Marie-Odile Habert, Marie Chupin, Habib Benali, and Olivier Colliot.

Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the adni database. *NeuroImage*, 56(2):766 – 781, 2011. Multivariate Decoding and Brain Reading.

[33] Silvio Ramos Bernardes da Silva Filho, Jeam Haroldo Oliveira Barbosa, Carlo Rondinoni, Antonio Carlos dos Santos, Carlos Ernesto Garrido Salmon, Nereida Kilza da Costa Lima, Eduardo Ferriolli, and Júlio César Moriguti. Neurodegeneration profile of Alzheimer's patients: A brain morphometry study. *NeuroImage: Clinical*, 15:15 – 24, 2017.

[34] Christos Davatzikos, Priyanka Bhatt, Leslie M. Shaw, Kayhan N. Batmanghelich, and John Q. Trojanowski. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging*, 32(12):2322.e19 – 2322.e27, 2011.

[35] Christos Davatzikos, Yong Fan, Xiaoying Wu, Dinggang Shen, and Susan M. Resnick. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. *Neurobiology of Aging*, 29(4):514 – 523, 2008.

[36] M.S. de Oliveira, M.L.F. Balthazar, A. D'Abreu, C.L. Yasuda, B.P. Damasceno, F. Cendes, and G. Castellano. MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease. *American Journal of Neuroradiology*, 32(1):60–66, 2011.

[37] Michaela Defrancesco, Karl Egger, Josef Marksteiner, Regina Esterhammer, Hartmann Hinterhuber, Eberhard A. Deisenhammer, and Michael Schocke. Changes in white matter integrity before conversion from mild cognitive impairment to alzheimer?s disease. *PLoS ONE*, 9(8):1–10, 08 2014.

[38] Flavio Dell'Acqua and J.-Donald Tournier. Modelling white matter with spherical deconvolution: How and why? *NMR in Biomedicine*, 32(4):e3945, 2019. e3945 NBM-17-0226.R1.

[39] R. S. Desikan, H. J. Cabral, C. P. Hess, W. P. Dillon, C. M. Glastonbury, M. W. Weiner, N. J. Schmansky, D. N. Greve, D. H. Salat, and R. L. et al. Buckner. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. *Brain*, 132(8):2048–2057, 2009.

[40] Thijs Dhollander, Remika Mito, David Raffelt, and Alan Connelly. Improved white matter response function estimation for 3-tissue constrained spherical deconvolution. In *International Society of Magnetic Resonance in Medicine*, 05 2019.

[41] Thijs Dhollander, David Raffelt, and Alan Connelly. Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image. In *ISMRM Workshop on Breaking the Barriers of Diffusion MRI*, Lisbon, Portugal, 09 2016.

[42] Thijs Dhollander, Rami Tabbara, Jonas Rosnarho-Tornstrand, Jacques-Donald Tournier, David Raffelt, and Alan Connelly. Multi-tissue log-domain intensity and inhomogeneity normalization for quantitative apparent fibre density. In *International Society of Magnetic Resonance in Medicine*, 2021.

[43] B.C. Dickerson, E. Fenstermacher, D.H. Salat, D.A. Wolk, R.P. Maguire, R. Desikan, J. Pacheco, B.T. Quinn, A. Van der Kouwe, D.N. Greve, D. Blacker, M.S. Albert, R.J. Killiany, and B. Fischl. Detection of cortical thickness correlates of cognitive performance: Reliability across MRI scan sessions, scanners, and field strengths. *NeuroImage*, 39(1):10 – 18, 2008.

[44] B.C. Dickerson, I. Goncharova, M.P. Sullivan, C. Forchetti, R.S. Wilson, D.A. Bennett, L.A. Beckett, and L. deToledo Morrell. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. *Neurobiology of Aging*, 22(5):747 – 754, 2001.

[45] Nhat Trung Doan, Andreas Engvig, Karin Persson, Dag Alnæs, Tobias Kaufmann, Jaroslav Rokicki, Aldo Córdova-Palomera, Torgeir Moberget, Anne Brækhus, Maria Lage Barca, Knut Engedal, Ole A. Andreassen, Geir Selbæk, and Lars T. Westlye. Dissociable diffusion MRI patterns of white matter microstructure and connectivity in Alzheimer's disease spectrum. *Scientific Reports*, 7:45131, mar 2017.

[46] Aoyan Dong, Jon B. Toledo, Nicolas Honnorat, Jimit Doshi, Erdem Varol, Aristeidis Sotiras, David Wolk, John Q. Trojanowski, and Christos Davatzikos and. Heterogeneity of neuroanatomical patterns in prodromal Alzheimer's disease: links to cognition, progression and biomarkers. *Brain*, page aww319, dec 2017.

[47] Gwenaëlle Douaud, Saâd Jbabdi, Timothy E.J. Behrens, Ricarda A. Menke, Achim Gass, Andreas U. Monsch, Anil Rao, Brandon Whitcher, Gordon Kindlmann, Paul M. Matthews, and Stephen Smith. DTI measures in crossing-fibre areas: Increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *NeuroImage*, 55(3):880 – 890, 2011.

[48] Bruno Dubois, Howard H Feldman, Claudia Jacova, Jeffrey L Cummings, Steven T DeKosky, Pascale Barberger-Gateau, André Delacourte, Giovanni Frisoni, Nick C Fox, Douglas Galasko, Serge Gauthier, Harald Hampel, Gregory A Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Marie Sarazin, Leonardo C de Souza, Yaakov Stern, Pieter J Visser, and Philip Scheltens. Revising the definition of Alzheimer's disease: a new lexicon. *The Lancet Neurology*, 9(11):1118 – 1127, 2010.

[49] Bruno Dubois, Howard H Feldman, Claudia Jacova, Steven T DeKosky, Pascale Barberger-Gateau, Jeffrey Cummings, André Delacourte, Douglas Galasko, Serge Gauthier, Gregory Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Yaakov Stern, Pieter J Visser, and Philip Scheltens. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *The Lancet Neurology*, 6(8):734 – 746, 2007.

[50] Bruno Dubois, Howard H Feldman, Claudia Jacova, Harald Hampel, José Luis Molinuevo, Kaj Blennow, Steven T DeKosky, Serge Gauthier, Dennis Selkoe, Randall Bateman, Stefano Cappa, Sebastian Crutch, Sebastiaan Engelborghs, Giovanni B Frisoni, Nick C Fox, Douglas Galasko, Marie-Odile Habert, Gregory A Jicha, Agneta Nordberg, Florence Pasquier, Gil Rabinovici, Philippe Robert, Christopher Rowe, Stephen Salloway, Marie Sarazin, Stéphane Epelbaum, Leonardo C de Souza, Bruno Vellas, Pieter J Visser, Lon Schneider, Yaakov Stern, Philip Scheltens, and Jeffrey L Cummings. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *The Lancet Neurology*, 13(6):614 – 629, 2014.

[51] Bruno Dubois, Harald Hampel, Howard H. Feldman, Philip Scheltens, Paul Aisen, Sandrine Andrieu, Hovagim Bakardjian, Habib Benali, Lars Bertram, Kaj Blennow, Karl Broich, Enrica Cavedo, Sebastian Crutch, Jean-François Dartigues, Charles Duyckaerts, Stéphane Epelbaum, Giovanni B. Frisoni, Serge Gauthier, Remy Genthon, Alida A. Gouw, Marie-Odile Habert, David M. Holtzman, Miia Kivipelto, Simone Lista, José-Luis Molinuevo, Sid E. O'Bryant, Gil D. Rabinovici, Christopher Rowe, Stephen Salloway, Lon S. Schneider, Reisa Sperling, Marc Teichmann, Maria C. Carrillo, Jeffrey Cummings, and Cliff R. Jack. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimer's & Dementia*, 12(3):292–323, mar 2016.

[52] Matthieu Dumont, Maggie Roy, Pierre-Marc Jodoin, Felix C. Morency, Jean-Christophe Houde, Zhiyong Xie, Cici Bauer, Tarek A. Samad, Koene R. A. Van Dijk, James A. Goodman, Maxime Descoteaux, and Alzheimer's Disease Neuroimaging Initiative . Free water in white matter differentiates MCI and AD from control subjects. *Frontiers in Aging Neuroscience*, 11:270, 2019.

[53] Tim B. Dyrby, Henrik Lundell, Mark W. Burke, Nina L. Reislev, Olaf B. Paulson, Maurice Ptito, and Hartwig R. Siebner. Interpolation of diffusion weighted imaging datasets. *NeuroImage*, 103:202–213, 2014.

[54] Emily C. Edmonds, Lisa Delano-Wood, Lindsay R. Clark, Amy J. Jak, Daniel A. Nation, Carrie R. McDonald, David J. Libon, Rhoda Au, Douglas Galasko, David P. Salmon, and Mark W. Bondi. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimer's & Dementia*, 11(4):415 – 424, 2015.

[55] Emily C. Edmonds, Joel Eppig, Mark W. Bondi, Kelly M. Leyden, Bailey Goodwin, Lisa Delano-Wood, and Carrie R. McDonald. Heterogeneous cortical atrophy patterns in MCI not captured by conventional diagnostic criteria. *Neurology*, 87(20):2108–2116, 2016.

[56] Emily C. Edmonds, Carrie R. McDonald, Anisa Marshall, Kelsey R. Thomas, Joel Eppig, Alexandra J. Weigand, Lisa Delano-Wood, Douglas R. Galasko, David P. Salmon, and Mark W. Bondi. Early versus late MCI: Improved MCI staging using a neuropsychological approach. *Alzheimer's & Dementia*, 15(5):699 – 708, 2019.

[57] Juan José Egozcue, Vera Pawlowsky-Glahn, Glòria Mateu-Figueras, and Carles Barcelo-Vidal. Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3):279–300, 2003.

[58] Joel S. Eppig, Emily C. Edmonds, Laura Campbell, Mark Sanderson-Cimino, Lisa Delano-Wood, and Mark W. Bondi. Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: A latent profile analysis. *Journal of the International Neuropsychological Society*, 23(7):564–576, 2017.

[59] Yong Fan, Susan M. Resnick, Xiaoying Wu, and Christos Davatzikos. Structural and functional biomarkers of prodromal Alzheimer's disease: A high-dimensional pattern classification study. *NeuroImage*, 41(2):277 – 285, 2008.

[60] Marshal F. Folstein, Susan E. Folstein, and Paul R. McHugh. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 1975.

[61] Mark A. Fraser, Marnie E. Shaw, and Nicolas Cherbuin. A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *NeuroImage*, 112(Supplement C):364 – 374, 2015.

[62] P. A. Freeborough and N. C. Fox. MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease. *IEEE Transactions on Medical Imaging*, 17(3):475–478, June 1998.

[63] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[64] Giovanni B. Frisoni, Marina Boccardi, Frederik Barkhof, Kaj Blennow, Stefano Cappa, Konstantinos Chiotis, Jean-Francois Démonet, Valentina Garibotto, Panteleimon Giannakopoulos, Anton Gietl, Oskar Hansson, Karl Herholz, Clifford R. Jr Jack, Flavio Nobili, Agneta Nordberg, Heather M. Snyder, Mara Ten Kate, Andrea Varrone, Emiliano Albanese, Stefanie Becker, Patrick Bossuyt, Maria C. Carrillo, Chiara Cerami, Bruno Dubois, Valentina Gallo, Ezio Giacobini, Gabriel Gold, Samia Hurst, Anders Lönneborg, Karl-Olof Lovblad, Niklas Mattsson, José-Luis Molinuevo, Andreas U. Monsch, Urs Mosimann, Alessandro Padovani, Agnese Picco, Corinna Porteri, Osman Ratib, Laure Saint-Aubert, Charles Scerri, Philip Scheltens, Jonathan M. Schott, Ida Sonni, Stefan Teipel, Paolo Vineis, Pieter Jelle Visser, Yutaka Yasui, and Bengt Winblad. Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *The Lancet Neurology*, 16(8):661–676, 07 2017.

[65] Giovanni B. Frisoni, Nick C. Fox, Clifford R. Jack, Philip Scheltens, and Paul M. Thompson. The clinical use of structural MRI in Alzheimer disease. *Nature Reviews Neurology*, 6(2):67–77, 2010.

[66] Leticia Garcia-Alvarez, Jesus J. Gomar, Amber Sousa, Maria P. Garcia-Portilla, and Terry E. Goldberg. Breadth and depth of working memory and executive function compromises in mild cognitive impairment and their relationships to frontal lobe morphometry and functional competence. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 11:170–179, December 2019.

[67] Serge Gauthier, Marilyn Albert, Nick Fox, Michel Goedert, Miia Kivipelto, Jorge Mestre-Ferrandiz, and Lefkos T. Middleton. Why has therapy development for dementia failed in the last two decades? *Alzheimer's & Dementia*, 12(1):60–64, 2016.

[68] Emilie Gerardin, Gaël Chételat, Marie Chupin, Rémi Cuingnet, Béatrice Desgranges, Ho-Sung Kim, Marc Niethammer, Bruno Dubois, Stéphane Lehéricy, Line Garnero, Francis Eustache, and Olivier Colliot. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *NeuroImage*, 47(4):1476 – 1486, 2009.

[69] Laura E. Gibbons, Adam C. Carle, R. Scott Mackin, Danielle Harvey, Shubhabrata Mukherjee, Philip Insel, S. McKay Curtis, Dan Mungas, and Paul K. Crane. A composite score for executive functioning, validated in Alzheimer's disease neuroimaging initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging and Behavior*, 6(4):517–527, 2012.

[70] Diana Giraldo, Hanne Struyfs, David Raffelt, Paul Parizel, Sebastiaan Engel-borghs, Eduardo Romero, Jan Sijbers, and Ben Jeurissen. Fixel-based analysis of Alzheimer's disease using multi-tissue constrained spherical deconvolution of multi-shell diffusion MRI. In *International Society of Magnetic Resonance in Medicine*, Honolulu, HI, USA, 2017.

[71] Diana L. Giraldo, Juan D. García-Arteaga, Simón Cárdenas-Robledo, and Eduardo Romero. Characterization of brain anatomical patterns by comparing region intensity distributions: Applications to the description of Alzheimer's disease. *Brain and Behavior*, 8(4):e00942, 2018.

[72] Diana L. Giraldo, Jan Sijbers, and Eduardo Romero. Quantifying cognition and behavior in normal aging, mild cognitive impairment, and Alzheimer's disease. In *13th International Conference on Medical Information Processing and Analysis*, volume 10572, pages 125 – 134, 2017.

[73] Diana L. Giraldo, Jan Sijbers, Eduardo Romero, and for the Alzheimer's Disease Neuroimaging Initiative. Quantification of cognitive impairment to characterize heterogeneity of patients at risk of developing Alzheimer's disease dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 13(1):e12237, 2021.

[74] Brian T. Gold, Nathan F. Johnson, David K. Powell, and Charles D. Smith. White matter integrity and vulnerability to Alzheimer's disease: Preliminary findings and future directions. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1822(3):416 – 422, 2012. Imaging Brain Aging and Neurodegenerative Disease.

[75] Tamar H. Gollan, Gali H. Weissberger, Elin Runnqvist, Rosa I. Montoya, and Cynthia M. Cera. Self-ratings of spoken language dominance: A multilingual naming test (MINT) and preliminary norms for young and aging spanish–english bilinguals. *Bilingualism: Language and Cognition*, 15(3):594–615, August 2011.

[76] Teresa Gomez-Isla and Matthew P. Frosch. The challenge of defining Alzheimer disease based on biomarkers in the absence of symptoms. *JAMA Neurology*, 76(10):1143–1144, 10 2019.

[77] H. Goodglass and E. Kaplan. *The Assessment of Aphasia and Related Disorders*. Number v. 2 in The Assessment of Aphasia and Related Disorders. Lea & Febiger, 1983.

[78] Devendra Goyal, Donna Tjandra, Raymond Q. Migrino, Bruno Giordani, Zeeshan Syed, and Jenna Wiens. Characterizing heterogeneity in the progression of Alzheimer's disease using longitudinal clinical and neuroimaging biomarkers. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10:629 – 637, 2018.

[79] James W. Grice. Computing and evaluating factor scores. *Psychological Methods*, 6, 2001.

[80] Joseph H. Grochowalski, Ying Liu, and Karen L. Siedlecki. Examining the reliability of ADAS-Cog change scores. *Aging, Neuropsychology, and Cognition*, 23(5):513–529, 2016. PMID: 26708116.

[81] Murray Grossman, Corey McMillan, Peachie Moore, Lijun Ding, Guila Glosser, Melissa Work, and James Gee. What's in a name: voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain*, 127(3):628–649, 2004.

[82] Robert M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, November 1973.

[83] John A. Hardy and Gerald A. Higgins. Alzheimer's disease: The amyloid cascade hypothesis. *Science*, 256(5054):184–185, 1992.

[84] Rolf A. Heckemann, Shiva Keihaninejad, Paul Aljabar, Katherine R. Gray, Casper Nielsen, Daniel Rueckert, Joseph V. Hajnal, and Alexander Hammers. Automatic morphometry in Alzheimer's disease and mild cognitive impairment. *NeuroImage*, 56(4):2024 – 2037, 2011.

[85] Judith Henf, Michel J. Grothe, Katharina Brueggen, Stefan Teipel, and Martin Dyrba. Mean diffusivity in cortical gray matter in Alzheimer's disease: The importance of partial volume correction. *NeuroImage: Clinical*, 17(Supplement C):579 – 586, 2018.

[86] Karl Herrup. The case for rejecting the amyloid cascade hypothesis. *Nature Neuroscience*, 18(6):794–799, May 2015.

[87] Andrew R. Hoy, Martina Ly, Cynthia M. Carlsson, Ozioma C. Okonkwo, Henrik Zetterberg, Kaj Blennow, Mark A. Sager, Sanjay Asthana, Sterling C. Johnson, Andrew L. Alexander, and Barbara B. Bendlin. Microstructural white matter alterations in preclinical Alzheimer's disease detected using free water elimination diffusion tensor imaging. *PLOS ONE*, 12(3):1–21, 03 2017.

[88] Yifan Huang, Kaori Ito, Clare B. Billing, and Richard J. Anziano. Development of a straightforward and sensitive scale for MCI and early AD clinical trials. *Alzheimer's & Dementia*, 11(4):404 – 414, 2015.

[89] Bradley T. Hyman, Creighton H. Phelps, Thomas G. Beach, Eileen H. Bigio, Nigel J. Cairns, Maria C. Carrillo, Dennis W. Dickson, Charles Duyckaerts, Matthew P. Frosch, Eliezer Masliah, Suzanne S. Mirra, Peter T. Nelson, Julie A. Schneider, Dietmar Rudolf Thal, Bill Thies, John Q. Trojanowski, Harry V. Vinters, and Thomas J. Montine. National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's & Dementia*, 8(1):1 – 13, 2012.

[90] Jeremy D Isaacs and Marianne Boenink. Biomarkers for dementia: too soon for routine clinical use. *The Lancet Neurology*, 19(11):884–885, November 2020.

[91] C. R. Jack, R. C. Petersen, P. C. O'Brien, and E. G. Tangalos. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology*, 42(1):183–183, 1992.

[92] Clifford R. Jack, David A. Bennett, Kaj Blennow, Maria C. Carrillo, Billy Dunn, Samantha Budd Haeberlein, David M. Holtzman, William Jagust, Frank Jessen, Jason Karlawish, Enchi Liu, Jose Luis Molinuevo, Thomas Montine, Creighton Phelps, Katherine P. Rankin, Christopher C. Rowe, Philip Scheltens, Eric Siemers,

Heather M. Snyder, Reisa Sperling, Cerise Elliott, Eliezer Masliah, Laurie Ryan, and Nina Silverberg. NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4):535–562, apr 2018.

[93] Clifford R Jack, David S Knopman, William J Jagust, Ronald C Petersen, Michael W Weiner, Paul S Aisen, Leslie M Shaw, Prashanthi Vemuri, Heather J Wiste, Stephen D Weigand, Timothy G Lesnick, Vernon S Pankratz, Michael C Donohue, and John Q Trojanowski. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, 12(2):207–216, feb 2013.

[94] Jr Jack, Clifford R., Terry M. Therneau, Stephen D. Weigand, Heather J. Wiste, David S. Knopman, Prashanthi Vemuri, Val J. Lowe, Michelle M. Mielke, Rosebud O. Roberts, Mary M. Machulda, Jonathan Graff-Radford, David T. Jones, Christopher G. Schwarz, Jeffrey L. Gunter, Matthew L. Senjem, Walter A. Rocca, and Ronald C. Petersen. Prevalence of Biologically vs Clinically Defined Alzheimer Spectrum Entities Using the National Institute on Aging–Alzheimer's Association Research Framework. *JAMA Neurology*, 76(10):1174–1183, 10 2019.

[95] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002.

[96] Mark Jenkinson, Christian F. Beckmann, Timothy E.J. Behrens, Mark W. Woolrich, and Stephen M. Smith. FSL. *NeuroImage*, 62(2):782–790, 2012. 20 YEARS OF fMRI.

[97] Jens H. Jensen, Joseph A. Helpern, Anita Ramani, Hanzhang Lu, and Kyle Kaczynski. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53(6):1432–1440, 2005.

[98] Ben Jeurissen, Alexander Leemans, Jacques-Donald Tournier, Derek K. Jones, and Jan Sijbers. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping*, 34(11):2747–2766, 2013.

[99] Ben Jeurissen, Jacques-Donald Tournier, Thijs Dhollander, Alan Connelly, and Jan Sijbers. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*, 103:411 – 426, 2014.

[100] David T. Jones, David S. Knopman, Jeffrey L. Gunter, Jonathan Graff-Radford, Prashanthi Vemuri, Bradley F. Boeve, Ronald C. Petersen, Michael W. Weiner, Clifford R. Jr. Jack, and on behalf of the Alzheimer's Disease Neuroimaging Initiative. Cascading network failure across the Alzheimer's disease spectrum. *Brain*, 139(2):547–562, 11 2015.

[101] K. Kantarci, R. C. Petersen, B. F. Boeve, D. S. Knopman, S. D. Weigand, P. C. O'Brien, M. M. Shiung, G. E. Smith, R. J. Ivnik, and E. G. et al. Tangalos. Dwi predicts future progression to alzheimer disease in amnestic mild cognitive impairment. *Neurology*, 64(5):902–904, 2005.

[102] E. Kaplan, H. Goodglass, and S. Weintraub. *The Boston Naming Test*. Lea and Febiger, 1989.

[103] Elias Kellner, Bibek Dhital, Valerij G. Kiselev, and Marco Reisert. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, 76(5):1574–1581, 2016.

[104] Kasper Planeta Kepp. Ten challenges of the amyloid hypothesis of Alzheimer's disease. *Journal of Alzheimer's Disease*, 55(2):447–457, November 2016.

[105] Stefan Klöppel, Cynthia M. Stonnington, Carlton Chu, Bogdan Draganski, Rachael I. Scahill, Jonathan D. Rohrer, Nick C. Fox, Clifford R. Jack, John Ashburner, and Richard S. J. Frackowiak. Automatic classification of MR scans in Alzheimer's disease. *Brain*, 131(3):681–689, 2008.

[106] Jacqueline K. Kueper, Mark Speechley, and Manuel Montero-Odasso. The Alzheimer's disease assessment scale–cognitive subscale (ADAS-Cog): Modifications and responsiveness in pre-dementia populations. A narrative review. *Journal of Alzheimer's Disease*, 63(2):423–444, April 2018.

[107] Florian Kurth, Christian Gaser, and Eileen Luders. A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nature Protocols*, 10(2):293–304, January 2015.

[108] Denis Le Bihan. Apparent diffusion coefficient and beyond: What diffusion mr imaging can tell us about tissue structure. *Radiology*, 268(2):318–322, 2013. PMID: 23882093.

[109] S.-H. Lee, J.-P. Coutu, P. Wilkens, A. Yendiki, H.D. Rosas, and D.H. Salat. Tract-based analysis of white matter degeneration in Alzheimer's disease. *Neuroscience*, 301:79 – 89, 2015.

[110] Fangda Leng and Paul Edison. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here? *Nature Reviews Neurology*, 17(3):157–172, December 2020.

[111] Peter J. Rousseeuw Leonard Kaufman. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1 edition, 2005.

[112] Jason P. Lerch, Jens Pruessner, Alex P. Zijdenbos, D. Louis Collins, Stefan J. Teipel, Harald Hampel, and Alan C. Evans. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. *Neurobiology of Aging*, 29(1):23 – 30, 2008.

[113] Sung-Han Lin, Wen-Chuin Hsu, Shu-Hang Ng, Jur-Shan Cheng, Oleksandr Khegai, Chin-Chang Huang, Yao-Liang Chen, Yi-Chun Chen, and Jiun-Jie Wang. Increased water diffusion in the parcellated cortical regions from the patients with amnestic mild cognitive impairment and alzheimer's disease. *Frontiers in Aging Neuroscience*, 8:325, 2017.

[114] Manhua Liu, Fan Li, Hao Yan, Kundong Wang, Yixin Ma, Li Shen, and Mingqing Xu. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease. *NeuroImage*, 208:116459, 2020.

[115] Daniel A. Llano, Genevieve Laforet, and Viswanath Devanarayan. Derivation of a new ADAS-cog composite using tree-based multivariate analysis. *Alzheimer Disease & Associated Disorders*, 25(1):73–84, 2011.

[116] David A. Loewenstein, Raymond M.D. Ownby, Lynn Schram, Amarilis Acevedo, Mark Rubert, and Trinidad Argelles. An evaluation of the NINCDS-ADRDA neuropsychological criteria for the assessment of Alzheimers disease: A confirmatory factor analysis of single versus multi-factor models. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, 23(3):274–284, jun 2001.

[117] Benoît Magnin, Lilia Mesrob, Serge Kinkingnéhun, Mélanie Pélégrini-Issac, Olivier Colliot, Marie Sarazin, Bruno Dubois, Stéphane Lehéricy, and Habib Benali. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology*, 51, 02 2009.

[118] Ian B. Malone, David Cash, Gerard R. Ridgway, David G. MacManus, Sebastien Ourselin, Nick C. Fox, and Jonathan M. Schott. Miriad—public release of a multiple time point Alzheimer's MR imaging dataset. *NeuroImage*, 70:33–36, 2013.

[119] Ian B. Malone, Kelvin K. Leung, Shona Clegg, Josephine Barnes, Jennifer L. Whitwell, John Ashburner, Nick C. Fox, and Gerard R. Ridgway. Accurate automatic estimation of total intracranial volume: A nuisance variable with less nuisance. *NeuroImage*, 104:366 – 372, 2015.

[120] Daniel S. Marcus, Tracy H. Wang, Jamie Parker, John G. Csernansky, John C. Morris, and Randy L. Buckner. Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9):1498–1507, 09 2007.

[121] Razvan V. Marinescu, Neil P. Oxtoby, Alexandra L. Young, Esther E. Bron, Arthur W. Toga, Michael W. Weiner, Frederik Barkhof, Nick C. Fox, Arman Eshaghi, Tina Toni, Marcin Salaterski, Veronika Lunina, Manon Ansart, Stanley Durrleman, Pascal Lu, Samuel Iddi, Dan Li, Wesley K. Thompson, Michael C. Donohue, Aviv Nahon, Yarden Levy, Dan Halbersberg, Mariya Cohen, Huiling Liao, Tengfei Li, Kaixian Yu, Hongtu Zhu, Jose G. Tamez-Pena, Aya Ismail, Timothy Wood, Hector Corrada Bravo, Minh Nguyen, Nanbo Sun, Jiashi Feng, B. T. Thomas Yeo, Gang Chen, Ke Qi, Shiyang Chen, Deqiang Qiu, Ionut Buciuman, Alex Kelner, Raluca Pop, Denisa Rimocea, Mostafa M. Ghazi, Mads Nielsen, Sebastien Ourselin, Lauge Sorensen, Vikram Venkatraghavan, Keli Liu, Christina Rabe, Paul Manser, Steven M. Hill, James Howlett, Zhiyue Huang, Steven Kiddle, Sach Mukherjee, Anais Rouanet, Bernd Taschler, Brian D. M. Tom, Simon R. White, Noel Faux, Suman Sedai, Javier de Velasco Oriol, Edgar E. V. Clemente, Karol Estrada, Leon Aksman, Andre Altmann, Cynthia M. Stonnington, Yalin Wang, Jianfeng Wu, Vivek Devadas, Clementine Fourrier, Lars Lau Raket, Aristeidis Sotiras, Guray Erus, Jimit Doshi, Christos Davatzikos, Jacob Vogel, Andrew Doyle, Angela Tam, Alex Diaz-Papkovich, Emmanuel Jammeh, Igor Koval, Paul Moore, Terry J. Lyons, John Gallacher, Jussi Tohka, Robert Ciszek, Bruno Jedynak, Kruti Pandya, Murat Bilgel, William Engels, Joseph Cole, Polina Golland, Stefan Klein, and Daniel C. Alexander. The Alzheimer's disease prediction of longitudinal evolution (TADPOLE) challenge: Results after 1 year follow-up, 2020.

[122] Chantel D. Mayo, Erin L. Mazerolle, Lesley Ritchie, John D. Fisk, and Jodie R. Gawryluk. Longitudinal changes in microstructural white matter metrics in Alzheimer's disease. *NeuroImage: Clinical*, 13(Supplement C):330 – 338, 2017.

[123] Linda K. McEvoy, Christine Fennema-Notestine, J. Cooper Roddey, Jr Donald J. Hagler, Dominic Holland, David S. Karow, Christopher J. Pung, James B. Brewer, and Anders M. Dale. Alzheimer disease: Quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology*, 251(1):195–205, 2009. PMID: 19201945.

[124] Guy McKhann, David Drachman, Marshall Folstein, Robert Katzman, Donald Price, and Emanuel M. Stadlan. 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*, 34(7):939, 1984.

[125] Dev Mehta, Robert Jackson, Gaurav Paul, Jiong Shi, and Marwan Sabbagh. Why do trials for Alzheimer's disease drugs keep failing? a discontinued drug perspective for 2010-2015. *Expert Opinion on Investigational Drugs*, 26(6):735–739, 2017. PMID: 28460541.

[126] Chandan Misra, Yong Fan, and Christos Davatzikos. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI. *NeuroImage*, 44(4):1415 – 1422, 2009.

[127] Remika Mito, Thijs Dhollander, Ying Xia, David Raffelt, Olivier Salvado, Leonid Churilov, Christopher C. Rowe, Amy Brodtmann, Victor L. Villemagne, and Alan Connelly. In vivo microstructural heterogeneity of white matter lesions in healthy elderly and Alzheimer's disease participants using tissue compositional analysis of diffusion MRI data. *NeuroImage: Clinical*, 28:102479, 2020.

[128] Remika Mito, David Raffelt, Thijs Dhollander, David N Vaughan, J-Donald Tournier, Olivier Salvado, Amy Brodtmann, Christopher C Rowe, Victor L Villemagne, and Alan Connelly. Fibre-specific white matter reductions in Alzheimer's disease and mild cognitive impairment. *Brain*, 141(3):888–902, 2018.

[129] Elaheh Moradi, Antonietta Pepe, Christian Gaser, Heikki Huttunen, and Jussi Tohka. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *NeuroImage*, 104:398 – 412, 2015.

[130] J. C. Morris. The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, 43, 11 1993.

[131] J. C. Morris, A. Heyman, R. C. Mohs, J. P. Hughes, G. van Belle, G. Fillenbaum, E. D. Mellits, and C. Clark. The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assesment of Alzheimer's disease. *Neurology*, 39, 09 1989.

[132] Melissa E Murray, Neill R Graff-Radford, Owen A Ross, Ronald C Petersen, Ranjan Duara, and Dennis W Dickson. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *The Lancet Neurology*, 10(9):785–796, sep 2011.

[133] Ziad S. Nasreddine, Natalie A. Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L. Cummings, and Howard Chertkow. The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4):695–699, 2005.

[134] Jasmine Nettiksimmons, Charles DeCarli, Susan Landau, and Laurel Beckett. Biological heterogeneity in ADNI amnestic mild cognitive impairment. *Alzheimer's & Dementia*, 10(5):511 – 521.e1, 2014.

[135] Emma Nichols, Cassandra E I Szoeke, Stein Emil Vollset, Nooshin Abbasi, Foad Abd-Allah, Jemal Abdela, Miloud Taki Eddine Aichour, Rufus O Akinyemi, Fares Alahdab, Solomon W Asgedom, Ashish Awasthi, Suzanne L Barker-Collo, Bernhard T Baune, Yannick Béjot, Abate B Belachew, Derrick A. Bennett, Belete Biadgo, Ali Bijani, Muhammad Shahdaat Bin Sayeed, Carol Brayne, David O Carpenter, Félix Carvalho, Ferrán Catalá-López, Ester Cerin, Jee-Young J Choi, Anh Kim Dang, Meaza G Degefa, Shirin Djalalinia, Manisha Dubey, Eyasu Ejeta Duken, David Edvardsson, Matthias Endres, Sharareh Eskandarieh, Andre Faro, Farshad Farzadfar, Seyed-Mohammad Fereshtehnejad, Eduarda Fernandes, Irina Filip, Florian Fischer, Abadi K Gebre, Demeke Geremew, Maryam Ghasemi-Kasman, Elena V. Gnedovskaya, Rajeev Gupta, Vladimir Hachinski, Tekleberhan B Hagos, Samer Hamidi, Graeme J. Hankey, Josep M Haro, Simon I Hay, Seyed Sina N Irvani, Ravi P Jha, Jost B. Jonas, Rizwan Kalani, André Karch, Amir Kasaeian, Yousef Saleh Khader, Ibrahim A Khalil, Ejaz Ahmad Khan, Tripti Khanna, Tawfik A M Khoja, Jagdish Khubchandani, Adnan Kisa, Katarzyna Kissimova-Skarbek, Mika Kivimäki, Ai Koyanagi, Kristopher J Krohn, Giancarlo Logroscino, Stefan Lorkowski, Marek Majdan, Reza Malekzadeh, Winfried März, João Massano, Getnet Mengistu, Atte Meretoja, Moslem Mohammadi, Maryam Mohammadi-Khanaposhtani, Ali H Mokdad, Stefania Mondello, Ghobad Moradi, Gabriele Nagel, Mohsen Naghavi, Gurudatta Naik, Long H Nguyen, Trang H Nguyen, Yirga L Nirayo, Molly R Nixon, Richard Ofori-Asenso, Felix A Ogbo, Andrew T Olagunju, Mayowa O Owolabi, Songhomitra Panda-Jonas, Valéria M de Azeredo Passos, David M. Pereira, Gabriel D Pinilla-Monsalve, Michael A Piradov, Constance D Pond, Hossein Poustchi, Mostafa Qorbani, Amir Radfar, Robert C Reiner, Stephen R Robinson, Gholamreza Roshandel, Ali Rostami, Tom C Russ, Perminder S Sachdev, Hosein Safari, Saeid Safiri, Ramesh Sahathevan, Yahya Salimi, Maheswar Satpathy, Monika Sawhney, Mete Saylan, Sadaf G. Sepanlou, Azadeh Shafieesabet, Masood A Shaikh, Mohammad Ali Sahraian, Mika Shigematsu, Rahman Shiri, Ivy Shiue, João P Silva, Mari Smith, Soheila Sobhani, Dan J Stein, Rafael Tabarés-Seisdedos, Marcos R Tovani-Palone, Bach X Tran, Tung Thanh Tran, Amanuel T Tsegay, Irfan Ullah, Narayanaswamy Venketasubramanian, Vasily Vlassov, Yuan-Pang Wang, Jordan Weiss, Ronny Westerman, Tissa Wijeratne, Grant M.A. Wyper, Yuichiro Yano, Ebrahim M Yimer, Naohiro Yonemoto, Mahmoud Yousefifard, Zoubida Zaidi, Zohreh Zare, Theo Vos, Valery L. Feigin, and Christopher J L Murray. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet Neurology*, 18(1):88–106, January 2019.

[136] Julie Ottoy, Ellis Niemantsverdriet, Jeroen Verhaeghe, Ellen De Roeck, Hanne Struyfs, Charisse Somers, Leonie wyffels, Sarah Ceyssens, Sara Van Mossevelde, Tobi Van den Bossche, Christine Van Broeckhoven, Annemie Ribbens, Maria Bjerke, Sigrid Stroobants, Sebastiaan Engelborghs, and Steven Staelens. Association of

short-term cognitive decline and MCI-to-AD dementia conversion with CSF, MRI, amyloid- and 18F-FDG-PET imaging. *NeuroImage: Clinical*, 22:101771, 2019.

[137] Jong-Yun Park, , Han Kyu Na, Sungsoo Kim, Hyunwook Kim, Hee Jin Kim, Sang Won Seo, Duk L. Na, Cheol E. Han, and Joon-Kyung Seong. Robust identification of Alzheimer's disease subtypes based on cortical atrophy patterns. *Scientific Reports*, 7(1), mar 2017.

[138] Lovingly Quitania Park, Alden L. Gross, Donald G. McLaren, Judy Pa, Julene K. Johnson, Meghan Mitchell, and Jennifer J. Manly. Confirmatory factor analysis of the ADNI neuropsychological battery. *Brain Imaging and Behavior*, 6(4):528–539, 2012.

[139] Ofer Pasternak, Nir Sochen, Yaniv Gur, Nathan Intrator, and Yaniv Assaf. Free water elimination and mapping from diffusion MRI. *Magnetic Resonance in Medicine*, 62(3):717–730, 2009.

[140] Telma Pereira, Francisco L. Ferreira, Sandra Cardoso, Dina Silva, Alexandre de Mendonça, Manuela Guerreiro, Sara C. Madeira, and for the Alzheimer's Disease Neuroimaging Initiative. Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease: a feature selection ensemble combining stability and predictability. *BMC Medical Informatics and Decision Making*, 18(1):137, Dec 2018.

[141] Telma Pereira, Luís Lemos, Sandra Cardoso, Dina Silva, Ana Rodrigues, Isabel Santana, Alexandre de Mendonça, Manuela Guerreiro, and Sara C. Madeira. Predicting progression of mild cognitive impairment to dementia using neuropsychological data: a supervised learning approach using time windows. *BMC Medical Informatics and Decision Making*, 17(1):110, Jul 2017.

[142] Jessica Peter, Ahmed Abdulkadir, Christoph Kaller, Dorothee Kümmerer, Michael Hüll, Werner Vach, and Stefan Klöppel. Subgroups of Alzheimer's disease: Stability of empirical clusters over time. *Journal of Alzheimer's Disease*, 42(2):651 – 661, 2014.

[143] R. C. Petersen. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256(3):183–194, 2004.

[144] Ronald C. Petersen. Mild cognitive impairment. *New England Journal of Medicine*, 364(23):2227–2234, 2011. PMID: 21651394.

[145] R. I. Pfeffer, T. T. Kurosaki, C. H. Harrah, J. M. Chance, and S. Filos. Measurement of functional activities in older adults in the community. *Journal of Gerontology*, 37, 05 1982.

[146] Maximilian Pietsch, Daan Christiaens, Jana Hutter, Lucilio Cordero-Grande, Anthony N. Price, Emer Hughes, A. David Edwards, Joseph V. Hajnal, Serena J. Counsell, and J-Donald Tournier. A framework for multi-component analysis of diffusion mri data over the neonatal period. *NeuroImage*, 186:321–337, 2019.

[147] Claudia Plant, Stefan J. Teipel, Annahita Oswald, Christian Böhm, Thomas Meindl, Janaina Mourao-Miranda, Arun W. Bokde, Harald Hampel, and Michael Ewers. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. *NeuroImage*, 50(1):162 – 174, 2010.

[148] Martin James Prince, Anders Wimo, Maelenn Mari Guerchet, Gemma Claire Ali, Yu-Tzu Wu, and Matthew Prina. *World Alzheimer Report 2015 - The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends*. Alzheimer's Disease International, August 2015.

[149] Olivier Querbes, Florent Aubry, Jérémie Pariente, Jean-Albert Lotterie, Jean-François Démonet, Véronique Duret, Michèle Puel, Isabelle Berry, Jean-Claude Fort, and Pierre Celsis. Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve. *Brain*, 132(8):2036, 2009.

[150] Joaquim Radua, Erick Jorge Canales-Rodríguez, Edith Pomarol-Clotet, and Raymond Salvador. Validity of modulation and optimal settings for advanced voxel-based morphometry. *NeuroImage*, 86:81 – 90, 2014.

[151] David Raffelt, Thijs Dhollander, Jacques-Donald Tournier, Rami Tabbara, Robert Smith, Eric Pierre, and Alan Connelly. Bias field correction and intensity normalisation for quantitative analysis of apparent fibre density. In *International Society of Magnetic Resonance in Medicine*, page 3541, 04 2017.

[152] David Raffelt, J-Donald Tournier, Jurgen Fripp, Stuart Crozier, Alan Connelly, and Olivier Salvado. Symmetric diffeomorphic registration of fibre orientation distributions. *NeuroImage*, 56(3):1171 – 1180, 2011.

[153] David Raffelt, J.-Donald Tournier, Stephen Rose, Gerard R. Ridgway, Robert Henderson, Stuart Crozier, Olivier Salvado, and Alan Connelly. Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*, 59(4):3976 – 3994, 2012.

[154] David A. Raffelt, Robert E. Smith, Gerard R. Ridgway, J-Donald Tournier, David N. Vaughan, Stephen Rose, Robert Henderson, and Alan Connelly. Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. *NeuroImage*, 117:40 – 55, 2015.

[155] David A. Raffelt, J.-Donald Tournier, Robert E. Smith, David N. Vaughan, Graeme Jackson, Gerard R. Ridgway, and Alan Connelly. Investigating white matter fibre density and morphology using fixel-based analysis. *NeuroImage*, 144, Part A:58 – 73, 2017.

[156] Nandini Raghavan, Mahesh N. Samtani, Michael Farnum, Eric Yang, Gerald Novak, Michael Grundman, Vaibhav Narayan, Allitia DiBernardo, and Alzheimer's Disease Neuroimaging Initiative. The ADAS-Cog revisited: Novel composite scales based on ADAS-Cog to improve efficiency in MCI and early AD trials. *Alzheimer's & Dementia*, 9(1S):S21–S31, 2013.

[157] Saima Rathore, Mohamad Habes, Muhammad Aksam Iftikhar, Amanda Shacklett, and Christos Davatzikos. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage*, pages –, 2017.

[158] R.M. Reitan and D. Wolfson. *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. Halstead-Reitan neuropsychological test battery for adults. Neuropsychology Press, 1985.

[159] A. P. Reynolds, G. Richards, B. de la Iglesia, and V. J. Rayward-Smith. Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*, 5(4):475–504, Dec 2006.

[160] Shannon L. Risacher, Wesley H. Anderson, Arnaud Charil, Peter F. Castelluccio, Sergey Shcherbinin, Andrew J. Saykin, and Adam J. Schwarz and. Alzheimer disease brain atrophy subtypes are associated with cognition and rate of decline. *Neurology*, 89(21):2176–2186, oct 2017.

[161] Gilberto Rojas-Vite, Ricardo Coronado-Leija, Omar Narvaez-Delgado, Alonso Ramírez-Manzanares, José Luis Marroquín, Ramsés Noguez-Imm, Marcos L. Aranda, Benoit Scherrer, Jorge Larriva-Sahd, and Luis Concha. Histological validation of per-bundle water diffusion metrics within a region of fiber crossing following axonal degeneration. *NeuroImage*, 201:116013, 2019.

[162] WG Rosen, RC. Mohs, and KL. Davis. A new rating scale for Alzheimer's disease. *American Journal of Psychiatry*, 141(11):1356–1364, 1984. PMID: 6496779.

[163] Yves Rosseel. lavaan: An R package for structural equation modeling. *Journal of Statistical Software, Articles*, 48(2):1–36, 2012.

[164] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.

[165] Andrea Rueda, Fabio Gonzalez, and Eduardo Romero. Extracting salient brain patterns for imaging-based classification of neurodegenerative diseases. *IEEE Transactions on Medical Imaging*, 33(6):1262–1274, 2014.

[166] Marwan N. Sabbagh, Suzanne Hendrix, and John E. Harrison. FDA position statement "Early Alzheimer's disease: Developing drugs for treatment, guidance for industry". *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 5:13 – 19, 2019.

[167] David H. Salat, Douglas N. Greve, Jennifer L. Pacheco, Brian T. Quinn, Karl G. Helmer, Randy L. Buckner, and Bruce Fischl. Regional white matter volume differences in nondemented aging and Alzheimer's disease. *NeuroImage*, 44(4):1247–1258, 2009.

[168] Jorge Samper-González, Ninon Burgos, Simona Bottani, Sabrina Fontanella, Pascal Lu, Arnaud Marcoux, Alexandre Routier, Jérémy Guillon, Michael Bacci, Junhao Wen, Anne Bertrand, Hugo Bertin, Marie-Odile Habert, Stanley Durrleman, Theodoros Evgeniou, and Olivier Colliot. Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. *NeuroImage*, 183:504 – 521, 2018.

[169] P Scheltens, F Barkhof, J Valk, PR Algra, RG van der Hoop, J Nauta, and EC Wolters. White matter lesions on magnetic resonance imaging in clinically diagnosed Alzheimer's disease. evidence for heterogeneity. *Brain*, 115(3):735–748, 1992.

[170] P Scheltens, D Leys, F Barkhof, D Huglo, H C Weinstein, P Vermersch, M Kuiper, M Steinling, E C Wolters, and J Valk. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *Journal of Neurology, Neurosurgery & Psychiatry*, 55(10):967–972, 1992.

[171] Philip Scheltens, Kaj Blennow, Monique M B Breteler, Bart de Strooper, Giovanni B Frisoni, Stephen Salloway, and Wiesje Maria Van der Flier. Alzheimer's disease. *The Lancet*, pages –, 2016.

[172] Maartje H. N. Schermer and Edo Richard. On the reconceptualization of Alzheimer's disease. *Bioethics*, 33(1):138–145, 2019.

[173] Daniel Schmitter, Alexis Roche, Bénédicte Maréchal, Delphine Ribes, Ahmed Abdulkadir, Meritxell Bach-Cuadra, Alessandro Daducci, Cristina Granziera, Stefan Klöppel, Philippe Maeder, Reto Meuli, and Gunnar Krueger. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical*, 7:7 – 17, 2015.

[174] Norbert Schuff, Duygu Tosun, Philip S. Insel, Gloria C. Chiang, Diana Truran, Paul S. Aisen, Clifford R. Jack, and Michael W. Weiner. Nonlinear time course of brain volume loss in cognitively normal and impaired elders. *Neurobiology of Aging*, 33(5):845–855, May 2012.

[175] Chris Seiffert, Taghi M Khoshgoftaar, Jason Van Hulse, and Amri Napolitano. RUSboost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(1):185–197, 2010.

[176] Claire E. Sexton, Ukwuori G. Kalu, Nicola Filippini, Clare E. Mackay, and Klaus P. Ebmeier. A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging*, 32(12):2322.e5 – 2322.e18, 2011.

[177] Hiral Shah, Emiliano Albanese, Cynthia Duggan, Igor Rudan, Kenneth M Langa, Maria C Carrillo, Kit Yee Chan, Yves Joanette, Martin Prince, Martin Rossor, Shekhar Saxena, Heather M Snyder, Reisa Sperling, Mathew Varghese, Huali Wang, Marc Wortmann, and Tarun Dua. Research priorities to reduce the global burden of dementia by 2025. *The Lancet Neurology*, 15(12):1285 – 1294, 2016.

[178] Robert Smith, Dennis Dimond, David Vaughan, Donna Parker, Thijs Dhollander, Graeme Jackson, and Alan Connelly. Intrinsic non-stationarity correction for fixel-based analysis. In *Proc OHBM*, 06 2019.

[179] Robert E. Smith, Jacques-Donald Tournier, Fernando Calamante, and Alan Connelly. SIFT: Spherical-deconvolution informed filtering of tractograms. *NeuroImage*, 67:298 – 312, 2013.

[180] Stephen M. Smith and Thomas E. Nichols. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, 44(1):83 – 98, 2009.

[181] Charisse Somers, Hanne Struyfs, Joery Goossens, Ellis Niemantsverdriet, Jill Luyckx, Naomi De Roeck, Ellen De Roeck, Bart De Vil, Patrick Cras, Jean-Jacques Martin, Peter-Paul De Deyn, Maria Bjerke, and Sebastiaan Engelborghs. A decade of cerebrospinal fluid biomarkers for Alzheimer's disease in Belgium. *Journal of Alzheimer's Disease*, 54(1):383–395, August 2016.

[182] Hanne Struyfs, Diana Maria Sima, Melissa Wittens, Annemie Ribbens, Nuno Pedrosa de Barros, Thanh Vân Phan, Maria Ines Ferraz Meyer, Lene Claes, Ellis Niemantsverdriet, Sebastiaan Engelborghs, Wim Van Hecke, and Dirk Smeets. Automated MRI volumetry as a diagnostic tool for Alzheimer's disease: Validation of icobrain dm. *NeuroImage: Clinical*, 26:102243, 2020.

[183] Hanne Struyfs, Wim Van Hecke, Jelle Veraart, Jan Sijbers, Sylvie Slaets, Maya De Belder, Laura Wuyts, Benjamin Peters, Kristel Sleegers, and Caroline et al. Robberecht. Diffusion kurtosis imaging: A possible MRI biomarker for AD diagnosis? *Journal of Alzheimer's Disease*, 48(4):937–948, 2015.

[184] L. Sørensen, A. Pai, C. Anker, I. Balas, M. Lillholm, C. Igel, and M. Nielsen. Dementia diagnosis using MRI cortical thickness, shape, texture, and volumetry. *Proc MICCAI Workshop Challenge on Computer-Aided Diagnosis of Dementia Based on Structural MRI Data*, pages 111–118, 01 2014.

[185] Lauge Sørensen, Christian Igel, Naja Liv Hansen, Merete Osler, Martin Lauritzen, Egill Rostrup, Mads Nielsen, for the Alzheimer's Disease Neuroimaging Initiative, the Australian Imaging Biomarkers, and Lifestyle Flagship Study of Ageing. Early detection of Alzheimer's disease using MRI hippocampal texture. *Human Brain Mapping*, 37(3):1148–1161, 2016.

[186] Curtis Tatsuoka, Huiyun Tseng, Judith Jaeger, Ferenc Varadi, Mark A Smith, Tomoko Yamada, Kathleen A Smyth, and Alan J Lerner and. Modeling the heterogeneity in risk of progression to Alzheimer's disease across cognitive profiles in mild cognitive impairment. *Alzheimer's Research & Therapy*, 5(2):14, 2013.

[187] Alexander N.W. Taylor, Lana Kambeitz-Ilankovic, Benno Gesierich, Lee Simon-Vermot, Nicolai Franzmeier, Miguel Á. Araque Caballero, Sophia Müller, Liu Hesheng, Birgit Ertl-Wagner, Katharina Bürger, Michael W. Weiner, Martin Dichgans, Marco Duering, Michael Ewers, and Alzheimer's Disease Neuroimaging Initiative (ADNI). Tract-specific white matter hyperintensities disrupt neural network function in Alzheimer's disease. *Alzheimer's & Dementia*, 13(3):225–235, 2017.

[188] Stefan Teipel, Alexander Drzezga, Michel J Grothe, Henryk Barthel, Gaël Chételat, Norbert Schuff, Pawel Skudlarski, Enrica Cavedo, Giovanni B Frisoni, Wolfgang Hoffmann, Jochen René Thyrian, Chris Fox, Satoshi Minoshima, Osama Sabri, and Andreas Fellgiebel. Multimodal imaging in Alzheimer's disease: validity and usefulness for early detection. *The Lancet Neurology*, 14(10):1037–1053, October 2015.

[189] Stefan J. Teipel, Michel J. Grothe, Massimo Filippi, Andreas Fellgiebel, Martin Dyrba, Giovanni B. Frisoni, Thomas Meindl, Arun L.W. Bokde, Harald Hampel, and Stefan Kloppel. Fractional anisotropy changes in Alzheimer's disease depend on the underlying fiber tract architecture: A multiparametric DTI study using joint independent component analysis. *Journal of Alzheimer's Disease*, 41(1):69–83, 2014.

[190] Mara ten Kate, Ellen Dicks, Pieter Jelle Visser, Wiesje M van der Flier, Charlotte E Teunissen, Frederik Barkhof, Philip Scheltens, Betty M Tijms, and Alzheimer's Disease Neuroimaging Initiative. Atrophy subtypes in prodromal Alzheimer's disease are associated with cognitive decline. *Brain*, 141(12):3443–3456, 10 2018.

[191] Mara ten Kate, Alberto Redolfi, Enrico Peira, Isabelle Bos, Stephanie J. Vos, Rik Vandenberghe, Silvy Gabel, Jolien Schaeverbeke, Philip Scheltens, Olivier Blin, Jill C. Richardson, Regis Bordet, Anders Wallin, Carl Eckerstrom, José Luis Molinuevo, Sebastiaan Engelborghs, Christine Van Broeckhoven, Pablo Martinez-Lage, Julius Popp, Magdalini Tsolaki, Frans R. J. Verhey, Alison L. Baird, Cristina Legido-Quigley, Lars Bertram, Valerija Dobricic, Henrik Zetterberg, Simon Lovestone,

Johannes Streffer, Silvia Bianchetti, Gerald P. Novak, Jerome Revillard, Mark F. Gordon, Zhiyong Xie, Viktor Wottschel, Giovanni Frisoni, Pieter Jelle Visser, and Frederik Barkhof. MRI predictors of amyloid pathology: results from the EMIF-AD multimodal biomarker discovery study. *Alzheimer's Research & Therapy*, 10(1), September 2018.

[192] Kelsey R. Thomas, Joel S. Eppig, Alexandra J. Weigand, Emily C. Edmonds, Christina G. Wong, Amy J. Jak, Lisa Delano-Wood, Douglas R. Galasko, David P. Salmon, Steven D. Edland, and Mark W. Bondi. Artificially low mild cognitive impairment to normal reversion rate in the Alzheimer's Disease Neuroimaging Initiative. *Alzheimer's & Dementia*, 15(4):561 – 569, 2019.

[193] Paul M. Thompson, Kiralee M. Hayashi, Greig de Zubicaray, Andrew L. Janke, Stephen E. Rose, James Semple, David Herman, Michael S. Hong, Stephanie S. Dittmer, David M. Doddrell, and Arthur W. Toga. Dynamics of gray matter loss in Alzheimer's disease. *The Journal of Neuroscience*, 23(3):994–1005, 2003.

[194] Betty M Tijms and Pieter Jelle Visser. Capturing the Alzheimer's disease pathological cascade. *The Lancet Neurology*, 17(3):199–200, mar 2018.

[195] Matthew Toews, William Wells III, D. Louis Collins, and Tal Arbel. Feature-based morphometry: Discovering group-related anatomical patterns. *NeuroImage*, 49(3):2318 – 2327, 2010.

[196] Manuela Tondelli, Gordon K. Wilcock, Paolo Nichelli, Celeste A. De Jager, Mark Jenkinson, and Giovanna Zamboni. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. *Neurobiology of Aging*, 33(4):825.e25 – 825.e36, 2012.

[197] J-Donald Tournier. Diffusion mri in the brain – theory and concepts. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 112-113:1 – 16, 2019.

[198] J-Donald Tournier, Fernando Calamante, and Alan Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459 – 1472, 2007.

[199] J.-Donald Tournier, Fernando Calamante, David G. Gadian, and Alan Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.

[200] J-Donald Tournier, Robert Smith, David Raffelt, Rami Tabbara, Thijs Dhollander, Maximilian Pietsch, Daan Christiaens, Ben Jeurissen, Chun-Hung Yeh, and Alan Connelly. Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, 202:116137, 2019.

[201] Jacques-Donald Tournier, Fernando Calamante, and Allan Connelly. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. In *International Society of Magnetic Resonance in Medicine*, 2010.

[202] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, jun 2010.

[203] Dallas P. Veitch, Michael W. Weiner, Paul S. Aisen, Laurel A. Beckett, Nigel J. Cairns, Robert C. Green, Danielle Harvey, Clifford R. Jack, William Jagust, John C. Morris, Ronald C. Petersen, Andrew J. Saykin, Leslie M. Shaw, Arthur W. Toga, and John Q. Trojanowski. Understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer's disease neuroimaging initiative. *Alzheimer's & Dementia*, 15(1):106 – 152, 2019.

[204] Prashanthi Vemuri, Jeffrey L. Gunter, Matthew L. Senjem, Jennifer L. Whitwell, Kejal Kantarci, David S. Knopman, Bradley F. Boeve, Ronald C. Petersen, and Clifford R. Jack Jr. Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies. *NeuroImage*, 39(3):1186 – 1197, 2008.

[205] Jelle Veraart, Els Fieremans, and Dmitry S. Novikov. Diffusion MRI noise mapping using random matrix theory. *Magnetic Resonance in Medicine*, 76(5):1582–1593, 2016.

[206] Sjoerd B. Vos, Derek K. Jones, Max A. Viergever, and Alexander Leemans. Partial volume effect as a hidden covariate in DTI analyses. *NeuroImage*, 55(4):1566 – 1576, 2011.

[207] Christian Wachinger, David H. Salat, Michael Weiner, Martin Reuter, and for the Alzheimer's Disease Neuroimaging Initiative. Whole-brain analysis reveals increased neuroanatomical asymmetries in dementia for hippocampus and amygdala. *Brain*, 139(12):3253–3266, 10 2016.

[208] D. Wechsler. *Wechsler Memory Scale-Revised: Manual*. Harcourt Brace Jovanovich, 1987.

[209] Michael W. Weiner, Dallas P. Veitch, Paul S. Aisen, Laurel A. Beckett, Nigel J. Cairns, Robert C. Green, Danielle Harvey, Clifford R. Jack, William Jagust, John C. Morris, Ronald C. Petersen, Andrew J. Saykin, Leslie M. Shaw, Arthur W. Toga, and John Q. Trojanowski. Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials. *Alzheimer's & Dementia*, 13(4):e1 – e85, 2017.

[210] Sandra Weintraub, Alissa H. Wicklund, and David P. Salmon. The neuropsychological profile of Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, 2(4), 2012.

[211] Eric Westman, Andrew Simmons, Yi Zhang, J-Sebastian Muehlboeck, Catherine Tunnard, Yawu Liu, Louis Collins, Alan Evans, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Kłoszewska, Hilkka Soininen, Simon Lovestone, Christian Spenger, and Lars-Olof Wahlund. Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *NeuroImage*, 54(2):1178 – 1187, 2011.

[212] Philip S.J. Weston, Ivor J.A. Simpson, Natalie S. Ryan, Sebastien Ourselin, and Nick C. Fox. Diffusion imaging changes in grey matter in Alzheimer's disease: a potential marker of early neurodegeneration. *Alzheimer's Research & Therapy*, 7(1):1–8, 2015.

[213] Jennifer L Whitwell, Dennis W Dickson, Melissa E Murray, Stephen D Weigand, Nirubol Tosakulwong, Matthew L Senjem, David S Knopman, Bradley F Boeve,

Joseph E Parisi, and Ronald C et al. Petersen. Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. *The Lancet Neurology*, 11(10):868–877, 2012.

[214] Anderson M. Winkler, Gerard R. Ridgway, Matthew A. Webster, Stephen M. Smith, and Thomas E. Nichols. Permutation inference for the general linear model. *NeuroImage*, 92:381 – 397, 2014.

[215] Mandy Melissa Jane Wittens, Diana Maria Sima, Ruben Houbrechts, Annemie Ribbens, Ellis Niemantsverdriet, Erik Fransen, Christine Bastin, Florence Benoit, Bruno Bergmans, Jean-Christophe Bier, Peter-Paul De Deyn, Olivier Deryck, Bernard Hanseeuw, Adrian Ivanoiu, Jean-Claude Lemper, Eric Mormont, Gaëtane Picard, Ezequiel de la Rosa, Eric Salmon, Kurt Segers, Anne Sieben, Dirk Smeets, Hanne Struyfs, Evert Thiery, Jos Tournoy, Eric Triau, Anne-Marie Vanbinst, Jan Versijpt, Maria Bjerke, and Sebastiaan Engelborghs. Diagnostic performance of automated MRI volumetry by icobrain dm for Alzheimer's disease in a clinical setting: A REMEMBER study. *Journal of Alzheimer's Disease*, page 1–17, July 2021.

[216] Robin Wolz, Valtteri Julkunen, Juha Koikkalainen, Eini Niskanen, Dong Ping Zhang, Daniel Rueckert, Hilkka Soininen, Jyrki Lötjönen, and the Alzheimer's Disease Neuroimaging Initiative. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. *PLoS ONE*, 6(10):1–9, 10 2011.

[217] Daoqiang Zhang, Yaping Wang, Luping Zhou, Hong Yuan, and Dinggang Shen. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage*, 55(3):856 – 867, 2011.

[218] Jing Zhang, Chunshui Yu, Guilian Jiang, Weifang Liu, and Longzheng Tong. 3D texture analysis on MRI images of Alzheimer's disease. *Brain Imaging and Behavior*, 6(1):61–69, 2011.

[219] Lin Zhuang, Perminder S. Sachdev, Julian N. Trollor, Simone Reppermund, Nicole A. Kochan, Henry Brodaty, and Wei Wen. Microstructural white matter changes, not hippocampal atrophy, detect early amnestic mild cognitive impairment. *PLOS ONE*, 8(3):1–10, 03 2013.

[220] Chen Zu, Biao Jie, Mingxia Liu, Songcan Chen, Dinggang Shen, and Daoqiang Zhang. Label-aligned multi-task feature learning for multimodal classification of Alzheimer's disease and mild cognitive impairment. *Brain Imaging and Behavior*, pages 1–12, 2015.