Neuroimaging biomarkers in genetic frontotemporal dementia : towards a timely diagnosis

Feis, R.A.

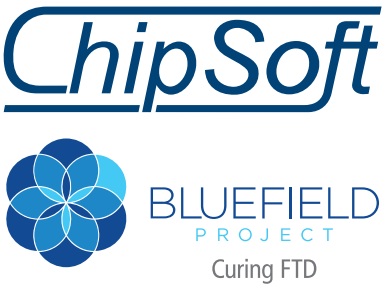
Neuroimaging biomarkers in genetic frontotemporal dementia: towards a timely diagnosis

Rogier Feis

Cover design Rogier Feis
Layout Rogier Feis
Printed by Gildeprint
ISBN 9789464021035

The research described in this dissertation was funded by the Leiden University Medical Centre and the Netherlands Organisation for Scientific Research (NWO).

Printing of this dissertation was funded by Alzheimer Nederland, ChipSoft & The Bluefield Project.



Neuroimaging biomarkers in genetic frontotemporal dementia: towards a timely diagnosis

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op 14–10–2020
klokke 15:00 uur

door

Rogier Alexander Feis

geboren te 's-Gravenhage
in 1991

Promotores

Prof. dr. S.A.R.B. Rombouts

Prof. dr. M.A. van Buchem

Copromotor

Dr. J. van der Grond

Promotiecommissie

Prof. dr. N.J.A. van der Wee	
Prof. dr. H.A.M. Middelkoop	
Prof. dr. J.C. van Swieten	Erasmus MC Rotterdam
Prof. dr. M. Smits	Erasmus MC Rotterdam
Prof. dr. C.E. Mackay	University of Oxford

Dit proefschrift is opgedragen aan mijn ouders.

Table of contents

Chapter 1

General introduction

Background

Frontotemporal dementia (FTD) is a common and devastating form of dementia that often occurs in the presenile population, i.e., aged under 65 years (Ratnavalli et al., 2002; Harvey et al., 2003; Hogan et al., 2016). It was first described in 1892 by Arnold Pick in a patient with aphasia and asymmetric temporal lobe atrophy (Pick, 1892). Today, FTD represents a diverse spectrum of clinical syndromes. The disease is typically characterised by progressive degeneration of the frontal and temporal lobes, leading to behavioural disorders and language deficits (Gorno-Tempini et al., 2011; Rascovsky et al., 2011). However, FTD patients may additionally develop symptoms of amyotrophic lateral sclerosis (ALS; Lomen-Hoerth et al., 2002), corticobasal syndrome (CBS), or progressive supranuclear palsy (PSP; Josephs et al., 2006). Moreover, considerable clinical overlap exists with other dementias and psychiatric disorders (Krudop et al., 2015). For example, Alzheimer's disease (AD) patients may exhibit behavioural symptoms (Johnson et al., 1999), whereas FTD patients may present with memory deficits (Graham et al., 2005). As such, clinical diagnosis is challenging, especially in the initial stages of the disease.

Early-stage and accurate diagnosis is important for establishing a prognosis and organising patient care. Furthermore, early-stage diagnosis is crucial for the development and application of disease modifying treatments. Most dementia treatments aim to reverse the neuropathological pathways that lead up to neuronal cell death, for example through inhibition of pathological proteins' accumulation (Cummings et al., 2018). The ability to diagnose patients before they develop marked cerebral atrophy is therefore essential to increase the window of opportunity to test these potential treatments. This dissertation aims to establish neuroimaging biomarkers that are sensitive to early-stage FTD.

Frontotemporal dementia heterogeneity

FTD is a many-faced disease. There are multiple distinct clinical FTD syndromes, and several types of underlying neuropathological substrates. Furthermore, FTD has symptomatic overlap with psychiatric diseases and different dementia types. To illustrate FTD heterogeneity, **Figure 1.1** shows a simplification of the relationships between the clinical FTD syndromes, the underlying pathology, and genetics.

Clinical heterogeneity

The most common FTD manifestation is behavioural variant FTD (bvFTD), which accounts for 50–80% of all FTD cases (Johnson et al., 2005; Seelaar et al., 2008; Hogan et al., 2016). Typically, bvFTD patients show signs of disinhibition, apathy, loss of empathy, compulsive behaviour, dietary changes, and/or executive dysfunction. On magnetic resonance imaging (MRI), bvFTD is characterised by frontal and/or anterior temporal atrophy (Rascovsky et al., 2011).

The other main FTD manifestation is primary progressive aphasia (PPA), which can be subdivided into three variants: non-fluent variant PPA (nfvPPA), semantic variant PPA (svPPA), and logopenic variant PPA (lvPPA). The non-fluent variant is characterised by agrammatism, apraxia of speech, and, in later stages, mutism. However, nfvPPA patients usually have spared single-word comprehension and object knowledge. In these patients, brain atrophy is predominantly seen in the left posterior fronto-insula on MRI (Gorno-Tempini et al., 2011; Spinelli et al., 2017). Patients with svPPA typically have impaired confrontation naming, single-word comprehension, object knowledge, and surface dyslexia or dysgraphia. Repetition and speech production in terms of

grammar and motor speech are usually spared. These patients show bilateral atrophy of the ventral and lateral parts of the anterior temporal lobes on MRI, though the left side is usually more strongly afflicted (Gorno-Tempini et al., 2011). The logopenic variant is the most recently described PPA variant, and is characterised by impaired single-word retrieval, sentence repetition deficits, and phonologic errors. However, single-word comprehension, object knowledge, and motor speech are preserved. Patients with lvPPA usually show brain atrophy in left temporo-parietal junction areas, such as the posterior temporal, supramarginal, and angular gyri (Gorno-Tempini et al., 2011; Spinelli et al., 2017).

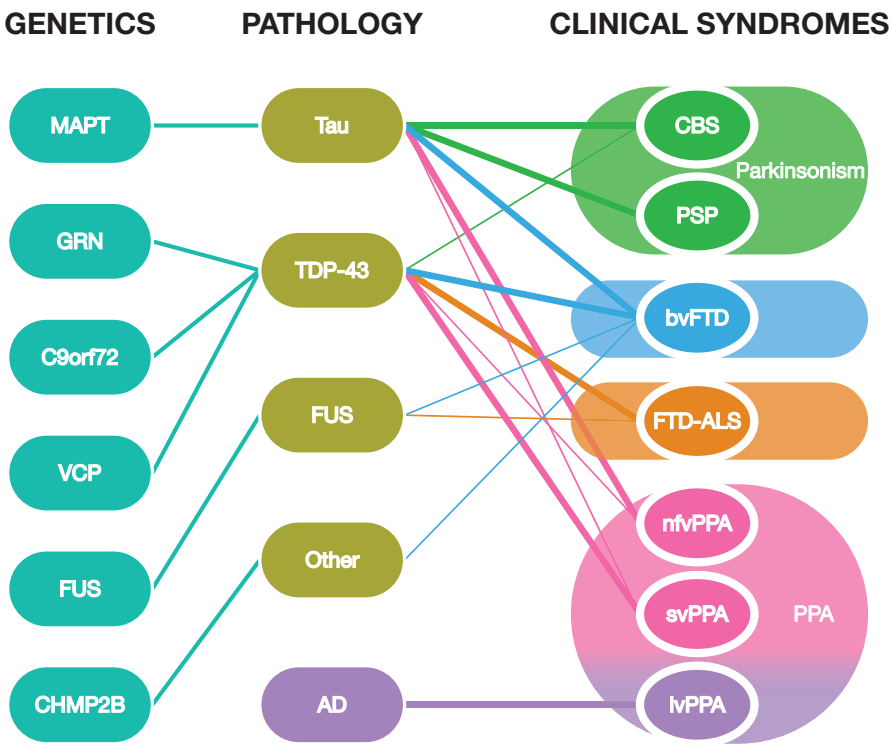


Figure 1.1 Heterogeneity in frontotemporal dementia.
This diagram features the complex relationships between the clinical FTD syndromes, the underlying pathology, and genetic causes. Thicker lines denote stronger relationships, e.g., svPPA patients usually have TDP-43 pathology, although patients may also show tau pathology (Josephs et al., 2011).
Genetics: *MAPT*, microtubule-associated protein tau; *GRN*, progranulin; *C9orf72*, chromosome 9 open reading frame 72; *VCP*, valosin-containing protein; *FUS*, fused in sarcoma gene; *CHMP2B*, charged multivesicular body protein 2b.
Pathology: TDP-43, transactive response DNA binding protein 43 kDa; FUS, fused in sarcoma protein; AD, Alzheimer's disease pathology.
Clinical syndromes: CBS, corticobasal syndrome; PSP, progressive supranuclear palsy; FTD, frontotemporal dementia; bvFTD, behavioural variant FTD; FTD-ALS, FTD with amyotrophic lateral sclerosis; PPA, primary progressive aphasia; nfvPPA, non-fluent variant PPA; svPPA, semantic variant PPA; lvPPA, logopenic variant PPA.
Modified with permission from Rohrer & Rosen (2013), and reprinted by permission of the publisher Taylor & Francis Ltd.

In addition to cognitive disorders, some FTD patients develop atypical parkinsonian syndromes, such as CBS or PSP, while others develop ALS. Conversely, patients with CBS, PSP, and ALS may develop behavioural or language dysfunction (Kertesz et al., 2005). Clinically, CBS is characterised by unilateral rigidity, apraxia, myoclonus, and the alien hand phenomenon (Armstrong et al., 2013). PSP patients typically develop vertical gaze palsy, falls, axial rigidity, and pseudobulbar palsy (Boxer et al., 2017). ALS is a motor neuron disease characterised by progressive loss of upper and/or lower motor neurons, resulting in muscle weakness, fasciculations, and eventually immobility, dysphagia, and respiratory complications (Brooks et al., 2000).

Neuropathological heterogeneity

Underlying these clinical syndromes are three groups of neuropathological substrates, named after the protein inclusions found at autopsy: tau, transactive response DNA binding protein 43 kDa (TDP-43), and RNA binding protein fused in sarcoma (FUS). Together, tau, TDP-43, and FUS constitute the neuropathological umbrella term frontotemporal lobar degeneration (FTLD). As shown in **Figure 1.1**, FTLD pathologies generally have multiple associated clinical syndromes, and vice versa, though some clinicopathological associations are stronger than others. Therefore, determining which pathology underlies a certain clinical syndrome remains challenging (Josephs et al., 2011). Interestingly, lvPPA is usually caused by AD pathology, even though its clinical syndrome may be similar to nfvPPA and svPPA (Chare et al., 2014). This dissertation will primarily focus on the clinical FTD syndromes with FTLD pathology, i.e., bvFTD, nfvPPA, and svPPA.

Genetic heterogeneity

Around 25–50% of FTD patients have a positive family history for dementia or a related disorder, such as ALS or parkinsonism. Moreover, 10–30% of FTD patients have an autosomal dominant inheritance (Seelaar et al., 2008; Rohrer et al., 2009; Benussi et al., 2015). Several gene mutations have been identified that are associated with monogenic FTD. Most commonly, genetic FTD patients have a mutation in the microtubule-associated protein tau (*MAPT*) or progranulin (*GRN*) genes, or a chromosome 9 open reading frame 72 (*C9orf72*) repeat expansion. Other, more infrequently affected genes include valosin-containing protein (*VCP*), charged multivesicular body protein 2B (*CHMP2B*), and fused in sarcoma (*FUS*; Benussi et al., 2015). These gene mutations have an autosomal dominant inheritance pattern and are all highly penetrant. Therefore, roughly 50% of subjects from genetic FTD families will eventually develop FTD. Though pathogenic FTD mutations each lead to a specific FTLD pathology, patients with the same mutation may develop different clinical FTD syndromes (Benussi et al., 2015).

The significance of a timely diagnosis

Currently, there are no United States Food and Drug Administration (FDA-)approved therapies for FTD. In fact, most pharmacological treatments for FTD involve off-label use of psychotropic medication for symptomatic relief (Pressman & Miller, 2014). Recently, advances in the understanding of FTLD pathophysiology have led to increased efforts to develop disease modifying treatments. These treatments aim to inhibit, or reverse, underlying neuropathological processes to prevent neuronal cell death. For example, disease modifying treatments may aim to inhibit tau protein aggregation, stabilise microtubules, or increase progranulin levels. A comprehensive review of past and current clinical trials until 2019 involving the FTLD spectrum has been published recently (Logroscino et al., 2019). Due to their working mechanism, disease modifying treatments

have an optimal potential in early FTD disease stages, before atrophy occurs. However, inclusion criteria for clinical trials typically include FTD diagnosis based on the relevant diagnostic criteria (Gorno-Tempini et al., 2011; Rascovsky et al., 2011), in which brain atrophy on MRI is included as diagnostic criterion. FTD patients included in clinical trials therefore often have marked atrophy, reducing the potential benefit of treatments that aim to prevent cell death. The insensitivity of current diagnostic criteria to early disease stages therefore hinders the efforts to develop disease modifying treatments.

Diagnostic delay, calculated as the time between symptom onset and diagnosis, has been estimated between three and six years for bvFTD, around three years for nfvPPA, and around five years for svPPA (Hodges et al., 2003; Pasquier et al., 2004; Roberson et al., 2005). Given that the mean survival from symptom onset is estimated at roughly eight years for bvFTD, nfvPPA, and svPPA (Kansal et al., 2016), this diagnostic delay significantly shortens the window of opportunity for FTD disease modifying treatments. Multiple factors contribute to this delay. Most importantly, neuropsychological symptoms in early stages are not specific for FTD, and may be misinterpreted as manifestations of psychiatric disease (Pijnenburg et al., 2004), especially when the patient is relatively young. FTD patients are typically younger than patients with other types of dementia; the mean age of symptom onset is roughly 58 years for bvFTD, 64 years for nfvPPA, and 60 years for svPPA (Kansal et al., 2016). In other patients, symptoms may overlap with AD (Graham et al., 2005). Consequently, diagnosis may be delayed by one or two years even after patients are first referred to a neurologist or memory clinic (Pijnenburg et al., 2004). Diagnostic delay is also partly caused by patient delay. For example, bvFTD patients typically lack disease insight and can be reluctant to seek professional help (Neary et al., 1998; Pijnenburg et al., 2004).

In order to improve clinical trial efficacy through timely diagnosis, biomarkers are necessary that are sensitive to early-stage FTD-related pathological changes. Additionally, differential diagnostic biomarkers could power clinical trials through trial stratification (e.g., based on clinical FTD variant or underlying pathology), and disease progression biomarkers could facilitate more accurate evaluation of disease modifying treatments' effects.

MRI biomarkers in genetic frontotemporal dementia

Due to the autosomal dominant inheritance and high penetrance of pathogenic FTD mutations, genetic FTD families provide an ideal population to study the pathological mechanisms that underlie FTD in early stages of the disease, and to develop biomarkers for early FTD detection. Specifically, it is possible to follow mutation carriers from such families over time, while they are still in the 'presymptomatic' stage, i.e., before symptom onset.

In recent years, presymptomatic FTD mutation carriers of *MAPT*, *GRN*, and *C9orf72* repeat expansion have been increasingly investigated to find early changes with potential MRI-based (Borroni et al., 2008; Miyoshi et al., 2010; Whitwell et al., 2011a; Dopper et al., 2014; Rohrer et al., 2015; Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018; Cash et al., 2018; Jiskoot et al., 2018a, 2019; Panman et al., 2019), cognitive (Dopper et al., 2014; Rohrer et al., 2015; Jiskoot et al., 2016, 2018b; Papma et al., 2017; Bertrand et al., 2018), fluid (Meeter et al., 2016a, 2016b, 2018a, 2018b; Lehmer et al., 2017; Galimberti et al., 2018), and positron emission tomography (PET; Miyoshi et al., 2010; Jacova et al., 2013) biomarkers. Notably, application of MRI-based biomarkers in the clinic seems viable due to the non-invasive nature and ready availability of MRI. MRI-based biomarker candidates include grey matter and white matter structure, as well as functional connectivity. Grey matter structure can be analysed using voxel-based morphometry on structural MRI scans (Good et al., 2001; Douaud et al., 2007). White matter structure is usually tested using

diffusion tensor imaging (DTI), which measures the diffusion of water inside the brain (Smith et al., 2006). The directionality of water diffusion is often used as a proxy for white matter integrity, although this interpretation remains controversial (Wheeler-Kingshott & Cercignani, 2009). Functional connectivity is calculated from resting-state functional MRI (rs-fMRI), and sheds light on how different brain regions coactivate as distinct functional networks (Beckmann & Smith, 2004). Below, we briefly outline the reported presymptomatic changes in grey matter structure, white matter diffusion, and functional connectivity in the three most common pathogenic FTD mutations.

Microtubule-associated protein tau

Presymptomatic atrophy in *MAPT* mutation carriers is located in the temporal pole and hippocampus (Panman et al., 2019). In the white matter, presymptomatic *MAPT* mutation carriers predominantly show cross-sectional DTI abnormalities in the uncinate fasciculus (Dopper et al., 2014; Jiskoot et al., 2018a). These abnormalities increase over time in the uncinate fasciculus, and expand to the left anterior thalamic radiation and left inferior fronto-occipital fasciculus (Panman et al., 2019). Functional connectivity with the default mode network, a functional brain network involved in emotional processing, self-reference, and memory (Raichle, 2015), is decreased in the lateral temporal lobe and medial prefrontal cortex in presymptomatic *MAPT* mutation carriers, and is increased in the medial parietal lobe (Whitwell et al., 2011a).

Progranulin

Presymptomatic *GRN* mutation carriers do not typically show presymptomatic atrophy (Borroni et al., 2008; Dopper et al., 2014; Cash et al., 2018; Panman et al., 2019). White matter DTI abnormalities occur in the uncinate fasciculus, inferior fronto-occipital fasciculus (Borroni et al., 2008), the splenium, and in the anterior and posterior internal capsule (Jiskoot et al., 2018a). Functional connectivity is reduced in presymptomatic *GRN* mutation carriers in the default mode network and in the salience network, which is involved in emotional processing (Dopper et al., 2014).

Chromosome 9 open reading frame 72 repeat expansion

Presymptomatic differences in *C9orf72* repeat expansion carriers are more abundant and present at an earlier stage. Grey matter atrophy predominantly occurs in the thalamus, cerebellum, and insula, but also extends to other frontotemporal regions (Rohrer et al., 2015; Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018; Cash et al., 2018; Panman et al., 2019). White matter DTI abnormalities also include widespread areas, such as the corpus callosum, anterior thalamic radiation, corticospinal tract, and the superior and inferior longitudinal fasciculi (Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018; Jiskoot et al., 2018a; Panman et al., 2019). Functional connectivity is reduced in the salience network and the default mode network (Lee et al., 2017).

Clinical translation of biomarkers using machine learning

The group differences described above show that MRI-based biomarkers may be promising for the early detection of FTD. However, the clinical translation towards a diagnostic tool requires that biomarkers have discriminative value on the individual level, not merely on group-level. To harness the full potential of biomarkers on the individual level, artificial intelligence, or more specifically machine learning, can be applied. In machine learning, labelled data is used to train

a classifier to distinguish between groups, for instance between FTD patients and controls. Once the classifier is trained, it can be used to classify new subjects whose group label is unknown. An important characteristic is that classifiers can be trained on multiple inputs to increase performance by combining complementary information (Schouten et al., 2016). While machine learning allows for many different inputs, classifiers typically generate a single score per subject, allowing for a straightforward clinical implementation in the diagnostic work-up. Recently, MRI-based classifiers have been used to study the classification of FTD patients and controls. Classifiers based on grey matter volume (Raamana et al., 2014; Klöppel et al., 2015; Koikkalainen et al., 2016; Bron et al., 2017; Canu et al., 2017; Meyer et al., 2017; Bouts et al., 2018), white matter diffusion (Bron et al., 2017; Canu et al., 2017; Bouts et al., 2018), and functional connectivity (Canu et al., 2017; Bouts et al., 2018) perform well in the classification of FTD patients and controls. However, these classification studies were performed in patients with symptomatic FTD, and it is unknown whether these results generalise to earlier FTD disease stages. The accurate classification of early-stage FTD patients or presymptomatic FTD mutation carriers could facilitate early inclusion into clinical trials with disease modifying treatments, but is currently lacking.

Aims and outline of this dissertation

Based on group differences between FTD mutation carriers and non-carriers, MRI-based biomarkers in genetic FTD show promise for early FTD detection, which may in turn facilitate clinical trial efficiency. The main goal of this dissertation is to further the translation of potential MRI-based biomarkers for FTD towards clinical application by applying them in a single-subject classification setting analogous to a diagnostic test.

In **chapter 2**, we study MRI-based classification in a cross-sectional sample of presymptomatic FTD mutation carriers and non-carriers. To assess whether a classification model based on full-blown FTD generalises to the presymptomatic disease stage, we use a bvFTD classification model (Bouts et al., 2018) to classify presymptomatic FTD mutation carriers and non-carriers. Furthermore, we comprehensively evaluate the classification performance for different MRI features by training and testing unimodal and multimodal (i.e., based on multiple MRI features) carrier-control classification models.

In **chapter 3**, our aim is to evaluate the sensitivity of MRI-based classification to FTD-related changes over time. We apply the same bvFTD classification model on longitudinal data of FTD mutation carriers and non-carriers to obtain classification scores on multiple time points. We test whether longitudinal patterns differ between non-carriers, mutation carriers that remain presymptomatic within our study's follow-up time, and mutation carriers that 'convert' to FTD (i.e., develop FTD symptoms) during follow-up.

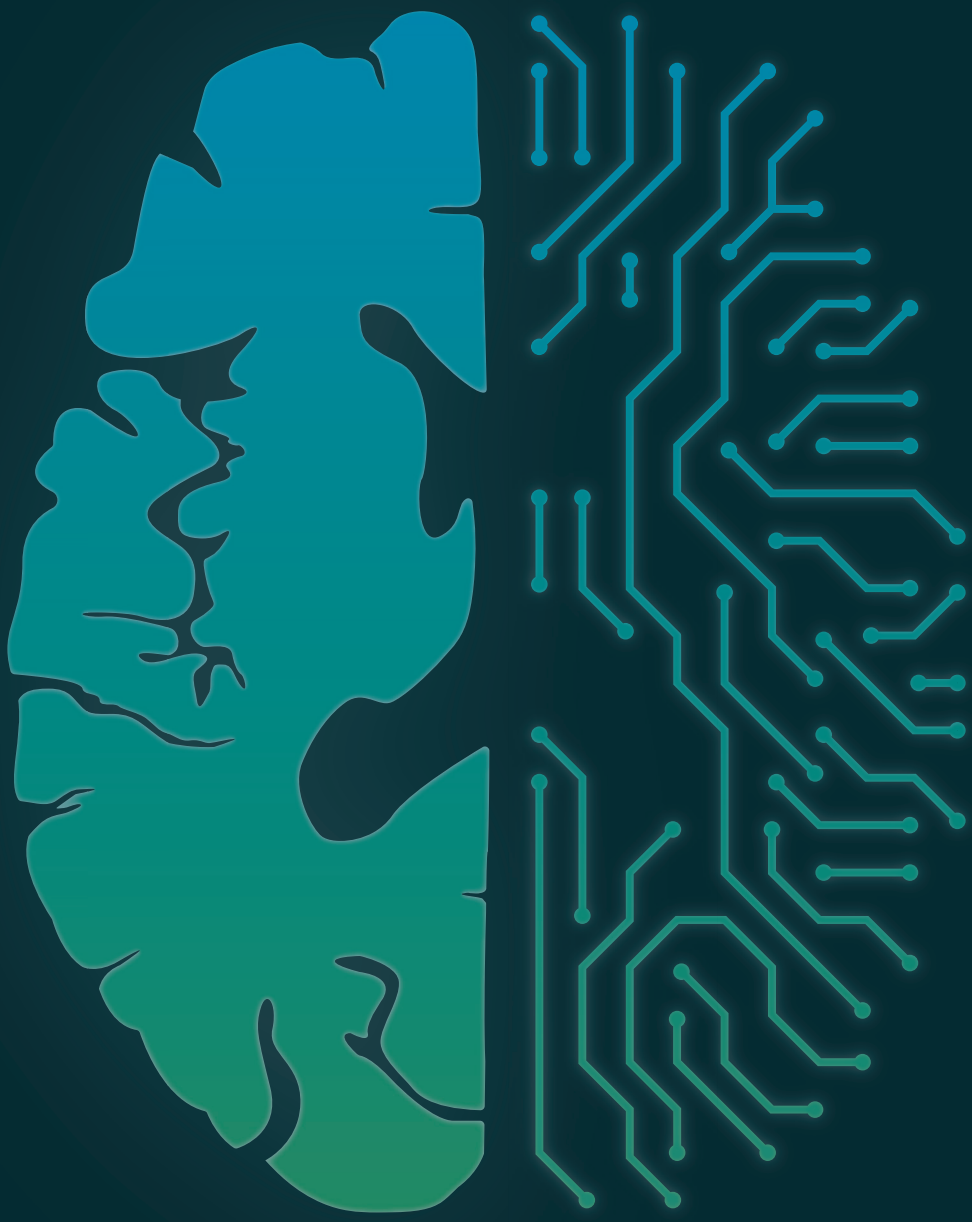
Chapter 4 describes the use of MRI-based classification to predict which FTD mutation carriers will convert to FTD within the timespan of four years, and which mutation carriers will remain presymptomatic. We aim to evaluate the prognostic value of MRI-based biomarkers to predict symptom onset in genetic FTD, since an accurate prognostic test for symptom onset could facilitate early inclusion of FTD mutation carriers into clinical trials.

The prevalence of FTD mutation carriers is relatively low. The combination of existing MRI data across centres would facilitate the pooling of subjects into larger cohorts, as well as unique study contrasts with rare groups. However, combining data from different scanners is problematic due to hardware, software, and environmental differences. In **chapter 5**, we aim to facilitate the

combination of existing rs-fMRI data from different centres through the reduction of structured noise. We apply a new noise removal method on rs-fMRI data of two groups of healthy controls from different sites to reduce scan site bias.

Another important step to develop MRI measures into specific biomarkers is to establish whether MRI changes in presymptomatic FTD mutation carriers are unique for this group, or whether they also exist in subjects at risk for different dementias, such as AD. In **chapter 6**, we jointly analyse *MAPT/GRN* mutation carriers vs. non-carriers, and apolipoprotein E ε4 (*APOE4*; the most important risk gene for AD) carriers vs. non-carriers to investigate whether early changes in risk mutation carriers of FTD and AD follow different patterns.

To conclude, we summarise and discuss the main findings of this dissertation in **chapter 7** and provide methodological considerations as well as recommendations for future research.



Chapter 2

Frontotemporal dementia mutation carriers: presymptomatic MRI-based classification

Published in *NeuroImage: Clinical 2019;22:101718* as:

Single-subject classification of presymptomatic frontotemporal dementia mutation carriers using multimodal MRI

Rogier A. Feis, Mark J.R.J. Bouts, Jessica L. Panman, Lize C. Jiskoot, Elise G.P. Dopper, Tijn M. Schouten, Frank de Vos, Jeroen van der Grond, John C. van Swieten, Serge A.R.B. Rombouts

Abstract

Classification models based on magnetic resonance imaging (MRI) may aid early diagnosis of frontotemporal dementia (FTD) but have only been applied in established FTD cases. Detection of FTD patients in earlier disease stages, such as presymptomatic mutation carriers, may further advance early diagnosis and treatment. In this study, we aim to distinguish presymptomatic FTD mutation carriers from controls on an individual level using multimodal MRI-based classification.

Anatomical MRI, diffusion-weighted imaging (DWI) and resting-state functional MRI data were collected in 55 presymptomatic FTD mutation carriers (8 microtubule-associated protein tau, 35 progranulin, and 12 chromosome 9 open reading frame 72 repeat expansion) and 48 familial controls. We calculated grey and white matter density features from anatomical MRI scans, diffusivity features from DWI, and functional connectivity features from resting-state functional MRI. These features were applied in a recently introduced multimodal behavioural variant FTD (bvFTD) classification model, and were subsequently used to train and test unimodal and multimodal carrier-control models. Classification performance was quantified using area under the receiver operating characteristic curves (AUC).

The bvFTD model was not able to separate presymptomatic mutation carriers from controls beyond chance level (AUC = 0.582, $p = 0.078$). In contrast, one unimodal and several multimodal carrier-control models performed significantly better than chance level. The unimodal model included the radial diffusivity feature and had an AUC of 0.642 ($p = 0.032$). The best multimodal model combined radial diffusivity and white matter density features (AUC = 0.684, $p = 0.004$).

FTD mutation carriers can be separated from controls with a modest AUC even before symptom onset, using a newly created carrier-control classification model, while this was not possible using a recent bvFTD classification model. A multimodal MRI-based classification score may therefore be a useful biomarker to aid earlier FTD diagnosis. The exclusive selection of white matter features in the best performing model suggests that the earliest FTD-related pathological processes occur in white matter.

Keywords: frontotemporal dementia; *MAPT* protein, human; *GRN* protein, human; *C9orf72*, human; diffusion tensor imaging; resting-state functional MRI; multimodal MRI; classification; machine learning

Introduction

Frontotemporal lobar degeneration is a common cause of early-onset dementia with a similar prevalence to Alzheimer's disease in the presenile population (Ratnavalli et al., 2002; Harvey et al., 2003; Rabinovici & Miller, 2010; Rascovsky et al., 2011; Seelaar et al., 2011). Although there are clinical disease criteria for the different clinical variants of frontotemporal dementia (FTD; Gorno-Tempini et al., 2011; Rascovsky et al., 2011), diagnosis is often complicated and delayed by clinical heterogeneity. This hinders clinicians in providing accurate prognosis, effective disease management, and developing new treatments (Mohs et al., 2001; Mendez et al., 2007; Mendez, 2009; Pressman & Miller, 2014).

Multimodal magnetic resonance imaging (MRI) has been suggested as a promising biomarker to improve on diagnostic standards in FTD. In FTD patients, MRI revealed specific patterns of neurodegeneration, involving grey matter and white matter atrophy (Whitwell & Jack, 2005; Chao et al., 2007; Rabinovici et al., 2008; Seeley et al., 2009; Zhang et al., 2011; Pan et al., 2012; Whitwell et al., 2012, 2015; Risacher & Saykin, 2013; Frings et al., 2014; Möller et al., 2015a), differences in diffusion tensor imaging (DTI) measures (Zhang et al., 2009, 2011; Agosta et al., 2012; McMillan et al., 2012, 2014; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016), and differences in functional connectivity (Zhou et al., 2010; Farb et al., 2013; Lee et al., 2014; Zhou & Seeley, 2014; Hafkemeijer et al., 2015, 2016).

These patterns have subsequently been utilised on an individual level to create MRI-based classification algorithms that can discriminate between FTD patients and control subjects (Davatzikos et al., 2008; McMillan et al., 2014; Raamana et al., 2014; Koikkalainen et al., 2016; Wang et al., 2016; Bron et al., 2017; Meyer et al., 2017; Bouts et al., 2018). Accurate classification of FTD patients using MRI measures is an important step towards a more substantiated diagnostic standard. However, most classification models are based on established FTD cases, limiting generalisability in patients who are at an earlier disease stage. Still, detection of these early-stage FTD cases is necessary to facilitate precise subject recruitment into clinical trials and potential early treatment with disease-modifying drugs (Huey et al., 2008).

In order to characterise FTD pathophysiology at an earlier stage, presymptomatic carriers of autosomal dominant FTD gene mutations were compared to controls in MRI group analyses (Borroni et al., 2008, 2012; Whitwell et al., 2011a; Rohrer et al., 2013, 2015; Dopper et al., 2014; Premi et al., 2014; Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018; Cash et al., 2018). Carriers of the three most common FTD gene mutations microtubule-associated protein tau (*MAPT*), progranulin (*GRN*), and chromosome 9 open reading frame 72 (*C9orf72*) repeat expansion show brain alterations on MRI, even well before symptom onset. In these subjects, white matter diffusivity changes (Borroni et al., 2008; Dopper et al., 2014; Papma et al., 2017) and functional connectivity changes (Whitwell et al., 2011a; Borroni et al., 2012; Dopper et al., 2014; Premi et al., 2014) are often, but not exclusively (Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018), found in the absence of grey matter atrophy, suggesting that changes in the functional architecture and white matter tracts may precede structural deterioration in the grey matter (Rohrer et al., 2013). Nonetheless, multicentre analyses of a large international cohort show grey matter loss in *MAPT* mutation carriers, *GRN* mutation carriers, and *C9orf72* repeat expansion carriers even before 'conversion' to symptomatic FTD (Rohrer et al., 2015; Cash et al., 2018). Although these presymptomatic group differences give insight into the pathophysiological mechanisms of FTD, individual heterogeneity complicates its utility in FTD diagnosis. Therefore, translation from group differences to single-subject classification models is imperative.

The present study brings two research areas together: we combine machine learning with presymptomatic FTD mutation carriers to study individual classification of FTD-pathology at an early stage. Our aim is to distinguish individual presymptomatic FTD mutation carriers from healthy controls using multimodal MRI.

Methods

Design

In order to distinguish presymptomatic FTD mutation carriers from controls, we applied two models. First, we applied a recent behavioural variant FTD (bvFTD-)control classification model (Bouts et al., 2018) to our MRI data to investigate whether the model separates presymptomatic mutation carriers from controls. We shall refer to this model as the 'bvFTD model'. In a second analysis, we trained new classification models on the presymptomatic mutation carriers and controls' data using cross-validation. We shall refer to these models as 'carrier-control models'. MRI preprocessing, feature selection and classification were performed identically to previous work (Bouts et al., 2018).

Participants

This retrospective study partially included previously published (Dopper et al., 2014; Jiskoot et al., 2016; Papma et al., 2017) and newly acquired data from the Erasmus Medical Centre and Leiden University Medical Centre. Participants and clinical investigators were blinded to the participants' DNA status. The study was conducted in accordance with regional regulations and the Declaration of Helsinki. The Erasmus Medical Centre and Leiden University Medical Centre local medical ethics committees approved the study, and every participant provided written informed consent.

For the current study, we included 55 presymptomatic FTD mutation carriers (8 *MAPT*, 35 *GRN*, and 12 *C9orf72* repeat expansion) and 48 healthy familial controls (6 *MAPT* family, 31 *GRN* family, and 11 *C9orf72* family) between May 2010 and March 2016. These subjects were recruited from a cohort of healthy first-degree relatives of FTD patients with either an *MAPT* mutation, *GRN* mutation, or *C9orf72* repeat expansion (FTD-Risk Cohort; FTD-RisC), and visited the Erasmus Medical Centre for a one-day assessment in order to ascertain asymptomatic status, collect clinical data, and determine DNA status as described before (Dopper et al., 2014; Jiskoot et al., 2016; Papma et al., 2017). Participants were considered asymptomatic in the absence of (1) behavioural, cognitive, or neuropsychiatric change reported by the participant or knowledgeable informant, (2) cognitive disorders on neuropsychiatric tests, (3) motor neuron disease signs on neurologic examination, and (4) other FTD (Gorno-Tempini et al., 2011; Rascovsky et al., 2011) or amyotrophic lateral sclerosis (Ludolph et al., 2015) criteria. Healthy controls were assumed to have equal FTD risk as the general population. For a more detailed description of the recruitment protocol, see earlier work (Dopper et al., 2014; Jiskoot et al., 2016; Papma et al., 2017). Inclusion criteria for the current study were: age between 40 and 70 years, and availability of a 3-dimensional T_1-weighted MRI ($3DT_1w$) scan, a diffusion-weighted imaging (DWI) data set, and a resting-state functional T_2*-weighted MRI (rs-fMRI) scan. Exclusion criteria were: current or past neurologic or psychiatric disorders, history of drug abuse, large image artefacts, and gross brain pathology other than atrophy.

For details on the sample on which the bvFTD model was trained, please refer to Bouts et al. (2018). In short, 23 bvFTD patients and 35 controls between 40 and 80 years old were included to undergo a clinical assessment and MRI between November 2009 and November 2012. The MRI acquisition protocol was similar to the protocol applied in the current sample of mutation carriers and controls. Image processing steps were identical to processing steps in the current sample.

MRI data acquisition

All subjects were scanned at the Leiden University Medical Centre using a 3 T MRI scanner (Achieva, Philips Medical Systems, Best, The Netherlands) with an 8-channel SENSE head coil. The imaging protocol included a whole-brain near-isotropic $3DT_1w$ sequence for cortical and subcortical tissue type segmentation, a DWI sequence for assessments of white matter diffusion metrics, and an rs-fMRI for the calculation of functional connectivity measures. Participants were instructed to lie still with their eyes closed and not to fall asleep during rs-fMRI. For scan parameters, see **Table 2.1**.

Table 2.1 MRI sequence parameter settings

	$3DT_1w$	DWI[a]	rs-fMRI
Slices, *n*	140	70	38
TR, ms	9.8	8,250	2,200
TE, ms	4.6	80	30
Flip angle, °	8	90	80
Matrix, mm	256 × 256	128 × 128	80 × 80
Voxel size, mm	0.88 × 0.88 × 1.20	2.00 × 2.00 × 2.00	2.75 × 2.75 × 2.99[b]
Duration, min	4.57	8.48	7.28

Scan protocol of whole-brain near-isotropic 3-dimensional T_1-weighted ($3DT_1w$), diffusion-weighted imaging (DWI), and resting-state functional T_2*-weighted MRI (rs-fMRI) on a 3 T scanner at the Leiden University Medical Centre.
TR, repetition time; TE, echo time.
[a] 60 directions, b = 1,000, one b0 image.
[b] Including 10% interslice gap.

Image preprocessing

For $3DT_1w$ images, the following preprocessing steps were performed: bias field correction (N4ITK; Tustison et al., 2010), brain extraction (FSL BET; Smith, 2002), nonlinear registration to the MNI152 2 × 2 × 2 mm T1 template (FNIRT; Anderson et al., 2007), tissue type segmentation (SPM12; Friston et al., 2007) and segmentation of deep grey matter structures, including the bilateral thalamus, caudate nucleus, putamen, globus pallidum, nucleus accumbens, amygdala, and hippocampus (FIRST; Patenaude et al., 2011).

Preprocessing for DWI data sets included correction of motion and eddy-current induced distortion (eddy correct; Leemans & Jones, 2009), and calculation of voxel-wise measures of fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AxD, largest eigenvalue), and radial diffusivity (RD, average of the two remaining eigenvalues, DTIFIT; Smith et al., 2004). A global mean FA image was created by nonlinearly registering FA maps to the FMRIB58_FA template, and tract-based spatial statistics (FSL TBSS; Smith et al., 2006) was used to extract FA, MD, AxD, and RD values using the standard FSL TBSS skeleton. The skeleton was thresholded at 0.2 to ensure skeleton extracted values originate from white matter.

For rs-fMRI data, preprocessing included motion correction (Jenkinson et al., 2002), brain extraction, spatial smoothing using a Gaussian kernel with a full width at half maximum of 3 mm, grand mean intensity normalisation, motion artefact removal, and high-pass temporal filtering (cut-off frequency = 0.01 Hz). Motion artefacts were removed using a single-session independent

component analysis (ICA) to decompose the rs-fMRI data into distinct statistically independent components. Subsequently, motion-related components were automatically identified and removed using the ICA-based automatic removal of motion artefacts (ICA-AROMA, version 0.3-beta; Pruim et al., 2015) procedure. Registration to standard space was performed in two steps. First, a temporal mean image calculated from the 4D rs-fMRI volume was registered to the $3DT_1$w image using boundary-based registration (Greve & Fischl, 2009). Next, resulting registration parameters were concatenated to the $3DT_1$w-to-MNI152 template registration parameters to obtain the final registration parameters.

All registration and segmentation steps were critically reviewed and errors were corrected accordingly.

Feature selection

Cortical grey matter density (GMD) and white matter density (WMD) were calculated as a weighted average of their respective regional white matter or grey matter probability (SPM segmentation) weighted by the probability of a voxel being part of that specific tract or region. The latter probabilities were derived from the 48 Harvard-Oxford probabilistic anatomical brain atlas cortical regions (split into left and right) and from the Johns Hopkins University white matter tractography atlas for 20 white matter tract regions. Voxels with region probability values less than 25% were excluded. This provided a measure of brain atrophy of a specific grey matter region or white matter tract. For deep grey matter regions, GMD values were calculated as the regions' volume (FIRST segmentations) divided by total intracranial volume. This resulted in a feature vector of 110 average GMD values (48 left cortical, 48 right cortical, and 14 deep grey matter regions), and a feature vector of 20 average WMD values per subject.

DTI-based features were calculated by projecting each subject's FA, MD, AxD, and RD values onto the TBSS group skeleton on a voxel-wise basis. Like the WMD features, the 20 white matter tracts of the probabilistic Johns Hopkins University white matter tractography atlas were then used to calculate a weighted mean value per tract per subject. This resulted in 4×20 feature vectors of mean FA, MD, AxD, and RD values per subject.

In order to calculate the functional connectivity features, all processed rs-fMRI images were combined in a temporally concatenated ICA (Beckmann & Smith, 2004), with dimensionality fixed at 70 components and an ICA threshold of 0.99 (Smith et al., 2013). This meant that each voxel included in the ICA map was 99 times more likely to be part of that component than to be caused by Gaussian background noise. For each subject, we calculated the mean time course for each component, weighted by the ICA weight map and grey matter probability of that component's region. These mean time courses were subsequently used to determine the functional connectivity of a component with the 69 other components. Functional connectivity was either expressed as full correlations (FCor) or as sparse, L1-regularised, partial correlations (PCor) between the components' time courses. Partial correlations were calculated using the graphical lasso algorithm (J. Friedman et al., 2008). The functional connectivity measures resulted in two feature vectors of each $(70 \times 69) \div 2 = 2{,}415$ (partial) correlations per subject. Finally, we concatenated all feature vectors into one vector per subject.

BvFTD model

For our first analysis, a bvFTD patient-control classification model (Bouts et al., 2018) was applied to each subject's extracted feature vector. We applied the best performing, multimodal model that

discriminated bvFTD patients from controls, which included the features GMD, FCor, FA, and MD (Bouts et al., 2018), as well as age and sex. Each subject's feature vector was fed into the model, resulting in a classification score from 0 to 1, where 0 represents a control subject and 1 represents a bvFTD patient. Extrapolated to our subjects, these scores showed how alike our presymptomatic FTD mutation carriers and healthy controls are to bvFTD patients.

Carrier-control models

For the second analysis, feature vectors were used to train a logistic elastic net regression algorithm (Zou & Hastie, 2005; Friedman et al., 2010; Schouten et al., 2016; Bouts et al., 2018). The elastic net regression procedure estimates a sparse regression model that includes only a subset of the provided features by imposing a penalty for including features and for the weight of each feature. This way, elastic net provides a solution for the imbalance between the large number of features and the small number of subjects. Age and sex were included into the model without penalty to ensure that estimated feature regression coefficients were conditional on subject age and sex. Here, a classification score of 0 represented a control subject and 1 represented a presymptomatic FTD mutation carrier.

Cross-validation

Similarly to previous work (Schouten et al., 2016; Bouts et al., 2018), we trained our carrier-control model in a nested 10-fold cross-validation scheme to reduce classification bias. One part of the data (e.g., 10%) was set apart as a test set and served to test the generalised classification performance of the elastic net regression model. The remaining parts (90%) were used to train the model. However, in addition to the classification performance, we also wanted to determine the optimal penalty size without overestimating classification performance (Varma & Simon, 2006; Kriegeskorte et al., 2009). To this end, we used a second, nested, 10-fold cross-validation loop on the training set over a grid of hyperparameters to determine the optimal penalty. In the nested loop, we estimated the model's hyperparameters that corresponded with the lowest binomial deviance, a goodness-of-fit measure that evaluates the difference between the predicted and actual observations. Next, these hyperparameters and corresponding penalties were used to train a model using the training set of the outer loop. Finally, the classification performance was tested on the test set of the outer loop. This process was repeated ten times to make sure that each subject was part of the test set at least once. Since the test set of the outer loop was neither used for model training, nor for parameter optimisation, potential prediction bias was reduced as much as possible (Kriegeskorte et al., 2009). The entire classification procedure was repeated 50 times to average classification outcome variability resulting from random partitioning in training and test folds. All classification analyses and evaluations were implemented in R version 3.3.2 (R core 2016, GLMnet package; Friedman et al., 2010).

Classification performance

For both analyses, we quantified classification performances using receiver operating characteristic (ROC) curves. ROC curves were calculated by shifting the threshold for classifying an individual as patient (bvFTD model analysis) or mutation carrier (carrier-control model analysis) from 0 to 1, and plotting the true positive rate (sensitivity) vs. the false positive rate (1 – specificity) for each intermediate point. The area under this ROC curve (AUC) is a measure of classification performance insensitive to the distribution between the groups (Fawcett, 2006). Additionally, we

calculated the optimal operating point on the curve to calculate the model's sensitivity, specificity, and classification accuracy, given equal class distribution, and equal penalty for false positive and false negative predictions. For the carrier-control model analysis, we averaged AUC, accuracy, sensitivity, and specificity values from the 50 times repeated nested cross-validations.

Multimodal classification

To obtain the best multimodal carrier-control model using several feature vectors, we performed step-wise feature concatenation as previously described (Schouten et al., 2016; Bouts et al., 2018). First, we assessed classification performance for each feature separately. Subsequently, we added a new feature to the best performing feature combination (i.e., highest AUC) of the previous step until all features were included in the model. The best performing feature combination will be referred to as the multimodal carrier-control model.

Statistical analysis

Statistical analyses of non-imaging data were performed using R (R Core 2016, Vienna, Austria). We tested for carrier-control differences using independent *t*-tests (age and education), the Mann-Whitney U test (mini-mental state examination [MMSE] scores [0–30]) and the chi-square test (sex distribution). Classification scores were compared using Mann-Whitney U tests for overall carrier-control contrasts, and Kruskal Wallis H tests and Dunn post-hoc tests for comparisons between all four groups (*MAPT*, *GRN*, *C9orf72*, and controls). To compare models' AUC values against chance level, we used permutation tests (N = 5,000; Noirhomme et al., 2014). In order to correct for multiple comparisons, we took the maximum AUC difference of the family of tests for each permutation. Then we compared the observed AUC difference to the new distribution of maximum AUC differences to get a family-wise error rate corrected p-value. The alpha level required for statistical significance was set at 0.05.

Results

Demographics

In total, 103 subjects met the inclusion criteria (**Table 2.2**). Mean age was similar for mutation carriers (52.0 ± 8.6 years) and healthy controls (54.2 ± 7.5 years). The proportion of female participants between mutation carriers (67%) and healthy controls (58%) was not different (p = 0.35). Education level was similar between groups (mutation carriers, 13.6 ± 2.9 years; healthy controls, 13.2 ± 2.4 years). MMSE was similarly distributed between groups (median [min–max], mutation carriers: 30 [24–30], healthy controls: 29 [24–30]).

Table 2.2 Participant demographics

	Mutation carriers (n = 55)	**Controls (n = 48)**	***p*-value**
Gene, *n MAPT* / *GRN* / *C9orf72*	8 / 35 / 12	–	–
Age, mean (SD) years	52.0 (8.6)	54.2 (7.5)	0.18
Sex, *n* (%) ♀	37 (67%)	28 (58%)	0.35
Education, mean (SD) years[a]	13.6 (2.9)	13.2 (2.4)	0.42
MMSE, median (range) points	30 (24–30)	29 (24–30)	0.53

C9orf72, chromosome 9 open reading frame 72; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; MMSE, mini-mental state examination.

[a] Education values were missing for one mutation carrier and two controls.

BvFTD model

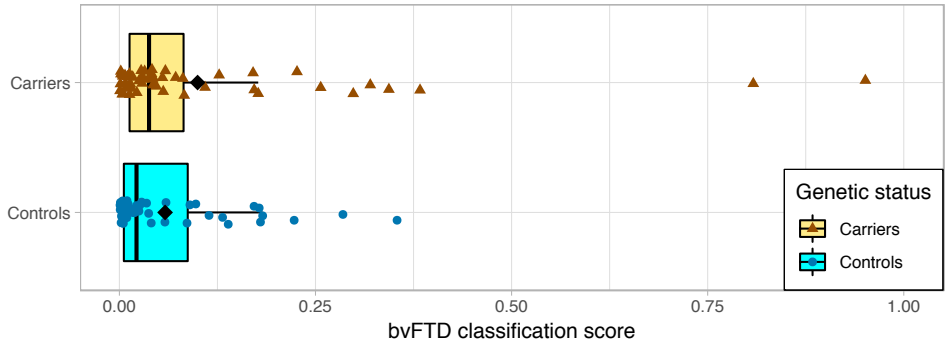
Application of the bvFTD model resulted in low bvFTD classification scores for most subjects (**Figure 2.1A**), and the bvFTD classification scores were not significantly different between presymptomatic mutation carriers (median = 0.038) and controls (median = 0.022, p = 0.15). ROC analysis of the bvFTD classification scores resulted in an AUC of 0.582, which was not significantly better than chance level (p = 0.078). Separated by gene (**Figure 2.1B**), there were no differences between the four groups' bvFTD classification scores (p = 0.37). BvFTD classification scores of the original patients and controls used for cross-validation of the bvFTD model were added as reference (**Figure 2.1C**, data courtesy of Bouts et al., 2018).

Figure 2.1 Classification results bvFTD model (opposite page)

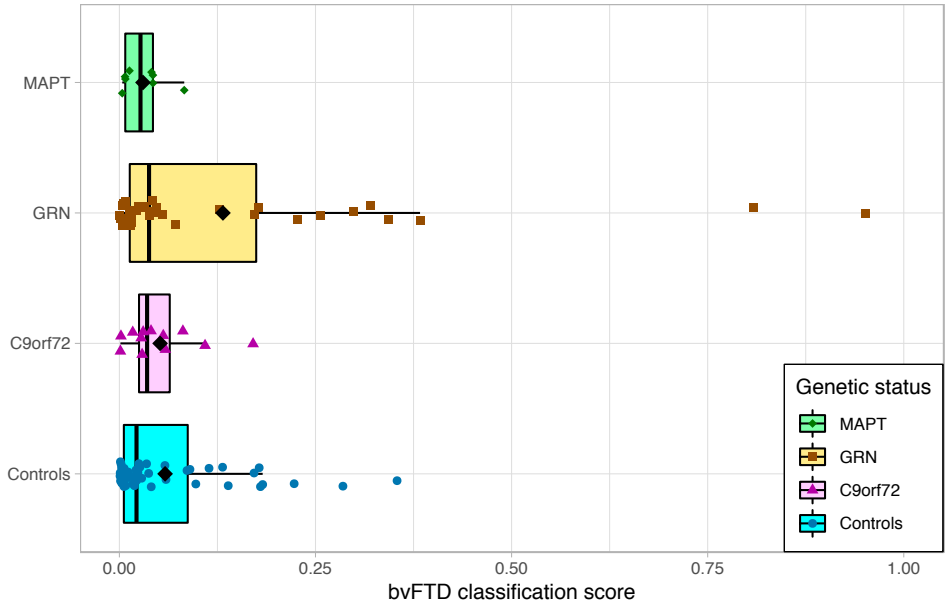
Box and scatter plot of each subject's bvFTD classification score on a scale from 0 (representing control) to 1 (representing bvFTD patient) after application of the bvFTD model. Groups are defined by mutation carrier status (**A**) and genetic status (**B**). Classification scores were not significantly different for mutation carriers and controls (p = 0.15), and did not differ between the four genetic groups (p = 0.37). Classification score results of the bvFTD patients and controls on which the bvFTD model was trained were added for reference (**C**, data courtesy of Bouts et al., 2018).

BvFTD, behavioural variant frontotemporal dementia; *C9orf72*, chromosome 9 open reading frame 72; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau.

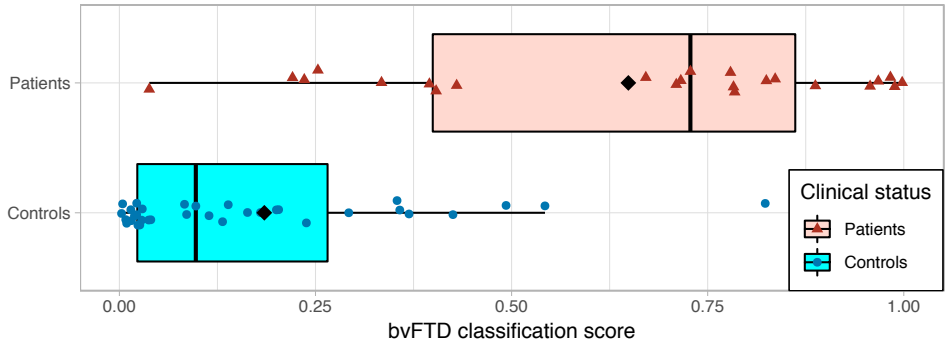
A. Classification by carrier status



B. Classification by genetic status



C. Classification by disease status



Carrier-control models

The best performing unimodal carrier-control models included RD, WMD, and MD, with AUCs of 0.642, 0.592, and 0.587, respectively. Of these models, only the RD model outperformed chance after family-wise error rate correction ($p = 0.032$, **Table 2.3**). Step-wise concatenation resulted in the best performing multimodal model, which included the features RD and WMD, and outperformed chance with an AUC of 0.684 ($p = 0.004$). Classification performance did not improve when additional features were added to this model (**Table 2.4**). Interestingly, all multimodal models that outperformed chance level used exclusively white matter features from diffusion-weighted and/or structural scans: RD, WMD, MD, and AxD).

Application of the best performing multimodal carrier-control model resulted in mutation carrier classification scores (**Figure 2.2**), which were different between mutation carriers (median = 0.589) and controls (median = 0.435, $p < 0.001$, **Figure 2.2A**). Furthermore, there was a difference between the four groups' mutation carrier classification scores ($p = 0.008$) when separated by gene (**Figure 2.2B**). Post-hoc tests revealed that *GRN* mutation carriers had higher mutation carrier classification scores than controls (Bonferroni family-wise error rate corrected $p = 0.009$). The other groups did not differ from each other.

Table 2.3 ROC characteristics

Modality	AUC	Min – max	Sensitivity	Specificity	Accuracy	FWERC *p*-value (AUC > chance)
GMD	0.502	0.427 – 0.554	0.538	0.514	0.554	0.895
WMD	0.592	0.547 – 0.656	0.595	0.585	0.663	0.190
FA	0.494	0.408 – 0.554	0.528	0.504	0.535	0.897
MD	0.587	0.523 – 0.629	0.591	0.579	0.642	0.240
AxD	0.554	0.500 – 0.598	0.568	0.559	0.578	0.448
RD	*0.642*	*0.583 – 0.673*	*0.637*	*0.629*	*0.669*	*0.032*
FCor	0.509	0.454 – 0.554	0.537	0.527	0.590	0.818
PCor	0.505	0.457 – 0.551	0.542	0.492	0.540	0.836
Multimodal	*0.684*	*0.629 – 0.722*	*0.664*	*0.640*	*0.735*	*0.004*

Presymptomatic mutation carriers vs. controls classification. Multimodal represents the best combination from our step-wise multimodal procedure (i.e., RD and WMD). Italic: mean AUC significantly higher than chance level after family-wise error rate correction.

AUC, area under the receiver operating characteristic curve; AxD, axial diffusivity; FA, fractional anisotropy; FCor, full correlations between ICA components; FWERC, family-wise error rate corrected; GMD, grey matter density; ICA, independent component analysis; MD, mean diffusivity; PCor, L1-regularised partial correlations between ICA components; RD, radial diffusivity; WMD, white matter density.

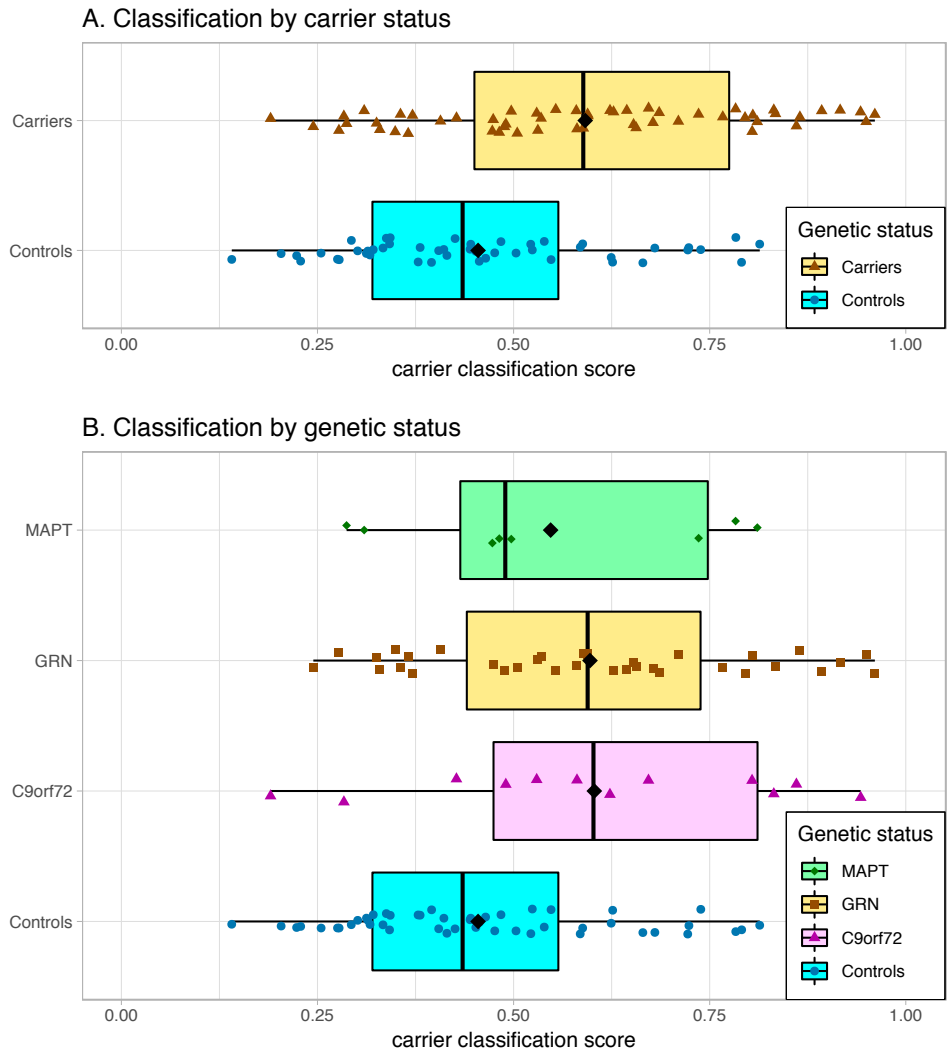


Figure 2.2 Classification results carrier-control model

Box and scatter plot of each subject's mutation carrier classification score on a scale from 0 (representing control) to 1 (representing presymptomatic mutation carrier) after application of the best performing carrier-control model including the features RD and WMD. Mutation carriers had significantly higher scores than controls (**A**, $p < 0.001$). Furthermore, there was an omnibus difference between the four genetic groups (**B**, $p = 0.008$), and post-hoc tests revealed higher scores for *GRN* mutation carriers than for controls ($p = 0.009$).

C9orf72, chromosome 9 open reading frame 72; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; RD, radial diffusivity; WMD, white matter density.

Table 2.4 Multimodal classification performance

Step\combined with:	**RD**	**WMD**	**MD**	**AxD**	**GMD**	**FA**	**FCor**	**PCor**
1: –	*0.642*	0.592	0.587	0.554	0.502	0.494	0.509	0.505
2: RD	–	***0.684***	0.621	0.600	0.565	0.564	0.516	0.501
3: RD + WMD	–	–	*0.658*	*0.640*	0.596	0.607	0.520	0.501
4: RD + WMD + MD	–	–	–	0.627	0.612	0.614	0.524	0.502
5: RD + WMD + MD + AxD	–	–	–	–	0.612	0.609	0.531	0.507
6: RD + WMD + MD + AxD + GMD	–	–	–	–	–	0.585	0.534	0.509
7: RD + WMD + MD + AxD + GMD + FA	–	–	–	–	–	–	0.530	0.500
8: RD + WMD + MD + AxD + GMD + FA + FCor	–	–	–	–	–	–	–	0.510

Mean AUC values from 50 repetitions. Multimodal models result from step-wise addition of measures to the best performing classification model of the previous step, starting with the best performing single measure (i.e., RD). Bold: best performing model. Italic: mean AUC significantly higher than chance level after family-wise error rate correction.

AUC, area under the receiver operating characteristic curve; AxD, axial diffusivity; FA, fractional anisotropy; FCor, full correlations between ICA components; GMD, grey matter density; ICA, independent component analysis; MD, mean diffusivity; PCor, L1-regularised partial correlations between ICA components; RD, radial diffusivity; WMD, white matter density.

Discussion

This study investigated whether presymptomatic FTD mutation carriers with mutations in *MAPT*, *GRN*, or with a *C9orf72* repeat expansion can be individually distinguished from healthy controls using MRI. Using a recently introduced MRI-based classification model trained on established bvFTD patients and controls, nearly all FTD mutation carriers and controls had low classification scores. The bvFTD model was therefore not able to separate mutation carriers from controls beyond chance level. However, MRI-based classification models that were trained on our own sample were able to separate mutation carriers from controls better than chance level. In our carrier-control model, the RD feature proved sufficient to separate mutation carriers from controls better than chance, but the best performing model used RD in combination with the WMD feature. All models that outperformed chance used exclusively white matter features, such as DTI features and WMD, supporting the hypothesis that white matter alterations are the first to appear in preclinical FTD pathology.

In an effort to improve on the FTD diagnostic criteria, single-subject classification using MRI measures has recently received significant attention (Davatzikos et al., 2008; McMillan et al., 2014; Raamana et al., 2014; Möller et al., 2015b; Koikkalainen et al., 2016; Wang et al., 2016; Bron et al., 2017; Meyer et al., 2017; Bouts et al., 2018). A recent multimodal classification study incorporated structural, DTI, and arterial spin labelling data to classify FTD (behavioural and language variants) from cognitively normal controls, and achieved an AUC of 0.96 (Bron et al., 2017). Another classification study included tissue density, DTI, and rs-fMRI measures, and achieved an AUC of 0.92 for bvFTD vs. cognitively normal controls (Bouts et al., 2018). These high classification performances are promising, but they are based on established FTD cases. It is unclear how FTD patient models generalise to earlier FTD stages, where brain alterations are less distinct. To test this, we applied a bvFTD model (Bouts et al., 2018) on FTD mutation carriers in a presymptomatic stage. We hypothesised that if the bvFTD model would be able to recognise early-stage FTD pathology, our presymptomatic FTD mutation carriers would have higher classification scores than controls. We found that it was not possible to separate mutation carriers from controls significantly better than chance using this model, as most mutation carriers and controls had very low bvFTD classification scores. This could indicate that presymptomatic differences present in FTD mutation carriers are too subtle to be picked up by a classification model that was trained established bvFTD patients. However, it could also mean that most of our mutation carriers were still too far from conversion to have significant FTD-related changes. Since the bvFTD model was trained on patients, it stands to reason that classification of mutation carriers and controls becomes more accurate as mutation carriers approach conversion. Vice versa, one might expect the mutation carriers with high classification scores to be closer to symptom onset than mutation carriers with lower classification scores. Although it was not statistically significant, there was a trend towards older age in mutation carriers with a bvFTD classification score higher than 0.25 than in the rest of the mutation carrier group (data not shown). It can therefore not be entirely ruled out that age is partly associated with a higher bvFTD score. Longitudinal research is warranted to formally test whether this model captures presymptomatic FTD-related changes as mutation carriers approach conversion.

By training classifiers on presymptomatic FTD mutation carriers and controls, we obtained a unimodal carrier-control model based on the RD feature and several multimodal carrier-control models that significantly outperformed chance level. This suggests that classification models should be trained using early-stage FTD patients or presymptomatic FTD mutation carriers instead of

advanced FTD cases, in order to be sensitive to early-stage FTD pathology. Furthermore, our carrier-control models demonstrate that MRI-based machine learning is powerful enough to detect subtle pathological changes associated with FTD even before symptom onset and on a single-subject level. Although classification performance beyond chance level is an important finding, it must be noted that AUCs of 0.642 and 0.684 are modest and far from sufficient for diagnostic use in the clinic. This is at least partly explained by our heterogeneous sample, as we included mutation carriers of several genes in order to obtain sufficient sample size for robust cross-validation. Heterogeneity further arose from the uncertain time to onset in our sample. Investigating a uniform population a few years before symptom onset might lead to higher classification performance, but these data were not available to us.

On a pathological level, it has been argued that neurodegeneration in FTD starts in the white matter (Rohrer et al., 2013; Suri et al., 2014; Möller et al., 2016; Canu et al., 2017). Our results support this hypothesis, as the only unimodal model that outperformed chance was based on the RD feature, which was furthermore included in all multimodal models that significantly outperformed chance. Additionally, all models that outperformed chance level were based on DTI features, while three models also included WMD. This means that our carrier-control model was able to combine subtle white matter differences from the diffusion-weighted scans and the structural $3DT_1w$ scan to classify a subject as mutation carrier or healthy control. In our sample, inclusion of GMD and functional connectivity features did not aid classification, but in fact drove down classification performance. For instance, adding FCor or PCor to the model resulted in classification scores near 0.5. This reinforces the notion that feature selection is important for MRI-based machine learning.

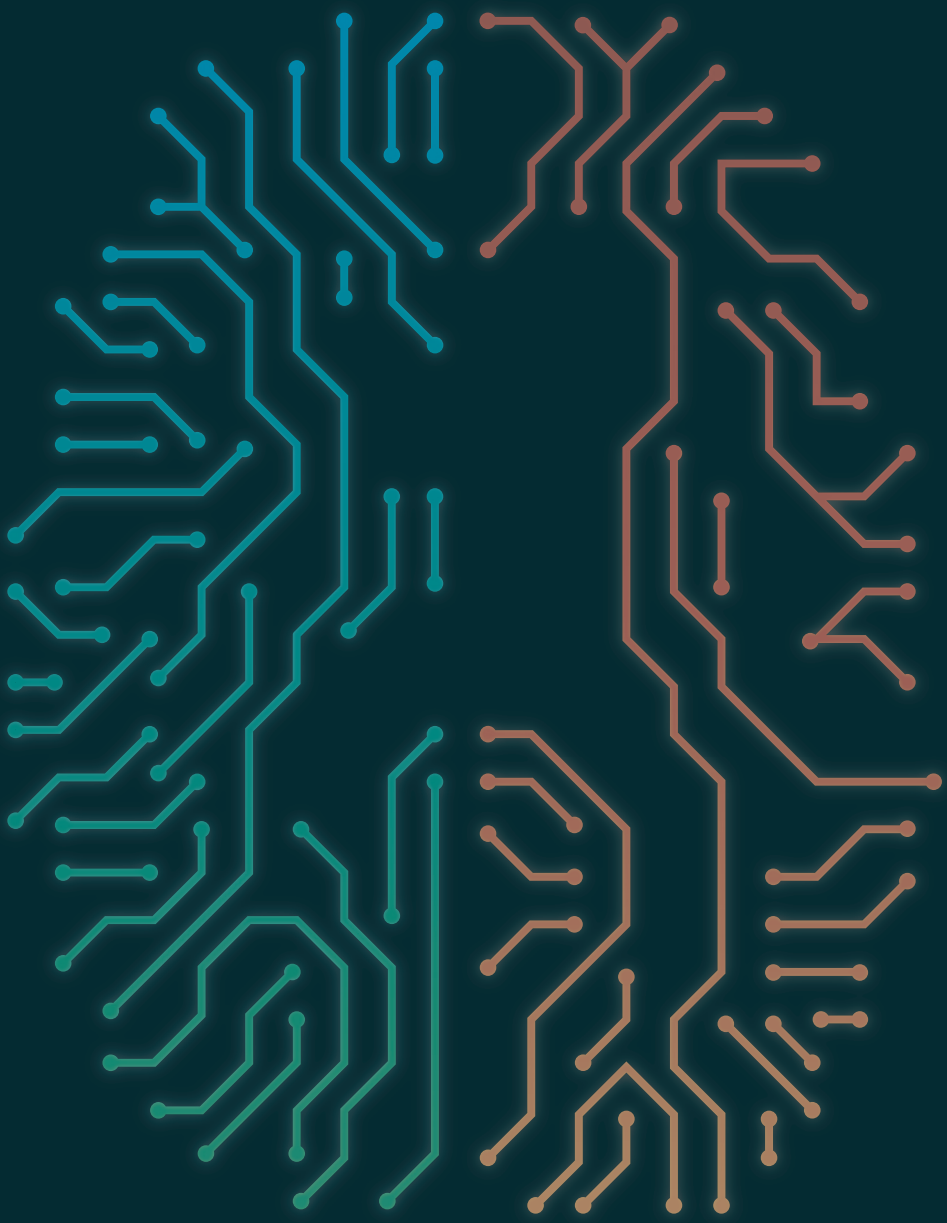
In addition to the uncertain time to onset, there were several other limitations. Firstly, the bvFTD model was trained on a relatively small sample of 23 bvFTD patients and 35 controls. A model based on a larger sample might capture the heterogeneity of bvFTD pathology more completely, which could benefit generalisation to our presymptomatic sample. Furthermore, the model was trained on sporadic bvFTD patients, while it was applied to *MAPT* mutation carriers, *GRN* mutation carriers, and *C9orf72* repeat expansion carriers. Since correlations between genetics, pathology and phenotype are not fully elucidated (Mann & Snowden, 2017), care must be taken not to overinterpret our results. Specifically, pathological changes associated with non-behavioural variants (Seelaar et al., 2011) may be insufficiently recognised by the bvFTD model. Lastly, we used nested cross-validation to estimate out-of-sample performance for the carrier-control model, which minimises prediction bias (Kriegeskorte et al., 2009). Still, measuring performance on a separate validation cohort would further increase the validity of this study.

Conclusion

Our data show that presymptomatic FTD mutation carriers can be distinguished from healthy controls on an individual level using a new multimodal MRI-based carrier-control classification model, while this was not possible using a recent bvFTD classification model. A multimodal MRI-based classification score may therefore be a useful biomarker to aid earlier FTD diagnosis. Successful single-subject recognition of early-stage or presymptomatic FTD may facilitate more precise subject recruitment into clinical trials. Furthermore, our multimodal MRI-based carrier-control classification model supports the hypothesis that FTD-related neurodegeneration starts in the white matter.

Acknowledgements

Our appreciation goes out to all FTD-RisC participants and families. The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to RAF), ZonMw programme Memorabel project 733050103, JPND PreFrontAls consortium project 733051042 (to JCvS), and a VICI grant 016–130–667 from The Netherlands Organisation for Scientific Research (NWO; to SARBR). The views expressed are those of the authors and not necessarily those of the funding sources. The funding sources were not involved in the design of the study; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The authors report no conflict of interest.



Chapter 3

Frontotemporal dementia mutation carriers: longitudinal effects in MRI-based classification scores

Published in *Journal of Neurology, Neurosurgery & Psychiatry 2019; 90(11):1207–14* as:

A multimodal MRI-based classification signature emerges just prior to symptom onset in frontotemporal dementia mutation carriers

Rogier A. Feis, Mark J.R.J. Bouts, Frank de Vos, Tijn M. Schouten, Jessica L. Panman, Lize C. Jiskoot, Elise G.P. Dopper, Jeroen van der Grond, John C. van Swieten, Serge A.R.B. Rombouts

Abstract

Multimodal magnetic resonance imaging (MRI-)based classification may aid early frontotemporal dementia (FTD) diagnosis. Recently, presymptomatic FTD mutation carriers, who have a high risk of developing FTD, were separated beyond chance level from controls using MRI-based classification. However, it is currently unknown how these scores from classification models progress as mutation carriers approach symptom onset. In this longitudinal study, we investigated multimodal MRI-based classification scores between presymptomatic FTD mutation carriers and controls. Furthermore, we contrasted mutation carriers that converted during follow-up ('converters') and non-converting mutation carriers ('non-converters').

We acquired anatomical MRI, diffusion-weighted imaging, and resting-state functional MRI in 55 presymptomatic FTD mutation carriers and 48 healthy controls at baseline, and at two, four, and six years follow-up as available. At each time point, FTD classification scores were calculated using a behavioural variant FTD classification model. Classification scores were tested in a mixed effects model for mean differences and differences over time.

Presymptomatic mutation carriers did not have a stronger classification score increase over time than controls ($p = 0.15$), although mutation carriers had higher FTD classification scores than controls on average ($p = 0.032$). However, converters ($n = 6$) showed a stronger classification score increase over time than non-converters ($p < 0.001$).

Our findings imply that presymptomatic FTD mutation carriers may remain similar to controls in terms of MRI-based classification scores until they are close to symptom onset. This proof-of-concept study shows the promise of longitudinal MRI data acquisition in combination with machine learning to contribute to early FTD diagnosis.

Keywords: frontotemporal dementia; *MAPT* protein, human; *GRN* protein, human; *C9orf72*, human; diffusion tensor imaging; resting-state functional MRI; multimodal MRI; classification; machine learning

Introduction

Magnetic resonance imaging (MRI-)based classification is a novel tool that may aid early frontotemporal dementia (FTD) diagnosis. Classification algorithms using structural T_1-weighted scans (Raamana et al., 2014; Meyer et al., 2017), or multimodal imaging (Klöppel et al., 2015; Koikkalainen et al., 2016; Bron et al., 2017; Canu et al., 2017; Bouts et al., 2018) accurately distinguish FTD patients from controls. However, detection of FTD patients in earlier disease stages is necessary to further advance future treatment through increasing recruitment accuracy and clinical trial efficiency (Tsai & Boxer, 2016).

To detect FTD-related neuropathology in an earlier stage, we applied MRI-based classification on a sample of presymptomatic FTD mutation carriers, who have a nearly full risk of developing FTD, and controls (Feis et al., 2019a). We found that it was possible to separate presymptomatic FTD mutation carriers from healthy controls beyond chance level, showing the promise of MRI-based machine learning to detect early-stage FTD-related changes before symptom onset in a single-subject setting. However, classification performance was modest, partly due to the cross-sectional design of the study. Specifically, the time to symptom onset in presymptomatic mutation carriers was unknown and mutation carriers who were still far from conversion may have had less FTD-related pathological changes than those who were close to symptom onset. To understand how classification scores develop over time, it is therefore necessary to apply classification models on longitudinal MRI data in presymptomatic FTD populations.

Here, we use longitudinal data of presymptomatic FTD mutation carriers and healthy controls to relate FTD disease progression with classification scores from an MRI-based classification model. First, we explore the progression of FTD classification scores over time between presymptomatic FTD mutation carriers and controls. Secondly, we contrast classification score trajectories of mutation carriers that developed symptoms of FTD during the course of the study ('converters') and mutation carriers who did not ('non-converters'). Our aim is to test whether the MRI-based classification model is sensitive to early FTD-related changes in genetic FTD that occur before and around symptom onset.

Methods

Participants

The FTD-Risk Cohort (FTD-RisC; Dopper et al., 2014; Jiskoot et al., 2016, 2018b, 2019; Papma et al., 2017; Feis et al., 2019a) is a longitudinal study that follows healthy, 50% at-risk family members of genetic FTD patients on a 2-year basis. Subjects were assigned to mutation (microtubule-associated protein tau [*MAPT*], progranulin [*GRN*], or chromosome 9 open reading frame 72 [*C9orf72*] repeat expansion) carrier or control groups by genotyping, as described in previous work (Dopper et al., 2014; Jiskoot et al., 2016). For the current study, we included the baseline data of 55 presymptomatic FTD mutation carriers (eight *MAPT*, 35 *GRN*, and 12 *C9orf72*) and 48 healthy familial controls (six *MAPT* family, 31 *GRN* family, and 11 *C9orf72* family), who entered FTD-RisC between May 2010 and March 2016 and were between 40 and 70 years old at inclusion (Feis et al., 2019a). At study entry, all subjects were asymptomatic in accordance with established diagnostic criteria for behavioural variant FTD (bvFTD; Rascovsky et al., 2011), primary progressive aphasia (PPA; Gorno-Tempini et al., 2011), and amyotrophic lateral sclerosis (Ludolph et al., 2015). We included follow-up data as available at 2, 4, and 6 years (**Figure 3.1**). Exclusion criteria were: current or past neurologic or primary psychiatric disorders, history of drug abuse, large image artefacts, and gross brain pathology other than atrophy. As such, we included 103 subjects at baseline, 98 subjects after two years follow-up, 56 subjects after four years follow-up, and 43 subjects after six years follow-up (**Figure 3.1**). However, 20 subjects had incomplete data (n = 10) or artefacts (n = 10) at a certain time point, leaving 55 mutation carriers and 48 controls at baseline, 45 mutation carriers and 42 controls after two years follow-up, 27 mutation carriers and 23 controls at four years follow-up, and 20 mutation carriers and 20 controls at six years follow-up.

For details on the sample on which the FTD classification model was trained, please refer to Bouts et al. (2018). In short, 23 sporadic bvFTD patients and 35 controls between 40 and 80 years old were included to undergo a clinical assessment and MRI between November 2009 and November 2012. The MRI acquisition protocol was similar to the protocol applied in the current sample of mutation carriers and controls. Imaging processing steps were identical to processing steps in the current analysis.

Standard protocol approvals, registrations, and patient consents

Participants and clinical investigators were blinded to the participants' genetic status, except for those that underwent predictive testing at their own request. For converters, genetic counselling was offered to the patient and family members, and genetic status was unblinded to confirm the presence of the pathogenic mutation. The study was conducted in accordance with regional regulations and the Declaration of Helsinki. The Erasmus Medical Centre and Leiden University Medical Centre local medical ethics committees approved the study, and every participant provided written informed consent.

Conversion

Conversion was determined in a multidisciplinary consensus meeting of the Erasmus MC FTD Expertise Centre, involving neurologists (JCvS), neuropsychologists (JLP, LCJ), neuroradiologists, geriatricians, a clinical geneticist, and a care consultant. In the consensus meetings, anamnestic and heteroanamnestic information, neuropsychological assessment, and MRI of the brain were

reviewed, and genetic status was unblinded. For detailed information on conversion criteria, see Jiskoot et al. (2018b, 2019).

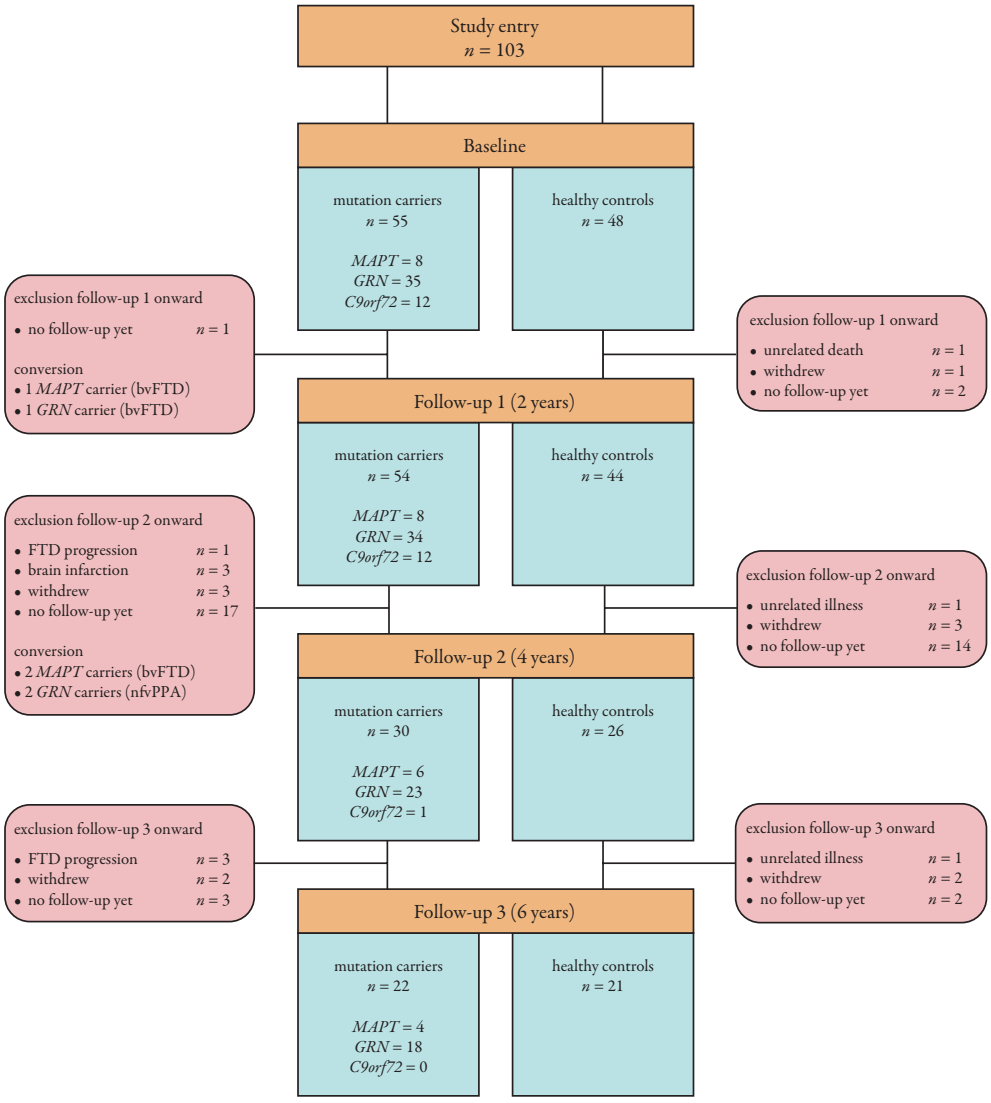


Figure 3.1 Study flow diagram
Blue fields show number of subjects included at each time point. Red fields show excluded subjects, reasons for exclusion, and conversion information.
C9orf72, chromosome 9 open reading frame 72; bvFTD, behavioural variant FTD; FTD, frontotemporal dementia; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; nfvPPA, non-fluent variant primary progressive aphasia.

MRI data acquisition

All subjects were scanned at the Leiden University Medical Centre using a 3 T MRI scanner (Achieva, Philips Medical Systems, Best, The Netherlands) with an 8-channel SENSE head coil. The imaging protocol included a whole-brain near-isotropic 3-dimensional T_1-weighted MRI ($3DT_1w$) sequence for cortical and subcortical tissue-type segmentation, a diffusion-weighted imaging (DWI) sequence for assessments of white matter diffusivity, and a resting-state functional MRI (rs-fMRI) scan for the calculation of functional connectivity measures. Participants were instructed to lie still with their eyes closed and not to fall asleep during rs-fMRI. For scan parameters, see **Table 3.1**. Although MRI sequence parameters were fixed over time, a routine MRI software upgrade was performed at our MRI site by the manufacturer in September 2015. We added a covariate to all our imaging analyses to correct for this.

Table 3.1 MRI sequence parameter settings

	$3DT_1w$	DWI[a]	rs-fMRI
Slices, *n*	140	70	38
TR, ms	9.8	8,250	2,200
TE, ms	4.6	80	30
Flip angle, °	8	90	80
Matrix, mm	256 × 256	128 × 128	80 × 80
Voxel size, mm	0.88 × 0.88 × 1.20	2.00 × 2.00 × 2.00	2.75 × 2.75 × 2.99[b]
Duration, min	4.57	8.48	7.28

Scan protocol of whole-brain near-isotropic 3-dimensional T_1-weighted ($3DT_1w$), diffusion-weighted imaging (DWI), and resting-state functional T_2*-weighted MRI (rs-fMRI) on a 3 T scanner at the Leiden University Medical Centre.
TR, repetition time; TE, echo time.
[a] 60 directions, b = 1,000, one b0 image.
[b] Including 10% interslice gap.

Image preprocessing

Image preprocessing was performed identically to earlier work (Feis et al., 2019a). For a detailed description, please see the Supplemental material. All registration and segmentation steps were critically reviewed and errors were corrected accordingly.

Feature selection

We selected MRI features for classification based on earlier work, in which bvFTD patients were separated from controls with high accuracy (Bouts et al., 2018). We selected the best performing combination of MRI features, including grey matter density (GMD), functional correlations (FCor), fractional anisotropy (FA), and mean diffusivity (MD). Features were calculated using the same procedures as in previous work (Schouten et al., 2016; Bouts et al., 2018; Feis et al., 2019a). For a detailed description, please see the Supplemental material. The GMD feature included 96 cortical and 14 subcortical values, the FCor feature consisted of 2,415 functional correlation values between 70 independent components, and the FA and MD features each contained 20 values.

Classification model application

We trained the bvFTD model on the bvFTD patients' and controls' features (i.e., GMD, FCor, FA, and MD features, as well as sex) using an elastic net logistic regression classifier (Bouts et al., 2018). Elastic net hyperparameters were optimised in a 10-fold cross-validation. Next, we applied the resulting bvFTD classification models on our longitudinal data of FTD mutation carriers and controls data. As such, we obtained a predicted probability at each time point for each subject, which we call bvFTD classification scores. Normally, these classification scores are transformed using a logit link function to range between 0 (representing control) and 1 (representing bvFTD patient). However, such a transformation introduces nonlinearity to the scores, which makes them unsuited for repeated measures analyses. Therefore, we use the raw scores, which in this case ranged between −8.9 and 8.4. Here, negative values represent controls and positive values represent bvFTD patients.

Statistical analysis

To investigate the bvFTD classification scores' progression over time, we used a repeated measures linear mixed effects model. Mixed effects models estimate both fixed and random effects of predictor variables, in order to determine global effects and account for between-subject variance, respectively (Laird & Ware, 1982; Sullivan et al., 1999). We used a random intercept to account for between-subject variation in classification score. Our main covariate of interest was the interaction between label and time point (i.e., the progression over time contrasted between mutation carriers and controls). Other covariates included time point, label, age at baseline, sex, and a covariate to correct for the MRI scanner software upgrade. Additionally, we added an interaction between label and age at baseline, to see whether the effect of age at baseline was different for mutation carriers and controls. Covariates were tested for significance using likelihood ratio chi-square tests. In the presymptomatic carrier-control analysis, we excluded data points from subjects who had at that time developed symptoms.

In a subgroup analysis within the mutation carrier group, we contrasted converters and non-converters. This subgroup analysis was performed using the same mixed model covariates as the carrier-control analysis and included presymptomatic as well as symptomatic data points.

Statistical analyses of baseline demographic variables included independent *t*-tests for age and education, a chi-square test for sex distribution, and a Mann-Whitney *U* test for mini-mental state examination (MMSE) scores (0–30). All statistical analyses were performed using R (R Core 2016, Vienna, Austria). For the mixed effects model analyses, we used the nlme package (version 3.1.128; Pinheiro et al., 2018).

Data availability

Raw data were generated at the Leiden University Medical Centre. The derived data, as well as scripts, that support the findings of this study are available from the corresponding author on reasonable request.

Results

Demographics

Presymptomatic mutation carriers and controls

At baseline, there were no differences between mutation carriers and controls in terms of age, sex, education, or MMSE (**Table 3.2**), nor were there differences at any follow-up time point. The proportion of scans acquired after the MRI software upgrade was installed varied per time point, but did not differ between mutation carriers and controls at each time point. The total follow-up time of all subjects was 376 years, and the average follow-up time was 3.7 years.

Table 3.2 Participant baseline demographics

	Mutation carriers (*n* = 55)	**Controls (*n* = 48)**	***p*-value**
Gene, *n MAPT* / *GRN* / *C9orf72*	8 / 35 / 12	–	–
Age, mean (SD) years	52.0 (8.6)	54.2 (7.5)	0.18
Sex, *n* (%) ♀	37 (67%)	28 (58%)	0.35
Education, mean (SD) years[a]	13.6 (2.9)	13.2 (2.4)	0.42
MMSE, median (range) points	30 (24–30)	29 (24–30)	0.53

C9orf72, chromosome 9 open reading frame 72; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; MMSE, mini-mental state examination.

[a] Education values were missing for one mutation carrier and two controls.

Converters and non-converters

During the study, six mutation carriers converted to FTD. One *MAPT* mutation carrier and one *GRN* mutation carrier converted to bvFTD after two years and two *MAPT* mutation carriers converted to bvFTD after four years. These subjects presented with progressive behavioural deterioration, functional decline, and frontal and/or temporal lobe atrophy on structural MRI, fulfilling the international diagnostic consensus criteria for bvFTD with definite frontotemporal lobar degeneration pathology (Rascovsky et al., 2011). Two *GRN* mutation carriers converted after four years and presented with isolated language difficulties and no impairment in daily living activities. Both showed a non-fluent, halting speech with sound errors and agrammatism, fulfilling the diagnostic criteria for non-fluent variant primary progressive aphasia (nfvPPA; Gorno-Tempini et al., 2011). For a full description of these subjects' clinical profile including neuropsychological assessment, see Jiskoot et al. (2018b, 2019). Baseline demographics for the converters and non-converters are shown in **Table 3.3**, and a timeline for each converter, including bvFTD classification scores and time of conversion, is given in **Figure 3.2**. Other than higher education for converters than non-converters ($p = 0.003$), there were no significant demographic differences at baseline. However, at four years follow-up, converters had significantly lower MMSE than non-converters ($p = 0.001$). At that time, all converters had converted to the symptomatic phase.

Table 3.3 Converter baseline demographics

	Converters (*n* = 6)	Non-converters (*n* = 49)	*p*-value
Gene, *n MAPT* / *GRN* / *C9orf72*	3 / 3 / 0	5 / 32 / 12	–
Age, mean (SD) years	52.7 (8.1)	51.9 (8.8)	0.84
Sex, *n* (%) ♀	4 (67%)	33 (67%)	0.97
Education, mean (SD) years[a]	15.2 (0.8)	13.5 (3.0)	0.003
MMSE, median (range) points	29.5 (28–30)	30 (24–30)	1.0

C9orf72, chromosome 9 open reading frame 72; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; MMSE, mini-mental state examination.

[a] The education value of one non-converter was missing.

Classification scores over time

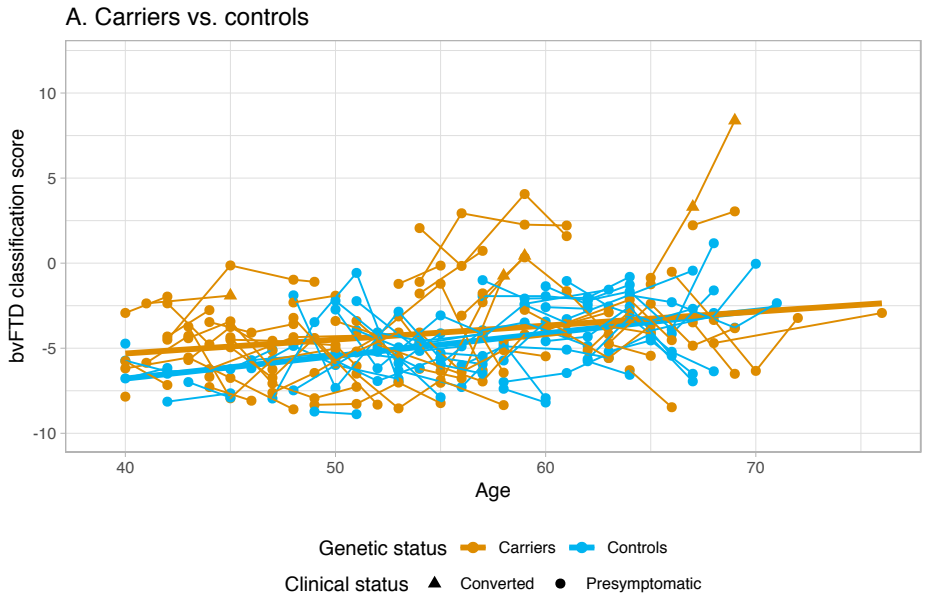
Presymptomatic mutation carriers and controls

Application of the bvFTD model resulted in a classification score for each included time point. **Figure 3.3A** shows the results for all mutation carriers and controls; **Supplemental Figure S3.1A-C** shows the results split per gene. Data points at which mutation carriers had developed symptoms (triangles) were not included in this presymptomatic analysis. Classification scores of mutation carriers did not show a stronger increase over time than controls' scores ($p = 0.15$). However, on average, mutation carriers' classification scores were higher than controls' scores ($p = 0.032$). There was a strong effect of age at baseline on the classification score ($p < 0.001$), but this effect was not different between mutation carriers and controls ($p = 0.72$). Male participants had higher classification scores than female participants ($p < 0.001$).

Converters and non-converters

The bvFTD classification scores for all converters and non-converters are shown in **Figure 3.3B**. In **Supplemental Figure S3.1D–F**, the results are shown split per gene. Classification scores of converters increased more strongly over time than non-converters' scores ($p < 0.001$), though they were similar on average ($p = 0.33$). There was a strong effect of age at baseline on classification score ($p = 0.006$), which was more pronounced in converters than in non-converters ($p = 0.035$). Again, male participants had higher classification scores than female participants ($p = 0.001$).

Figure 3.2 Converter timelines
Converter timelines including baseline demographic information, bvFTD classification scores, clinical follow-up information, and data exclusion information. For reference, classification scores in our population ranged from −8.9 to 8.4, where higher scores represent greater resemblance to bvFTD patients according to the classification model. Blue panels indicate presymptomatic status; red panels indicate subjects have converted.
DWI, diffusion-weighted imaging; bvFTD, behavioural variant frontotemporal dementia; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; nfvPPA, non-fluent variant primary progressive aphasia.



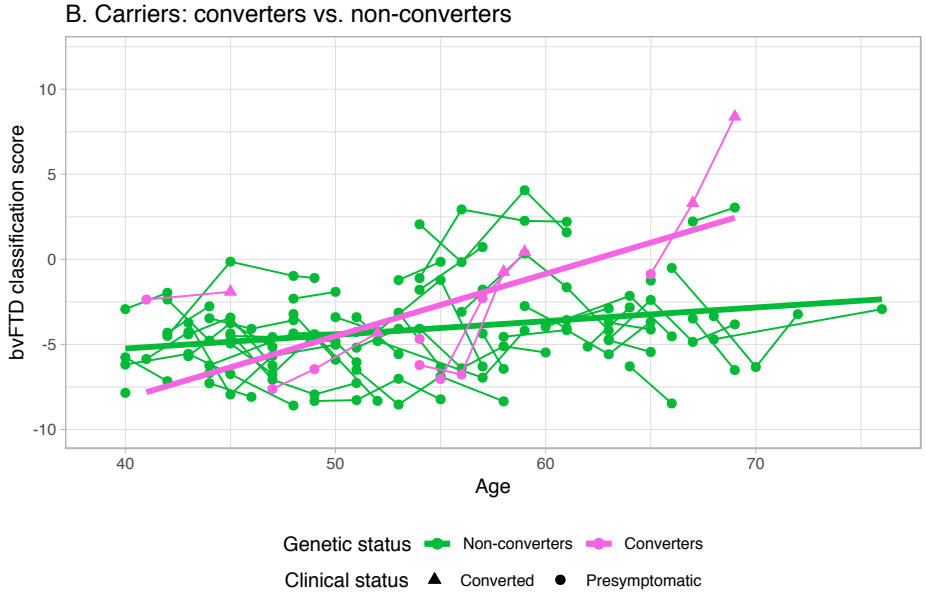


Figure 3.3 BvFTD classification scores
Values indicate bvFTD classification scores as determined by the bvFTD classification model (Bouts et al., 2018). Classification scores are shown (**A**) for all mutation carriers (yellow) vs. controls (blue), and (**B**) for converters (pink) vs. non-converters (green). Triangles indicate converted status at appropriate time points. Regression lines were based on predicted values from the repeated measures linear mixed effects model.
bvFTD, behavioural variant frontotemporal dementia.

Discussion

In this study, we aimed to assess the sensitivity of MRI-based classification to FTD-related changes in the brain before and around FTD symptom onset. We used an MRI-based bvFTD classification model (Bouts et al., 2018) on longitudinal MRI data of presymptomatic FTD mutation carriers and healthy controls with a follow-up of up to six years. Overall, mutation carriers had higher bvFTD classification scores on average than controls, but did not differ from controls in longitudinal score development. However, a subgroup analysis within mutation carriers showed that mutation carriers who converted during follow-up had a stronger score increase over time than non-converting mutation carriers. Converters showed a stronger increase of classification scores between time points, as well as a stronger effect of age at baseline. Our longitudinal findings highlight that MRI-based bvFTD classification is sensitive to FTD-related changes in the brain that arise around symptom onset in FTD mutation carriers.

Carrier-control analysis revealed that over the course of two to six years of follow-up, mutation carriers did not have significantly stronger increases in bvFTD classification scores over time than controls. These results are in line with our previous work, where we showed that the bvFTD model could not separate presymptomatic mutation carriers from controls beyond chance level at baseline (Feis et al., 2019a). That work was limited by its cross-sectional design, which greatly reduced power to find differences in a population with uncertain time to symptom onset. However, even in our current longitudinal design, the mutation carrier group as a whole did not differ from controls in bvFTD classification score increase over time. On the other hand, mutation carriers that converted to FTD during the follow-up of this study did show stronger classification score increases over time. These findings may imply that presymptomatic mutation carriers remain quite similar to controls in terms of individually detectable changes in MRI features until they are within a few years of symptom onset. A direct correlation between MRI-based classification scores and neuropathology has yet to be established. Nonetheless, our results seem in line with the hypothesis that FTD-related neuropathological processes considerably accelerate in the final years before conversion, which was previously postulated in studies focusing on cognitive, neuroimaging, and fluid biomarkers. Specifically, FTD mutation carriers that convert to FTD typically start to show signs of cognitive decline in the language or executive domains two years before diagnosis (Rohrer et al., 2008; Jiskoot et al., 2018b). Furthermore, in longitudinal neuroimaging studies, grey matter atrophy (Rohrer et al., 2008; Jiskoot et al., 2019) and white matter FA changes (Jiskoot et al., 2019) were also first picked up two years before diagnosis, after which they rapidly expanded. In a large multicentre study (Meeter et al., 2016a) of the Genetic Frontotemporal Dementia Initiative (GENFI; Rohrer et al., 2015), neurofilament light chain levels in cerebrospinal fluid and serum were much higher in symptomatic mutation carriers than in presymptomatic mutation carriers, who had similar levels to non-carriers. Additionally, longitudinal samples showed an three- to four-fold increase in neurofilament light chain levels in the cerebrospinal fluid in mutation carriers that converted during the study (Meeter et al., 2016a). All of these results suggest that the presymptomatic stage remains stable until a couple of years before symptom onset, when neurodegeneration rapidly develops.

Strengths of this study include its unique longitudinal design, the inclusion of multimodal MRI data of presymptomatic FTD mutation carriers, and the inclusion of mutation carriers that converted during follow-up. Even so, the length of follow-up may have been insufficient to capture the differences between mutation carriers and controls over time using classification models. A limitation to our study was the heterogeneity of our sample, as we had to pool *MAPT* mutation carriers, *GRN* mutation carriers, and *C9orf72* repeat expansion carriers in order to obtain sufficient

sample size for robust analyses. Also, the classification model was trained on bvFTD patients and controls, and may therefore not be sensitive to MRI changes associated with non-behavioural variants (Seelaar et al., 2011). The greater part (i.e., 35 out of 55) of FTD mutation carriers were *GRN* mutation carriers, who usually develop bvFTD, but also frequently convert to nfvPPA, or may develop atypical parkinsonian syndromes (Benussi et al., 2015). If MRI changes associated with these other diseases are not picked up by the bvFTD model, it might explain why we found no differences between mutation carriers and controls. For example, bvFTD patients show diffuse atrophy and white matter diffusivity changes in bilateral frontotemporal regions, while in nfvPPA patients, these changes are predominantly left hemispheric and are more specifically located (Zhang et al., 2013). Lastly, since the FTD-RisC is an ongoing study, the variable follow-up time is a limitation. Although the repeated measures model is robust for differences in follow-up time, the shorter follow-up time in some mutation carriers could explain the limited number of converters. Specifically, there were no converters in the *C9orf72* repeat expansion carrier group, which was included in the FTD-RisC study at a later time than the *MAPT* and *GRN* groups. Therefore, the results of our converters' analysis should not be generalised to *C9orf72* repeat expansion carriers. Lastly, it should be noted that we chose to not interpret the bvFTD classification model's beta values, even though it might seem interesting to investigate which features drove the classification. Contrary to explanatory regression models, the beta values of our classification model do not designate direct relationships between the features and classification score, nor do they reflect mean differences between the groups (Shmueli, 2010). For example, each feature's effect is conditional on the effects of all other features in the model, and multicollinearity between features may result in suppression of the effect of some features. Furthermore, the model was forced to be sparse, which may have resulted in the suppression or exclusion of true effects from the model. Given the large feature space (i.e., 110 GMD values, 2,415 FCor values, 20 FA values, and 20 MD values) on which the bvFTD classification model was trained, these effects are likely to have occurred, obscuring the true meaning of the model's beta values.

Despite these limitations, we were able to show that FTD-related MRI changes can be recognised by an MRI-based classification model around symptom onset. This proof-of-concept study further emphasises the importance of longitudinal data acquisition such as the FTD-RisC and GENFI. With longer follow-up times and more converters, it may eventually be possible to find exclusively presymptomatic changes on an individual level, which would be an important step towards early and accurate FTD diagnosis. Larger sample sizes might also facilitate stratification according to the different gene mutations (i.e., *MAPT*, *GRN*, and *C9orf72* repeat expansion) or different clinical variants, which was currently infeasible. In order to increase the clinical relevance of MRI-based machine learning for dementia, future studies should focus on replicating classification models in larger samples, validating them in separately established samples, and finally applying them in undiagnosed memory clinic populations to establish—with the use of follow-up information or post-mortem examination—whether they are useful in the earliest phases of the disease.

Conclusion

To conclude, we showed that FTD-related changes are measurable by MRI-based classification in years around symptom onset in FTD mutation carriers. This indicates that MRI-based classification is a reasonable candidate to aid early diagnosis of FTD, and could contribute to improved development of disease modifying treatments that can slow down or possibly reverse the underlying neuropathological processes.

Acknowledgements

Our appreciation goes out to all FTD-RisC participants and families. The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to RAF), ZonMw programme Memorabel project 733050103, JPND PreFrontAls consortium project 733051042 (to JCvS), and a VICI grant 016–130–667 from The Netherlands Organisation for Scientific Research (NWO; to SARBR). The views expressed are those of the authors and not necessarily those of the funding sources. The funding sources were not involved in the design of the study; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The authors report no conflict of interest.

Supplemental material

Image preprocessing

For $3DT_1w$ images, we performed bias field correction (N4ITK; Tustison et al., 2010), brain extraction (FSL BET; Smith, 2002), nonlinear registration to the MNI152 2 × 2 × 2 mm T_1 template (FNIRT; Anderson et al., 2007), tissue type segmentation (SPM12; Friston et al., 2007), and segmentation of deep grey matter structures, including the bilateral thalamus, caudate nucleus, putamen, globus pallidum, nucleus accumbens, amygdala, and hippocampus (FIRST; Patenaude et al., 2011).

Preprocessing of DWI data sets included correction of motion and eddy-current induced distortions (eddy correct; Leemans & Jones, 2009), and calculation of voxel-wise measures of FA and MD (DTIFIT; Smith et al., 2004). A global mean FA image was created by nonlinearly registering FA maps to the FMRIB58_FA template, and tract-based spatial statistics (FSL TBSS; Smith et al., 2006) was used to extract FA and MD values using the standard FSL TBSS skeleton. The skeleton was thresholded at 0.2 to ensure that extracted values originate from white matter.

For rs-fMRI data, preprocessing consisted of motion correction (Jenkinson et al., 2002), brain extraction, spatial smoothing using a Gaussian kernel with a full width at half maximum of 3 mm, grand mean intensity normalisation, motion artefact removal, and high-pass temporal filtering (cut-off frequency = 0.01 Hz). Motion artefacts were removed using a single-session independent component analysis (ICA) to decompose the rs-fMRI data into distinct statistically independent components, followed by automatic identification and removal using the ICA-based automatic removal of motion artefacts (ICA-AROMA, version 0.3 beta; Pruim et al., 2015) procedure. Registration to standard space was performed in two steps. First, a temporal mean image calculated from the 4D rs-fMRI volume was registered to the $3DT_1w$ image using boundary-based registration (Greve & Fischl, 2009). Next, resulting registration parameters were concatenated to the $3DT_1w$-to-MNI152 template registration parameters to obtain the final registration parameters.

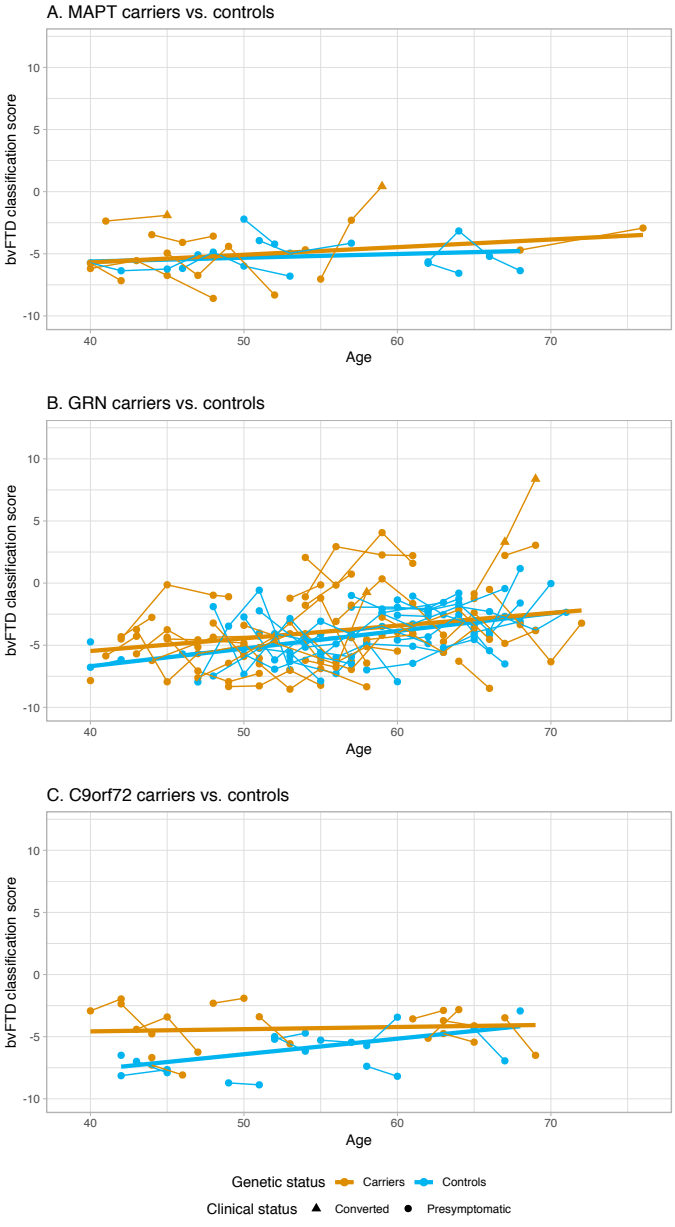
Feature selection

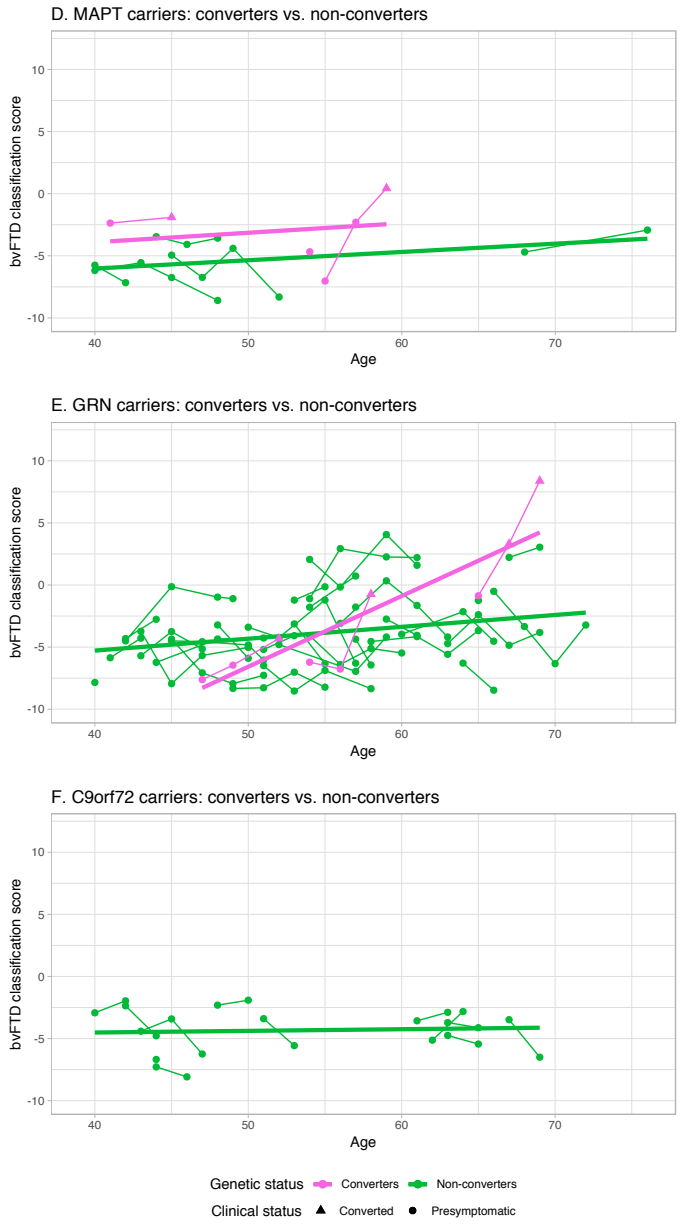
Cortical grey matter density (GMD) was calculated as weighted mean of the grey matter probability (SPM segmentation) weighted by the probability of a voxel belonging to a specific cortical region, derived from the 48 Harvard-Oxford probabilistic anatomical brain atlas (split into left and right). Voxels with atlas probability values under 25% were excluded. For deep grey matter regions, GMD values were calculated as the regions' volume (FIRST segmentations) divided by the total intracranial volume. This resulted in a grey matter atrophy feature vector of 110 weighted mean GMD values (48 left cortical, 48 right cortical, and 14 deep grey matter regions) per subject.

The FA and MD features were calculated as weighted mean value per tract per subject. First, we projected each subject's values onto the TBSS group skeleton on a voxel-wise basis. Next, we weighted the mean values per tract by the probability of a voxel belonging to that specific tract, derived from 20 tracts of the Johns-Hopkins University white matter tractography atlas. Voxels with atlas probability values under 25% were excluded. This resulted in two feature vectors, each with 20 weighted mean FA or MD values per subject.

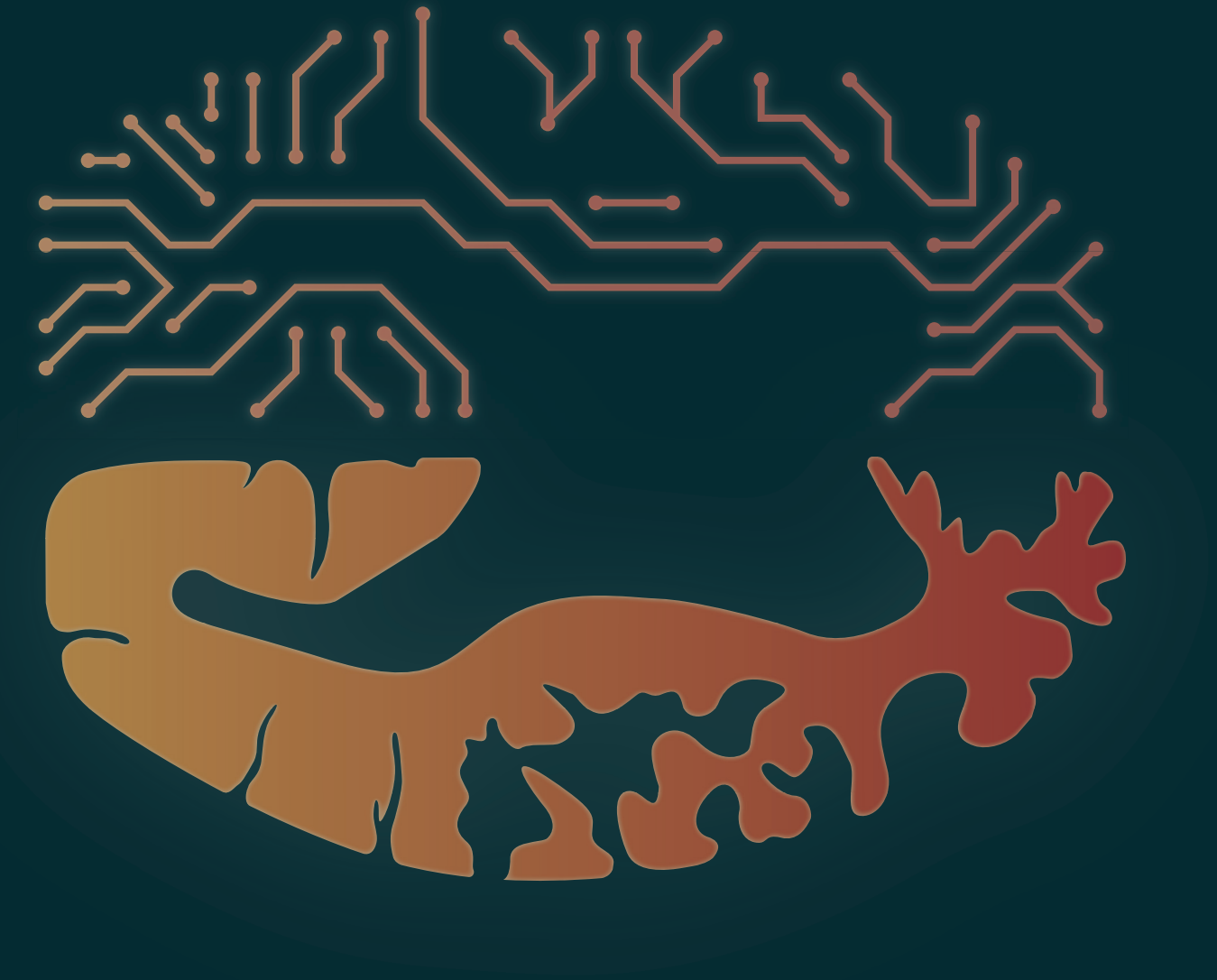
To calculate full correlations between ICA components (FCor), we used the 70 group ICA (MELODIC; Beckmann & Smith, 2004) components' spatial maps belonging to the data on which the bvFTD model was trained (Bouts et al., 2018). For each subject, we calculated the mean time course in the rs-fMRI data for each of these components, weighted by the ICA weight map and grey

matter probability of that component's regions. Full correlations between the mean time courses of all pairs of components were subsequently calculated, resulting in an FCor vector of (70 × 69) ÷ 2 = 2,415 correlations per subject.





Supplemental Figure S3.1 BvFTD classification scores per gene
Values indicate bvFTD classification scores as determined by the bvFTD classification model (Bouts et al., 2018). Classification scores of mutation carriers (yellow) vs. controls (blue) are shown for *MAPT* (**A**), *GRN* (**B**), and *C9orf72* repeat expansion (**C**). Classification scores of converters (pink) vs. non-converters (green) are shown for *MAPT* (**D**), *GRN* (**E**), and *C9orf72* repeat expansion (**F**). Triangles indicate converted status at appropriate time points. Regression lines were based on predicted values from the repeated measures linear mixed effects model.
bvFTD, behavioural variant frontotemporal dementia; *C9orf72*, chromosome 9 open reading frame 72; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau.



Chapter 4

Frontotemporal dementia mutation carriers: symptom onset prediction using MRI-based and cognitive features

Published in *Brain Communications 2020;2(2):fcaa079* as:

Classification using fractional anisotropy predicts conversion in genetic frontotemporal dementia, a proof of concept

Rogier A. Feis, Jeroen van der Grond, Mark J.R.J. Bouts, Jessica L. Panman, Jackie M. Poos, Tijn M. Schouten, Frank de Vos, Lize C. Jiskoot, Elise G.P. Dopper, Mark A. van Buchem, John C. van Swieten, Serge A.R.B. Rombouts

Abstract

Frontotemporal dementia (FTD) is a highly heritable and devastating neurodegenerative disease. About 10–20% of all FTD is caused by known pathogenic mutations, but a reliable tool to predict clinical conversion in mutation carriers is lacking. In this retrospective proof-of-concept case-control study, we investigate whether magnetic resonance imaging (MRI-)based and cognition-based classifiers can predict which mutation carriers from genetic frontotemporal dementia families will develop symptoms ('convert') within 4 years.

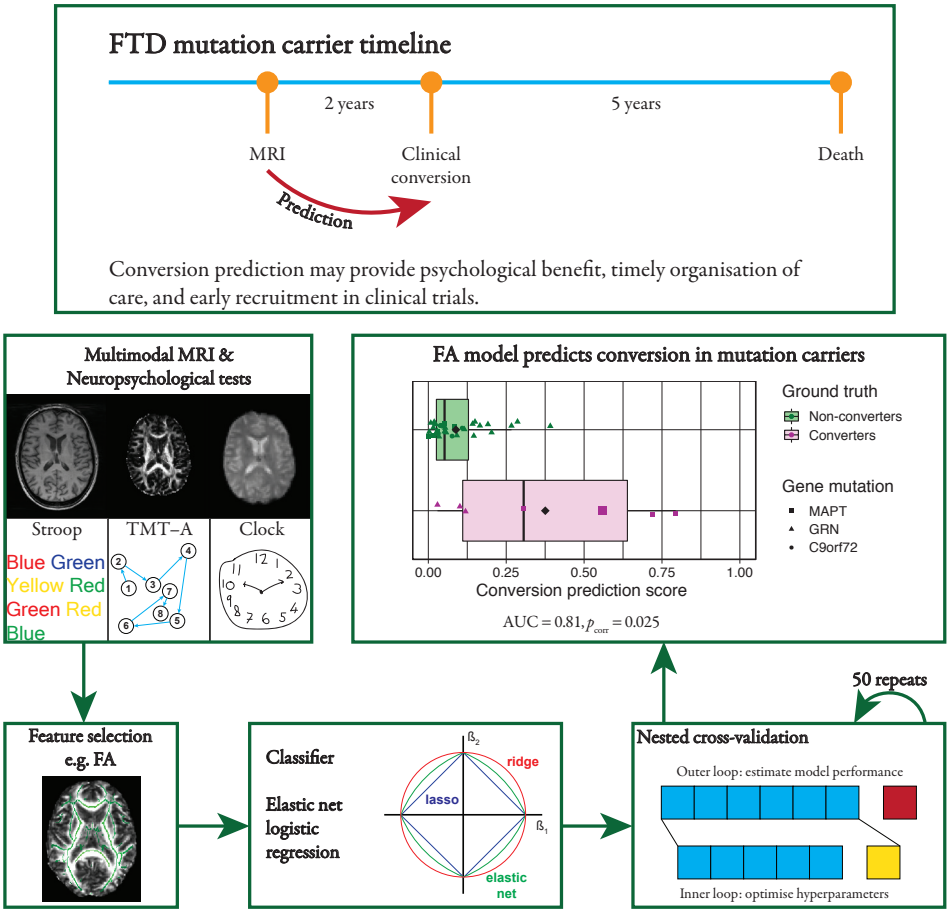
From genetic FTD families, we included 42 presymptomatic FTD mutation carriers. We acquired anatomical, diffusion-weighted imaging, and resting-state functional MRI, as well as neuropsychological data. After 4 years, seven mutation carriers had converted to FTD ('converters'), while 35 had not ('non-converters'). We trained regularised logistic regression models on baseline MRI and cognitive data to predict conversion to FTD within 4 years, and quantified prediction performance using area under the receiver operating characteristic curves (AUC).

The prediction model based on fractional anisotropy, with highest contribution of the forceps minor, predicted conversion to FTD beyond chance level (AUC = 0.81, family-wise error corrected $p = 0.025$ vs. chance level). Other MRI-based and cognitive features did not outperform chance level.

Even in a small sample, fractional anisotropy predicted conversion in presymptomatic FTD mutation carriers beyond chance level. After validation in larger data sets, conversion prediction in genetic FTD may facilitate early recruitment into clinical trials.

Keywords: frontotemporal dementia; *MAPT* protein, human; *GRN* protein, human; *C9orf72*, human; diffusion tensor imaging; resting-state functional MRI; multimodal MRI; classification; machine learning

Graphical abstract



AUC, area under the receiver operating characteric curve; C9orf72, chromosome 9 open reading frame 72; FA, fractional anisotropy; FTD, frontotemporal dementia; GRN, progranulin; MAPT, microtubule-associated protein tau; MRI, magnetic resonance imaging; p_{corr}, family-wise error rate corrected p-value; TMT-A, trail making test A.

Introduction

Frontotemporal dementia (FTD) is a highly heritable and devastating neurodegenerative disease that often occurs at a presenile age (Ratnavalli et al., 2002; Harvey et al., 2003; Hogan et al., 2016). Patients typically present with behavioural symptoms (behavioural variant FTD; bvFTD; Rascovsky et al., 2011) or with language disorders (primary progressive aphasia; PPA; Gorno-Tempini et al., 2011), but may also develop amyotrophic lateral sclerosis (ALS; Lomen-Hoerth et al., 2002) or parkinsonian syndromes such as corticobasal syndrome (CBS) and progressive supranuclear palsy (PSP; Josephs et al., 2006). About 10–20% of all FTD is caused by three known pathogenic mutations: microtubule-associated protein tau (*MAPT*), progranulin (*GRN*), and chromosome 9 open reading frame 72 (*C9orf72*) repeat expansion (Seelaar et al., 2008; Rohrer et al., 2009; Benussi et al., 2015). These mutations have a high penetrance and an autosomal dominant inheritance pattern. As such, roughly 50% of family members from a genetic FTD pedigree will develop an FTD variant.

Uncertainty about if and when a person will develop FTD is a major burden (Riedijk et al., 2009; Cohn-Hokke et al., 2018), and some subjects undergo genetic testing to eliminate this uncertainty. However, when a mutation is found, this raises the question when symptoms will emerge. A reliable prognosis may provide psychological benefit, and contributes to the timely organisation of adequate care. Moreover, conversion prediction in genetic FTD may facilitate early recruitment into clinical trials for disease modifying treatments, which could lead to improved trial efficacy (Tsai & Boxer, 2016). However, accurate conversion prediction is currently not possible. Expected time to symptom onset, calculated using the ages of FTD onset of family members, has been used as proxy in research settings (Rohrer et al., 2015; Jiskoot et al., 2018a), but these estimates are inaccurate due to the large variation in age of onset between and within genetic FTD families (van Swieten & Heutink, 2008; Jiskoot et al., 2018a). A more reliable clinical tool is needed to predict conversion in genetic FTD.

Magnetic resonance imaging (MRI-)based classification is a promising diagnostic biomarker for FTD (Raamana et al., 2014; Klöppel et al., 2015; Koikkalainen et al., 2016; Bron et al., 2017; Meyer et al., 2017; Bouts et al., 2018), and distinguishes presymptomatic FTD mutation carriers from controls beyond chance (Feis et al., 2019a). Classification with cognitive measures has also been advocated as diagnostic FTD biomarker (Wang et al., 2016). Therefore, MRI- and cognition-based classification may be suitable candidates for conversion prediction in genetic FTD. In recent observational studies, FTD mutation carriers were followed up until conversion, and multimodal MRI data were analysed using mass univariate techniques such as voxel-based morphometry and diffusion tensor imaging (DTI; Jiskoot et al., 2019), or using classification scores calculated from multimodal MRI and a classification model based on FTD patients (Feis et al., 2019b). These studies suggest that brain changes as found on MRI appear relatively explosively in the final years before symptom onset. However, it is yet unknown whether conversion in such subjects could have been predicted based on pre-conversion MRI scans or neuropsychological data.

Here, we study whether MRI-based features and cognitive measures have predictive value for FTD conversion in genetic FTD. To this end, we train prediction models on MRI-based features and cognitive features to predict which presymptomatic FTD mutation carriers develop symptoms within 4 years ('converters'), and which mutation carriers do not ('non-converters').

Methods

Participants

This retrospective case-control study included FTD mutation carriers from the FTD-Risk Cohort (FTD-RisC; Dopper et al., 2014; Meeter et al., 2016a; Papma et al., 2017; Jiskoot et al., 2016, 2018b, 2019), a longitudinal study that follows healthy, 50% at-risk family members of genetic FTD patients on a 2-year basis. For the current study, we included the data of 43 FTD mutation carriers, who entered FTD-RisC between May 2010 and November 2014. Inclusion criteria were age at MRI scan between 40 and 70 years old, availability of a 3-dimensional T_1-weighted MRI ($3DT_1w$) scan, a diffusion-weighted imaging (DWI) data set, a resting-state functional MRI T_2*-weighted (rs-fMRI) scan, and availability of clinical follow-up after 2 and 4 years. Genotyping, as described in previous work (Dopper et al., 2014; Jiskoot et al., 2016), revealed nine *MAPT* mutation carriers, 28 *GRN* mutation carriers, and six *C9orf72* repeat expansion carriers. For seven FTD mutation carriers who had converted to FTD (four *MAPT* and three *GRN* mutation carriers; 'converters'; see 'Conversion' section below), we included the MRI data of their last visit before conversion. For one converter (*MAPT* mutation carrier), these MRI data were excluded due to the presence of artefacts, and we included MRI data of the previous visit. The time between MRI and conversion ranged between 11 and 41 months (mean = median = 22 months, SD = 10.5 months). More details on this group's demographics are provided in there results. The remaining mutation carriers were confirmed to be asymptomatic after 4 years follow-up in accordance with established diagnostic criteria for bvFTD (Rascovsky et al., 2011), PPA (Gorno-Tempini et al., 2011), and ALS (Ludolph et al., 2015), and were termed 'non-converters'. Follow-up after 6 years was not complete for this group: 25 non-converters were followed up after 6 years and all remained asymptomatic at that point. For non-converters, we included the FTD-RisC baseline MRI data. One non-converter with a *GRN* mutation was excluded due to incomplete neuropsychological data. Other exclusion criteria were: current or past neurologic (other than dementia) or primary psychiatric disorders, history of drug abuse, large image artefacts, and gross brain pathology other than atrophy.

Standard protocol approvals, registrations, and patient consents

Participants and clinical investigators were blinded to the participants' genetic status, except for those that underwent predictive testing at their own request. For converters, genetic counselling was offered to the patient and family members, and genetic status was unblinded to confirm the presence of the pathogenic mutation. The study was conducted in accordance with regional regulations and the Declaration of Helsinki. The Erasmus Medical Centre and Leiden University Medical Centre local medical ethics committees approved the study, and every participant provided written informed consent (Feis et al., 2019b).

Conversion

Conversion was determined in a multidisciplinary consensus meeting of the Erasmus Medical Centre FTD Expertise Centre, involving neurologists (JCvS), neuropsychologists (JLP, JMP, LCJ), neuroradiologists, a clinical geneticist, and a care consultant (Feis et al., 2019b). In the consensus meetings, information from the medical history, neuropsychological assessment, and MRI of the brain were reviewed. The timing of symptom onset was estimated from heteroanamnestic information provided by knowledgeable informants (e.g., siblings, spouses). After the clinical diagnosis was

made, and if the subject and family agreed, genetic status was unblinded for confirmation. Four *MAPT* mutation carriers and one *GRN* mutation carrier converted to bvFTD, while two *GRN* mutation carriers converted to non-fluent variant PPA (nfvPPA). BvFTD converters presented with progressive behavioural deterioration, functional decline, and frontal and/or temporal lobe atrophy on structural MRI, fulfilling the international diagnostic consensus criteria for bvFTD with definite frontotemporal lobar degeneration pathology (Rascovsky et al., 2011). The two nfvPPA converters presented with isolated language difficulties and no impairment in daily living activities. Both showed a non-fluent, halting speech with sound errors and agrammatism, fulfilling the diagnostic criteria for nfvPPA (Gorno-Tempini et al., 2011). For more detailed information on conversion criteria, and for a full description of the converters' clinical profile, see Jiskoot et al. (2018b, 2019).

MRI data acquisition

Subjects were scanned at the Leiden University Medical Centre using a 3 T MRI scanner (Achieva, Philips Medical Systems, Best, The Netherlands) with an 8-channel SENSE head coil (Feis et al., 2019a). The imaging protocol included a whole-brain near-isotropic $3DT_1w$ sequence for cortical and subcortical tissue-type segmentation, a DWI sequence for assessments of white matter diffusivity, and rs-fMRI sequence for the calculation of functional connectivity measures. Participants were instructed to lie still with their eyes closed and not to fall asleep during rs-fMRI. Scan parameters are provided in **Table 4.1**.

Table 4.1 MRI sequence parameter settings

	$3DT_1w$	DWI[a]	rs-fMRI
Slices, *n*	140	70	38
TR, ms	9.8	8,250	2,200
TE, ms	4.6	80	30
Flip angle, °	8	90	80
Matrix, mm	256 × 256	128 × 128	80 × 80
Voxel size, mm	0.88 × 0.88 × 1.20	2.00 × 2.00 × 2.00	2.75 × 2.75 × 2.99[b]
Duration, min	4.57	8.48	7.28

Scan protocol of whole-brain near-isotropic 3-dimensional T_1-weighted ($3DT_1w$), diffusion-weighted imaging (DWI), and resting-state functional T_2*-weighted MRI (rs-fMRI) on a 3 T scanner at the Leiden University Medical Centre.
TR, repetition time; TE, echo time.
[a] 60 directions, b = 1,000, one b0 image.
[b] Including 10% interslice gap.

Image preprocessing

Preprocessing of MRI scans was performed similarly to previous work (Bouts et al., 2018; Feis et al., 2019a, b). All registration and segmentation steps were critically reviewed and errors were corrected accordingly.

For $3DT_1w$ images, we performed bias field correction (N4ITK; Tustison et al., 2010), brain extraction (FSL BET; Smith, 2002), nonlinear registration to the MNI152 2 × 2 × 2 mm T_1 template (FNIRT; Anderson et al., 2007), tissue-type segmentation (SPM12; Friston et al., 2007),

and segmentation of deep grey matter structures, including the bilateral thalamus, caudate nucleus, putamen, globus pallidum, nucleus accumbens, amygdala, and hippocampus (FIRST; Patenaude et al., 2011). Our choice to use SPM segmentation was based on the segmentation tool comparison by Kazemi & Noorizadeh, 2014.

Preprocessing of DWI data sets included correction of motion and eddy-current induced distortions (eddy correct; Leemans & Jones, 2009), and voxel-wise calculation of the measures fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AxD; largest eigenvalue), and radial diffusivity (RD; average of the two remaining eigenvalues) using DTIFIT (Smith et al., 2004). A global mean FA image was created by nonlinearly registering FA maps to the FMRIB58_FA template, and tract-based spatial statistics (FSL TBSS; Smith et al., 2006) was used to extract FA, MD, AxD, and RD values using the standard FSL TBSS skeleton. The skeleton was thresholded at 0.2 to ensure that extracted values originate from white matter.

For rs-fMRI data, preprocessing consisted of motion correction (Jenkinson et al., 2002), brain extraction, spatial smoothing using a Gaussian kernel with a full width at half maximum of 3 mm, grand mean intensity normalisation, motion artefact removal, and high-pass temporal filtering (cut-off frequency = 0.01 Hz). Motion artefacts were removed using a single-session independent component analysis (ICA) to decompose the rs-fMRI data into distinct statistically independent components, followed by automatic identification and removal of motion artefacts using ICA-AROMA (version 0.3 beta; Pruim et al., 2015). Registration to standard space was performed in two steps. First, a temporal mean image calculated from the 4D rs-fMRI volume was registered to the $3DT_1$w image using boundary-based registration (Greve & Fischl, 2009). Next, resulting registration parameters were concatenated to the $3DT_1$w-to-MNI152 template registration parameters to obtain the final registration parameters.

MRI feature selection

Selection of MRI features was performed as described previously (Bouts et al., 2018; Feis et al., 2019a).

Cortical grey matter density (GMD) and white matter density (WMD) were calculated as weighted means of their respective regional grey matter or white matter probability (SPM segmentation) weighted by the probability of a voxel belonging to a specific cortical region. Cortical probabilities were derived from the 48 Harvard-Oxford probabilistic anatomical brain atlas (split into left and right), and white matter probabilities were derived from the Johns-Hopkins University white matter tractography atlas for 20 white matter tracts. Voxels with atlas probability values under 25% were excluded. We estimated these features in native space by transforming the atlas masks to the images' native space using the inverted non-linear registration parameters. For deep grey matter regions, GMD values were calculated as the regions' volume (FIRST segmentations) divided by the total intracranial volume. This resulted in a GMD feature vector of 110 weighted mean GMD values (48 left cortical, 48 right cortical, and 14 deep grey matter regions) and in a feature vector of 20 weighted mean WMD values per subject. While grey matter volume, as estimated by voxel-based morphometry, would also be a useful feature for prediction, we chose the simpler GMD for generalisation purposes, as it is not reliant on a study-specific template.

DTI features were calculated as weighted mean FA, MD, AxD, or RD values per tract per subject. First, we projected each subject's FA, MD, AxD, and RD values onto the TBSS group skeleton on a voxel-wise basis. Next, we weighted the mean values per tract by the probability of a voxel belonging to that specific tract, derived from 20 tracts of the Johns-Hopkins University white matter tractography atlas. Voxels with atlas probability values under 25% were excluded. This resulted in

four feature vectors of each 20 weighted mean values per subject.

To calculate functional connectivity features, all processed rs-fMRI images were combined in a temporally concatenated group-level ICA (MELODIC; Beckmann & Smith, 2004), with dimensionality fixed at 20 components and an ICA threshold of 0.99 (Smith et al., 2013). This means that each voxel included in the ICA map was 99 times more likely to be part of that component than to be caused by Gaussian background noise. For each subject, we calculated the mean time course for each component, weighted by the ICA weight map and grey matter probability of that component's regions. Correlations between the mean time courses of all pairs of components were subsequently calculated. Functional connectivity was calculated as full correlations (FCor), and as sparse, L1-regularised, partial correlations (PCor) between the mean time courses of all pairs of components. Partial correlations were calculated using the graphical lasso algorithm (J. Friedman et al., 2008). This procedure resulted in two feature vectors of each $(20 \times 19) \div 2 = 190$ (partial) correlations per subject.

Finally, we created one multimodal feature vector by concatenating all feature vectors together. As such, we had nine feature vectors in total: two $3DT_1w$ feature vectors, four DTI feature vectors, two rs-fMRI feature vectors, and one multimodal imaging feature vector.

Neuropsychological assessment and cognitive feature selection

We screened global cognitive functioning by means of the Mini-Mental State Examination (MMSE; Folstein et al., 1975). Experienced neuropsychologists (JLP, JMP, LCJ) administered neuropsychological tests within six cognitive domains: language, attention and mental processing speed, executive functioning, social cognition, memory, and visuoconstruction. The language domain was assessed using the 60-item Boston naming test (BNT; Kaplan et al., 1978), verbal semantic association test (SAT; Visch-Brink et al., 2005), ScreeLing phonology (Doesborgh et al., 2003), and categorical fluency (Thurstone & Thurstone, 1962). The domain of attention and mental processing speed was evaluated using the trail making test (TMT-)A (Army Individual Test Battery, 1944), the mean of Stroop colour-word tests I and II (Stroop, 1935), Wechsler adult intelligence scale III (WAIS-III) digit span forwards (Wechsler, 2005), and letter digit substitution test (LDST; Jolles et al., 1995). We measured the executive functioning domain using TMT-B (Army Individual Test Battery, 1944), Stroop colour-word test III (Stroop, 1935), WAIS-III digit span backwards (Wechsler, 2005), modified Wisconsin card sorting test (WCST) concepts (Nelson, 1976), letter fluency (Thurstone & Thurstone, 1962), and WAIS-III similarities (Wechsler, 2005). The social cognition domain was assessed by means of Happé theory of mind cartoons, Happé non-theory of mind cartoons (Happé et al., 1999), and Ekman faces (Ekman & Friesen, 1976). We evaluated the memory domain using the immediate response and recall of the Dutch Rey auditory verbal learning test (RAVLT; Rey, 1958), and the visual association test (VAT; Lindeboom et al., 2002). Lastly, we tested the visuoconstruction domain by means of clock drawing (Royall et al., 1998) and WAIS-III block design (Wechsler, 2005).

Raw test scores were added into one cognitive feature vector per domain. This resulted in a language feature vector with four features, an attention feature vector with four features, an executive feature vector with six features, a social feature vector with three features, a memory feature vector with three features, and a visuoconstruction feature vector with two features. Finally, we added these feature vectors together to create a multidomain feature vector, bringing the total number of feature vectors on seven: language, attention, execution, social, memory, visuoconstruction, and multidomain.

Classifier

In order to identify future converters (n = 7) from non-converters (n = 35) at baseline, we trained prediction models using MRI features (n = 9) and, separately, cognitive features (n = 7). Feature vectors were used to train a logistic elastic net regression algorithm (Zou & Hastie, 2005; Friedman et al., 2010; Schouten et al., 2016; Bouts et al., 2018; Feis et al., 2019a). The elastic net regression procedure estimates a sparse regression model that includes only a subset of the provided features by imposing a penalty for including features (i.e., L1 penalty) and for the sum of the squared value of the coefficients (i.e., L2 penalty). This way, elastic net provides a solution for the imbalance between the large number of features and the small number of subjects. Age and sex were included in the model without penalty to ensure that estimated feature regression coefficients were conditional on subject age and sex. A prediction score of 0 represented a non-converter and 1 represented a converter.

Cross-validation

We trained our conversion prediction models in a stratified nested seven-fold cross-validation scheme to make sure that the proportion of converters and non-converters was the same in each fold (i.e., one converter and five non-converters per fold). In the outer loop, one part of the data (i.e., one of the seven folds) was set apart as a test set and served to test the generalised prediction performance of the elastic net regression model. The remaining parts (six of the seven folds) were used to train the model. Within the training set of the outer loop, we performed a nested cross-validation to optimise the model's hyperparameters without overestimating prediction performance (Varma & Simon, 2006; Kriegeskorte et al., 2009). The resulting optimal hyperparameters were used in the training set of the outer loop to train the model, and the prediction performance was then tested in the test set of the outer loop. This process was repeated seven times to make sure that each subject was part of the test set exactly once. Since the test set of the outer loop was neither used for model training, nor for parameter optimisation, we reduced the risk of overestimating the generalisation performance as much as possible (Kriegeskorte et al., 2009; Schouten et al., 2016; Bouts et al., 2018; Feis et al., 2019a). The entire prediction procedure was repeated 50 times to average prediction outcome variability resulting from random partitioning in training and test folds. All prediction analyses and evaluations were implemented in R version 3.3.2 (R core 2016, GLMnet package; Friedman et al., 2010).

Prediction performance

For both analyses, we quantified prediction performances using receiver operating characteristic curves. Receiver operating characteristic curves were calculated by shifting the threshold for predicting an individual as converter from 0 to 1, and plotting the true positive rate (sensitivity) vs. the false positive rate (1 – specificity) for each intermediate point. The area under this receiver operating characteristic curve (AUC) is a measure of prediction performance insensitive to the distribution between the groups (Fawcett, 2006). Additionally, we calculated the optimal operating point on the curve to calculate the model's sensitivity, specificity, and prediction accuracy, given equal class distribution and equal penalty for false positive and false negative predictions. We averaged AUC, accuracy, sensitivity, and specificity values from the 50 times repeated nested cross-validations (Schouten et al., 2016; Bouts et al., 2018; Feis et al., 2019a).

Statistical analysis

Statistical group analyses of demographic data were performed using R (R Core 2016, Vienna, Austria). We tested for differences between converters and non-converters using unpaired *t*-tests (age and education), the Mann-Whitney U test (MMSE scores [0–30]) and the chi-square test (sex distribution).

To compare MRI and cognitive prediction models' AUC values vs. chance level, we used permutation tests (N = 5,000; Noirhomme et al., 2014). We used the maximum *t*-statistic method to correct for multiple comparisons within, respectively, the nine MRI-based and the seven cognitive prediction analyses. For each permutation, we calculated the maximum absolute *t*-statistic within the family of tests, which resulted in a maximum *t*-distribution of 5,000 maximum absolute *t*-statistics. The observed *t*-statistic of each analysis was then compared to this maximum *t*-distribution in order to obtain a family-wise error rate corrected *p*-value. The alpha level required for statistical significance was set at 0.05.

Data availability

Raw data were generated at the Leiden University Medical Centre. The derived data, as well as scripts, that support the findings of this study are available from the corresponding author upon request.

Results

Demographics

Seven converters and 35 non-converters met the inclusion criteria (**Table 4.2**). At the time of MRI scan, converters did not differ from non-converters in terms of age ($p = 0.85$), sex distribution ($p = 0.40$), and MMSE scores ($p = 0.50$). However, converters had higher levels of education than non-converters ($p = 0.015$). More information on the converter group is shown in **Table 4.3**.

Table 4.2 Participant demographics

	Converters ($n = 7$)[a]	Non-converters ($n = 35$)	p-value
Age, mean (SD) years	51.7 (8.8)	51.0 (8.4)	0.85
Sex, *n* (%) ♀	4 (57%)	26 (74%)	0.40
Education, mean (SD) years[b]	15.2 (0.8)	13.7 (2.9)	0.015
MMSE, median (range) points	29 (27–30)	30 (24–30)	0.50

MMSE, Mini-Mental State Examination.
[a] 4 microtubule-associated protein tau mutation carriers, 3 progranulin mutation carriers.
[b] Education values were missing for one converter.

Table 4.3 Converter demographics

	Gene	Mutation	Age at MRI	Sex	Months to conversion	FTD variant	Prediction score[a]
Converter 1	*GRN*	S82VfsX174	65	Female	23	bvFTD	0.12
Converter 2	*MAPT*	P301L	54	Female	28	bvFTD	0.31
Converter 3	*MAPT*	P301L	57	Male	18	bvFTD	0.79
Converter 4	*GRN*	S82VfsX174	56	Female	11	nfvPPA	0.10
Converter 5	*GRN*	S82VfsX174	49	Female	22	nfvPPA	0.03
Converter 6	*MAPT*	G272V	41	Male	41[b]	bvFTD	0.56
Converter 7	*MAPT*	G272V	42	Male	11	bvFTD	0.72

bvFTD, behavioural variant FTD; FTD, frontotemporal dementia; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau; MRI, magnetic resonance imaging; nfvPPA, non-fluent variant primary progressive aphasia.
[a] Prediction scores based on the FA prediction model range from 0 to 1, with 0 representing non-converters and 1 representing converters.
[b] MRI data nearer to symptom onset was excluded due to artefacts for this subject.

Prediction performance

Prediction with MRI features yielded mixed results (**Table 4.4**). The model based solely on FA features (FA model) significantly outperformed chance level with an AUC of 0.81 ($p = 0.025$ vs. chance level). For the FA model, subjects' prediction scores are shown in **Figure 4.1**. Converters with a *MAPT* mutation seemed to have higher conversion prediction scores than those with a *GRN* mutation. Similarly, male mutation carriers seemed to have higher prediction scores than female

mutation carriers. There seemed to be no clear correlation (Pearson's $r = 0.02$) between conversion prediction scores and time from MRI to conversion. However, the above three observations should be noted with care, since the small sample size did not permit for meaningful statistical testing.

The FA model's beta weights for the 50 cross-validation repeats are shown in **Figure 4.2**. The highest beta weights were found in the forceps minor. Other white matter tracts with above average beta weights included the forceps major, right corticospinal tract, right inferior longitudinal fasciculus, right anterior thalamic radiation, and right uncinate fasciculus. For the two WM tracts with the largest beta weights (i.e., forceps minor and major), we plotted subjects' mean FA value across the tract in **Figure 4.3**. Corresponding with the beta weights, the differences between converters and non-converters were larger for the forceps minor than for the forceps major.

The other DTI features (i.e., MD, AxD, RD) did not outperform chance level, nor did grey and white matter density features, and functional connectivity features (i.e., FCor, PCor). Concatenating all features together in a multimodal MRI model did not improve performance compared to the FA model (AUC = 0.63, $p = 0.55$ vs. chance level).

Prediction with cognitive features resulted in poor performances (**Table 4.5**). The best performing cognitive model was based on the executive domain (AUC = 0.35). Adding multiple domains together in a multidomain cognitive model did not improve performance (AUC = 0.34).

Table 4.4 MRI features' performance

MRI modality	AUC	Min – max	Sensitivity	Specificity	Accuracy	FWERC *p*-value (AUC > chance)
GMD	0.668	0.563 – 0.731	0.746	0.599	0.623	0.375
WMD	0.438	0.306 – 0.567	0.571	0.517	0.526	0.976
FA	*0.812*	*0.608 – 0.906*	*0.769*	*0.761*	*0.762*	*0.025*
MD	0.677	0.616 – 0.743	0.711	0.693	0.696	0.322
AxD	0.435	0.327 – 0.563	0.626	0.491	0.514	0.977
RD	0.625	0.465 – 0.722	0.689	0.630	0.640	0.590
FCor	0.336	0.204 – 0.449	0.446	0.513	0.501	1.000
PCor	0.403	0.233 – 0.502	0.514	0.527	0.525	0.996
Multimodal	0.626	0.547 – 0.710	0.557	0.781	0.744	0.552

Converters (seven presymptomatic FTD-RisC mutation carriers that developed symptoms within 4 years after assessment) vs. non-converters (35 presymptomatic FTD-RisC mutation carriers that remained cognitively healthy after 4 years). Multimodal represents a combination of all MRI features. Min – max AUC values represent the variance across the 50 repeats. Italic: mean AUC significantly higher than chance level after family-wise error rate correction.

AUC, area under the receiver operating characteristic curve; FA, fractional anisotropy; FCor, full correlations between 20 ICA components; FTD-RisC, frontotemporal dementia Risk Cohort; FWERC, family-wise error rate corrected; GMD, grey matter density; ICA, independent component analysis; MD, mean diffusivity; PCor, L1-regularised partial correlations between 20 ICA components; RD, radial diffusivity; WMD, white matter density.

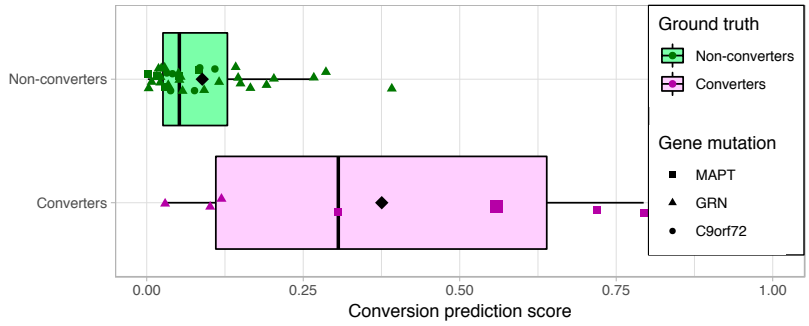


Figure 4.1 Converters' and non-converters' conversion prediction scores
Box and scatter plot of each subject's conversion prediction score on a scale from 0 (representing non-converter) to 1 (representing converter) after applying the FA model. Different gene mutations were represented with different shapes. The converter with the longest time between MRI and conversion (i.e., 41 months) was annotated with increased size. These conversion prediction scores result in a performance of 0.81 AUC for the FA model ($p = 0.025$ vs. chance level).
C9orf72, chromosome 9 open reading frame 72; FA, fractional anisotropy; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau.

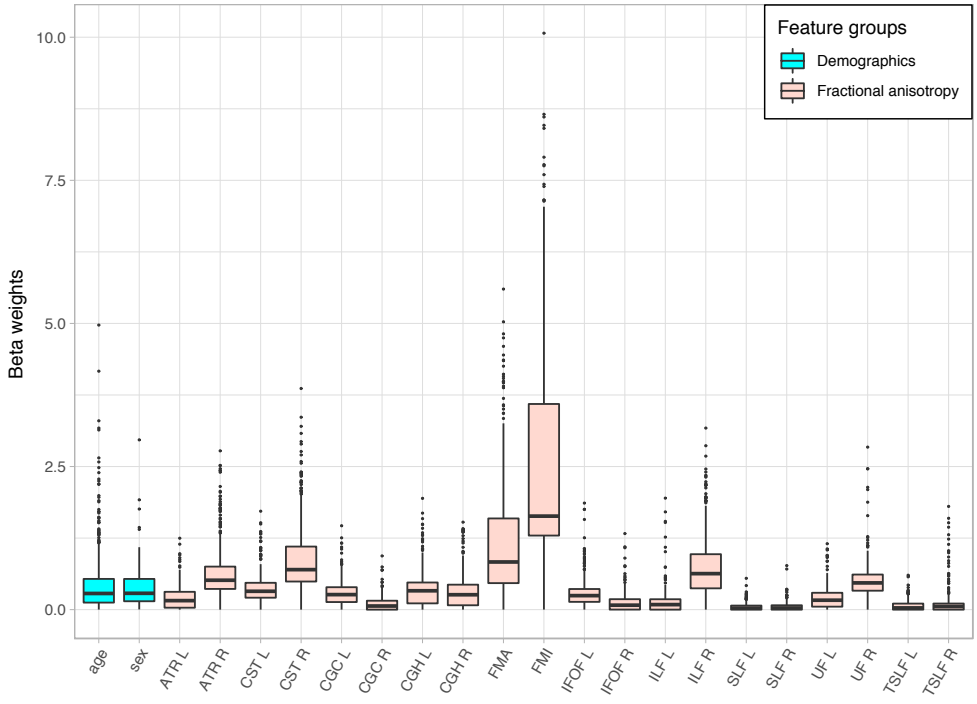


Figure 4.2 FA model beta weights
Box plots show the FA model's standardised beta weights for the 50 cross-validation repeats. Demographics (age and sex, in blue) were included in the model without penalty, while the FA features (red) were regularised.
ATR, anterior thalamic radiation; CST, corticospinal tract; CGC, cingulum in the cingulate gyrus area; CGH, cingulum in the hippocampal area; FA, fractional anisotropy; FMA, forceps major; FMI, forceps minor; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, left; R, right; SLF, superior longitudinal fasciculus; TSLF, temporal projection of the SLF; UF, uncinate fasciculus.

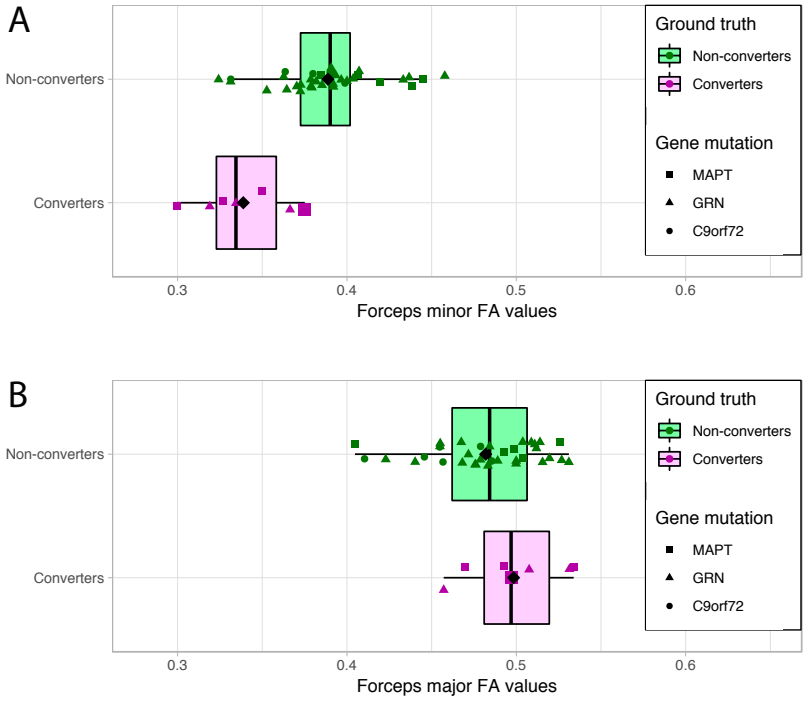


Figure 4.3 Mean FA values in two white matter tracts
Box and scatter plot of each subject's mean FA value in the forceps minor (**A**), and forceps major (**B**). Different FTD gene mutations were represented with different shapes. The converter with the longest time between MRI and conversion (i.e., 41 months) was annotated with increased size.
C9orf72, chromosome 9 open reading frame 72; FA, fractional anisotropy; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau.

Table 4.5 Cognitive features' performance

Cognitive domain	AUC	Min – max	Sensitivity	Specificity	Accuracy	FWERC *p*-value (AUC > chance)
Language	0.282	0.180 – 0.371	0.394	0.515	0.495	0.999
Attention	0.343	0.208 – 0.441	0.517	0.451	0.462	0.992
Executive	0.349	0.184 – 0.482	0.534	0.439	0.455	0.987
Social	0.344	0.208 – 0.449	0.474	0.488	0.486	0.992
Memory	0.272	0.143 – 0.359	0.491	0.414	0.427	0.999
Visuoconstruction	0.288	0.186 – 0.424	0.429	0.477	0.469	0.999
Multidomain	0.340	0.208 – 0.449	0.477	0.481	0.480	0.994

Converters (seven presymptomatic FTD-RisC mutation carriers that developed symptoms within 4 years after assessment) vs. non-converters (35 presymptomatic FTD-RisC mutation carriers that remained cognitively healthy after 4 years). Multidomain represents a combination of all cognitive features. Min – max AUC values represent the variance across the 50 repeats.
AUC, area under the receiver operating characteristic curve; FTD-RisC, frontotemporal dementia Risk Cohort; FWERC, family-wise error rate corrected.

Discussion

This study describes conversion prediction within 4 years in FTD mutation carriers using MRI-based and cognitive measures. We found that it was possible to distinguish converters from non-converters beyond chance level with an AUC of 0.81 using FA. Other MRI-based and cognitive measures did not outperform chance level.

The FA model's performance reaffirms the potential of diffusion scans for early FTD diagnosis. Group differences in FA and other DTI metrics exist in presymptomatic FTD mutation carriers compared to controls (Borroni et al., 2008; Dopper et al., 2014; Pievani et al., 2014; Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018; Jiskoot et al., 2018a, 2019). More specifically with regard to conversion, two recent FTD-RisC studies observed how mass univariate MRI measures and MRI-based classification scores develop as subjects approach conversion, as compared to subjects who did not convert during follow-up. Jiskoot et al. found that grey matter atrophy and white matter DTI changes become noticeable around two years before symptom onset, and longitudinal FA decline was primarily located in the genu of the corpus callosum (Jiskoot et al., 2019), the area with the higest beta values in our current study. Feis et al. combined multimodal MRI to calculate a single classification score at each time point using a classification model that was trained on bvFTD patients and controls. They found that the classification scores of presymptomatic mutation carriers and non-carriers did not differ over time, but that the classification scores of converters rose faster than those of non-converting mutation carriers (Feis et al., 2019b). Together, these studies suggest that FTD-related deterioration on MRI accelerates when subjects are near conversion. However, they fail to address the most important question: is it possible, using presymptomatic cross-sectional data, to predict future conversion in individual FTD mutation carriers? We expand on this by showing that individual differences in FA have predictive value for FTD onset within 4 years in FTD mutation carriers.

Within the FA model, the forceps minor feature had the highest beta weights across 50 repeats, suggesting that this area was important for prediction. Accordingly, the difference in mean FA between converters and non-converters was larger in the forceps minor than in other tracts, e.g., the forceps major. The genu of the corpus callosum (i.e., part of the forceps minor) was recently found to be the most consistent white matter region in terms of DTI changes across FTD subtypes (Elahi et al., 2017). As our converter group included both bvFTD and nfvPPA converters, this might explain why the forceps minor had high beta weights, and could provide biological support for our model. Forceps minor changes may therefore be a promising marker for early FTD detection, although it might not discriminate between the different FTD subtypes. Other white matter tracts are affected more specifically in the different clinical syndromes. BvFTD is generally characterised by DTI changes in the anterior thalamic radiation, cingulum, and uncinate fasciculus (Zhang et al., 2011; Mahoney et al., 2014; Möller et al., 2015; Daianu et al., 2016), while the superior longitudinal fasciculus and corticospinal tract are typically affected in nfvPPA (Whitwell et al., 2010; Galantucci et al., 2011; Omer et al., 2017), and the inferior longitudinal and uncinate fasciculi show DTI changes in svPPA (Whitwell et al., 2010; Galantucci et al., 2011; Tu et al., 2015). Future studies should investigate whether this combination of shared and different white matter DTI changes between the FTD subtypes facilitate hierarchical classification. Although it is interesting to examine which features were important for prediction, care must be taken not to overinterpret the beta weights. Contrary to explanatory regression models, the beta weights of our regularised prediction model do not designate direct relationships between the features and prediction score, nor do they reflect mean differences between the groups (Shmueli, 2010). Each feature's effect is

conditional on the effects of all other features in the model, and multicollinearity between features may result in suppression of the effect of some features. For example, we cannot be sure whether the higher beta weights in the FA model for right hemispheric white matter tracts point towards asymmetric effects—such as are reported in *GRN* mutation carriers (Jiskoot et al., 2018a)—or whether the corresponding regions in the left hemisphere added redundant information, and were therefore suppressed in the model.

Apart from the FA model, no other MRI- or cognition-based prediction model significantly outperformed chance. However, it would be inappropriate to draw conclusions based on the non-significant results. Due to our small sample size, it is unclear whether these features were uninformative, or whether the study was underpowered to find significant results. Grey matter atrophy occurs at a later stage than white matter changes in *MAPT* and *GRN* mutation carriers (Borroni et al., 2008; Agosta et al., 2012; Rohrer & Rosen, 2013; Feis et al., 2019a; Jiskoot et al., 2019), but the question whether or not that precludes the use of grey matter features for early diagnosis requires further investigation. Similarly, we cannot comment on why the remaining diffusion features MD, AxD, and RD, which often yield comparable results to FA, did not outperform chance level. None of the cognitive features predicted conversion within 4 years. However, it should be noted that our selection of neuropsychological tests was arbitrary, and higher performance might be possible with different tests. Also, longitudinal cognitive assessments may be more specific and predictive of conversion than cross-sectional differences (Jiskoot et al., 2018b).

Important strengths of this study include the unique population and follow-up, which enabled the prediction of conversion in genetic FTD. Our clinical follow-up demonstrated that non-converters remained asymptomatic even after 4 (and in 25 cases after 6) years, ensuring that the labels 'converter' and 'non-converter' were truly separate. Furthermore, our methods have been validated in previous studies concerning classification in AD and FTD (Schouten et al., 2016; Bouts et al., 2018). The most important limitation to this study was sample size. Our index group consisted of seven converters, which is a small number in neuroimaging studies, and even more so in machine learning. This resulted in performance estimates with a large degree of uncertainty, which means that relatively high performances were necessary to significantly outperform chance level. For example, the AUC needed to significantly outperform chance level after family-wise error rate correction was 0.78 for the MRI analyses, and 0.74 for cognitive analyses. Another limitation to this study was the sample's heterogeneity. We included subjects from *MAPT*, *GRN*, and *C9orf72* families in order to boost our sample size, even though there is evidence that each mutation has its own pattern of neurodegeneration over time due to different underlying pathology (Seelaar et al., 2011; Whitwell et al., 2012; Mann & Snowden, 2017; Jiskoot et al., 2018a). Indeed, converters with a *MAPT* mutation seemed to have higher conversion prediction scores than did converters with a *GRN* mutation, though this may also be due to sex differences. Stratification was not feasible due to our small sample size, but is necessary in future studies to show whether conversion prediction is possible in all FTD gene mutations or solely in *MAPT*. Just as different FTD mutations have different neurodegenerative profiles, so do the different clinical syndromes that constitute FTD (Seelaar et al., 2011; Whitwell et al., 2012; Rohrer & Rosen, 2013). Five subjects developed bvFTD, while two developed nfvPPA, increasing heterogeneity in our analyses. Despite these sources of heterogeneity, the FA-based model predicted conversion in 4 years beyond chance level, which is an important proof of concept that conversion prediction is possible in genetic FTD. Larger sample sizes might in the future facilitate hierarchical (Kim et al., 2019) or multilabel (Raamana et al., 2014; Klöppel et al., 2015; Koikkalainen et al., 2016; Bron et al., 2017; Canu et al., 2017) prediction to deal with these sources of heterogeneity. While the acquisition of a larger, similar cohort is costly and time-

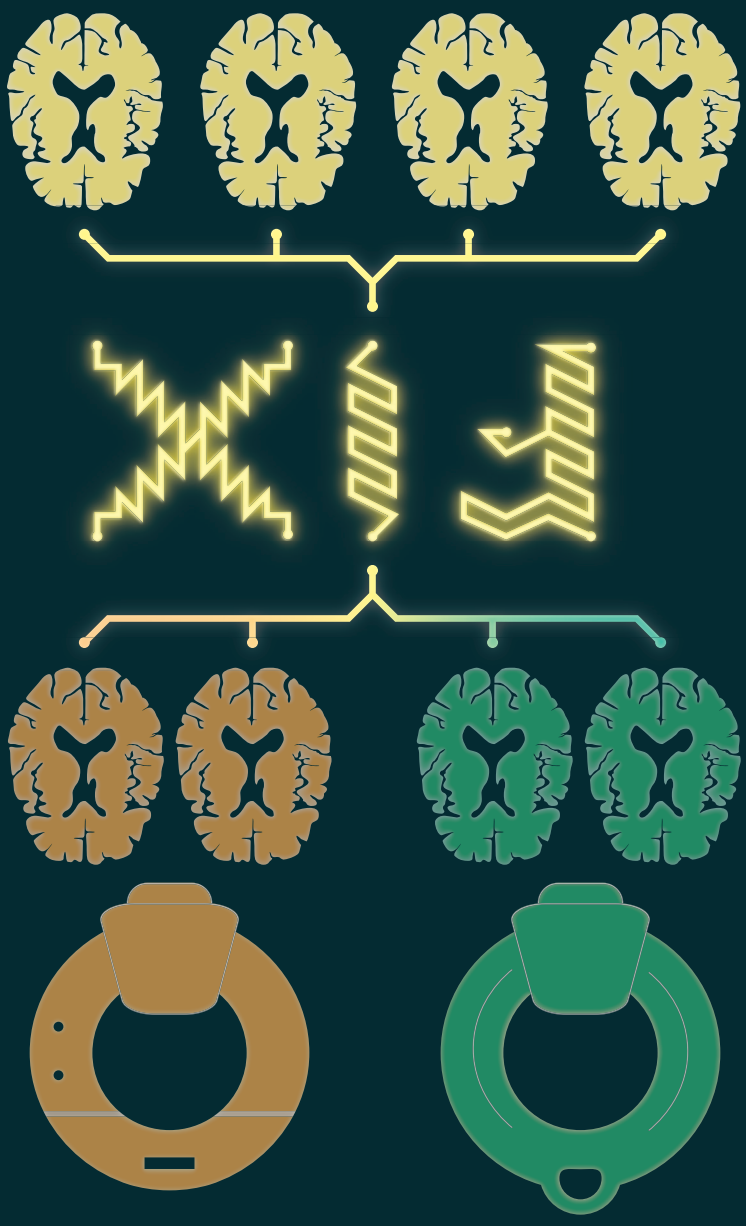
consuming, our current results indicate that the acquisition of such data is meaningful and may lead to more accurate conversion prediction for genetic FTD. Lastly, the timing of conversion, and, therefore, the time between MRI and conversion, was based on heteroanamnestic information from knowledgeable informants, and as such may not be fully accurate. However, it should be noted that an exact time of conversion is near impossible to estimate in neurodegenerative diseases due to the gradual increase in symptomatology, and this inaccuracy is unlikely to have influenced our conclusions.

Conclusion

To conclude, we showed that FA predicted conversion within 4 years in presymptomatic FTD mutation carriers. This proof-of-concept study underlines the potential of MRI-based prediction in genetic FTD to contribute to a reliable early-stage FTD diagnosis, and should be replicated when larger sample sizes become available to corroborate our results.

Acknowledgements

Our appreciation goes out to all FTD-RisC participants and families. The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to RAF), ZonMw programme Memorabel project 733050103, JPND PreFrontAls consortium project 733051042 (to JCvS), and a VICI grant 016–130–667 from The Netherlands Organisation for Scientific Research (NWO; to SARBR). The views expressed are those of the authors and not necessarily those of the funding sources. The funding sources were not involved in the design of the study; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The authors report no conflict of interest.



Chapter 5

Methods: multicentre data harmonisation through artefact removal in resting-state functional MRI

Published in *Frontiers in Neuroscience 2015;9:395* as:

ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI

Rogier A. Feis, Stephen M. Smith, Nicola Filippini, Gwenaëlle Douaud, Elise G.P. Dopper, Verena Heise, Aaron J. Trachtenberg, John C. van Swieten, Mark A. van Buchem, Serge A.R.B. Rombouts, Clare E. Mackay

Abstract

Resting-state functional magnetic resonance imaging (rs-fMRI) has shown considerable promise in providing potential biomarkers for diagnosis, prognosis, and drug response across a range of diseases. Incorporating rs-fMRI into multicentre studies is becoming increasingly popular, imposing technical challenges on data acquisition and analysis, as fMRI data is particularly sensitive to structured noise resulting from hardware, software, and environmental differences.

Here, we investigated whether a novel clean-up tool for structured noise was capable of reducing centre-related rs-fMRI differences between healthy subjects. We analysed 3 T rs-fMRI data from 72 subjects, half of whom were scanned with eyes closed in a Philips Achieva system in The Netherlands, and half of whom were scanned with eyes open in a Siemens Trio system in the UK. After prestatistical processing and individual independent component analysis (ICA), FMRIB's ICA-based X-noiseifier (FIX) was used to remove noise components from the data. Group-level ICA and dual regression were run, and non-parametric statistics were used to compare spatial maps between groups before and after applying FIX.

Large significant differences were found in all resting-state networks between study sites before using FIX, most of which were reduced to non-significant after applying FIX. The between-centre difference in the medial/primary visual network, presumably reflecting a between-centre difference in protocol, remained statistically significant.

FIX helps facilitate multicentre rs-fMRI research by diminishing structured noise from rs-fMRI data. In doing so, it improves combination of existing data from different centres in new settings and comparison of rare diseases and risk genes for which adequate sample size remains a challenge.

Keywords: resting-state functional MRI, multicentre analysis, independent component analysis, dual regression, structured noise reduction

Introduction

Resting-state functional magnetic resonance imaging (rs-fMRI) has become an important tool in neuroimaging research to examine resting-state networks (RSNs) in normal brains, during the ageing process, and in various neurological disorders (Greicius et al., 2003; Fox et al., 2005; De Luca et al., 2006; Fox & Raichle, 2007; Littow, 2010). One of the techniques used for this purpose is independent component analysis (ICA)—a data-driven technique that facilitates comparison of functional networks in the brain without requiring a priori selected seed regions (Beckmann & Smith, 2004).

(Rs-)fMRI research has certain challenges, such as problems regarding sample size in clinical and at-risk populations. Multicentre analysis may help to solve these limitations, but has been shown to be difficult to perform for (rs-)fMRI. Specifically, differences between groups may not always be attributable to the feature of interest, such as disease or gene mutation carrier status, but may also be secondary to scanner hardware differences (manufacturer, head coil), software differences (filters, k-space acquisition method, scan parameters) and environmental differences (radio-frequency noise; Casey et al., 1998; Zivadinov & Cox, 2008). Confounding centre effects also manifest as noise in multicentre analysis, reducing power.

Whilst several studies have investigated and provided guidelines and recommendations for these difficulties for fMRI (Zou et al., 2005; Costafreda et al., 2007; L. Friedman et al., 2008; Wegner et al., 2008; Zivadinov & Cox, 2008; Glover et al., 2012), multicentre research using rs-fMRI is still an underdeveloped field. Long et al. (2008) were able to cross-validate the default mode network in a multicentre study even though scanner parameters were not controlled. Biswal et al. (2010) demonstrated that functional connectivity has a universal architecture in an extensive study with 1,414 subjects. However, they also found many differences due to centre-related variability. As data sharing becomes more important in imaging research, e.g., the Genetic Frontotemporal dementia Initiative (GENFI; Rohrer et al., 2013); 1,000 Functional Connectome Project (Biswal et al., 2010); ADHD 200 Consortium data set (The ADHD-200 Consortium, 2012); and Autism Brain Imaging Data Exchange (ABIDE; Di Martino et al., 2014), methods for reducing scan site differences must be developed.

In the current study, a novel tool for the clean-up of structured noise-components from ICA was used to study whether rs-fMRI data from different scan sites become more comparable in a multicentre analysis. The Functional Magnetic Resonance Imaging of the Brain Centre's (FMRIB's) ICA-based X-noiseifier (FIX) is a plug-in to FMRIB Software Library (FSL) that is able to automatically classify and remove structured noise-components (e.g., motion-effects, scanner artefacts, [non-neuronal] physiological noise, etc.) from rs-fMRI data, once it has been trained through hand-classifications (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). FIX has been used before to clean up structured noise in order to heighten the quality of rs-fMRI data (Salimi-Khorshidi et al., 2014), but this is the first time FIX is used to diminish scanner differences in a multicentre study.

Methods

Participants

In this study, MRI data was included from subjects scanned at the Leiden University Medical Centre (LUMC) and from subjects scanned at the Oxford Centre for Clinical Magnetic Resonance Research (OCMR). The LUMC data (referred to here as the 'Dutch sample') consisted of 36 subjects from the control group of an earlier rs-fMRI study investigating the effect of microtubule-associated protein tau (*MAPT*) and progranulin (*GRN*), risk genes for Frontotemporal Dementia (FTD), on the brain (Dopper et al., 2014). The OCMR data (referred to here as the 'English sample') consisted of 36 subjects from control groups of earlier rs-fMRI studies investigating the effect of apolipoprotein E ε4 (*APOE4*), a risk gene for Alzheimer's Disease (AD), on the brain (Filippini et al., 2009, 2011; Heise et al., 2011; Trachtenberg et al., 2012a, 2012b).

The English subjects were selected from a larger cohort in order to match the Dutch subjects in age, sex and sample size.

For a detailed description of the recruitment protocols, see Dopper et al. (2014) for the Dutch data, and Filippini et al. (2009, 2011) and Trachtenberg et al. (2012a) for the English data. In short, 36 *MAPT* and *GRN* non-carriers were selected from a pool of 160 healthy first-degree relatives of FTD patients with either an *MAPT* or *GRN* mutation. It is assumed that non-carriers from these families have the same risk for dementia as the general population. Thirty-six *APOE4* non-carriers, scanned at the OCMR, were selected from the general population and the data from most (30/36) were reported in previous studies (Filippini et al., 2009, 2011; Trachtenberg et al., 2012a).

Pre-scan exclusion criteria included MRI contraindications, history of drug abuse, and current or past neurologic or psychiatric disorders for the Dutch sample, and head injury, substance abuse (including alcohol), corticosteroid therapy, youth diabetes therapy, memory complaints, and current or past neurologic or psychiatric disorders for the English sample.

All participants provided written informed consent, and ethical approval for data acquisition was obtained from National Research Ethics Service Committee South Central—Oxford C (Oxford data), and the Medical Ethical Committees in Rotterdam and Leiden (Leiden data).

Image acquisition

LUMC scans were acquired using a Philips 3 T Achieva MRI scanner with an 8-channel SENSE head coil. OCMR scans were acquired using a Siemens 3 T Trio scanner with a 12-channel head coil. Participants were instructed to keep their eyes closed (LUMC) or open (OCMR), to think of nothing in particular (OCMR), and to remain awake. The scan parameters used for the high-resolution 3-dimensional anatomical T_1-weighted and for the rs-fMRI T_2*-weighted images are shown in **Table 5.1**.

Image analysis

FSL (http://www.fmrib.ox.ac.uk/fsl) tools were used for all data analyses (Smith et al., 2004; Woolrich et al., 2009; Jenkinson et al., 2012).

Prestatistical processing

Individual preprocessing included motion correction (Jenkinson et al., 2002), brain extraction (Smith, 2002), spatial smoothing using a Gaussian kernel of 6 mm full width at half maximum,

4D grand-mean scaling, and high-pass temporal filtering corresponding to a period of 150 s (~ 0.007 Hz). FMRI volumes were registered to MNI152 standard space (Montreal Neurologic Institute average T_1-weighted image created from 152 normal subjects' T_1 scans). Boundary-based registration (Jenkinson & Smith, 2001; Jenkinson et al., 2002; Greve & Fischl, 2009) was used to register each individual's echo-planar imaging volumes onto their respective high-resolution T_1-weighted structural images. T_1-weighted structural scans were aligned to MNI152 standard space using nonlinear image registration (Anderson et al., 2007; Jenkinson et al., 2012). The resulting registration matrices were then used to register the echo-planar imaging volumes onto MNI152 standard space. Individual ICA was carried out and voxel-by-voxel intensity normalisation was performed manually, dividing each voxel by its mean value across time and multiplying by 10,000.

Table 5.1 Structural and functional scan parameters per scan site

Parameters	Structural		Resting-state	
	LUMC	**OCMR**[a]	**LUMC**	**OCMR**
Slices, *n*	140	192	38	34
TR, ms	9.8	2,040	2,200	2,000
TE, ms	4.6	4.7	30	28
Flip angle, °	8	8	80	89
Volumes, *n*	–	–	200	180
Voxel size, mm	0.88 × 0.88 × 1.20	1.0 × 1.0 × 1.0	2.75 × 2.75 × 2.99[b]	3.0 × 3.0 × 3.5
Duration, min	5	6	8	6

FOV, field of view; LUMC, Leiden University Medical Centre; OCMR, Oxford Centre for Clinical Magnetic Resonance Research; TE, echo time; TR, repetition time.

[a] Structural scanning at OCMR was done using a magnetisation-prepared rapid gradient echo sequence (MPRAGE).

[b] Including 10% interslice gap.

FIX

Network components obtained from the individual ICA were visually judged and were labeled as signal, noise, or unknown for 12 subjects from each group. Manual classification was done by looking, firstly, at their spatial maps (typically thresholded abs(Z) > 2.3), then at the temporal power spectrum, and lastly at their time series. Unthresholded spatial maps were examined when necessary (Salimi-Khorshidi et al., 2014).

Using these classifications, the FIX classifier was trained and a training file was created. As described by Salimi-Khorshidi et al. (2014), FIX uses over 180 features, capturing components' spatial and temporal characteristics, which are fed into a multilevel classifier (built around several different classifiers). Temporal features include autoregressive properties, distributional properties, jump amplitudes, the Fourier transform, and the time series' correlation with grey matter-, white matter-, cerebrospinal fluid- and head motion-derived time series. Spatial features include clusters' sizes and spatial distribution, voxel intensity information indicating whether voxels are grey matter or, e.g., blood vessels, percent on brain boundary, hand-created mask-based features for components that have signal-like spatiotemporal characteristics (such as sagittal sinus, cerebrospinal fluid, and white matter), and other spatial features such as spatial smoothness.

Next, a leave-one-out test was run in order to control the quality of the classifier algorithm by estimating the level of agreement of the hand-labeled classifications and the classifier's classifications. The accordance was measured as a true-positive rate (TPR), a true-negative rate (TNR), and a composite measure, i.e., $(3 \times TPR + TNR) \div 4$, for a range of thresholds (used to determine the binary classification of components since FIX's output is probabilistic). After checking the TPR, TNR, and the composite measure, the optimal threshold (i.e., 20) was chosen and the classifier was applied to all subjects' data using this threshold in order to classify and remove the structured noise components from the data (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014).

Group-level ICA

After prestatistical processing and FIX, three group-level ICA (GICA) analyses were run using MELODIC. In order to qualitatively compare FIX's effect on GICA components, GICA was run on combined English and Dutch data before application of FIX (GICA-1), and on combined English and Dutch data after application of FIX (GICA-2). For statistical analysis of FIX's effect on the multicentre differences, GICA was carried out on all data combined (GICA-3). Consequently, the data used for this analysis (GICA-3) contained four subgroups: Dutch subjects with and without use of FIX, and English subjects with and without use of FIX. Rs-fMRI data were temporally concatenated across individuals to create a single 4D data set. The data were whitened and principle component analysis was used to project the data into a 25-dimensional subspace, matching many previous rs-fMRI studies (Filippini et al., 2009; Smith et al., 2009; Cocozza et al., 2015; Gaudio et al., 2015). By optimising for non-Gaussian source estimates through a fixed-point iteration technique, we obtained component maps (Hyvärinen, 1999). After transforming the component maps to *Z*-maps, Gaussian/Gamma Mixture Models were fitted to them in order to obtain 25 independent spatial maps defining functional connectivity patterns across the participants (Beckmann & Smith, 2004). The GICA-derived spatial maps were then judged by eye and divided into RSN and noise components.

Dual regression

Analysis of group differences was performed using FSL's dual regression, a regression technique that allows for voxel-wise comparisons of rs-fMRI (Filippini et al., 2009; Veer et al., 2010). All spatial maps derived from GICA-3 (using English and Dutch FIX and non-FIX data) were regressed against each individual's preprocessed rs-fMRI data, resulting in a time course for each component and subject. The produced time courses were regressed against the same individual's preprocessed rs-fMRI data, resulting in subject-specific spatial maps for parameter estimates (PEs) and *Z*-stats. GICA noise component maps were disregarded, and RSN component maps were collected across subjects into 4D files (one per ICA component, with the fourth dimension being subject identification) and were tested voxel-wise for statistically significant differences between groups. We used a general linear model equivalent to an independent *t*-tests to test the PE- and *Z*-stat-driven spatial maps for differences between Dutch and English groups before use of FIX, differences between Dutch and English groups after use of FIX, and the interaction between the use of FIX and group differences (by comparing the differences before and after use of FIX to each other). Age and years of education were added to the analysis as confound regressors. Non-parametric permutation-based testing was done by running 5,000 random permutations using the randomise algorithm, a tool based on the Freedman-Lane methods within FSL (Winkler et al., 2014). Afterwards, threshold-free cluster enhancement (TFCE), a method for finding clusters in data without defining

clusters in a binary way, was applied (Smith & Nichols, 2009), and a family-wise error corrected cluster significance threshold of $p < 0.05$ was used. In a more qualitative approach, non-family-wise error corrected results and raw *t*-stat maps were also investigated.

Result masking

In order to fully appreciate the impact that FIX has on the data, results of the differences between groups for all components were thresholded, binarised, and merged. The resulting imaging volumes display the total number significant voxels for all different components together, with colour variation showing the number of components with significant change in each voxel.

Statistical analysis

Statistics of non-imaging variables were performed using SPSS version 20 (SPSS, Chicago, IL). Demographic variables were tested using independent *t*-tests for continuous variables and chi-square tests for categorical variables.

Results

Sample demographics

Demographic information for the Dutch and English subjects is shown in **Table 5.2**. Age and sex were matched across groups.

Table 5.2 Participant demographics

	OCMR (*n* = 36)	LUMC (*n* = 36)	*p*-value
Age, mean (SD) years	49.9 (11.5)	49.8 (11.3)	0.94
Sex, *n* (%) ♀	19 (52.8)	18 (50.0)	1.0
Education, mean (SD) years[a]	16.6 (3.2)	12.6 (2.9)	<0.001

LUMC, Leiden University Medical Centre; OCMR, Oxford Centre for Clinical Magnetic Resonance Research.

[a] Scores of education level in years were missing for two individuals (both LUMC subjects).

Individual ICA and FIX

Table 5.3 shows the number of extracted independent components by individual ICA for each group (OCMR and LUMC), as well as the number of components classified as noise and RSN by FIX. Significantly more independent components were extracted from Dutch data, compared to English data. Also, significantly more components from Dutch data were classified as noise by FIX. The number of components classified as RSN by FIX was not found to be different between groups.

Table 5.3 FIX classifications

	OCMR (*n* = 36)	LUMC (*n* = 36)	*p*-value
ICs, mean (SD)	36.1 (4.8)	44.3 (7.9)	<0.001
Noise ICs, mean (SD)	23.6 (3.9)	31.8 (8.2)	<0.001
RSN ICs, mean (SD)	12.6 (3.0)	12.7 (3.0)	0.88

FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; IC(A), independent component (analysis); LUMC, Leiden University Medical Centre; OCMR, Oxford Centre for Clinical Magnetic Resonance Research; RSN, resting-state network.

Group-level ICA

Figure 5.1 shows spatial maps derived from GICA for data before (GICA-1, **Figure 5.1A**) and after (GICA-2, **Figure 5.1B**) application of FIX (numbers in text correspond to numbers in figure). RSN components are shown with a green frame, whereas noise components are shown with a red frame. FIX's effect on GICA seems to be two-fold: some noise components are eliminated (i.e., motion artefacts [**5.1A**: numbers 6, 22, 23], brain stem/vascular artefacts [**5.1A**: numbers 14, 18, 25], and sagittal sinus artefacts [**5.1A**: numbers 8, 19]) and others are 'pushed back' (i.e., have a higher index number after the use of FIX: white matter [**5.1A**: number 4, **5.1B**: number 23] and frontal sinus susceptibility noise [**5.1A**: number 9, **5.1B**: number 21]). Both observations rely on the same mechanism: FIX removes variance explained by noise components from the data. As

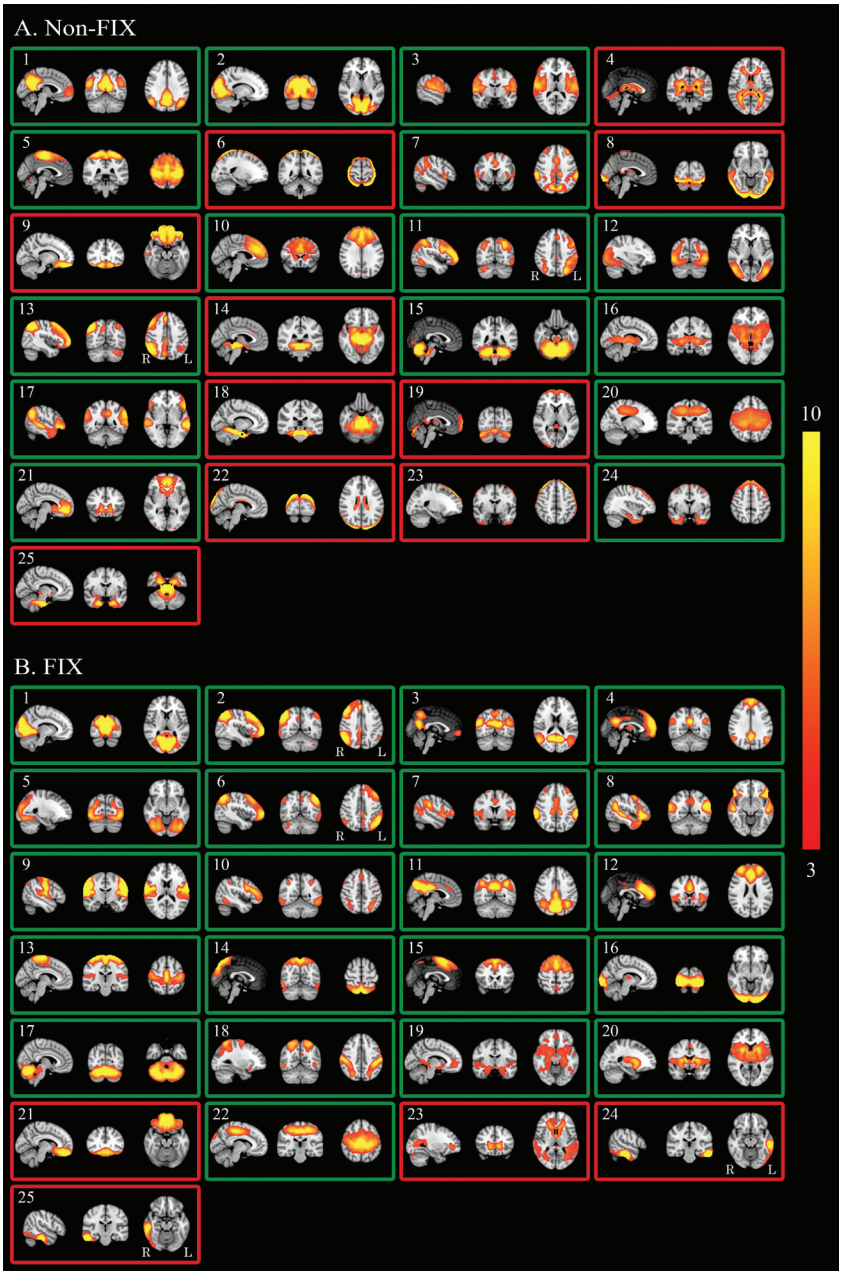


Figure 5.1 GICA spatial maps before and after FIX
Maps illustrate the 25 GICA networks' most informative orthogonal slices before (**A**, GICA-1) and after (**B**, GICA-2) applying FIX. Green frames indicate RSNs; red frames indicate noise networks. Colour bar represents *Z*-scores.
FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; GICA, group-level ICA; ICA, independent component analysis.

MELODIC shows components in order of decreasing explained variance, the removal of variance explained by noise components results in higher component numbers or even exclusion.

Spatial maps that were used as spatial regressors for dual regression (GICA-3) are shown in **Figure 5.2**. Identified RSNs were the default mode network (1), primary/medial (2) and lateral (7, 13) visual networks, lateralised higher order cognitive networks involved with working memory (3, 5), a network showing the dorsal attention network combined with the salience network (4), the auditory network (6), a network combining features of the default mode network and the ventral stream (8), the executive control network (9), networks that describe different parts of the sensorimotor network (10, 11), cerebellar network (14), a network describing the basal ganglia (17), and a network showing frontal default mode network features as well as features from the executive control network (21).

Dual regression

All RSNs' combined results based on PE-driven spatial maps are shown for family-wise error corrected group differences before the use of FIX (**Figure 5.3A**), group differences after the use FIX (**Figure 5.3B**), and for the interaction between applying FIX and group differences (**Figure 5.3C**). Dual regression results for each RSN are shown separately in **Supplemental Figure S5.1** (numbers in **Supplemental Figure S5.1** correspond with numbers in **Figure 5.2**).

Before the use of FIX, large areas of statistically significant differences were shown in all (15) RSNs. After applying FIX, the size and number of areas with significant differences between groups was strongly reduced: only 7/15 RSNs showed statistically significant differences and the number of significantly different voxels was reduced by 98%. The RSN with the largest area of significant differences after using FIX was the primary/medial visual network (PVN), containing 85% of all significantly different voxels after applying FIX. This network is associated with a difference in scan protocol (eyes open vs. closed) and showed greater activation in English than in Dutch subjects.

The interaction between the use of FIX and site differences was significant in 13/15 RSNs.

Dual regression results based on *Z*-stat-driven spatial maps were similar on visual inspection.

Additionally, for a more qualitative view of the results, **Supplemental Figure S5.2** shows dual regression results without family-wise error correction for each component. Another point of view on FIX's effect is offered in **Supplemental Figure S5.3**, demonstrating a reduction in raw *t*-stats for group differences in each component after applying FIX.

Figure 5.2 GICA spatial maps for statistical analysis
Maps illustrate the 25 GICA networks' most informative orthogonal slices of data before and after applying FIX combined (GICA-3). Green frames indicate RSNs; red frames indicate noise networks. Colour bar represents *Z*-scores.
FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; GICA, group-level ICA; ICA, independent component analysis.

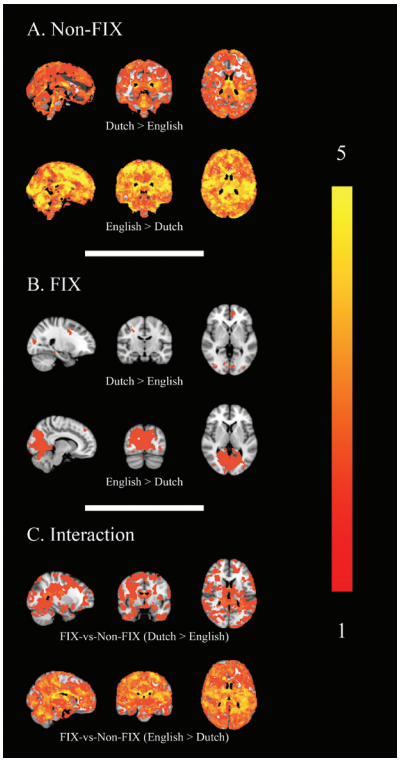


Figure 5.3 Combined group differences
Maps show statistically significant ($p < 0.05$) differences between groups: without the use of FIX (**A**), after the use of FIX (**B**), and the interaction between FIX and group differences (**C**) in all (15) RSNs combined. Colour bar represents the number of significantly differing networks.
FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; ICA, independent component analysis; RSN, resting-state network.

Discussion

In this study, FIX was found to be helpful in the comparison of multicentre rs-fMRI data. FIX significantly reduced structured noise resulting from hardware, software, and environmental differences in a multicentre group comparison, as demonstrated by **Figure 5.3** and **Supplemental Figures S5.1, S5.2**. Additionally, **Supplemental Figure S5.3** shows an intra- and inter-component reduction in raw *t*-stat variability after applying FIX. The significant interaction between the application of FIX and group differences (**Figure 5.3** and **Supplemental Figure S5.1**) shows that site differences are not just pushed below significance threshold, but are significantly changed by applying FIX. Importantly, the remaining differences between sites after FIX (**Figure 5.3** and **Supplemental Figure S5.1**) are primarily confined to the primary/medial visual cortex, which reflects differences in experimental design (Dutch participants had eyes closed, whereas English participants had eyes open). This implies that FIX removes structured noise, but retains physiologically driven differences.

Dual regression is usually run using PE-driven spatial maps; alternatively, *Z*-stat-driven spatial maps can be used. Our results using PE- and *Z*-stat-driven spatial maps were similar on visual inspection, suggesting that the use of FIX is of value for both types of analysis. However, in order to assess whether FIX works better for either one, a more specific analysis is required.

Structured noise in fMRI has various origins: hardware differences (e.g., scanner manufacturer, type of head coil), software differences (filters, k-space acquisition methods, and scan parameters), and radio-frequency noise (Casey et al., 1998; Zivadinov & Cox, 2008). As demonstrated in **Figure 5.3** and **Supplemental Figures S5.1–3**, FIX helps to deal with noise from these origins, inasmuch as they present themselves as separate noise components in individual subjects' ICA. Still, it cannot account for all potential between-site differences. For example, it cannot deal with differences that present themselves within RSN components such as differences in sensitivity to RSNs based on hardware configurations or RSN spatial variability relating to head coils. However, due to the nature of ICA, the most striking differences caused by structured noise are presented as separate noise components. Therefore, intra-component variability is much smaller than inter-component variability, implying that the scope of this drawback is altogether limited. Also, FIX cannot account for differences in the magnitude of the blood-oxygen-level-dependent effect. Voxel-wise intensity normalisation may help to reduce this problem, but site-wise confound regressors should be used when they do not correlate with the regressors of interest.

Although eyes-open and eyes-closed differences cannot be mathematically disentangled from site/scanner differences, we suggest that the remaining differences in the PVN after using FIX are a manifestation of this protocol discrepancy since the differences in all other networks are substantially reduced. Recently, a number of studies have investigated the effect of eyes-open vs. eyes-closed conditions on functional connectivity without yet reaching a clear consensus. For example, reduced activation (Feige et al., 2005), amplitude of low frequency fluctuations (ALFF; Yang et al., 2007; Yan et al., 2009; Liu et al., 2013; Liang et al., 2014; Yuan et al., 2014), and regional homogeneity (Liu et al., 2013) have all been reported in eyes-closed relative to eyes-open conditions. Conversely, other studies showed higher blood-oxygen-level-dependent response (McAvoy et al., 2008) and higher mean ALFF (Jao et al., 2013) for the eyes-closed condition, or no difference in seed-based correlations (Patriat et al., 2013). Aside from these local changes in functional connectivity between conditions, Jao et al. (2013) discovered that the mean ALFF of the whole brain was greater in eyes-closed vs. eyes-open conditions. Some of these studies also reported functional connectivity differences in other networks than the PVN, including the sensorimotor, default mode, and

auditory networks. The family-wise error corrected changes found in this study in non-PVN networks were small, scattered and generally did not follow the independent component's anatomy closely. Therefore, it is difficult to infer if these changes are due to the experimental design or if they reflect a small quantity of remaining noise. The changes we found in independent component 2 are extensive and follow the PVN anatomy well. As it is unlikely that false positive results or leftover noise would take this form and since similar PVN differences between eyes-open and eyes-closed conditions have been described in rs-fMRI multiple times before, it seems reasonable to assume that this effect is due to reported differences in experimental design.

Whilst groups were matched for age and sex, there was a significant difference in level of education. This may be attributable to the recruitment protocols. The English recruitment protocol selected subjects from the general population near Oxford, a relatively highly educated region (Filippini et al., 2009, 2011; Trachtenberg et al., 2012a), whereas the Dutch sample was recruited from known FTD families (Dopper et al., 2014). In order to account for this, demeaned education values were added to the general linear model as a regressor of no interest.

Previous studies on multicentre fMRI primarily focused on data collected using standardised protocols. Glover et al. (2012) argue that hardware, software and procedural aspects should be carefully matched and managed in order to successfully perform multicentre fMRI research. Zivadinov & Cox (2008) suggest the use of quality assurance methods and careful subject selection and matching across centres, such as used by Wegner et al. (2008), in order to control for scan site by adding centre as covariate in the analysis. Whereas these recommendations are naturally important for the correct set-up of a new multicentre study, our results suggest reanalysis of existing non-standardised rs-fMRI data may also be possible across sites. Additionally, although it would be interesting to see how these different sources of structured noise are dealt with individually by FIX (whilst controlling for the others), this study importantly shows that even with more of these problems present simultaneously, FIX adequately diminishes structured noise.

Conclusion

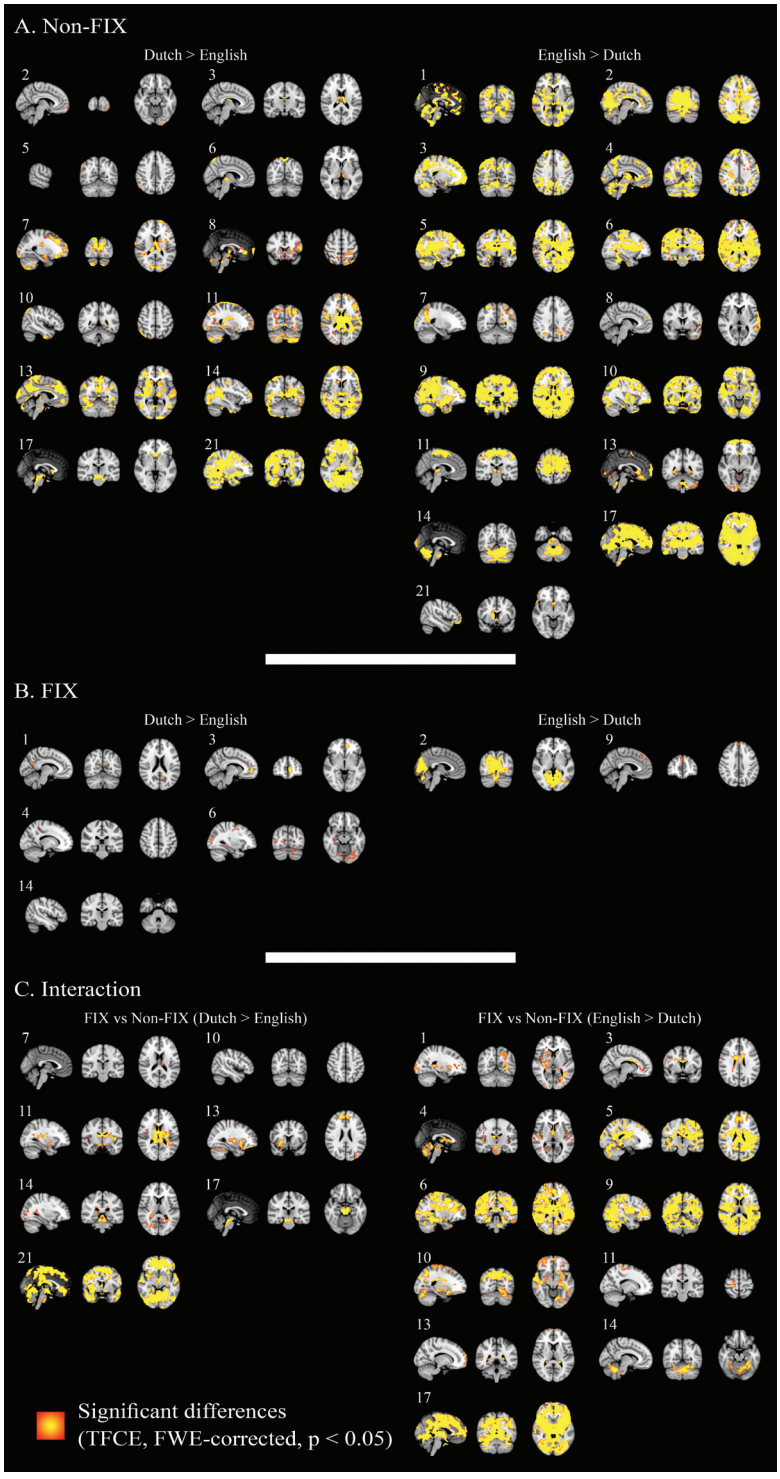
Previous studies using FIX have considered the theoretical and practical use of spatial ICA, classifier training and noise detection (Salimi-Khorshidi et al., 2014), and denoising (Griffanti et al., 2014). They showed that FIX is a useful tool for noise clean-up and therefore helps in making data more sensitive to changes related to neuronal activity. This study is the first to show FIX's additional value in multicentre rs-fMRI analysis. By improving multicentre fMRI research and efficient reanalysis of acquired data, comparisons of rare diseases and at-risk populations will be more efficient and convenient, leading to a better insight in neurological disorders. Furthermore, as free data sharing is an upcoming way to create large rs-fMRI data sets (Biswal et al., 2010; The ADHD-200 Consortium, 2012; Di Martino et al., 2014), FIX may be a valuable tool to ensure valid comparison of data acquired at different centres.

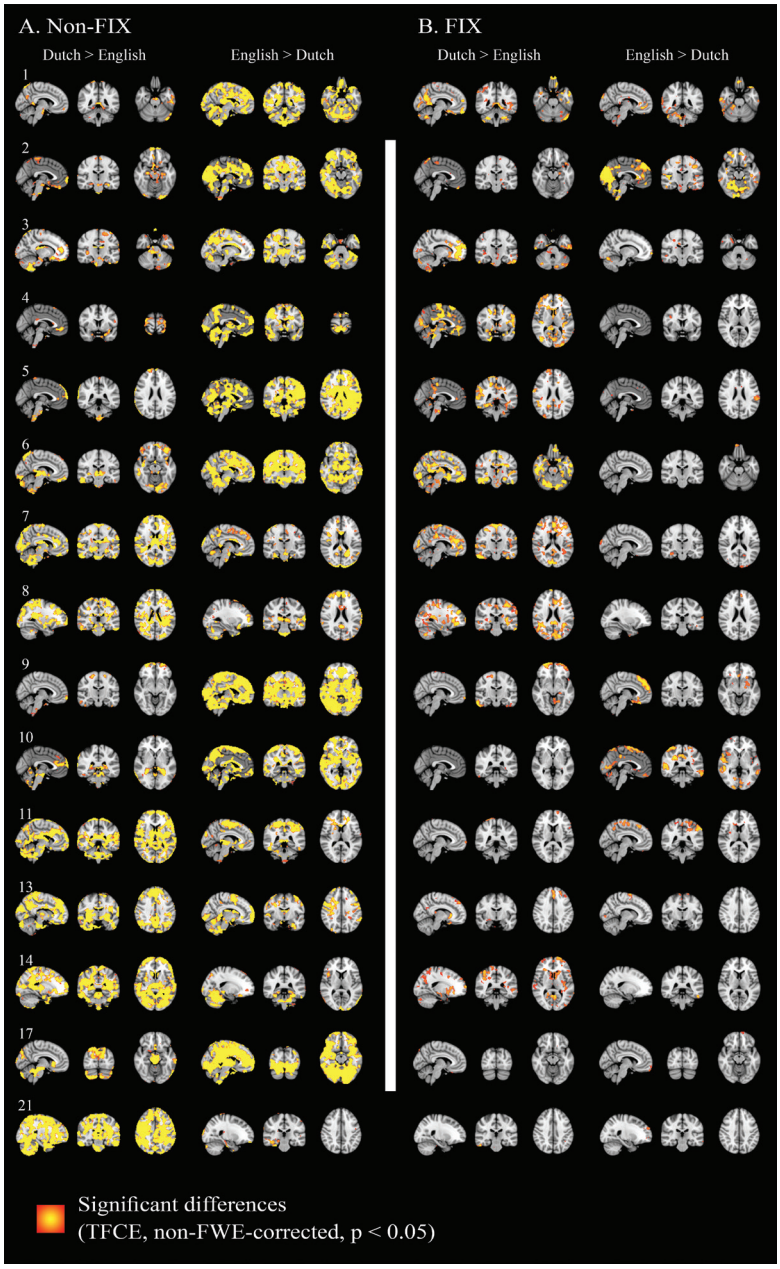
Acknowledgements

Our appreciation goes out to all study participants. The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to RAF); Wellcome Trust (to SMS); HDH Wills 1965 charitable trusts (to NF, English Charity Register 1117747); the Medical Research Council (MRC) UK (MR/K006673/1; to GD); the Dioraphte Foundation (grant 09–02–03–00), the Association for Frontotemporal Dementias Research Grant 2009, and The Netherlands Alzheimer Foundation (to EGPD and JCvS); The Netherlands Organisation for Scientific Research (NWO) grant HCMI 056–013–018 (to EGPD) and VICI project (grant number 016–130–667, to SARBR); an Alzheimer's Research UK studentship (to VH, English Charity Register 1077089); a Rhodes scholarship (to AJT); and the National Institute for Health Research (NIHR), UK as part of the Oxford Biomedical Research Centre (BRC; to CEM). The views expressed are those of the authors and not necessarily those of the NWO, the NHS or the NIHR. The funding sources were not involved in the design of the study; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The authors report no conflict of interest.

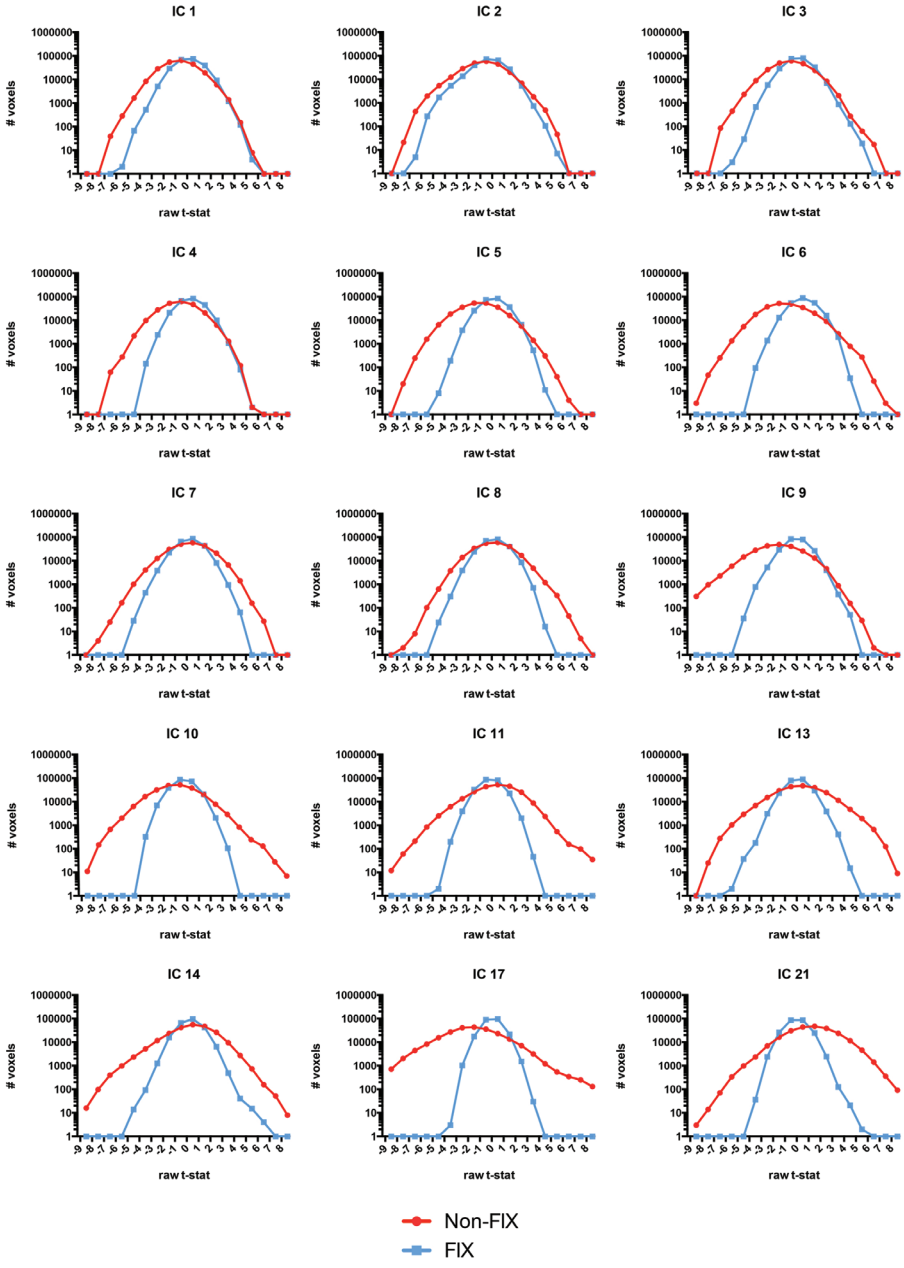
Supplemental material

Supplemental Figure S5.1 Separate FWE corrected group differences (opposite page)
Maps show FWE corrected statistically significant ($p < 0.05$) differences between groups: without the use of FIX (**A**), after the use of FIX (**B**), and the interaction between FIX and group differences (**C**) for each of 15 RSNs. Colour bar represents significance.
FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; FWE, family-wise error; ICA, independent component analysis; RSN, resting-state network; TFCE, Threshold-free cluster enhancement.


A. Non-FIX
Dutch > English
English > Dutch
B. FIX
Dutch > English
English > Dutch
C. Interaction
FIX vs Non-FIX (Dutch > English)
FIX vs Non-FIX (English > Dutch)
Significant differences
(TFCE, FWE-corrected, p < 0.05)



Supplemental Figure S5.2 Separate uncorrected group differences
Maps show uncorrected statistically significant ($p < 0.05$) differences between groups: without the use of FIX (**A**) and after the use of FIX (**B**) for each of 15 RSNs. Colour bar represents significance.
FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; ICA, independent component analysis; RSN, resting-state network; TFCE, Threshold-free cluster enhancement.



Supplemental Figure S5.3 Raw *t*-stats variability
Graphs show raw *t*-stats between groups (positive *t*-stats signify Dutch > English, negative *t*-stats signify English > Dutch) before and after FIX for each of 15 RSNs on a logarithmic scale.
FIX, FMRIB's ICA-based X-noiseifier; FMRIB, Functional Magnetic Resonance Imaging of the Brain Centre; ICA, independent component analysis; RSN, resting-state network.

Chapter 6

Frontotemporal dementia vs. Alzheimer's disease: specificity of multimodal MRI in at-risk groups

Published in *BMC Neurology 2019;19(1):343* as:

Multimodal MRI of grey matter, white matter, and functional connectivity in cognitively healthy mutation carriers at risk for frontotemporal dementia and Alzheimer's disease

Rogier A. Feis, Mark J.R.J. Bouts, Elise G.P. Dopper, Nicola Filippini, Verena Heise, Aaron J. Trachtenberg, John C. van Swieten, Mark A. van Buchem, Jeroen van der Grond, Clare E. Mackay, Serge A.R.B. Rombouts

Abstract

Frontotemporal dementia (FTD) and Alzheimer's disease (AD) are associated with divergent differences in grey matter volume, white matter diffusion, and functional connectivity. However, it is unknown at what disease stage these differences emerge. Here, we investigate whether divergent differences in grey matter volume, white matter diffusion, and functional connectivity are already apparent between cognitively healthy carriers of pathogenic FTD mutations, and cognitively healthy carriers at increased AD risk.

We acquired multimodal magnetic resonance imaging (MRI) brain scans in cognitively healthy subjects with (n = 39) and without (n = 36) microtubule-associated protein tau (*MAPT*) or progranulin (*GRN*) mutations, and with (n = 37) and without (n = 38) apolipoprotein E ε4 (*APOE4*) allele. We evaluated grey matter volume using voxel-based morphometry, white matter diffusion using tract-based spatial statistics (TBSS), and region-to-network functional connectivity using dual regression in the default mode network and salience network. We tested for differences between the respective mutation carriers and controls, as well as for divergence of those differences. For the divergence contrast, we additionally performed region-of-interest TBSS analyses in known areas of white matter diffusion differences between FTD and AD (i.e., uncinate fasciculus, forceps minor, and anterior thalamic radiation).

MAPT/GRN mutation carriers did not differ from controls in any modality. *APOE4* carriers had lower fractional anisotropy than controls in the callosal splenium and right inferior fronto-occipital fasciculus, but did not show grey matter volume or functional connectivity differences. We found no divergent differences between both carrier-control contrasts in any modality, even in region-of-interest analyses.

Concluding, we could not find differences suggestive of divergent pathways of underlying FTD and AD pathology in asymptomatic risk mutation carriers. Future studies should focus on asymptomatic mutation carriers that are closer to symptom onset to capture the first specific signs that may differentiate between FTD and AD.

Keywords: microtubule-associated protein tau; progranulin; apolipoprotein E4; voxel-based morphometry; diffusion tensor imaging (DTI); tract-based spatial statistics (TBSS); functional connectivity; dual regression; frontotemporal dementia; Alzheimer's disease

Introduction

Frontotemporal dementia (FTD) and Alzheimer's disease (AD) are two of the most common causes of dementia (Lobo et al., 2000; Plassman et al., 2007; Seelaar et al., 2008; Vieira et al., 2013). In addition to distinct clinical features (Gorno-Tempini et al., 2011; McKhann et al., 2011; Rascovsky et al., 2011; Seelaar et al., 2011; Galimberti & Scarpini, 2012), FTD and AD demonstrate different patterns of functional and structural neurodegeneration on magnetic resonance imaging (MRI; Seeley et al., 2007; Zhang et al., 2009, 2011; Zhou et al., 2010; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016; Tuovinen et al., 2017). Atrophy is more pronounced in FTD than in AD in frontotemporal areas such as the anterior cingulate cortex, fronto-insula, and inferior frontal cortex (Seeley et al., 2007; Zhang et al., 2011; Möller et al., 2015b). Conversely, AD patients have more atrophy in the occipital gyrus and precuneus than FTD patients (Zhang et al., 2011). In terms of white matter diffusion tensor imaging (DTI) alterations, FTD patients have reduced fractional anisotropy (FA) and increased radial diffusivity (RD) compared to AD patients in the uncinate fasciculi, forceps minor, and anterior thalamic radiation, whereas AD patients do not show FA decreases or RD increases compared to FTD patients (Zhang et al., 2009, 2011; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016). Furthermore, functional connectivity is inversely affected in FTD and AD. In FTD patients, functional connectivity with the salience network is disrupted, while functional connectivity with the default mode network is increased. Vice versa, functional connectivity with the default mode network is disrupted in AD patients, while functional connectivity with the salience network is increased (Zhou et al., 2010; Tuovinen et al., 2017).

Despite these different patterns of neurodegeneration, the differentiation between FTD and AD is often demanding when patients first present in the memory clinic. For example, FTD patients may present with memory deficits (Graham et al., 2005; Le Ber et al., 2008), and as such may be misdiagnosed as AD patients (Johnson et al., 1999). Conversely, AD patients may be misdiagnosed as FTD patients due to the presentation of behavioural symptoms. Indeed, 13% of initial FTD diagnoses were corrected to AD after two years follow-up (Mendez et al., 2007), while 10–30% of clinical FTD patients were found to have AD pathology upon autopsy (Forman et al., 2006; Knibb et al., 2006; Alladi et al., 2007). The current criteria for behavioural variant FTD (bvFTD; Rascovsky et al., 2011), and language FTD variants (Gorno-Tempini et al., 2011) lack specificity to distinguish early-stage FTD patients from early-stage AD patients (McKhann et al., 2011). This diagnostic problem delays effective disease management (Mohs et al., 2001; Mendez et al., 2007; Mendez, 2009; Pressman & Miller, 2014), and frustrates the development of new treatments. Considering that the potential of disease modifying drugs is highest in the stage before atrophy occurs, the identification of early-stage dementia patients is crucial for patient selection in clinical trials (Rabinovici & Miller, 2010).

To assess whether FTD- and AD-related pathological changes are present even before symptom onset, carriers of FTD and AD risk mutations have been studied using structural, diffusion-weighted, and functional MRI (fMRI). For example, mutations in microtubule-associated protein tau (*MAPT*), progranulin (*GRN*), and repeat expansions in chromosome 9 open reading frame 72 (*C9orf72*) are known causes of genetic FTD. Presymptomatic carriers of these mutations have therefore been regularly studied to investigate early-stage FTD-related pathology (Whitwell et al., 2011a, b; Borroni et al., 2012; Dopper et al., 2014; Pievani et al., 2014). Similarly, mutations in presenilin 1, presenilin 2, and amyloid precursor protein are known causes of genetic AD. However, due to its higher prevalence, apolipoprotein E ε4 (*APOE4*), the strongest risk factor for sporadic AD, has been more extensively used to study early-stage AD-related pathology (Nierenberg et al.,

2005; Cherbuin et al., 2008; Agosta et al., 2009; Filippini et al., 2009, 2011; Honea et al., 2009; Heise et al., 2011; Machulda et al., 2011; Trachtenberg et al., 2012; Matura et al., 2014).

Contrary to findings in clinical FTD and AD (Zhang et al., 2009, 2011; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016), differences in diffusion metrics associated with asymptomatic *APOE4* (Persson et al., 2006; Gold et al., 2010; Smith et al., 2010; Heise et al., 2011; Adluru et al., 2014; Lyall et al., 2014; Laukka et al., 2015; Cavedo et al., 2017; Operto et al., 2018) are more widespread than diffusion differences associated with asymptomatic *MAPT/GRN* mutation carriers (Dopper et al., 2014; Pievani et al., 2014). Functional connectivity differences have also been shown in these asymptomatic groups (Machulda et al., 2011; Dopper et al., 2014). However, a comparison between these presymptomatic patterns of change in risk mutation carriers for FTD and AD is lacking, even though early-stage differences between these dementias may aid early differential diagnosis.

To this end, we investigated multimodal MRI in asymptomatic subjects at risk for FTD and AD. First, we aimed to replicate early carrier-control differences found between *MAPT*/*GRN* mutation carriers and controls, and between *APOE4* carriers and controls, respectively, by assessing whole-brain grey matter volume, white matter DTI measures, and functional connectivity in the default mode network and salience network. Secondly, we investigated whether *MAPT/GRN* carrier-control differences diverged from *APOE4* carrier-control differences, similar to FTD-AD differences. For the latter analysis, we additionally evaluated a priori selected white matter tracts known to be affected more strongly in FTD than AD (i.e., uncinate fasciculus, forceps minor, and anterior thalamic radiation). We hypothesised that the differences in grey matter volumes, DTI measures, and functional connectivity seen in FTD and AD patients (Seeley et al., 2007; Zhang et al., 2009, 2011; Zhou et al., 2010; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016; Tuovinen et al., 2017) may also be present to a smaller extent before symptom onset in risk mutation carriers.

Methods

Participants

Subjects were included retrospectively from studies carried out at the Leiden University Medical Centre (LUMC), The Netherlands, and at the Functional Magnetic Resonance Imaging of the Brain Centre (FMRIB), Oxford, UK.

The Dutch sample included 39 *MAPT/GRN* mutation carriers (11 *MAPT*, 28 *GRN*) and 36 controls, recruited from a pool of 160 healthy first-degree relatives of FTD patients with either *MAPT* or *GRN* mutation (Dopper et al., 2014). Participants were considered asymptomatic in the absence of (1) behavioural, cognitive, or neuropsychiatric change reported by the participant or knowledgeable informant, (2) cognitive disorders on neuropsychiatric tests, (3) motor neuron disease signs on neurologic examination, and (4) other FTD (Gorno-Tempini et al., 2011; Rascovsky et al., 2011) or amyotrophic lateral sclerosis (Ludolph et al., 2015) criteria. Asymptomatic non-carriers from these families and the general population were assumed to have equal risk of developing dementia. *MAPT/GRN* mutation carriers and controls were not tested for *APOE4* alleles.

Data from 37 *APOE4* carriers (30 apolipoprotein E ε3/ε4 heterozygotes, 7 apolipoprotein E ε4/ε4 homozygotes) and 38 controls (all apolipoprotein E ε3/ε3 homozygotes) were collected in Oxford from the general population in Oxfordshire and were selected to match the Dutch sample in terms of age and sex. Due to the limited sample size, it was not possible to match the groups' education level. Middle-aged and elderly *APOE4* carriers and controls underwent a pre-screening cognitive test (Addenbrooke's Cognitive Examination-revised version; Filippini et al., 2011; Heise et al., 2011) to assure asymptomatic status. *APOE4* carriers and controls were not tested for *MAPT/GRN* mutations.

In both cohorts, participants were between 21 and 70 years old. A priori exclusion criteria included MRI contraindications, head injury, current or past neurologic or psychiatric disorders, (history of) substance abuse including alcohol, corticosteroid therapy, type I diabetes therapy, and memory complaints.

The study was conducted in accordance with regional regulations and the Declaration of Helsinki. Written informed consent was received from all participants, and ethical approval for data acquisition was provided by the Medical Ethical Committees in Rotterdam and Leiden for *MAPT/GRN* data, and the National Research Ethics Service Committee South Central—Oxford C for *APOE4* data. For further details regarding the recruitment protocols, see Dopper et al. (2014) for the Dutch sample and Filippini et al. (2011) for the English sample.

Image acquisition

MRI data were acquired with a Philips 3 T Achieva MRI scanner using an 8-channel SENSE head coil (*MAPT/GRN* mutation carriers and controls) or on a Siemens 3 T Trio scanner with a 12-channel head coil (*APOE4* carriers and controls). T_1-weighted data were acquired with TR = 9.8 ms, TE = 4.6 ms, flip angle = 8°, 140 axial slices, and voxel size = 0.88 × 0.88 × 1.20 mm for *MAPT/GRN* mutation carriers and controls, and using a magnetisation-prepared rapid gradient echo sequence (MPRAGE; TR = 2,040 ms, TE = 4.7 ms, flip angle = 8°, 192 axial slices, voxel size = 1 × 1 × 1 mm) in *APOE4* carriers and controls. Diffusion-weighted images were acquired in 62 directions with TR = 8,250–9,300 ms, TE = 80–94 ms, b-value = 1,000 s/mm^2, flip angle = 90°, 65–70 axial slices, and voxel size = 2 × 2 × 2 mm. For the resting-state functional MRI (rs-fMRI) scan, subjects were instructed to remain awake and keep their eyes closed (*MAPT/GRN* mutation

carriers and controls) or open (*APOE4* carriers and controls), and to think of nothing in particular. We acquired 180–200 volumes with TR = 2,000–2,200 ms, TE = 28–30 ms, flip angle = 80–89°, and voxel size = 2.75 × 2.75 × 2.75 mm + 10% interslice gap or 3 × 3 × 3.5 mm.

Image analysis

FMRIB Software Library (FSL, http://www.fmrib.ox.ac.uk/fsl) tools were used for all data analyses (Jenkinson et al., 2012).

Grey matter volume analyses

Whole-brain voxel-wise structural analysis was carried out with FSL-VBM (Douaud et al., 2007), an optimised voxel-based morphometry protocol (Good et al., 2001) using FSL tools (Smith et al., 2004). First, we performed brain extraction and grey matter segmentation, and registered images to the MNI152 standard space using linear (FLIRT) and nonlinear registration (FNIRT; Anderson et al., 2007). The resulting images were averaged and flipped along the x-axis to create a study-specific grey matter template. Native grey matter images were then re-registered to this template, modulated using the field-warp Jacobian, and smoothed using an isotropic Gaussian kernel with a sigma of 2.5 mm (~ 6 mm full width at half maximum).

Diffusion tensor imaging

Diffusion-weighted imaging scans were processed using FMRIB's Diffusion Toolbox (FDT, http://www.fmrib.ox.ac.uk/fsl/fdt). First, we aligned raw diffusion weighted images to the b0-volume using 'eddy correct' to correct for movement and eddy currents. Next, we fitted the diffusion tensor model to the images at each voxel to create modality-specific images for fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AxD), and radial diffusivity (RD). For voxel-wise analysis of these images, we used tract-based spatial statistics (TBSS; Smith et al., 2006). After brain extraction, subjects' individual FA images were transformed to standard space using FNIRT. A mean FA image was then created and thinned to generate a whole-brain mean FA skeleton, representing the centres of all white matter tracts common to all subjects. Individual aligned FA images were projected onto this skeleton for group analysis. Similar analyses were performed on MD, AxD, and RD maps using the spatial transformation parameters that were estimated in the FA analysis. For our region-of-interest analyses, we masked the whole-brain skeleton with the combined masks of the uncinate fasciculi, forceps minor, and the bilateral anterior thalamic radiations, which have been shown to differ between FTD and AD patients in terms of DTI metrics (Zhang et al., 2009, 2011; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016).

Resting-state functional MRI

Prestatistical processing of resting-state data consisted of motion correction (Jenkinson et al., 2002), brain extraction, spatial smoothing using a Gaussian kernel of 6 mm full width at half maximum, 4D grand-mean scaling and high-pass temporal filtering corresponding to a period of 150s (~ 0.007 Hz). Registration to MNI152 standard space was carried out in two steps. We registered echo-planar images onto their respective T_1-weighted structural images using FLIRT and boundary-based registration (Jenkinson & Smith, 2001; Jenkinson et al., 2002; Greve & Fischl, 2009). Next, we used FNIRT to align T_1-weighted structural images to MNI152 standard space, and concatenated the resulting registration matrices to register echo-planar images directly to standard space. Next, we performed individual independent component analysis (ICA) and voxel-wise intensity normalisation (i.e., by dividing all voxels by their time series' mean values and

multiplying by 10,000).

We used FMRIB's ICA-based X-noiseifier (FIX; Griffanti et al., 2014; Salimi-Khorshidi et al., 2014; Feis et al., 2015) to clean up noise components and reduce rs-fMRI scan site bias. For a detailed description and validation of FIX as a multicentre bias reduction method, see Feis et al. (2015). In short, we classified the individual ICA components of a subset of the subjects as signal, noise, or unknown, trained the FIX classifier, and used a leave-one-out test to control the algorithm's quality. All subjects' data were then classified using the optimal threshold (i.e., 20—true-positive rate 95.1%, true-negative rate 91.4%), and structured noise components were removed.

After processing and application of FIX, rs-fMRI data were temporally concatenated and decomposed into 25 components using FSL's group-level ICA tool (Hyvärinen, 1999; Beckmann & Smith, 2004; Beckmann et al., 2005) in order to identify large-scale patterns of functional connectivity. The resulting group-level ICA spatial maps were compared to previously described resting-state networks (Beckmann et al., 2005; Damoiseaux et al., 2006; Rytty et al., 2013; Tian et al., 2013; Bey et al., 2015), and we selected default mode network components and salience network components for dual regression analyses. The default mode network is disrupted in AD and enhanced in FTD, while the salience network is disrupted in FTD and enhanced in AD (Zhou et al., 2010; Tuovinen et al., 2017). Components that included the precuneus, posterior cingulate cortex, angular gyrus, medial prefrontal cortex, and hippocampus were regarded as parts of the default mode network. Components featuring the anterior cingulate cortex, supplementary motor area, and insula were considered linked to the salience network. We found three networks resembling the default mode network (e.g., the anterior, inferior, and posterior default mode network, **Figure 6.1A–C**) and two networks resembling the salience network (e.g., the anterior and posterior salience network, **Figure 6.1D–E**). For these five resulting resting-state networks of interest, we performed dual regression to identify subject-specific spatial maps corresponding to the resting-state networks of interest (Beckmann et al., 2009; Filippini et al., 2009). First, the spatial maps derived from group-level ICA were used as a spatial regressor in each subjects' rs-fMRI data to obtain subject-specific time series describing the temporal dynamics for each component (**Supplemental Figure S6.1**, step 1). Next, the time series found by spatial regression were used as a temporal regressor to find voxels associated with those time series for each subject (**Supplemental Figure S6.1**, step 2). As such, we used the group-level ICA networks of interest to obtain subject-specific spatial maps that allow for voxel-wise comparison. Statistical analysis of region-to-network functional connectivity group differences was then carried out by testing for the functional connectivity between the five resting-state networks of interest and all other grey matter voxels.

Statistical analysis

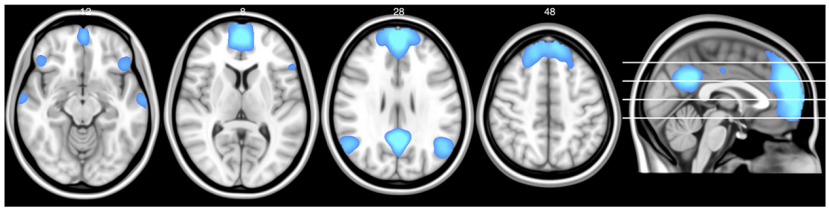
Statistical analysis of grey matter volume, DTI features, and rs-fMRI data was performed using general linear models, including age and education as confound regressors. Additionally, we added a voxel-wise covariate for grey matter volume to the functional connectivity analyses. We tested for differences between *MAPT/GRN* mutation carriers and controls, and for differences between *APOE4* carriers and controls, respectively. Additionally, we tested for the differences between these respective carrier-control contrasts to evaluate whether these gene mutations have divergent effects on the brain in cognitively healthy carriers that might reflect early substrates of FTD or AD pathology. Since possible centre effects are equivalent for carriers and controls at each site, these effects cancel out when we compared the carrier-control effect at one site to the carrier-control effect at the other site. Consequently, unknown confounding factors such as scanner and population differences should have minimal influence on our results.

Pooling *MAPT* and *GRN* mutation carriers, and *APOE4* heterozygotes and homozygotes in our carrier samples may have increased heterogeneity in our groups. To account for this possibility, we performed additional analyses with covariates encoding the difference between *MAPT* and *GRN* mutations, and between *APOE4* hetero- and homozygosity.

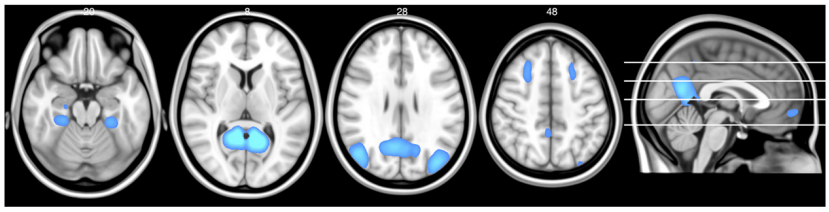
Voxel-wise application of these general linear models to the data was performed using FSL randomise, a permutation-based non-parametric test (5,000 permutations). We set the family-wise error rate at 5% across space by using threshold-free cluster enhancement (Winkler et al., 2014) in all analyses. The alpha level required for statistical significance was set at 0.025 for all imaging analyses, which corresponds to an alpha level of 0.05 in a two-sided *t*-test, since randomise performs the permutation equivalent of a one-sided *t*-test. Minimal cluster size for significant results was set at 10 voxels.

SPSS version 24 (SPSS, Chicago, IL) was used for statistics performed on non-imaging (demographic) variables. Analysis of variance (ANOVA) tests were performed on normally distributed continuous variables (age and education) and included Bonferroni post-hoc tests. A chi-square test was performed for sex. The alpha level required for statistical significance was set at 0.05.

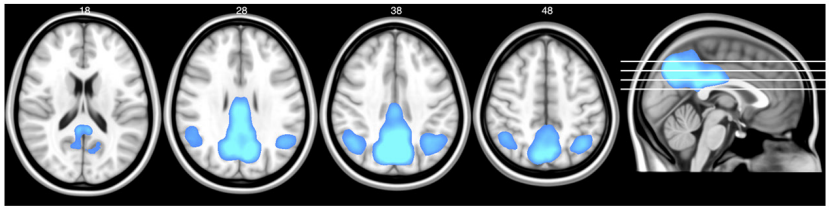
A. Anterior default mode network



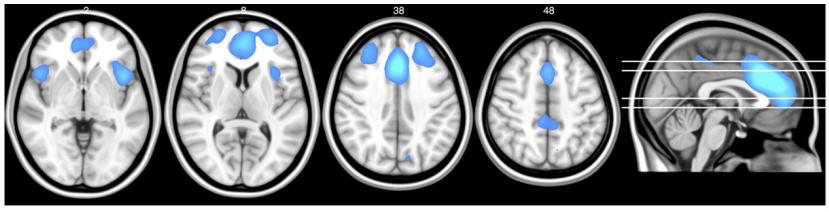
B. Inferior default mode network



C. Posterior default mode network



D. Anterior salience network



E. Posterior salience network

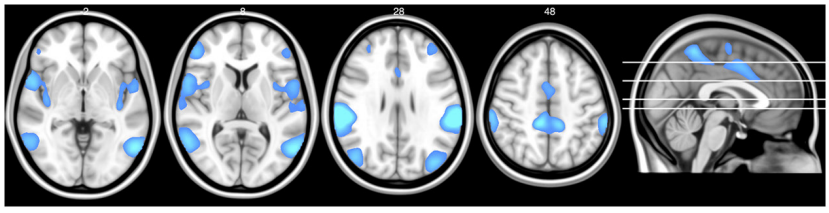


Figure 6.1 Resting-state networks
Maps illustrate the most informative slices of resting-state networks of interest that featured known default mode network and salience network regions and that were used for statistical testing after dual regression.

Results

Demographics

Demographic data for all groups are shown in **Table 6.1**. Age and sex did not differ between groups. Bonferroni post-hoc tests revealed significantly lower education level in years for *MAPT/GRN* mutation carriers than *APOE4* controls ($p = 0.001$), for *MAPT/GRN* controls than *APOE4* controls ($p < 0.001$), and for *MAPT/GRN* controls than *APOE4* carriers ($p = 0.001$).

Table 6.1 Participant demographics

	MAPT/GRN		***APOE4***		
	Carriers ***n* = 39**[a]	**Controls** ***n* = 36**	**Carriers** ***n* = 37**[b]	**Controls** ***n* = 38**	***p*-value**
Age, mean (SD) years	50.5 (10.0)	49.8 (11.3)	48.6 (10.3)	50.05 (10.5)	0.86
Sex, *n* (%) ♀	23 (59%)	18 (50%)	20 (54%)	20 (53%)	0.89
Education, mean (SD) years[c]	14.0 (2.5)	12.6 (2.9)	15.5 (3.7)	16.8 (3.2)	<0.001

APOE4, apolipoprotein E ε4; *GRN*, progranulin; *MAPT*, microtubule-associated protein tau.
[a] 11 *MAPT*, 28 *GRN*.
[b] 30 heterozygotes, 7 homozygotes.
[c] Scores of education level in years were missing for three *MAPT/GRN* mutation carriers and two *MAPT/GRN* controls.

Grey matter volume

We found no grey matter volume differences in *MAPT/GRN* mutation carriers compared to controls, in *APOE4* carriers and compared to controls, nor were there differences between both contrasts.

White matter diffusion

Tract-based spatial statistics revealed no FA, MD, AxD, or RD differences between *MAPT/GRN* mutation carriers and controls. However, we found four clusters of FA reductions in *APOE4* carriers compared to controls (**Table 6.2, Figure 6.2**). Three clusters were located in the forceps major, more specifically in right side of the callosal splenium, and one cluster was located in the right inferior fronto-occipital fasciculus. We found no significant differences between the *MAPT/GRN* and *APOE4* carrier-control contrasts in our whole-brain analysis, nor in our region-of-interest analyses.

Functional connectivity

We found no differences in region-to-network functional connectivity in *MAPT/GRN* mutation carriers compared to controls, in *APOE4* carriers compared to controls, nor between the two carrier-control contrasts in any of the five resting-state networks.

Table 6.2 Cluster information

Cluster	Size	Max *t*-statistic	MNI coordinates			L/R	Area (peak voxel)
			x	y	z		
1	64	4.14	54	101	72	R	IFOF
2	44	3.19	71	79	95	R	Splenium
3	32	3.58	63	73	87	R	Splenium
4	22	4.19	74	87	98	R	Splenium

Cluster information for significant clusters of reduced FA in *APOE4* carriers compared to controls. Minimum cluster size was 10.
APOE4, apolipoprotein E ε4; FA, fractional anisotropy; IFOF, inferior fronto-occipital fasciculus.

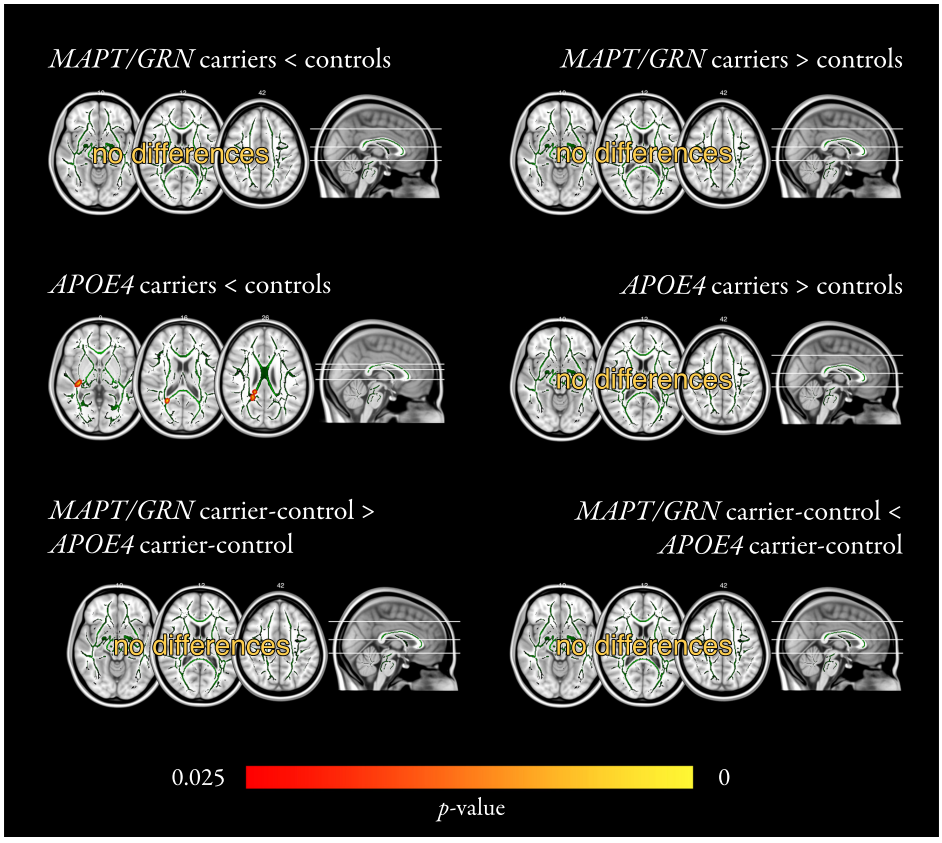


Figure 6.2 White matter FA analysis
Differences in FA (or lack thereof) are shown for each contrast (e.g., *MAPT/GRN* mutation carriers greater or smaller than controls; *APOE4* carriers greater or smaller than controls; *MAPT/GRN* carrier-control differences greater or smaller than *APOE4* carrier-control differences). Mean skeleton maps are shown in green; skeletonised significant results were thickened for better visualisation. Four clusters of FA reductions were found in *APOE4* carriers compared to controls (middle left panel). Colour bar represents significance.
APOE4, apolipoprotein E ε4; FA, fractional anisotropy; *MAPT/GRN*, microtubule-associated protein tau / progranulin.

Heterogeneity analyses

Analyses including covariates for the difference between *MAPT* and *GRN* mutations, and between *APOE4* hetero- and homozygosity yielded results similar to our main analyses. There were no grey matter volume differences between *MAPT/GRN* mutation carriers and controls, *APOE4* carriers and controls, nor between the two carrier-control contrasts. *APOE4* carriers had reduced FA in compared to controls (**Supplemental Figure S6.2**), though only one of the four clusters remained significant. We found no DTI differences between *MAPT/GRN* mutation carriers and controls, nor between the two carrier-control contrasts. We found no differences in region-to-network functional connectivity in *MAPT/GRN* mutation carriers compared to controls, *APOE4* carriers compared to controls, nor between the carrier-control contrasts in any of the five resting-state networks.

Data availability

All non-thresholded statistical images for grey matter volume, white matter diffusion, and functional connectivity results of our default analysis can be found on NeuroVault (Gorgolewski et al., 2015): https://neurovault.org/collections/NXLXKVCZ/.

Discussion

Differences in atrophy, white matter diffusion, and functional connectivity patterns have been repeatedly shown between FTD and AD patients (Zhang et al., 2009, 2011; Zhou et al., 2010; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016), and between asymptomatic mutation carriers at risk for these diseases and controls, e.g., *MAPT* and *GRN* mutation carriers (Whitwell et al., 2011a, b; Borroni et al., 2012; Dopper et al., 2014; Pievani et al., 2014); *APOE4* carriers (Nierenberg et al., 2005; Cherbuin et al., 2008; Agosta et al., 2009; Filippini et al., 2009, 2011; Honea et al., 2009; Heise et al., 2011; Machulda et al., 2011; Trachtenberg et al., 2012; Matura et al., 2014). However, comparisons between groups at risk for FTD and groups at risk for AD have been lacking, even though early-stage differences between these dementias are key to improve on diagnostic standards. In this study, we aimed to replicate previously found differences in asymptomatic mutation carriers at risk for FTD and AD compared to their respective control groups. More importantly, we investigated whether carrier-control differences diverged, similar to the divergences that exist between FTD and AD. While we could replicate some of the previously reported fractional anisotropy reductions in asymptomatic *APOE4* carriers, we found no evidence of divergence between *MAPT/GRN* carrier-control differences and *APOE4* carrier-control differences, even when restricting our DTI analysis to regions which are known to differ between FTD and AD patients. This may suggest that the neuroimaging biomarkers measured in this study are not sufficiently specific to differentiate between FTD-related pathology and pathology possibly related to AD at this early stage.

Our lack of differences between groups in grey matter volume were unsurprising. In asymptomatic risk mutation carriers, one would not expect dementia-related atrophy unless the mutation carrier would be close to symptom onset. Indeed, grey matter volume differences have not been reported in asymptomatic *MAPT/GRN* mutation carriers (Borroni et al., 2012; Dopper et al., 2014), though reports in asymptomatic *APOE4* carriers have been conflicting. While some groups report no grey matter volume differences in asymptomatic *APOE4* carriers (Cherbuin et al., 2008; Filippini et al., 2011; Heise et al., 2011; Matura et al., 2014), others found reduced grey matter volume in the hippocampus (Honea et al., 2009; Cacciaglia et al., 2018), lingual gyrus (Honea et al., 2009), precuneus (Honea et al., 2009; ten Kate et al., 2016), insula (ten Kate et al., 2016), caudate nucleus, precentral gyrus, and cerebellar crus (Cacciaglia et al., 2018). These conflicting findings may in part result from methodological differences, sample sizes, and the different age ranges between studies. Since disease modifying treatments aim to prevent atrophy, one would ideally aim to diagnose dementia patients before atrophy occurs to maximise potential treatment effect. Accordingly, biomarker research should focus on detecting substrates of neurodegeneration that precede atrophy and that may be reversible by future disease modifying treatments.

White matter diffusion analyses yielded areas of reduced FA in *APOE4* carriers compared to controls in the splenium of the corpus callosum, and in the right inferior fronto-occipital fasciculus. These results concur with previous reports in *APOE4* carriers. FA reductions were most often reported in the corpus callosum, cingulum, and inferior fronto-occipital fasciculi (Persson et al., 2006; Gold et al., 2010; Smith et al., 2010; Heise et al., 2011; Adluru et al., 2014; Lyall et al., 2014; Laukka et al., 2015; Cavedo et al., 2017; Operto et al., 2018), while FA differences in the corticospinal tract (Heise et al., 2011; Laukka et al., 2015; Cavedo et al., 2017; Operto et al., 2018) and superior longitudinal fasciculi (Heise et al., 2011; Cavedo et al., 2017; Operto et al., 2018) were less frequently reported. We found no diffusion differences in *MAPT/GRN* mutation carriers compared to controls, in contrast to earlier work (Dopper et al., 2014; Pievani et al., 2014).

However, this might be explained by differences in methodology. One study found significant FA reductions only within certain pre-specified tracts, and, similar to our current study, found no whole-brain differences (Dopper et al., 2014). The other study found differences at $p < 0.005$ uncorrected for multiple comparisons across space. Our analyses were performed with a more restrictive significance level, as we corrected for multiple comparisons across space using threshold-free cluster enhancement, and used the statistical threshold appropriate for a two-sided test, which is not a standard procedure in neuroimaging (Chen et al., 2019). Interestingly, DTI alterations are larger in FTD patients than in AD patients (Zhang et al., 2009, 2011; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016), while preclinical alterations in *APOE4* carriers (Persson et al., 2006; Gold et al., 2010; Smith et al., 2010; Heise et al., 2011; Adluru et al., 2014; Lyall et al., 2014; Laukka et al., 2015; Cavedo et al., 2017; Operto et al., 2018) are more widespread than in *MAPT/GRN* mutation carriers (Dopper et al., 2014; Pievani et al., 2014). Recently, it has been postulated that white matter DTI differences in genetic FTD develop rather explosively in the years just prior to symptom onset (Feis et al., 2019b; Jiskoot et al., 2019). This might explain why in our sample, we found DTI differences in *APOE4* carriers, but no DTI differences in *MAPT/GRN* mutation carriers. Although there were FA reductions in *APOE4* carriers compared to controls, the difference was not strong enough to result in a difference between the *MAPT/GRN* carrier-control contrast and the *APOE4* carrier-control contrast. We also performed region-of-interest analyses in the uncinate fasciculi, forceps minor, and bilateral anterior thalamic radiations, which were found to have FA reductions and RD increases in FTD patients compared to AD patients (Zhang et al., 2009, 2011; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016). However, even in these regions of interest, we could not find DTI differences between the *MAPT/GRN* carrier-control contrast and the *APOE4* carrier-control contrast. As such, we could not conclude that *MAPT/GRN* mutation carriership had a different effect on white matter diffusion metrics than *APOE4* carriership.

It has been previously argued that the default mode network and the salience network are inversely correlated and both play a role in AD and FTD. Specifically, functional connectivity in the default mode network was reported to be reduced in AD patients and increased in FTD patients, whereas functional connectivity in the salience network was reported to be inversely affected: reduced in FTD patients and increased in AD patients (Zhou et al., 2010; Tuovinen et al., 2017). In asymptomatic *APOE4* carriers, this inverse correlation was also shown. Functional connectivity with the default mode network was decreased and functional connectivity with the salience network was enhanced in *APOE4* carriers compared to controls (Machulda et al., 2011). In asymptomatic *MAPT* and *GRN* mutation carriers, functional connectivity was reduced in the salience network, but no differences in the default mode network were found (Dopper et al., 2014). Based on these results, we hypothesised that functional connectivity in the default mode network and salience network would be ideal candidates to screen for early changes in asymptomatic risk mutation carriers. However, we found no evidence of functional connectivity differences, either between the respective carrier and control groups or divergent differences between the carrier-control contrasts. This might in part be a power issue but could also be explained by population and methodological differences. For example, our sample was on average younger and had a broader age range than the *APOE4* sample investigated by Machulda et al. (2011). Furthermore, we performed data-driven dual regression analyses, whereas both Machulda et al. (2011) and Dopper et al. (2014) performed seed-based analyses. While small seed areas are arbitrarily placed and may be subject to registration mismatch, dual regression networks are less sensitive to these issues due to their data-driven origin. Indeed, dual regression is amongst the best functional MRI analysis techniques in

terms of test-retest reliability (Zuo et al., 2010; Zuo & Xing, 2014). Therefore, the most likely explanation of our functional connectivity results is that our groups were on average too far from symptom onset for functional connectivity alterations in the default mode network and salience network to robustly appear.

Strengths of this study include its unique design to pick up differences between FTD- and AD-related pathology in asymptomatic populations, and the inclusion of control groups from both sites to deal with potential scan site bias. We performed specific region-of-interest analyses to increase power to find differences in DTI metrics. Furthermore, we used FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) to clean up structured noise (e.g., motion, artefacts) from rs-fMRI data to reduce scanner-based functional connectivity differences (Feis et al., 2015) and increase the signal-to-noise ratio. To account for possible heterogeneity resulting from pooling *MAPT* and *GRN* mutation carriers, and *APOE4* hetero- and homozygotes, we performed additional analyses including covariates for the different mutation types. The results of these analyses were very similar to our main results, suggesting that the effect of genetic heterogeneity in our main analyses was altogether limited. Limitations must also be considered. Firstly, differences in penetrance and age of onset exist between *MAPT/GRN* and *APOE4*. *MAPT* and *GRN* mutations have an autosomal dominant inheritance pattern, and are highly penetrant (van Swieten et al., 2000; van Swieten & Heutink, 2008). On the other hand, *APOE4* has a dose-dependent effect on lifetime AD risk. Heterozygous *APOE4* carriers have an estimated lifetime risk for AD of approximately 25%, while *APOE4* homozygosity is associated with an estimated lifetime risk of around 55% (Genin et al., 2011). Therefore, it is unlikely that all *APOE4* carriers from our sample will develop AD, which reduced our power to detect AD-related differences. For the same reason, it cannot be entirely ruled out that some of the differences associated with the *APOE4* carriers do not reflect presymptomatic AD-related pathology. Information on *MAPT/GRN* mutation carriership was not available for *APOE4* carriers and controls, and information on *APOE4* carriership was not available for *MAPT/GRN* mutation carriers and controls. Due to the infrequency of *MAPT* and *GRN* mutations, it is unlikely that *APOE4* carriers or controls had an *MAPT* or *GRN* mutation. However, the frequency of the *APOE4* allele in Caucasian populations is around 14% (Eisenberg et al., 2010), and it is likely that some of the *MAPT* and *GRN* mutation carriers and controls had an *APOE4* allele. As *MAPT/GRN* mutation carriers and controls were from the same families, the frequency of the *APOE4* alleles within these groups was most likely similar. Therefore, the effect of *APOE4* on our *MAPT/GRN* analyses is presumably small. The broad age range in our groups presents another limitation. FTD- or AD-related pathology may be absent or present in a lesser degree in young mutation carriers than in older mutation carriers, who are closer to symptom onset. However, even though a broad age range was present in our sample, physiological brain ageing effects are unlikely to have influenced our results. The four groups were matched for age, and age was added as confound covariate to the model. Therefore, physiological brain ageing effects should be equally distributed across groups and were accounted for in the model. In order to increase power, future neuroimaging research comparing FTD- and AD-related pathology in asymptomatic risk groups should contain clinical follow-up and conversion information, which will enable the inclusion of a time to onset variable to the model.

Conclusion

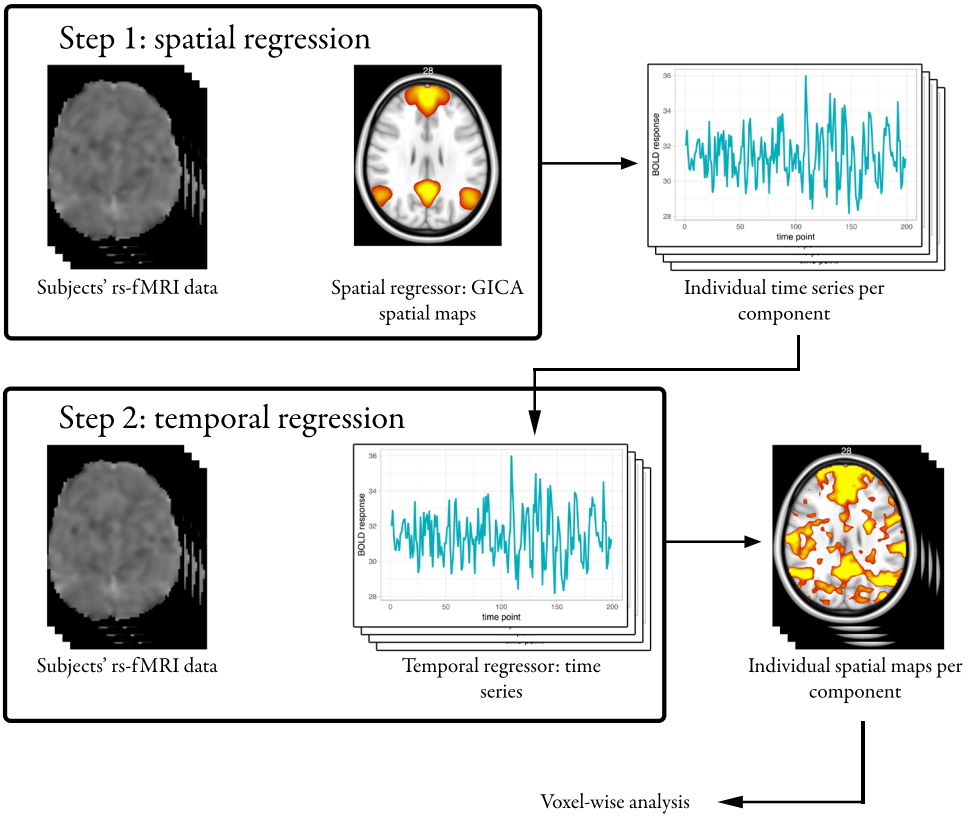
Dementias are relentlessly progressive diseases for which no adequate treatments currently exist, and differentiation between various forms of dementia is clinically challenging. Recently, MRI has shown different patterns of grey matter atrophy, DTI alterations, and functional connectivity differences in

AD and FTD patients (Zhang et al., 2009, 2011; Zhou et al., 2010; Mahoney et al., 2014; Möller et al., 2015b; Daianu et al., 2016; Tuovinen et al., 2017). However, early differential identification of at-risk groups is key to study pathophysiological processes, develop disease modulating drugs and, eventually, identify patient groups that may benefit from these treatments. In the current study, we could not find differences suggestive of divergent pathways of underlying FTD and AD pathology in asymptomatic risk mutation carriers.

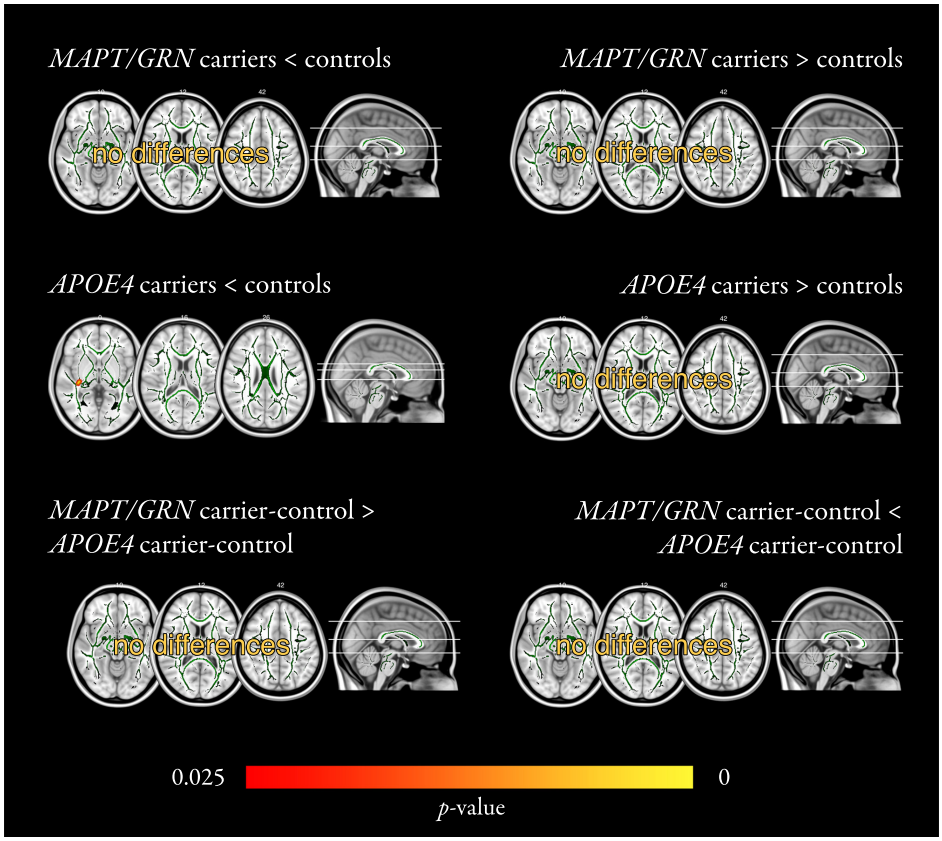
Acknowledgements

Our appreciation goes out to all study participants. The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to RAF); HDH Wills 1965 charitable trusts (to NF, English Charity Register 1117747); an Alzheimer's Research UK studentship (to VH, English Charity Register 1077089); a Rhodes scholarship (to AJT); ZonMw programme Memorabel project 733050103, JPND PreFrontAls consortium project 733051042 (to JCvS), the National Institute for Health Research (NIHR), UK as part of the Oxford Biomedical Research Centre (BRC; to CEM) and a VICI grant 016–130–667 from The Netherlands Organisation for Scientific Research (NWO; to SARBR). The views expressed are those of the authors and not necessarily those of the NWO, the NHS or the NIHR. The funding sources were not involved in the design of the study; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The authors report no conflict of interest.

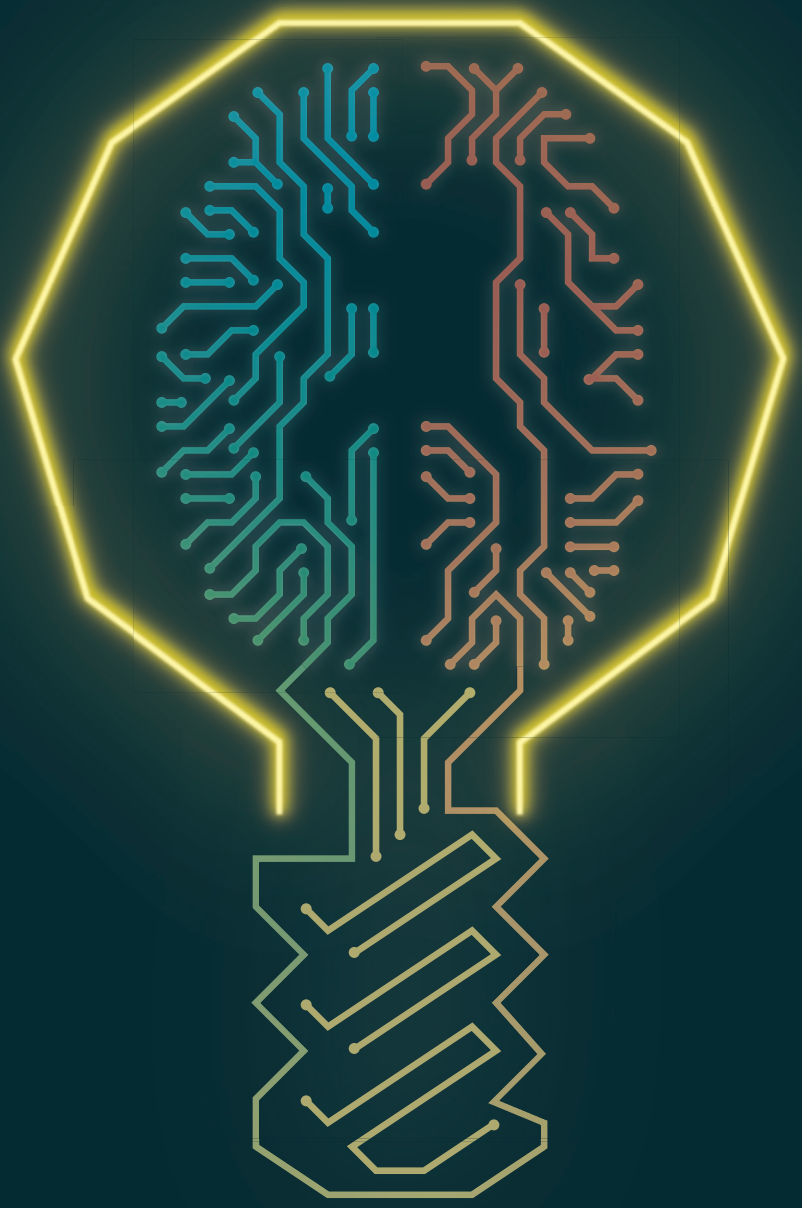
Supplemental material



Supplemental Figure S6.1 Dual regression
Subject-specific spatial maps for statistical testing are acquired from GICA spatial maps in two steps. First, GICA spatial maps are used as spatial regressor on each subject's rs-fMRI data to obtain time series associated with those GICA components (step 1). Next, these time series are used as temporal regressor to obtain subject-specific spatial maps for each component (step 2). These maps are then used for voxel-wise statistical testing.
GICA, group-level independent component analysis; rs-fMRI, resting-state functional magnetic resonance imaging.



Supplemental Figure S6.2 White matter FA analysis with mutation covariates
In this analysis, covariates for the differences between *MAPT* and *GRN* mutations, and between *APOE4* hetero- and homozygosity were added to account for genetic heterogeneity. Differences in FA (or lack thereof) are shown for each contrast (e.g., *MAPT/GRN* mutation carriers greater or smaller than controls; *APOE4* carriers greater or smaller than controls; *MAPT/GRN* carrier-control differences greater or smaller than *APOE4* carrier-control differences). Mean skeleton maps are shown in green; skeletonised significant results were thickened for better visualisation. One cluster of FA reductions was found in *APOE4* carriers compared to controls (middle left panel). Colour bar represents significance.
APOE4, apolipoprotein E ε4; FA, fractional anisotropy; *MAPT/GRN*, microtubule-associated protein tau / progranulin.



Chapter 7

General discussion

Main findings

In this dissertation, we aimed to advance the development of potential magnetic resonance imaging (MRI-)based biomarkers for the early detection of frontotemporal dementia (FTD). We focused on machine learning with multimodal MRI features to study potential early FTD biomarkers in the single-subject setting that corresponds with the clinical diagnostic process. Furthermore, we aimed to facilitate the combination of existing resting-state functional MRI (rs-fMRI) data from different centres through structured noise reduction. Lastly, we tested whether presymptomatic MRI changes were different between cognitively healthy subjects at genetic risk for respectively FTD and Alzheimer's disease (AD). In this chapter, the main findings of the studies in this dissertation are summarised and discussed. Critical considerations and future directions for neuroimaging research in FTD are also examined.

Clinical translation of MRI-based biomarkers for early FTD diagnosis

The translation of promising biomarkers to clinical practice requires validation in a single-subject setting. Accordingly, classification studies with FTD patients and controls are becoming increasingly important in the FTD neuroimaging field. Classifiers based on grey matter volume (Raamana et al., 2014; Klöppel et al., 2015; Koikkalainen et al., 2016; Bron et al., 2017; Canu et al., 2017; Meyer et al., 2017; Bouts et al., 2018), white matter diffusion (Bron et al., 2017; Canu et al., 2017; Bouts et al., 2018), and functional connectivity (Canu et al., 2017; Bouts et al., 2018) distinguish FTD patients from controls with high accuracy. However, previous classification studies relied on the current diagnostic criteria for patient inclusion, and may therefore not generalise to early FTD disease stages. On the other hand, multimodal MRI differences between groups of presymptomatic FTD mutation carriers and controls have been investigated (Borroni et al., 2008, 2012; Whitwell et al., 2011a; Rohrer et al., 2013, 2015; Dopper et al., 2014; Premi et al., 2014; Lee et al., 2017; Papma et al., 2017; Bertrand et al., 2018; Cash et al., 2018; Jiskoot et al., 2018a; Panman et al., 2019), though single-subject classification and prediction studies were lacking in the preclinical stage.

In **chapter 2**, we studied classification in presymptomatic FTD mutation carriers (microtubule-associated protein tau [*MAPT*] and progranulin [*GRN*] mutation carriers, and chromosome 9 open reading frame 72 [*C9orf72*] repeat expansion carriers) and familial non-carriers using MRI-based classification models. We found that several carrier-control models, which were based on white matter features, separated mutation carriers from controls beyond chance level. Importantly, this shows that single-subject classification is possible in the presymptomatic stage of genetic FTD. However, a classification model trained on symptomatic bvFTD patients and controls (Bouts et al., 2018) did not outperform chance level when applied to the presymptomatic sample. This model included grey matter density (GMD), white matter diffusion, and functional connectivity features, highlighting that features that are important for the classification of symptomatic FTD patients are not necessarily useful in earlier stages.

In **chapter 3**, we investigated how MRI-based classification scores develop over time as FTD mutation carriers approach symptom onset. We calculated MRI-based classification scores at each time point, and used linear mixed effects models to compare presymptomatic mutation carriers with non-carriers. We found that the progression of classification scores over time was similar for presymptomatic FTD mutation carriers and non-carriers. Within the mutation carrier group, we compared mutation carriers that developed symptoms during the study's follow up (termed 'converters') with mutation carriers that remained presymptomatic ('non-converters'). We found

that classification scores in the converter group increased significantly more than in the non-converter group. The results in **chapter 3** show that MRI-based classification is sensitive to FTD-related changes over time. Also, our results suggest that presymptomatic FTD mutation carriers remain relatively stable over time, until they are within a few years of symptom onset, when brain changes begin to rapidly develop.

Since the converters had a large increase in classification scores over time compared to non-converters, we hypothesised that it might be possible to use MRI features to predict future symptom onset in FTD mutation carriers. In **chapter 4**, we included MRI data of 42 presymptomatic FTD mutation carriers, and trained classifiers to separate those that developed symptoms within four years post-MRI from those that remained presymptomatic after four years. Although only seven subjects converted within the four years follow-up, it was still possible to separate these converters from non-converters beyond chance level using the fractional anisotropy (FA) feature. This proof of concept implies that accurate symptom onset prediction may be possible in genetic FTD once larger cohorts become available.

Our results in **chapter 2–4** are consistent with the hypothesis that white matter changes predate and exceed grey matter atrophy in FTD. In presymptomatic FTD mutation carriers, differences in white matter diffusion tensor imaging (DTI) metrics are found in absence of grey matter atrophy (Borroni et al., 2008; Dopper et al., 2014), or exceed grey matter atrophy in magnitude (Pievani et al., 2014; Cash et al., 2018; Jiskoot et al., 2018a; Panman et al., 2019). In the years around symptom onset, grey matter atrophy starts increasing (Jiskoot et al., 2019), though white matter DTI changes remain more extensive than grey matter atrophy in clinical FTD (Agosta et al., 2012; Zhang et al., 2013; Mahoney et al., 2014). Although we did not statistically compare the different classification models to each other, it is striking that white matter features were the only ones to outperform chance level in our presymptomatic studies (**chapter 2, 4**). The best unimodal classification model to separate FTD mutation carriers from controls was based on the radial diffusivity (RD) feature (**chapter 2**), while multimodal models that outperformed chance included RD, white matter density (WMD), mean diffusivity (MD), and/or axial diffusivity (AxD). FA was the only feature that could predict which FTD mutation carriers would develop symptoms in four years' time (**chapter 4**). A model trained on bvFTD patients and controls that also included the GMD feature could not discriminate between presymptomatic mutation carriers and controls, neither cross-sectionally (**chapter 2**), nor longitudinally (**chapter 3**). However, it did show a different classification score progression between mutation carriers that developed symptoms during follow-up and mutation carriers that remained presymptomatic (**chapter 3**). This result might be explained by the notion that grey matter atrophy only becomes detectable around the time of symptom onset. Together, our results and previous studies indicate that white matter DTI changes are likely the most promising MRI biomarker for early diagnosis in genetic FTD.

Resting-state functional MRI data harmonisation across centres

In **chapter 5**, we applied 'FIX' (FMRIB's ICA-based X-noiseifier), a novel clean-up tool for structured noise, on two MRI data sets of healthy controls from different centres to investigate whether it can improve rs-fMRI data harmonisation between centres. We showed that FIX removes structured noise, while retaining physiologically driven functional connectivity differences. FIX reduced nearly all functional connectivity differences to non-significant levels, with the exception of functional connectivity differences in the visual network. These visual network differences were likely physiological differences secondary to the experimental design, as one sample was scanned with eyes open, while the other was scanned with eyes closed.

Recently, several other data harmonisation tools have been proposed, such as ICA-AROMA (Pruim et al., 2015) and ComBat (Johnson et al., 2007; Yu et al., 2018). FIX and ICA-AROMA are both specifically designed for the denoising of rs-fMRI data using independent component analysis (ICA), while ComBat has been validated for DTI (Fortin et al., 2017), cortical thickness (Fortin et al., 2018), and functional connectivity (Yu et al., 2018) analyses. Further research should establish whether the combination of FIX and ComBat further enhances the quality and reproducibility of rs-fMRI, or whether combining these steps would prove redundant.

In new multicentre studies, the standardisation of protocols across centres is important to reduce noise as much as possible (Wegner et al., 2008; Zivadinov & Cox, 2008; Glover et al., 2012). However, our results in **chapter 5** and other data harmonisation studies (Pruim et al., 2015; Fortin et al., 2017, 2018; Yu et al., 2018) suggest that reanalysis of existing non-standardised rs-fMRI data may also be possible across sites. This possibility will especially be important to facilitate the valid comparisons of rare diseases and at-risk populations.

Neuroimaging biomarkers for differential diagnosis of FTD and AD

On MRI, FTD and AD are characterised by divergent patterns of structural and functional neurodegeneration in terms of grey matter atrophy (Zhang et al., 2011; Möller et al., 2015b), white matter degeneration (Zhang et al., 2011; Möller et al., 2015b; Daianu et al., 2016), and functional connectivity (Zhou et al., 2010). Grey matter atrophy is greater in FTD patients in the orbitofrontal, inferior and medial frontal gyrus, the anterior cingulate gyrus, caudate nucleus, and the nucleus accumbens (Zhang et al., 2011; Möller et al., 2015b), while AD patients have more grey matter atrophy in the bilateral occipital gyri and the left precuneus (Zhang et al., 2011). White matter changes are more abundant in FTD patients than AD patients. FTD patients have stronger diffusion abnormalities than AD patients in the uncinate fasciculi, forceps minor, and anterior thalamic radiation, while there are no areas that clearly show the opposite pattern (Zhang et al., 2011; Möller et al., 2015b; Daianu et al., 2016). Furthermore, functional connectivity in the default mode network and salience network are inversely affected in AD and FTD. In AD, functional connectivity in the default mode network is disrupted, while functional connectivity in the salience network is increased. Conversely, functional connectivity in the salience network is disrupted in FTD, while functional connectivity in the default mode network is increased (Zhou et al., 2010).

If present in early FTD and AD stages, these diverging patterns might serve as biomarkers to differentiate between FTD and AD. In **chapter 6**, we investigated grey matter atrophy, white matter diffusion, and functional connectivity in cognitively healthy subjects at risk for FTD (i.e., *MAPT* and *GRN* mutation carriers) and AD (i.e., apolipoprotein E ε4 [*APOE4*] carriers). However, we found no evidence for diverging patterns between the at-risk groups for FTD and AD in our cross-sectional sample. To gain more insight into the early differences between FTD and AD neuropathological processes, longitudinal studies in genetic at-risk groups may be required. Alternatively, memory clinic cohorts, especially patients without dementia diagnosis at first presentation, may provide the means to longitudinally study early stages of sporadic FTD and AD (Handels et al., 2012).

Explosive onset of neuropathology in genetic FTD

In genetic FTD, widespread presymptomatic changes were found in large samples of mutation carriers (Rohrer et al., 2015; Cash et al., 2018; Jiskoot et al., 2018a). However, in smaller samples, similar presymptomatic changes were generally not found (Borroni et al., 2008; Whitwell et al.,

2011a; Dopper et al., 2014; Panman et al., 2019), unless subjects were near symptom onset (Jiskoot et al., 2019). These results have led to the hypothesis that detectable neurodegenerative processes may start relatively shortly before symptom onset in genetic FTD. Our results further corroborate this hypothesis. We did not find convincing presymptomatic differences on multimodal MRI between presymptomatic FTD mutation carriers and non-carriers (**chapter 6**), and classification performance was modest though above chance level (**chapter 2**). Moreover, classification scores progressed similarly for presymptomatic mutation carriers and controls, unless the mutation carriers were near symptom onset, in which case the scores rose steeply (**chapter 3**). Finally, it was possible to predict symptom onset in presymptomatic mutation carriers within a timespan of four years beyond chance level, even in a small sample (**chapter 4**). These results might seem slightly at odds with large group studies in the Genetic Frontotemporal dementia Initiative (GENFI; Rohrer et al., 2013), which reported grey matter volume loss from up to 10 years before estimated symptom onset (Cash et al., 2018), and white matter DTI changes from up to 30 years before estimated symptom onset (Jiskoot et al., 2018a). Indeed, the acquisition of large data sets enables the detection of very small differences in group studies. However, these early and small differences are not necessarily useful for classification in the single-subject setting due to small effect sizes and large individual variance.

So far, few FTD mutation carriers have developed symptomatic FTD in longitudinal studies (**chapter 3, 4**). Given the explosive rate in which neurodegeneration develops in the years prior to symptom onset, continuation of these studies is crucial to develop more accurate and robust classification and prediction models.

Critical considerations and future directions

Sample size

For machine learning purposes (**chapters 2, 3,** and particularly **chapter 4**), we had relatively few data. Cross-validation crucially minimises overfitting in classification analyses, but leads to model uncertainty, especially in small sample sizes (Varoquaux, 2018). This resulted in low power to find classification models that statistically outperformed chance level, which may have resulted in false negative findings in **chapters 2, 4**. On the other hand, *MAPT* and *GRN* mutations, and *C9orf72* repeat expansion are each associated with its distinct neurodegenerative pattern (Seelaar et al., 2011; Whitwell et al., 2012; Jiskoot et al., 2018a), and one might argue that pooling them may have biased our outcomes towards the largest group, in our case the *GRN* mutation carriers, and introduced additional heterogeneity. As such, our small sample size led to methodological imperfections from both clinical and statistical perspectives. We balanced these two interests to present a unique proof of concept that presymptomatic changes in FTD mutation carriers can be detected on the single-subject level.

Longitudinal MRI data acquisition is expensive and presents a psychological burden for these rare mutation carriers. Therefore, our results are important as a proof of concept that longitudinal MRI acquisition may aid the development of early diagnostic FTD biomarkers. When more data become available, stratification of classification analyses across mutations, underlying frontotemporal lobar degeneration (FTLD) pathologies, and clinical FTD syndromes may lead to more accurate and robust models through increased homogeneity. Moreover, the combination of overlapping (Elahi et al., 2017) and diverging (Whitwell et al., 2010; Galantucci et al., 2011; Zhang et al., 2011; Mahoney et al., 2014; Tu et al., 2015; Daianu et al., 2016; Omer et al., 2017) neuroimaging changes associated with the different FTD syndromes may facilitate hierarchical classification in presymptomatic FTD mutation carriers, as has already been done in symptomatic AD and FTD variants (Kim et al., 2019). The low prevalence of FTD mutation carriers and heterogeneity within and between mutations are considerable challenges for the validation of potential biomarkers. However, advances in data harmonisation techniques will make it easier to combine neuroimaging data (**chapter 5**; Fortin et al., 2017, 2018; Yu et al., 2018) within and perhaps between large consortia, such as GENFI (Rohrer et al., 2013; Europe and Canada), and the Advancing Research and Treatment for Frontotemporal Lobar Degeneration / Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects (ARTFL/LEFFTDS; Staffaroni et al., 2019; USA and Canada) cohorts.

Diffusion imaging

In this dissertation, especially **chapters 2, 4**, we speculate that white matter diffusion metrics are the most promising biomarkers for early diagnosis in genetic FTD. We investigated white matter diffusion characteristics using the DTI model (Basser et al., 1994) in conjunction with tract-based spatial statistics (TBSS; Smith et al., 2006). However, both methods have their limitations. For example, when using a study-specific TBSS skeleton (e.g., in **chapter 6**), white matter pathology in the tract centres of patients or mutation carriers may influence the skeletonisation step, thus biasing results (Bach et al., 2014). In **chapters 2–4**, we used a predefined white matter skeleton to circumvent this problem. More general problems with DTI include the inability of DTI metrics such as FA to differentiate between intrinsic white matter properties and fibre coherence (Dell'Acqua & Tournier, 2019). Recent advances in diffusion MR acquisition, such as multiband excitation (Feinberg et al., 2010; Sotiropoulos et al., 2013), and modelling, such as spherical deconvolution (Tournier et al.,

2007; Dell'Acqua et al., 2010; Zhang et al., 2012; Raffelt et al., 2015; Dell'Acqua & Tournier, 2019), provide solutions to these problems, but have not been extensively used yet in clinical populations (Mito et al., 2018). We believe it is crucial to adopt these new diffusion MRI techniques in future studies involving FTD patients and presymptomatic mutation carriers to validate our results using novel white matter metrics that better capture microstructural white matter changes.

Classification analyses

For our classification analyses in **chapters 2–4**, we used elastic net (Zou & Hastie, 2005; Friedman et al., 2010), a type of regularised logistic regression, on multimodal MRI features in a repeated nested cross-validation scheme. Here, we review some strengths and limitations to this approach, and discuss the features we used.

Elastic net combines L1 (lasso; Tibshirani, 1996) and L2 (ridge; Hoerl & Kennard, 1970) penalties to regularise feature inclusion and the weight of each feature. The resulting models only include a subset of all features, which is essential when the number of features exceeds the number of subjects. However, though regularisation aids model performance and generalisation, it hinders the interpretability of the classification model (Shmueli, 2010). When redundant information is encoded in multiple correlated features, regularised regression models will minimise the sum of prediction error and penalty by including only one of these features in the classification model. Therefore, the features' beta weights in the classification model do not reliably indicate how meaningful the included features are. In neuroimaging classification studies, high beta weights in certain features, such as hippocampal atrophy, are sometimes claimed to verify the model's biological validity. In our work, we have tried to refrain from bold claims based on the classification models' beta weights.

We performed all classification analyses in a nested cross-validation scheme to ensure we did not overfit our classification models (Kriegeskorte et al., 2009). We used the inner loop of the cross-validation to tune the hyperparameters, and the outer loop to fit and test the classification model. As such, the outer loop test set is used for neither hyperparameter tuning, nor training. The drawback of cross-validation is that data partitioning increases variance, resulting in increased model uncertainty, especially in smaller samples (Varoquaux, 2018). We repeated the entire cross-validation procedure 50 times to reduce the variability arising from the random partitioning in training and test folds.

In addition to applying a nested cross-validation approach, we took further care to avoid overfitting by using feature selection procedures that reduced the data's dimensionality without relying on class differences in our sample. To define grey and white matter features, we used anatomical atlases, while we used an unsupervised data-driven ICA (Beckmann & Smith, 2004) approach to define functional connectivity features. Our grey matter classification features consisted of GMD in 96 regions defined by the Harvard-Oxford cortical atlas, and grey matter volume in 14 subcortical volumes using FSL FIRST segmentations. White matter features included WMD and diffusion metrics (i.e., FA, MD, AxD, and RD) in 20 Johns-Hopkins university white matter atlas tracts. Functional connectivity features comprised full and partial (i.e., L1-regularised; J. Friedman et al., 2008) correlations between 70 (**chapter 2, 3**) or 20 (**chapter 4**) ICA components.

We chose these specific features based on earlier work, in which AD patients were classified from controls (Schouten et al., 2016; Bouts et al., 2018), and FTD patients were classified from controls and AD patients (Bouts et al., 2018). In the FTD field, no comparisons have been made between different features of grey matter structure, white matter diffusion, or functional connectivity. Therefore, higher performances than reported in this dissertation might be possible with a different

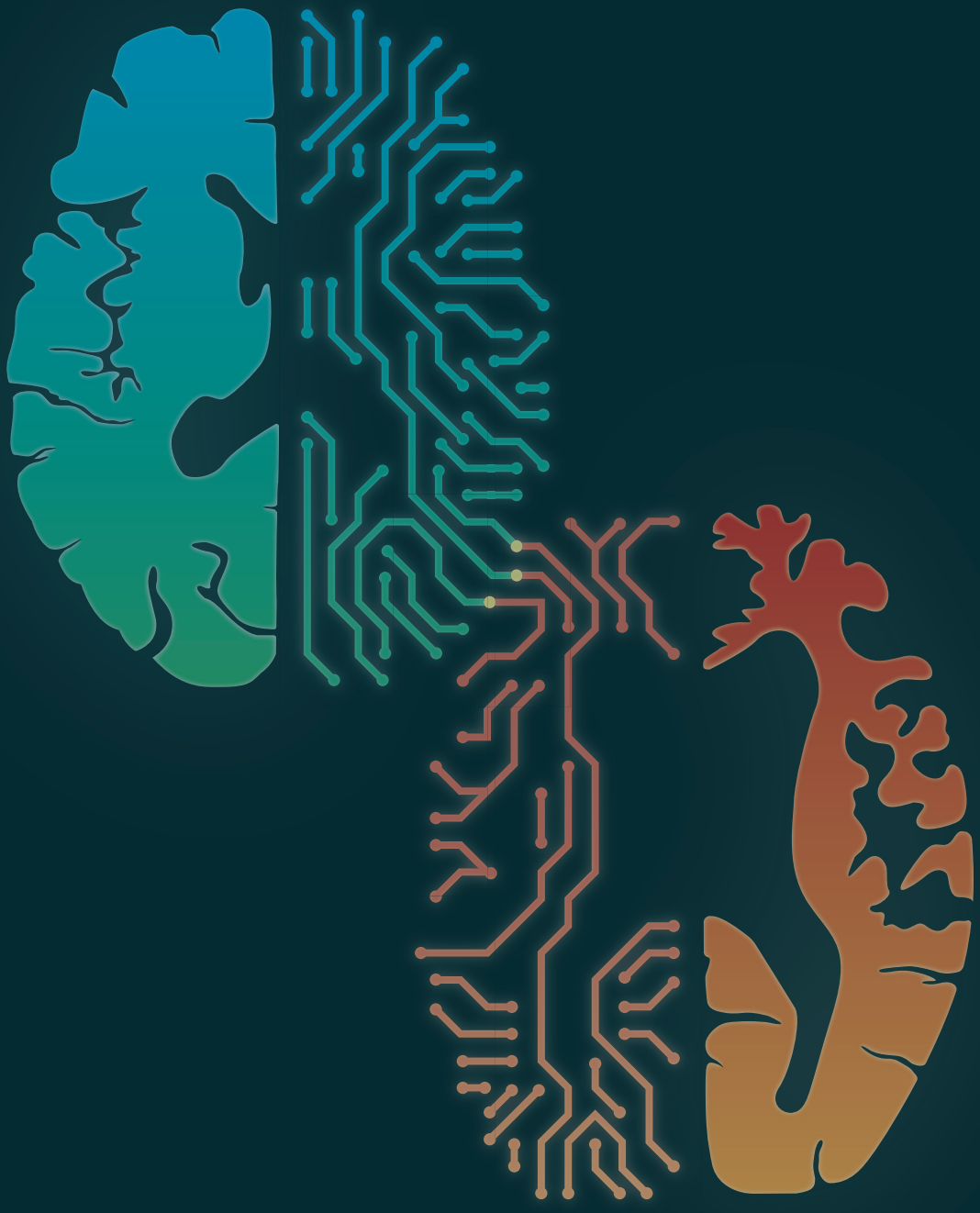
feature selection. However, in comparative AD classification studies for grey matter structure (de Vos et al., 2016), white matter diffusion (Schouten et al., 2017), and functional connectivity (de Vos et al., 2018) features, the features used here performed generally well.

Diagnostic uncertainty

In **chapters 3, 4**, we included the data of *MAPT* and *GRN* mutation carriers that had converted to FTD. Conversion was carefully determined by a multidisciplinary team according to the newest criteria for bvFTD (Rascovsky et al., 2011), primary progressive aphasia variants (Gorno-Tempini et al., 2011), and amyotrophic lateral sclerosis (Ludolph et al., 2015), which are also used in the clinic. The use of these diagnostic criteria may have introduced some circularity to our results, as atrophy on neuroimaging is one of the diagnostic criteria on which conversion was defined. However, a strength of this approach is that it facilitates the comparison to other studies and simplifies replication.

Conclusions

Neuroimaging group analyses in presymptomatic FTD mutation carriers have taught us much about the early pathways of FTLD neuropathology. This dissertation aimed to advance the development of presymptomatic MRI-based biomarkers for FTD towards clinical application through the use of machine learning. Importantly, eventual clinical implementation of machine learning seems straightforward, as machine learning can combine multiple complementary inputs (e.g., different MRI-based features) to calculate a single diagnostic score per subject. We showed that multimodal MRI-based biomarkers can identify presymptomatic FTD mutation carriers on the individual level, and are sensitive to increasing FTLD pathology over time. Moreover, we found that diffusion MRI can predict symptom onset within 4 years in genetic FTD. Finally, we showed that the reduction of structured noise in rs-fMRI aids the combination of such data in multicentre studies. Though our findings require replication in larger cohorts, this dissertation provides the optimistic prospect that neuroimaging FTD biomarkers may aid timely FTD diagnosis in individual patients.



Chapter 8

Samenvatting & summary

Samenvatting

Frontotemporale dementie (FTD) is een heterogene neurodegeneratieve ziekte die zich kenmerkt door degeneratie van de frontale en temporale hersenkwabben, resulterend in gedragsstoornissen (gedragsvariant FTD; Engelse afkorting bvFTD) en taalstoornissen (primaire progressieve afasie; PPA). Daarnaast ontwikkelen sommige FTD-patiënten symptomen van andere neurologische ziekten zoals atypische parkinsonismen of amyotrofe lateraalsclerose. FTD-patiënten ontwikkelen hun eerste symptomen vaak voor de leeftijd van 65 jaar, en zijn daarmee in het algemeen jonger dan patiënten met andere vormen van dementie zoals de ziekte van Alzheimer. In vroege stadia is de diagnose FTD echter lastig te stellen vanwege de grote variatie in symptomen, welke veelal overlappen met psychiatrische ziektes en andere dementies.

Momenteel zijn er geen bewezen effectieve therapieën om FTD te remmen of te genezen. Niettemin is de kennis over de ontstaanswijze van FTD in de afgelopen jaren fors toegenomen, en wordt steeds meer onderzoek gedaan naar *ziektemodificerende therapieën*. Dit zijn medicijnen die onderliggende pathologische processen van FTD (zoals eiwitstapeling) trachten te remmen, om te voorkomen dat hersencellen afsterven. Een probleem bij de ontwikkeling van ziektemodificerende therapieën is dat FTD-patiënten ten tijde van diagnose vaak al onomkeerbare hersenschade (ook wel hersenatrofie genoemd) hebben, terwijl deze medicijnen juist bedoeld zijn om atrofie te voorkomen. Voor de ontwikkeling van ziektemodificerende therapieën is het daarom cruciaal om FTD-patiënten in een vroeger stadium te kunnen diagnosticeren.

Ongeveer 10–30% van alle FTD-patiënten heeft een familiaire vorm, welke meestal wordt veroorzaakt door mutaties in de genen microtubule-associated protein tau (*MAPT*), progranuline (*GRN*), of een repeat-expansie in het gen chromosoom 9 open reading frame 72 (*C9orf72*). Deze genetische vormen van FTD bieden de unieke mogelijkheid om mutatiedragers te onderzoeken in een *presymptomatisch* stadium, waarin wel al pathologische veranderingen optreden, maar mutatiedragers nog geen symptomen hebben. Op deze manier kunnen *biomarkers* ontwikkeld worden die het mogelijk maken om FTD in een vroeg stadium op te sporen, ziekteprogressie te volgen, of het effect van medicatie te evalueren. Dit proefschrift stelt zich ten doel de ontwikkeling van *magnetic resonance imaging* (MRI-)biomarkers voor de vroege opsporing van FTD te bevorderen.

Hoofdstuk 1 geeft een algemene introductie over FTD en beschrijft de doelen van de verschillende studies binnen dit proefschrift.

In **hoofdstukken 2–4** onderzochten we of MRI-biomarkers voor FTD op individueel niveau diagnostische waarde hebben om FTD-gerelateerde verschillen aan te tonen in het presymptomatische stadium van de ziekte. Traditionele groepsstudies hebben inzicht verschaft in de processen die in vroege stadia van FTD optreden, en hebben op basis daarvan potentiële biomarkers aangedragen. De translatie van deze potentiële biomarkers naar de klinische praktijk vergt validatie in een individuele setting. Dat wil zeggen dat biomarkers een betrouwbare diagnostische uitspraak moeten kunnen doen over een individu.

In **hoofdstuk 2** gebruikten we MRI-maten om te classificeren of een persoon drager is van een FTD-mutatie (*MAPT* mutatie, *GRN* mutatie of *C9orf72* repeat-expansie) of niet-drager is. De classificatie overtrof kansniveau, hoewel deze nog niet accuraat genoeg was voor klinisch gebruik. De best presterende classificatiemodellen waren gebaseerd op maten van de witte stof.

Vervolgens hebben we in **hoofdstuk 3** hetzelfde cohort een aantal jaren gevolgd om te onderzoeken hoe de *classificatiescore* (i.e. de uitkomst van een classificatiemodel) zich zou ontwikkelen in mutatiedragers en controles. De classificatiescores van mutatiedragers die gezond bleven tijdens de studie en niet-dragers gedroegen zich niet verschillend mettertijd. Een aantal mutatiedragers

converteerden echter gedurende de studie naar FTD: bij hen werd symptomatische FTD vastgesteld. Vergeleken met de gezonde mutatiedragers lieten deze converterende mutatiedragers gaandeweg een sterke stijging in classificatiescores zien. Dit resultaat onderschreef ons vermoeden dat FTD-mutatiedragers—voor wat betreft MRI-afwijkingen—lange tijd stabiel blijven tot enkele jaren voor conversie, en vervolgens op explosieve wijze afwijkingen ontwikkelen.

Gezien dit explosieve karakter van MRI-veranderingen vroegen we ons af of het mogelijk zou zijn om, met behulp van MRI-maten, toekomstige conversie te voorspellen in FTD-mutatiedragers op een moment dat ze nog gezond waren. In **hoofdstuk 4** onderzochten we MRI-data van mutatiedragers die na vier jaar nog gezond waren, en van mutatiedragers die na vier jaar geconverteerd waren naar FTD. Het bleek mogelijk om beter dan kans te voorspellen wie wel en niet binnen vier jaar zou converteren met behulp van diffusiegewogen MRI van de witte stof. Hoewel deze studie in grotere onderzoekspopulaties herhaald moet worden ter bevestiging, biedt dit resultaat perspectief om FTD-mutatiedragers te includeren in geneesmiddelenonderzoek voordat ze hersenatrofie hebben.

In **hoofdstuk 5** beoogden we de vergelijkbaarheid van *resting-state functional MRI* (rs-fMRI; MRI waarmee wordt gemeten hoe verschillende hersengebieden samenwerken als functionele netwerken) data tussen MRI-scanners te bevorderen, zodat data van verschillende centra gezamenlijk geanalyseerd kunnen worden. Hiertoe pasten we *FIX* (*FMRIB's ICA-based X-noiseifier*) toe op rs-fMRI data van verschillende scanners. FIX is een methode om gestructureerde ruis, afkomstig van hardware- en softwareverschillen tussen scanners, te reduceren. Vrijwel alle verschillen in functionele connectiviteit tussen de twee centra waren na toepassing van FIX verminderd tot niet-significante niveaus. De overgebleven significante verschillen bevonden zich voornamelijk in het visuele netwerk en waren zeer waarschijnlijk een gevolg van een verschil in scanprotocol. De ene groep was namelijk met de ogen open gescand, terwijl de andere groep met de ogen dicht was gescand. Onze resultaten demonstreren dat FIX gestructureerde ruis reduceert, terwijl fysiologische verschillen in functionele connectiviteit behouden blijven. FIX lijkt daarom een geschikte methode om rs-fMRI data van verschillende centra te harmoniseren.

Tot slot onderzochten we potentiële biomarkers voor de differentiatie tussen vroege stadia van FTD en de ziekte van Alzheimer. FTD-patiënten hebben een ander patroon van MRI-afwijkingen in de grijze stof, witte stof en qua functionele connectiviteit dan patiënten met de ziekte van Alzheimer. In **hoofdstuk 6** onderzochten we of deze verschillende patronen van afwijkingen ook aanwezig zijn in gezonde groepen met genetisch risico op deze ziektes. Hoewel we veranderingen vonden in de witte stof van dragers van apolipoprotein E ε4 (*APOE4*; risicoallel voor de ziekte van Alzheimer) ten opzichte van hun controlegroep, waren deze veranderingen niet sterk genoeg om een uiteenlopend patroon aan te tonen ten opzichte van een groep *MAPT*- en *GRN*-dragers met genetisch risico op FTD. Gezien het transversale karakter van de studie was geen informatie beschikbaar over het klinisch verloop bij deze personen. Het is aannemelijk dat veel dragers nog te jong waren om al afwijkingen te vertonen gerelateerd aan FTD of de ziekte van Alzheimer. Om te onderzoeken in welk stadium patronen van MRI-afwijkingen onderscheidend worden tussen FTD en de ziekte van Alzheimer is daarom longitudinaal onderzoek nodig in genetische risicogroepen.

Hoofdstuk 7 beschrijft de conclusies van dit proefschrift, plaatst kritische noten bij de gebruikte methoden en doet aanbevelingen voor toekomstig onderzoek. Concluderend demonstreerden we dat MRI-biomarkers presymptomatische FTD-mutatiedragers op individueel niveau kunnen identificeren, dat deze biomarkers sensitief zijn voor een toename van FTD-gerelateerde MRI-veranderingen over de tijd en zelfs kunnen voorspellen welke FTD-mutatiedragers binnen vier jaar zullen converteren. Verder onderzoek in grotere steekproeven zal moeten aantonen of deze classificatiemodellen nauwkeuriger worden als ze worden toegespitst op één type genmutatie of op

één klinische FTD-variant. Daarnaast levert dit proefschrift aanvullend bewijs voor de hypothesen dat de ontwikkeling van MRI-veranderingen in het brein kort voor de conversie explosief accelereert, en dat diffusie-MRI de meest veelbelovende maat is om vroege veranderingen in genetische FTD op te sporen. Tot slot toonden we aan dat de reductie van gestructureerde ruis het mogelijk maakt om rs-fMRI data van meerdere centra te combineren. Hoewel onze bevindingen bevestiging nodig hebben in grotere steekproeven, biedt dit proefschrift het optimistische vooruitzicht dat MRI-biomarkers kunnen helpen om FTD in een vroeger stadium te diagnosticeren.

Summary

Frontotemporal dementia (FTD) is a heterogeneous neurodegenerative disease characterised by the progressive degeneration of the frontal and temporal lobes, which results in behavioural (behavioural variant FTD; bvFTD) and language (primary progressive aphasia; PPA) disorders. Additionally, FTD patients may develop concurring symptoms of atypical parkinsonism or amyotrophic lateral sclerosis. Patients with FTD are generally younger than patients with different types of dementia, as the typical age of FTD onset is below 65 years. Still, FTD is difficult to diagnose in early stages, as symptoms may vary and overlap with psychiatric disease and/or other dementias.

No effective therapies currently exist to cure FTD or slow disease progression. However, our knowledge of FTD pathophysiology has increased substantially over the past decades, and increasing efforts are currently made to develop *disease modifying treatments*. These treatments aim to inhibit or reverse underlying pathological FTD processes, e.g., protein accumulation, in order to prevent neuronal cell death. A major challenge to developing disease modifying treatments is the presence of atrophy (i.e., irreversible brain damage) at the time of FTD diagnosis, which limits the window of opportunity for treatment. Therefore, reliably diagnosing FTD at an earlier stage (specifically, before atrophy occurs) is crucial for the development of disease modifying treatments.

Approximately 10–30% of all FTD patients have a familial form, which is often caused by mutations in the genes microtubule-associated protein tau (*MAPT*), progranulin (*GRN*), or a repeat expansion in the gene chromosome 9 open reading frame 72 (*C9orf72*). Genetic FTD families offer the unique opportunity to study mutation carriers in a *presymptomatic* stage, where early pathological changes may already occur, but subjects are cognitively healthy. Thus, *biomarkers* may be developed for early detection of FTD, disease progression, and treatment effect evaluation. This dissertation aims to advance the development of *magnetic resonance imaging* (MRI-)based biomarkers for the early detection of FTD.

In **chapter 1**, a general introduction of FTD is given and the goals of the studies within this dissertation are described.

Chapters 2–4 describe studies in which we investigated whether MRI biomarkers for FTD have diagnostic value on the single-subject level to detect FTD-related differences in the presymptomatic disease stage. Traditional group studies have provided insight in the processes that occur in early stages of FTD, and have suggested potential biomarkers for early detection. The translation of these potential biomarkers requires validation in the individual setting, i.e., biomarkers need to be able to make reliable diagnoses in individuals.

Chapter 2 describes the use of MRI measures to classify whether or not a subject is carrier of a pathogenic FTD gene mutation (*MAPT* mutation, *GRN* mutation, or *C9orf72* repeat expansion). The classification was beyond chance level, though not sufficiently accurate for clinical use. The best performing classification models were based on white matter measures.

Next, we followed the same cohort over time in **chapter 3** to study how the *classification score*, i.e., the outcome of the classification model, develops over time in FTD mutation carriers and non-carriers. The classification scores of mutation carriers who remained cognitively healthy during the study and the classification scores of non-carriers did not differ over time. However, some mutation carriers *converted* during the study: they were diagnosed with symptomatic FTD. Compared to the healthy mutation carriers, these converting mutation carriers showed a strong increase in classification score over time. This result confirmed our suspicion that FTD mutation carriers, in terms of MRI changes, remain relatively stable until a couple of years before conversion, after which neuropathological changes progressively develop.

Given the accelerating nature of MRI changes leading up to symptom onset, we wondered whether we could use MRI measures to predict conversion in FTD mutation carriers at a time when they were still healthy. In **chapter 4**, we studied MRI data of FTD mutation carriers who were still healthy after four years of follow-up, and of mutation carriers that had converted to FTD after four years. We found it was possible to predict beyond chance level who would and would not convert to FTD within four years, using diffusion-weighted MRI of the white matter. While this study should be replicated in a larger sample for confirmation, this result offers perspective to include FTD mutation carriers in clinical trials before brain atrophy occurs.

In **chapter 5**, we aimed to advance the combination of *resting-state functional MRI* (rs-fMRI; MRI that measures how brain regions coactivate in functional networks) data between MRI scanners, in order for multicentre data to be collectively analysed. To that end, we applied *FIX* (*FMRIB's ICA-based X-noiseifier*) on multicentre rs-fMRI data from different scanners. FIX is a method to reduce structured noise resulting from hardware and software differences from rs-fMRI data. Nearly all differences in functional connectivity between the centres were reduced to non-significant levels after the application of FIX. The leftover differences were predominantly situated in the visual network and were presumably a consequence of a difference in scan protocol, as one group was scanned with eyes open, while the other group was scanned with eyes closed. Our results demonstrate that FIX reduces structured noise, while physiological differences in functional connectivity are preserved. Therefore, FIX seems a suitable method to harmonise rs-fMRI data from different centres.

Lastly, we studied potential biomarkers for the differentiation between early stages of FTD and Alzheimer's disease. FTD patients have a different pattern of MRI changes in grey matter, white matter, and functional connectivity than patients with Alzheimer's disease. In **chapter 6**, we studied whether such patterns of differences were also apparent in cognitively healthy subjects at genetic risk for these diseases. Although we found differences in white matter diffusion measures of apolipoprotein E ε4 (*APOE4*; risk allele for Alzheimer's disease) carriers compared to non-carriers, these differences were not strong enough to result in a divergent pattern compared to a group of *MAPT* and *GRN* carriers (at genetic risk for FTD). Due to the cross-sectional design of the study, no clinical follow-up was available for these subjects. It is plausible that some mutation carriers were still too young to show MRI changes related to FTD or Alzheimer's disease. Longitudinal research in genetic risk groups is therefore necessary to investigate at which stage patterns of MRI changes start to differ between FTD and Alzheimer's disease.

Chapter 7 describes the conclusions of this dissertation, provides methodological considerations, and presents recommendations for future research. To conclude, we demonstrated that MRI biomarkers can identify presymptomatic FTD mutation carriers at an individual level. We found that these biomarkers are sensitive to an increase in FTD-related MRI changes over time, and can furthermore predict which FTD mutation carriers will convert within four years. Further research in larger samples will have to show whether classification models become more accurate when specifically applied on one type of gene mutation or one clinical FTD variant. This dissertation additionally provides evidence for the hypotheses that the development of MRI changes in the brain accelerates shortly before conversion, and suggests that diffusion-weighted MRI is the most promising measure to detect early changes in genetic FTD. Finally, we showed that the reduction of structured noise facilitates the combination of rs-fMRI data across centres. While our findings require replication in larger samples, this dissertation offers the optimistic prospect that MRI biomarkers may aid early-stage FTD diagnosis.

Appendix

Abbreviations

Bibliography

List of publications

Dankwoord

Curriculum vitae

Abbreviations

$3DT_1w$	3-dimensional T_1-weighted MRI
ABIDE	Autism Brain Imaging Data Exchange
AD	Alzheimer's disease
ALFF	amplitude of low frequency fluctuations
ALS	amyotrophic lateral sclerosis
APOE4	apolipoprotein E ε4
ARTFL	Advancing Research and Treatment for FTLD
ATR	anterior thalamic radiation
AxD	axial diffusivity
AUC	area under the ROC curve
BNT	Boston naming test
bvFTD	behavioural variant FTD
C9orf72	chromosome 9 open reading frame 72
CBS	corticobasal syndrome
CGC	cingulum in the cingulate gyrus area
CGH	cingulum in the hippocampal area
CHMP2B	charged multivesicular body protein 2b
CST	corticospinal tract
DNA	deoxyribonucleic acid
DTI	diffusion tensor imaging
DWI	diffusion-weighted imaging
FA	fractional anisotropy
FCor	full correlations between ICA components' time courses
FDA	United States Food and Drug Administration
FIX	FMRIB's ICA-based X-noiseifier
FMA	forceps major
FMI	forceps minor
fMRI	functional MRI
FMRIB	Functional Magnetic Resonance Imaging of the Brain Centre
FOV	field of view
FSL	FMRIB Software Library
FTD	frontotemporal dementia
FTD-RisC	FTD Risk Cohort
FTLD	frontotemporal lobar degeneration
FUS	fused in sarcoma protein
FUS	fused in sarcoma gene
FWE(RC)	family-wise error (rate corrected)
GENFI	Genetic Frontotemporal Dementia Initiative
GICA	group-level ICA
GMD	grey matter density
GRN	progranulin
IC(A)	independent component (analysis)
IFOF	inferior fronto-occipital fasciculus
ILF	inferior longitudinal fasciculus

L	left
LDST	letter digit substitution test
LEFFTDS	Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects
LUMC	Leiden University Medical Centre
lvPPA	logopenic variant PPA
MAPT	microtubule-associated protein tau
MD	mean diffusivity
MMSE	mini-mental state examination
MRI	magnetic resonance imaging
nfvPPA	non-fluent variant PPA
OCMR	Oxford Centre for Clinical Magnetic Resonance Research
PCor	L1-regularised partial correlations between ICA components' time courses
PE	parameter estimate
PET	positron emission tomography
PPA	primary progressive aphasia
PSP	progressive supranuclear palsy
PVN	primary/medial visual network
R	right
RAVLT	Rey auditory verbal learning test
RD	radial diffusivity
RNA	ribonucleic acid
ROC	receiver operating characteristic
rs-fMRI	resting-state fMRI
RSN	resting-state network
SAT	semantic association test
SLF	superior longitudinal fasciculus
svPPA	semantic variant PPA
TBSS	tract-based spatial statistics
TDP-43	transactive response DNA binding protein 43 kDa
TE	echo time
TFCE	threshold-free cluster enhancement
TMT	trail making test
TNR	true-negative rate
TPR	true-positive rate
TR	repetition time
TSLF	temporal projection of the SLF
UF	uncinate fasciculus
VAT	visual association test
VCP	valosin-containing protein
WAIS	Wechsler adult intelligence scale
WCST	Wisconsin card sorting test
WMD	white matter density

Bibliography

Adluru, N., Destiche, D.J., Lu, S.Y.-F., Doran, S.T., Birdsill, A.C., Melah, K.E., Okonkwo, O.C., Alexander, A.L., Dowling, N.M., Johnson, S.C., Sager, M.A., and Bendlin, B.B. (2014). White matter microstructure in late middle-age: Effects of apolipoprotein E4 and parental family history of Alzheimer's disease. *NeuroImage: Clinical*, 4:730–42.

Agosta, F., Vossel, K.A., Miller, B.L., Migliaccio, R., Bonasera, S.J., Filippi, M., Boxer, A.L., Karydas, A., Possin, K.L., and Gorno-Tempini, M.L. (2009). Apolipoprotein E epsilon4 is associated with disease-specific effects on brain atrophy in Alzheimer's disease and frontotemporal dementia. *Proceedings of the National Academy of Sciences of the USA*, 106(6):2018–22.

Agosta, F., Scola, E., Canu, E., Marcone, A., Magnani, G., Sarro, L., Copetti, M., Caso, F., Cerami, C., Comi, G., Cappa, S.F., Falini, A., and Filippi, M. (2012). White matter damage in frontotemporal lobar degeneration spectrum. *Cerebral Cortex*, 22(12):2705–14.

Anderson, J.L.R., Jenkinson, M., and Smith, S.M. (2007). Non-linear registration aka Spatial normalisation. FMRIB Technical Report TR07JA2.

Armstrong, M.J., Litvan, I., Lang, A.E., Bak, T.H., Bhatia, K.P., Borroni, B., Boxer, A.L., Dickson, D.W., Grossman, M., Hallett, M., Josephs, K.A., Kertesz, A., Lee, S.E., Miller, B.L., Reich, S.G., Riley, D.E., Tolosa, E., Tröster, A.I., Vidailhet, M., and Weiner, W.J. (2013). Criteria for the diagnosis of corticobasal degeneration. *Neurology*, 80(5):496–503.

Army Individual Test Battery (1944). Manual of directions and scoring. *War Department, Adjunct General's Office, Washington, DC*.

Bach, M., Laun, F.B., Leemans, A., Tax, C.M.W., Biessels, G.J., Stieltjes, B., and Maier-Hein, K.H. (2014). Methodological considerations on tract-based spatial statistics (TBSS). *NeuroImage*, 100:358–69.

Basser, P.J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–67.

Beckmann, C.F., and Smith, S.M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 23(2):137–52.

Beckmann, C.F., DeLuca, M., Devlin, J.T., and Smith, S.M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B, Biological Sciences*, 360(1457):1001–13.

Beckmann, C.F., Mackay, C.E., Filippini, N., and Smith, S.M. (2009). Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. *OHBM*.

Benussi, A., Padovani, A., and Borroni, B. (2015). Phenotypic Heterogeneity of Monogenic Frontotemporal Dementia. *Frontiers in Aging Neuroscience*, 7:171.

Bertrand, A., Wen, J., Rinaldi, D., Houot, M., Sayah, S., Camuzat, A., Fournier, C., Fontanella, S., Routier, A., Couratier, P., Pasquier, F., Habert, M.-O., Hannequin, D., Martinaud, O., Caroppo, P., Levy, R., Dubois, B., Brice, A., Durrleman, S., Colliot, O., and Le Ber, I. (2018). Early Cognitive, Structural, and Microstructural Changes in Presymptomatic C9orf72

Carriers Younger Than 40 Years. *JAMA Neurology*, 75(2):236–45.

Bey, K., Montag, C., Reuter, M., Weber, B., and Markett, S. (2015). Susceptibility to everyday cognitive failure is reflected in functional network interactions in the resting brain. *NeuroImage*, 121:1–9.

Biswal, B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kötter, R., Li, S., Lin, C., Lowe, M.J., Mackay, C.E., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.A.R.B., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.-F., Zhang, H., Castellanos, F.X., and Milham, M.P. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the USA*, 107(10):4734–9.

Borroni, B., Alberici, A., Premi, E., Archetti, S., Garibotto, V., Agosti, C., Gasparotti, R., Di Luca, M., Perani, D., and Padovani, A. (2008). Brain magnetic resonance imaging structural changes in a pedigree of asymptomatic progranulin mutation carriers. *Rejuvenation Research*, 11(3):585–95.

Borroni, B., Alberici, A., Cercignani, M., Premi, E., Serra, L., Cerini, C., Cosseddu, M., Pettenati, C., Turla, M., Archetti, S., Gasparotti, R., Caltagirone, C., Padovani, A., and Bozzali, M. (2012). Granulin mutation drives brain damage and reorganization from preclinical to symptomatic FTLD. *Neurobiology of Aging*, 33(10):2506–20.

Bouts, M.J.R.J., Möller, C., Hafkemeijer, A., van Swieten, J.C., Dopper, E.G.P., van der Flier, W.M., Vrenken, H., Wink, A.M., Pijnenburg, Y.A.L., Scheltens, P., Barkhof, F., Schouten, T.M., de Vos, F., Feis, R.A., van der Grond, J., de Rooij, M., and Rombouts, S.A.R.B. (2018). Single Subject Classification of Alzheimer's Disease and Behavioral Variant Frontotemporal Dementia Using Anatomical, Diffusion Tensor, and Resting-State Functional Magnetic Resonance Imaging. *Journal of Alzheimer's Disease*, 62(4):1827–39.

Boxer, A.L., Yu, J.-T., Golbe, L.I., Litvan, I., Lang, A.E., and Höglinger, G.U. (2017). Advances in progressive supranuclear palsy: new diagnostic criteria, biomarkers, and therapeutic approaches. *The Lancet Neurology*, 16(7):552–63.

Bron, E.E., Smits, M., Papma, J.M., Steketee, R.M.E., Meijboom, R., de Groot, M., van Swieten, J.C., Niessen, W.J., and Klein, S. (2017). Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using structural and advanced MRI. *European Radiology*, 27(8):3372–82.

Brooks, B.R., Miller, R.G., Swash, M., Munsat, T.L., and World Federation of Neurology Research Group on Motor Neuron Diseases (2000). El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*, 1(5):293–9.

Cacciaglia, R., Molinuevo, J.L., Falcón, C., Brugulat-Serrat, A., Sánchez-Benavides, G., Gramunt, N., Esteller, M., Morán, S., Minguillón, C., Fauria, K., and Gispert, J.D. (2018). Effects of

APOE-ε4 allele load on brain morphology in a cohort of middle-aged healthy individuals with enriched genetic risk for Alzheimer's disease. *Alzheimer's & Dementia*, 14(7):902–12.

Canu, E., Agosta, F., Mandic-Stojmenovic, G., Stojković, T., Stefanova, E., Inuggi, A., Imperiale, F., Copetti, M., Kostic, V.S., and Filippi, M. (2017). Multiparametric MRI to distinguish early onset Alzheimer's disease and behavioural variant of frontotemporal dementia. *NeuroImage: Clinical*, 15:428–38.

Casey, B.J., Cohen, J.D., O'Craven, K., Davidson, R.J., Irwin, W., Nelson, C., Noll, D.C., Hu, X., Lowe, M.J., Rosen, B.R., Truwitt, C.L., and Turski, P.A. (1998). Reproducibility of fMRI results across four institutions using a spatial working memory task. *NeuroImage*, 8(8):249–61.

Cash, D.M., Bocchetta, M., Thomas, D.L., Dick, K.M., van Swieten, J.C., Borroni, B., Galimberti, D., Masellis, M., Tartaglia, M.C., Rowe, J.B., Graff, C., Tagliavini, F., Frisoni, G.B., Laforce, R., Finger, E., de Mendonça, A., Sorbi, S., Rossor, M.N., Ourselin, S., and Rohrer, J.D. (2018). Patterns of gray matter atrophy in genetic frontotemporal dementia: results from the GENFI study. *Neurobiology of Aging*, 62:191–6.

Cavedo, E., Lista, S., Rojkova, K., Chiesa, P.A., Houot, M., Brueggen, K., Blautzik, J., Bokde, A.L.W., Dubois, B., Barkhof, F., Pouwels, P.J.W., Teipel, S., and Hampel, H. (2017). Disrupted white matter structural networks in healthy older adult APOE ε4 carriers – An international multicenter DTI study. *Neuroscience*, 357:119–33.

Chao, L.L., Schuff, N., Clevenger, E.M., Mueller, S.G., Rosen, H.J., Gorno-Tempini, M.L., Kramer, J.H., Miller, B.L., and Weiner, M.W. (2007). Patterns of White Matter Atrophy in Frontotemporal Lobar Degeneration. *Archives of Neurology*, 64(11):1619–24.

Chare, L., Hodges, J.R., Leyton, C.E., McGinley, C., Tan, R.H., Kril, J.J., and Halliday, G.M. (2014). New criteria for frontotemporal dementia syndromes: clinical and pathological diagnostic implications. *Journal of Neurology, Neurosurgery, and Psychiatry*, 85(8):865–70.

Chen, G., Cox, R.W., Glen, D.R., Rajendra, J.K., Reynolds, R.C., and Taylor, P.A. (2019). A tail of two sides: Artificially doubled false positive rates in neuroimaging due to the sidedness choice with t-tests. *Human Brain Mapping*, 40(3):1037–43.

Cherbuin, N., Anstey, K.J., Sachdev, P.S., Maller, J.J., Meslin, C., Mack, H.A., Wen, W., and Easteal, S. (2008). Total and regional gray matter volume is not related to APOE*E4 status in a community sample of middle-aged individuals. *The Journals of Gerontology: Series A*, 63(5):501–4.

Cocozza, S., Saccà, F., Cervo, A., Marsili, A., Russo, C.V., Maria Delle Acque Giorgio, S., De Michele, G., Filla, A., Brunetti, A., and Quarantelli, M. (2015). Modifications of resting state networks in spinocerebellar ataxia type 2. *Movement Disorders*, 30(10):1382–90.

Cohn-Hokke, P.E., van Swieten, J.C., Pijnenburg, Y.A.L., Tibben, A., Meijers-Heijboer, H., and Kievit, A. (2018). The Effect of Predictive Testing in Adult-Onset Neurodegenerative Diseases on Social and Personal Life. *Journal of Genetic Counseling*, 27(4):947–54.

Costafreda, S.G., Brammer, M.J., Vêncio, R.Z.N., Mourão, M.L., Portela, L.A.P., de Castro, C.C., Giampietro, V.P., and Amaro, E. (2007). Multisite fMRI reproducibility of a motor task

using identical MR systems. *Journal of Magnetic Resonance Imaging*, 26(4):1122–6.

Cummings, J., Ritter, A., and Zhong, K. (2018). Clinical Trials for Disease-Modifying Therapies in Alzheimer's Disease: A Primer, Lessons Learned, and a Blueprint for the Future. *Journal of Alzheimer's Disease*, 64:3–22.

Daianu, M., Mendez, M.F., Baboyan, V.G., Jin, Y., Melrose, R.J., Jimenez, E.E., and Thompson, P.M. (2016). An advanced white matter tract analysis in frontotemporal dementia and early-onset Alzheimer's disease. *Brain Imaging and Behavior*, 10(4):1038–53.

Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., and Beckmann, C.F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of the USA*, 103(37):13848–53.

Davatzikos, C., Resnick, S.M., Wu, X., Parmpi, P., and Clark, C.M. (2008). Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. *NeuroImage*, 41(4):1220–7.

De Luca, M., Beckmann, C.F., De Stefano, N., Matthews, P.M., and Smith, S.M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage*, 29(4):1359–67.

de Vos, F., Schouten, T.M., Hafkemeijer, A., Dopper, E.G.P., van Swieten, J.C., de Rooij, M., van der Grond, J., and Rombouts, S.A.R.B. (2016). Combining multiple anatomical MRI measures improves Alzheimer's disease classification. *Human Brain Mapping*, 37(5):1920–9.

de Vos, F., Koini, M., Schouten, T.M., Seiler, S., van der Grond, J., Lechner, A., Schmidt, R., de Rooij, M., and Rombouts, S.A.R.B. (2018). A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. *NeuroImage*, 167:62–72.

Dell'Acqua, F., Scifo, P., Rizzo, G., Catani, M., Simmons, A., Scotti, G., and Fazio, F. (2010). A modified damped Richardson-Lucy algorithm to reduce isotropic background effects in spherical deconvolution. *NeuroImage*, 49(2):1446–58.

Dell'Acqua, F., and Tournier, J.-D. (2019). Modelling white matter with spherical deconvolution: How and why? *NMR in Biomedicine*, 32(4):e3945.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D.A., Gallagher, L., Kennedy, D.P., Keown, C.L., Keysers, C., Lainhart, J.E., Lord, C., Luna, B., Menon, V., Minshew, N.J., Monk, C.S., Mueller, S., Müller, R.-A., Nebel, M.B., Nigg, J.T., O'Hearn, K., Pelphrey, K.A., Peltier, S.J., Rudie, J.D., Sunaert, S., Thioux, M., Tyszka, J.M., Uddin, L.Q., Verhoeven, J.S., Wenderoth, N., Wiggins, J.L., Mostofsky, S.H., and Milham, M.P. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19(6):659–67.

Doesborgh, S.J.C., van de Sandt-Koenderman, W.M.E., Dippel, D.W.J., van Harskamp, F., Koudstaal, P.J., and Visch-Brink, E.G. (2003). Linguistic deficits in the acute phase of stroke. *Journal of Neurology*, 250(8):977–82.

Dopper, E.G.P., Rombouts, S.A.R.B., Jiskoot, L.C., den Heijer, T., de Graaf, J.R.A., de Koning, I., Hammerschlag, A.R., Seelaar, H., Seeley, W.W., Veer, I.M., van Buchem, M.A., Rizzu, P., and

van Swieten, J.C. (2014). Structural and functional brain connectivity in presymptomatic familial frontotemporal dementia. *Neurology*, 83(2):e19-26.

Douaud, G., Smith, S.M., Jenkinson, M., Behrens, T., Johansen-Berg, H., Vickers, J., James, S., Voets, N., Watkins, K., Matthews, P.M., and James, A. (2007). Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain*, 130(9):2375–86.

Ekman, P., and Friesen, W.V. (1976). Pictures of facial affect. *Consulting Psychologists Press, Palo Alto*.

Elahi, F.M., Marx, G., Cobigo, Y., Staffaroni, A.M., Kornak, J., Tosun, D., Boxer, A.L., Kramer, J.H., Miller, B.L., and Rosen, H.J. (2017). Longitudinal white matter change in frontotemporal dementia subtypes and sporadic late onset Alzheimer's disease. *NeuroImage: Clinical*, 16:595–603.

Farb, N.A.S., Grady, C.L., Strother, S., Tang-Wai, D.F., Masellis, M., Black, S., Freedman, M., Pollock, B.G., Campbell, K.L., Hasher, L., and Chow, T.W. (2013). Abnormal network connectivity in frontotemporal dementia: Evidence for prefrontal isolation. *Cortex*, 49(7):1856–73.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–74.

Feige, B., Scheffler, K., Esposito, F., Francesco, D.S., Hennig, J., and Seifritz, E. (2005). Cortical and subcortical correlates of electrocephalographic alpha rhythm modulation. *Journal of Neurophysiology*, 93:2864–72.

Feinberg, D.A., Moeller, S., Smith, S.M., Auerbach, E., Ramanna, S., Glasser, M.F., Miller, K.L., Ugurbil, K., and Yacoub, E. (2010). Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI and Fast Diffusion Imaging. *PLOS One*, 5(12):e15710.

Feis, R.A., Smith, S.M., Filippini, N., Douaud, G., Dopper, E.G.P., Heise, V., Trachtenberg, A.J., van Swieten, J.C., van Buchem, M.A., Rombouts, S.A.R.B., and Mackay, C.E. (2015). ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI. *Frontiers in Neuroscience*, 9:395.

Feis, R.A., Bouts, M.J.R.J., Panman, J.L., Jiskoot, L.C., Dopper, E.G.P., Schouten, T.M., de Vos, F., van der Grond, J., van Swieten, J.C., and Rombouts, S.A.R.B. (2019a). Single-subject classification of presymptomatic frontotemporal dementia mutation carriers using multimodal MRI. *NeuroImage: Clinical*, 22:101718.

Feis, R.A., Bouts, M.J.R.J., de Vos, F., Schouten, T.M., Panman, J.L., Jiskoot, L.C., Dopper, E.G.P., van der Grond, J., van Swieten, J.C., and Rombouts, S.A.R.B. (2019b). A multimodal MRI-based classification signature emerges just prior to symptom onset in frontotemporal dementia mutation carriers. *Journal of Neurology, Neurosurgery, and Psychiatry*, 90(11):1207–14.

Filippini, N., MacIntosh, B.J., Hough, M.G., Goodwin, G.M., Frisoni, G.B., Smith, S.M., Matthews, P.M., Beckmann, C.F., and Mackay, C.E. (2009). Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. *Proceedings of the National Academy of Sciences of the USA*, 106(17):7209–14.

Filippini, N., Ebmeier, K.P., MacIntosh, B.J., Trachtenberg, A.J., Frisoni, G.B., Wilcock, G.K., Beckmann, C.F., Smith, S.M., Matthews, P.M., and Mackay, C.E. (2011). Differential effects of the APOE genotype on brain function across the lifespan. *NeuroImage*, 54(1):602–10.

Folstein, M.F., Folstein, S.E., and McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–98.

Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M.A., Ruparel, K., Roalf, D.R., Satterthwaite, T.D., Gur, R.C., Gur, R.E., Schultz, R.T., Verma, R., and Shinohara, R.T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, 161:149–70.

Fortin, J.-P., Cullen, N., Sheline, Y.I., Taylor, W.D., Aselcioglu, I., Cook, P.A., Adams, P., Cooper, C., Fava, M., McGrath, P.J., McInnis, M., Phillips, M.L., Trivedi, M.H., Weissman, M.M., and Shinohara, R.T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 167:104–20.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., and Raichle, M.E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the USA*, 102(27):9673–8.

Fox, M.D., and Raichle, M.E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–11.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–41.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1):1–22.

Friedman, L., Stern, H., Brown, G.G., Mathalon, D.H., Turner, J., Glover, G.H., Gollub, R.L., Lauriello, J., Lim, K.O., Cannon, T., Greve, D.N., Bockholt, H.J., Belger, A., Mueller, B., Doty, M.J., He, J., Wells, W., Smyth, P., Pieper, S., Kim, S., Kubicki, M., Vangel, M., and Potkin, S.G. (2008). Test-retest and between-site reliability in a multicenter fMRI study. *Human Brain Mapping*, 29(8):958–72.

Frings, L., Yew, B., Flanagan, E., Lam, B.Y.K., Hüll, M., Huppertz, H.-J., Hodges, J.R., and Hornberger, M. (2014). Longitudinal Grey and White Matter Changes in Frontotemporal Dementia and Alzheimer's Disease. *PLOS One*, 9(3):e90814.

Friston, K.J., Ashburner, J., Kiebel, S., Nichols, T., and Penny, W.D. (2007). Statistical parametric mapping: the analysis of funtional brain images. London: Elsevier/Academic Press.

Galantucci, S., Tartaglia, M.C., Wilson, S.M., Henry, M.L., Filippi, M., Agosta, F., Dronkers, N.F., Henry, R.G., Ogar, J.M., Miller, B.L., and Gorno-Tempini, M.L. (2011). White matter damage in primary progressive aphasias: a diffusion tensor tractography study. *Brain*, 134(10):3011–29.

Galimberti, D., and Scarpini, E. (2012). Clinical phenotypes and genetic biomarkers of FTLD. *Journal of Neural Transmission*, 119(7):851–60.

Galimberti, D., Fumagalli, G.G., Fenoglio, C., Cioffi, S.M.G., Arighi, A., Serpente, M., Borroni, B., Padovani, A., Tagliavini, F., Masellis, M., Tartaglia, M.C., van Swieten, J.C., Meeter, L.H.H., Graff, C., de Mendonça, A., Bocchetta, M., Rohrer, J.D., Scarpini, E., Andersson, C., Archetti, S., Arighi, A., Benussi, L., Binetti, G., Black, S.E., Cash, D.M., Cosseddu, M., Dick, K.M., Fallström, M., Ferreira, C., Finger, E., Fox, N.C., Freedman, M., Frisoni, G.B.,

Gazzina, S., Ghidoni, R., Grisoli, M., Jelic, V., Jiskoot, L.C., Keren, R., Laforce, R., Lombardi, G., Maruta, C., Mead, S., van Minkelen, R., Nacmias, B., Öijerstedt, L., Ourselin, S., Panman, J.L., Pievani, M., Polito, C., Prioni, S., Rademakers, R., Redaelli, V., Rogaeva, E., Rossi, G., Besta, C., Rossor, M.N., Rowe, J.B., Sorbi, S., Tang-Wai, D.F., Thomas, D.L., Thonberg, H., Tiraboschi, P., Verdelho, A., and Warren, J.D. (2018). Progranulin plasma levels predict the presence of GRN mutations in asymptomatic subjects and do not correlate with brain atrophy: results from the GENFI study. *Neurobiology of Aging*, 62:245.e9-12.

Gaudio, S., Piervincenzi, C., Beomonte Zobel, B., Romana Montecchi, F., Riva, G., Carducci, F., and Quattrocchi, C.C. (2015). Altered resting state functional connectivity of anterior cingulate cortex in drug naïve adolescents at the earliest stages of anorexia nervosa. *Scientific Reports*, 5:10818.

Genin, E., Hannequin, D., Wallon, D., Sleegers, K., Hiltunen, M., Combarros, O., Bullido, M.J., Engelborghs, S., De Deyn, P., Berr, C., Pasquier, F., Dubois, B., Tognoni, G., Fiévet, N., Brouwers, N., Bettens, K., Arosio, B., Coto, E., Del Zompo, M., Mateo, I., Epelbaum, J., Frank-Garcia, A., Helisalmi, S., Porcellini, E., Pilotto, A., Forti, P., Ferri, R., Scarpini, E., Siciliano, G., Solfrizzi, V., Sorbi, S., Spalletta, G., Valdivieso, F., Vepsäläinen, S., Alvarez, V., Bosco, P., Mancuso, M., Panza, F., Nacmias, B., Bossù, P., Hanon, O., Piccardi, P., Annoni, G., Seripa, D., Galimberti, D., Licastro, F., Soininen, H., Dartigues, J.-F., Kamboh, M.I., Van Broeckhoven, C., Lambert, J.C., Amouyel, P., and Campion, D. (2011). APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Molecular Psychiatry*, 16(9):903–7.

Glover, G.H., Mueller, B.A., Turner, J.A., van Erp, T.G.M., Liu, T.T., Greve, D.N., Voyvodic, J.T., Rasmussen, J., Brown, G.G., Keator, D.B., Calhoun, V.D., Lee, H.J., Ford, J.M., Mathalon, D.H., Diaz, M., O'Leary, D.S., Gadde, S., Preda, A., Lim, K.O., Wible, C.G., Stern, H.S., Belger, A., McCarthy, G., Ozyurt, B., and Potkin, S.G. (2012). Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies. *Journal of Magnetic Resonance Imaging*, 36(1):39–54.

Gold, B.T., Powell, D.K., Andersen, A.H., and Smith, C.D. (2010). Alterations in multiple measures of white matter integrity in normal women at high risk for Alzheimer's disease. *NeuroImage*, 52(4):1487–94.

Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N.A., Friston, K.J., and Frackowiak, R.S.J. (2001). A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. *NeuroImage*, 14(1):21–36.

Gorgolewski, K.J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S.S., Maumet, C., Sochat, V. V., Nichols, T.E., Poldrack, R.A., Poline, J.-B., Yarkoni, T., and Margulies, D.S. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, 9:8.

Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S.F., Ogar, J.M., Rohrer, J.D., Black, S., Boeve, B.F., Manes, F., Dronkers, N.F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R., Mesulam, M.M., and Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11):1006–14.

Graham, A., Davies, R., Xuereb, J., Halliday, G., Kril, J., Creasey, H., Graham, K., and Hodges,

J. (2005). Pathologically proven frontotemporal dementia presenting with severe amnesia. *Brain*, 128(3):597–605.

Greicius, M.D., Krasnow, B., Reiss, A.L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the USA*, 100(1):253–8.

Greve, D.N., and Fischl, B. (2009). Accurate and Robust Brain Image Alignment using Boundary-based Registration. *NeuroImage*, 48(1):63–72.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., Moeller, S., Xu, J., Yacoub, E., Baselli, G., Ugurbil, K., Miller, K.L., and Smith, S.M. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*, 95:232–47.

Hafkemeijer, A., Möller, C., Dopper, E.G.P., Jiskoot, L.C., Schouten, T.M., van Swieten, J.C., van der Flier, W.M., Vrenken, H., Pijnenburg, Y.A.L., Barkhof, F., Scheltens, P., van der Grond, J., and Rombouts, S.A.R.B. (2015). Resting state functional connectivity differences between behavioral variant frontotemporal dementia and Alzheimer's disease. *Frontiers in Human Neuroscience*, 9:474.

Hafkemeijer, A., Möller, C., Dopper, E.G.P., Jiskoot, L.C., van den Berg-Huysmans, A.A., van Swieten, J.C., van der Flier, W.M., Vrenken, H., Pijnenburg, Y.A.L., Barkhof, F., Scheltens, P., van der Grond, J., and Rombouts, S.A.R.B. (2016). Differences in structural covariance brain networks between behavioral variant frontotemporal dementia and Alzheimer's disease. *Human Brain Mapping*, 37(3):978–88.

Handels, R.L., Aalten, P., Wolfs, C.A., OldeRikkert, M., Scheltens, P., Visser, P.J., Joore, M.A., Severens, J.L., and Verhey, F.R. (2012). Diagnostic and economic evaluation of new biomarkers for Alzheimer's disease: the research protocol of a prospective cohort study. *BMC Neurology*, 12(1):72.

Happé, F., Brownell, H., and Winner, E. (1999). Acquired "theory of mind" impairments following stroke. *Cognition*, 70(3):211–40.

Harvey, R.J., Skelton-Robinson, M., and Rossor, M.N. (2003). The prevalence and causes of dementia in people under the age of 65 years. *Journal of Neurology, Neurosurgery, and Psychiatry*, 74(9):1206–9.

Heise, V., Filippini, N., Ebmeier, K.P., and Mackay, C.E. (2011). The APOE ε4 allele modulates brain white matter integrity in healthy adults. *Molecular Psychiatry*, 16(9):908–16.

Hodges, J.R., Davies, R., Xuereb, J., Kril, J., and Halliday, G. (2003). Survival in frontotemporal dementia. *Neurology*, 61(3):349–54.

Hoerl, A.E., and Kennard, R.W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1):55–67.

Hogan, D.B., Jetté, N., Fiest, K.M., Roberts, J.I., Pearson, D., Smith, E.E., Roach, P., Kirk, A., Pringsheim, T., and Maxwell, C.J. (2016). The Prevalence and Incidence of Frontotemporal Dementia: a Systematic Review. *Canadian Journal of Neurological Sciences / Journal Canadien*

des Sciences Neurologiques, 43 Suppl 1:S96-109.

Honea, R.A., Vidoni, E., Harsha, A., and Burns, J.M. (2009). Impact of APOE on the Healthy Aging Brain: A Voxel-Based MRI and DTI Study. *Journal of Alzheimer's Disease*, 18:553–64.

Huey, E.D., Armstrong, N., Momeni, P., and Grafman, J. (2008). Challenges and new opportunities in the investigation of new drug therapies to treat frontotemporal dementia. *Expert Opinion on Therapeutic Targets*, 12(11):1367–76.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–34.

Jacova, C., Hsiung, G.-Y.R., Tawankanjanachot, I., Dinelle, K., McCormick, S., Gonzalez, M., Lee, H., Sengdy, P., Bouchard-Kerr, P., Baker, M., Rademakers, R., Sossi, V., Stoessl, A.J., Feldman, H.H., and Mackenzie, I.R. (2013). Anterior brain glucose hypometabolism predates dementia in progranulin mutation carriers. *Neurology*, 81(15):1322–31.

Jao, T., Vértes, P.E., Alexander-Bloch, A.F., Tang, I.-N., Yu, Y.-C., Chen, J.-H., and Bullmore, E.T. (2013). Volitional eyes opening perturbs brain dynamics and functional connectivity regardless of light input. *NeuroImage*, 69:21–34.

Jenkinson, M., and Smith, S.M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143–56.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S.M. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–41.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., and Smith, S.M. (2012). FSL. *NeuroImage*, 62(2):782–90.

Jiskoot, L.C., Dopper, E.G.P., den Heijer, T., Timman, R., van Minkelen, R., van Swieten, J.C., and Papma, J.M. (2016). Presymptomatic cognitive decline in familial frontotemporal dementia: A longitudinal study. *Neurology*, 87(4):384–91.

Jiskoot, L.C., Bocchetta, M., Nicholas, J.M., Cash, D.M., Thomas, D., Modat, M., Ourselin, S., Rombouts, S.A.R.B., Dopper, E.G.P., Meeter, L.H.H., Panman, J.L., van Minkelen, R., van der Ende, E.L., Donker Kaat, L., Pijnenburg, Y.A.L., Borroni, B., Galimberti, D., Masellis, M., Tartaglia, M.C., Rowe, J., Graff, C., Tagliavini, F., Frisoni, G.B., Laforce, R., Finger, E., de Mendonça, A., Sorbi, S., Papma, J.M., van Swieten, J.C., and Rohrer, J.D. (2018a). Presymptomatic white matter integrity loss in familial frontotemporal dementia in the GENFI cohort: A cross-sectional diffusion tensor imaging study. *Annals of Clinical and Translational Neurology*, 5(9):1025–36.

Jiskoot, L.C., Panman, J.L., van Asseldonk, L., Franzen, S., Meeter, L.H.H., Donker Kaat, L., van der Ende, E.L., Dopper, E.G.P., Timman, R., van Minkelen, R., van Swieten, J.C., van den Berg, E., and Papma, J.M. (2018b). Longitudinal cognitive biomarkers predicting symptom onset in presymptomatic frontotemporal dementia. *Journal of Neurology*, 265(6):1381–92.

Jiskoot, L.C., Panman, J.L., Meeter, L.H.H., Dopper, E.G.P., Donker Kaat, L., Franzen, S., van der Ende, E.L., van Minkelen, R., Rombouts, S.A.R.B., Papma, J.M., and van Swieten, J.C. (2019). Longitudinal multimodal MRI as prognostic and diagnostic biomarker in presymptomatic

familial frontotemporal dementia. *Brain*, 142(1):193–208.

Johnson, J.K., Head, E., Kim, R., Starr, A., and Cotman, C.W. (1999). Clinical and Pathological Evidence for a Frontal Variant of Alzheimer Disease. *Archives of Neurology*, 56(10):1233–9.

Johnson, J.K., Diehl, J., Mendez, M.F., Neuhaus, J., Shapira, J.S., Forman, M., Chute, D.J., Roberson, E.D., Pace-Savitsky, C., Neumann, M., Chow, T.W., Rosen, H.J., Forstl, H., Kurz, A., and Miller, B.L. (2005). Frontotemporal Lobar Degeneration. *Archives of Neurology*, 62(6):925–30.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, 8(1):118–27.

Jolles, J., Houx, P.J., van Boxtel, M.P.J., and Ponds, R. (1995). Maastricht Aging Study: determinants of cognitive aging. *Neuropsych Publishers, Maastricht*.

Josephs, K.A., Petersen, R.C., Knopman, D.S., Boeve, B.F., Whitwell, J.L., Duffy, J.R., Parisi, J.E., and Dickson, D.W. (2006). Clinicopathologic analysis of frontotemporal and corticobasal degenerations and PSP. *Neurology*, 66(1):41–8.

Josephs, K.A., Hodges, J.R., Snowden, J.S., Mackenzie, I.R., Neumann, M., Mann, D.M., and Dickson, D.W. (2011). Neuropathological background of phenotypical variability in frontotemporal dementia. *Acta Neuropathologica*, 122(2):137–53.

Kansal, K., Mareddy, M., Sloane, K.L., Minc, A.A., Rabins, P. V., McGready, J.B., and Onyike, C.U. (2016). Survival in Frontotemporal Dementia Phenotypes: A Meta-Analysis. *Dementia and Geriatric Cognitive Disorders*, 41(1–2):109–22.

Kaplan, E., Goodglass, H., and Weintraub, S. (1978). The Boston Naming Test. *Lea & Febiger, Philadelphia*.

Kazemi, K., and Noorizadeh, N. (2014). Quantitative Comparison of SPM, FSL, and Brainsuite for Brain MR Image Segmentation. *Journal of Biomedical Physics & Engineering*, 4(1):13–26.

Kertesz, A., McMonagle, P., Blair, M., Davidson, W., and Munoz, D.G. (2005). The evolution and pathology of frontotemporal dementia. *Brain*, 128(9):1996–2005.

Kim, J.P., Kim, J., Park, Y.H., Park, S.B., Lee, J.S., Yoo, S., Kim, E.-J., Kim, H.J., Na, D.L., Brown, J.A., Lockhart, S.N., Seo, S.W., and Seong, J.-K. (2019). Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. *NeuroImage: Clinical*, 23:101811.

Klöppel, S., Peter, J., Ludl, A., Pilatus, A., Maier, S., Mader, I., Heimbach, B., Frings, L., Egger, K., Dukart, J., Schroeter, M.L., Perneczky, R., Häussermann, P., Vach, W., Urbach, H., Teipel, S., Hüll, M., Abdulkadir, A., and Alzheimer's Disease Neuroimaging Initiative (2015). Applying Automated MR-Based Diagnostic Methods to the Memory Clinic: A Prospective Study. *Journal of Alzheimer's Disease*, 47(4):939–54.

Koikkalainen, J., Rhodius-Meester, H., Tolonen, A., Barkhof, F., Tijms, B., Lemstra, A.W., Tong, T., Guerrero, R., Schuh, A., Ledig, C., Rueckert, D., Soininen, H., Remes, A.M., Waldemar, G., Hasselbalch, S., Mecocci, P., van der Flier, W., and Lötjönen, J. (2016). Differential diagnosis of neurodegenerative diseases using structural MRI data. *NeuroImage: Clinical*, 11:435–49.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., and Baker, C.I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, 12(5):535–40.

Krudop, W.A., Kerssens, C.J., Dols, A., Prins, N.D., Möller, C., Schouws, S., van der Flier, W.M., Scheltens, P., Sikkes, S., Stek, M.L., and Pijnenburg, Y.A.L. (2015). Identifying bvFTD Within the Wide Spectrum of Late Onset Frontal Lobe Syndrome: A Clinical Approach. *The American Journal of Geriatric Psychiatry*, 23(10):1056–66.

Laird, N.M., and Ware, J.H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics*, 38(4):963–74.

Laukka, E.J., Lövdén, M., Kalpouzos, G., Papenberg, G., Keller, L., Graff, C., Li, T.-Q., Fratiglioni, L., and Bäckman, L. (2015). Microstructural White Matter Properties Mediate the Association between APOE and Perceptual Speed in Very Old Persons without Dementia. *PLOS One*, 10(8):e0134766.

Lee, S.E., Khazenzon, A.M., Trujillo, A.J., Guo, C.C., Yokoyama, J.S., Sha, S.J., Takada, L.T., Karydas, A.M., Block, N.R., Coppola, G., Pribadi, M., Geschwind, D.H., Rademakers, R., Fong, J.C., Weiner, M.W., Boxer, A.L., Kramer, J.H., Rosen, H.J., Miller, B.L., and Seeley, W.W. (2014). Altered network connectivity in frontotemporal dementia with C9orf72 hexanucleotide repeat expansion. *Brain*, 137(11):3047–60.

Lee, S.E., Sias, A.C., Mandelli, M.L., Brown, J.A., Brown, A.B., Khazenzon, A.M., Vidovszky, A.A., Zanto, T.P., Karydas, A.M., Pribadi, M., Dokuru, D., Coppola, G., Geschwind, D.H., Rademakers, R., Gorno-Tempini, M.L., Rosen, H.J., Miller, B.L., and Seeley, W.W. (2017). Network degeneration and dysfunction in presymptomatic C9orf72 expansion carriers. *NeuroImage: Clinical*, 14:286–97.

Leemans, A., and Jones, D.K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. *Magnetic Resonance in Medicine*, 61(6):1336–49.

Lehmer, C., Oeckl, P., Weishaupt, J.H., Volk, A.E., Diehl-Schmid, J., Schroeter, M.L., Lauer, M., Kornhuber, J., Levin, J., Fassbender, K., Landwehrmeyer, B., Schludi, M.H., Arzberger, T., Kremmer, E., Flatley, A., Feederle, R., Steinacker, P., Weydt, P., Ludolph, A.C., Edbauer, D., and Otto, M. (2017). Poly-GP in cerebrospinal fluid links C9orf72-associated dipeptide repeat expression to the asymptomatic phase of ALS/FTD. *EMBO Molecular Medicine*, 9(7):859–68.

Liang, B., Zhang, D., Wen, X., Xu, P., Peng, X., Huang, X., Liu, M., and Huang, R. (2014). Brain spontaneous fluctuations in sensorimotor regions were directly related to eyes open and eyes closed: evidences from a machine learning approach. *Frontiers in Human Neuroscience*, 8:645.

Lindeboom, J., Schmand, B., Tulner, L., Walstra, G., and Jonker, C. (2002). Visual association test to detect early dementia of the Alzheimer type. *Journal of Neurology, Neurosurgery, and Psychiatry*, 73(2):126–33.

Littow, H. (2010). Age-related differences in functional nodes of the brain cortex – a high model order group ICA study. *Frontiers in Systems Neuroscience*, 4:32.

Liu, D.-Q., Dong, Z., Zuo, X.-N., Wang, J., and Zang, Y.-F. (2013). Eyes-open/eyes-closed dataset sharing for reproducibility evaluation of resting state fMRI data analysis methods.

Neuroinformatics, 11(4):469–76.

Lobo, A., Launer, L.J., Fratiglioni, L., Andersen, K., Di Carlo, A., Breteler, M.M.B., Copeland, J.R.M., Dartigues, J.-F., Jagger, C., Martinez-Lage, J., Soininen, H., Hofman, A., and Group for the Neurologic Diseases in the Elderly Research (2000). Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. *Neurology*, 54(11 Suppl 5):S4-9.

Logroscino, G., Imbimbo, B.P., Lozupone, M., Sardone, R., Capozzo, R., Battista, P., Zecca, C., Dibello, V., Giannelli, G., Bellomo, A., Greco, A., Daniele, A., Seripa, D., and Panza, F. (2019). Promising therapies for the treatment of frontotemporal dementia clinical phenotypes: from symptomatic to disease-modifying drugs. *Expert Opinion on Pharmacotherapy*, 20(9):1091–107.

Lomen-Hoerth, C., Anderson, T., and Miller, B. (2002). The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology*, 59(7):1077–9.

Long, X.-Y., Zuo, X.-N., Kiviniemi, V.J., Yang, Y., Zou, Q.-H., Zhu, C.-Z., Jiang, T.-Z., Yang, H., Gong, Q.-Y., Wang, L., Li, K.-C., Xie, S., and Zang, Y.-F. (2008). Default mode network as revealed with multiple methods for resting-state functional MRI analysis. *Journal of Neuroscience Methods*, 171(2):349–55.

Ludolph, A., Drory, V., Hardiman, O., Nakano, I., Ravits, J., Robberecht, W., Shefner, J., and The WFN Research Group On ALS/MND (2015). A revision of the El Escorial criteria – 2015. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 16(5–6):291–2.

Lyall, D.M., Harris, S.E., Bastin, M.E., Muñoz Maniega, S., Murray, C., Lutz, M.W., Saunders, A.M., Roses, A.D., del Carmen Valdés Hernández, M., Royle, N.A., Starr, J.M., Porteous, D.J., Wardlaw, J.M., and Deary, I.J. (2014). Alzheimer's disease susceptibility genes APOE and TOMM40, and brain white matter integrity in the Lothian Birth Cohort 1936. *Neurobiology of Aging*, 35(6):1513.e25-33.

Machulda, M.M., Jones, D.T., Vemuri, P., McDade, E., Avula, R., Przybelski, S., Boeve, B.F., Knopman, D.S., Petersen, R.C., and Jack, C.R. (2011). Effect of APOE ε4 Status on Intrinsic Network Connectivity in Cognitively Normal Elderly Subjects. *Archives of Neurology*, 68(9):1131–6.

Mahoney, C.J., Ridgway, G.R., Malone, I.B., Downey, L.E., Beck, J., Kinnunen, K.M., Schmitz, N., Golden, H.L., Rohrer, J.D., Schott, J.M., Rossor, M.N., Ourselin, S., Mead, S., Fox, N.C., and Warren, J.D. (2014). Profiles of white matter tract pathology in frontotemporal dementia. *Human Brain Mapping*, 35(8):4163–79.

Mann, D.M.A., and Snowden, J.S. (2017). Frontotemporal lobar degeneration: Pathogenesis, pathology and pathways to phenotype. *Brain Pathology*, 27(6):723–36.

Matura, S., Prvulovic, D., Jurcoane, A., Hartmann, D., Miller, J., Scheibe, M., O'Dwyer, L., Oertel-Knöchel, V., Knöchel, C., Reinke, B., Karakaya, T., Fußer, F., and Pantel, J. (2014). Differential effects of the ApoE4 genotype on brain structure and function. *NeuroImage*, 89:81–91.

McAvoy, M., Larson-Prior, L., Nolan, T.S., Vaishnavi, S.N., Raichle, M.E., and D'Avossa, G. (2008). Resting states affect spontaneous BOLD oscillations in sensory and paralimbic cortex.

Journal of Neurophysiology, 100(2):922–31.

McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C., Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., and Phelps, C.H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3):263–9.

McMillan, C.T., Brun, C., Siddiqui, S., Churgin, M., Libon, D., Yushkevich, P., Zhang, H., Boller, A., Gee, J., and Grossman, M. (2012). White matter imaging contributes to the multimodal diagnosis of frontotemporal lobar degeneration. *Neurology*, 78(22):1761–8.

McMillan, C.T., Avants, B.B., Cook, P., Ungar, L., Trojanowski, J.Q., and Grossman, M. (2014). The power of neuroimaging biomarkers for screening frontotemporal dementia. *Human Brain Mapping*, 35(9):4827–40.

Meeter, L.H.H., Dopper, E.G.P., Jiskoot, L.C., Sanchez-Valle, R., Graff, C., Benussi, L., Ghidoni, R., Pijnenburg, Y.A.L., Borroni, B., Galimberti, D., Laforce, R., Masellis, M., Vandenberghe, R., Le Ber, I., Otto, M., van Minkelen, R., Papma, J.M., Rombouts, S.A.R.B., Balasa, M., Öijerstedt, L., Jelic, V., Dick, K.M., Cash, D.M., Harding, S.R., Jorge Cardoso, M., Ourselin, S., Rossor, M.N., Padovani, A., Scarpini, E., Fenoglio, C., Tartaglia, M.C., Lamari, F., Barro, C., Kuhle, J., Rohrer, J.D., Teunissen, C.E., and van Swieten, J.C. (2016a). Neurofilament light chain: a biomarker for genetic frontotemporal dementia. *Annals of Clinical and Translational Neurology*, 3(8):623–36.

Meeter, L.H.H., Patzke, H., Loewen, G., Dopper, E.G.P., Pijnenburg, Y.A.L., van Minkelen, R., and van Swieten, J.C. (2016b). Progranulin Levels in Plasma and Cerebrospinal Fluid in Granulin Mutation Carriers. *Dementia and Geriatric Cognitive Disorders Extra*, 6(2):330–40.

Meeter, L.H.H., Gendron, T.F., Sias, A.C., Jiskoot, L.C., Russo, S.P., Donker Kaat, L., Papma, J.M., Panman, J.L., van der Ende, E.L., Dopper, E.G.P., Franzen, S., Graff, C., Boxer, A.L., Rosen, H.J., Sanchez-Valle, R., Galimberti, D., Pijnenburg, Y.A.L., Benussi, L., Ghidoni, R., Borroni, B., Laforce, R., Del Campo, M., Teunissen, C.E., van Minkelen, R., Rojas, J.C., Coppola, G., Geschwind, D.H., Rademakers, R., Karydas, A.M., Öijerstedt, L., Scarpini, E., Binetti, G., Padovani, A., Cash, D.M., Dick, K.M., Bocchetta, M., Miller, B.L., Rohrer, J.D., Petrucelli, L., van Swieten, J.C., and Lee, S.E. (2018a). Poly(GP), neurofilament and grey matter deficits in C9orf72 expansion carriers. *Annals of Clinical and Translational Neurology*, 5(5):583–97.

Meeter, L.H.H., Vijverberg, E.G., Del Campo, M., Rozemuller, A.J.M., Donker Kaat, L., de Jong, F.J., van der Flier, W.M., Teunissen, C.E., van Swieten, J.C., and Pijnenburg, Y.A.L. (2018b). Clinical value of neurofilament and phospho-tau/tau ratio in the frontotemporal dementia spectrum. *Neurology*, 90(14):e1231-9.

Mendez, M.F., Shapira, J.S., McMurtray, A., Licht, E., Miller, B.L., and Trojanowski, J.Q. (2007). Accuracy of the Clinical Evaluation for Frontotemporal Dementia. *Archives of Neurology*, 64(6):830–5.

Mendez, M.F. (2009). 'Frontotemporal Dementia: Therapeutic Interventions', in: Giannakopoulos, P., and Hof, P.R. (Eds.), Dementia in Clinical Practice. KARGER, Basel, pp. 168–78.

Meyer, S., Mueller, K., Stuke, K., Bisenius, S., Diehl-Schmid, J., Jessen, F., Kassubek, J., Kornhuber, J., Ludolph, A.C., Prudlo, J., Schneider, A., Schuemberg, K., Yakushev, I., Otto, M., and Schroeter, M.L. (2017). Predicting behavioral variant frontotemporal dementia with pattern classification in multi-center structural MRI data. *NeuroImage: Clinical*, 14:656–62.

Mito, R., Raffelt, D., Dhollander, T., Vaughan, D.N., Tournier, J.-D., Salvado, O., Brodtmann, A., Rowe, C.C., Villemagne, V.L., and Connelly, A. (2018). Fibre-specific white matter reductions in Alzheimer's disease and mild cognitive impairment. *Brain*, 141(3):888–902.

Miyoshi, M., Shinotoh, H., Wszolek, Z.K., Strongosky, A.J., Shimada, H., Arakawa, R., Higuchi, M., Ikoma, Y., Yasuno, F., Fukushi, K., Irie, T., Ito, H., and Suhara, T. (2010). In vivo detection of neuropathologic changes in presymptomatic MAPT mutation carriers: A PET and MRI study. *Parkinsonism & Related Disorders*, 16(6):404–8.

Mohs, R.C., Doody, R.S., Morris, J.C., Ieni, J.R., Rogers, S.L., Perdomo, C.A., Pratt, R.D., and "312" Study Group (2001). A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. *Neurology*, 57(3):481–8.

Möller, C., Dieleman, N., van der Flier, W.M., Versteeg, A., Pijnenburg, Y.A.L., Scheltens, P., Barkhof, F., and Vrenken, H. (2015a). More Atrophy of Deep Gray Matter Structures in Frontotemporal Dementia Compared to Alzheimer's Disease. *Journal of Alzheimer's Disease*, 44(2):635–47.

Möller, C., Hafkemeijer, A., Pijnenburg, Y.A.L., Rombouts, S.A.R.B., van der Grond, J., Dopper, E.G.P., van Swieten, J.C., Versteeg, A., Pouwels, P.J.W., Barkhof, F., Scheltens, P., Vrenken, H., and van Der Flier, W.M. (2015b). Joint assessment of white matter integrity, cortical and subcortical atrophy to distinguish AD from behavioral variant FTD: A two-center study. *NeuroImage: Clinical*, 9:418–29.

Möller, C., Pijnenburg, Y.A.L., van der Flier, W.M., Versteeg, A., Tijms, B., de Munck, J.C., Hafkemeijer, A., Rombouts, S.A.R.B., van der Grond, J., van Swieten, J.C., Dopper, E.G.P., Scheltens, P., Barkhof, F., Vrenken, H., and Wink, A.M. (2016). Alzheimer Disease and Behavioral Variant Frontotemporal Dementia: Automatic Classification Based on Cortical Atrophy for Single-Subject Diagnosis. *Radiology*, 279(3):838–48.

Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P.H., Albert, M., Boone, K., Miller, B.L., Cummings, J., and Benson, D.F. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, 51(6):1546–54.

Nelson, H.E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, 12(4):313–24.

Nierenberg, J., Pomara, N., Hoptman, M.J., Sidtis, J.J., Ardekani, B.A., and Lim, K.O. (2005). Abnormal white matter integrity in healthy apolipoprotein E epsilon4 carriers. *Neuroreport*, 16(12):1369–72.

Noirhomme, Q., Lesenfants, D., Gomez, F., Soddu, A., Schrouff, J., Garraux, G., Luxen, A., Phillips, C., and Laureys, S. (2014). Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. *NeuroImage: Clinical*, 4:687–94.

Omer, T., Finegan, E., Hutchinson, S., Doherty, M., Vajda, A., McLaughlin, R.L., Pender, N., Hardiman, O., and Bede, P. (2017). Neuroimaging patterns along the ALS-FTD spectrum: a multiparametric imaging study. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 18(7–8):611–23.

Operto, G., Cacciaglia, R., Grau-Rivera, O., Falcon, C., Brugulat-Serrat, A., Ródenas, P., Ramos, R., Morán, S., Esteller, M., Bargalló, N., Molinuevo, J.L., and Gispert, J.D. (2018). White matter microstructure is altered in cognitively normal middle-aged APOE-ε4 homozygotes. *Alzheimer's Research & Therapy*, 10(1):48.

Pan, P.L., Song, W., Yang, J., Huang, R., Chen, K., Gong, Q.-Y., Zhong, J.G., Shi, H.C., and Shang, H.F. (2012). Gray matter atrophy in behavioral variant frontotemporal dementia: a meta-analysis of voxel-based morphometry studies. *Dementia and Geriatric Cognitive Disorders*, 33(2–3):141–8.

Panman, J.L., Jiskoot, L.C., Bouts, M.J.R.J., Meeter, L.H.H., van der Ende, E.L., Poos, J.M., Feis, R.A., Kievit, A.J.A., van Minkelen, R., Dopper, E.G.P., Rombouts, S.A.R.B., van Swieten, J.C., and Papma, J.M. (2019). Gray and white matter changes in presymptomatic genetic frontotemporal dementia: a longitudinal MRI study. *Neurobiology of Aging*, 76:115–24.

Papma, J.M., Jiskoot, L.C., Panman, J.L., Dopper, E.G.P., den Heijer, T., Donker Kaat, L., Pijnenburg, Y.A.L., Meeter, L.H.H., van Minkelen, R., Rombouts, S.A.R.B., and van Swieten, J.C. (2017). Cognition and gray and white matter characteristics of presymptomatic C9orf72 repeat expansion. *Neurology*, 89(12):1256–64.

Pasquier, F., Richard, F., and Lebert, F. (2004). Natural History of Frontotemporal Dementia: Comparison with Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, 17:253–7.

Patenaude, B., Smith, S.M., Kennedy, D.N., and Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage*, 56(3):907–22.

Patriat, R., Molloy, E.K., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., and Birn, R.M. (2013). The effect of resting condition on resting-state fMRI reliability and consistency: A comparison between resting with eyes open, closed, and fixated. *NeuroImage*, 78:463–73.

Persson, J., Lind, J., Larsson, A., Ingvar, M., Cruts, M., Van Broeckhoven, C., Adolfsson, R., Nilsson, L.-G., and Nyberg, L. (2006). Altered brain white matter integrity in healthy carriers of the APOE epsilon4 allele: a risk for AD? *Neurology*, 66(7):1029–33.

Pick, A. (1892). Über die Beziehungen der senilen Hirnatrophie zur Aphasie. *Prager Medicinische Wochenschrift*, 17:165–7.

Pievani, M., Paternicò, D., Benussi, L., Binetti, G., Orlandini, A., Cobelli, M., Magnaldi, S., Ghidoni, R., and Frisoni, G.B. (2014). Pattern of structural and functional brain abnormalities in asymptomatic granulin mutation carriers. *Alzheimer's & Dementia*, 10(5 Suppl):S354-63.

Pijnenburg, Y.A.L., Gillissen, F., Jonker, C., and Scheltens, P. (2004). Initial Complaints in Frontotemporal Lobar Degeneration. *Dementia and Geriatric Cognitive Disorders*, 17(4):302–6.

Pinheiro, J., Bates, D., DebRoy S, Sarkar D, and R Core Team (2018). nlme: Linear and Nonlinear Mixed Effects Models [R package nlme version 3.1-137].

Plassman, B.L., Langa, K.M., Fisher, G.G., Heeringa, S.G., Weir, D.R., Ofstedal, M.B., Burke, J.R., Hurd, M.D., Potter, G.G., Rodgers, W.L., Steffens, D.C., Willis, R.J., and Wallace, R.B. (2007). Prevalence of dementia in the United States: The aging, demographics, and memory study. *Neuroepidemiology*, 29(1–2):125–32.

Premi, E., Cauda, F., Gasparotti, R., Diano, M., Archetti, S., Padovani, A., and Borroni, B. (2014). Multimodal fMRI Resting-State Functional Connectivity in Granulin Mutations: The Case of Fronto-Parietal Dementia. *PLOS One*, 9(9):e106500.

Pressman, P.S., and Miller, B.L. (2014). Diagnosis and Management of Behavioral Variant Frontotemporal Dementia. *Biological Psychiatry*, 75(7):574–81.

Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., and Beckmann, C.F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, 112:267–77.

Raamana, P.R., Rosen, H.J., Miller, B., Weiner, M.W., Wang, L., and Beg, M.F. (2014). Three-Class Differential Diagnosis among Alzheimer Disease, Frontotemporal Dementia, and Controls. *Frontiers in Neurology*, 5:71.

Rabinovici, G.D., Seeley, W.W., Kim, E.J., Gorno-Tempini, M.L., Rascovsky, K., Pagliaro, T.A., Allison, S.C., Halabi, C., Kramer, J.H., Johnson, J.K., Weiner, M.W., Forman, M.S., Trojanowski, J.Q., DeArmond, S.J., Miller, B.L., and Rosen, H.J. (2008). Distinct MRI atrophy patterns in autopsy-proven Alzheimer's disease and frontotemporal lobar degeneration. *American Journal of Alzheimer's Disease & Other Dementias*, 22(6):474–88.

Rabinovici, G.D., and Miller, B.L. (2010). Frontotemporal Lobar Degeneration. *CNS Drugs*, 24(5):375–98.

Raffelt, D.A., Smith, R.E., Ridgway, G.R., Tournier, J.-D., Vaughan, D.N., Rose, S., Henderson, R., and Connelly, A. (2015). Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. *NeuroImage*, 117:40–55.

Raichle, M.E. (2015). The Brain's Default Mode Network. *Annual Review of Neuroscience*, 38(1):433–47.

Rascovsky, K., Hodges, J.R., Knopman, D., Mendez, M.F., Kramer, J.H., Neuhaus, J., van Swieten, J.C., Seelaar, H., Dopper, E.G.P., Onyike, C.U., Hillis, A.E., Josephs, K.A., Boeve, B.F., Kertesz, A., Seeley, W.W., Rankin, K.P., Johnson, J.K., Gorno-Tempini, M.L., Rosen, H.J., Prioleau-Latham, C.E., Lee, A., Kipps, C.M., Lillo, P., Piguet, O., Rohrer, J.D., Rossor, M.N., Warren, J.D., Fox, N.C., Galasko, D., Salmon, D.P., Black, S.E., Mesulam, M., Weintraub, S., Dickerson, B.C., Diehl-Schmid, J., Pasquier, F., Deramecourt, V., Lebert, F., Pijnenburg, Y.A.L., Chow, T.W., Manes, F., Grafman, J., Cappa, S.F., Freedman, M., Grossman, M., and Miller, B.L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9):2456–77.

Ratnavalli, E., Brayne, C., Dawson, K., and Hodges, J.R. (2002). The prevalence of frontotemporal dementia. *Neurology*, 58(11):1615–21.

Rey, A. (1958). L'examen clinique en psychologie. *Presses Universitaires de France, Paris*.

Riedijk, S.R., Niermeijer, M.F.N., Dooijes, D., and Tibben, A. (2009). A Decade of Genetic Counseling in Frontotemporal Dementia Affected Families: Few Counseling Requests and much Familial Opposition to Testing. *Journal of Genetic Counseling*, 18(4):350–6.

Risacher, S., and Saykin, A. (2013). Neuroimaging Biomarkers of Neurodegenerative Diseases and Dementia. *Seminars in Neurology*, 33(4):386–416.

Roberson, E.D., Hesse, J.H., Rose, K.D., Slama, H., Johnson, J.K., Yaffe, K., Forman, M.S., Miller, C.A., Trojanowski, J.Q., Kramer, J.H., and Miller, B.L. (2005). Frontotemporal dementia progresses to death faster than Alzheimer disease. *Neurology*, 65(5):719–25.

Rohrer, J.D., Warren, J.D., Barnes, J., Mead, S., Beck, J., Pepple, T., Boyes, R., Omar, R., Collinge, J., Stevens, J.M., Warrington, E.K., Rossor, M.N., and Fox, N.C. (2008). Mapping the progression of progranulin-associated frontotemporal lobar degeneration. *Nature Clinical Practice Neurology*, 4(8):455–60.

Rohrer, J.D., Guerreiro, R., Vandrovcova, J., Uphill, J., Reiman, D., Beck, J., Isaacs, A.M., Authier, A., Ferrari, R., Fox, N.C., Mackenzie, I.R.A., Warren, J.D., de Silva, R., Holton, J., Revesz, T., Hardy, J., Mead, S., and Rossor, M.N. (2009). The heritability and genetics of frontotemporal lobar degeneration. *Neurology*, 73(18):1451–6.

Rohrer, J.D., and Rosen, H.J. (2013). Neuroimaging in frontotemporal dementia. *International Review of Psychiatry*, 25(2):221–9.

Rohrer, J.D., Warren, J.D., Fox, N.C., and Rossor, M.N. (2013). Presymptomatic studies in genetic frontotemporal dementia. *Revue Neurologique*, 169(10):820–4.

Rohrer, J.D., Nicholas, J.M., Cash, D.M., van Swieten, J.C., Dopper, E.G.P., Jiskoot, L.C., van Minkelen, R., Rombouts, S.A.R.B., Cardoso, M.J., Clegg, S., Espak, M., Mead, S., Thomas, D.L., De Vita, E., Masellis, M., Black, S.E., Freedman, M., Keren, R., MacIntosh, B.J., Rogaeva, E., Tang-Wai, D., Tartaglia, M.C., Laforce, R., Tagliavini, F., Tiraboschi, P., Redaelli, V., Prioni, S., Grisoli, M., Borroni, B., Padovani, A., Galimberti, D., Scarpini, E., Arighi, A., Fumagalli, G., Rowe, J.B., Coyle-Gilchrist, I., Graff, C., Fallström, M., Jelic, V., Ståhlbom, A.K., Andersson, C., Thonberg, H., Lilius, L., Frisoni, G.B., Binetti, G., Pievani, M., Bocchetta, M., Benussi, L., Ghidoni, R., Finger, E., Sorbi, S., Nacmias, B., Lombardi, G., Polito, C., Warren, J.D., Ourselin, S., Fox, N.C., and Rossor, M.N. (2015). Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the Genetic Frontotemporal dementia Initiative (GENFI) study: a cross-sectional analysis. *The Lancet Neurology*, 14(3):253–62.

Royall, D.R., Cordes, J.A., and Polk, M. (1998). CLOX: an executive clock drawing task. *Journal of Neurology, Neurosurgery, and Psychiatry*, 64(5):588–94.

Rytty, R., Nikkinen, J., Paavola, L., Abou Elseoud, A., Moilanen, V., Visuri, A., Tervonen, O., Renton, A.E., Traynor, B.J., Kiviniemi, V., and Remes, A.M. (2013). GroupICA dual regression analysis of resting state networks in a behavioral variant of frontotemporal dementia. *Frontiers in Human Neuroscience*, 7:461.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., and Smith, S.M.

(2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*, 90:449–68.

Schouten, T.M., Koini, M., de Vos, F., Seiler, S., van der Grond, J., Lechner, A., Hafkemeijer, A., Möller, C., Schmidt, R., de Rooij, M., and Rombouts, S.A.R.B. (2016). Combining anatomical, diffusion, and resting state functional magnetic resonance imaging for individual classification of mild and moderate Alzheimer's disease. *NeuroImage: Clinical*, 11:46–51.

Schouten, T.M., Koini, M., de Vos, F., Seiler, S., de Rooij, M., Lechner, A., Schmidt, R., van den Heuvel, M., van der Grond, J., and Rombouts, S.A.R.B. (2017). Individual classification of Alzheimer's disease with diffusion magnetic resonance imaging. *NeuroImage*, 152:476–81.

Seelaar, H., Kamphorst, W., Rosso, S.M., Azmani, A., Masdjedi, R., de Koning, I., Maat-Kievit, J.A., Anar, B., Donker Kaat, L., Breedveld, G.J., Dooijes, D., Rozemuller, J.M., Bronner, I.F., Rizzu, P., and van Swieten, J.C. (2008). Distinct genetic forms of frontotemporal dementia. *Neurology*, 71(16):1220–6.

Seelaar, H., Rohrer, J.D., Pijnenburg, Y.A.L., Fox, N.C., and van Swieten, J.C. (2011). Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. *Journal of Neurology, Neurosurgery, and Psychiatry*, 82(5):476–86.

Seeley, W.W., Allman, J.M., Carlin, D.A., Crawford, R.K., Macedo, M.N., Greicius, M.D., DeArmond, S.J., and Miller, B.L. (2007). Divergent social functioning in behavioral variant frontotemporal dementia and Alzheimer disease: reciprocal networks and neuronal evolution. *Alzheimer Disease & Associated Disorders*, 21(4):S50-7.

Seeley, W.W., Crawford, R.K., Zhou, J., Miller, B.L., and Greicius, M.D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron*, 62(1):42–52.

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3):289–310.

Smith, C.D., Chebrolu, H., Andersen, A.H., Powell, D.A., Lovell, M.A., Xiong, S., and Gold, B.T. (2010). White matter diffusion alterations in normal women at risk of Alzheimer's disease. *Neurobiology of Aging*, 31(7):1122–31.

Smith, S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–55.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., and Matthews, P.M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1:S208-19.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., and Behrens, T.E.J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31(4):1487–505.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., and Beckmann, C.F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of*

Sciences of the USA, 106(31):13040–5.

Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, 44(1):83–98.

Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., Kelly, M., Laumann, T., Miller, K.L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A.Z., Vu, A.T., Woolrich, M.W., Xu, J., Yacoub, E., Uğurbil, K., Van Essen, D.C., and Glasser, M.F. (2013). Resting-state fMRI in the Human Connectome Project. *NeuroImage*, 80:144–68.

Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D.A., Yacoub, E., Lenglet, C., Van Essen, D.C., Ugurbil, K., and Behrens, T.E.J. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*, 80:125–43.

Spinelli, E.G., Mandelli, M.L., Miller, Z.A., Santos-Santos, M.A., Wilson, S.M., Agosta, F., Grinberg, L.T., Huang, E.J., Trojanowski, J.Q., Meyer, M., Henry, M.L., Comi, G., Rabinovici, G., Rosen, H.J., Filippi, M., Miller, B.L., Seeley, W.W., and Gorno-Tempini, M.L. (2017). Typical and atypical pathology in primary progressive aphasia variants. *Annals of Neurology*, 81(3):430–43.

Staffaroni, A.M., Bajorek, L., Casaletto, K.B., Cobigo, Y., Goh, S.-Y.M., Wolf, A., Heuer, H.W., Elahi, F.M., Ljubenkov, P.A., Dever, R., Kornak, J., Appleby, B., Bove, J., Bordelon, Y., Brannelly, P., Brushaber, D., Caso, C., Coppola, G., Dheel, C., Dickerson, B.C., Dickinson, S., Dominguez, S., Domoto-Reilly, K., Faber, K., Ferrall, J., Fields, J.A., Fishman, A., Fong, J., Foroud, T., Forsberg, L.K., Gavrilova, R., Gearhart, D., Ghazanfari, B., Ghoshal, N., Goldman, J., Graff-Radford, J., Graff-Radford, N., Grant, I., Grossman, M., Haley, D., Hsiung, G.-Y., Huey, E.D., Irwin, D.J., Jones, D.T., Jones, L., Kantarci, K., Karydas, A., Kaufer, D.I., Kerwin, D.R., Knopman, D.S., Kraft, R., Kremers, W.K., Kukull, W.A., Litvan, I., Lucente, D., Lungu, C., Mackenzie, I.R., Maldonado, M., Manoochehri, M., McGinnis, S.M., McKinley, E., Mendez, M.F., Miller, B.L., Multani, N., Onyike, C.U., Padmanabhan, J., Pantelyat, A., Pearlman, R., Petrucelli, L., Potter, M., Rademakers, R., Ramos, E.M., Rankin, K.P., Rascovsky, K., Roberson, E.D., Rogalski, E., Sengdy, P., Shaw, L.M., Syrjanen, J., Tartaglia, M.C., Tatton, N., Taylor, J., Toga, A., Trojanowski, J.Q., Weintraub, S., Wang, P., Wong, B., Wszolek, Z., Boxer, A.L., Boeve, B.F., Kramer, J.H., and Rosen, H.J. (2019). Assessment of executive function declines in presymptomatic and mildly symptomatic familial frontotemporal dementia: NIH-EXAMINER as a potential clinical trial endpoint. *Alzheimer's & Dementia*, 16(1):11–21.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18:643–62.

Sullivan, L.M., Dukes, K.A., and Losina, E. (1999). An introduction to hierarchical linear modelling. *Statistics in Medicine*, 18(7):855–88.

Suri, S., Topiwala, A., Mackay, C.E., Ebmeier, K.P., and Filippini, N. (2014). Using Structural and Diffusion Magnetic Resonance Imaging To Differentiate the Dementias. *Current Neurology and Neuroscience Reports*, 14(9):475.

ten Kate, M., Sanz-Arigita, E.J., Tijms, B.M., Wink, A.M., Clerigue, M., Garcia-Sebastian, M., Izagirre, A., Ecay-Torres, M., Estanga, A., Villanua, J., Vrenken, H., Visser, P.J., Martinez-Lage, P., and Barkhof, F. (2016). Impact of APOE-ε4 and family history of dementia on gray matter atrophy in cognitively healthy middle-aged adults. *Neurobiology of Aging*, 38:14–20.

The ADHD-200 Consortium (2012). The ADHD-200 Consortium: A Model to Advance the Translational Potential of Neuroimaging in Clinical Neuroscience. *Frontiers in Systems Neuroscience*, 6:62.

Thurstone, L.L.T., and Thurstone, T.G. (1962). Primary mental abilities. *Science Research Associates, Chicago*.

Tian, L., Kong, Y., Ren, J., Varoquaux, G., Zang, Y.-F., and Smith, S.M. (2013). Spatial vs. Temporal Features in ICA of Resting-State fMRI – A Quantitative and Qualitative Investigation in the Context of Response Inhibition. *PLOS One*, 8(6):e66572.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, 58(1):267–88.

Tournier, J.-D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–72.

Trachtenberg, A.J., Filippini, N., Cheeseman, J., Duff, E.P., Neville, M.J., Ebmeier, K.P., Karpe, F., and Mackay, C.E. (2012a). The effects of APOE on brain activity do not simply reflect the risk of Alzheimer's disease. *Neurobiology of Aging*, 33(3):618.e1–13.

Trachtenberg, A.J., Filippini, N., Ebmeier, K.P., Smith, S.M., Karpe, F., and Mackay, C.E. (2012b). The effects of APOE on the functional architecture of the resting brain. *NeuroImage*, 59(1):565–72.

Tsai, R.M., and Boxer, A.L. (2016). Therapy and clinical trials in frontotemporal dementia: past, present, and future. *Journal of Neurochemistry*, 138:211–21.

Tu, S., Leyton, C.E., Hodges, J.R., Piguet, O., and Hornberger, M. (2015). Divergent Longitudinal Propagation of White Matter Degradation in Logopenic and Semantic Variants of Primary Progressive Aphasia. *Journal of Alzheimer's Disease*, 49(3):853–61.

Tuovinen, T., Rytty, R., Moilanen, V., Abou Elseoud, A., Veijola, J., Remes, A.M., and Kiviniemi, V.J. (2017). The Effect of Gray Matter ICA and Coefficient of Variation Mapping of BOLD Data on the Detection of Functional Connectivity Changes in Alzheimer's Disease and bvFTD. *Frontiers in Human Neuroscience*, 10:680.

Tustison, N.J., Avants, B.B., Cook, P.A., Yuanjie Zheng, Egan, A., Yushkevich, P.A., and Gee, J.C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–20.

van Swieten, J.C., Rosso, S.M., and Heutink, P. (2000). 'MAPT-Related Disorders', in: Adam, M.P., Ardinger, H.H., Pagon, R.A., Wallace, S.E., Bean, L.J.H., Stephens, K., and Amemiya, A. (Eds.), GeneReviews®.

van Swieten, J.C., and Heutink, P. (2008). Mutations in progranulin (GRN) within the spectrum

of clinical and pathological phenotypes of frontotemporal dementia. *The Lancet Neurology*, 7(10):965–74.

Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7:91.

Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*, 180:68–77.

Veer, I.M., Beckmann, C.F., van Tol, M.-J., Ferrarini, L., Milles, J., Veltman, D.J., Aleman, A., van Buchem, M.A., van der Wee, N.J., and Rombouts, S.A.R.B. (2010). Whole brain resting-state analysis reveals decreased functional connectivity in major depression. *Frontiers in Systems Neuroscience*, 4:41.

Vieira, R.T., Caixeta, L., Machado, S., Silva, A.C., Nardi, A.E., Arias-Carrión, O., and Carta, M.G. (2013). Epidemiology of early-onset dementia: a review of the literature. *Clinical Practice & Epidemiology in Mental Health*, 9:88–95.

Visch-Brink, E., Stronks, D., and Denes, G. (2005). SAT: Semantische Associatie Test. *Swets & Zeitlinger, Lisse*.

Wang, J., Redmond, S.J., Bertoux, M., Hodges, J.R., and Hornberger, M. (2016). A Comparison of Magnetic Resonance Imaging and Neuropsychological Examination in the Diagnostic Distinction of Alzheimer's Disease and Behavioral Variant Frontotemporal Dementia. *Frontiers in Aging Neuroscience*, 8:119.

Wechsler, D. (2005). WAIS-III Nederlandse Bewerking. Technische Handleiding. *Harcourt Test Publishers, Lisse*.

Wegner, C., Filippi, M., Korteweg, T., Beckmann, C.F., Ciccarelli, O., De Stefano, N., Enzinger, C., Fazekas, F., Agosta, F., Gass, A., Hirsch, J., Johansen-Berg, H., Kappos, L., Barkhof, F., Polman, C., Mancini, L., Manfredonia, F., Marino, S., Miller, D.H., Montalban, X., Palace, J., Rocca, M., Ropele, S., Rovira, A., Smith, S.M., Thompson, A., Thornton, J., Yousry, T., and Matthews, P.M. (2008). Relating functional changes during hand movement to clinical parameters in patients with multiple sclerosis in a multi-centre fMRI study. *European Journal of Neurology*, 15(2):113–22.

Wheeler-Kingshott, C.A.M., and Cercignani, M. (2009). About "axial" and "radial" diffusivities. *Magnetic Resonance in Medicine*, 61(5):1255–60.

Whitwell, J.L., and Jack, C.R. (2005). Comparisons Between Alzheimer Disease, Frontotemporal Lobar Degeneration, and Normal Aging With Brain Mapping. *Topics in Magnetic Resonance Imaging*, 16(6):409–25.

Whitwell, J.L., Avula, R., Senjem, M.L., Kantarci, K., Weigand, S.D., Samikoglu, A., Edmonson, H.A., Vemuri, P., Knopman, D.S., Boeve, B.F., Petersen, R.C., Josephs, K.A., and Jack, C.R. (2010). Gray and white matter water diffusion in the syndromic variants of frontotemporal dementia. *Neurology*, 74(16):1279–87.

Whitwell, J.L., Josephs, K.A., Avula, R., Tosakulwong, N., Weigand, S.D., Senjem, M.L., Vemuri, P., Jones, D.T., Gunter, J.L., Baker, M., Wszolek, Z.K., Knopman, D.S., Rademakers, R., Petersen, R.C., Boeve, B.F., and Jack, C.R. (2011a). Altered functional connectivity in asymptomatic

MAPT subjects: a comparison to bvFTD. *Neurology*, 77(9):866–74.

Whitwell, J.L., Weigand, S.D., Gunter, J.L., Boeve, B.F., Rademakers, R., Baker, M., Knopman, D.S., Wszolek, Z.K., Petersen, R.C., Jack, C.R., and Josephs, K.A. (2011b). Trajectories of brain and hippocampal atrophy in FTD with mutations in MAPT or GRN. *Neurology*, 77(4):393–8.

Whitwell, J.L., Weigand, S.D., Boeve, B.F., Senjem, M.L., Gunter, J.L., DeJesus-Hernandez, M., Rutherford, N.J., Baker, M., Knopman, D.S., Wszolek, Z.K., Parisi, J.E., Dickson, D.W., Petersen, R.C., Rademakers, R., Jack, C.R., and Josephs, K.A. (2012). Neuroimaging signatures of frontotemporal dementia genetics: C9orf72, tau, progranulin and sporadics. *Brain*, 135(3):794–806.

Whitwell, J.L., Boeve, B.F., Weigand, S.D., Senjem, M.L., Gunter, J.L., Baker, M.C., Dejesus-Hernandez, M., Knopman, D.S., Wszolek, Z.K., Petersen, R.C., Rademakers, R., Jack, C.R., and Josephs, K.A. (2015). Brain atrophy over time in genetic and sporadic frontotemporal dementia: A study of 198 serial magnetic resonance images. *European Journal of Neurology*, 22(5):745–52.

Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., and Nichols, T.E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92:381–97.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C.F., Jenkinson, M., and Smith, S.M. (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45(1 Suppl 1):S173-86.

Yan, C.-G., Liu, D.-Q., He, Y., Zou, Q.-H., Zhu, C.-Z., Zuo, X.-N., Long, X.-Y., and Zang, Y.-F. (2009). Spontaneous brain activity in the default mode network is sensitive to different resting-state conditions with limited cognitive load. *PLOS One*, 4(5):e5743.

Yang, H., Long, X.-Y., Yang, Y., Yan, H., Zhu, C.-Z., Zhou, X.-P., Zang, Y.-F., and Gong, Q.-Y. (2007). Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional MRI. *NeuroImage*, 36(1):144–52.

Yu, M., Linn, K.A., Cook, P.A., Phillips, M.L., McInnis, M., Fava, M., Trivedi, M.H., Weissman, M.M., Shinohara, R.T., and Sheline, Y.I. (2018). Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Human Brain Mapping*, 39(11):4213–27.

Yuan, B.-K., Wang, J., Zang, Y.-F., and Liu, D.-Q. (2014). Amplitude differences in high-frequency fMRI signals between eyes open and eyes closed resting states. *Frontiers in Human Neuroscience*, 8:503.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., and Alexander, D.C. (2012). NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4):1000–16.

Zhang, Y., Schuff, N., Du, A.-T., Rosen, H.J., Kramer, J.H., Gorno-Tempini, M.L., Miller, B.L., and Weiner, M.W. (2009). White matter damage in frontotemporal dementia and Alzheimer's disease measured by diffusion MRI. *Brain*, 132(9):2579–92.

Zhang, Y., Schuff, N., Ching, C., Tosun, D., Zhan, W., Nezamzadeh, M., Rosen, H.J., Kramer, J.H.,

Gorno-Tempini, M.L., Miller, B.L., and Weiner, M.W. (2011). Joint assessment of structural, perfusion, and diffusion MRI in Alzheimer's disease and frontotemporal dementia. *International Journal of Alzheimer's Disease*, 546871.

Zhang, Y., Tartaglia, M.C., Schuff, N., Chiang, G.C., Ching, C., Rosen, H.J., Gorno-Tempini, M.L., Miller, B.L., and Weiner, M.W. (2013). MRI signatures of brain macrostructural atrophy and microstructural degradation in frontotemporal lobar degeneration subtypes. *Journal of Alzheimer's Disease*, 33(2):431–44.

Zhou, J., Greicius, M.D., Gennatas, E.D., Growdon, M.E., Jang, J.Y., Rabinovici, G.D., Kramer, J.H., Weiner, M., Miller, B.L., and Seeley, W.W. (2010). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain*, 133(5):1352–67.

Zhou, J., and Seeley, W.W. (2014). Network Dysfunction in Alzheimer's Disease and Frontotemporal Dementia: Implications for Psychiatry. *Biological Psychiatry*, 75(7):565–73.

Zivadinov, R., and Cox, J.L. (2008). Is functional MRI feasible for multi-center studies on multiple sclerosis? *European Journal of Neurology*, 15(2):109–10.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–20.

Zou, K.H., Greve, D.N., Wang, M., Pieper, S.D., Warfield, S.K., White, N.S., Manandhar, S., Brown, G.G., Vangel, M.G., Kikinis, R., and Wells, W.M. (2005). Reproducibility of functional MR imaging: preliminary results of prospective multi-institutional study performed by Biomedical Informatics Research Network. *Radiology*, 237(3):781–9.

List of publications

2020

Feis, R.A., van der Grond, J., Bouts, M.J.R.J., Panman, J.L., Poos, J.M., Schouten, T.M., de Vos, F., Jiskoot, L.C., Dopper, E.G.P., van Buchem, M.A., van Swieten, J.C., and Rombouts, S.A.R.B. (2020). Classification using fractional anisotropy predicts conversion in genetic frontotemporal dementia, a proof of concept. *Brain Communications*, 2(2):fcaa079.

de Vos, F., Schouten, T.M., Koini, M., Bouts, M.J.R.J., **Feis, R.A.**, Lechner, A., Schmidt, R., van Buchem, M.A., Verhey, F.R., Olde Rikkert, M.G., Scheltens, P., de Rooij, M., van der Grond, J., and Rombouts, S.A.R.B. (2020). Pre-trained MRI-based Alzheimer's disease classification models to classify memory clinic patients. *NeuroImage: Clinical*, 27:102303.

2019

Feis, R.A., Bouts, M.J.R.J., Dopper, E.G.P., Filippini, N., Heise, V., Trachtenberg A.J., van Swieten, J.C., van Buchem M.A., van der Grond J., Mackay, C.E., and Rombouts, S.A.R.B. (2019c). Multimodal MRI of grey matter, white matter, and functional connectivity in cognitively healthy mutation carriers at risk for frontotemporal dementia and Alzheimer's disease. *BMC Neurology*, 19(1):343.

Feis, R.A., Bouts, M.J.R.J., de Vos, F., Schouten, T.M., Panman, J.L., Jiskoot, L.C., Dopper, E.G.P., van der Grond, J., van Swieten, J.C., and Rombouts, S.A.R.B. (2019b). A multimodal MRI-based classification signature emerges just prior to symptom onset in frontotemporal dementia mutation carriers. *Journal of Neurology, Neurosurgery, and Psychiatry*, 90(11):1207–14.

Feis, R.A., Bouts, M.J.R.J., Panman, J.L., Jiskoot, L.C., Dopper, E.G.P., Schouten, T.M., de Vos, F., van der Grond, J., van Swieten, J.C., and Rombouts, S.A.R.B. (2019a). Single-subject classification of presymptomatic frontotemporal dementia mutation carriers using multimodal MRI. *NeuroImage: Clinical*, 22:101718.

Bouts, M.J.R.J., Grond, J., Vernooij, M.W., Koini, M., Schouten, T.M., Vos, F., **Feis, R.A.**, Cremers, L.G.M., Lechner, A., Schmidt, R., Rooij, M., Niessen, W.J., Ikram, M.A., and Rombouts, S.A.R.B. (2019). Detection of mild cognitive impairment in a community-dwelling population using quantitative, multiparametric MRI-based classification. *Human Brain Mapping*, 40(9):2711–22.

Schouten, T.M., de Vos, F., van Rooden, S., Bouts, M.J.R.J., van Opstal, A.M., **Feis, R.A.**, Terwindt, G.M., Wermer, M.J.H., van Buchem, M.A., Greenberg, S.M., de Rooij, M., Rombouts, S.A.R.B., and van der Grond, J. (2019). Multiple Approaches to Diffusion Magnetic Resonance Imaging in Hereditary Cerebral Amyloid Angiopathy Mutation Carriers. *Journal of the American Heart Association*, 8(3):e011288.

Panman, J.L., Jiskoot, L.C., Bouts, M.J.R.J., Meeter, L.H.H., van der Ende, E.L., Poos, J.M., **Feis,**

R.A., Kievit, A.J.A., van Minkelen, R., Dopper, E.G.P., Rombouts, S.A.R.B., van Swieten, J.C., and Papma, J.M. (2019). Gray and white matter changes in presymptomatic genetic frontotemporal dementia: a longitudinal MRI study. *Neurobiology of Aging*, 76:115–24.

2018

Bouts, M.J.R.J., Möller, C., Hafkemeijer, A., van Swieten, J.C., Dopper, E.G.P., van der Flier, W.M., Vrenken, H., Wink, A.M., Pijnenburg, Y.A.L., Scheltens, P., Barkhof, F., Schouten, T.M., de Vos, F., **Feis, R.A.**, van der Grond, J., de Rooij, M., and Rombouts, S.A.R.B. (2018). Single Subject Classification of Alzheimer's Disease and Behavioral Variant Frontotemporal Dementia Using Anatomical, Diffusion Tensor, and Resting-State Functional Magnetic Resonance Imaging. *Journal of Alzheimer's Disease*, 62(4):1827–39.

2015

Feis, R.A., Smith, S.M., Filippini, N., Douaud, G., Dopper, E.G.P., Heise, V., Trachtenberg, A.J., van Swieten, J.C., van Buchem, M.A., Rombouts, S.A.R.B., and Mackay, C.E. (2015). ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI. *Frontiers in Neuroscience*, 9:395.

Dankwoord

Het behalen van een doctoraat is een absoluut voorrecht en was niet mogelijk geweest zonder de hulp en steun die ik van velen heb mogen ontvangen.

Allereerst wil ik de studiedeelnemers bedanken, met name de deelnemers aan FTD-RisC. Door de periodieke deelname aan deze studie worden zij telkens geconfronteerd met de impact van frontotemporale dementie op henzelf en hun familie. Toch weerhoudt dat hen er niet van om met grote overtuiging deel te nemen. Zonder hun onbaatzuchtige inzet was dit werk niet mogelijk geweest. Het is mijn hoop dat dit proefschrift bijdraagt aan een toekomst waarin frontotemporale dementie tijdig opgespoord en behandeld kan worden.

Mijn dank gaat verder uit naar mijn promotoren, Serge Rombouts en Mark van Buchem, en copromotor Jeroen van der Grond. Serge, ontzettend bedankt voor je fijne begeleiding. Je liet me de vrijheid om een koers uit te zetten die bij me past, maar je deur stond altijd open als ik om raad verlegen zat. Mark, bedankt voor je vertrouwen in mij, mijn project en de medische toepassing van artificiële intelligentie in het algemeen. Jeroen, bedankt voor je frisse blik, kritische noten en je humor tijdens onze besprekingen.

Prof. dr. Hogendoorn wil ik graag bedanken voor zijn onmiddellijke vertrouwen en steun, die hebben geleid tot een prachtige wetenschapsstage in Oxford.

At FMRIB, I would like to thank Clare Mackay for taking me into her research group and giving me the opportunity to develop into a true neuroscientist. Nico Filippini and Verena Heise, thank you for your patience, lessons in neuroimaging, and for your continued support after I left Oxford. Steve Smith, thank you for your collaboration on the FIX paper; it was a privilege to work with you. Other colleagues at FMRIB, thank you for your help and for the great time I had.

Ik ben mijn collega's in het LUMC en FSW erg dankbaar. Anne, jij hebt mij de eerste scripts geleerd nog voor ik naar Oxford ging, en had altijd een antwoord of juist die ene wedervraag paraat, dankjewel daarvoor. Mark, dankjewel voor de gezelligheid, voor je uitleg over classificatie en voor het gebruik van je analyse pipeline. FrankenTijn, bedankt voor het delen van jullie statistische kennis en voor alle komische gesprekken in Leiden en op congres. Zelfs in dit dankwoord zijn jullie onafscheidelijk. Andere collega's van K4.44 en de afdeling Methode & Statistiek, bedankt voor de gezellige momenten tijdens de koffie en lunch.

Mijn collega's in het Erasmus MC wil ik heel hartelijk danken dat ik gebruik mocht maken van de door hen verzamelde data. Dit tijdrovende werk is mij bespaard gebleven dankzij hun enorme inzet en generositeit. Jessica Panman en Lize Jiskoot wil ik in het bijzonder bedanken voor hun hulp bij het uitzoeken van de data en hun uitleg over de neuropsychologische meetinstrumenten. Prof. dr. Van Swieten, beste John, bedankt voor je heldere uitleg in treffende en gevatte metaforen. Jouw klinische blik op ons werk heeft geholpen om altijd de patiënt in het vizier te houden.

Een promotietraject vergt intensieve en langdurige concentratie op een kleine wetenschappelijke niche. Ik ben mijn vrienden daarom erg dankbaar voor hun steun en brede interesse. Zij hebben mij geholpen om te relativeren en om uit mijn wetenschappelijke cocon te breken. Sietse, Martijn en Benjamin, bedankt voor jullie langdurige vriendschap; ik hoop dat we nog vaak op avontuur gaan met een net niet toereikende voorbereiding. Ravage, dank voor jullie onvoorwaardelijke steun en voor de discussies over alles (of vooral niets). De Civitas wil ik bij voorbaat graag bedanken voor het bedanken voor het bedanken. Jullie hebben geholpen om tijdens het promoveren Gewoon Rogier te blijven. (Oud-)Teamgenoten, bedankt voor alle hilarische momenten binnen en buiten het veld.

Speciale dank gaat uit naar mijn familie. Papa, helaas heb je mijn passie voor de neurowetenschap nooit mogen meemaken. Toch kan er geen twijfel bestaan dat jij me hebt geïnspireerd om dit

avontuur aan te gaan. Mama, ontzettend bedankt voor je steun en liefde, die ook in moeilijke tijden zowel onvoorwaardelijk als onuitputtelijk bleken. Marieke, jouw doorzettingsvermogen en integriteit zijn een lichtend voorbeeld voor mij en voor de wetenschap in het algemeen. Oom Eric, bedankt voor je scherpzinnigheid en vriendschap.

Tot slot, liefste Els, jij kunt van de meest simpele dingen een feest maken. Dankjewel voor je stralende aanwezigheid, humor en liefde. Helaas kan ik, zelfs na drie jaar classificatieonderzoek, nog niet goed voorspellen welke woordgrappen jij wel en niet kan waarderen; daar is vrees ik meer data voor nodig. Ik kan niet wachten om met jou te trouwen en samen beetje bij beetje de wereld te ontdekken.

Curriculum vitae

Rogier Alexander Feis was born on the first of May in 1991. He grew up in his birthplace, The Hague, where he graduated from the Christelijk Gymnasium Sorghvliet cum laude in 2009. During the last two years of secondary school, he enrolled in an extracurricular track called Pre-University College at Leiden University, where he first experienced academic research during an internship at the Leiden Academic Centre for Drug Research.

Rogier went on to study Medicine at the Leiden University Medical Centre, and obtained his Bachelor's degree in 2012. During his Master's degree, he spent eight months in the United Kingdom at the Oxford Centre for Functional MRI of the Brain for a neuroimaging research internship. This internship, performed in collaboration with the department of Radiology of the Leiden University Medical Centre, formed the basis for the research presented in this dissertation. Rogier continued with neuroimaging research alongside his medical clinical internships, and graduated from Medical school cum laude in 2016.

To further advance his research project as a PhD candidate, he received an MD/PhD-grant from the Leiden University Medical Centre, and started his PhD project under supervision of prof. dr. Serge Rombouts, prof. dr. Mark van Buchem, and dr. Jeroen van der Grond at the department of Radiology. The work in this dissertation was performed in collaboration with the department of Methodology and Statistics, institute of Psychology, faculty of Social and Behavioural Sciences, Leiden University; the department of Neurology, Erasmus Medical Centre; the Oxford Centre for Functional MRI of the Brain.

After obtaining his PhD, Rogier aims to specialise in the field of Neurology. In March 2020, he started his clinical career as house officer at the Internal Medicine department of the Alrijne hospital in Leiderdorp.