Automated subcortical brain segmentation using multispectral MRI for improved AD diagnosis and disease tracking

Emily Ndakola Manning

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Dementia Research Centre
University College London

March 9, 2018

I, Emily Ndakola Manning, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

This thesis is a detailed investigation into subcortical changes in Alzheimer's disease (AD). Hippocampal volumes, shapes and diffusion metrics were investigated in different disease stages and presentations, and the ability of these metrics for disease group classification investigated. A new method for automated thalamic segmentation using multimodal imaging was developed and applied to two different datasets.

Hippocampal volumes were found to be disproportionately affected by the apolipoprotein (APOE) $\varepsilon 4$ allele. Hippocampal volumes were also found to be reduced in subjects with the posterior cortical atrophy variant of AD. These changes were localized in the hippocampal tail region and hippocampal shape metrics were found to be superior to hippocampal volumes in differentiating these subjects from controls.

The manual thalamic segmentation protocol developed was found to have good reliability, and a template library of thalamic segmentations was generated for use in automated pipelines. The manual segmentation protocol used both T1-weighted and diffusion magnetic resonance imaging (MRI) scans for improved segmentation accuracy. The template library was used for automatic segmentation of the thalamus and leave-one-out cross-validation revealed good segmentation reliability, better than that reported by the most widely-used automated thalamic segmentation techniques. The thalami from subjects from the Alzheimer's disease neuroimaging initiative (ADNI)-GO/2 datasets, which includes control subjects, subjects with subjective memory complaints, with early mild cognitive impairment (EMCI), late mild cognitive impairment

(LMCI) and AD, were segmented using the automated thalamic pipeline. Subjects with AD and mild cognitive impairment (MCI) were found to have lower thalamic volumes, as well as lower hippocampal volumes suggesting early thalamic involvement.

Differences in diffusion metrics were found and some diffusion metrics were associated with subsequently higher atrophy rates. The inclusion of hippocampal and thalamic diffusion metrics, in addition to volumes were found to improve disease group classification.

In summary, this work in this thesis extends existing knowledge about how the hippocampi and thalami are affected in Alzheimer's disease.

Acknowledgements

Firstly I would like to thank all the individuals who volunteered to be scanned as part of the various studies included in this thesis, without whom none of the work presented in this thesis would be possible.

I would like to thank Jo Barnes, Jorge Cardoso and Jonathan Schott, my PhD supervisors. Thank you Jo, you have been incredibly supportive over the last few years, even making time to travel to London for supervisions during your maternity leave. I have really appreciated having you as my primary supervisor, have really enjoyed working with you and feel that I have been very lucky in that regard. Thank you Jorge, particularly for all your support and input whilst Jo was on maternity leave. And thanks to Jon for your input into the various studies in this thesis and for providing me with access to the data from the AVID and YOAD studies. I would also like to thank Nick Fox for enabling me to do a PhD at the Dementia Research Centre and for all his helpful input into the various studies in this thesis. A big thanks to all those who directly helped with the work that formed this thesis: Kelvin Leung, Jonathan Bartlett, Jennifer Nicholas, Chris Frost, Ian Malone, Manja Lehmann, Kate Macdonald, Tracey Pepple, Jonathan Young, Dave Cash, Marc Modat, Carole Sudre, Maria Zuluaga and Seb Ourselin. Thanks to Susie Henley and Marc Modat for reading and examining my upgrade thesis and to Anne Parnell for her administrative support. Thanks also to my colleagues (and friends) from the imaging trials team who I've worked closely with over the years: Laila Ahsan, Liz Gordon, Ian Malone, Casper Nielsen and Shona Clegg. Thanks also to Cassy Fiford and Tom Veale for being awesome desk buddies.

I'd like to thank my family, Peter, Barbara and Danny who have been incredibly supportive.

Finally, thank you Andrea, for putting up with me over the last few months, and for all your emotional support. Without you this would have all been much harder.

Contents

1 Thesis aims 20

2 Introduction 24

2.1 Background 24

2.2 Biomarkers in Alzheimer's disease 26

2.2.1 Biomarkers of amyloid and tau deposition in the brain . 27

2.2.2 Structural brain imaging as a biomarker 28

2.2.3 Neuropsychology 34

2.2.4 Genetics 34

2.2.5 Ordering of biomarker changes in Alzheimer's disease . . 34

2.3 Measuring volumes and volume change using MRI scans 35

2.3.1 Manual segmentation 35

2.3.2 Automated segmentation 36

2.3.3 Indirect measures of change 36

2.3.4 Direct measures of change 36

2.3.5 Structures commonly segmented in AD 38

2.3.6 Multimodal segmentation: a way of investigating subcortical structures? 38

3 Methods 40

3.1 Subjects 40

3.1.1 ADNI 40

3.1.2 Young-onset Alzheimer's Disease 41

3.1.3 AVID cohort 41
3.1.4 Posterior Cortical Atrophy cohort 41
3.2 MRI acquisition 42
3.2.1 ADNI 42
3.2.2 young onset Alzheimer's disease (YOAD) 43
3.2.3 AVID 44
3.2.4 Posterior cortical atrophy 44
3.3 Image processing 45
3.3.1 Diffusion processing pipeline 45
3.4 Brain region segmentation 46
3.4.1 Brains 46
3.4.2 Ventricles 46
3.4.3 Hippocampi 47
3.4.4 Total intracranial volume 47
3.5 Volume change measurement 48
3.5.1 Brains 48
3.5.2 Ventricles 48
3.5.3 Hippocampi 48
3.6 Software 48
3.6.1 For manual or semi-automated region segmentation . . . 48
3.6.2 Stata 49
3.6.3 Matlab 49

4 Assessment of image quality in accelerated T1-weighted scans 50
4.1 Introduction 50
4.2 Materials and Methods 51
4.2.1 Image data and acquisition 51
4.2.2 Quality control (quality control (QC)) of MRI scans . . . 52
4.2.3 Quality Comparison of accelerated and unaccelerated scans 54
4.2.4 Segmentation and volume measurement 55
4.2.5 Volume Change Measurement 55

4.2.6 Comparison of accelerated and unaccelerated volumes and atrophy rates 55
4.2.7 Comparison of subject characteristics between passed and failed scans 57
4.2.8 Sample size estimates 57
4.3 Results 58
4.3.1 Quality Control 58
4.3.2 Scan quality differences between accelerated and unaccelerated T1 MRI scans 60
4.3.3 Baseline volumes 60
4.3.4 Atrophy rates 61
4.3.5 Comparison of subject characteristics between those with failed scan pairs due to motion and those with passed scan pairs 63
4.3.6 Sample size estimates 63
4.4 Discussion 66

5 Investigation into hippocampal atrophy in APOE ε4 carriers 70
5.1 Introduction 70
5.2 Methods 71
5.2.1 Subjects 71
5.2.2 Image acquisition and analysis 71
5.2.3 Statistical Analysis 73
5.3 Results 74
5.3.1 Baseline cross-sectional results 75
5.3.2 Longitudinal Results 75
5.4 Discussion 78

6 Investigation into hippocampal shape and volume in Posterior Cortical Atrophy 86
6.1 Introduction 86

6.2 Methods 87
6.2.1 Subjects 87
6.2.2 Image acquisition 87
6.2.3 Image processing 87
6.2.4 Hippocampal shape analysis 88
6.2.5 Statistics: demographics 90
6.2.6 Statistics: brain and hippocampal volume analyses . . . 90
6.2.7 Statistics: hippocampal shape analysis 91
6.2.8 Classification of subjects using hippocampal shape features 92
6.3 Results 93
6.3.1 Participant demographics 93
6.3.2 Brain and hippocampal volume analysis 95
6.3.3 Hippocampal shape analysis 95
6.4 Disease classification using SPHerical HARMonics (SPHARM) coefficients 96
6.5 Discussion 100

7 Hippocampal template library comparison 106
7.1 Introduction 106
7.2 Methods 109
7.2.1 Subjects 109
7.2.2 Hippocampal segmentation 110
7.2.3 Validation of template library 2 111
7.2.4 Statistical analyses 112
7.3 Results 113
7.3.1 Leave-one-out cross validation of template library 2 . . . 113
7.3.2 Disease group classification comparison 113
7.4 Discussion 115

8 Manual and automated thalamic segmentation 117
8.1 Introduction 117

8.2 Methods 119
8.2.1 Subjects 119
8.2.2 MRI scan acquisition and processing 119
8.2.3 Manual segmentation protocol 120
8.2.4 Testing reliability of manual segmentations 127
8.2.5 Testing reliability of automated segmentations 128
8.3 Results 128
8.3.1 Reliability of manual segmentations 128
8.3.2 Reliability of automated segmentations 129
8.4 Discussion 134

9 Investigation into thalamic volume and diffusion metrics in mild cognitive impairment and Alzheimer's disease 136
9.1 Introduction 136
9.2 Methods 137
9.2.1 Study Participants 137
9.2.2 MRI Acquisition 137
9.2.3 Quality control 138
9.2.4 Total intracranial volume, whole-brain, hippocampal and thalamic region segmentation 138
9.2.5 Extraction of diffusion metrics within regions-of-interest (region of interest (ROI)s) 138
9.2.6 Whole-brain and hippocampal atrophy rates 138
9.2.7 Statistical Analyses 138
9.3 Results 141
9.3.1 Quality control 141
9.3.2 Participant Demographics 143
9.3.3 Baseline volumes 144
9.3.4 Diffusion metrics 149
9.3.5 Disease classification using diffusion metrics in addition to volumetric metrics 154

9.3.6 Predictors of atrophy rate 156
9.4 Discussion 157

10 Investigation in to thalamic and hippocampal volumes and diffusion metrics in early-onset Alzheimer's disease 162
10.1 Introduction 162
10.1.1 Methods 163
10.1.2 Results 164
10.2 Discussion 167

11 General Conclusions 171

12 Publications 174

13 Appendix: Statistical models used in APOE ε4 study 175
13.1 Cross-sectional analysis 175
13.2 Longitudinal analysis 176

Acronyms 178

Bibliography 180

List of Figures

2.1 Sagittal, coronal and axial slices of a T1-weighted scan, T2-weighted scan and coloured fractional anisotropy map. 29

2.2 Sagittal, coronal and axial views of T1-weighted MRI scans in a healthy individual, MCI and AD 30

2.3 Baseline and month 36 follow-up scan illustrating purpose of registration. 33

2.4 Illustration of the boundary shift integral. 37

4.1 Breakdown of no. of subjects at each stage 59

5.1 Subject selection process 72

5.2 Effect of APOE ε4 on baseline hippocampal volumes. 76

5.3 Effect of APOE ε4 on hippocampal atrophy rates. 79

5.4 Effect of APOE ε4 on whole-brain atrophy rates 80

5.5 Difference in hippocampal atrophy rates: ε4+ vs ε4-. 81

6.1 Summary of shape analysis procedure 91

6.2 Nested 10-fold cross-validation procedure used for model tuning and evaluation. 94

6.3 Hippocampal shape difference after adjusting for age, gender, and total intracranial volume (TIV) in (a) posterior cortical atrophy (PCA) vs. controls, (b) typical Alzheimer's disease (tAD) vs. Controls, (c) PCA vs. tAD. 97

6.4 Hippocampal shape difference after adjusting for age, gender, and hippocampal volume in (a) PCA vs. controls, (b) tAD vs. Controls, (c) PCA vs. tAD. 98

6.5 Hippocampal shape difference in PCA vs. tAD after adjusting for (a) age, gender, mini-mental state exam (MMSE) score, disease duration, and head size and (b) age, gender, MMSE score, disease duration, and hippocampal volume. 99

7.1 Example of the first three coronal slices of an automated hippocampal segmentation generated from the template library without tails illustrating the arbitrary nature of the starting slices. 108

7.2 Sagittal cross-section through a hippocampus segmented manually using (a) the protocol used in template library 1 and (b) the protocol used in template library 2. 112

7.3 Receiver operator curves for disease group classification using template library 1 and template library 2 114

8.1 Coronal slice showing the location of the thalamus and it's surrounding structures 120

8.2 Axial slice showing the location of the thalamus and it's surrounding structures 121

8.3 Three most superior axial slices through the thalamus 122

8.4 Axial slice through the thalamus (outlined in red) where white matter of fornix is visible posterior to the thalamus. 122

8.5 Coronal, sagittal and axial view through an example brain with cross-hairs indicating the fornix appearing as white matter posterior to thalamus. 123

8.6 Coronal, sagittal and axial view through the thalamus with cross-hairs indicating the habenula appearing as white matter medial to thalamus. 124

8.7 Axial section through brain showing location of superior colliculus with respect to the thalamus (outlined in red). 125
8.8 Coronal, sagittal and axial view through thalamus with cross-hairs indicating the top of the red nucleus, inferior to the thalamus. 125
8.9 Axial slice through the brain at the level of the cerebral peduncle 126
8.10 Axial slice through thalamus on T1-weighted scan (left) and coloured fractional anisotropy (FA) map (right) 127
8.11 Best and worst automated thalamic segmentation (according to dice score) when using T1-weighted scans only. 131
8.12 Best and worst automated thalamic segmentation (according to dice score) when using FA maps in addition to T1-weighted scans only. 132
8.13 Scan with the biggest difference in dice scores between the automatically generated thalamic region using T1-weighted scans and the automatically generated thalamic region using both FA maps and T1-weighted scans. 133

9.1 Examples of scans with motion artefacts severe enough for exclusion from the study. 141
9.2 Examples of scans of good quality with little motion evident. . . 141
9.3 Subject with likely incidental arachnoid cyst in left thalamus. Arrows indicate location of arachnoid cyst. 142
9.4 Comparison of baseline characteristics and volumes between those who had diffusion imaging available at baseline and those who did not 144
9.5 Baseline mean adjusted volumes in the full ADNI-2/GO cohort. 146
9.6 Baseline mean adjusted volumes in the subset of subjects with diffusion imaging available. 148
9.7 Mean mean diffusivity (MD) in left and right hippocampus . . . 149
9.8 Mean MD in left and right thalamus 151
9.9 Mean FA in left and right hippocampus. 152

9.10 Mean FA in left and right thalamus 153
9.11 receiver operator curve (ROC) curve comparison for disease group classification using thalamic and hippocampal volumes and diffusion metrics 155
9.12 Baseline hippocampal MD vs hippocampal atrophy rate 156
9.13 Baseline thalamic MD and whole-brain atrophy rate 157
9.14 Baseline thalamic FA vs whole-brain atrophy rate 157

List of Tables

3.1 MRI scan parameters by diagnostic group 45

4.1 Subject Demographics 53

4.2 Baseline volumes (mean (sd) unless otherwise stated) 60

4.3 Comparison of whole-brain ventricular and hippocampal atrophy rates calculated from accelerated and non-accelerated scan pairs using the boundary shift integral in subjects whose accelerated and non-accelerated scan pairs both passed QC. 62

4.4 Comparison of subject characteristics between those with failed scan pairs due to motion and those with passed scan pairs . . . 64

4.5 Atrophy rates by diagnosis in all subjects (regardless of whether scan pairs passed or failed QC) 65

4.6 Sample size estimates for a 25% reduction in atrophy rate (with bootstrap 95% CIs and 80% power) 66

5.1 Baseline demographics and image summary statistics by clinical group. Mean(SD) unless otherwise stated) 75

5.2 Adjusted mean baseline hippocampal volumes for $\varepsilon 4$ non-carriers and adjusted mean differences in total (left and right summed) baseline hippocampal volumes between $\varepsilon 4$ carriers and non-carriers in controls, stable MCI, MCI progressors and AD . 76

5.3 Adjusted mean difference in whole-brain and hippocampal atrophy rate (ml) [95% CI] for $\varepsilon 4$ carriers compared with non-carriers in controls, stable MCI, MCI progressors and AD (+ve means atrophy rate is higher in $\varepsilon 4$+). 77

6.1 Participant Demographics 94

6.2 Brain and hippocampal volumes (adjusted for age, gender and head-size) 95

6.3 Comparison of performance of support vector machine (SVM) classifer using SPHARM coefficients vs SVM classifier using left and right hippocampal volumes as features) 100

7.1 Demographics of subjects included in template library 1 109

7.2 Demographics of subjects included in template library 2 110

8.1 Demographics of subjects used in template library 119

8.2 Similarity metrics for repeated manual segmentations. Mean [95% confidence interval] shown. 129

8.3 Similarity metrics for automated segmentations using T1-weighted scans only vs manual segmentations and automated segmentations using FA maps in addition to T1-weighted scans vs manual segmentations. Mean [95% confidence interval] shown 130

9.1 Baseline characteristics of ADNI-2/GO cohort. Mean (sd) shown unless otherwise indicated. 143

9.2 Area under ROC curves for disease group classification using thalamic and hippocampal volumes and diffusion metrics 154

10.1 Baseline characteristic of local young-onset AD cohort. Mean (standard deviation (SD)) shown. †Significant difference with respect to controls ($p<0.05$). 165

10.2 Mean adjusted baseline volumes [95% confidence interval] by diagnostic group. Volumes are adjusted for mean age, mean head-size and assume an equal gender split. 166
10.3 Mean adjusted [95% confidence interval] FA and MD metrics in the left and right thalamus and hippocampus. Metrics were adjusted for mean age and assume an equal gender split. 167

Chapter 1

Thesis aims

Change in volume of whole brains, ventricles and hippocampi are well-established markers of structural brain degeneration in Alzheimer's disease (AD). Whole brain and hippocampal atrophy rates and ventricular expansion rates are elevated in AD and mild cognitive impairment (MCI) and have been used as outcome markers in clinical trials of potentially disease-modifying therapies [Fox et al., 2005, Salloway et al., 2009]. Rates of hippocampal atrophy are usually considered more specific markers of AD progression and have been shown to differentiate mildly cognitively impaired subjects from healthy controls at an early disease stage [Henneman et al., 2009]. Presentations of AD are heterogeneous however, and in some variants, such as posterior cortical atrophy (PCA) the hippocampus remains relatively spared [Crutch et al., 2012]. In addition, there is evidence that the hippocampi may be affected differently in subjects who are carriers of the apolipoprotein (APOE) $\varepsilon 4$ allele [Agosta et al., 2009, Geroldi et al., 1999, Lehtovirta, 1995, Pievani et al., 2009], the biggest genetic risk factor for sporadic AD [van Es and van den Berg, 2009].

In order to quantify changes in volumes using magnetic resonance imaging (MRI), the structures or regions require segmentation. Baseline and follow-up scans can then be registered and change in volume can be measured directly from registered scan pairs using techniques such as the boundary shift integral (boundary shift integral (BSI)) [Freeborough et al., 1997]. In addition to volume and volume change, shape metrics can be extracted from brain structures.

There is evidence that shape metrics provide additional information to volume that may aid in differential diagnosis [Gerardin et al., 2009].

Most segmentation algorithms require a single imaging modality: T1-weighted structural MRI scans. Although whole-brain, ventricular and hippocampal volumes are widely used as AD biomarkers, relatively little attention has been given to the role of other subcortical grey matter structures in AD. However, there is evidence to suggest that structures such as the parahippocampal gyrus [Braak and Braak, 1995,Echávarri et al., 2011], amydala [Vogt et al., 1990, Poulin et al., 2011], thalamus [Braak and Braak, 1991, De Jong et al., 2008] and putamen [De Jong et al., 2008] also undergo atrophy in AD. Some of the boundaries of these structures are difficult to distinguish from surrounding tissue on T1-weighted MRI scans meaning the accurate segmentation of these regions remains a challenge.

It may be that other types of MRI scan which provide complimentary information on brain structure (such as diffusion tensor imaging (DTI)) improve the segmentation accuracy of some of these deep brain structures. However, including more types of scan has the disadvantage of increasing scan time and hence cost and patient discomfort. Patient discomfort may result in motion artefacts and although data is routinely discarded from studies due to motion, a thorough investigation into how this may bias the data is warranted. Further, the more scan types required for a segmentation to be successful, the more likely one of these will have artefacts, such as motion, which will render the scan unusable.

The aims of this thesis are therefore to:

1. Investigate whether accelerated T1-weighted MRI scans can be used in place of non-accelerated scans for measuring volume and atrophy rates in established structural biomarkers of AD and whether bias is introduced by excluding scans which appear to suffer from motion artefacts.

   Hypothesis: That accelerated T1-weighted scans can be used in place of unaccelerated for measures of brain volume and atrophy and that fewer

accelerated scans suffer from motion artefacts.

2. To explore in depth the heterogeneity of hippocampal atrophy patterns in Alzheimer's disease using existing techniques:

   (a) To compare hippocampal and whole-brain atrophy rates in APOE $\varepsilon 4$ carriers and non-carriers and to investigate whether APOE $\varepsilon 4$ carriers show disproportionate hippocampal atrophy.
   Hypothesis: That not only do APOE $\varepsilon 4$ carriers show higher rates of hippocampal atrophy compared with non-carriers, but that the hippocampi are disproportionately affected in APOE $\varepsilon 4$ carriers.

   (b) To investigate if and how the hippocampi of subjects with the PCA variant of AD are affected using volume and shape metrics.
   Hypothesis: That subjects with the PCA variant of AD have reduced hippocampal volumes compared with controls and that the posterior portion of the hippocampus is likely more affected than the anterior portion (given the that the posterior portion of the brain is more affected in PCA subjects).

3. To investigate whether automated hippocampal segmentation can be improved upon by using a template library which includes the full extent of the hippocampal tail.
   Hypothesis: That the hippocampal tail also atrophies in Alzheimer's disease and that a template library with segmentations that include the full extent of the hippocampal tail will improve disease group classification.

4. To investigate whether the use of diffusion-weighted imaging in addition to T1-weighted imaging improves segmentation reliability of the thalamus and to generate a template library of thalamic regions for automated segmentation pipelines.
   Hypothesis: That diffusion-weighted imaging provides useful additional structural information about the thalamus and will improve the reliability of thalamic segmentation.

5. To investigate whether there are thalamic and hippocampal volume and diffusion differences in:

   (a) The earliest stages of sporadic Alzheimer's disease - subjective memory complainers and mild cognitive impairment, as well as in established Alzheimer's disease.

   (b) Young-onset Alzheimer's disease.

   Hypothesis: That thalamic volumes are reduced in subjects with MCI and that group differentiation can be improved by including thalamic and hippocampal diffusion metrics. That thalamic volumes are reduced in young-onset Alzheimer's disease compared with controls.

Chapter 2

Introduction

2.1 Background

In the UK an estimated 850 thousand people are living with dementia [M. Knapp et al., 2014] whilst globally, the number is estimated to be more than 46 million [Prince et al., 2015]. There is evidence to suggest that the prevalence of dementia has actually fallen slightly in recent years in some high-income countries [Langa et al., 2017,Wu et al., 2016], possibly due to improved education and lifestyle factors. Increasing life expectancy in these countries means that the numbers of people living with dementia will continue to rise significantly nonetheless. Low- and middle-income countries, on the other hand, have seen a significant rise in the prevalance of dementia in recent years and the numbers are predicted to continue rising dramatically [Prince et al., 2015]. By 2050, an estimated 131.5 million people will be living with dementia across the world [Prince et al., 2015].

Dementia is a general term that describes a syndrome which results in a decline in cognitive function severe enough to affect activities of daily living, usually associated with progressive neuronal loss. According to the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV), dementia is a syndrome that may be caused or characterized by:

Multiple cognitive deficits, which include memory impairment and at least one of the following: aphasia, apraxia, agnosia or disturbance in executive

functioning. Social or occupational function is also impaired. A diagnosis of dementia should not be made during the course of a delirium. (A dementia and a delirium may both be diagnosed if the dementia is present at times when the delirium is not present.)

Alzheimer's disease is the most common underlying pathology for dementia, accounting for 60-80% of cases followed by vascular dementia, Lewy body dementia, and frontotemporal dementia [Barker et al., 2002, Wilson et al., 2012]. A majority of those with Alzheimer's disease pathology actually have some form of mixed pathology, such as Alzheimer's disease together with vascular dementia [Schneider et al., 2007].

Although dementia is typically thought of as a disease of old-age, approximately 5% of those with dementia are under the age of 65 at symptom onset [M. Knapp et al., 2014]. Dementia in those under the age of 65 is known as young-onset dementia with Alzheimer's disease thought to be the underlying pathology in 30-35% of cases [Harvey and Skelton-Robinson, 2003]. Although some genetic mutations cause young-onset Alzheimer's disease, in the vast majority of cases, there is no known genetic cause.

Previous studies have shown that there is more heterogeneity in symptom presentation [Barnes et al., 2015] and atrophy patterns in those who are affected by the Alzheimer's at a younger age [Mendez et al., 2012]. A larger proportion of patients with young-onset Alzheimer's disease have a non-amnestic presentation than older-onset cases [Mendez et al., 2012].

Posterior cortical atrophy (PCA) is a clinicoradiological syndrome characterized by impairment of visuoperceptual, visuospatial and other posterior cognitive functions and atrophy of the occipital, parietal and occipito-temporal cortices [Benson et al., 1988, Crutch et al., 2012]. PCA is most commonly caused by Alzheimer's disease (AD) [Alladi et al., 2007, Renner et al., 2004, Tang-Wai et al., 2004] and is probably the most common atypical clinical presentation of AD [Dubois et al., 2014a]. Although AD is the most common underlying pathology in PCA (>80%) [Alladi et al., 2007, Renner

et al., 2004, Tang-Wai et al., 2004] a number of cases have other underlying pathologies such as dementia with Lewy bodies (DLB), corticobasal degeneration, prion disease and subcortical gliosis [Crutch et al., 2012]. Unlike typical Alzheimer's disease (tAD), where memory loss is one of the earliest and most prominent symptoms, episodic memory function in PCA is initially relatively well preserved [Mendez et al., 2002, Tang-Wai et al., 2004]. Previous cross-sectional magnetic resonance imaging (MRI) studies have shown distinct atrophy patterns in subjects with PCA as compared to tAD with PCA showing greater atrophy in the right occipital cortex compared with tAD subjects and tAD subjects showing greater atrophy in the left temporal lobe compared to PCA [Lehmann et al., 2011, Whitwell et al., 2007].

After a diagnosis of Alzheimer's disease, a person can live for several years, or even decades, for many of these requiring full-time care. It is a disease that can be devastating on an individual level. On a global level it is already having a huge impact on health services and society and as the number of people affected by the disease continues to grow, it will become an increasingly pressing burden.

As such, therapies that slow down or stop the disease are urgently needed.

2.2 Biomarkers in Alzheimer's disease

A biomarker is 'a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention' [Atkinson et al., 2001].

Biomarkers are required for accurate diagnoses and for objective measurements to test the effect of putative Alzheimer's therapies on the brain as well as for understanding more about disease progression.

In Alzheimer's disease, there is no single biomarker than can be used for a definitive diagnosis of AD during life. A definitive diagnosis of AD can only be made when the hallmark features of the disease - amyloid plaques and neurofibrillary tangles - are identified in the brain, most often using histopathol-

ogy at post-mortem. This means that in many of the clinical trials that have taken place to date, a percentage of participants may not have actually had AD pathology [Jellinger and Attems, 2015]. Another problem clinical trials in AD have faced is that by the time a person has symptoms of AD , they may have already lost a substantial amount of brain tissue. Complicating matters further, it is a disease that can present with a variety of symptoms and the disease can initially affect different areas of the brain in different individuals.

In order to find an effective treatment for Alzheimer's disease, it is necessary to test potential disease-modifying therapies in those who genuinely have AD pathology, to target the disease at a stage when extensive damage has not already taken place and to take into consideration some of the heterogeneity in disease presentation when using biomarkers as outcome measures in clinical trials.

The importance of recognising preclinical AD , in order that individuals can be targeted at this early stage, before tissue is lost and symptoms are accrued has resulted in published guidelines [Knopman et al., 2012, Dubois et al., 2016]. The guidelines in these publications do not establish diagnostic criteria for preclinical AD , but propose that additional research is undertaken in order to establish which biomarkers are better able to identify dementia-related changes in the brain at a preclinical stage.

A number of biomarkers are currently in use in Alzheimer's disease and have complimentary roles. Although the focus of this thesis is on structural brain imaging as well as diffusion tensor imaging, a brief summary of some of the most common biomarkers in use in Alzheimer's disease is given below.

2.2.1 Biomarkers of amyloid and tau deposition in the brain

Abnormal levels of amyloid in the brain can be measured using cerebrospinal fluid (CSF) [Shaw et al., 2009] and positron emission tomography (PET) scans [Klunk and Mathis, 2008]. Although normal amyloid levels can be used to exclude those without underlying Alzheimer's disease pathology, there are many individuals who are found to have numerous amyloid plaques post-

mortem who had no other evidence of the disease during life [Dickson et al., 1992]. Therefore, evidence of amyloid in the brain alone, is not enough for a diagnosis of Alzheimer's disease.

Tau protein levels can also be measured in CSF [Blennow et al., 2010], and, more recently, imaging of tau pathology using PET tracers has become possible [Villemagne et al., 2015]. Evidence suggests that tau deposition in the brain is more tightly coupled to neurodegeneration than amyloid. Tau imaging is still far from being widely available in the clinic however.

2.2.2 Structural brain imaging as a biomarker

MRI plays a key role in diagnosis, research and clinical trials in dementia. Single time-point structural MRI of the brain allows for the visualisation of atrophy patterns, which can aid diagnosis in neurodegenerative diseases. Recent diagnostic criteria recommend structural imaging not only to exclude treatable causes (e.g. tumour) but to support diagnosis (including variants of AD) [McKhann et al., 2011, Dubois et al., 2014b, Hyman et al., 2012]. Volumetric MRI allows for quantification of brain volumes. Serial 3D structural MRI scans of the brain allows for the visualisation and quantification of progressive tissue loss, which may be helpful in particularly difficult cases, and notably as objective assessments in clinical trials. Reductions in brain volume are thought to predominantly reflect neuronal loss in AD. Differences in hydration can impact brain volume measures however [Duning et al., 2005]. In addition, anti-amyloid trials have reported increased brain volume loss in antibody responders [Fox et al., 2005] [Vellas et al., 2009]. It is thought that this loss in volume was due to the removal of amyloid-plaques or changes in fluid balance, rather than increased tissue loss in these individuals [Fox et al., 2005].

2.2.2.1 T1-weighted MRI

An example of a T1-weighted MRI scan is shown in the first column of figure 2.1. In T1-weighted MRI scans, grey matter appears grey, white matter ap-

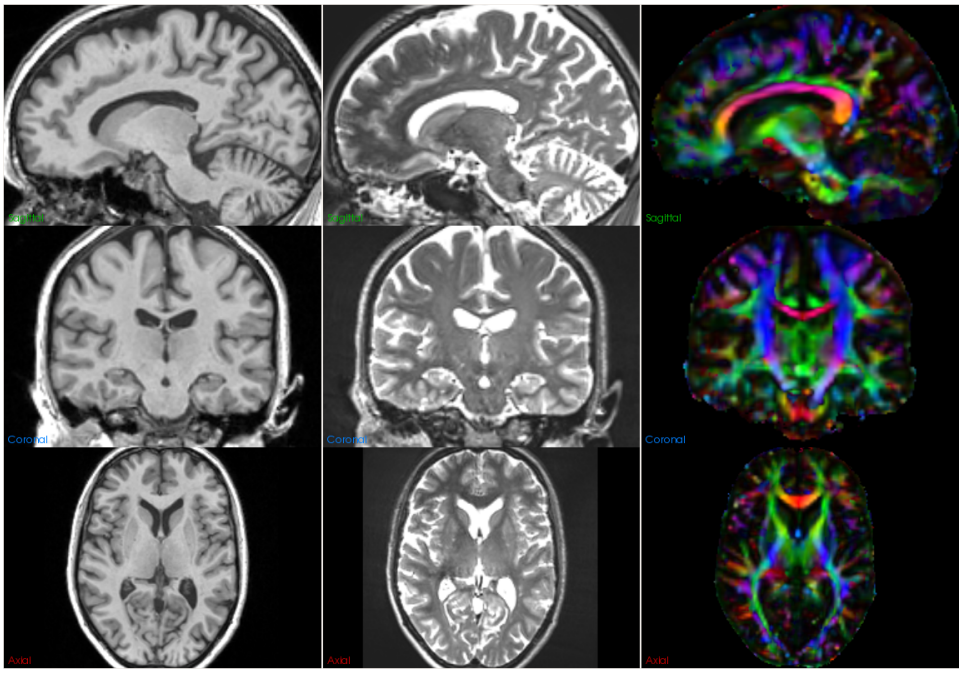


Figure 2.1: Sagittal, coronal and axial slices of a T1-weighted scan (column 1), T2-weighted scan (column 2) and coloured fractional anisotropy map (column 3) through the brain of the same individual. The colours indicate the diffusion direction (red = diffusion along the left-right axis, blue = diffusion along the inferior/superior axis, green = diffusion along the anterior/posterior axis.) and intensity is proportional to fractional anisotropy.

pears as a lighter grey and cerebrospinal fluid appears dark. T1-weighted scans are useful for visualising atrophy patterns. Examples of T1-weighted scans in a healthy participant, an individual with mild cognitive impairment and an individual with AD are shown in figure 2.2. Several visual rating scales have been developed for assessing levels of atrophy on single time-point T1-weighted scans [Scheltens and van de Pol, 2012, Koedam et al., 2011]. Volumetric 3D scans with high resolution and good contrast between grey and white matter can be acquired for measuring volumes and progressive tissue loss.

2.2.2.2 T2-weighted MRI

An example of a T2-weighted MRI scan is given in the second column of figure 2.1. In T2-weighted scans, CSF appears bright, white matter appears dark grey, and grey matter appears light grey. T2-weighted scans are useful for imaging pathology such as oedemas and for assessing white matter damage.

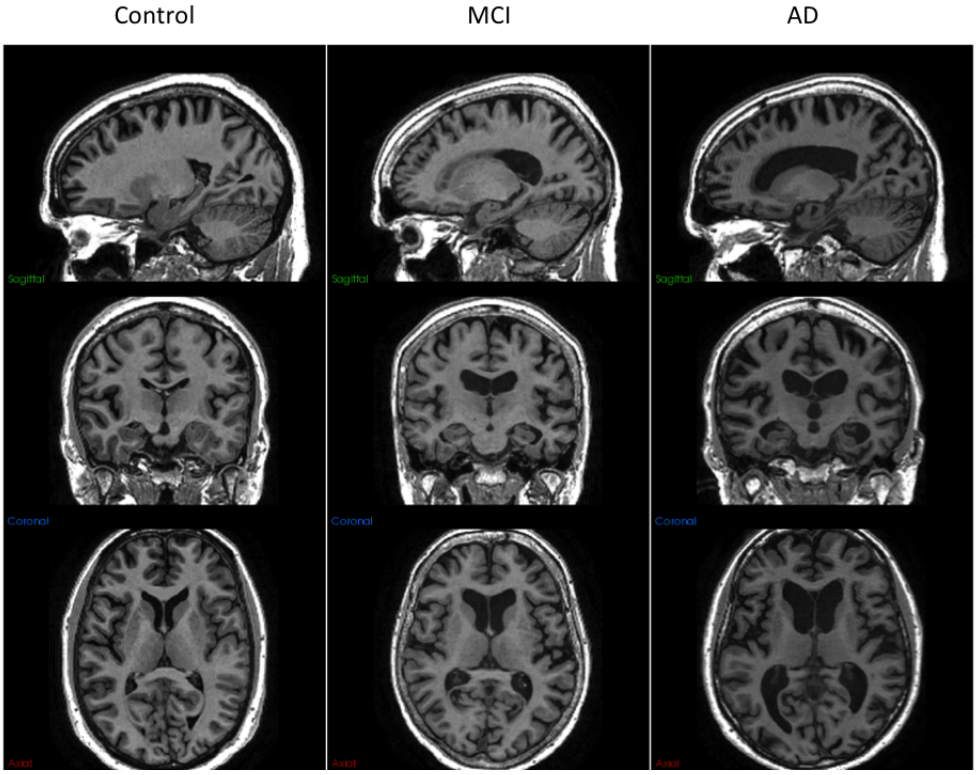


Figure 2.2: Sagittal, coronal and axial views of T1-weighted MRI scans in a healthy individual (column 1), an individual with mild cognitive impairment (MCI) (column 2) and Alzheimer's disease (AD) (column 3).

T1-weighted scans have been preferentially used for measuring brain volumes . Recent advances in MRI pulse sequencing has made volumetric T2-weighted scans feasible only recently [Busse et al., 2008] and so, although it may be possible to use T2-weighted scans in place of T1-weighted scans for assessment of atrophy, volumetric T2-weighted imaging has a less established role in Alzheimer's disease than volumetric T1-weighted imaging.

2.2.2.3 Computer tomography

computer tomography (CT) scans can also be used for visualisation of atrophy patterns and for exclusion of other causes of cognitive deficits. The contrast between grey and white matter is much better on MRI scans than CT scans however, making MRI scans preferable for delineating subcortical structures in the brain and for longitudinal assessment of brain atrophy patterns. An additional advantage of MRI scans over CT scans is that MRI scans do not

expose individuals to ionising radiation.

2.2.2.4 Diffusion tensor imaging

Multiple types of scan with complimentary information can be acquired using an MRI scanner. In addition to structural MRI scans, such as T1-weighted or T2-weighted scans, scans that give information about the diffusion of water molecules in the brain can be acquired. Diffusion tensor imaging (DTI) is a type of MRI scan that gives information about the orientation, location and anisotropy of the brain's white matter tracts. Metrics such as fractional anisotropy (FA) and mean diffusivity (MD) can be derived from diffusion tensor imaging, providing measures of altered tissue structure. FA represents the degree of diffusion anisotropy and is a scalar value between 0 and 1, where 0 represents isotropic diffusion (unrestricted in all directions) and 1 represents diffusion along only one axis. FA is a measure of microstructural integrity and reflects a number of tissue properties such as axonal count and density, degree of myelination and fibre organization [Winston, 2012]. MD, also a scalar value, represents reflects the rate of water diffusion within a voxel, independent of direction. Diffusion weighted imaging has a number of limitations such as it's relatively low resolution and it's inability to correctly characterise diffusion in areas where there is a complex fibre architecture. For example, in a voxel where fibres cross, the FA would be lower than in a voxel where there are fibres oriented in only one direction [Van Hecke et al., 2016]. An example of a coloured FA map, showing both the magnitude and direction of diffusion of water molecules in the brain, is given in column 3 of figure 2.1.

Differences in DTI metrics have been shown in AD when compared to controls [Canu et al., 2012] and may even precede volume loss [Douaud et al., 2011]. Previous work has shown that the addition of DTI metrics to grey matter morphometry in the hippocampus and parahippocampal gyrus improved the ability to distinguish AD subjects from controls [Kantarci et al., 2010]. Other work has shown reduced white matter volumes and fractional anisotropy in the parahippocampal gyrus in AD patients [Wang et al., 2012, Kalus et al.,

2006] and in MCI [Rose et al., 2006] suggesting that disruption in flow of information to the hippocampus may exist at early stages of disease.

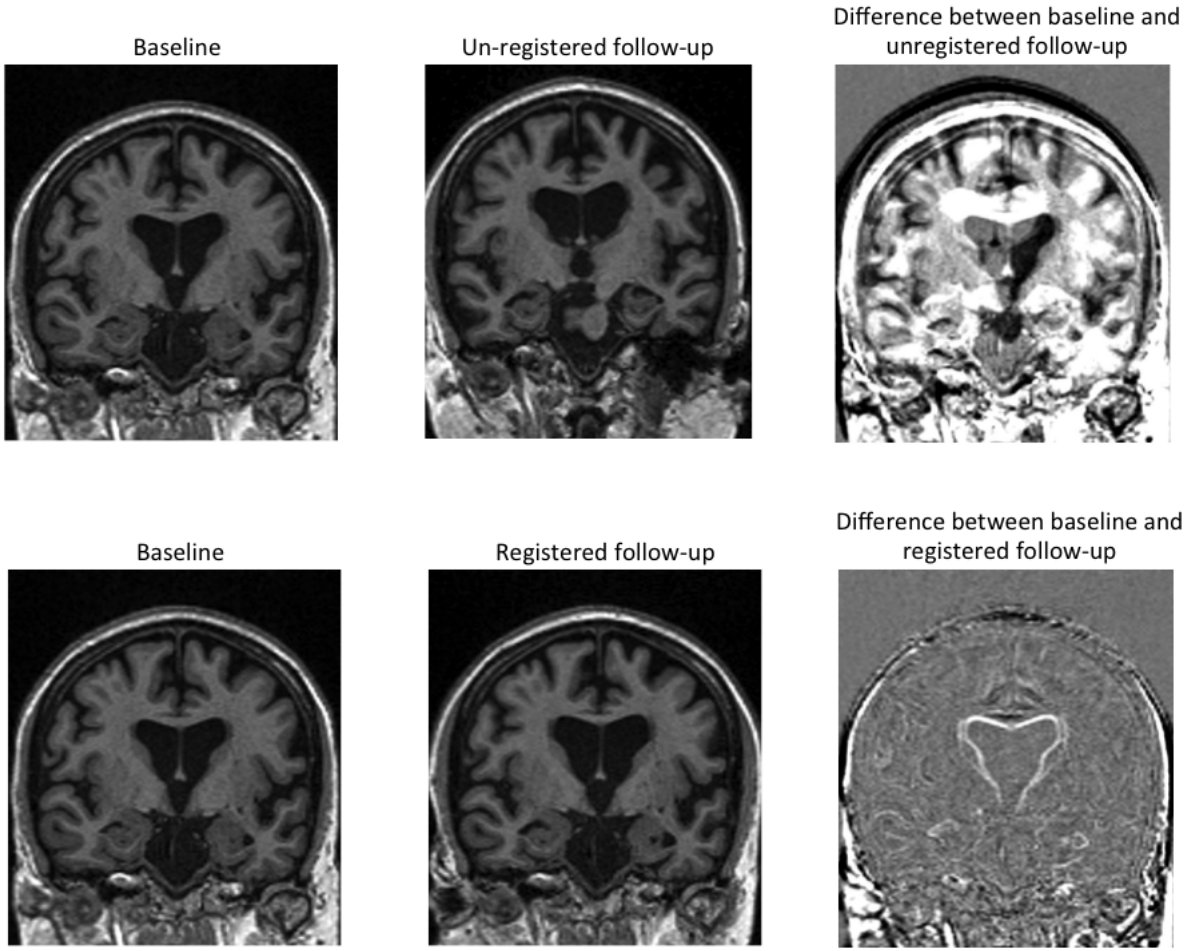


Figure 2.3: Baseline and month 36 follow-up scan. The first row shows unregistered follow-up scan and the difference between the baseline and unregistered follow-up. The second row shows the follow-up scan affinely registered to the baseline and the difference between baseline and registered follow-up scans. By registering the follow-up scan to the baseline, tissue loss can be localised and quantified using the boundary shift integral.

2.2.3 Neuropsychology

Cognitive deficits of Alzheimer's disease can be assessed using neuropsychological tests. Results from these assessments are usually primary outcome measures for clinical trials. Although somewhat subjective, they theoretically more closely associate with symptomatic loss of skills and increase in deficits experienced by patients. Neuropsychological tests have a number of limitations. For example, the mini-mental state exam (MMSE) [Folstein et al., 1975], widely-used in research and the clinic suffers from floor and ceiling effects. The MMSE assigns individuals a score out of 30. Controls and those at the early stages of the disease often score 30 on this test. In addition, test scores such as the MMSE, which is biased towards memory testing, may not accurately capture disease severity in individuals who have atypical forms of Alzheimer's disease where memory problems are less likely to be apparent.

2.2.4 Genetics

Although only a small percentage of Alzheimer's disease cases have a known genetic cause, there are a number of genes that increase the likelihood of developing sporadic Alzheiemer's disease. Arguably, the most important genetic risk factor for sporadic AD is the $\varepsilon 4$ variant of the apolipoprotein (APOE) gene [van Es and van den Berg, 2009]. Of the three common alleles of the APOE gene, $\varepsilon 3$ is most frequent with $\varepsilon 4$ less common and $\varepsilon 2$ relatively rare [Eisenberg et al., 2010]. The $\varepsilon 4$ allele increases the risk of AD and lowers the age of disease onset [Corder et al., 1993]. There is also evidence that the topography of atrophy in $\varepsilon 4$ carriers ($\varepsilon 4$+) may be different from non-carriers ($\varepsilon 4$-) in AD [Agosta et al., 2009, Geroldi et al., 1999, Lehtovirta, 1995, Pievani et al., 2009] although not all studies have confirmed this [Drzezga et al., 2009].

2.2.5 Ordering of biomarker changes in Alzheimer's disease

Changes in biomarkers occur a different disease stages. A hypothetical model for the ordering of biomarker changes in Alzheimer's disease is given in [Jack et al., 2013].

Changes in levels of amyloid in the brain are thought to be the earliest sign of Alzheimer's disease. Abnormal levels of amyloid occur before the loss of brain tissue and possibly many decades before symptom onset. These levels likely approach plateau at an early stage in the disease course [Jack et al., 2013], so may not be the most suitable biomarkers for tracking disease progression once the disease is manifest. This is closely followed by abnormal levels of the tau protein.

Increased brain atrophy occurs after the deposition of amyloid and tau in the brain, but years before clinical onset [Dickerson et al., 2011]. Volumetric MRI scans are therefore ideally placed for tracking disease progression.

Clinical symptoms and cognitive deficits of Alzheimer's disease are thought to become apparent only after the brain starts to atrophy. Neuropsychological tests in Alzheimer's disease are useful diagnostic tools of manifest disease, but are less well placed for disease tracking at an early stage.

2.3 Measuring volumes and volume change using MRI scans

In order to measure regional volumes and to quantify changes in volumes using MRI , the structures or regions require segmentation.

2.3.1 Manual segmentation

Manual segmentation is considered the gold standard technique for region segmentation, however, current databases of patients and control subjects are of a size (total ADNI2 dataset > 700 subjects with multiple scans) where such practices are difficult to implement for the whole study. This problem can be circumvented by using a large number of manual segmentors, however, training can take three months and inter-rater variability may confound the results. In addition, manual segmentations are not always accurate - humans make errors.

2.3.2 Automated segmentation

Some of the most widely-used techniques for automated subcortical brain structure segmentation are atlas-based (FIRST, freesurfer) [Patenaude et al., 2011, Fischl et al., 2002], where a manually labelled brain atlas or atlases are non-linearly warped to the unsegmented target image and the labels are used segment the target brain. Various combinations of different types of atlas (topological [Collins et al., 1998], probabilistic [Shattuck et al., 2008, Hammers et al., 2003]), registration techniques and segmentation strategies (label propagation [Iosifescu et al., 1997, Fischl et al., 2002], multi-atlas label propagation [Aljabar et al., 2009], probabilistic atlas-based segmentation [Awate et al., 2006] etc) are employed in the different segmentation algorithms. The segmentation accuracy of atlas-based algorithms is dependent on the quality of the manual segmentations used in the atlases, as well as the variety of morphology represented in the atlas database and the types of pathologies under investigation. Most existing automated subcortical segmentation methods rely on manual segmentation of T1-weighted images to create the atlases used in such algorithms. However, since not all boundaries of subcortical structures are clearly visible on T1-weighted scans, the manual segmentation accuracy is limited as is the accuracy of the resulting automated subcortical segmentation methods based on these T1-weighted atlases.

2.3.3 Indirect measures of change

Volume change between two scans can be measured indirectly by segmenting the structure of interest on the follow-up and baseline scans, measuring the volume at each time point and subtracting one from the other. The problem with measuring atrophy with this method is that the measurement errors (in the segmentation) can be of the order of the volume changes being measured.

2.3.4 Direct measures of change

Volume change can be measured directly by registering baseline and follow-up scans. Change in volume can be measured directly from registered scan pairs

using techniques such as the boundary shift integral (BSI) [Freeborough et al., 1997]. The boundary shift integral reduces the measurement errors by using the segmentations at baseline and follow-up to register the structure of interest and then calculates how much the boundary of the structure of interest has shifted using intensity windowing directly on the voxel intensities. A graphical representation of the BSI is shown in figure 2.4.

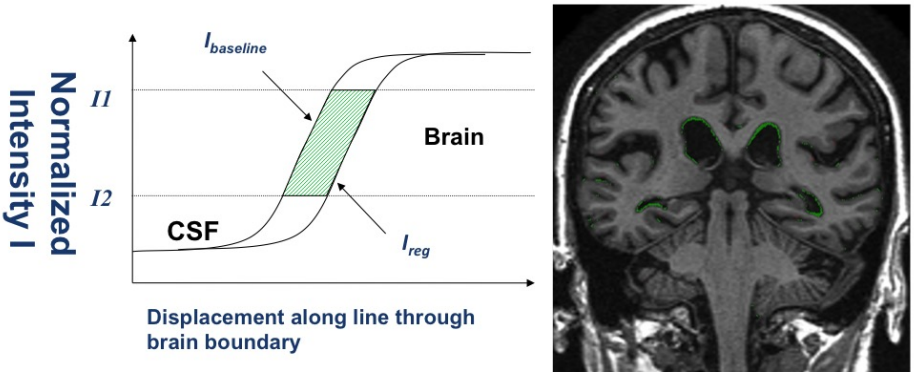


Figure 2.4: The BSI determines the change in volume directly from voxel intensities on registered pairs of MRI scans. Hypothetical intensity profiles of baseline and follow-up scans are shown along a line through the brain boundary. The green hatched area represents the BSI on both the graph and the image of a brain.

Another direct method for measuring change in brain volume is SIENA (Structural Image Evaluation, using Normalisation, of Atrophy) [Smith et al., 2002]. In this method, brain regions are automatically extracted from baseline and follow-up scans and they are then registered. The surface motion (between the two time points) is estimated at brain surface points, to subvoxel accuracy. A percentage brain volume change estimate can then be calculated from the mean perpendicular edge motion across the entire brain surface. This method of measuring brain volume change has been shown to have strong correspondence to the BSI [Smith et al., 2007].

Non-linear registration methods can also be used to measure brain volume change directly [Boyes et al., 2006]. In these methods, an affine registration is followed by a non-linear registration. From the non-linear registration, jacobian matrices for each voxel are obtained. The determinant of each of these

matrices represents the expansion or contraction of each voxel. These can then be integrated over a brain region to produce an estimate of atrophy.

By treating the baseline scan as a reference scan for follow-up scans, some techniques for measuring brain volume change directly have been found to be biased and show asymmetry. A number of techniques for ensuring symmetrical measures of atrophy have since been employed [Fox et al., 2011].

A head-to-head comparison of various direct methods for measuring volume change over time and an assesment of their bias is reported in [Cash et al., 2015].

2.3.5 Structures commonly segmented in AD

Although whole-brain, ventricular and hippocampal volumes are widely-used as AD biomarkers, little attention has been given to the role of other subcortical grey matter structures in AD but there is evidence to suggest that structures such as the parahippocampal gyrus [Braak and Braak, 1995, Echávarri et al., 2011], amydala [Vogt et al., 1990, Poulin et al., 2011], thalamus [Braak and Braak, 1991, De Jong et al., 2008] and putamen [De Jong et al., 2008] also atrophy in AD. Some of the boundaries of these structures are difficult to distinguish from surrounding tissue on T1-weighted MRI scans meaning the accurate segmentation of these regions remains a challenge.

2.3.6 Multimodal segmentation: a way of investigating subcortical structures?

Different MRI modalities provide complimentary information about tissue properties and its microstructure including DTI. Subcortical regions, such as the thalamus, have well-structured fibre connections to defined cortical and subcortical areas. DTI allows the delineation of axonal tracts within white matter, unlike T1-weighted MRI. As a result, DTI imaging may offer enhanced intrinsic contrast in some subcortical areas such as the thalamus. Incorporation of DTI into manual segmentation protocols may allow for more accurate segmentations of subcortical regions and sub-regions by combining

information from DTI and T1-weighted scans. Indeed fusion of DTI and T1 weighted scans has shown promising results for the improvement of the segmentation of grey matter, white matter and cerebrospinal fluid (CSF) [Awate et al., 2008]. In addition, several reports have demonstrated techniques for segmenting sub-regions of the thalamus [Mang et al., 2012, Behrens et al., 2003, Jbabdi et al., 2009, Johansen-Berg et al., 2005, Unrath et al., 2008, Wiegell et al., 2003b, Ziyan et al., 2006, Ziyan and Westin, 2008] and the sub-regions of the amygdala [Solano-Castiella et al., 2010, Solano-Castiella et al., 2011] using information from DTI. These techniques focused on the segmentation of sub-regions however and used masks (either manually or automatically generated) based on T1-weighted scans to define the limiting outer boundaries of those structures. Only one study to date has used information from both T1-weighted scans and DTI scans to automatically segment subcortical structures [Marrakchi-Kacem et al., 2010]. This study describes a hybrid multicontrast deformable model segmentation algorithm that incorporates information from both T1-weighted and DTI scans. They reported increased segmentation accuracy when including information from DTI into the algorithm but only compared the segmentation results with and without DTI data to manual segmentations performed on T1-weighted scans alone. Whether information from both DTI and T1-weighted MRI scans can be successfully combined to improve the segmentation of the thalamus remains to be thoroughly investigated.

Chapter 3

Methods

3.1 Subjects

3.1.1 Alzheimer's disease neuroimaging initiative (ADNI)

Data used in chapters 4, 7 and 9 were downloaded from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see `www.adni-info.org`. Participants underwent baseline and periodically repeated clinical and neuropsychometric assessments and MRI. Detailed inclusion criteria for the ADNI study can be found at `http://www.adni-info.org/Scientists/doc/ADNI_Protocol_Extension_A2_091908.pdf`.

There are several phases to the ADNI study. ADNI-1 was the first phase of the ADNI study. Participants with mild AD, MCI and controls were enrolled in this phase of the study and most subjects recruited in this phase of the study were scanned on 1.5T MRI scanners. ADNI-GO was the second phase of the ADNI study. In this phase, as well as continuing to collect data from subjects recruited as part of ADNI-1, 200 subjects with early mild cognitive impairment

(EMCI) were recruited and were scanned using 3T MRI scanners. ADNI-2 was the third phase of the ADNI study. In this phase new controls, EMCI, late mild cognitive impairment (LMCI) and mild AD subjects were recruited and scanned on 3T scanners. Wechsler Memory Scale Logical Memory II scores (adjusted for years of education) were used to define individuals with MCI as early or late MCI. For details see `http://www.adni-info.org/Scientists/doc/ADNI2_Protocol_A3_17Oct2014_CLEAN.pdf`.

3.1.2 Young onset Alzheimer's disease (YOAD)

The patients in this cohort were recruited prospectively from a specialist Cognitive Disorders clinic between 2013 to 2015. A total of 45 patients, who all had a symptom onset at less than 65 years and who met consensus criteria for probable AD [McKhann et al., 2011] were recruited. None had a known mutation or family history suggestive of autosomal dominant inheritance. Patients were classified as having a typical [McKhann et al., 2011] or atypical (posterior cortical atrophy) [Tang-Wai et al., 2004]) AD phenotype according to published criteria. One patient had an atypical frontal variant of Alzheimer's disease. 24 age- and gender- matched control subjects were also recruited.

3.1.3 AVID cohort

The patients in this cohort were recruited as part of a pilot study to demonstrate brain amyloid using Florbetapir F18, between 2012 and 2013. A total of 22 participants were recruited, with a range of pathologies: 5 with semantic dementia (SD), 5 with posterior cortical atrophy (PCA), 3 with logopenic aphasia (LPA), 4 with progressive nonfluent aphasia (PNFA) and 5 healthy age-matched controls.

3.1.4 Posterior Cortical Atrophy cohort

124 subjects were identified retrospectively from a clinical database at the Dementia Research Centre (PCA (n=47), typical AD (tAD; n=29), and control subjects (n=48)). All subjects have been described in a previous study of grey matter volume and cortical thickness in PCA and typical Alzheimer's disease

(tAD) [Lehmann et al., 2011]. Subjects required at least one T1 weighted volumetric MRI scan to be included in the study. All PCA patients fulfilled the clinical criteria for posterior cortical atrophy proposed by [Mendez et al., 2002] and [Tang-Wai et al., 2004] and more recent [Dubois et al., 2014a] criteria for atypical AD, including evidence of posterior cortical dysfunction on neuropsychological assessment and atrophy on MRI. In addition subjects were only included in the PCA group if there was no indication of another underlying pathology (such as dementia with Lewy bodies (DLB)). Although the neuropsychological tests completed were not identical across all individuals, all PCA patients showed evidence of deficits (scored <5th percentile) in at least two tasks sensitive to parietal dysfunction - object perception (VOSP Object Decision test [Warrington and James, 1991b, Warrington and James, 1991a]), spelling (Graded Difficulty Spelling test [Baxter and Warrington, 1994]), space perception (VOSP Number Location test [Warrington and James, 1991b]) and calculation (Graded Difficulty Arithmetic test [Jackson and Warrington, 1986]) - and had relatively preserved episodic memory (>5th percentile on verbal and or visual Recognition Memory Tests [Tosun et al., 2016, Warrington, 1996]). Those included in the typical AD group fulfilled revised NINCDS-ADRDA criteria for probable AD [Dubois et al., 2007, McKhann et al., 1984] and had significant episodic memory impairments (namely gradual and progressive change in memory function, objective evidence of significantly impaired episodic memory, and presence of medial temporal lobe atrophy) with episodic memory impairments quantified as performance <5th percentile on the verbal and visual Recognition Memory Tests [Tosun et al., 2016, Warrington, 1996]).

3.2 MRI acquisition

3.2.1 ADNI

3.2.1.1 ADNI-1

The ADNI-1 MRI protocol used in this study is described elsewhere [Jack et al., 2008]. Most participants in this phase of the study were scanned on

1.5T scanners. Two (unaccelerated) T1-weighted MRI scans (Magnetization Prepared RApid Gradient Echo (MPRAGE)) were acquired at each session. The image of highest quality (as assessed by a single quality control centre) was selected. Pre-processing corrections were then applied depending on the scanner manufacturer and head coil used: 1) correction for image geometry distortion due to gradient non-linearity (gradwarp) [Jovicich et al., 2006], 2) B1 non-uniformity correction [Narayana et al., 1988] and 3) intensity non-uniformity correction (N3 histogram peak sharpening) [Sled et al., 1998]. All pre-processing was performed by the Mayo clinic before being downloaded locally. After pre-processing, the scans were additionally visually inspected at the Dementia Research Centre for motion artefacts.

3.2.1.2 ADNI-GO/2

Some changes to the imaging protocol used were made in the ADNI-GO and -2 phases of the ADNI study. Participants enrolled in this phase of the study were scanned on 3T scanners. All participants had both accelerated and non-accelerated T1-weighted MRI scans acquired at each session. Those who were scanned on GE scanners also had diffusion weighted imaging scans acquired. Details of imaging protocols can be found at `http://adni.loni.usc.edu/methods/documents/mri-protocols/`. The same pre-processing corrections applied to ADNI-1 scans, described above, were applied to ADNI-GO and -2 scans.

3.2.2 YOAD

All subjects were scanned on a single Siemens Magnetom Trio 3T MRI scanner using a 32-channel phased array head coil. 3D MPRAGE T1-weighted volumetric MRI was acquired for each participant (acquisition time 9min 23s, echo time (TE)/TR/TI = 2.9/2200/900ms, dimensions 256 x 256 x 208, voxel size 1.1 x 1.1 x 1.1mm). Two identical diffusion-weighted imaging acquisitions were performed using a single-shot, spin-echo echo planar imaging sequence (64 diffusion-weighted directions, b = 1000 s/mm2; 9 b = 0 s/mm2 images

(referred to as 'b0' images); 55 slices; voxel size 2.5 × 2.5 × 2.5 mm3; TR/TE = 6900/91 ms;).

3.2.3 AVID

MRI scans were acquired on a single 3T Siemens TIM Trio scanner using a 32-channel phased array head-coil. Anatomical data included a sagittal 3D MPRAGE T1-weighted volumetric MRI (acquisition time 9min 23s, TE/TR/TI = 2.9/2200/900ms, dimensions 256 x 256 x 208, voxel size 1.1 x 1.1 x 1.1mm) and a coronal T2 Fluid-Attenuated Inversion Recovery (FLAIR) sequence (TE/TR/TI = 87/9000/2500ms, voxel size of 0.9375 x 0.9375 x 5mm). Two identical diffusion-weighted imaging acquisitions were performed using a single-shot, spin-echo echo planar imaging sequence (64 diffusion-weighted directions, b = 1000 s/mm2; 9 b = 0 s/mm2 images (referred to as 'b0' images); 55 slices; voxel size 2.5 × 2.5 × 2.5 mm3; TR/TE = 6900/91 ms;).

3.2.4 Posterior cortical atrophy

T1 weighted volumetric MR scans were acquired for all subjects on 1.5T Signa scanners (General Electric, Milwaukee). All scans used an inversion recovery sequence and all but 7 of the scans consisted of 124 contiguous 1.5 mm coronal slices through the head. The remaining 7 consisted of 120 contiguous 1.5 mm coronal slices (5 PCA subjects and 2 tAD subjects). Since this was a retrospective cohort, there was some variation in the scan parameters and in-plane resolutions of the MRI scans; See Table 3.1 for a breakdown of the imaging parameters by diagnostic group. The majority of subjects in each diagnostic group had an in-plane resolution of 0.9 mm x 0.9 mm (including all 7 subjects with 120 coronal slices).

Table 3.1: MRI scan parameters by diagnostic group

MRI parameters		Controls[a] (n=48)	PCA (n=47)	tAD (n=29)
field of view (FOV) (mm)		200 - 280	200 - 280	200 - 280
TR(ms)		11.7 - 15	11.7 - 15	13.6 - 15
TE (ms)		4.2 - 5.4	4.2 - 6.4	4.2 - 5.4
TI (ms)		650	650	650
Flip angle (degree)		13 - 20	13 - 20	15 - 20
% Phase FOV		75 - 100	75 - 100	75 - 100
Slice thickness (mm)		1.5	1.5	1.5
No. of subjects by in-plane resolution (mm)	0.9 x 0.9	35	31	21
	1.1 x 1.1	4	1	3
	0.8 x 1.0	7	14	5
	0.8 x 1.3	0	1	0

[a] Scan parameters not available for 2 subjects.

3.3 Image processing

3.3.1 Diffusion processing pipeline

In order to generate fractional anisotropy and mean diffusivity maps from the acquired diffusion weighted images (diffusion weighted imaging (DWI)s) the following steps were taken. An iterative groupwise registration and averaging of the B=0 images was performed using NiftyReg [Modat et al., 2010], followed by registration of individual DWIs to the average B=0 image (for motion/eddy current correction) and B vectors adjusted for rotation. Echo-planar imaging (EPI) susceptibility distortion correction was performed using the field map images [Daga et al., 2013], and all DWI and B=0 images resampled to the average B=0 space in a single step using a composition of the registration and susceptibility correction transforms. Finally, tensor fitting was performed using weighted least squares fitting implemented in NiftyFit [Melbourne et al., 2016].

An intracranial mask was generated for the T1-weighted scans using SPM12's segmentation toolbox and combining grey matter, white matter and cerebrospinal fluid classes. This was propagated to the average B=0 space by

affine registration of the T1 and the B=0 image upsampled to 1.5mm isotropic using NiftyFit, followed by nonlinear registration of the mask to the average B0 and combination of the affine and non-linearly transformed masks (to ensure complete brain coverage before and after susceptibility correction). This mask was then used during the susceptibility correction and tensor fitting steps.

To produce diffusion images matching the T1 image, the fitted diffusion tensor and average B=0 images were resampled using the inverse of the T1 to B=0 affine registration. Diffusion metrics for the T1 space were derived from the resampled tensor volume.

3.4 Brain region segmentation

3.4.1 Brains

Brain regions were delineated by Multi-Atlas Propagation and Segmentation (Multi-Atlas Propagation and Segmentation (MAPS)) [Leung et al., 2011]. The whole-brain MAPS technique uses a template library of semi-automatically segmented whole-brain regions (comprised of grey and white matter containing voxels with the brain-stem included up until the most inferior slice containing cerebellum). The MAPS technique works by comparing the target image to these templates and the best-matched templates are then combined to generate the segmentation of the target image. Segmentations were performed in native space.

3.4.2 Ventricles

Ventricles (including the temporal horn and excluding the third and fourth ventricles) were segmented using a semi-automated technique. First the scans were registered to standard space (using a 9-dof-6 approach: 9 degrees of freedom (dof) registration from which the rigid body transform is extracted and applied to image). An upper threshold of 60% of the mean brain intensity was then applied to separate brain tissue from the cerebrospinal fluid. Manual editing was then performed to delineate the ventricular boundaries where required.

3.4.3 Hippocampi

Two methods were used for automatic hippocampal segmentation in this thesis. Both methods use the same template library of 110 manually segmented hippocampi (right and left from 55 subjects), unless otherwise stated. Details of the template library used is given in [Barnes et al., 2008]. In both methods, baseline scans are first registered to standard space and any follow-up scans are then affinely registered to their baseline.

3.4.3.1 hippocampal multi-atlas propagation and segmentation (HMAPS)

HMAPS [Leung et al., 2010a], or hippocampal multi-atlas propagation and segmentation, then uses non-linear registration of the best matched templates from the template library to generate multiple segmentations, and combines them using the STAPLE [Warfield et al., 2004] algorithm.

3.4.3.2 Similarity and Truth Estimation for Propagated Segmentations (STEPS)

The STEPS, or similarity and truth estimation for propagated segmentations, pipeline is an improvement to the HMAPS pipeline described above [Cardoso et al., 2013]. Rather than using a global metric for ranking similarity, as HMAPS does, a local similarity metric on a voxel-by-voxel basis is used for improved segmentation accuracy.

3.4.4 Total intracranial volume

Total intra-cranial volume (total intracranial volume (TIV)) was estimated by summing the volumes of grey matter, white matter, and cerebrospinal fluid (cerebrospinal fluid (CSF)) segmentations using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8). Brain-to-TIV ratio was calculated by dividing the extracted whole-brain volumes by the extracted TIVs.

3.5 Volume change measurement

3.5.1 Brains

For whole brain volume change, follow-up scans were registered to the baseline scans using affine registration (12dof) and differential bias correction was applied. Volume change between follow-up and baseline was calculated using the robust boundary shift integral (robust boundary shift integral (KN-BSI)) [Leung et al., 2010b].

3.5.2 Ventricles

Likewise, for measuring ventricular volume change, follow-up scans were registered to the baseline scans (which were registered to standard space prior to segmentation) using affine registration (12dof). Volume change between follow-up and baseline was calculated using the fixed window boundary shift integral (BSI) for the ventricles [Freeborough et al., 1997].

3.5.3 Hippocampi

Local 6-dof registration was performed separately for left and right side hippocampi after segmentation using the hippocampus regions dilated by 2 voxels. Volume change between follow-up and baseline was calculated using the double window BSI for the hippocampi [Leung et al., 2010b].

3.6 Software

3.6.1 For manual or semi-automated region segmentation

MIDAS (Medical Information Display and Analysis System) software [Freeborough et al., 1997] was used for manual segmentations of brains, ventricles and hippocampi used in some of the chapters in this thesis. This software allows for the simultaneous multi-planer display of 3D data.

NiftyMIDAS, a new software tool under development by the Centre for Medical Image Computing at UCL, was used for the manual segmentation of the thalami in this thesis. In addition to being able to view 3D data in all 3 planes, this software also allows for the simultaneous segmentation of multi-

modal data. This feature was utilised for the manual thalamic segmentations, where T1-weighted and fractional anisotropy maps were simultaneously displayed and segmented.

3.6.2 Stata

Stata version 12 (Stata Corporation, College Station, TX, USA) was used for most of the statistical analyses performed in this thesis.

3.6.3 Matlab

Matlab (matrix laboratory) was used, in conjunction with the SurfStat [Worsley et al., 2009] toolbox for statistical shape analysis in chapter 6.

Chapter 4

Assessment of image quality in accelerated T1-weighted scans

4.1 Introduction

Atrophy rates, as measured using the boundary shift integral (BSI) have been shown to be sensitive biomarkers of Alzheimer's disease (AD) and have been used as outcome measures in a number of clinical trials [Fox et al., 2005, Jack et al., 2008, Salloway et al., 2014]. However, the quality of each scan in the longitudinal series is important to provide robust and accurate results. Poor quality scans can be caused by patient-related factors such as movement during the scanning process.

A typical 3T 3D structural brain magnetic resonance imaging (MRI) with 1 mm resolution takes approximately 9 minutes but some subjects have difficulty remaining still for this time period, often resulting in unusable scans. Acquisition times can be reduced by employing parallel imaging techniques (e.g. with scan times 5 minutes), potentially reducing motion artefacts, scanning costs, as well as increasing patient comfort and compliance. Reducing scanning time may mean that a higher proportion of subjects recruited to a study or clinical trial have usable MRI scans.

However, accelerated acquisitions alter scan characteristics such as signal-to-noise ratio, noise distribution and tissue contrast. Before accelerated T1

scans can be used in place of unaccelerated T1 scans in clinical trials, it is essential that we understand how accelerated acquisitions may affect cross-sectional volumes and longitudinal atrophy rate measures. In addition, despite poor quality MRI scans routinely being excluded from studies (for reasons such as head coil failure, geometric distortions or patient motion), the question of whether subject characteristics differ between those who have usable MRI scans with those who have poor quality scans (due to patient motion) needs to be evaluated.

Therefore, the aims in this study were to: 1) compare whole brain, ventricular and hippocampal volumes at baseline and atrophy rates (measured using the BSI) over 6-month and 12-month intervals in accelerated and unaccelerated scans, 2) investigate whether there was a difference in the number of good quality accelerated and unaccelerated scans, 3) assess whether subject characteristics differed between subjects whose scans were considered unusable due to motion compared with those whose scans were of good quality and 4) compare estimated sample size requirements for a hypothetical clinical trial when using accelerated and unaccelerated scans.

4.2 Materials and Methods

4.2.1 Image data and acquisition

Data used in this chapter were downloaded from Alzheimer's disease neuroimaging initiative (ADNI) (as described in 3). ADNI-GO and ADNI-2 included both accelerated and non-accelerated acquisition protocols for each participant. At the time of downloading baseline MRI scans were available for a total of 884 subjects, 6 month scans for 572 subjects, and 12 month scans for 384 subjects. All subjects used in the preparation of this chapter were scanned on 3T scanners and were new recruits to ADNI-GO/2. Image preprocessing included post-acquisition correction of gradient warping [Jovicich et al., 2006] and intensity non-uniformity correction using N3 [Sled et al., 1998] and SPM5 with tissue priors from a custom template consisting of 400

elderly subjects (200 NC and 200 AD) from the first phase of ADNI. Scanners from three different manufacturers (Philips, Siemens and General Electric) were in use across the different sites. The three different scanner manufacturers use different acceleration protocols; details of the various MRI protocols are listed on the ADNI website (`http://adni.loni.usc.edu/methods/documents/mri-protocols/`). Non-accelerated scans were always acquired prior to the accelerated scans during the same scanning session. Subject demographics for all subjects with both accelerated and non-accelerated scans available at baseline, 6- and 12- months are shown in Table 4.1.

4.2.2 Quality control (quality control (QC)) of MRI scans

First, imaging sites were instructed to immediately assess the quality of T1 weighted scans and to re-acquire if necessary; therefore in some cases more than one accelerated or non-accelerated T1 weighted scan was acquired in a particular session. QC was performed on all MRI scans at the Mayo Clinic before being released for download. This QC procedure entailed assessment of MRI scans for adherence to the ADNI scanning protocol, medical abnormalities and severe artefacts (such as metal-induced artefacts, head-coil failure etc.). Only scans that passed this initial QC are available for download and so all analyses in this paper were only performed on the subset that passed initial QC by Mayo. At the dementia research centre (DRC), UCL, further quality assessment of the MRI scans available to download was performed: individual scans were assessed visually for artefacts and very poor quality scans were excluded from further analysis (e.g. due to severe motion artefacts); once the scans were segmented and scan pairs registered (as detailed below), a single rater (EM), blinded to BSI values, assessed the quality of registered scan pairs. If a scan failed at this QC stage (from here on referred to as DRC QC), it was considered as failed for all BSI measures.

The BSI measures how much the boundary of a brain region has shifted between successive scans using normalised voxel intensities of co-registered images. Blurring due to patient motion can change the intensities of voxels

Table 4.1: Subject Demographics

	cognitively normal (CN)	early mild cognitive impairment (EMCI)	late mild cognitive impairment (LMCI)	AD
Baseline				
Subjects, n	231	310	173	126
Age, years	73.1 (6.1)	70.8 (7.3)	72.2 (7.5)	75.1 (7.7)
MMSE score/30	29 (1)	28 (2)	28 (2)	23 (2)
Education, years	17 (3)	16 (3)	17 (2)	16 (3)
% female	54%	46%	46%	43%
% apolipoprotein (APOE) 4 carriers	31%	46%	61%	71%
Baseline to 6 months				
Subjects, n	157	235	129	52
Age, years	73.6 (6.1)	71.0 (7.2)	72.0 (7.6)	75.6 (8.0)
MMSE score/30	29 (1)	28 (2)	28 (2)	23 (2)
Education, years	17 (3)	16 (3)	17 (2)	15 (3)
% female	50%	45%	48%	38%
% APOE 4 carriers	29%	42%	57%	77%
Baseline to 12 months				
Subjects, n	122	122	103	37
Age, years	74.0 (6.0)	70.7 (7.1)	71.9 (7.4)	75.9 (7.9)
MMSE score/30	29 (1)	28 (2)	28 (2)	23 (2)
Education, years	17 (3)	16 (3)	17 (2)	15 (3)
% female	52%	44%	47%	27%
% APOE 4 carriers	27%	44%	56%	78%

Note: mean (sd) shown unless otherwise indicated
CN = cognitively normal, EMCI = early mild cognitive impairment, LCMI = late mild cognitive impairment, AD = Alzheimer's disease, MMSE = mini-mental state examination (at baseline)

at the brain boundary and may therefore render the BSI measure unreliable [Preboske et al., 2006]. Since subjects had more opportunity to move during the longer non-accelerated scan a strict quality control process was used (i.e. erring on the side of excluding scans) as the first aim was to investigate whether the image acquisition, rather than differences in motion, might influence BSI measures. If motion, geometric distortions (due to different positioning in the scanner) or significant quality differences between baseline and follow-up scans were found, they were rated as unusable. In addition, any subjects whose follow-up scans were performed on a different scanner from the baseline scans were excluded.

4.2.3 Quality Comparison of accelerated and unaccelerated scans

To evaluate quality the proportion of scan pairs that passed DRC QC was examined. Patient motion is one of the major reasons a registered scan pair may fail but other reasons, such as geometric distortions due to positional differences may occur. It was hypothesised that there may be more motion artefacts in non-accelerated scans than accelerated due to the longer time required for the non-accelerated scan, but that other reasons for failing QC such as geometric distortions due to positional differences would be independent of acquisition type. Therefore, in order to compare scan quality both the non-accelerated and accelerated scan pair were excluded if either were failed for reasons other than motion. Within the remaining scan pairs, a paired comparison of the proportion of accelerated scan pairs that passed DRC QC and the proportion of non-accelerated scan pairs that passed DRC QC was conducted using the McNemar test [McNemar, 1947]. Where more than one scan was acquired per session, the first scan to be acquired was used in this comparison (in the BSI comparisons the best scan acquired was used).

4.2.4 Segmentation and volume measurement

Whole-brains were segmented using Multi-Atlas Propagation and Segmentation (MAPS) as described in chapter 3. They were visually checked and manually edited if necessary (edits included removing skull inclusions, spillages into non-brain tissue around the temporal lobes and cutting-off the brain stem at the most inferior slice of the cerebellum). Differences in the numbers of scans with brain, ventricle and hippocampal volumes available were due to the work-flow at the Dementia Research Centre, whereby scans are released for segmentation in batches and all regions are manually checked and edited where necessary. This means that not all available scans had hippocampal regions segmented and checked at the time of writing. All available scans were included to maximise the numbers for the analyses.

4.2.5 Volume Change Measurement

Whole brain, ventricular and hippocampal volume change in 6- and 12- month follow-up scans was calculated using the methods described in chapter 3.

4.2.6 Comparison of accelerated and unaccelerated volumes and atrophy rates

Paired t-tests were used to compare baseline mean volumes and 6- and 12-month atrophy rates between accelerated and unaccelerated scans. Differences in variances were tested for using Pitman's test for equality of variance in paired samples.

To compare atrophy rates between accelerated and non-accelerated scan pairs a family of linear mixed models developed for the analysis of repeated measures of direct change [Frost et al., 2004] were used. Only scans that passed DRC QC were included in this analysis. These models account for the correlation between repeated atrophy measures and permit inclusion of all available atrophy measures in the analysis under the assumption that missing values are missing at random. The dependent variables in separate models were the ml loss of brains, ventricles and hippocampi as calculated by the BSI

(see equation below). Time (years) between baseline and follow-up scans was included as a fixed-effect. Interactions terms between scan type and time were included to allow atrophy rate to vary with scan type. Random effects for visit were included to allow participants to have visit specific deviations from a linear trajectory, and random effect for time allow for participant specific differences from the mean rate of atrophy. The random effect of time was allowed to differ by diagnostic group to allow for differing between subject heterogeneity in atrophy rate. The model can be written as equation (1):

$$y_{ijk} = (\beta_0 + \beta_1 scantype + \beta_2 EMCI + \beta_3 LMCI + \beta_4 AD + b_{i,C} + b_{i,EMCI} + b_{i,LMCI} + b_{i,AD})t_{ijk} - u_{ij} + u_{ik} + \varepsilon_{a,ijk} + \varepsilon_{na,ijk}$$

$$b_{i,C} \sim N(0, \sigma^2_{b,C}), b_{i,EMCI} \sim N(0, \sigma^2_{b,EMCI}),$$
$$b_{i,LMCI} \sim N(0, \sigma^2_{b,LMCI}), b_{i,AD} \sim N(0, \sigma^2_{b,AD})$$

$$u_{ij} \sim N(0, \sigma^2_u), u_{ik} \sim N(0, \sigma^2_u), \varepsilon_{a,ijk} \sim N(0, \sigma^2_{a,\varepsilon}), \varepsilon_{na,ijk} \sim N(0, \sigma^2_{na,\varepsilon})$$

Where y_{ijk} is the measured change between the j-th and k-th visit for the i-th individual, t_{ijk} is the time interval between visits j and k, scantype is a categorical variable representing scan type (0 if accelerated, 1 if non-accelerated), EMCI is a categorical variable representing EMCI status (0 if not diagnosed as EMCI, 1 if diagnosis = EMCI) and likewise for LMCI and AD, β_0 is the mean atrophy rate in accelerated scans in controls, is the difference in mean atrophy rate between accelerated and non-accelerated scans and are the fixed effects coefficients corresponding to a diagnosis of EMCI, LMCI or AD, terms are the random effect slope for subject i in the relevant diagnostic group and $\varepsilon_{a,ijk}$ and $\varepsilon_{na,ijk}$ are the error terms for accelerated and non-accelerated scans respectively. I hypothesised the residual error (ε_{ijk}) may be differ between non-accelerated scans and accelerated scans, as this represents the measurement error introduced in making the direct BSI measurement of atrophy. It

was assumed that all other variance components, such as between subject heterogeneity in their atrophy rates could not be plausibly influenced by the scan type. Therefore, in order to test for differences in variance between accelerated and non-accelerated scans the likelihood ratio test on nested models was used, differing only in that one model specified with common residual variance (ε_{ijk}) and one model had separate residual variances by scan type ($\varepsilon_{a,ijk}$;$\varepsilon_{na,ijk}$).

4.2.7 Comparison of subject characteristics between passed and failed scans

Baseline mini-mental state exam (MMSE), age and vascular burden- white matter hyperintensities (WMH) volumes (segmented on the Fluid-Attenuated Inversion Recovery (FLAIR) scans at UC Davis using an automated technique) [DeCarli et al., 1999] were compared in subjects with scan pairs that failed QC (where one or both of the 0-6 and 0-12 month scan pairs failed DRC QC due to motion) with those who had only passing scan pairs (either 0-6 or 0-12 months or both if available) using linear regression. For the MMSE score comparison age, gender, years of education and diagnosis were adjusted for. For the WMH volume comparison age, baseline diagnosis and head size were adjusted for. Since WMH volumes and MMSE scores are not normally distributed bootstrapping (with 2000 iterations) was used to calculate bias-corrected and accelerated 95% confidence intervals (confidence interval (CI)) for both of these analyses. For the age comparison baseline diagnosis was adjusted for as the AD subjects were older than the EMCI subjects. The proportion of subjects with cognitive impairment (EMCI, LMCI or AD) was compared between the groups (failed vs non-failed scans) using logistic regression and calculated risk ratios.

4.2.8 Sample size estimates

Sample size requirements for a clinical trial to measure a 25% reduction in whole-brain atrophy rates in accelerated and unaccelerated scans were estimated using the standard formula: sample size per arm = $(u + v)^2 \times (\sigma 1^2$

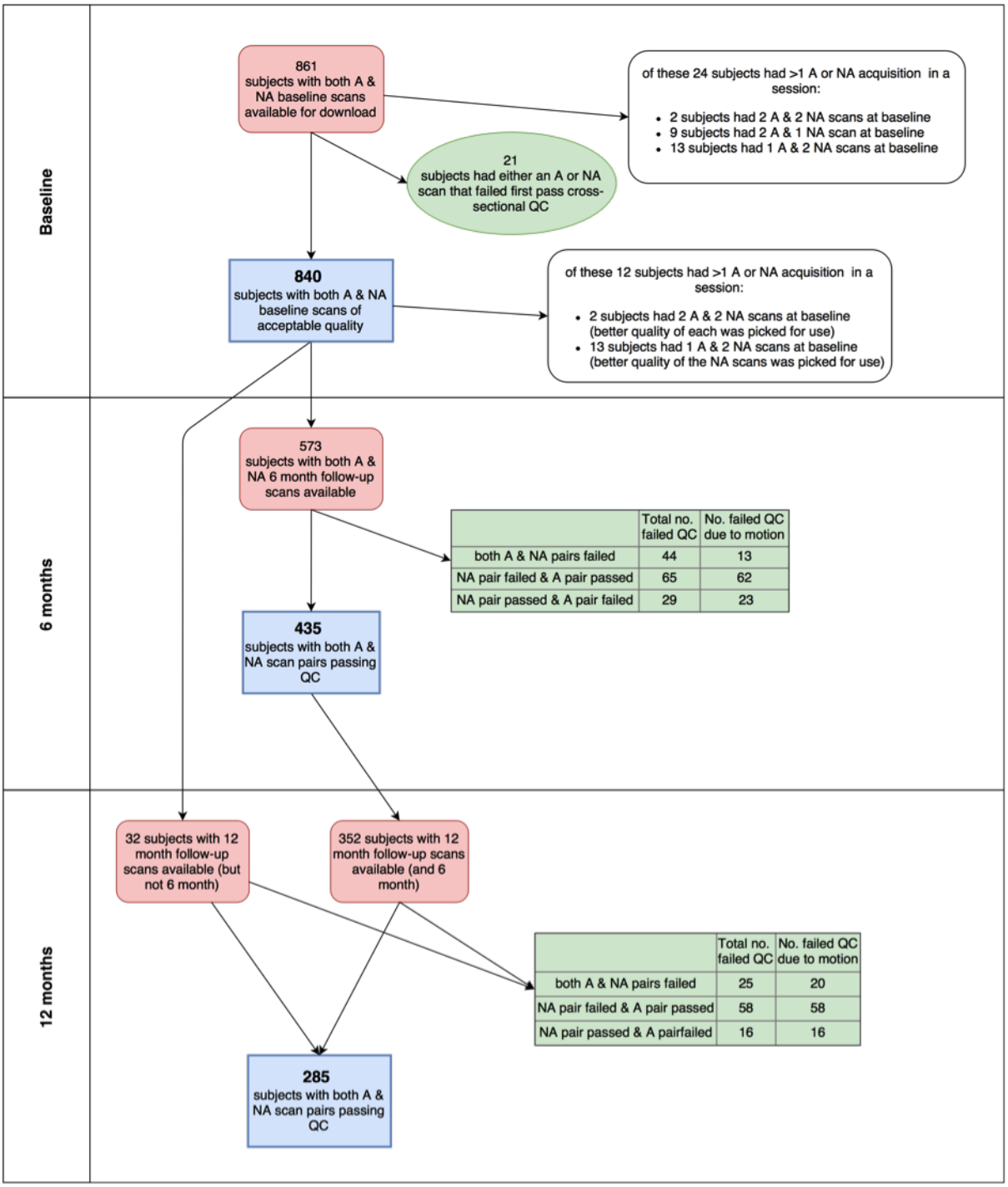
$+ \sigma 2^2)/(\mu 1 - \mu 2)^2$, where u = 0.84 to provide 80% power and v = 1.96 to test at the 5% level; μ and σ are the mean and SDs of rates of atrophy in the treatment and placebo groups (assumes $\mu 1 \approx \mu 2$). In order to directly compare the sample size requirements using accelerated and unaccelerated scans, all 0-6 and 0-12 month scan pairs were included, whether or not they failed DRC QC.

4.3 Results

4.3.1 Quality Control

The majority of scans failed by Mayo were failed due to non-adherence to the ADNI protocol, incomplete coverage of the brain, metal induced artefacts or pathology. Only one subject had scans failed by Mayo due to severe motion. A total of 861 subjects had baseline accelerated and non-accelerated scans available to download from LONI. Of these 840 subjects had both a baseline accelerated and non-accelerated scan that passed DRC first pass QC. See Figure 4.1 for a detailed breakdown of the number of subjects at each stage of the DRC QC.

Of the 576 subjects who had scans available at 6 months, scan pairs from 29 subjects were rated as unusable due to geometrical distortions due to positioning differences in the scanner being detected on either or both of the accelerated and unaccelerated scans. A further 138 subjects had either an accelerated or unaccelerated scan pair rated as unusable due to motion. Of the 385 subjects who had scans available at 12 months, geometric distortion was detected on either the accelerated or unaccelerated scan pairs or both for 5 subjects and were therefore rated as unusable and a further 99 subjects had either an accelerated or unaccelerated scan pair rated as unusable due to motion.



	Total no. failed QC	No. failed QC due to motion
both A & NA pairs failed	44	13
NA pair failed & A pair passed	65	62
NA pair passed & A pair failed	29	23

	Total no. failed QC	No. failed QC due to motion
both A & NA pairs failed	25	20
NA pair failed & A pair passed	58	58
NA pair passed & A pairfailed	16	16

Figure 4.1: Breakdown of no. of subjects at each stage (NA = non-accelerated scan, A = accelerated scan)

4.3.2 Scan quality differences between accelerated and unaccelerated T1 MRI scans

At both 6- and 12- month intervals, significantly more (twice as many) unaccelerated than accelerated scan pairs failed quality control due to motion artefacts (in one or other of the pairs); at 6 months 14% of unaccelerated scan pairs were failed due to motion vs 7% of accelerated scan pairs ($p \leq 0.001$); at 12 months 20% of unaccelerated scan pairs failed due to motion vs 9% of accelerated scan pairs ($p \leq 0.001$).

4.3.3 Baseline volumes

Table 4.2 summarises the mean baseline brain, ventricular and hippocampal volumes measured from the baseline MRI scans. There were no differences in the mean baseline brain or hippocampal volumes measured using accelerated scans as compared with unaccelerated scans. The measured baseline ventricular volume was on average 0.09ml lower for unaccelerated compared to accelerated scans. Although this difference was statistically significant ($p<0.001$) it was very small at only 0.2% of the mean ventricular volume. There were no significant differences in variances of any of any of the measures.

Table 4.2: Baseline volumes (mean (sd) unless otherwise stated)

	n=	Non-accelerated	Accelerated	Mean difference (non-accelerated - accelerated) [95% CI]	Limits of agreement (accelerated - non-accelerated)
Brain volume (ml)	840	1066 (106)	1066 (106)	0.3 [-0.03, 0.6], p=0.07	-9.22 to 9.83
Ventricular volume (ml)	840	39.2 (22.8)	39.3 (22.8)	-0.09 [-0.1, -0.06], p<0.001	-0.81 to 0.63
Hippocampal volume (ml)	645	5.2 (0.03)	5.2 (0.03)	-0.002 [-0.004, 0.008], p = 0.5	-0.15 to 0.15

4.3.4 Atrophy rates

The results of the comparison of annualised brain, ventricular and hippocampal atrophy rates calculated using the BSI from accelerated and non-accelerated MRI scan pairs are shown in Table 4.3. No significant differences in whole-brain, ventricular or hippocampal atrophy rates were found in any of the comparisons. Residual errors were generally slightly higher for non-accelerated scans, although the difference only reached statistical significance in the hippocampal comparison in Philips scanners.

Table 4.3: Comparison of whole-brain ventricular and hippocampal atrophy rates calculated from accelerated and non-accelerated scan pairs using the boundary shift integral in subjects whose accelerated and non-accelerated scan pairs both passed QC, mean (se) shown unless otherwise indicated

	No. scan pairs	Mean adjusted atrophy rate* in accelerated scans, ml/year	Mean adjusted atrophy rate* in non-accelerated scans, ml/year	Mean difference in atrophy rates* (non-accelerated - accelerated) [95% CI]	Residual error: accelerated scans	Residual error: non-accelerated scans	p-value of likelihood ratio test
Brains							
All	487	5.81 (0.88)	5.85 (0.88)	0.04 [-0.28 to 0.36], p=0.802	5.41 (0.57)	6.66 (0.60)	p=0.153
GE	91	6.52 (1.99)	7.12 (2.01)	0.60 [-0.65 to 1.86], p=0.347	12.71 (2.68)	19.16 (3.53)	p=0.153
Siemens	289	6.44 (1.23)	6.44 (1.23)	0.00 [-0.36 to 0.36], p=0.986	0.48 (0.90)	1.11 (0.95)	p=0.358
Philips	107	4.61 (1.99)	4.31 (1.99)	-0.30 [-0.72 to 0.13], p=0.168	2.25 (0.58)	2.91 (0.60)	p=0.506
Ventricles							
All	487	1.30 (0.22)	1.30 (0.22)	0.01 [-0.02 to 0.03], p=0.619	0.01 (0.01)	0.01 (0.01)	p=0.589
GE	91	0.86 (0.52)	0.92 (0.52)	0.06 [-0.03 to 0.15], p=0.225	0.01 (0.03)	0.04 (0.03)	p=0.483
Siemens	289	1.57 (0.27)	1.57 (0.27)	0.00 [-0.03 to 0.03], p=0.998	0.001 (0.005)	0.01 (0.01)	p=0.450
Philips	107	0.92 (0.53)	0.91 (0.53)	-0.01 [-0.07 to 0.04], p=0.647	0.02 (0.01)	0.00 (0.01)	p=0.492
Hippocampi							
All	227	0.05 (0.02)	0.04 (0.02)	-0.01 [-0.02 to 0.00], p=0.262	0.004 (0.000)	0.004 (0.0005)	p=0.795
GE	45	-0.06 (0.09)	-0.08 (0.09)	-0.02 [-0.05 to 0.00], p=0.100	0.003 (0.002)	0.002 (0.002)	p=0.347
Siemens	126	0.07 (0.01)	0.06 (0.01)	-0.01 [-0.02 to 0.00], p=0.125	0.001 (0.001)	0.002 (0.001)	p=0.381
Philips	56	0.03 (0.02)	0.04 (0.02)	0.01 [-0.00 to 0.03], p=0.121	0.002 (0.000)	0.003 (0.001)	p=0.045

* results are from the model with two separate residual errors by scan type

4.3.5 Comparison of subject characteristics between those with failed scan pairs due to motion and those with passed scan pairs

Unadjusted baseline WMH volumes, age, MMSE scores and change in MMSE over 12 months along with the results of the regression analyses comparing these variables in subjects who had one or more scan pairs that failed DRC QC due to motion with those whose scan pairs passed DRC QC are shown in table 4.4.

4.3.5.1 In accelerated scan pairs

Subjects who had a failed 0-6 or 0-12 month accelerated scan pair (due to motion) had a lower mean adjusted WMH volume (adjusted for age, diagnosis and head-size) and had a higher mean adjusted age (adjusted for diagnosis). There was no significant association between the proportion of cognitively impaired subjects and QC failure due to motion, nor was an association found between baseline MMSE score and QC scan pair failure.

4.3.5.2 In unaccelerated scan pairs

A higher proportion (11% higher) of subjects who had a failed 0-6 or 0-12 month unaccelerated scan pair (due to motion) were cognitively impaired compared with those whose scan pairs passed QC (p=0.01). Subjects who had a failed 0-6 or 0-12 month unaccelerated scan pair (due to motion) also had a lower mean adjusted MMSE score (adjusted for age, gender, education and diagnosis) ($p \leq 0.05$). No significant difference in WMH volumes or age was found.

4.3.6 Sample size estimates

Mean annualised brain atrophy rates by diagnostic group are shown in Table 4.5 and estimated sample sizes for subjects with LMCI and AD are shown in Table 4.6. No significant differences in sample size requirements were found when comparing accelerated and unaccelerated scan pairs for any of the atro-

phy measures.

Table 4.4: Comparison of subject characteristics between those with failed scan pairs due to motion and those with passed scan pairs

	Both (if available) 0-6 and 0-12 month scan pairs passed QC	One or both of the 0-6 and 0-12 month scan pairs failed QC	Adjusted difference [95% CI]
Accelerated scan pairs			
WMH volume (mm3)	6.55 (8.96), n=498	9.40 (9.71), n=66	2.01[a] [0.01, 4.62], $p \leq 0.05$
Age (years)	72.2 (7.4), n=513	74.3 (6.7), n=66	2.2[b] [0.4, 4.0], p=0.02
Baseline MMSE/30	27.8 (2.4), n=513	27.7 (2.1), n=66	0.1[c] [-0.3, 0.4], $p > 0.05$
% Subjects cognitively impaired	73%	78%	4% [-6, 15], p=0.45
Non-accelerated scan pairs			
WMH volume (mm3)	6.85 (9.28), n=444	7.01 (8.40), n=120	-0.32[a] [-1.82, 1.49], $p > 0.05$
Age (years)	72.3 (7.2), n=455	72.9 (7.6), n=124	0.9[b] [-0.6, 2.3], p=0.24
Baseline MMSE/30	27.9 (2.3), n=455	27.3 (2.5), n=124	-0.3[c] [-0.6, -0.1], $p \leq 0.05$
% Subjects cognitively impaired	71%	82%	11% [19, 3], p=0.01

[a] Difference in mean WMH volume adjusted for age, diagnosis and headsize using linear regression analysis. CI calculated using bootstrapping. [b] Difference in baseline age adjusted for diagnosis using linear regression analysis [c] Difference in baseline MMSE score adjusted for age, education and diagnosis. CI calculated using bootstrapping.

Table 4.5: Atrophy rates by diagnosis in all subjects (regardless of whether scan pairs passed or failed QC)

		0-6 month atrophy rates			0-12 month atrophy rates	
	n=	Non-accelerated scans	Accelerated scans	n=	Non-accelerated scans	Accelerated scans
Annualised whole-brain atrophy rates ml/year						
CN	157	7.21 (14.65)	6.74 (13.71)	122	7.18 (7.7)	7.37 (7.48)
EMCI	234	7.44 (13.89)	7.52 (14.15)	122	7.82 (8.92)	7.5 (8.35)
LMCI	129	11.48 (15.01)	11.16 (14.25)	103	10.84 (9.68)	11.01 (9.18)
AD	52	16 (15.97)	16.06 (15.06)	37	14.89 (8.24)	14.94 (9.2)
Annualised ventricular expansion rates ml/year						
CN	157	1.31 (2.45)	1.42 (2.59)	122	1.43 (1.54)	1.41 (1.52)
EMCI	234	1.4 (2.47)	1.39 (2.51)	122	1.77 (1.81)	1.76 (1.79)
LMCI	129	2.7 (3.17)	2.7 (3.09)	103	2.94 (2.59)	2.91 (2.59)
AD	52	4.2 (4.21)	4.39 (4.17)	37	4.14 (3.01)	4.16 (2.95)
Annualised hippocampal atrophy rates ml/year						
CN	115	0.08 (0.22)	0.08 (0.22)	92	0.06 (0.11)	0.07 (0.11)
EMCI	180	0.06 (0.17)	0.07 (0.17)	85	0.09 (0.14)	0.1 (0.14)
LMCI	95	0.15 (0.23)	0.14 (0.22)	76	0.13 (0.15)	0.14 (0.14)
AD	38	0.2 (0.16)	0.21 (0.21)	25	0.18 (0.12)	0.18 (0.11)

Note: mean (sd) shown unless otherwise indicated

Table 4.6: Sample size estimates for a 25% reduction in atrophy rate (with bootstrap 95% CIs and 80% power)

		non-accelerated scans, n [95% CI]	Accelerated scans, n [95% CI]	Difference (non-accelerated - accelerated), n [95% CI]
Baseline to 6 months				
LMCI	Brains	420 [257, 858]	414 [258, 759]	6 [-100, 146]
LMCI	Ventricles	349 [241, 692]	334 [226, 633]	15 [-38, 50]
LMCI	Hippocampi	673 [342, 1851]	612 [340, 1707]	61 [-226, 763]
AD	Brains	252 [146, 606]	223 [127, 529]	29 [-65, 248]
AD	Ventricles	254 [160, 466]	228 [145, 427]	26 [-3, 75]
AD	Hippocampi	145 [89, 270]	236 [136, 576]	-91 [-153, 19]
Baseline to 12 months				
LMCI	Brains	204 [139, 371]	177 [127, 277]	27 [-14, 76]
LMCI	Ventricles	202 [148, 320]	205 [149, 329]	-3 [-16, 10]
LMCI	Hippocampi	361 [221, 853]	234 [160, 415]	127 [-16, 435]
AD	Brains	77 [53, 126]	96 [56, 203]	-19 [-47, 23]
AD	Ventricles	134 [89, 244]	127 [85, 232]	7 [-1, 20]
AD	Hippocampi	110 [58, 328]	93 [40, 621]	17 [-64, 172]

4.4 Discussion

A significantly lower proportion of accelerated scan pairs than unaccelerated were found to suffer from poorer quality due to motion artefacts. Further whole-brain, ventricular and hippocampal volumes and atrophy rates calculated from accelerated MRI scans were shown to be comparable to those calculated from unaccelerated MRI scan pairs in a large cohort of control, MCI and AD subjects. Differences were also found in the characteristics of those who failed the accelerated and unaccelerated scan pairs: those who had failed unaccelerated scan pairs had a lower MMSEs at baseline (more severe) and a higher proportion of them were cognitively impaired whereas those whose

accelerated scan pairs failed DRC QC were older and had greater WMH volumes. The difference in WMH volumes remained even after adjustment for age suggesting that subjects with more extensive white matter damage are less able to remain sufficiently motionless even at the shorter scanning times. Therefore, disregarding data from subjects with unusable scans due to motion changes the characteristics of the participants included in any analysis, meaning that they may have different characteristics from those who were originally recruited to the study.

One previous study investigated the influence of parallel imaging acquisition on brain volume measurements in 4 healthy control subjects [Krueger et al., 2012]. They acquired 12 Magnetization Prepared RApid Gradient Echo (MPRAGE) volumes in each session with a range of acceleration factors. They then repeated the same sequences 8 weeks later in each of the 4 subjects and found no significant differences in BSI measures even at high acceleration factors.

A further two previous studies investigated the influence of using accelerated T1 acquisitions on different measures of brain atrophy rate [Ching et al., 2015, Vemuri et al., 2015]. Both of these studies used data from ADNI. One study found that measures of atrophy rate derived from tensor based morphometry were very similar between accelerated and unaccelerated acquisitions [Ching et al., 2015]. Another study [Vemuri et al., 2015], which used a different atrophy rate measure based on symmetric diffeomorphic image normalisation and tensor based morphometry (TBM-Syn), did find some significant differences, with accelerated scan pairs tending to show higher TBM-Syn scores (or lower rates of atrophy) compared with unaccelerated pairs.

There are some limitations to this study. First, subjects always had the unaccelerated scan prior to the accelerated scan. It could be that the finding of more motion artefacts in the unaccelerated scans was due to the ordering of the scans rather than the longer scan time. Secondly, the motion artefacts were visually rated and the rater was not blinded to the type of scan per-

formed, which may have introduced some bias. Thirdly, although the results of this study indicate that atrophy rates measured using the BSI in accelerated scans are not markedly different from those measured in unaccelerated scans, it may be that different techniques for measuring brain atrophy rates are more susceptible to the changes in scan characteristics introduced by parallel imaging techniques. Notably, a small, but statistically significant difference in hippocampal atrophy rates between accelerated and unaccelerated scan pairs at 12 months was found. The signal-to-noise ratio (SNR) ratios in accelerated scans are highest near the head coils, on the surface of the brain and get progressively worse, as the distance from the coils increases [Krueger et al., 2012]. The hippocampal BSI may therefore be more sensitive to the differences in accelerated and unaccelerated scans due to their location deep in the brain. It may be that with longer time intervals, these differences become more apparent and further studies would be required to investigate this.

In summary, the shorter scan time of accelerated scans may reduce the proportion of scans rendered unusable due to motion artefacts. Importantly, the use of accelerated T1 structural MRI scans in place of unaccelerated scans does not appear to have an impact on whole-brain, ventricular and hippocampal volume and atrophy rate (BSI) measures. Therefore it may be advantageous to use accelerated T1 MRI scans rather than unaccelerated scans for assessing brain volumes and atrophy rates (BSI) in clinical trials. In addition, the use of accelerated T1 instead of unaccelerated would permit more time to be devoted to other imaging types without having a negative impact on the derived metrics. Finally, differences in subject characteristics were observed in those subjects whose scan pairs passed DRC QC from those whose scan pairs failed DRC QC due to motion in both accelerated and unaccelerated scans. Disregarding data from subjects who are unable to keep sufficiently still during an MRI scan to produce quality data ultimately biases the characteristics of the subject group in some way, possibly excluding those subjects who are more severe or have a more vascular form of the disease. However, the use

of accelerated T1 volumetric scans rather than unaccelerated T1 volumetric scans means that higher quality longitudinal image can be obtained on a larger proportion of the original study population.

Chapter 5

Investigation into hippocampal atrophy in apolipoprotein (APOE) ε4 carriers

5.1 Introduction

Hippocampal atrophy rate has been proposed as an imaging biomarker for Alzheimer's disease (AD) progression [Barnes et al., 2004, van der Flier and Scheltens, 2009]. However, it is essential to understand how factors might affect hippocampal atrophy rates if this biomarker is to be used most effectively in clinical trials.

Numerous publications have attempted to elucidate whether APOE modifies hippocampal atrophy rates [Chiang et al., 2011, Cohen et al., 2001, Crivello et al., 2010, Hashimoto et al., 2009, Jack et al., 1998, Lo et al., 2011, Moffat et al., 2000, Mori et al., 2002, Morra et al., 2009, Risacher et al., 2010, Schuff et al., 2009, Spampinato et al., 2011, Van De Pol et al., 2007, Wang et al., 2009, Wolz et al., 2010]. Although some studies reported elevated hippocampal atrophy rates in ε4+ in AD, mild cognitive impairment (mild cognitive impairment (MCI)) and control groups, it is possible that the greater hippocampal rates observed could have been attributed to higher concurrent whole-brain atrophy rates and therefore faster disease progression.

To better understand the effect of the APOE ε4 allele on the progression of structural brain changes the aim of this study was to investigate whether

different whole-brain and hippocampal atrophy rates were observed in ε4+ compared with ε4- in AD, MCI and controls. Further, to investigate if there is evidence of higher hippocampal atrophy rates in ε4+ when adjusting for concurrent whole-brain atrophy rates, which has not been examined previously.

5.2 Methods

5.2.1 Subjects

Subjects from Alzheimer's disease neuroimaging initiative (ADNI)-1 who had a baseline magnetic resonance imaging (MRI) scan and at least 1 follow-up scan were included in this study. Details of the ADNI cohort are given in 3.1.1. Each subject underwent APOE genotyping at the screening visit. All demographic information, diagnoses, neuropsychological test scores and APOE genotype data were downloaded from the ADNI clinical data repository.

Since a proportion of MCI subjects will likely not progress to dementia caused by AD, this group is likely to be quite heterogeneous with respect to underlying pathology. As a result, the MCI subjects were dichotomised into those who were observed to progress to a clinical diagnosis of AD within 36 months of baseline and maintained that diagnosis (MCI-P) and those who were stable over the follow-up period (MCI-S). Subjects whose diagnosis changed from MCI to AD and subsequently reverted to MCI during the study were excluded as were subjects whose diagnosis changed from MCI to normal. ε2 carriers (i.e. ε2/ε2, ε2/ε3 and ε2/ε4 subjects) were also excluded from the study as they may have lower hippocampal atrophy rates [Chiang et al., 2010]. There were a total of 840 ADNI-1 subjects available at the time of this study, after exclusions this number reduced to 622 subjects. The number of subjects excluded at each exclusion stage is summarised in Figure 5.1.

5.2.2 Image acquisition and analysis

The ADNI-1 MRI protocol used in this study is described in 3.2.1.1. Those scans with significant motion artefacts were excluded from the current study. Whole-brain and hippocampi were automatically delineated using the

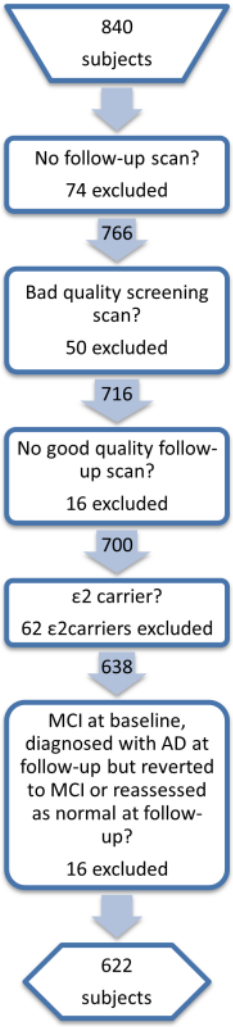


Figure 5.1: Subject selection process

Multi-Atlas Propagation and Segmentation (MAPS) technique from the preprocessed 1.5-T T1-weighted MRI scans at all available time-points [Leung et al., 2010a, Leung et al., 2011]. For details on the MAPS method used see 3.4.1 and 3.4.3.1. The change in the volumes of the whole-brain and hippocampi between follow-up and baseline were calculated using the methods described in chapter 3. Total intra-cranial volume (total intracranial volume (TIV)) was estimated by summing the volumes of grey matter, white matter, and cerebrospinal fluid (cerebrospinal fluid (CSF)) segmentations using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8). Brain-to-TIV ratio was calculated by dividing the extracted whole-brain volumes by the extracted

TIVs.

5.2.3 Statistical Analysis

APOE $\varepsilon 4$ carrier status was coded as 1 for carriers of 1 or 2 $\varepsilon 4$ alleles and 0 for those who did not carry an $\varepsilon 4$ allele. The effect of APOE $\varepsilon 4$ carrier-status on the volume of the sum of the left and right hippocampi at baseline adjusting for the level of overall whole-brain atrophy was investigated. To do this a linear regression was performed within each clinical group with bilateral hippocampal volume as the dependent variable and APOE $\varepsilon 4$ carrier-status, age, gender, mini-mental state exam (MMSE) score, TIV and brain-to-TIV ratio included as covariates. Age was included as a covariate as normal aging is associated with brain volume loss, TIV to control for variation in head size and gender to control for any differences in male-to-female ratio between the different genotype groups. MMSE score and brain-to-TIV ratio were included as covariates in order to assess the effect of the APOE $\varepsilon 4$ carrier-status above and beyond any global differences in cognitive impairment and whole-brain atrophy.

To analyse the effect of the APOE $\varepsilon 4$ carrier-status on the rate of atrophy of the hippocampi and whole-brain (as measured using the boundary shift integral (BSI)), joint linear mixed models were used. These models allow the random-effects dictating the trajectories of hippocampal and whole-brain atrophy to be correlated, thus permitting estimates of hippocampal atrophy rate adjusted for true whole-brain atrophy rate. They allow for repeated measures and accommodate missing values under the missing at random assumption. The dependent variables were the ml loss of hippocampi as calculated by the hippocampal-BSI and brain as calculated by the brain-BSI.

Interval (years) between baseline and follow-up scans was included as a fixed-effect and interactions terms between APOE $\varepsilon 4$ carrier-status and scan interval were included to allow hippocampal atrophy rate to vary with APOE $\varepsilon 4$ carrier-status. Interactions of interval with age, MMSE score, brain-to-TIV ratio, gender and TIV (all measured at baseline) were also included as fixed-

effects in the model. Interval was also included as a random-effect, to allow for between subject heterogeneity in atrophy rate. No constant terms (fixed or random) were included, consistent with the assumption that true (as opposed to measured) atrophy between two scans from the same time-point is zero. A single joint model was fitted to both hippocampal and whole brain losses, allowing distinct fixed and random effect parameters for the two processes. The two trajectories were linked through a correlation between the two random slopes. The difference in mean hippocampal rates between ε4+ and ε4- after adjustment for concurrent brain atrophy rate was then estimated. This was calculated as the difference in hippocampal rates (unadjusted for brain atrophy rate), minus the difference attributable due to differences in brain rates (based on the standard deviations of the random-slopes and their correlation in the joint model). See the appendix for the expressions of the statistical models used.

Since gender was included as a binary categorical variable in the analyses the mean adjusted values for a 50/50 split of males: females is presented in the Figures and Tables (adjusted for disease-group specific mean age, baseline brain-to-total intracranial volume ratio, MMSE score and total intracranial volume). The mean adjusted values for a 50/50 gender split were calculated by multiplying the coefficients for males and females by 0.5 and adding them together. Given that an interaction term between ε4 carrier-status and gender was not included in the analyses, the differences in whole-brain and hippocampal atrophy rates are the same for males and females.

5.3 Results

Table 1 shows demographics and imaging summary statistics for each clinical group used in this study. As previously shown [Leung et al., 2010b], the AD subjects had smaller mean hippocampal volumes at baseline than MCI subjects whose hippocampi were in turn smaller than control subjects (Table 5.1); the mean hippocampal volume for the AD subjects was 20% smaller than the

Table 5.1: Baseline demographics and image summary statistics by clinical group. Mean(SD) unless otherwise stated)

	Controls	MCI stable	MCI progressors	AD
No. Subjects (at 6m, at 12m, at 18m, at 24m, at 36m)	167 (165, 153, 0, 137, 115)	169 (157, 147, 125, 103, 66)	138 (133, 131, 116, 102, 69)	148 (143, 124, 1, 93, 1)
No. ε4 non-carriers (% total), No. ε4 heterozygotes (% total), No. ε4 homozygotes (% total)	118 (71%), 44 (26%), 5 (3%)	86 (51%), 68 (40%), 15 (9%)	42 (30%), 70 (51%), 26 (19%)	44 (30%), 70 (47%), 34 (23%)
% male	54%	66%	59%	55%
Age [years]	76.0 (5.1)	75.5 (7.2)	74.2 (6.9)	75.0 (7.6)
MMSE score	29.2 (0.9)	27.2 (1.8)	26.6 (1.7)	23.4 (1.9)
TIV [cm3]	1548 (143)	1558 (142)	1552 (156)	1537 (167)
Unadjusted mean bilateral baseline hippocampal volume [cm3]	5.2 (0.7)	4.6 (0.8)	4.2 (0.8)	3.9 (0.9)

controls with the MCI-P and MCI-S subjects having intermediate volumes.

5.3.1 Baseline cross-sectional results

Table 2 and Figure 5.2 show the results of the cross-sectional analysis of hippocampal volumes. In AD, after adjustment for age, gender, MMSE score, brain-to-TIV ratio and TIV, the mean baseline hippocampal volume of ε4+ was significantly smaller than that of ε4- (by 8%). There was no evidence of a difference in mean adjusted baseline hippocampal volume between ε4 carriers and non-carriers in MCI-P, MCI-S or controls.

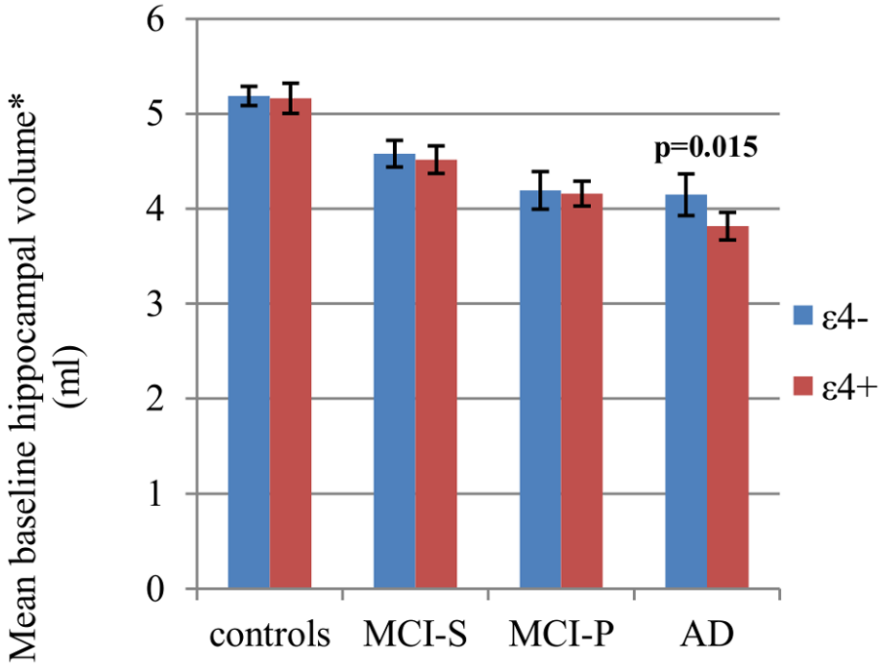
5.3.2 Longitudinal Results

Table 3 and Figure 5.3, Figure 5.4 and Figure 5.5 show the results of the longitudinal analyses of the differences in mean adjusted atrophy rates between ε4+ and ε4- in all subject groups.

Table 5.2: Adjusted mean baseline hippocampal volumes for $\varepsilon 4$ non-carriers and adjusted mean differences in total (left and right summed) baseline hippocampal volumes between $\varepsilon 4$ carriers and non-carriers in controls, stable MCI, MCI progressors and AD (-ve sign means $\varepsilon 4+ < \varepsilon 4-$).

	Controls ($\varepsilon 4-$ =118, $\varepsilon 4+$ = 49)	MCI-S ($\varepsilon 4-$ = 86, $\varepsilon 4+$ = 83)	MCI-P ($\varepsilon 4-$ = 42, $\varepsilon 4+$ = 96)	AD ($\varepsilon 4-$ = 44, $\varepsilon 4+$ = 104)
Mean adjusted* baseline hippocampal volume** in $\varepsilon 4-$ (cm3) [95% CI]	5.19 [5.08, 5.29]	4.58 [4.44, 4.72]	4.19 [4.00, 4.39]	4.15 [3.93, 4.37]
Difference in mean adjusted* baseline hippocampal volume** between $\varepsilon 4+$ and $\varepsilon 4-$ (cm3) [95% CI]	-0.02 [-0.21, 0.16] p=0.811	-0.06 [-0.26, 0.13] p=0.508	-0.03 [-0.27, 0.20] p=0.772	-0.33 [-0.59, -0.07] p=0.015

* all values are for a 50/50 gender split and are adjusted for disease-group specific mean age, baseline brain-to-total intracranial volume ratio, MMSE score, and total intracranial volume
** average of left and right



*adjusted for mean age, mean head-size, mean MMSE, mean level of prior atrophy (brain/TIV) and assuming an equal gender split
Error bars represent 95% confidence intervals

Figure 5.2: Effect of APOE $\varepsilon 4$ on baseline hippocampal volumes.*

Table 5.3: Adjusted mean difference in whole-brain and hippocampal atrophy rate (ml) [95% CI] for $\varepsilon 4$ carriers compared with non-carriers in controls, stable MCI, MCI progressors and AD (+ve means atrophy rate is higher in $\varepsilon 4+$).

		$\varepsilon 4$ carrier status	Controls ($\varepsilon 4$- =118, $\varepsilon 4+$ =49)	MCI stable ($\varepsilon 4$- = 86, $\varepsilon 4+$ = 83)	MCI progressors ($\varepsilon 4$- =42, $\varepsilon 4+$ = 96)	AD ($\varepsilon 4$- =44, $\varepsilon 4+$ = 104)
Whole-brain	Mean adjusted* atrophy rate (ml/year)	$\varepsilon 4$-	6.54 [5.88, 7.20]	7.91 [6.90, 8.93]	12.24 [10.47, 14.02]	14.11 [12.26, 15.96]
	Difference in mean adjusted* atrophy rate (ml/year)	$\varepsilon 4+$ vs $\varepsilon 4$-	0.05 [-1.15 1.25] p=0.938	2.57 [1.14, 4.00] p¡0.001	1.62 [-0.54, 3.77] p=0.142	1.58 [-0.65, 3.81] p=0.165
Hippocampus**	Mean adjusted* atrophy rate (ml/year)	$\varepsilon 4$-	0.069 [0.058, 0.079]	0.102 [0.085, 0.120]	0.151 [0.125, 0.177]	0.173 [0.145, 0.200]
	Difference in mean adjusted* atrophy rate (ml/year)	$\varepsilon 4+$ vs $\varepsilon 4$-	0.001 [-0.018, 0.021] p=0.881	0.036 [0.011, 0.061] p=0.005	0.045 [0.014, 0.076] p=0.004	0.043 [0.010, 0.076] p=0.011
	Difference in mean adjusted* atrophy rate after adjustment for concurrent whole-brain atrophy rate (ml/year)	$\varepsilon 4+$ vs $\varepsilon 4$-	0.001 [-0.014, 0.016] p=0.897	0.013 [-0.009, 0.036] p=0.250	0.031 [0.006, 0.056] p=0.014	0.029 [0.002, 0.057] p=0.037

* all values are for a 50/50 gender split and are adjusted for disease-group specific mean age, baseline brain-to-total intracranial volume ratio, MMSE score, and total intracranial volume ** average of left and right

There was statistically significant evidence that in AD, MCI-P and MCI-S subjects, after adjusting for age, gender, TIV, MMSE score and brain-to-TIV ratio, the mean hippocampal atrophy rates were higher in ε4+ compared with ε4- (see Figure 5.3). Mean adjusted brain atrophy rates were also higher in ε4+ compared with ε4-, but only significantly so in the MCI-S group (see Figure 5.4). After adjustment for concurrent whole-brain atrophy, the difference in atrophy rate between ε4+ and ε4-was reduced by 25% in AD, by 40% in MCI-P and by 75% in MCI-S (see Figure 5.5). Although the differences in mean adjusted hippocampal atrophy rates were reduced when additionally adjusting for concurrent whole-brain loss, differences between ε4+ and ε4- remained statistically significant in AD and MCI-P. In the control group there was no evidence that hippocampal or whole-brain atrophy rate differed between ε4+ and ε4- ($p>0.8$ for both).

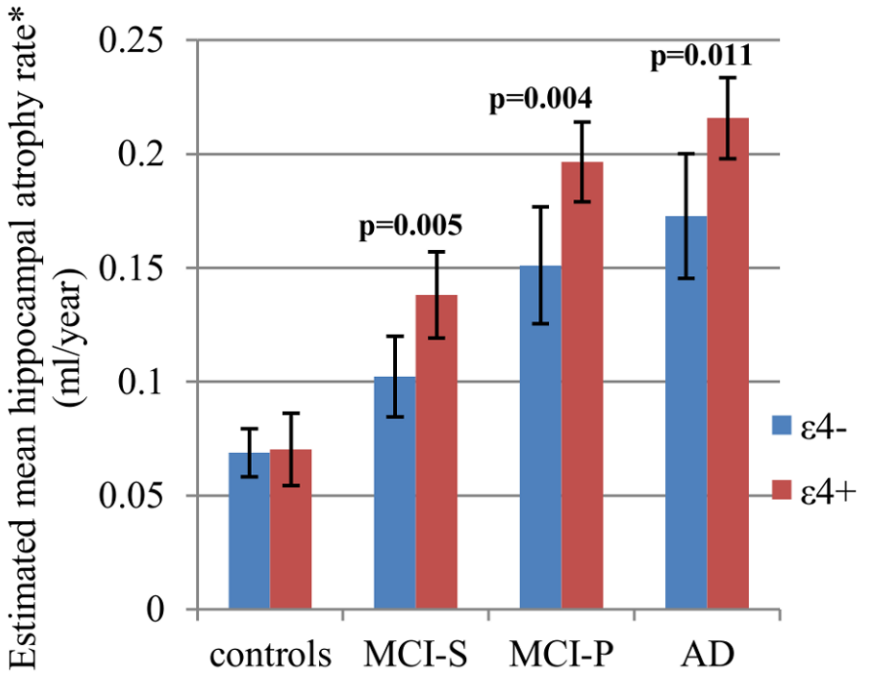
5.4 Discussion

This study examined the effect of APOE genotype on hippocampal volumes and hippocampal atrophy rates in AD, MCI and in controls, with and without adjusting for concurrent brain atrophy rates.

Cross-sectionally it was found that AD ε4+ had smaller (8%) mean hippocampal volumes at baseline than ε4- after adjusting for age, TIV, gender, MMSE score and brain-to-TIV ratio. There was no evidence that ε4+ had smaller hippocampal volumes than non-carriers in MCI-P, MCI-S or controls.

Longitudinally, mean adjusted hippocampal atrophy rates were found to be higher in ε4+ in AD, MCI-P and MCI-S but not in controls. There was also evidence that mean adjusted hippocampal atrophy rates were higher in ε4+ in AD and MCI-P after adjusting for concurrent whole-brain atrophy rates. The difference in hippocampal atrophy rates in MCI-S was no longer significant after adjustment for concurrent brain atrophy rate.

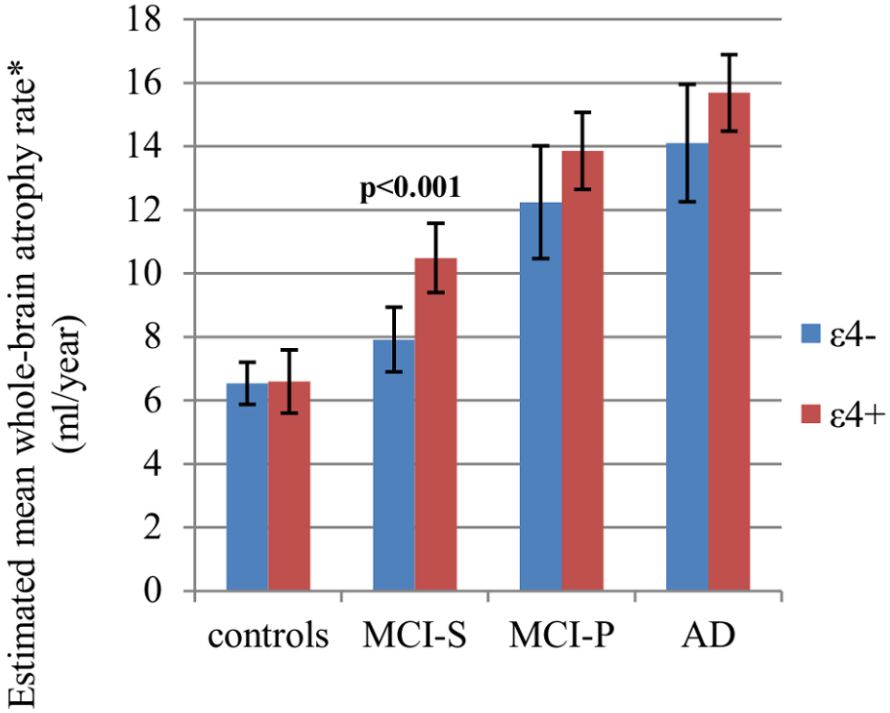
Taken together these results demonstrate that ε4 carriers with a clinical diagnosis of AD or of progressive MCI have a different pattern of atrophy - dis-



*at mean scan interval, with mean age, mean head-size, mean MMSE, mean level of prior atrophy (brain/TIV) and assuming an equal gender split
Error bars represent 95% confidence intervals

Figure 5.3: Effect of APOE $\varepsilon 4$ on hippocampal atrophy rates.*

proportionately greater hippocampal loss - than non-carriers. Cross-sectional studies have shown reduced hippocampal volumes in $\varepsilon 4+$ compared with $\varepsilon 4$- in AD. However, without investigating longitudinal changes in hippocampal volume, it is not possible to tell whether these findings could be perhaps explained by developmental differences. Indeed, there is evidence that there are some developmental differences with one study reporting higher Mental Development Index scores in 24 month old babies who were $\varepsilon 4+$ compared with those who were $\varepsilon 4$- [Wright et al., 2003]. There are few studies in healthy young people comparing hippocampal volumes in $\varepsilon 4+$ and $\varepsilon 4$-. One study in a large cohort of adolescents reported no significant difference in hippocampal volumes between $\varepsilon 4+$ and $\varepsilon 4$- [Khan et al., 2013] whilst another smaller study



*at mean scan interval, with mean age, mean head-size, mean MMSE, mean level of prior atrophy (brain/TIV) and assuming an equal gender split
Error bars represent 95% confidence intervals

Figure 5.4: Effect of APOE $\varepsilon 4$ on whole-brain atrophy rates
Effect of APOE $\varepsilon 4$ on whole-brain atrophy rates.*

in young adults reported significantly smaller hippocampi in $\varepsilon 4+$ [O'Dwyer et al., 2012]. However, the study in adolescents did not adjust for head size whilst the study in young adults did, which makes comparisons between the studies difficult. Further studies would be required to understand the developmental differences between $\varepsilon 4+$ and $\varepsilon 4$-.

In older adults previous longitudinal studies have reported higher hippocampal rates in $\varepsilon 4+$ compared with $\varepsilon 4$-. However, higher rates of hippocampal atrophy in $\varepsilon 4+$ could be potentially explained by higher rates of whole-brain atrophy (i.e. a more aggressive disease course with a more rapid loss of whole-brain tissue). In order to disentangle the effects of the $\varepsilon 4$ allele on global and local hippocampal atrophy it is necessary to adjust hippocampal

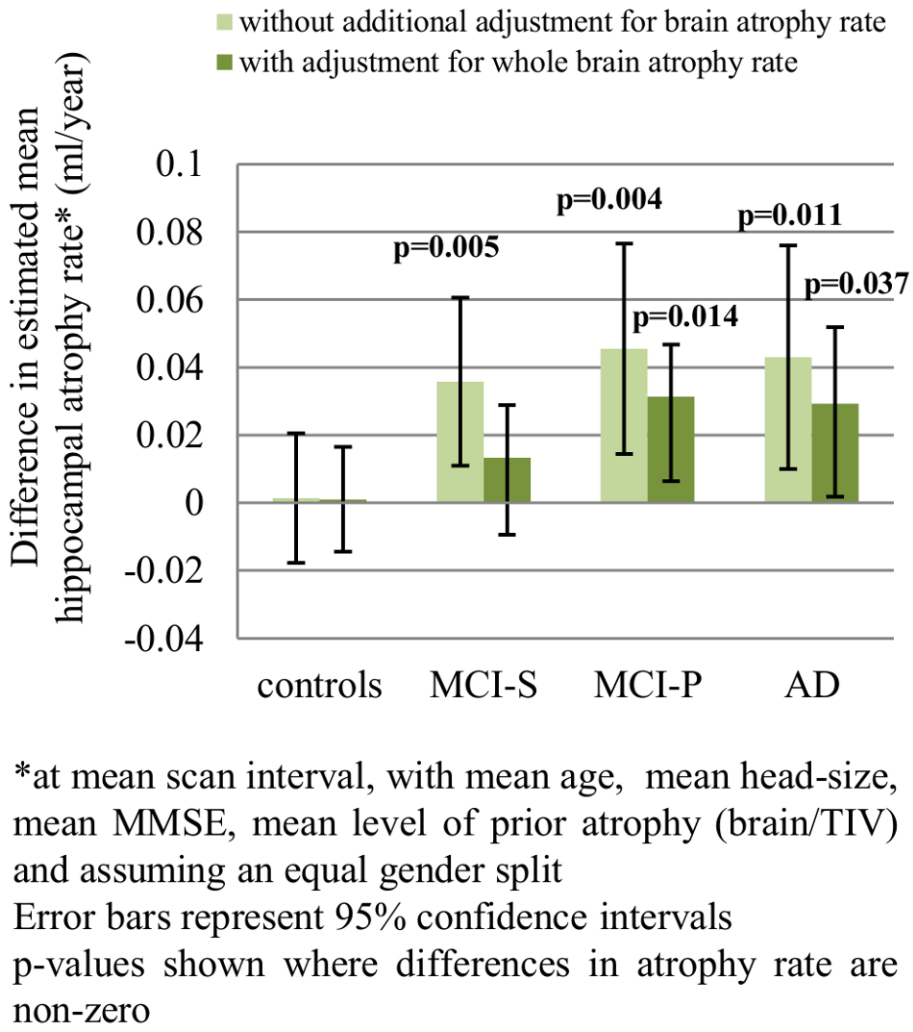


Figure 5.5: Difference in hippocampal atrophy rates*: ε4+ vs ε4-.

atrophy rates for global atrophy rates (whole-brain). In this study it was found that hippocampal atrophy rates were still higher in ε4+ in AD and progressive MCI following adjustment for whole-brain atrophy rates. This suggests that higher hippocampal atrophy rates found in ε4+ are unlikely to be simply due to a more aggressive disease with faster disease progression (as measured by generalised brain tissue loss) alone. It may be that AD associated with the ε4 allele is a different anatomical disease to AD without this allele, which should be considered when assessing the effect of potentially disease modifying treatments.

The finding of a lack of substantive differences between ε4+ and ε4- in

hippocampal volume and atrophy rate in healthy control subjects is in agreement with some previous findings [Jack et al., 1998, Lo et al., 2011, Schuff et al., 2009, Wang et al., 2009]. Conversely, a number of previous studies have reported increased hippocampal atrophy rates for ε4+ compared with ε4- controls [Chiang et al., 2011, Cohen et al., 2001, Crivello et al., 2010, Moffat et al., 2000, Morra et al., 2009, Risacher et al., 2010, Wolz et al., 2010, Jak et al., 2007, Lu et al., 2011]. However, inconsistencies in findings between this study and that of some of the others may be due to different recruitment strategies: some studies had less stringent inclusion criteria than ADNI by including some MCI subjects with controls [Moffat et al., 2000, Jak et al., 2007]; some had a majority of subjects with a 1st degree relative with a history of AD [Cohen et al., 2001]. Differences in study design may also explain inconsistencies: some studies measured atrophy over a longer period, thus increasing the power with which to estimate differences in atrophy rates [Crivello et al., 2010, Moffat et al., 2000, Lu et al., 2011]. In the largest longitudinal study to date, with over 200 ε4 heterozygotes, no evidence of a difference in rates between heterozygotes and non-carriers was found [Crivello et al., 2010], consistent with the findings of this study.

Interestingly, different studies using subsets of the controls in the ADNI cohort have reported conflicting findings. Some reported significant evidence of an association between APOE genotype and bilateral hippocampal atrophy rate [Chiang et al., 2011, Risacher et al., 2010]. One study that analysed the left and right sides separately reported a significantly higher rate of hippocampal atrophy on the right side hippocampus in ε4+ compared with ε4- [Wolz et al., 2010] another reported a significantly higher atrophy rate in the left hippocampi in ε4+ compared to ε4- [Morra et al., 2009]. Others found no such association [Lo et al., 2011, Schuff et al., 2009]. Differences between findings of these studies and this study may be due to inclusion of ε2 carriers in most studies since ε2 carriers have shown lower hippocampal atrophy rates compared with non-carriers [Chiang et al., 2010].

Reported results in MCI subjects are also mixed; a number of publications have shown a significantly greater hippocampal atrophy rate in ε4+ compared with ε4- [Chiang et al., 2011, Lo et al., 2011, Spampinato et al., 2011, Wolz et al., 2010]. One study reported a significantly greater atrophy rate in the left hippocampus [Morra et al., 2009]. Conversely other studies reported no significant difference between ε4+ and ε4- in hippocampal atrophy rate in MCI [Schuff et al., 2009, Wang et al., 2009].

In the majority of the studies using data from ADNI an association has been found between ε4 carrier-status and higher hippocampal atrophy rates in MCI much like this study. This is unsurprising in many ways since the MCI group has a high proportion of subjects who will progress to clinical AD; these subjects are more likely to be ε4+ and more likely to have increased hippocampal atrophy when compared with the MCI subjects who remain stable and may be less likely to have underlying AD pathology and less likely to be an ε4 carrier.

Other studies have examined hippocampal atrophy rates in MCI-S and MCI-P separately. One study, using voxel based morphometry (VBM) found increased hippocampal atrophy rates in MCI-P ε4+ compared with ε4- but not in MCI-S [Spampinato et al., 2011]. Another study, which used a number of hippocampal measures, found significantly higher rates in ε4+ in all measures in the MCI stable group [Risacher et al., 2010]. In MCI-P they only found significantly increased loss of hippocampal grey matter (grey matter (GM)) density and GM volume in ε4+ but not hippocampal volume (as measured by FreeSurfer). No evidence of a difference in hippocampal atrophy rates in the MCI-S group after adjusting for concurrent whole-brain atrophy rate was found.

The finding in this study in AD of smaller hippocampi in ε4+ at baseline compared with ε4- is in keeping with a previous study which reported evidence of a negative association between ε4 dose and normalised hippocampal volume in AD subjects when adjusting for other covariates such as MMSE

score [Hashimoto et al., 2009]. Further, the longitudinal findings in this study in AD of increased hippocampal atrophy rates in ε4+ compared with ε4- are in line with some previous studies [Lo et al., 2011, Mori et al., 2002, Schuff et al., 2009]. Other studies report mixed or negative results for this comparison which may depend on the image analysis methodology: one study reported increased hippocampal GM atrophy in ε4+ but no significant increase in hippocampal atrophy (as measured with FreeSurfer) or GM density changes [Risacher et al., 2010]; Others found no significant difference in hippocampal loss rates between ε4+ and ε4- in AD [Chiang et al., 2011, Morra et al., 2009].

A strength of this study was the relatively large number of subjects with data from multiple time-points (up to 36 months from baseline). ADNI has the advantage of being a prospective study with standardised follow-up times and high quality MRI imaging. The MAPS hippocampal segmentation technique was used which has been shown to have good accuracy when compared with manual segmentations [Leung et al., 2010b]. In addition, the analysis method has the advantage of a robust and direct longitudinal measure of hippocampal and whole brain change, the BSI.

This study also has a number of limitations. Much like in the previous chapter, the ADNI clinical diagnoses have not been pathologically confirmed and it may be that some AD diagnoses will prove to be caused by non-AD pathology at autopsy. Secondly, since the segmentation method used in this study (hippocampal-MAPS) excludes the hippocampal tail, and it is possible that atrophy rates differ across hippocampal sub-regions, early changes in control subjects positive for the ε4 allele could potentially have been missed and including this region in all subject groups may change the results. Thirdly, the longitudinal model assumes that the missing observations were missing at random, an assumption which cannot be empirically verified. Finally, subjects with an ε2 allele were excluded in order not to confound the results. It would be of particular interest to investigate hippocampal atrophy rates in ε2/ε4 subjects as compared with other genotypes to evaluate whether ε2 or ε4 has

greater influence on rates; however this genotype was rare in this dataset (only 3 controls, 2 MCI-S, 5 MCI-P and 2 ADs had the $\varepsilon 2/\varepsilon 4$ genotype).

In summary, the association of hippocampal volume and hippocampal atrophy rate with APOE genotype was investigated, while adjusting for age, gender, cognitive impairment (MMSE score), baseline atrophy level (brain-to-TIV ratio) and for head size in the case of cross-sectional analysis and longitudinal analysis modelling atrophy in terms of absolute volume loss. There was evidence that within the AD group $\varepsilon 4+$ had lower mean adjusted hippocampal volumes at baseline compared with $\varepsilon 4$-. AD, MCI-P and MCI-S $\varepsilon 4+$ had higher mean adjusted hippocampal atrophy rates compared with $\varepsilon 4$- and furthermore AD and MCI-P $\varepsilon 4$ carriers still showed higher mean adjusted hippocampal atrophy rates after adjustment for concurrent whole-brain atrophy rates (which, to my knowledge, has not be previously shown). Higher atrophy rates in $\varepsilon 4+$ suggest that the patterns of atrophy are not merely manifestations of developmental differences according to genotype. These results thus support the hypothesis that in AD the $\varepsilon 4$ allele influences disease phenotype with greater hippocampal involvement compared with non-carriers.

Chapter 6

Investigation into hippocampal shape and volume in Posterior Cortical Atrophy

6.1 Introduction

The hippocampus, known to play an important role in the formation of long-term, consciously accessible memories [Mayes et al., 2007, Squire et al., 2007], is one of the earliest structures to atrophy in typical Alzheimer's disease (tAD). Recently, hippocampal atrophy, visually or volumetrically assessed using magnetic resonance imaging, has been included in diagnostic criteria for Alzheimer's disease (AD) [Dubois et al., 2007, Hyman et al., 2012]. In addition, lower hippocampal volumes have been proposed as an enrichment strategy to select individuals at risk of developing clinical AD for trials of putative treatments [Hill et al., 2014]. Hippocampal atrophy has not been extensively studied in posterior cortical atrophy (PCA); one previous study reported reduced grey matter volume in the right hippocampus (but not left) in PCA as compared with controls using voxel-based morphometry [Whitwell et al., 2007]. Evaluating the extent to which the hippocampus is affected in PCA may contribute to efforts to improve our understanding of the factors driving phenotypic heterogeneity in AD. More practically, the value of biomarkers

such as hippocampal atrophy may also differ in PCA compared with typical Alzheimer's disease, and have a bearing upon the question of whether to include individuals with PCA in clinical trials in which study outcome measures have been selected for patients with more typical amnestic or global clinical presentations [Crutch et al., 2012]. In this study, the aims were to: investigate hippocampal volume differences between PCA, tAD and healthy controls; localise areas of hippocampal tissue loss; and investigate whether shape metrics give any additional group separation information above volume alone.

6.2 Methods

6.2.1 Subjects

This study included data from 124 subjects who were identified retrospectively from a clinical database at the Dementia Research Centre (PCA (n=47), typical AD (tAD; n=29), and control subjects (n=48)). The subjects included in this chapter are described in detail in 3.1.4.

6.2.2 Image acquisition

The magnetic resonance imaging (MRI) acquisition protocols used in this chapter are described in detail in 3.2.4.

6.2.3 Image processing

In-house segmentation software [Freeborough et al., 1997] was used to segment whole-brains and hippocampi.

Whole-brain regions were segmented in native space using a semi-automated technique [Freeborough et al., 1997] and were manually edited where necessary. These whole brain regions were used to generate a volume, and also to use in the subsequent registration step.

The MRI scans were then aligned to Montreal Neurological Institute (MNI) space and resampled to produce isotropic voxels of 1 mm x 1 mm x 1 mm. The left and right hippocampi were manually segmented by experienced image analysts. The hippocampi were manually delineated (by other image

analysts) using every coronal slice referencing a standard neuroanatomical atlas [Duvernoy, 2005] using a protocol that was largely similar to the EADC-ADNI Harmonized Hippocampal Protocol (HarP) [Boccardi et al., 2015a]. The protocol used in this study includes the head, body and full extent of the hippocampal tail. Two key differences between our protocol and HarP were 1) The white matter that separates the lateral ventricles from the grey matter of the hippocampus at the level of the hippocampal tail was excluded and 2) vertical digitations were excluded from the hippocampal head. In addition, a minimum threshold of 70% of the mean whole brain intensity (using the whole brain region transformed into MNI space) was used to determine the boundary between the cerebrospinal fluid (CSF) and hippocampus for improved consistency.

Estimated total intracranial volume (etotal intracranial volume (TIV)) was measured using Freesurfer [Buckner et al., 2004].

6.2.4 Hippocampal shape analysis

The hippocampal regions generated by manual segmentation were used to analyse differences in shape between the subject groups. In this study spherical harmonic (SPHARM) decomposition was used to represent hippocampal shape. Arbitrarily shaped but simply connected objects can be decomposed into a weighted series of SPHerical HARMonics (SPHARM) basis functions. SPHARM shape decompositions have the advantage of encapsulating both global and local shape features compactly. The SPHerical HARMonics - Point Distribution Models (SPHARM-PDM) (Spherical Harmonics-Point Distribution Model) toolbox was used to calculate the coefficients of the SPHARM basis functions of the hippocampi [Styner et al., 2006]. Any structure with spherical topology can be represented by a weighted sum of spherical harmonic (SPHARM) functions. For a perfect representation of the original shape an infinite number of SPHARM basis functions would be required. In practice, the number of SPHARM basis functions used to represent the original shape is determined by a user-defined parameter “Lmax,” the maximum degree of

the SPHARM expansion. The greater Lmax is, more basis functions will be used in the representation and the finer the surface representation becomes. Various values of Lmax were tested on a subset of hippocampi and visually inspected. We chose to set Lmax=12 as we felt that this provided a sufficient amount of detail.

The processing steps were as follows:

1. The hippocampi, that had been manually segmented in MNI space at an isotropic resolution of 1 mm, were binarized and resampled to an isotropic resolution of 0.5 mm. Interior holes were filled and a minimal smoothing operation was applied to ensure spherical topology (Fig 6.1a) [Styner et al., 2006]. As described by [Styner et al., 2006], the smoothing was a two-step process: first a binary closing operation was applied followed by anti-aliasing smoothing. The anti-aliasing smoothing operation used in the SPHARM-PDM package (ITK filter itk::AntiAliasBinaryImageFilter) smooths out jagged boundaries but uses the original binary surface as a constraint ensuring minimal loss in detail or structure (the smoothed surface is guaranteed to be within ±3 voxels of the original surface) [Whitaker, 2000].

2. These pre-processed binary segmentations were then transformed into raw surface meshes and spherical parametrisations for each of the hippocampal meshes were computed (Fig 6.1b).

3. From the raw surface meshes and their spherical parametrisations, SPHARM (spherical harmonic) descriptions were computed and corresponding triangulated surface meshes were generated (Fig 6.1c). These were all visually checked against the original manual segmentations to ensure that the segmentations were well represented by their SPHARM decompositions.

4. The triangulated surface meshes were then aligned to one (randomly selected) individual's hippocampal mesh using Procrustes alignment

(translation and rotation only) and the triangulated surface meshes were then regenerated such that the vertices corresponded between each individual's mesh and the chosen reference mesh of the single subject (Fig 6.1d). This was done for left and right sides separately.

5. A mean mesh was calculated from the aligned meshes (including the individual reference mesh used for alignment in the previous step) and the meshes were then aligned to the mean mesh using Procrustes alignment and again the individual meshes were regenerated such that the vertices corresponded between this mean mesh and each subject's meshes. Again, this was done for the left and right sides separately.

6. The meshes were loaded up side by side and the alignment was visually checked for alignment failures.

7. Finally, the triangulated surfaces meshes were converted from Visualization ToolKit (VTK) format to MNI object format and imported into Matlab.

6.2.5 Statistics: demographics

Linear regression analysis was used to compare age (at the time of the scan) between the diagnostic groups with age as the dependent variable and diagnostic group (PCA, tAD or controls) as the independent variable. Fisher's exact test was used to compare the gender distributions between the groups. An unpaired t-test was used to compare mini-mental state exam (MMSE) scores between tAD and PCA subjects.

6.2.6 Statistics: brain and hippocampal volume analyses

Linear regression analysis was used to compare whole brain and hippocampal volumes between the diagnostic groups. Brain or hippocampal volume was the dependent variable, diagnostic group (PCA, tAD or controls) was the independent variable and mean-centred age, gender and mean-centred head size were adjusted for.

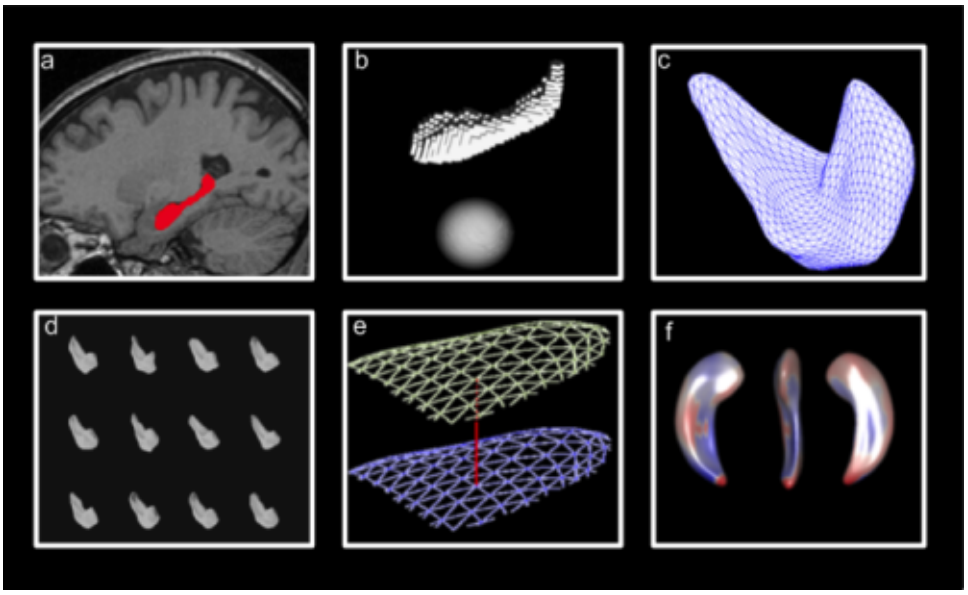


Figure 6.1: Summary of shape analysis procedure: (a) Hippocampi are segmented, holes filled and minimal smoothing applied, (b) raw surface meshes generated and spherical parametrisations are computed, (c) triangulated surface meshes are computed from spherical parametrisations, (d) surface meshes aligned (translation and rotation), (e) distance between mean mesh and individual meshes at each vertice is calculated, (f) statistical comparisons performed.

6.2.7 Statistics: hippocampal shape analysis

The SurfStat toolbox for Matlab was used to perform statistical comparisons on the hippocampal shapes [Worsley et al., 2009]. Two analyses on the hippocampal shapes were performed. In the first analysis, the distance between the surface of individual meshes and the mean mesh was the dependent variable, disease group was the independent variable and mean-centred age, gender and mean-centred head size were adjusted for. This was in order to visualise where there were any shape or volume differences in PCA subjects as compared to controls and typical AD. The second analysis was like the first analysis, but mean-centred hippocampal volume was adjusted for, instead of mean-centred head size. This was in order to visualise shape differences that were not due to volume differences. Additional analyses were run adjusting for mean-centred MMSE score and mean-centred disease duration as well as mean-centred age, gender and mean-centred head size for the PCA v tAD comparisons in order

to determine whether the differences observed in these comparisons were independent of these measures of disease severity. All comparisons were corrected for multiple comparisons (family wise error (FWE) correction). Maps showing where there were significant differences in hippocampal surface morphology were generated along with effect size maps.

6.2.8 Classification of subjects using hippocampal shape features

To quantify the extent to which hippocampal shape differences described group differences soft-margin support vector machine (SVM)s [Cortes and Vapnik, 1995] were used. The python package sci-kit learn, was used for this purpose [Pedregosa et al., 2011]. To find out whether SPHARM coefficients were better able to distinguish groups than hippocampal volume alone two SVMs were used for each group-wise comparison, one with the SPHARM coefficients as features and the other using just the left and right hippocampal volumes as features. Each subject had a total of 1014 SPHARM coefficients (from both the left and right hippocampi) since decomposition up to degree 12 was used. A nested cross-validation approach was taken whereby the sample was split into 10 mutually exclusive stratified sets of approximately equal size. Figure 6.2 summarizes the nested cross-validation process used. The classifier was trained and evaluated 10 times, once with each of the folds (data splits) as the test set and the remaining data used for training. In order to determine the best kernel (linear or radial basis function (RBF)) and kernel parameters, a grid search and stratified 10-fold cross-validation was used (experimental evidence suggests 10-fold cross-validation is the best method for model selection [Kohavi, 1995] on the training data each time. This sample splitting for kernel parameter choice was performed in order to avoid over-fitting the data. Once the kernel and hyperparameters were tuned, the resulting SVM was fitted to the training data. Finally, the SVM was used to predict the labels of the test set on which the SVM was not trained. This process was repeated for each of the 10 folds. The same folds for the SVM with SPHARM coefficients were

used as features and for the SVM with hippocampal volumes in order to be able to make comparisons. The accuracy (proportion of correctly classified subjects), sensitivity (proportion of true positives), and specificity (proportion of true negatives) were then calculated. The area under the receiver operator curve (ROC) curve (AUC) and f-score statistics are reported since they both have advantages over accuracy when assessing the performance of a classifier. The area under the curve (AUC) takes into account the decision value of the classifier which accuracy ignores. The f-score is appropriate for imbalanced classes (where one class is under-represented compared to another) and so is a better measure of the performance of a classifier than accuracy. Accuracy, sensitivity, and specificity are the same whether computing across all folds or whether taking the average of each of the folds. The AUC and f-measure are different however when computing across all folds or when averaging over the folds. Previous work suggests that taking the mean AUC of each of the cross-validation folds and computing the f-score across all folds (as opposed to averaging) are less biased [Forman and Scholz, 2010]. Therefore, mean AUC and f-score over all folds are reported in this study. Finally, in order to determine whether the SVM using the SPHARM coefficients was significantly better or worse at classification than the SVM using hippocampal volumes as features, the McNemar test [McNemar, 1947] was used since it has been shown to have a low type-1 error [Dietterich, 1998]

6.3 Results

6.3.1 Participant demographics

A total of 124 subjects were included in this study. See table 6.1 for a summary of participant demographics. The mean age of tAD subjects was higher by approximately 5 years than in controls and PCA subjects ($p<0.02$ in both comparisons). The mean MMSE scores was lower in tAD subjects than PCA subjects ($p=0.01$). There was no difference in disease duration or gender distributions between the diagnostic groups ($p>0.4$, both tests).

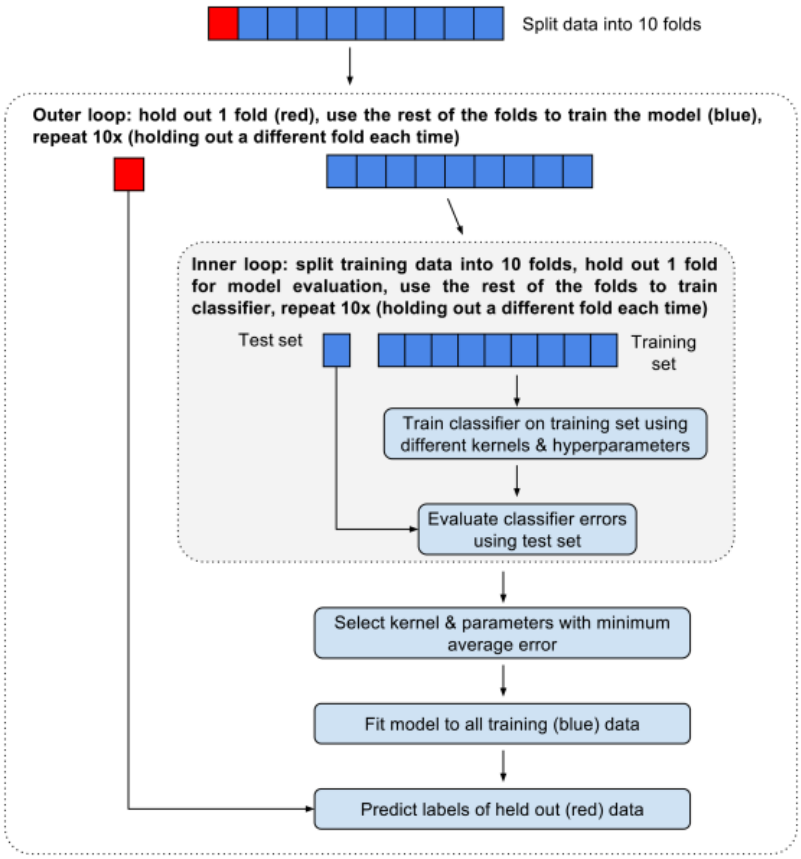


Figure 6.2: Nested 10-fold cross-validation procedure used for model tuning and evaluation.

Table 6.1: Participant Demographics

	Controls (n=48)	PCA (n=47)	tAD (n=29)	P-value
Age (years)	63.6 (9.7)	63.0 (7.0)	68.3 (8.4)	0.02 [a]
% Male	31%	40%	45%	0.5 [b]
MMSE score/30	N/A	21.2 (4.6)	18.3 (4.5)	0.01 [c]
Disease duration (years)	N/A	4.9 (2.7)	5.3 (3.1)	0.6 [c]

[a] Regression analysis. [b] Fisher's Exact Test. [c] Unpaired t-test.
Mean(SD) unless otherwise stated.

6.3.2 Brain and hippocampal volume analysis

Both PCA subjects and tAD subjects had significantly smaller mean adjusted brain volumes than the controls ($p<0.001$) but there was no significant difference in mean adjusted brain volume between the PCA subjects and the tADs ($p=0.3$). PCA subjects were found to have significantly smaller mean adjusted hippocampal volumes on both the right and left sides as compared to controls ($p \leq 0.002$, both comparisons) (see table 6.2). They were however significantly larger than those seen in tAD subjects ($p \leq 0.001$, both comparisons).

6.3.3 Hippocampal shape analysis

6.3.3.1 Comparison of PCA and controls

Significant differences in surface morphology were seen between PCA subjects and controls when adjusting for age, gender and head size: these were largely confined to the posterior hippocampus with inward deformations in the hippocampal tail region on both the right and left sides of the PCA subjects compared to controls (see the blue regions in Figure 1a). Significant inward deformations remained in the hippocampal tail regions of PCA subjects when adjusting for hippocampal volume rather than head size effectively identify-

Table 6.2: Brain and hippocampal volumes (adjusted for age, gender and head-size)

	Controls (n=48)	PCA (n=47)	tAD (n=29)
Mean adjusted brain volume (cm^3)	1131.7 [1111.5, 1151.9]	1005.8 [9851.0, 1026.5]	1021.9 [9957.5, 1048.1]
Mean adjusted left hippocampal volume (cm^3)	3.2 [3.1, 3.3]	3.0 [2.8, 3.1]	2.5 [2.3, 2.6]
Mean adjusted right hippocampal volume (cm^3)	3.3 [3.2, 3.4]	3.0 [2.9, 3.2]	2.6 [2.4, 2.7]

[a] $p \leq 0.001$ as compared with controls. [b] $p \leq 0.001$ as compared with PCA.
Mean [Confidence Interval]

ing areas of focal loss or deformation over and above the global hippocampal volume loss (see blue regions in Figure 2a).

6.3.3.2 Comparison of PCA and tAD

Outward deformations in the mean right and left hippocampal surfaces of PCA subjects were seen compared with tAD in large areas across the whole of the hippocampus (see yellow/red regions in Figure 1c). There were only very small regions where the mean surface of the tAD subjects had a significant outward deformation compared to PCA (see blue regions in Figure 1c). When adjusting for hippocampal volume rather than TIV no significant differences remained on the right side but significant differences in hippocampal shape were still seen in the left hippocampus in the left superior body with PCA outwardly deformed compared to tAD (see Figure 2c). The hippocampal shape differences observed between PCA and tAD subjects appear to be independent of disease severity as adjusting for MMSE score and disease duration made very little difference to the deformation patterns observed (see Figure 3).

6.3.3.3 Comparison of tAD and controls

In tAD, large areas of the mean left and right hippocampal surfaces were inwardly deformed as compared to controls when adjusting for age, gender and head size (see Figure 1b). The mean tAD hippocampal surface was inwardly deformed with respect to controls in most areas (see the blue regions in Figure 1b) with some small regions where there was an outward deformation of the mean surface in tAD compared to controls (see the red/yellow regions in Figure 1b). When adjusted for hippocampal volume however, only a small region of significant difference survived in the superior medial left hippocampal tail (see Figure 2b) and there were no significant differences on the right side.

6.4 Disease classification using SPHARM coefficients

The accuracies, sensitivities, specificities, mean AUCs, f-scores for each of the SVMs are shown in table 6.3. In the PCA-control comparison, by using SPHARM coefficients only a classification accuracy of 77% was achieved

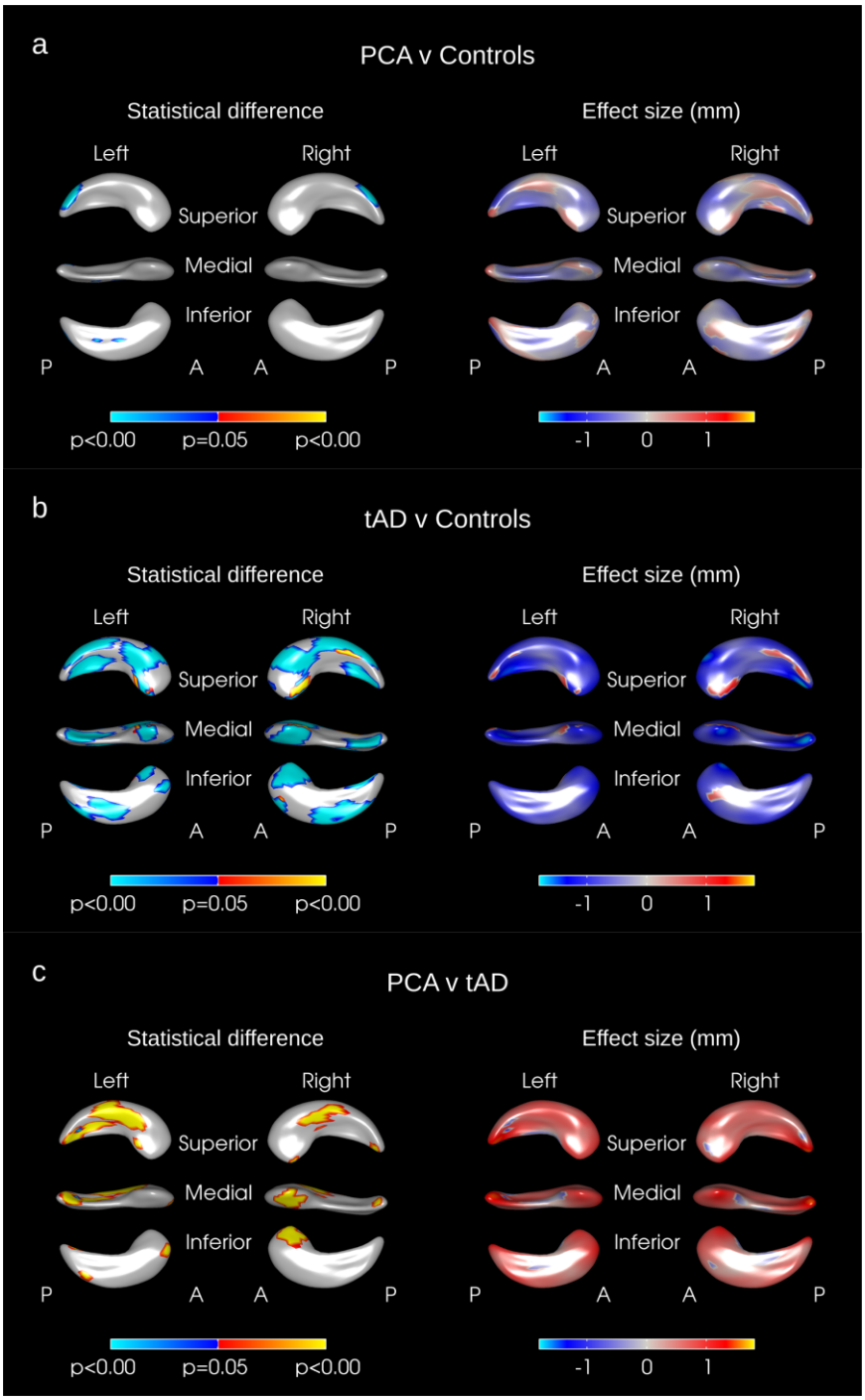


Figure 6.3: Hippocampal shape difference after adjusting for age, gender, and TIV in (a) PCA vs. controls, (b) tAD vs. Controls, (c) PCA vs. tAD. The colour scale for statistical difference represents the FWE-error corrected P-values at a threshold of P=0.05. Blue indicates areas where there was an inward deformation in (a) PCA as compared to tAD, (b) tAD as compared to controls, (c) PCA as compared to tAD whereas red/yellow indicates areas where there was an outward deformation. A=anterior, P=Posterior.

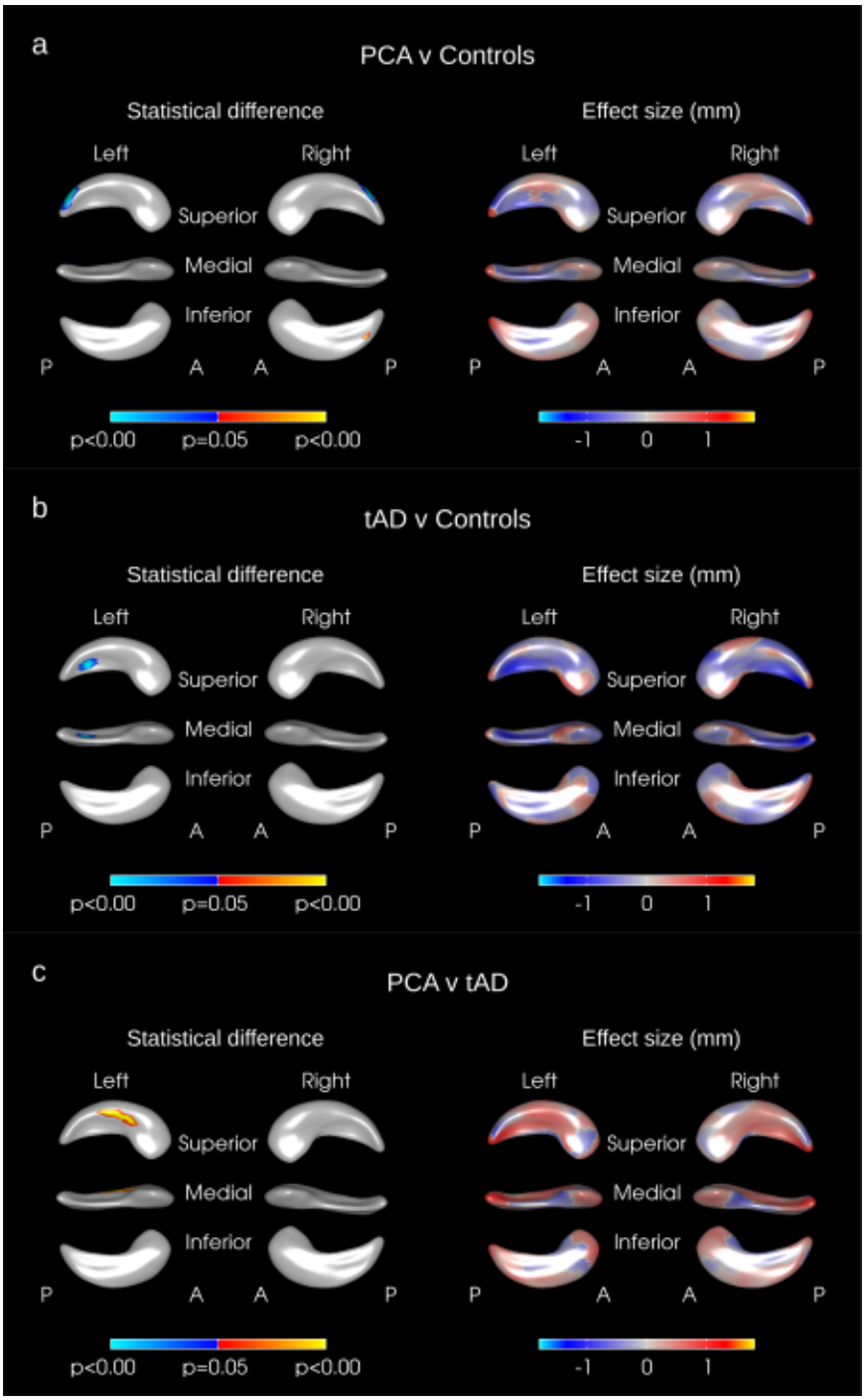


Figure 6.4: Hippocampal shape difference after adjusting for age, gender, and hippocampal volume in (a) PCA vs. controls, (b) tAD vs. Controls, (c) PCA vs. tAD. The colour scale for statistical difference represents the FWE-error corrected P-values at a threshold of P=0.05. Blue indicates areas where there was an inward deformation in (a) PCA as compared to tAD, (b) tAD as compared to controls, (c) PCA as compared to tAD whereas red/yellow indicates areas where there was an outward deformation. A=anterior, P=Posterior.

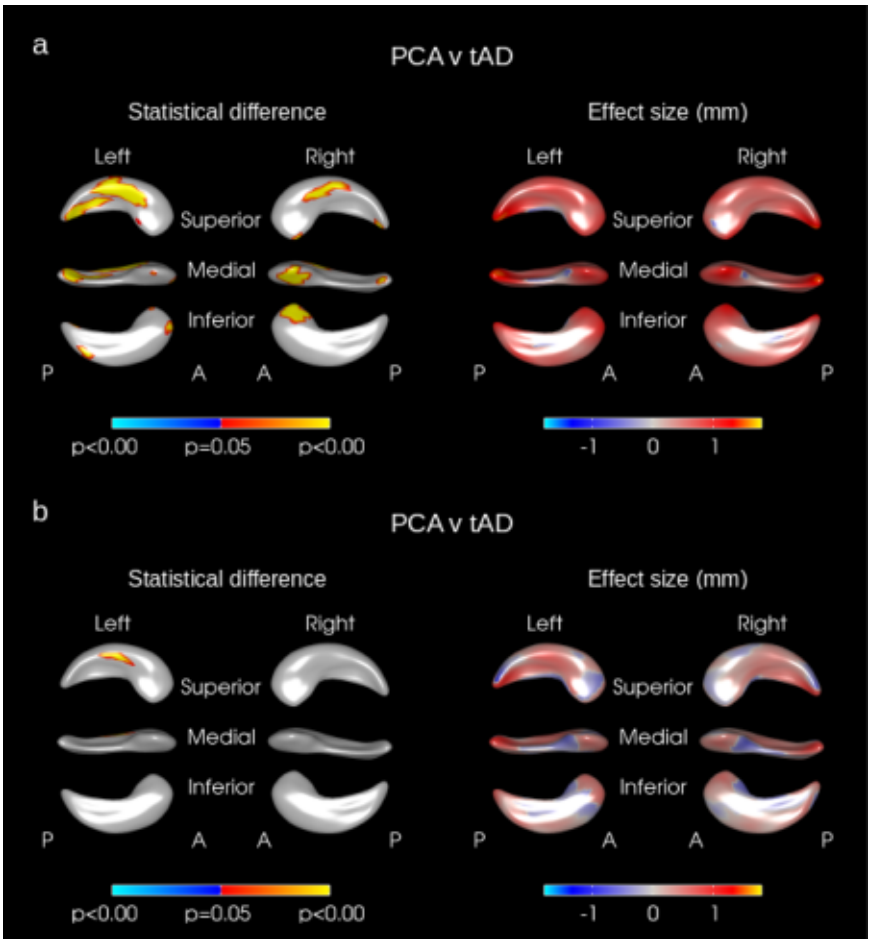


Figure 6.5: Hippocampal shape difference in PCA vs. tAD after adjusting for (a) age, gender, MMSE score, disease duration, and head size and (b) age, gender, MMSE score, disease duration, and hippocampal volume. The colour scale for statistical difference represents the FWE-error corrected P-values at a threshold of P=0.05. Blue indicates areas where there was an inward deformation in PCA as compared to tAD whereas red/yellow indicates areas where there was an outward deformation. A=anterior, P=Posterior.

compared to 56% when using hippocampal volume information alone. McNemar's test showed that the SVM classifier using SPHARM coefficients significantly outperformed the classifier using hippocampal volumes (p=0.002). In the Controls-tAD and PCA-tAD comparisons hippocampal volume alone was able to classify subjects as accurately as the SPHARM coefficients.

Table 6.3: Comparison of performance of SVM classifer using SPHARM coefficients vs SVM classifier using left and right hippocampal volumes as features)

	SVM features	Sensitivity	Specificity	Accuracy	Mean AUC	f-score	P-value (McNemar's test)
PCA vs. Controls	SPHARM coefficients only	0.74	0.8	0.77	0.84	0.76	0.002
	Hippocampal volumes only	0.59	0.55	0.56	0.57	0.56	
PCA vs. tAD	SPHARM coefficients only	0.68	0.82	0.77	0.82	0.69	0.819
	Hippocampal volumes only	0.67	0.81	0.76	0.83	0.67	
Controls vs. tAD	SPHARM coefficients only	0.77	0.91	0.86	0.93	0.80	0.763
	Hippocampal volumes only	0.77	0.94	0.87	0.91	0.81	

6.5 Discussion

PCA subjects had significantly reduced (8% lower) mean adjusted hippocampal volumes (adjusted for age, gender and head size) compared to controls; It is interesting to note that the PCA subjects had relatively preserved episodic memory function despite this volume loss. The shape analyses pointed to the differences in surface morphology in PCA being relatively localised posteriorly with inward deformations seen in the hippocampal tail regions in comparison with controls.

The loss of hippocampal volume in PCA was much lower than that seen in typical amnestic AD (25% smaller hippocampi than controls). PCA subjects had significantly larger hippocampal volumes than tAD subjects and that was

reflected in the shape differences reported. When adjusting for age, gender and head size, large areas of outward deformations, likely representing regions of relatively preserved hippocampal tissue, were found in PCA subjects compared to tAD; these were mostly seen in the superior hippocampal body with some more minor differences in the tail and the head portion of the subiculum.

When hippocampal volume was adjusted for rather than head size, some significant shape differences were still seen in PCA subjects as compared to tAD (PCA>tAD) over a small area in the left hippocampus. Given that the majority of difference was removed by adjusting for hippocampal volume it is unsurprising that SPHARM coefficients did not aid in the classification of PCA subjects from tAD.

To my knowledge, this is the first study to report shape differences in the hippocampi of PCA subjects. These results suggest that although the hippocampi in PCA subjects are relatively preserved as compared to tAD, there is some tissue loss occurring in the hippocampi of PCA subjects compared with controls. The tissue loss appears to be most significant in the superior lateral hippocampal tail region, fitting with the posterior pattern (or gradient) of atrophy seen in these subjects. In addition, when adjusting for hippocampal volume, significant differences in surface morphology were still seen in PCA subjects. Consistent with this, the hippocampal SPHARM coefficients were better able to classify PCA subjects from controls than volume alone. Taken together these data indicate that there is a distortion of the shape of the hippocampi in PCA, which could be due to focal atrophy in the hippocampus as well as the tissue to which it is connected. Although the exact functional organisation of the hippocampus remains unclear, it has been suggested that the posterior hippocampus supports detailed, context-rich spatial [Hirshhorn et al., 2012] and autobiographical [Addis et al., 2004] memories, whilst the anterior hippocampus supports more 'gist'-like memories [Strange et al., 2014]. To date there has been no detailed characterization of memory function in PCA, but the present findings of a posterior-anterior gradient of hippocampal

volume loss and shape change may predict qualitative as well as quantitative distinctions between memory processes in PCA and tAD.

As expected, significantly reduced hippocampal volumes were found in the tAD subjects as compared to controls: widespread significant inward deformations were seen across large areas of both the right and left hippocampi in tAD. Although it is difficult to precisely locate these inward deformations with respect to hippocampal subfields, in tAD these seem to approximate to the CA1 subfield as well as the anterior and posterior subiculum. A number of previous studies have compared hippocampal shapes in tAD and controls [Gerardin et al., 2009, Li et al., 2007, Lindberg et al., 2012, Shen et al., 2012, Thompson et al., 2004]. These findings are in keeping with two previous studies that found inward deformations in tAD subjects across large areas of the both the left and right hippocampi [Gerardin et al., 2009, Shen et al., 2012]. One study found large areas of inward deformations on the left hippocampus, particularly in the hippocampal head as well as the superior tail region but found no differences in the right hippocampus [Li et al., 2007]; another study reported localised inwards deformations in the hippocampal head in tAD subjects, particularly on the left side [Thompson et al., 2004] whilst another study found some inward deformations in the body of the left hippocampus and a small area of inward deformation on the medial part of the right hippocampal head [Lindberg et al., 2012]. Differences in the numbers of subjects, disease severities, shape analysis methods and hippocampal segmentation methods used may account for some of the different findings in these studies. When adjusting for hippocampal volume, no significant differences in shape were found on the right hippocampus and only a small region in the superior medial portion of the hippocampal tail on the left hippocampus. One other study [Shen et al., 2012] also investigated shape differences where the effect of volume was removed and, as in this study, found significant shape differences in the posterior hippocampus. The fact that most of the differences in shape were removed when adjusting for hippocampal volume suggests that in tAD there was generalized, diffuse

tissue loss across the whole of the hippocampus. Indeed, in this comparison, the SPHARM coefficients did not aid in the classification of tAD subjects from controls.

The fact that shape metrics helped separate PCA patients from controls suggests that they may be useful in addition to volume and could be explored in other diseases where diagnosis is difficult and subtle differences in atrophy patterns exist. In this study, shape metrics were no better than hippocampal volumes at distinguishing tAD subjects from controls. However, it is possible that the hippocampus does not atrophy uniformly during the tAD disease course. Indeed, previous studies have shown that the CA1 subfield is disproportionately affected in early AD [Chételat et al., 2008, Csernansky et al., 2005, La Joie et al., 2013, Mueller et al., 2010, Pluta et al., 2012, Wang et al., 2006] and that hippocampal subfields or hippocampal shape may be more sensitive at distinguishing mild cognitive impairment (MCI) or very mild AD subjects from controls than whole-hippocampal volume [Csernansky et al., 2005, La Joie et al., 2013, Mueller et al., 2010, Pluta et al., 2012]. Therefore SPHARM coefficients may prove to be more useful at distinguishing controls from tAD an earlier disease stage.

This study has a number of strengths. First the hippocampi were segmented manually, including the full extent of the structure from tail to head. Secondly, although PCA is an atypical variant of AD, there were a reasonable number of cases to include in these analyses. The mean MMSE score was lower in the tAD subjects than in the PCA subjects, this reflects the weighting of the questions towards memory and orientation and the relative lack of questions relating the visual deficits experienced by PCA subjects. Brain volumes in the PCA and tAD subjects were not significantly different however suggesting similar levels of overall brain atrophy between the groups.

There were several limitations to this study that warrant discussion. First, the SPHARM-PDM pipeline requires that the shapes being analysed have spherical topology. In the case of one of the hippocampi from one of the sub-

jects with tAD, the SPHARM-PDM processing failed, perhaps because this hippocampus did not have spherical topology (this subject was therefore excluded from all analyses and from the demographics table). It could be that the failure rate is higher when comparing subjects with particularly pronounced atrophy or by use of automated techniques where borders of the hippocampal masks may not adhere to the spherical topological description. Secondly, some smoothing was applied to the segmented regions before the spherical parameterization. Therefore, it may be that some of the differences that do in fact exist are not found using this method since they have been attenuated. Thirdly, pathological confirmation of AD was only available in 5 of the PCA subjects and it may be that some of the remaining PCA subjects actually have a different underlying disease [Crutch et al., 2012]. Fourthly, the type of registration is an important consideration in interpreting the results regarding localisation of tissue loss in any comparison. Other registration methods may align hippocampi differently and therefore localise deformations in other areas. Fifthly, the MRI scans used in this study were from a retrospective cohort with some variety in the scan parameters and in-plane resolutions; ideally, all subjects would have identical imaging parameters. Although it is unlikely that this would materially affect the results presented here, it cannot be excluded as a possibility and further studies using consistant imaging parameters would be required to confirm these findings. Sixthly, the images used were of limited resolution compared with the high-resolution temporal lobe imaging which is achievable [Winterburn et al., 2013]. Given that the hippocampi are relatively small structures it may be that using higher resolution scans would enable the detection of more subtle shape differences between groups. Finally, caution is required when interpreting the results from shape analysis studies - a recent study indicated that the SPHARM-PDM method of shape analysis might overestimate regions of significant difference [Gao et al., 2014]. Stringent statistical methods (family-wise error correction) were used in order to minimise false detection of differences where there were in fact none.

In conclusion, the hippocampal region is affected in PCA at a relatively early stage of the disease when memory is relatively preserved and produces posterior shape changes. Reduced hippocampal volumes were found in PCA subjects as compared to controls - intermediate between controls and tAD. Whereas the macroscopic differences between tAD and control subjects were governed by volume rather than shape, as were the differences between PCA and tAD, most of the differences between PCA and controls are governed by shape differences (PCA smaller in the tail). This was further evidenced by shape (SPHARM) coefficients that were better able to distinguish healthy controls from PCA subjects than hippocampal volume alone suggesting that shape metrics are important descriptors of hippocampal differences in PCA as compared with controls.

Chapter 7

Hippocampal template library comparison

7.1 Introduction

The automated hippocampal segmentation methods used to segment the hippocampi in chapters 4 and 5 both utilized the same template library of manually segmented hippocampi, described in detail in [Barnes et al., 2008]. The use of this template library in combination with the Similarity and Truth Estimation for Propagated Segmentations (STEPS) algorithm has been shown to have a good mean segmentation accuracy as compared to gold standard manually-segmented regions [Cardoso et al., 2015] with a mean dice score of 0.925 achieved using leave-one-out cross-validation. Details of the STEPS algorithm are given in chapter 3.

One limitation of this template library is that the protocol used for the manual hippocampal segmentations did not include the hippocampal tail region. The quality of segmentations in the tail region varies significantly when using this template library for automated segmentations (see figure 7.1). There is no biological plausibility for the volumes on these slices so this variation can only add noise to the volume measurements. There is evidence to suggest that at least some parts of the hippocampal tail [Frankó and Joly, 2013, Tang et al., 2015] are affected by atrophy in typical Alzheimer's disease and that

it is particularly affected in atypical forms of Alzheimer's disease (AD) such as posterior cortical atrophy (PCA) (see chapter 6). Using a template library that includes the full extent of the hippocampal tail in the manual segmentation protocol may improve the sensitivity of automated hippocampal volumes as a marker of disease. In addition, a recent international initiative, the hippocampal harmonization project, aimed to define a standard protocol for the manual segmentation of the hippocampus. Expert collaborators from around the world were involved in this project and agreed that the hippocampal tail should be included in the standard segmentation protocol [Boccardi et al., 2015b].

The aim of this study was therefore to investigate whether disease group classification (AD vs controls and mild cognitive impairment (MCI) vs controls) could be improved with the use of an alternative template library which included the full extent of the hippocampal tail.

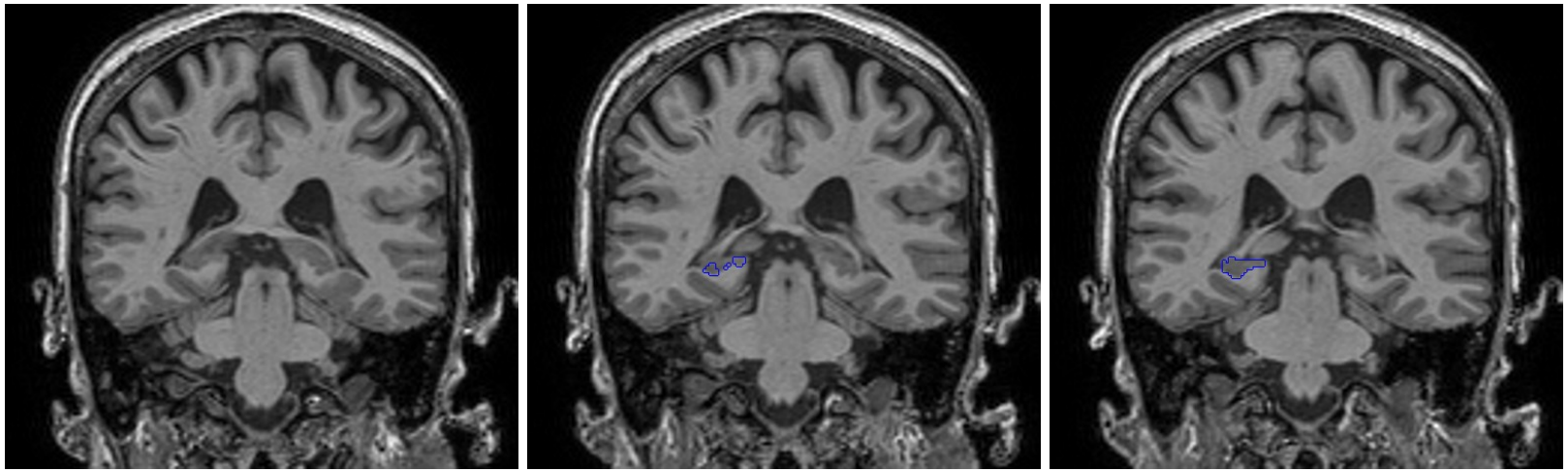


Figure 7.1: Example of the first three coronal slices of an automated hippocampal segmentation generated from the template library without tails illustrating the arbitrary nature of the starting slices.

7.2 Methods

7.2.1 Subjects

7.2.1.1 Template library 1

Template library 1 consisted of the manual segmentations described in [Barnes et al., 2008] and used in [Cardoso et al., 2015], those which did not include the full extent of the hippocampal tail. Manual segmentations of both right and left hippocampi (on T1-weighted magnetic resonance imaging (MRI) scans) for a total of 55 subjects were included in this template library: 33 with clinically diagnosed sporadic AD and 19 age matched healthy controls. See table 7.1 for a breakdown of the ages and genders of subjects used in template library 1. All manual segmentations were performed on scans registered to standard space.

Table 7.1: Demographics of subjects included in template library 1

	Controls (n = 19)	AD (n = 36)
Mean (SD) age in years	69 (7)	70 (7)
Sex (M:F)	9:10	14:22

All subjects were scanned on a 1.5T Signa MRI scanner (General Electric, Milwaukee) using an inversion recovery sequence. All the scans used in template library 1 had a voxel size of $0.9mm^3$.

7.2.1.2 Template library 2

Template library 2 consisted of manual segmentations which included the full extent of the hippocampal tail. Manual segmentations of the hippocampi (on T1-weighted MRI scans) for a total of 128 subjects were included in this template library: 30 with sporadic typical amnestic AD, 47 with PCA, 1 with familial AD, 2 with subjective memory complaints (worried well) and 48 healthy controls. Details of this cohort of subjects are given in 3.1.4. All manual segmentations were performed on scans registered to standard space.

The controls and the subjects with amnestic AD and PCA were included in the study in chapter 6. All subjects included in this template library were

Table 7.2: Demographics of subjects included in template library 2

	Controls (n=48)	Typical AD (n=30)	PCA (n=47)	Worried Well (n=2)	Familial AD (n=1)
Mean (SD) age in years	64 (10)	69 (9)	63 (7)	63 (N/A)	77 (N/A)
Sex (M:F)	15:33	13:17	19:28	2:0	1:0

N/A not applicable

scanned on a 1.5T Signa MRI scanner (General Electric, Milwaukee). There was a range of voxel sizes for the scans used in template library 2 (see chapter 3), but the majority had a voxel size of $0.9mm^3$.

7.2.1.3 Test data

To test for differences in the ability of the two template libraries to discriminate between controls and subjects with AD and MCI, baseline accelerated T1-weighted scans from a total of 834 subjects from Alzheimer's disease neuroimaging initiative (ADNI)-2/GO were used (192 controls, 320 early mild cognitive impairment (EMCI), 172 late mild cognitive impairment (LMCI) and 150 AD subjects). Accelerated scans were chosen over the non-accelerated scans due to the finding, reported in chapter 4, that accelerated scans suffer less from motion artefacts. All subjects in this study were scanned on a 3T scanner (details of MRI protocols used can be found here: http://adni.loni.usc.edu/methods/documents/mri-protocols/). See chapter 3 for details on the ADNI-2/GO cohort.

7.2.2 Hippocampal segmentation

7.2.2.1 Manual protocol used for template library 1

The segmentation protocol used for template library 1 is described in detail in [Barnes et al., 2008]. Importantly the posterior limit of the hippocampus in this protocol was defined as the coronal slice where the longest length of the

crus of the fornix was visible. All scans and their flipped mirror images (in the coronal plane such that the right side appeared on the left) were included as templates along with their associated segmentations, resulting in a template library of 110 hippocampal segmentations.

7.2.2.2 Manual protocol used for template library 2

A modification to the protocol used for template library 1 was made such that the full extent of the hippocampal tail was included in the segmentation. Briefly, the segmentation was continued on coronal slices posterior to coronal slice where the longest length of the crus of the fornix was visible, following the boundary between the grey and white matter inferiorly. Any visible white matter of the fornix was excluded and care was taken not to include the thalamus when it came into view. The most posterior slice was where the hippocampal tail was still clearly visible as an ovoid shaped grey matter mass. Segmentations were then checked and edited using the sagittal view. All scans and their flipped mirror images (in the coronal plane such that the right side appeared on the left) were included as templates along with their associated segmentations, resulting in a template library of 256 hippocampal segmentations.

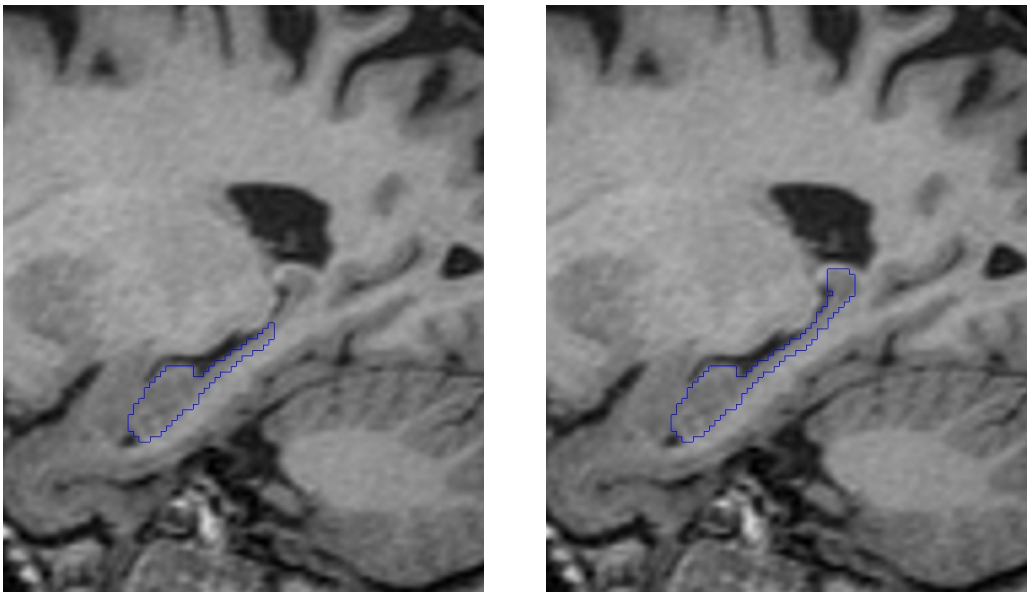
An example of a hippocampus manually segmented using protocol 1 and protocol 2 is shown in figure 7.2.

7.2.2.3 Automatic segmentation

The STEPS algorithm was used to segment the hippocampi of the ADNI-2/GO subjects using each template library. This resulted in a segmentation for each hippocampus using template library 1 and using template library 2.

7.2.3 Validation of template library 2

In order to test the reproducibility of template library 2, a leave-one-out cross validation was performed. The scans (original and flipped) and corresponding segmentations for each subject were removed from the template library in turn, and the remaining scans used to automatically segment the hippocampi



(a) (b)

Figure 7.2: Sagittal cross-section through a hippocampus segmented manually using (a) the protocol used in template library 1 and (b) the protocol used in template library 2.

for that individual.

7.2.4 Statistical analyses

To compare the classification performance when using template library 2 vs template library 1, logistic regression and receiver operator curve (ROC) curve analyses was used.

In the first model the automatically segmented left and right hippocampal volumes generated from the template library 1 (without tails) normalized for head-size, (by dividing hippocampal volume by head-size), were included as covariates. In the second model, the automatically segmented left and right hippocampal volumes generated from the template library two (normalized for head-size), were included as covariates.

7.3 Results

7.3.1 Leave-one-out cross validation of template library 2

When using a leave-one-out cross validation approach, a mean dice score of 0.918 (min: 0.838, max: 0.945) was achieved.

7.3.2 Disease group classification comparison

Figure 7.3 shows the ROC curves for the comparison of template libraries 1 and 2 for disease group classification. The area under the ROC curve was greater when using hippocampal volumes from template library 2 compared with template library 1 for AD vs controls (p=0.02) and LMCI vs controls (p=0.001). There was no evidence of a difference in the area under the ROC curve when comparing the two template libraries for classification of EMCI vs controls.

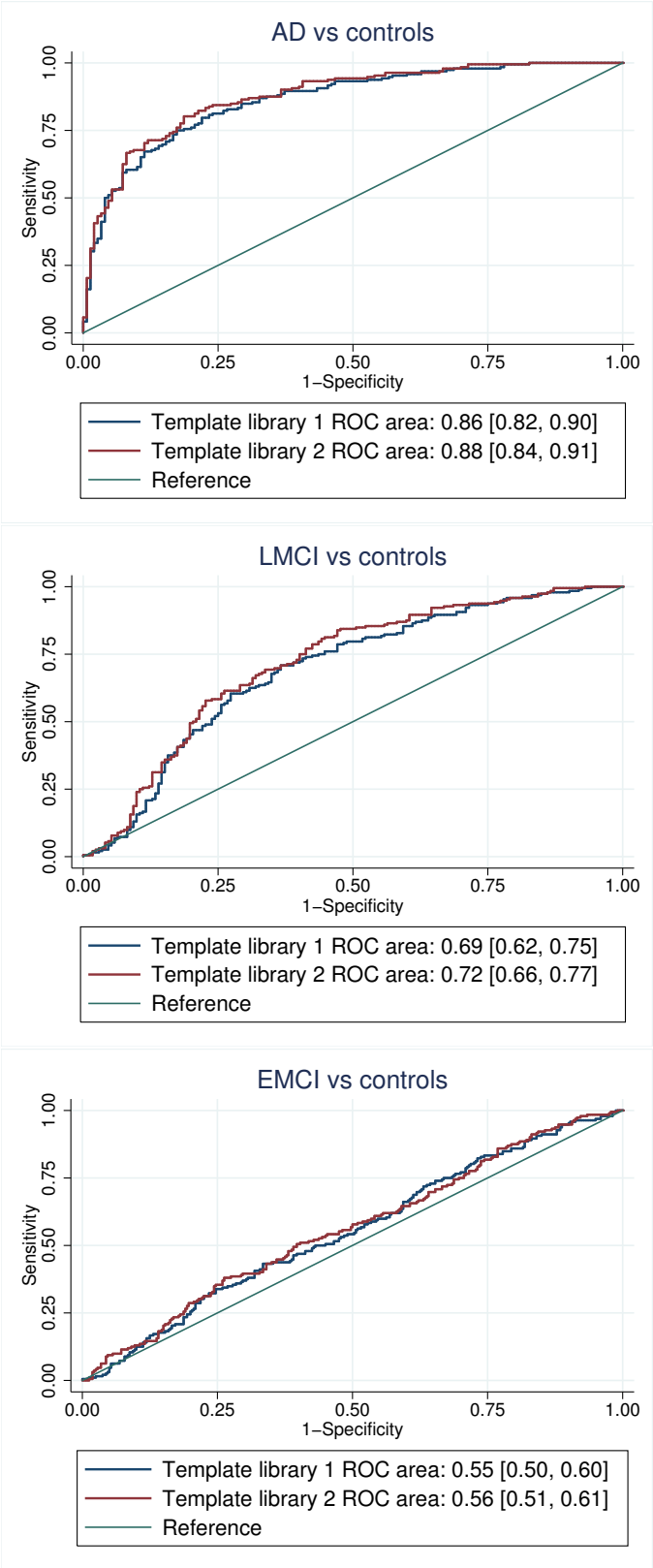


Figure 7.3: Receiver operator curves for disease group classification using template library 1 and template library 2. ROC area [95% confidence interval] shown in legend. AD: Alzheimer's disease, LMCI: late mild cognitive impairment, EMCI: early mild cognitive impairment.

7.4 Discussion

Cross-validation showed that template library 2 (which included the full extent of the hippocampal tail) had a good mean segmentation accuracy when comparing automated to manually-segmented regions. The mean dice score was 0.918 for template library 2, marginally lower than that achieved with template library 1 (0.925), but these are arguably not 100% comparable as each were tested within their respective template libraries. Despite this, the use of template library 2 had a higher classification accuracy of AD subjects from controls and LMCI subjects from controls compared with the use of template library 1. There was no difference between the two in terms of discriminative ability in EMCI from controls.

Reasons for the improved ability of template library 2 to distinguish between AD and controls and LMCI and controls could be due to a number of reasons. Firstly, the inclusion of the hippocampal tail in template library 2 may have reduced the noise of resultant segmentations. Secondly, the wider variety of morphologies used in template library 2 (due to the increased number of subjects and a wider variety of disease phenotypes) could have meant that for each new hippocampus to segmented, better matches could be made. Finally, differences in scan quality between the two template libraries could in part explain the improved results. Should scan quality be improved or a better match to the ADNI-GO/2 scans, then this might have an impact on the quality of the resultant segmentations..

This assessment has a number of limitations. The main limitation is that conclusions as to whether the inclusion of the tail or the quality or morphologies represented in the second template library, cannot be ascertained. In order to establish this, a comparison using the same set of template scans with two different sets of segmentations (with and without the hippocampal tail) would be required.

However, it can be concluded that a significant improvement in disease group classification was achieved by using template library 2 over template

library 1.

In addition to improving disease group classification, the segmentations used in template library 2 more closely match the segmentation protocol described by the recent hippocampal harmonization project [Boccardi et al., 2015b], where experts from around the globe came to a consensus on what should be included in hippocampal segmentations. The findings in this chapter suggest that template library 2 may be more appropriate for use in Alzheimer's disease studies than template library 1.

Chapter 8

Manual and automated thalamic segmentation

8.1 Introduction

This thesis has so far focussed on the hippocampus and hippocampal atrophy, a widely-established marker of Alzheimer's disease.

The hippocampus, part of the limbic system, is important for the formation of long-term memories [Scoville and Milner, 1957]. Other structures in the limbic system, such as the anterior thalamic nucleus, are also known to play an important role in episodic memory [Child and Benarroch, 2013] and there is evidence to suggest that the anterior thalamic nucleus is also affected by Alzheimer's disease pathology; Braak and Braak found that at post-mortem, whilst most of the thalamic nuclei were relatively mildy affected by Alzheimer's disease pathology, the limbic nuclei contained numerous neurofibrillary tangles and neurophil threads with the antero-dorsal nucleus severely affected [Braak and Braak, 1991]. Previous imaging studies have shown reduced thalamic volumes in subjects with Alzheimer's disease [De Jong et al., 2008].

The thalamus is an important hub in the brain, relaying information from different subcortical areas to the cortex. It is an ovoid structure in the diencephalon, mostly consisting of grey matter. The internal medullary lamina, a Y-shaped sheet of white matter, separates the thalamus into three main parts:

the anterior nuclei, the medial nuclei and the lateral nuclei. The superior surface of the thalamus is covered by a layer of white matter called the stratum zonale and the lateral surface by the external medullary lamina.

As discussed in chapter 2, thalamic segmentation from T1-weighted images is challenging due to the typically poor image contrast between the thalamus and some of its surrounding structures. Diffusion tensor imaging allows for the visualisation of white matter tracts, and may provide complimentary structural information with regards to thalamic boundaries.

The aims of this study were: 1) to investigate whether manual segmentation reproducibility could be improved by using both coloured fractional anisotropy maps and T1-weighted magnetic resonance imaging (MRI) scans together as opposed to T1-weighted MRI scans alone, 2) to generate a template library of manually-segmented thalami based on the segmentation protocol developed, for use in automated segmentation pipelines and 3) to test the ability of an automated pipeline [Cardoso et al., 2015] to accurately segment thalami using the generated manual template library.

8.2 Methods

8.2.1 Subjects

In order to generate a template library with a variety of brain morphologies a range of subjects with a mixture of pathologies was selected. For this purpose, data from the AVID and young onset Alzheimer's disease (YOAD) studies described in chapter 3 were used. A total of 73 subjects were included: 57 subjects from the YOAD cohort and 16 from the AVID cohort. Demographic information by cohort and diagnostic group is given in table 8.1.

Table 8.1: Demographics of subjects used in template library

	Diagnosis	No. subjects	Mean age (min,max)	Mean mini-mental state exam (MMSE) score [min, max]	Gender (M:F)
AVID1	Controls	3	68 [65,72]	30†[30,30]	2:1
	SD	4	68 [57,79]	23†[18,26]	2:2
	PNFA	4	68 [64,70]	15†[6,30]	1:3
	PCA	3	59 [57,62]	17 [13,23]	1:2
	LPA	2	65 [59,72]	28†	2:0
YOAD	Controls	20	60 [48,68]	30 [28, 30]	8:12
	PCA	11	61 [53,70]	22 [13, 29]	3:8
	tAD	25	61 [51,73]	21 [13, 28]	10:15
	FvAD	1	53	15	1:0
		73	62 [48,79]	73 [6,30]	30:43

SD: semantic dementia, PNFA: progressive nonfluent aphasia, posterior cortical atrophy (PCA): posterior cortical atrophy, LPA: logopenic aphasia, typical Alzheimer's disease (tAD): typical amnestic Alzheimer's disease, FvAD: subject with atypical Alzheimer's disease with a frontal presentation.
† MMSE score not available for one subject.

8.2.2 MRI scan acquisition and processing

For details on the MRI scan protocols used in this study see chapter 3. For details on the diffusion tensor imaging processing pipeline used to generate the coloured fractional anisotropy (FA) maps, see chapter 3.

All subjects included in this study had both T1-weighted and diffusion tensor imaging scans acquired in the same session. All T1-weighted scans were

aligned to Montreal Neurological Institute (MNI) space and FA maps were resampled in the space of the T1-weighted scans.

8.2.3 Manual segmentation protocol

A manual segmentation protocol of the thalamus was developed with reference to atlases [Duvernoy, 2005, Mori et al., 2002] and a previously published thalamic segmentation protocol [Power et al., 2015].

A coronal section through the thalamus with the surrounding structures labelled is shown in figure 8.1, fiigure 8.2 shows an axial section. The thalamus is bounded medially by cerebrospinal fluid, laterally by the internal capsule and inferiorly by the red nucleus.

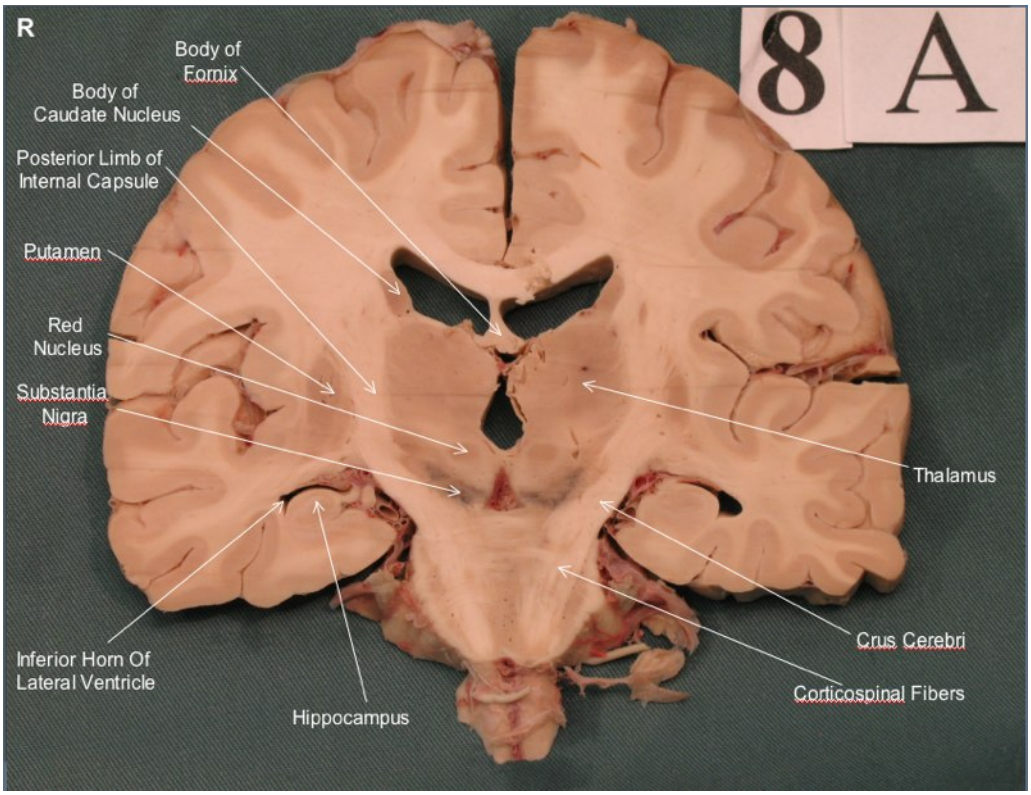
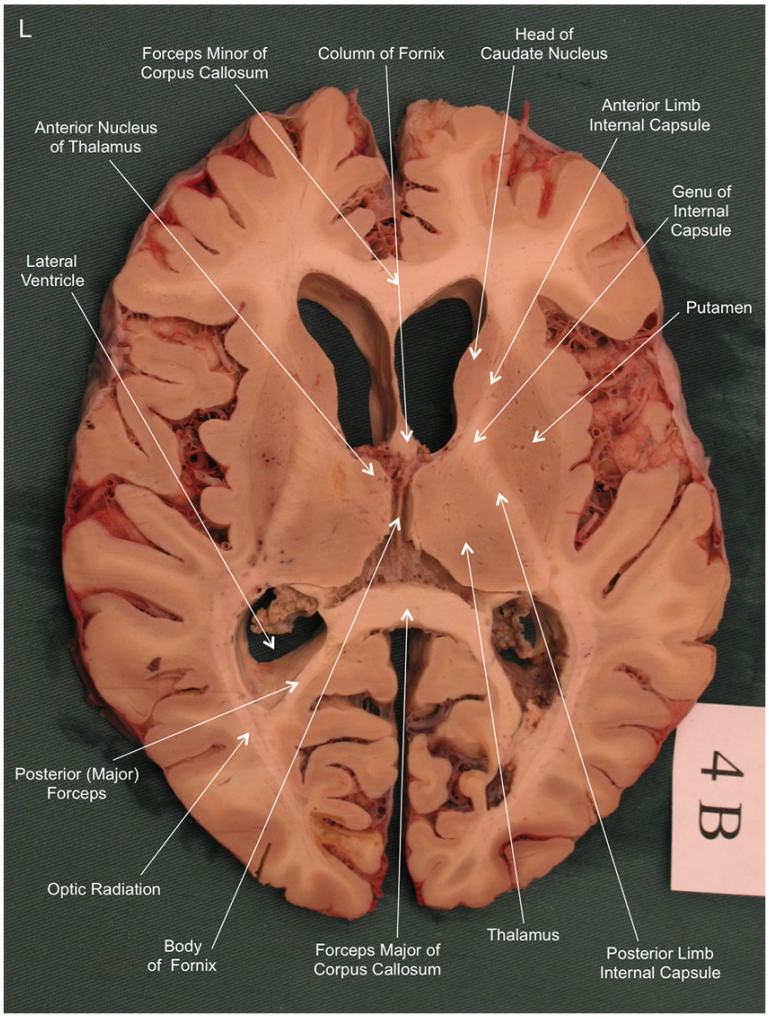


Figure 8.1: Coronal slice showing the location of the thalamus and it's surrounding structures. Photographed and labelled by Dr. Bruce Crawford and Kurt McBurney at the University of Victoria. This photograph is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 2.5 Canada License.

All manual segmentations were performed using the NiftyMidas software tool (see chapter 3). This tool allows for the simultaneous visualization in all



!ht]

Figure 8.2: Axial slice showing the location of the thalamus and it's surrounding structures.Photographed and labelled by Dr. Bruce Crawford and Kurt McBurney at the University of Victoria. This photograph is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 2.5 Canada License.

three planes whilst segmenting.

The anatomical landmarks used in the protocol developed in this study were largely similar to those described in the protocol developed by [Power et al., 2015]. In contrast to the protocol described in Power et al., the segmentation was initiated in the axial plane in the protocol developed in this study. In addition, a threshold of 70% of the mean brain intensity (derived from whole brain segmentations) was used to delineate the cerebrospinal fluid (CSF) from the thalamus.

8.2.3.1 Practical guide to thalamic segmentation when using T1-weighted scans only

In the most superior axial slices of the thalamus, the boundaries are defined by the internal capsule laterally and CSF medially (see figure 8.3).

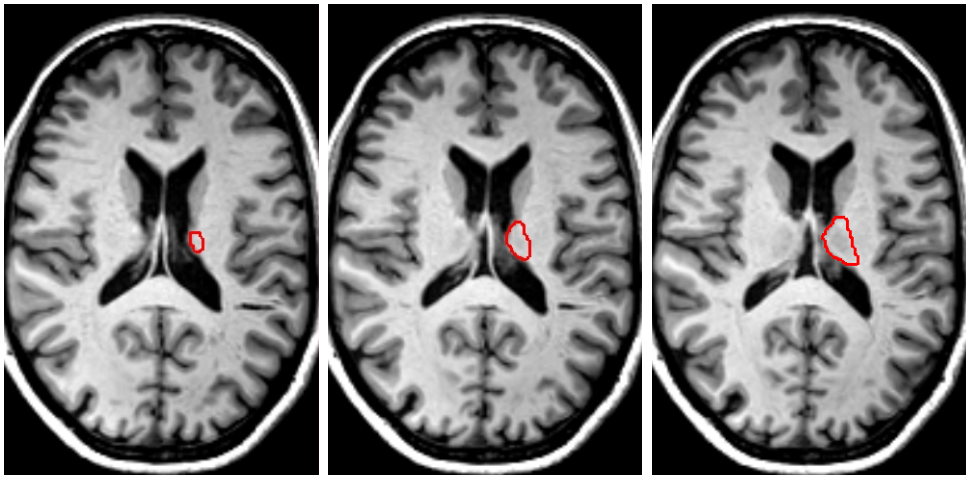


Figure 8.3: Three most superior axial slices through the thalamus (outlined in red). The most superior slice is on the left through to the most inferior on the right.

Moving inferiorly, the fornix comes into view, posterior to the thalamus. Care was taken to exclude the most posterior white matter visible posterior to the thalamus, which belongs to the fornix (using multiple views helped with this) see figure 8.5.

Figure 8.4: Axial slice through the thalamus (outlined in red) where white matter of fornix is visible posterior to the thalamus.

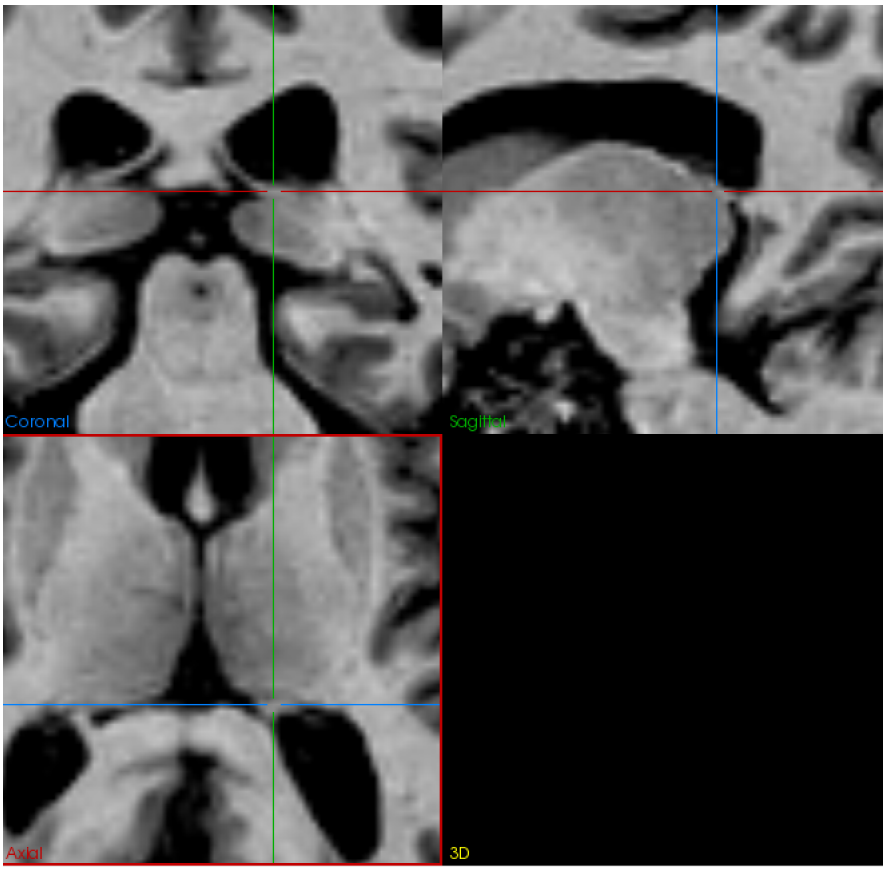


Figure 8.5: Coronal, sagittal and axial view through an example brain with cross-hairs indicating the fornix appearing as white matter posterior to thalamus.

Segmentation was continued, following the CSF boundary medially and the white matter boundary of the internal capsule laterally, on subsequent inferior axial slices. When the habenula came in to view (figure 8.6) on the medial side of the thalamus care was taken to exclude it from the segmentation. Figure 8.6 shows the habenula in all 3 views.

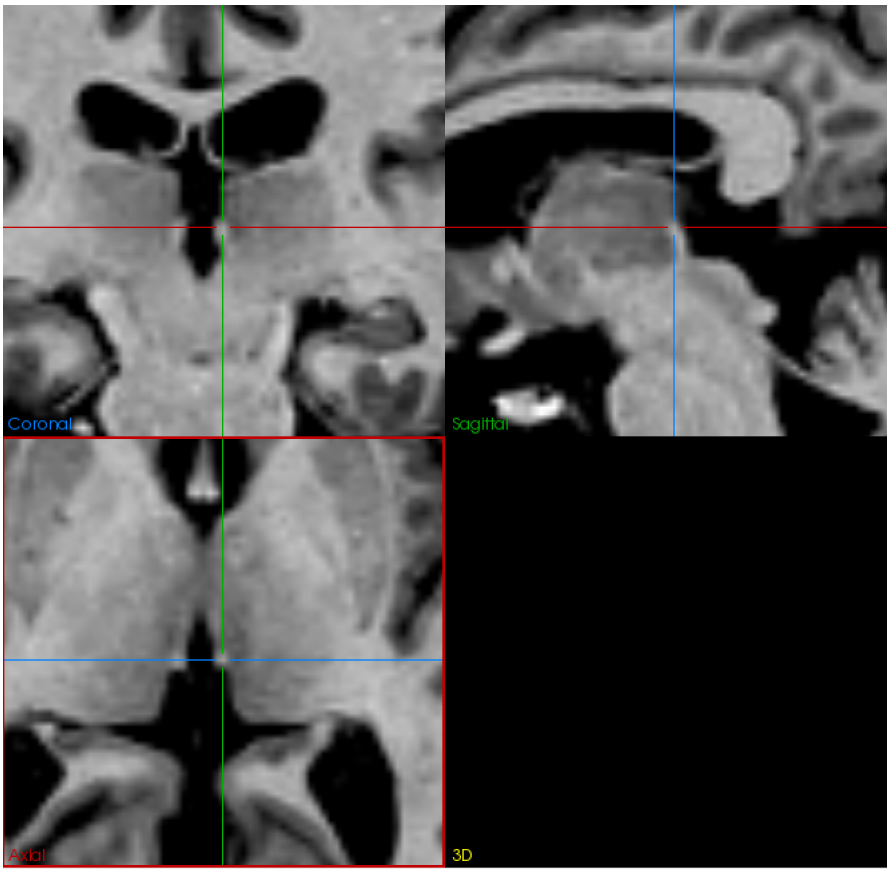


Figure 8.6: Coronal, sagittal and axial view through the thalamus with cross-hairs indicating the habenula appearing as white matter medial to thalamus.

Soon after the habenula is visible, the superior colliculus, medial to the thalamus, becomes visible, and care was taken to exclude this as well (see figure 8.7).

Moving inferiorly, the red nuclei become visible (see figure 8.8). In slices where the red nucleus was present, the top of the red nucleus, in coronal view, served as a landmark for the inferior boundary of the portion of the thalamus that is superior to the internal capsule.

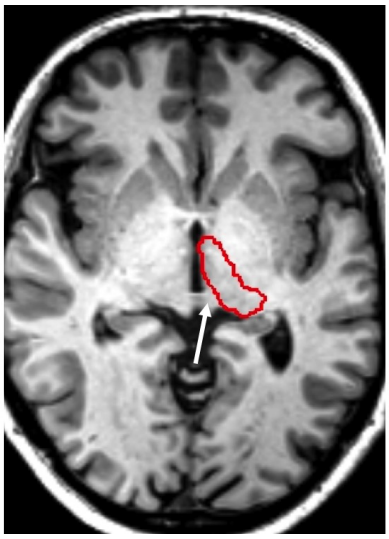


Figure 8.7: Axial section through brain showing location of superior colliculus with respect to the thalamus (outlined in red).

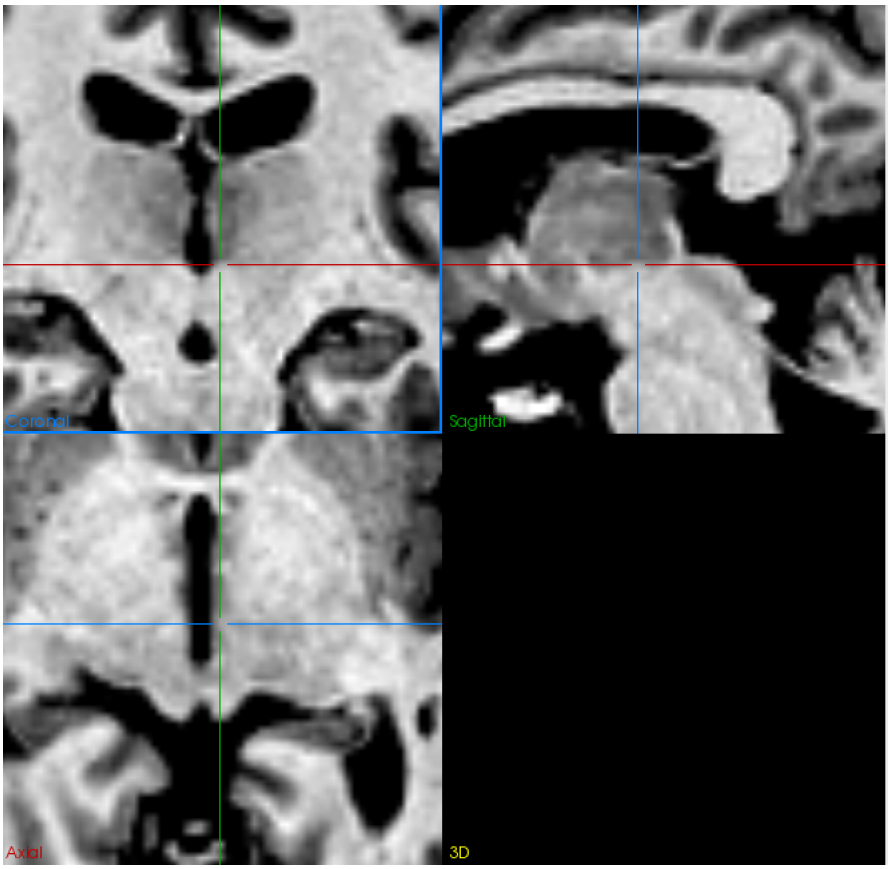


Figure 8.8: Coronal, sagittal and axial view through thalamus with cross-hairs indicating the top of the red nucleus, inferior to the thalamus.

Moving inferiorly, the lateral geniculate body is still visible as a grey mass, adjacent to the the cerebral peduncle in axial view (see figure 8.9). The lateral geniculate nuclei are inferior to the internal capsule/cerebral peduncle and may appear separated from the main mass of the thalamus in coronal view.

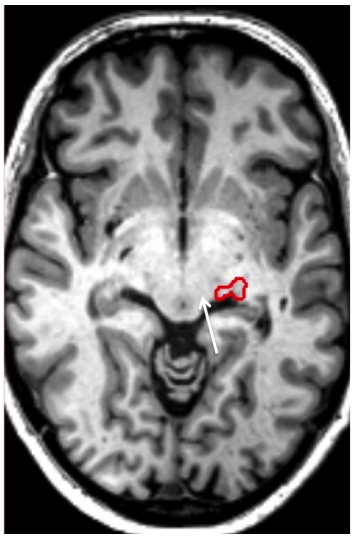


Figure 8.9: Axial slice through the brain at the level of the cerebral peduncle (indicated by a white arrow). The lateral geniculate nuclei of the thalamus are outlined in red.

Visualisation of the boundary between the internal capsule and thalamus can be challenging in some scans and on some scan slices. Increasing the contrast between grey and white matter on the T1-weighted scans helps with the visualisation of the boundary between the internal capsule and the thalamus. Where the boundary is not clearly visible, slices superior and inferior to the slice in question can be used to guide the placement of the boundary.

8.2.3.2 Practical guide to thalamic segmentation when using coloured FA maps in addition to T1-weighted scans

When using coloured fractional anisotropy maps in addition to T1-weighted scans, T1-weighted and coloured FA maps were overlaid in the window. In each slice, the thalamus was first segmented using the T1-weighted map according to the protocol outlined above, then the coloured FA map was used for refining the internal capsule/thalamic boundary. An example of a T1-weighted and

coloured FA map showing the boundary between the internal capsule and thalamus on both is shown in figure 8.10. The coloured FA map was found to be particularly useful for this purpose, but did not particularly aid in other areas.

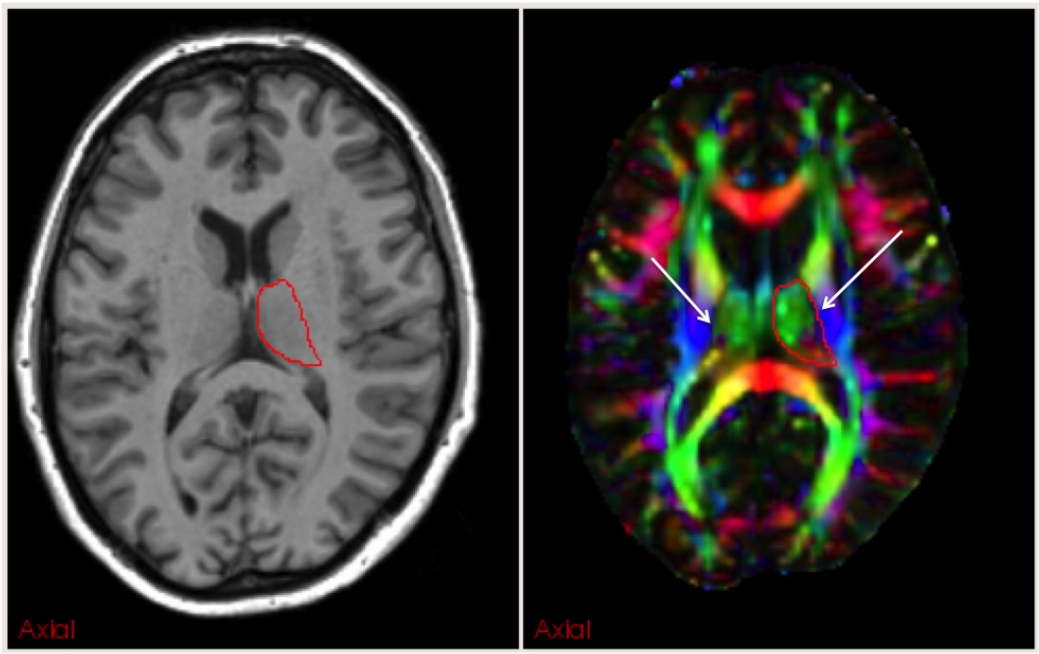


Figure 8.10: Axial slice through thalamus on T1-weighted scan (left) and coloured FA map (right). The left thalamus is outlined in red and arrows indicate the boundary between the internal capsule and thalamus, which can be more easily visualised on coloured FA maps than on T1-weighted scans.

8.2.4 Testing reliability of manual segmentations

In order to test the reliability of the manual segmentation protocol, a subset of 30 subjects from the YOAD dataset (10 with tAD, 10 with PCA and 10 controls) was selected. The left thalamus was segmented 4 times per subject; twice using T1-weighted scans only and twice using T1 and FA maps. This was performed in a random order and blinded to subject ID and diagnosis. For half of the subjects, only the T1-weighted scan was presented first. For the other half, both the FA map and T1-weighted scans were presented first.

Scans were presented in such a way that the scans for a particular subject did not get segmented more than once in the same day in order to prevent learning bias. To compare the reliability of the segmentations, dice overlap scores, mean surface distances and Hausdorff distances were calculated [Prados et al., 2017].

8.2.5 Testing reliability of automated segmentations

A leave-one-out cross-validation was performed using the template library of manual segmentations generated and their associated MRI scans. The scans (original and left-right flipped) and associated segmentations for each subject were excluded from the template library in turn, and the remaining scans and segmentations used to automatically segment the scan in question using the Similarity and Truth Estimation for Propagated Segmentations (STEPS) algorithm [Cardoso et al., 2015]. A brief description of how the STEPS algorithm works is given in chapter 3. Dice overlap scores, mean surface distances and Hausdorff distances between the manual segmentations for each subject and the automatically segmented regions were calculated.

To test whether the automated segmentation pipeline performed better with the inclusion of FA maps as compared with T1-weighted scans alone, two separate leave-one-out cross-validations were performed. The first used the labels with the T1-weighted scans only in the STEPS pipeline, the second used the labels with the T1-weighted scans and the FA maps in the STEPS pipeline.

8.3 Results

8.3.1 Reliability of manual segmentations

The mean dice score, mean surface distance and Hausdorff distance are shown in table 8.2. In none of the comparisons was there a statistically significant difference. When looking at all subjects combined, the use of FA maps in combination with T1-weighted scans resulted in a marginally better mean dice score and mean surface distance but a slightly worse Hausdorff distance compared with T1-weighted scans alone. Higher dice scores and lower mean surface

Table 8.2: Similarity metrics for repeated manual segmentations. Mean [95% confidence interval] shown.

		T1 only	T1 & diffusion	p-value
Dice	All subjects (n=30)	0.940 [0.935, 0.945]	0.942 [0.936, 0.948]	0.7
	Controls (n=10)	0.950 [0.946, 0.954]	0.953 [0.946, 0.960]	0.7
	PCA (n=10)	0.936 [0.926, 0.946]	0.944 [0.938, 0.950]	0.9
	AD (n=10)	0.934 [0.926, 0.942]	0.929 [0.915, 0.943]	0.3
Mean surface distance	All subjects (n=30)	0.346 [0.324, 0.369]	0.332 [0.307, 0.357]	0.2
	Controls (n=10)	0.304 [0.284, 0.325]	0.285 [0.254, 0.316]	0.2
	PCA (n=10)	0.361 [0.315, 0.407]	0.326 [0.298, 0.354]	0.1
	AD (n=10)	0.374 [0.332, 0.416]	0.385 [0.332, 0.439]	0.6
Hausdorff distance	All subjects (n=30)	2.455 [2.268, 2.641]	2.486 [2.240, 2.732]	0.6
	Controls (n=10)	2.192 [1.961, 2.424]	2.141 [1.715, 2.566]	0.4
	PCA (n=10)	2.475 [2.080, 2.870]	2.619 [2.179, 3.059]	0.7
	AD (n=10)	2.697 [2.337, 3.056]	2.699 [2.212, 3.186]	0.5

and Hausdorff distances were achieved in the control subjects than in either the PCA or tAD groups, none of these differences reached statistical significance however.

8.3.2 Reliability of automated segmentations

A summary of the results of the leave-one-out cross validation are shown in table 8.3. The use of both T1-weighted scans along with FA maps resulted in higher mean dice scores on both the right and left sides. The differences were subtle, but reached statistical significance.

Table 8.3: Similarity metrics for automated segmentations using T1-weighted scans only vs manual segmentations and automated segmentations using FA maps in addition to T1-weighted scans vs manual segmentations. Mean [95% confidence interval] shown

		T1 only	T1 & FA map	p-value
Left	Dice	0.940 [0.937, 0.944]	0.942 [0.939, 0.945]	0.02
	Mean surface distance	0.334 [0.32, 0.348]	0.329 [0.316, 0.342]	0.02
	Hausdorff distance	2.295 [2.153, 2.437]	2.302 [2.141, 2.463]	0.87
Right	Dice	0.941 [0.938, 0.944]	0.949 [0.947, 0.951]	<0.001
	Mean surface distance	0.329 [0.317, 0.341]	0.299 [0.289, 0.308]	<0.001
	Hausdorff distance	2.220 [2.105, 2.335]	2.081 [1.944, 2.218]	0.01

Examples of the best and worst automatic segmentation (according to the dice overlap score) when using T1-weighted scans only is shown in figure 8.3.2 and when using FA maps in addition to T1-weighted scans is shown in figure 8.12. The scan with the biggest difference in dice overlap scores between the region automatically segmented using T1-weighted scans only and the region automatically segmented using both FA-maps and T1-weighted scans is shown in figure 8.13.

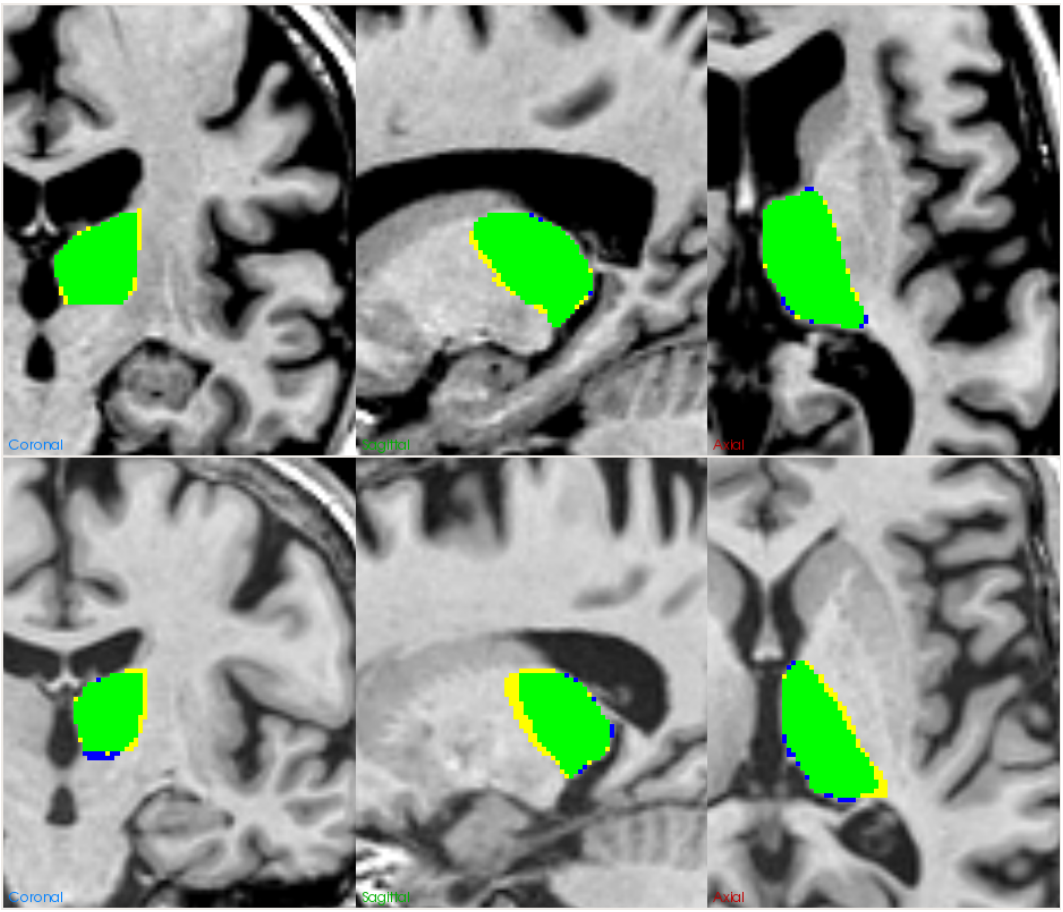


Figure 8.11: Best (top row) and worst (bottom row) automated segmentation (according to dice score) when using T1-weighted scans only. Blue: included in manual segmentation only, green: included in both manual and automated segmentations, yellow: included in automated segmentation only. The best segmentation had a dice score of 0.96 and the worst had a dice score of 0.89.

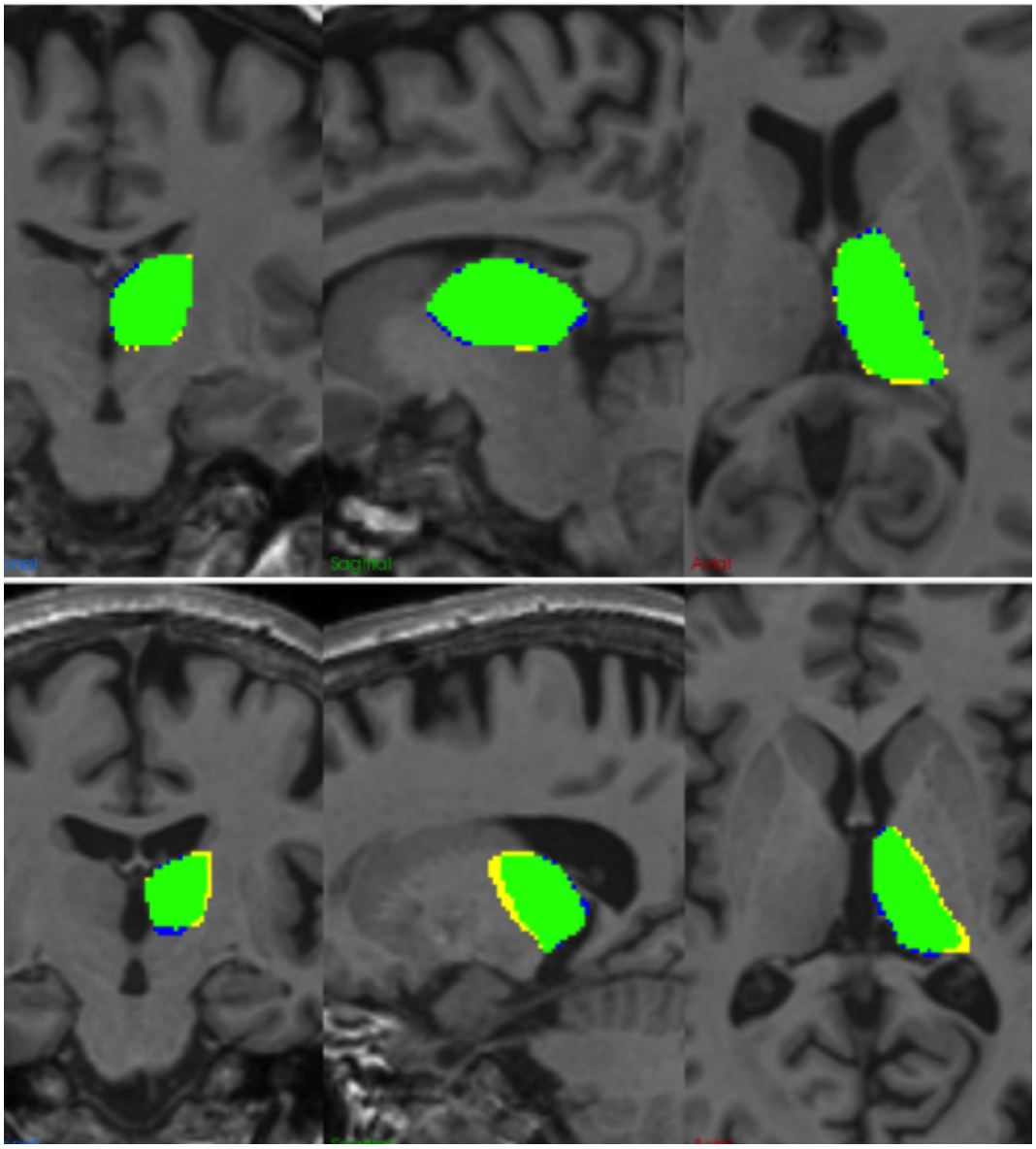


Figure 8.12: Best (top row) and worst (bottom row) automated segmentation (according to dice score) when using FA maps and T1-weighted scans. Blue: included in manual segmentation only, green: included in both manual and automated segmentations, yellow: included in automated segmentation only. The best segmentation had a dice score of 0.96 and the worst had a dice score of 0.90.

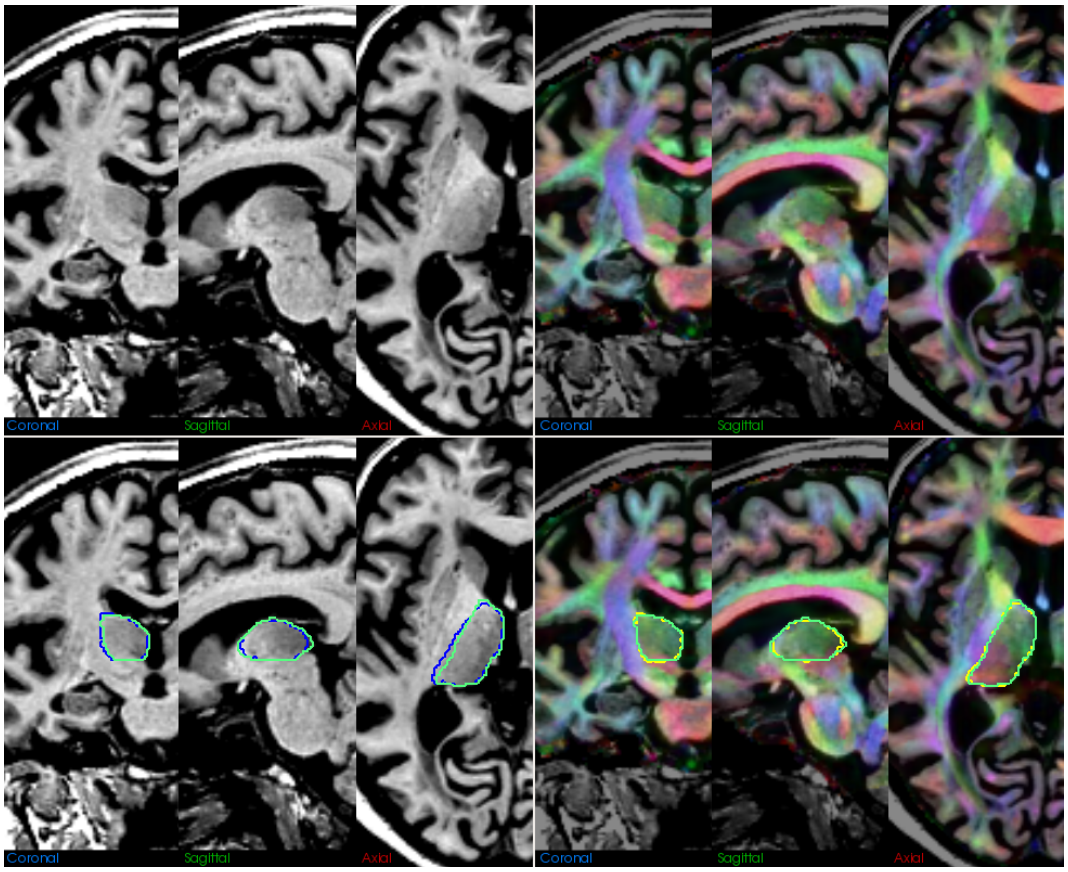


Figure 8.13: Scan with the biggest difference in dice scores between the automatically generated thalamic region using T1-weighted scans (shown in blue) and the automatically generated thalamic region using both FA maps and T1-weighted scans (shown in yellow). The manually segmented region is shown in green. The dice overlap between the manually segmented region and the region automatically segmented using T1-weighted scans only was 0.91 whilst for the manually segmented region and the automatically segmented region using both FA and T1-weighted scans it was 0.95.

8.4 Discussion

Good levels of reproducibility were achieved when segmenting thalami using T1-weighted scans only and when using coloured FA maps in addition to T1-weighted scans. Modest, but non-significant improvements in dice scores and mean surface distance were seen when comparing the manual segmentations based on T1-weighted scans alone with those based on T1-weighted and coloured FA maps.

The dice scores achieved from the automated segmentations in this study (0.94 when using T1-weighted scans only and 0.95 when using FA maps in addition to T1-weighed scans) were higher than dice scores reported in the literature for the thalamus. The most widely used thalamic segmentation method in AD studies, FIRST, reported a mean dice score of approximately 0.88 (number taken from graphical representation) [Patenaude et al., 2011]. The template library used by the FIRST algorithm includes scans from 336 subjects with range of pathologies, including Alzheimer's diease, as well as healthy controls. No information is given about the manual thalamic segmentation protocol that was followed nor the field strength of the MRI scanners used making comparisons difficult. The dice scores presented in this chapter were also higher than those achieved when using the STEPS algorithm in combination with a different template library 0.89 [Cardoso et al., 2015]. The template library used in [Cardoso et al., 2015] was a set of 30 manually-labelled brain scans from healthy volunteers aged between 20 and 54 years of age, described in detail in [Hammers et al., 2003,Hammers et al., 2007]. A description of the thalamic region definition used is given in [Hammers et al., 2003] and is accordance with what was included in the protocol developed in this chapter. Possible reasons for the superior dice score achieved in the study in this chapter could be the larger size of the template library used, better scan quality and tissue contrast (scans were acquired on 3T scanners in this study whilst those used in [Cardoso et al., 2015] were acquired on a 1.5T scanner [Hammers et al., 2003]) or more reliable manual segmentations (an example of one of the

manually labelled brains is given in [Cardoso et al., 2015] and although the slice shown does not include the thalamus, the regions shown appear to have somewhat irregular and biologically implausible boundaries). In addition, a modest, but statistically significant improvement in the mean dice score was observed when using both FA maps and T1-weighted maps in the automated pipeline as opposed to T1-weighted maps alone.

One recent study also utilized data from diffusion scans, in addition to T1-weighted and T2-weighted scans, to automatically segment the thalamus [Glaister et al., 2017]. In the method described by Glaister et al., a thalamic region of interest is first obtained, by using a multi-atlas segmentation approach. A set of features, incorporating information from diffusion, T1 and T2-weighted scans is then computed at each voxel within the ROI. This is then used to classify voxels as belonging to the thalamus, or not. Using this technique, the authors reported mean dice scores of 0.88 and 0.89 for the left and right thalami respectively, using leave-one-out cross validation. The dice scores achieved in leave-one-out cross-validation in this chapter were higher (0.94 using T1-weighted imaging only and 0.95 using FA-maps in addition to T1-weighted imaging), although a head-to-head comparison of the two techniques has not been performed.

In this study, a manual segmentation protocol with good reproducibility was developed. A template library of thalamic segmentations in subjects with a range of pathologies was generated based on this protocol. The use of this template library in combination with the STEPS algorithm was shown to produce good quality automatic thalamic segmentations.

Chapter 9

Investigation into thalamic volume and diffusion metrics in mild cognitive impairment and Alzheimer's disease

9.1 Introduction

As discussed in chapter 8, there is evidence that the thalamus may be affected by Alzheimer's disease pathology. Braak and Braak found that the anterodorsal nucleus was severely affected by Alzheimer's disease pathology post-mortem [Braak and Braak, 1991]. Imaging studies have shown reduced thalamic volumes in subjects with Alzheimer's disease [De Jong et al., 2008] and studies in pre-symptomatic mutation carriers for familial Alzheimer's disease have shown diffusion changes in the thalamus prior to symptom onset [Ryan et al., 2013]..

What is not known is at what stage in the disease process the thalami are affected in sporadic Alzheimer's disease, if there are thalamic diffusion differences in subjects with sporadic Alzheimer's disease (AD) and if diffusion metrics are predictive of subsequent atrophy. This may aid early diagnosis of disease and have implications for outcome measures of trials in early (preclinical) AD.

The aims of this study were to:

1. Investigate thalamic and hippocampal volumes in subjects with subjective memory complaints, early mild cognitive impairment, late mild cognitive impairment and established Alzheimer's disease. Hippocampal volumes were used as a comparator to assess whether the thalamus is affected at a similar stage to the hippocampus.

2. Investigate whether thalamic and hippocampal diffusion metrics can aid in group differentiation above volumetric measures.

3. Investigate whether diffusion metrics are predictive of subsequent whole-brain and hippocampal atrophy rates.

9.2 Methods

9.2.1 Study Participants

In this study, data from participants recruited to the the Alzheimer's disease neuroimaging initiative (ADNI)-GO and ADNI-2 phase of the ADNI study (described in chapter 3) were used. All newly-enrolled subjects in the ADNI-GO and ADNI-2 phases of the study were scanned on 3T magnetic resonance imaging (MRI) scanners. All newly-enrolled subjects with T1-weighted imaging were included in this study. A total of 973 subjects had accelerated T1-weighted MRI scans available at baseline. Accelerated scans were chosen over non-accelerated due to the findings reported in chapter 4. Subjects who were scanned on General Electric (GE) scanners had diffusion imaging available in addition to T1-weighted scans. A total of 218 subjects had both diffusion-weighted and T1-weighted imaging available at baseline.

9.2.2 MRI Acquisition

Details of the MRI protocol used in this study can be found in chapter 3. Diffusion weighted MRI scans were acquired in the same scanning session as T1-weighted volumetric images for subjects scanned on GE scanners. Details of the diffusion protocol and processing steps can be found in chapter 3.

9.2.3 Quality control

All scans were assessed for motion and any with severe motion were excluded from analyses.

9.2.4 Total intracranial volume, whole-brain, hippocampal and thalamic region segmentation

total intracranial volume (TIV)'s were calculated using the method described in chapter 3. Brains were segmented in native space using an automated method described in chapter 3.4.1. Hippocampi were automatically segmented using template library 2 as described in chapter 7. Thalami were automatically segmented using the method described in chapter 8.

9.2.5 Extraction of diffusion metrics within regions-of-interest (region of interest (ROI)s)

Both the thalamus and the hippocampus have boundaries with cerebrospinal fluid (CSF). In order to ensure that the diffusion metrics in the regions-of-interest were not biased by the inclusion of some CSF in the boundary voxels, all regions were eroded once using a 3x3x3 voxel kernel, the smallest possible sized kernel for symmetric erosion. All of the thalamic and hippocampal diffusion metrics reported in this chapter are the mean diffusion metrics in those eroded regions.

9.2.6 Whole-brain and hippocampal atrophy rates

Whole-brain and hippocampal atrophy rates were calculated using the methods described in chapter 3.

9.2.7 Statistical Analyses

All statistical analyses were performed in Stata (version 12).

9.2.7.1 Baseline demographics

Linear regression with age as the dependent variable and diagnostic group as the independent variable was used to compare age across diagnostic groups.

Linear regression with bootstrapping was used to compare mini-mental state exam (MMSE) scores across the diagnostic groups. Bootstrapping was used as MMSE is not normally distributed and has a maximum value of 30. Differences in the the proportions of male participants and apolipoprotein (APOE) $\varepsilon 4$ carriers were compared across the diagnositic groups using chi-square tests.

9.2.7.2 Comparison of subjects in diffusion subset with full ADNI-2/GO cohort

In order to acertain whether the group with diffusion imaging were a representative subset of the whole dataset group characteristics between those who had diffusion imaging available at baseline and those who did not were assessed. Linear regression analyses were used to compare baseline age, brain, hippocampal and thalamic volumes. Logistic regression was used to compare percentages of males and APOE $\varepsilon 4$ carriers and linear regression with bootstrapping to compare baseline MMSE scores as this score has a ceiling effect (subjects can score a maximum of 30).

9.2.7.3 Comparison of baseline volumes

In order to compare baseline whole-brain, thalamic and hippocampal volumes between diagnostic groups linear regression analyses were used. The dependent variable was the volume of interest with diagnostic group diagnostic group, age, gender and head-size (TIV) as independent variables.

9.2.7.4 Comparison of baseline hippocampal and thalamic diffusion metrics

Linear regression analyses were used to compare baseline diffusion metrics where the dependent variable was the mean diffusivity or fractional anisotropy in the ROI with the independent variable being diagnostic group and the models adjusted for age and gender. A second set of regression analyses were fitted, additionally adjusting for the (non-eroded) volume of the ROI. This was in order to see whether any significant differences by diagnostic group remained after accounting for volume differences.

9.2.7.5 Classification performance using diffusion metrics in addition to volumes

To investigate whether diffusion metrics improve diagnostic classification, logistic regression analyses were performed. Four binary variables were generated for the different disease group comparisons: AD vs controls, LMCI vs controls, early mild cognitive impairment (EMCI) vs controls and subjective memory complaints (SMC) vs controls. These were then used as the dependent variables in the logistic regression analyses. In order to elucidate whether the use of diffusion metrics in addition to volumetric measures improves diagnostic accuracy two logistic regression models for each disease group comparison were fitted, one with left and right hippocampal and thalamic volumes (divided by TIV) and the other with left and right hippocampal and thalamic mean diffusivity and fractional anisotropy metrics in addition to the normalised hippocampal and thalamic volumes. To compare the ability of the models to discriminate between cases and controls receiver operator curve (ROC) curves were plotted and the area under the ROC curve for each model compared.

9.2.7.6 Predictors of subsequent atrophy

To investigate whether baseline diffusion metrics in the thalami and hippocampi were associated with atrophy rates measured between baseline and 12 months, linear regression models were fitted with annualised whole-brain or hippocampal atrophy rate (as a percentage of the baseline volume) as the dependent variable, the diffusion metric of interest (mean diffusivity (MD) or fractional anisotropy (FA) in the bilateral thalami or hippocampi) as the independent variable and adjusted for age and gender and diagnostic group. An interaction term between diagnostic group and the diffusion metric of interest was included to allow for differences in the effect of the diffusion metric on atrophy rates between the diagnostic groups.

To investigate whether baseline hippocampal and thalamic volumes were associated with atrophy rates measured between baseline and 12 months, linear regression analyses were fitted with annualised whole-brain or hippocampal

atrophy rate as the dependent variable, the volume of interest (left or right hippocampal or thalamic volume) as the independent variable and adjusted for age, gender, diagnostic group and TIV. An interaction term between diagnostic group and the volume of interest was included to allow for differences in the effect of baseline volume on atrophy rates between the diagnostic groups.

9.3 Results

9.3.1 Quality control

Of the 973 subjects who had data available for download at baseline, 943 had T1-weighted scans that passed internal dementia research centre (DRC) quality control (QC) with 218 of these subjects having additional diffusion weighted scans available.

Examples of scans which had levels of motion deemed unacceptable for analysis are shown in figure 9.1. Examples of passing scans are shown in figure 9.2.

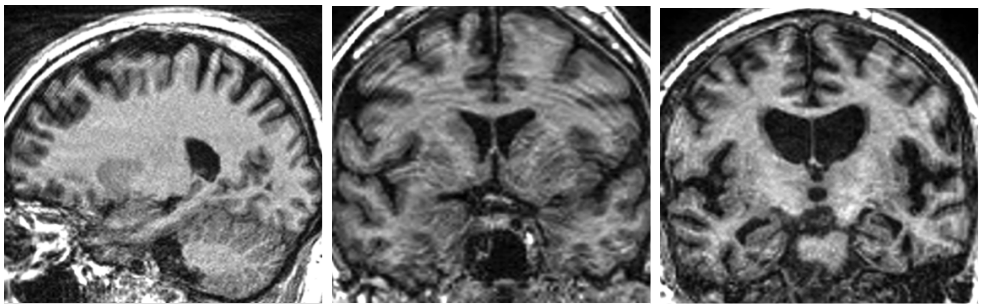


Figure 9.1: Examples of scans with motion artefacts severe enough for exclusion from the study.

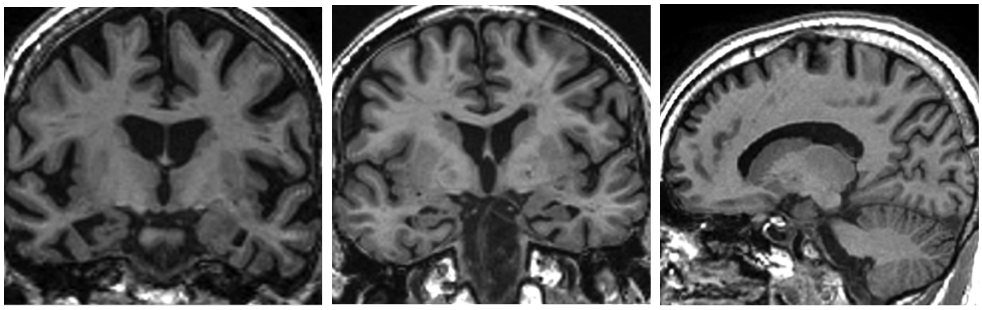


Figure 9.2: Examples of scans of good quality with little motion evident.

One control subject was excluded from analyses as pathology was noted in the left thalamus. This subject likely had an incidental thalamic arachnoid cyst (see figure 9.3).

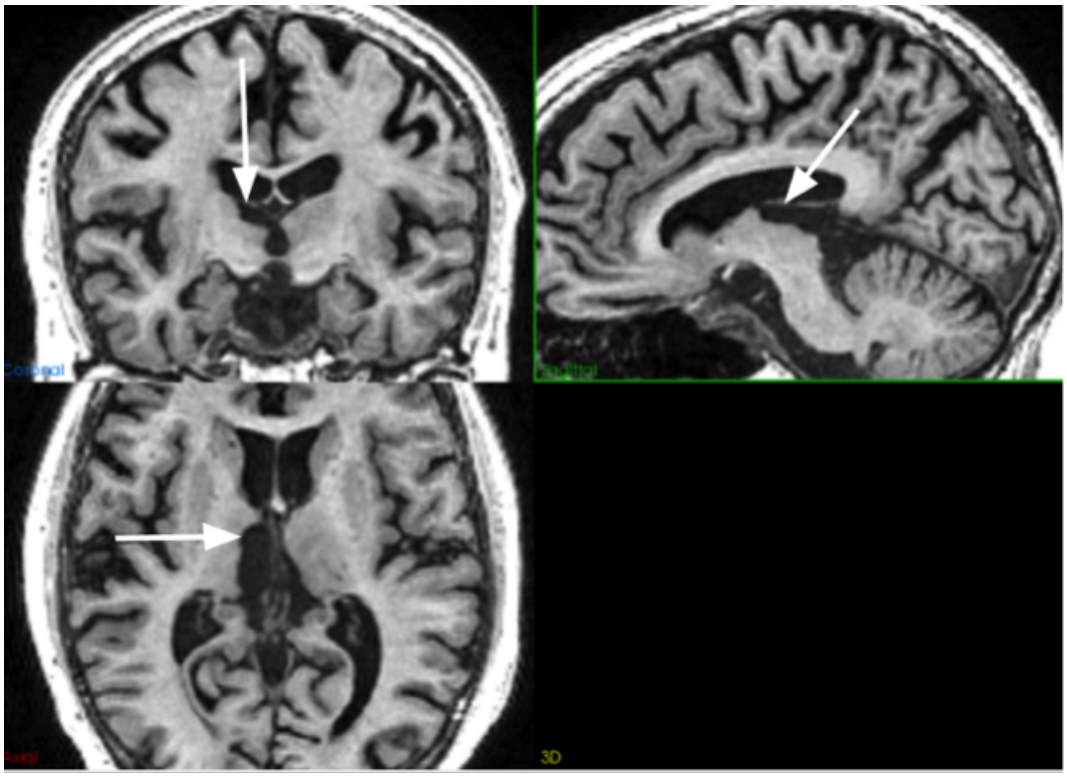


Figure 9.3: Subject with likely incidental arachnoid cyst in left thalamus. Arrows indicate location of arachnoid cyst.

9.3.2 Participant Demographics

Table 9.1 shows baseline characteristics by diagnostic group for all the ADNI-2/GO subjects with T1-weighted scans available at baseline.

Table 9.1: Baseline characteristics of ADNI-2/GO cohort.
Mean (sd) shown unless otherwise indicated.

	N	Age (years)	MMSE (/30)	Percent male	Percent APOE $\varepsilon 4$ carriers
NL	193	73.2 (6.3)	29.0 (1.2)	54	29
SMC	105	72.2 (5.6)	29.0 (1.3)	58	33
EMCI	321	71.0 (7.4)	28.3 (1.5)	47	42
LMCI	173	72.2 (7.6)	27.6 (1.8)	46	58
AD	150	74.9 (8.0)	23.1 (2.1)	42	67
P-value		<0.001[a]	<0.001[b]	0.05[c]	<0.001[d]

NL: controls, SMC: subjective memory complaints, EMCI: early mild cognitive impairment, LMCI: late mild cognitive impairment, AD: Alzheimer's disease.
[a] linear regression
[b] linear regression with bootstrapping
[c] chi-square test
[d] chi-square test

218 subjects had diffusion available (32 NL, 26 SMC, 65 EMCI, 36 LMCI and 48 AD). Figure9.4 shows a comparison of the baseline characteristics and brain volumes between those who had diffusion imaging available at baseline and those who did not. Those who had diffusion imaging available and were diagnosed as EMCI were significantly older than those who did not have diffusion imaging available at baseline. There was a significantly higher percentage of APOE $\varepsilon 4$ carriers in the late mild cognitive impairment (LMCI) group who had diffusion imaging compared with those who did not. There were no significant differences in baseline MMSE scores or gender ratios between any of the groups. Mean adjusted brain volumes were significantly smaller in the subset of subjects with diffusion imaging available in controls, SMC, EMCI and AD subjects than in those without, whilst there was no difference in baseline brain volume in LMCI subjects between the two groups.

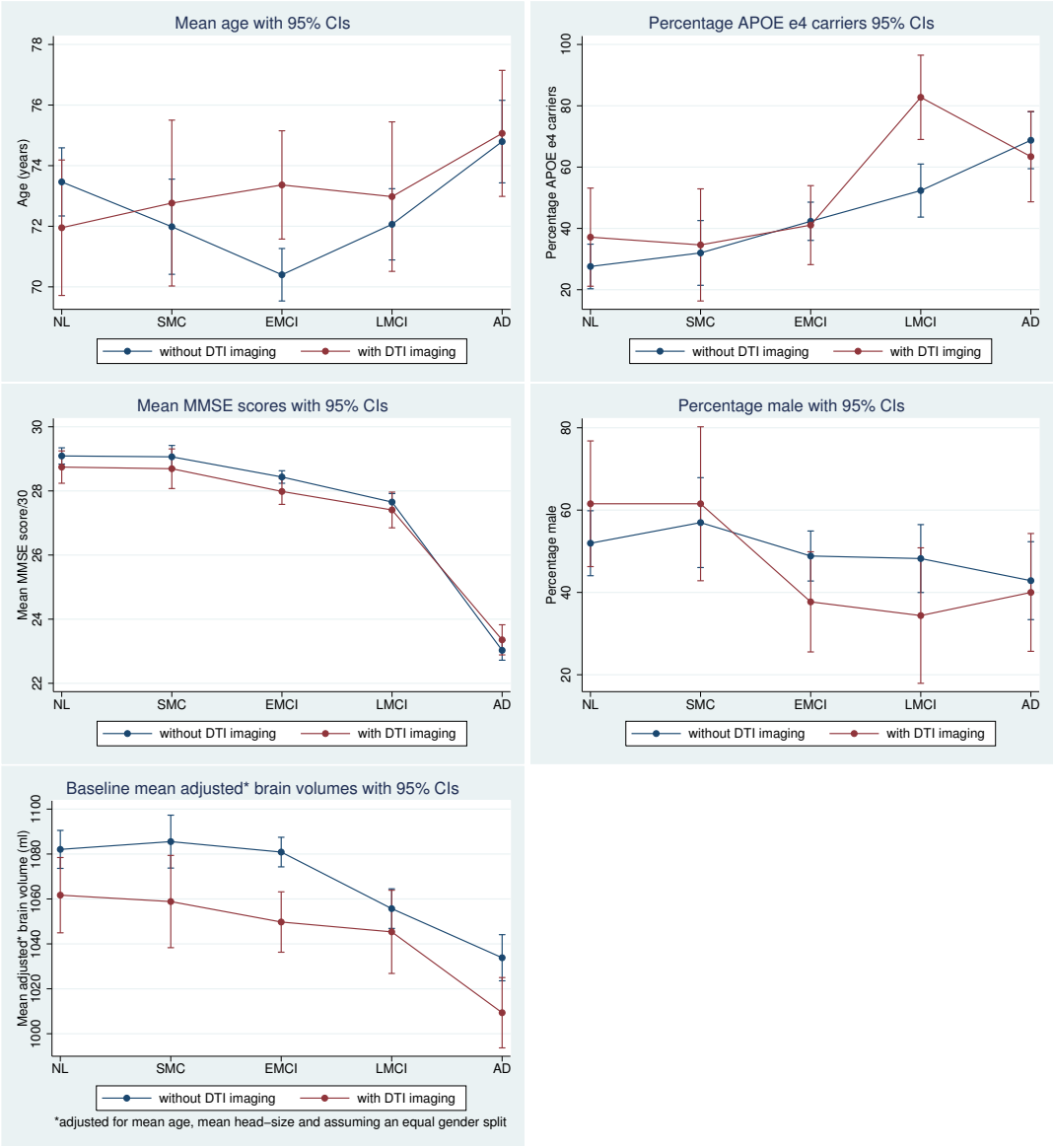


Figure 9.4: Comparison of baseline characteristics and volumes between those who had diffusion imaging available at baseline and those who did not. NL: controls, SMC: subjective memory complainers, EMCI: early mild cognitive impairment, LMCI: late mild cognitive impairment, AD: Alzheimer's disease.

9.3.3 Baseline volumes

9.3.3.1 In the whole dataset

When comparing volumes in the whole group, mean adjusted whole-brain volumes were found to be significantly smaller in LMCI and AD subjects as compared to controls, whilst there was no significant difference in brain volume

in EMCI subjects or SMC subjects with respect to controls. Mean adjusted hippocampal volumes were significantly smaller in EMCI, LMCI and AD subjects with respect to controls on both the right and left sides. Mean adjusted thalamic volumes were also smaller in EMCI, LMCI and AD on the left side and in LMCI and AD on the right side. Figure 9.5 shows the mean adjusted baseline volumes in the full ADNI-2/GO cohort.

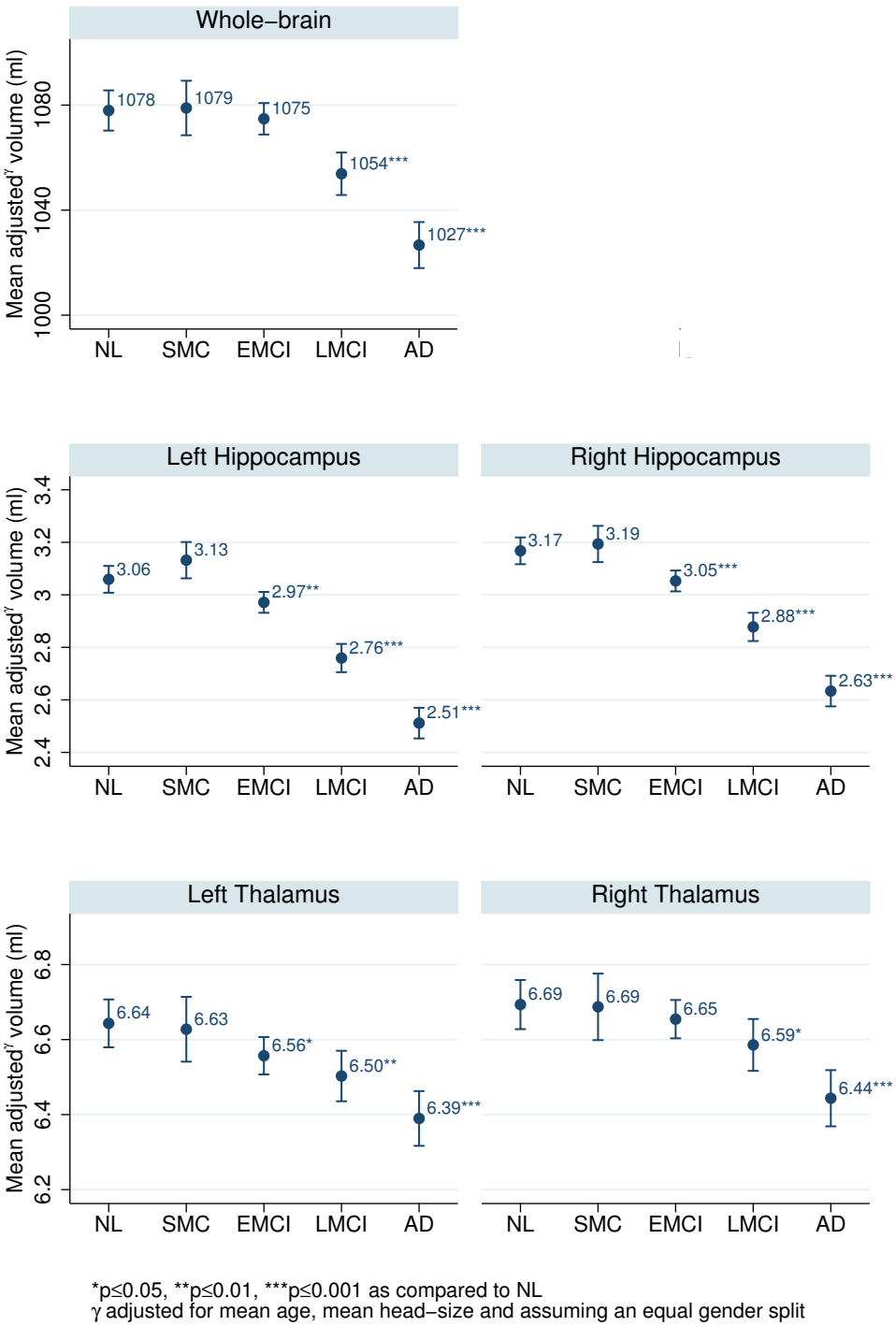


Figure 9.5: Baseline mean adjusted volumes in the full ADNI-2/GO cohort. Volumes are adjusted for mean age, mean head-size and assume an equal gender split. Error bars show 95% confidence intervals. Statistically significant differences are represented by asterisks and represent differences with respect to NL subjects. *p≤0.05, **p≤0.01, ***p≤0.001. NL=controls, SMC=subjective memory complaints, EMCI=early mild cognitive impairment, LMCI=late mild cognitive impairment, AD=Alzheimer's disease.

9.3.3.2 In the subset with diffusion imaging

In the subset of subjects who had diffusion imaging available at baseline, mean adjusted whole-brain volumes were lower in the LMCI and AD subjects. Lower hippocampal volumes were observed bilaterally in EMCI, LMCI and AD whilst lower thalamic volumes were only observed in the AD group compared with controls and no significant difference in volume was found between the LMCI, EMCI or SMC groups and controls. Figure 9.6 shows the baseline volumes in ADNI-2/GO subset with diffusion tensor imaging (DTI) imaging. Given the imbalance in APOE $\varepsilon 4$ carriers in the group with LMCI compared with the whole dataset, post-hoc adjustment for APOE $\varepsilon 4$ carrier status was performed, but no significant differences to the findings were observed. The results presented in figure 9.6 are without adjustment for APOE $\varepsilon 4$ as fewer subjects had APOE $\varepsilon 4$ carrier status available.

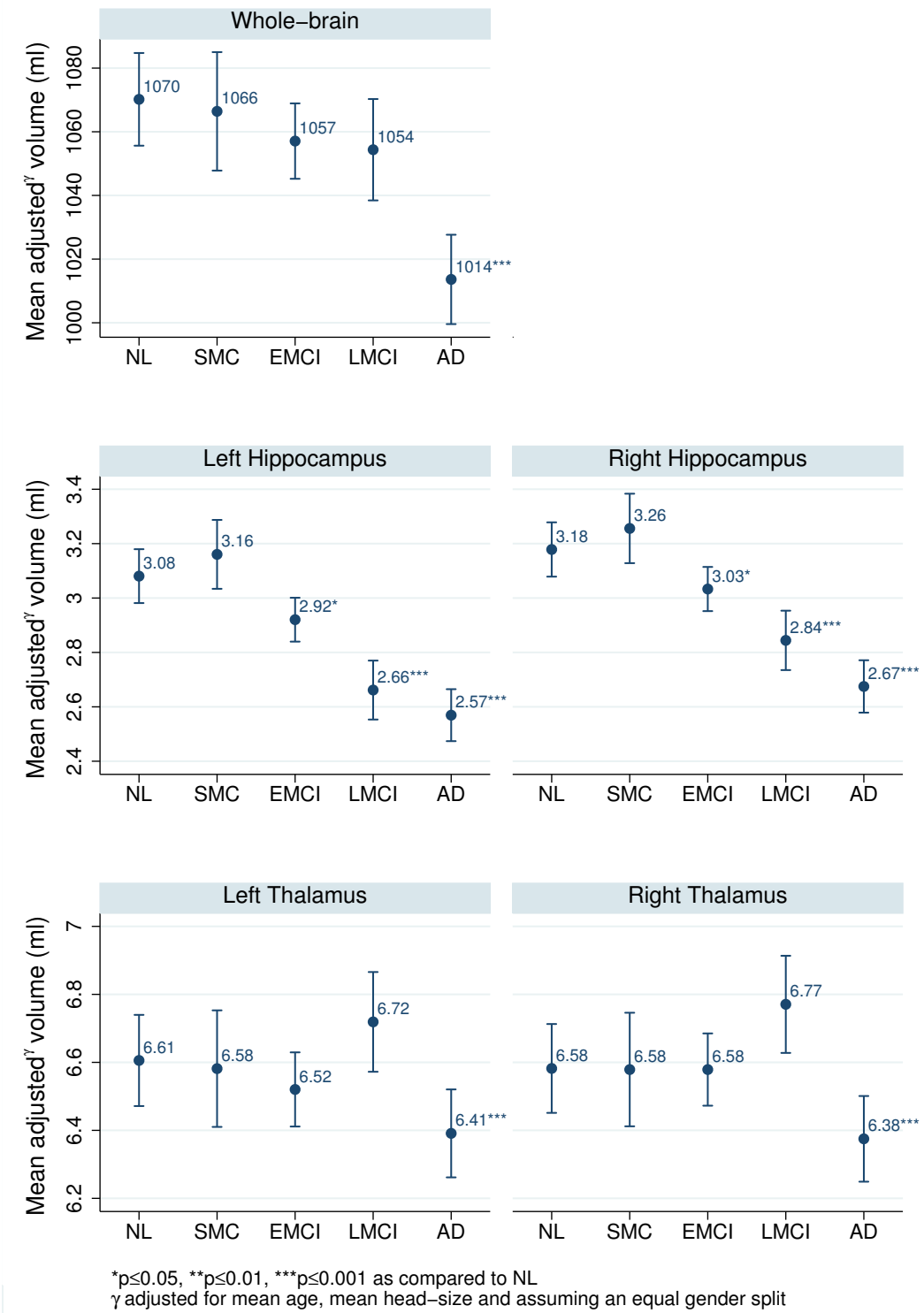
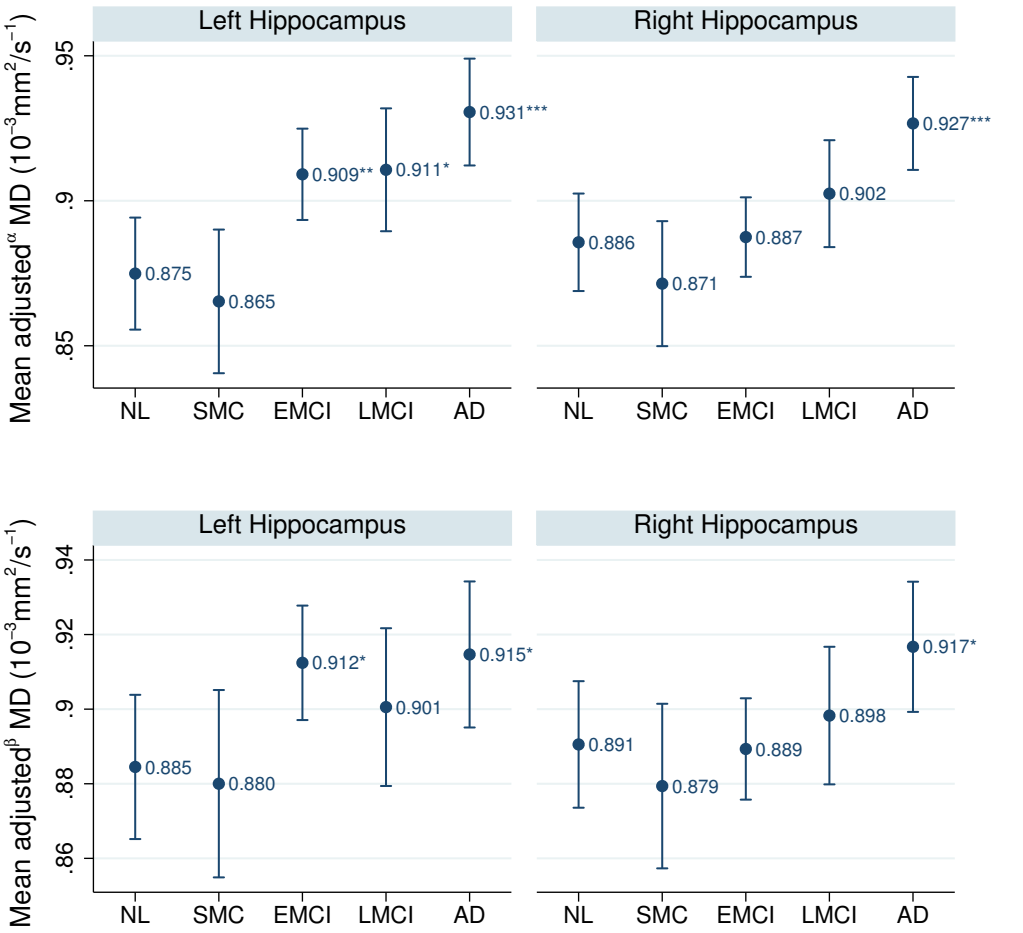


Figure 9.6: Baseline mean adjusted volumes in the subset of subjects with diffusion imaging available. Volumes are adjusted for mean age, mean head-size and assume an equal gender split. Error bars show 95% confidence intervals. Statistically significant differences are represented by asterisks and represent differences with respect to NL subjects. *p≤0.05, **p≤0.01, ***p≤0.001. NL=controls, SMC=subjective memory complaints, EMCI=early mild cognitive impairment, LMCI=late mild cognitive impairment, AD=Alzheimer's disease.

9.3.4 Diffusion metrics

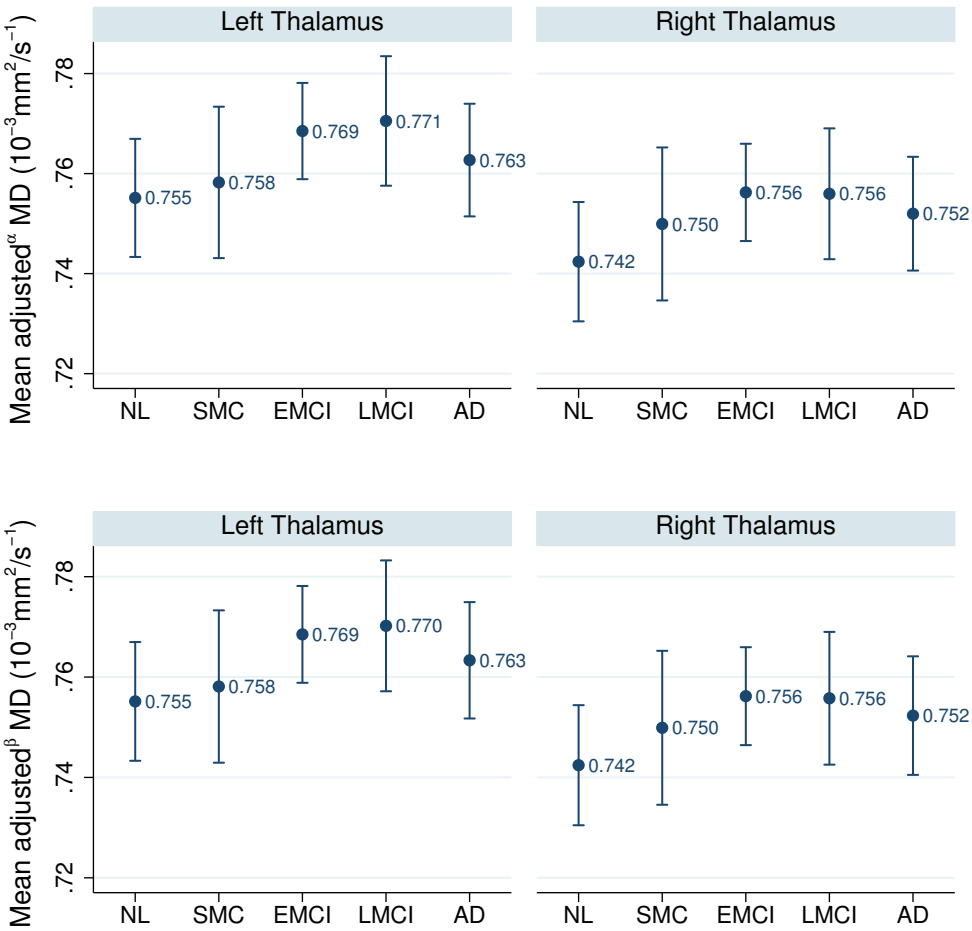
Significantly higher mean adjusted (for age and gender) MD values were observed in the left hippocampus in EMCI and LMCI subjects as compared with controls and in both the left and right hippocampi in AD subjects as compared with controls. When adjusting for hippocampal volume in addition to age and gender, EMCI subjects had significantly higher hippocampal mean diffusivity on left-side as compared with controls and AD subjects had significantly higher mean diffusivity in the left and right hippocampi. See figure 9.7.



α adjusted for mean age and assuming an equal gender split
β adjusted for mean age, mean volume (left or right hippocampus) and assuming an equal gender split
*p≤0.05 compared with controls

Figure 9.7: Mean MD in left and right hippocampus. Error bars show 95% confidence intervals.

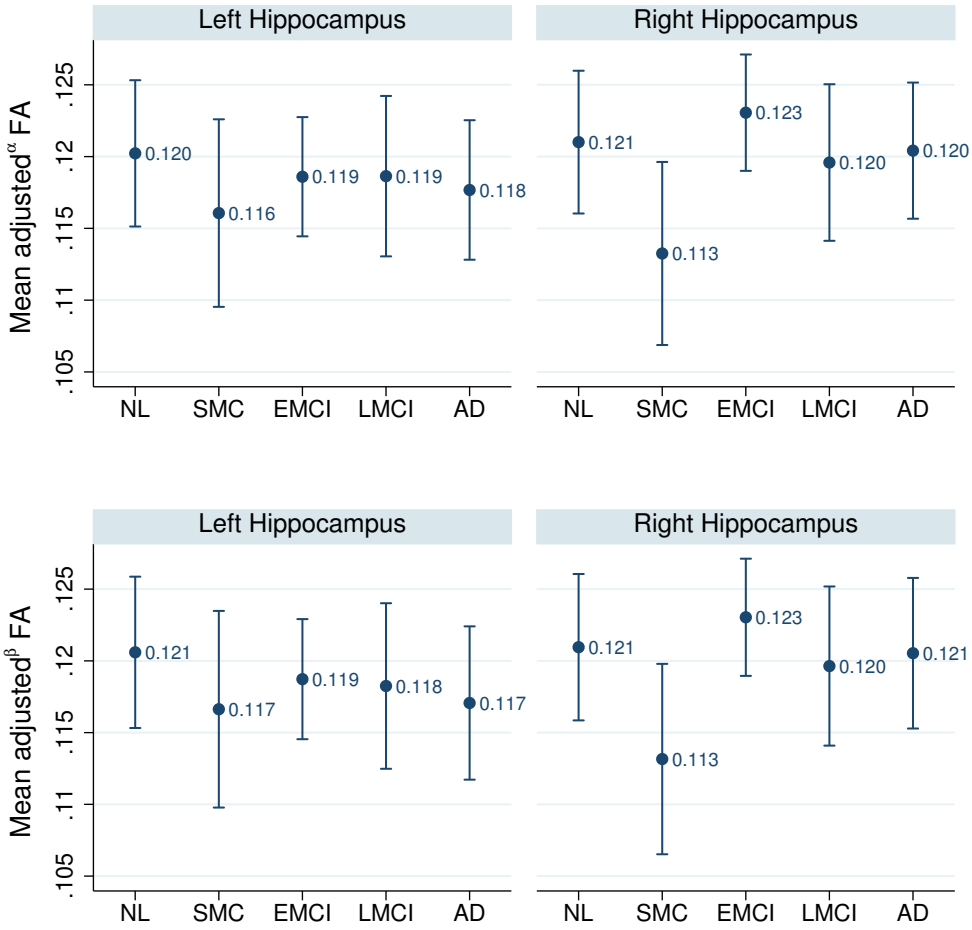
Subjects with SMC had slightly elevated mean adjusted MD in both the left and right thalami compared with controls, although the differences did not reach statistical significance. EMCI subjects had an elevated mean adjusted MD in the thalami and the difference was statistically significant on the right side and borderline significant on the left side (p=0.055). LMCI subjects had slightly elevated MD values as compared to controls, although the differences were smaller than in EMCI subjects and were not statistically significant. AD subjects again had slightly elevated MD values as compared with controls, and the differences again were not statistically significant. See figure 9.8.



α adjusted for mean age and assuming an equal gender split
β adjusted for mean age, mean volume (left or right thalamus) and assuming an equal gender split
*p≤0.05 compared with controls

Figure 9.8: Mean MD in left and right thalamus. Error bars show 95% confidence intervals.

There was no significant difference in mean adjusted FA values in the left or right side hippocampi in any diagnostic group (see figure 9.9). In the thalamus, AD subjects had significantly increased mean adjusted FA values compared with controls (see figure 9.10).

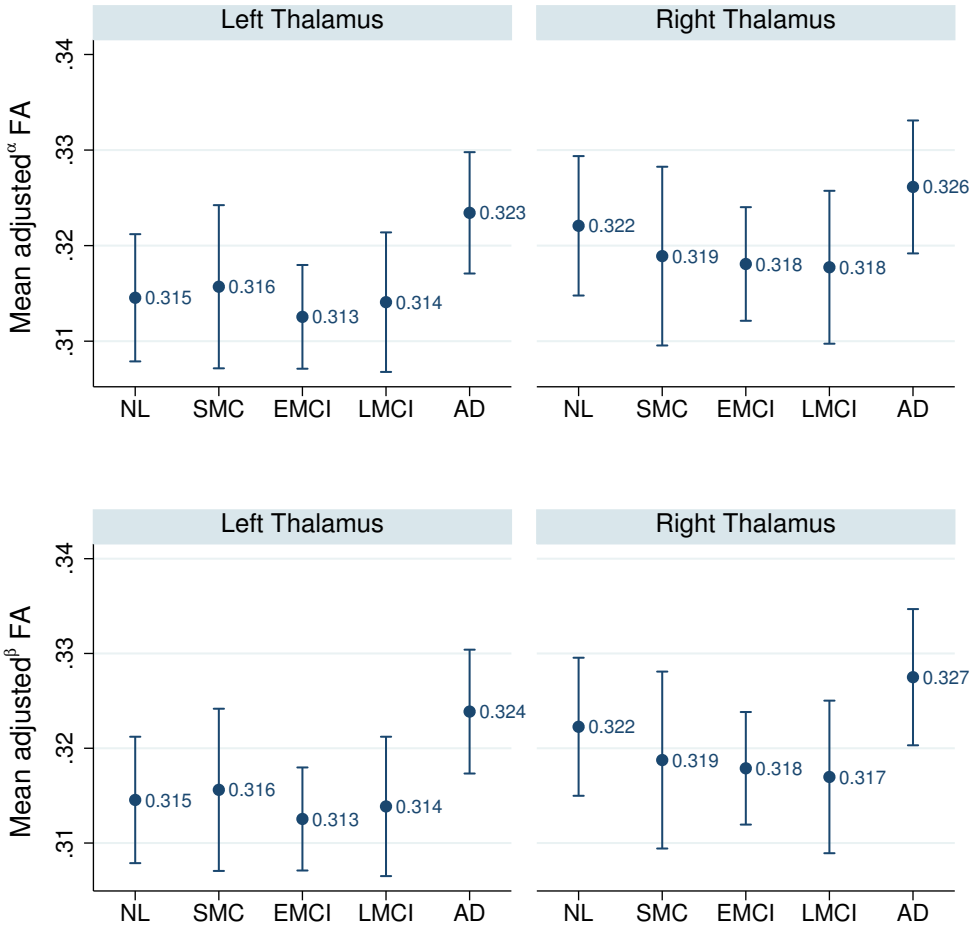


α adjusted for mean age and assuming an equal gender split

β adjusted for mean age, mean volume (left or right hippocampus) and assuming an equal gender split

*p≤0.05 compared with controls

Figure 9.9: Mean FA in left and right hippocampus. Error bars show 95% confidence intervals.



α adjusted for mean age and assuming an equal gender split
β adjusted for mean age, mean volume (left or right thalamus) and assuming an equal gender split
*p≤0.05 compared with controls

Figure 9.10: Mean FA in left and right thalamus. Error bars show 95% confidence intervals.

9.3.5 Disease classification using diffusion metrics in addition to volumetric metrics

When comparing classification performance of hippocampal and thalamic volumes alone or with additional diffusion metrics, it was found that classification performance was improved with diffusion metrics (hippocampal and thalamic FA and MD). This was significant in EMCI vs. controls and AD vs. controls. A comparison of the ROC curves with and without diffusion metrics are shown in figure 9.11. Table 9.2 summarizes the mean area under the ROC curve for each comparison.

p2cmC2cmC2cmC2cmC2cmC0.9cmC0.9cmC0.9cm

	SMC vs controls	EMCI vs controls	LMCI vs controls	AD vs controls
Model 1	0.58 [0.44, 0.73]	0.69 [0.59 ,0.79]	0.88 [0.820, 0.96]	0.91 [0.85, 0.97]
Model 2	0.72 [0.59, 0.85]	0.81* [0.73, 0.89]	0.92* [0.86, 0.99]	0.96* [0.93 ,0.99]

Table 9.2: Area under ROC curves [95% confidence intervals]. *$p \leq 0.05$ as compared with Model 1. Model 1: variables included left and right hippocampal and thalamic volumes (normalised for head-size). Model 2: variables included left and right hippocampal and thalamic MD and FA values in addition to volumes normalised for head-size.

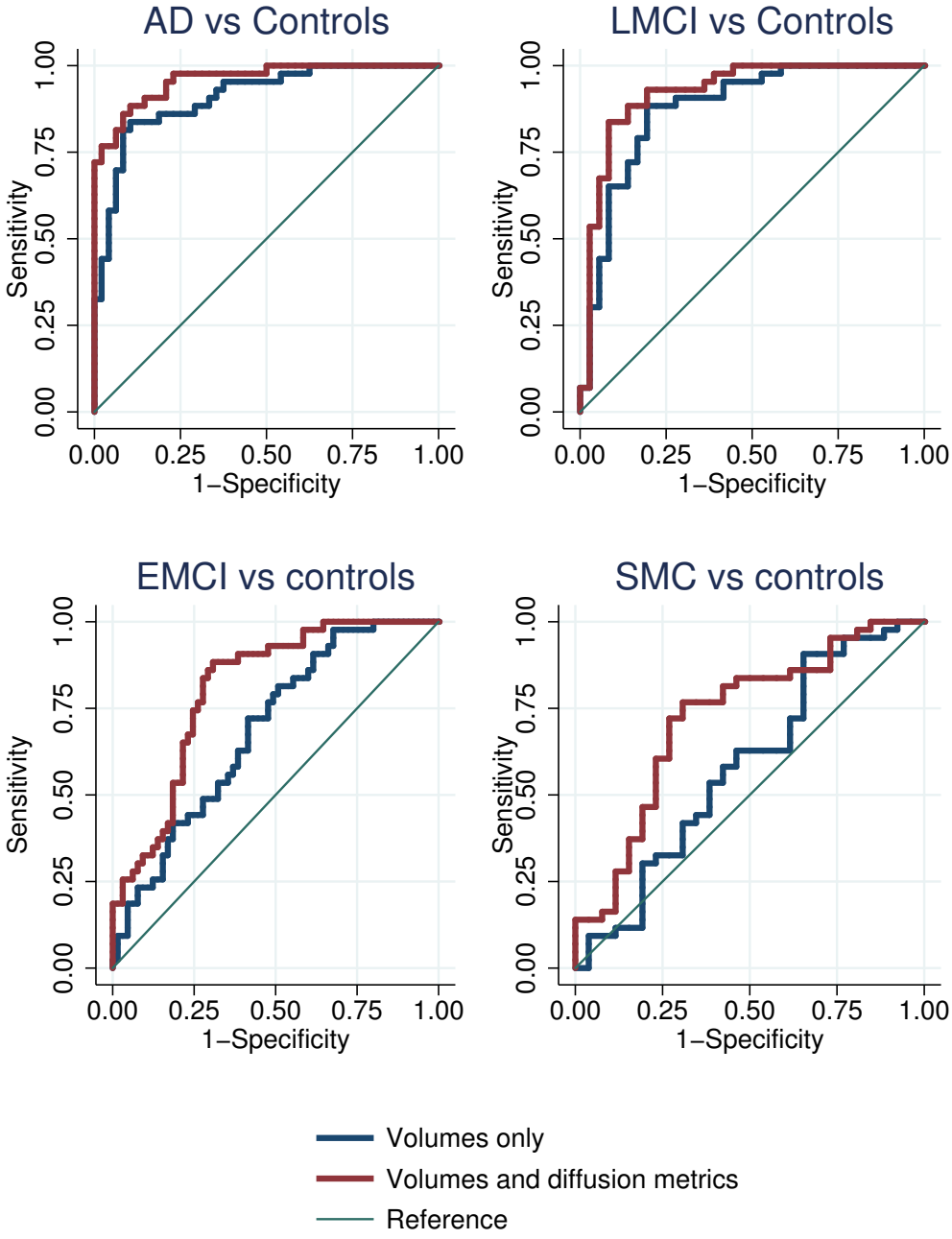


Figure 9.11: ROC curve comparison for disease group classification. Model 1, shown in blue, used left and right hippocampal and thalamic volumes (normalised for head-size) as predictor variables. Model 2, shown in red, included left and right hippocampal and thalamic diffusion metrics (FA and MD) in addition to the volumetric metrics included in Model 1.

9.3.6 Predictors of atrophy rate

In controls, increased bilateral hippocampal MD at baseline was associated with higher rates of hippocampal atrophy (an increase in hippocampal MD of 1 standard deviation was associated with a 2.2%/year increase in hippocampal atrophy rate, p=0.04). No significant association was found in the other diagnostic groups (p>0.2 for all other comparisons)).

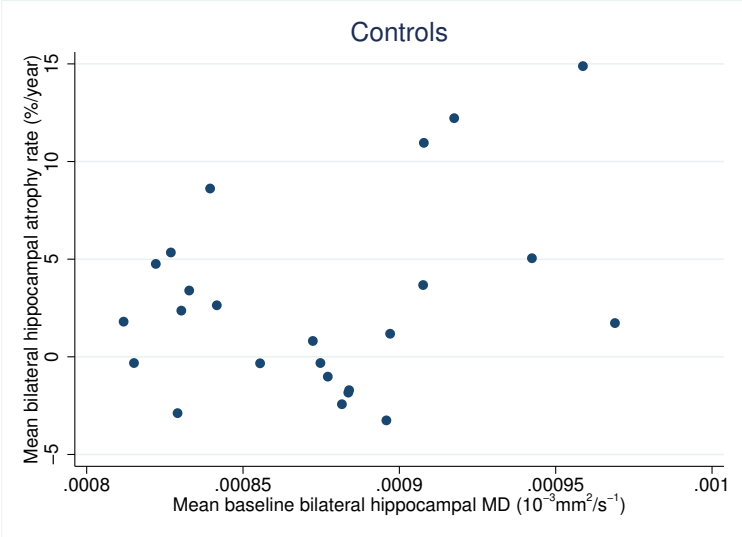


Figure 9.12: Baseline hippocampal MD vs hippocampal atrophy rate

In EMCI subjects, there was borderline evidence that increased baseline bilateral thalamic MD values were associated with lower rates of whole-brain atrophy (an increase in thalamic MD of 1 standard deviation was associated with a 0.25% decrease in whole-brain atrophy, p=0.051). See figure 9.13. No significant associations were found in the other diagnostic groups (p>0.5).

Increased bilateral thalamic FA at baseline was associated with increased subsequent whole-brain atrophy in subjects with EMCI (an increase in thalamic FA of 1 standard deviation was associated with an increase of 0.28%/year whole-brain atrophy rate, p=0.009), but not in any of the other disease groups (p>0.1). Baseline hippocampal FA metrics were not associated with whole-brain atrophy rates in any disease group (p>0.2).

Lower baseline left and right hippocampal volumes were associated with higher rates of whole-brain atrophy in LMCI subjects and there was borderline

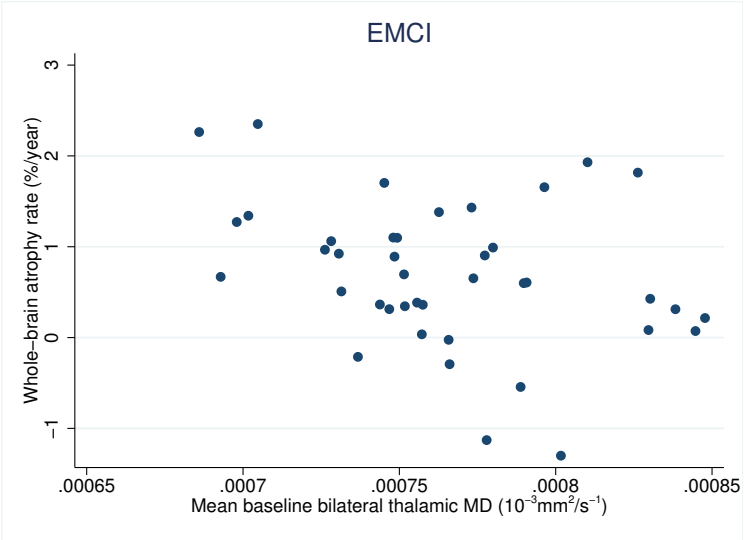


Figure 9.13: Baseline thalamic MD and whole-brain atrophy rate

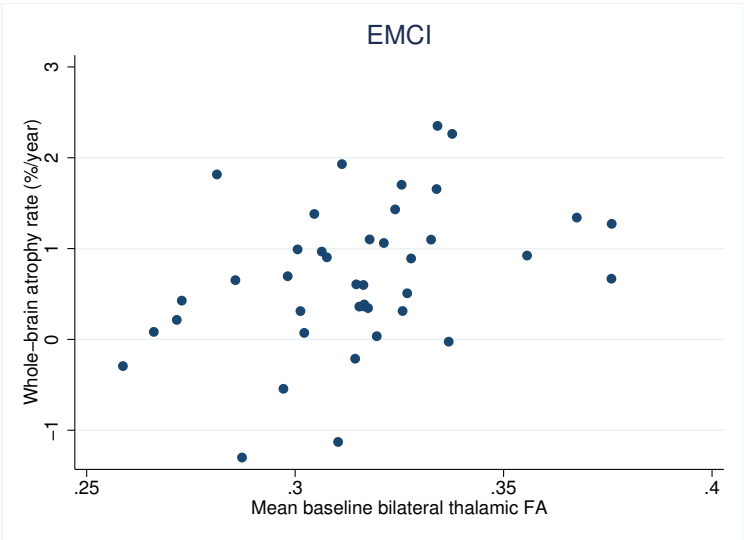


Figure 9.14: Baseline thalamic FA vs whole-brain atrophy rate

evidence that lower baseline left and right hippocampal volumes were associated with higher rates of hippocampal atrophy in EMCI and LMCI subjects (p=0.06 for both). Baseline thalamic volumes were not associated with either hippocampal or whole-brain atrophy rates in any of the disease groups.

9.4 Discussion

In this chapter thalamic and hippocampal volume and diffusion differences in subjects with SMC, EMCI, LMCI and AD were investigated using the ADNI-

2/GO dataset.

In the whole dataset, as expected, reduced hippocampal volumes were found in AD, LMCI and EMCI subjects. Thalamic volumes were significantly smaller in AD and LMCI subjects than in controls on both left and right sides and were significantly smaller than controls on the left side in EMCI subjects.

In the subset with diffusion imaging, hippocampal MD values were higher in AD, LMCI and EMCI than controls on the left side and higher in AD than controls on the right side. In the thalamus, MD values were slightly higher in EMCI and LMCI on the left side (trend level $p<0.1$) and in EMCI subjects on the right side (trend level $p< 0.1$). When adjusting for volume, some of the MD diffusion differences remained, suggesting that diffusion is providing different information compared with macroscopic T1-weighted volume differences. FA appeared to be a less promising measure in terms of group differentiation: there were no significant differences in the hippocampi and or the thalami (although thalamic FA values were slightly higher in AD subjects compared with controls on the left side, $p=0.06$). In terms of diagnostic value, there was evidence that diffusion metrics aided group differentiation in EMCI, LMCI and AD. Diffusion metrics also aided in group differentiation in SMC compared with controls although the difference in the area under the ROC curve did not reach statistical significance. Higher baseline MD in the hippocampi was associated with increased hippocampal atrophy rates in controls. It may be that the this represents the effects of normal ageing in this group of individuals, or possibly it could be an indication of very early dementia in those with higher baseline hippocampal MD, although further studies would be required to confirm this hypothesis. In EMCI subjects higher baseline thalamic FA was associated with higher whole-brain atrophy rates and there was borderline evidence that lower baseline thalamic MD was associated with higher rates of whole-brain atrophy whereas baseline thalamic volumes were not. These findings suggest that diffusion markers are useful prognostic markers early in the disease course. As expected, we found hippocampal volume to predict volume loss in LMCI

and EMCI subjects. However, extrapolating results from the diffusion subset to the whole of ADNI is problematic since there was also evidence that this was not a random subset: age differed in EMCI, APOE carrier numbers differed in LMCI.

There were some discrepancies in thalamic volume findings between the subset and the full ADNI2/GO cohort. Only the AD group showed significantly reduced mean adjusted thalamic volumes at baseline in the diffusion subset whilst reduced thalamic volumes were also seen in mild cognitive impairment (MCI) subjects in the full ADNI2/GO cohort. Only subjects who were scanned on GE scanners had DTI imaging available. Differences in baseline volumes between those scanned on GE scanners and those scanned on other scanners were found. This suggests that the subset with DTI imaging is not a random sample of the full the full ADNI2/GO cohort and results must be treated with necessary caution.

A number of previous studies have also reported reduced thalamic volumes in AD [Stěpán-Buksakowska et al., 2013, Yi et al., 2015, Möller et al., 2015, De Jong et al., 2008]. In terms of thalamic volumes in MCI, two studies with relatively small numbers of subjects (14 MCI subjects each) found singificantly reduced thalamic volumes in MCI as compared with controls [Eustache et al., 2016, Karas et al., 2004] as well a much larger, more recent study, with 201 MCI subjects [Yi et al., 2015]. The findings in this chapter are in keeping with these previous studies.

In the hippocampi, mean diffusivity was significantly higher in all disease groups as compared with controls whilst in the thalamus mean diffusivity was significantly higher in EMCI as compared with controls. Qualitatively, the relationship between disease stage and MD in the hippocampus and thalamus differ. MD in the hippocampus was higher with increased disease severity. MD in the thalamus was slightly higher in the early disease stages and relatively lower at later disease stages. These qualitative differences contrast with the largely similar volume patterns in both the hippocampus and thalamus

(later disease stages were associated with lower volumes). One previous study reported increased MD in the thalamus in AD compared to MCI (but they found no difference in MD between AD and controls nor between controls and MCI) [Cherubini et al., 2010]. Another previous study found no significant difference in thalamic MD between MCI and controls [Eustache et al., 2016].

The finding that thalamic FA was slightly higher in AD subjects as compared to controls (on the left side), but not significantly different in SMC, EMCI or LMCI subjects, was in keeping with the literature. Previous studies in familial AD have found increased FA in the thalamus in presymptomatic mutation carriers [Ryan et al., 2013]. The thalamus is organised into distinct nuclei with differing fibre orientations [Wiegell et al., 2003a]. If only specific nuclei are affected, as suggested in the literature [Aggleton et al., 2016], it is possible that the mean FA in the the thalamus as a whole actually increases. Since a proportion of the individuals included in the SMC, EMCI and LMCI groups will not go on to develop AD, it may be that the increased thalamic FA in these disease groups is attenuated.

Increased baseline left and right thalamic FA was associated with increased whole-brain atrophy at 12 months in subjects with EMCI and increased baseline hippocampal MD was associated with higher hippocampal atrophy rates in controls subjects. Previous longitudinal studies in subjects with MCI have found those who went on to convert to a diagnosis of AD had significantly higher MD in the hippocampi at baseline than those who did not [Müller et al., 2007,Kantarci et al., 2010,Douaud et al., 2011]. It is possible that increased baseline hippocampal MD could reflect a reduction in cellular integrity and be a very early indication of neurodegeneration in these controls, although to confirm this, a longer follow-up would be required in order to see who goes on to convert to a diagnosis of MCI or AD.

There were a number of limitations to this study. Firstly, the diffusion scans in the ADNI-2/GO cohort did not have field maps available, which may have had an impact on the quality of the registration with the T1 weighted

MRI scans. Secondly, this study was largely (apart from atrophy rates) cross-sectional in nature. In order to get a better insight into the dynamic changes occurring, in structures like the thalamus, a longitudinal study would be required. Thirdly, since the literature suggests that the anterior thalamic nuclei are more severely affected than the medial or lateral thalamic nuclei, it would be interesting to segment the thalami into sub-regions and investigate differences between the nuclei. This was not possible with the contrasts available on the T1-weighted scans and the FA maps from this study. Finally, in order to extract diffusion metrics that were not contaminated by voxels potentially containing some CSF, thalamic and hippocampal regions were eroded by a 3x3x3 kernel. This may have had differential effects in the thalamus and hippocampus; the thalamus is a relatively large 'bulb' shaped structure, whilst the hippocampus has a relatively thin body region. An erosion kernel of this size may mean that the diffusion metrics obtained from the body region of the hippocampus may only be coming from a limited number of voxels as a result. In order to overcome this limitation, it might be advantageous to instead take a partial volumes approach, such as that described in [Eaton-Rosen et al., 2016].

This study had a number of strengths. It was the largest study to date to investigate thalamic volumes and diffusion metrics in AD and MCI subjects, including EMCI subjects. In addition, a validated protocol for thalamic segmentation was used, which had better segmentation reliability than methods used in previous studies.

In summary, these findings give support to the hypothesis that the thalamus is implicated in Alzheimer's disease relatively early in the disease course and that diffusion metrics provide useful complimentary information to volume measures and may prove to be valuable in early diagnosis in Alzheimer's disease. In EMCI subjects, the inclusion of thalamic and hippocampal diffusion metrics significantly increased the ability to classify EMCI subjects from controls compared with volumetric measures alone.

Chapter 10

Investigation in to thalamic and hippocampal volumes and diffusion metrics in early-onset Alzheimer's disease

10.1 Introduction

As discussed in chapter 2, approximately 5% of those with dementia have an age of onset below 65 years of age, with Alzheimer's disease pathology being the underlying cause in about a third of those with young-onset dementia. Non-amnestic presentations are more common in young-onset Alzheimer's disease (AD), including posterior cortical atrophy (PCA) and logopenic aphasia (LPA) [Mendez et al., 2012]. Relatively little is known about the effects of AD on subcortical regions in younger onset cases. Studying different presentations gives better information about the different disease presentations of young-onset AD. In addition to macrostructural differences, microstructural differences may exist. This may lead to better biomarkers in the future for these different presentations.

There is evidence to suggest that amyloid deposition is greater in the thalami in young-onset Alzheimer's disease than in late-onset Alzheimer's dis-

ease [Cho et al., 2013b] and that young-onset Alzheimer's disease subjects have increased glucose hypometabolism in the thalamus compared with controls [Kim et al., 2005]. Lower thalamic volumes have been found in late-onset Alzheimer's disease [De Jong et al., 2008], and there is increasing evidence that the thalamus may play an important role in early stages of late-onset Alzheimer's disease [Aggleton et al., 2016].

The aim of this study was therefore to investigate thalamic and hippocampal volumes and diffusion metrics in subjects with young-onset Alzheimer's disease.

10.1.1 Methods

10.1.1.1 Subjects

The patients included in this chapter are described in chapter 3. Patients were classified as having a typical [McKhann et al., 2011] (herein referred to as typical Alzheimer's disease (tAD) or atypical (posterior cortical atrophy [Tang-Wai et al., 2004]) AD phenotype (herein referred to as PCA) according to published criteria. Twenty-four age and gender matched control subjects were also recruited.

10.1.1.2 Imaging

For details of the imaging protocol followed see chapter 3. Fractional anisotropy (fractional anisotropy (FA)) maps and mean diffusivity (mean diffusivity (MD)) maps were generated using the pipeline described in chapter 3.3.1.

10.1.1.3 Brain, thalamic, hippocampal and total-intracranial volume segmentation

Brains were segmented using an automated method as described in chapter 3.4.1. Hippocampi were segmented using the automated method described in chapter 7 with the template library which included the full extent of the hippocampal tail. Thalami were segmented manually, with reference to both the T1-weighted and registered colour FA maps, using the protocol described

in chapter 8. Total intracranial volume (TIV) was extracted automatically using the method described in chapter 3.

10.1.1.4 Extraction of thalamic and hippocampal diffusion metrics

The thalamic and hippocampal regions were eroded by a 3x3x3 voxel kernel and mean diffusivity and fractional anisotropy scores were calculated within the eroded regions. This was performed to ensure that boundary voxels which included cerebrospinal fluid were excluded.

10.1.1.5 Statistical Analyses

Linear regression was used to compare baseline ages between the disease groups. Linear regression with bootstrapping (200 repetitions) was used to compare baseline mini-mental state exam (MMSE) scores. A chi-squared test was used to compare gender ratios and t-test was used to compare disease duration between the PCA and tAD groups.

To compare baseline volumes, linear regression analyses were used with the volume of interest as the dependent variable and adjusting for diagnostic group, age, gender and head-size (TIV).

To compare baseline diffusion metrics, linear regression analyses were used with the diffusion metric of interest (thalamic or hippocampal fractional anisotropy (FA) or mean diffusivity (MD)) as the dependent variable and adjusting for diagnostic group, age and gender.

10.1.2 Results

10.1.2.1 Demographics

Table 10.1 shows the baseline characteristics of the subjects included in this study.

	N	Age	MMSE	% female	Disease duration
Controls	20	59.8 (6.1)	29.5 (0.7)	60	n/a
PCA	11	61.4 (6.0)	22.1†(5.1)	73	4.3 (4.9)
tAD	25	61.2 (5.2)	20.7†(4.4)	60	5.0 (2.8)

Table 10.1: Baseline characteristic of local young-onset AD cohort. Mean (standard deviation (SD)) shown. †Significant difference with respect to controls ($p<0.05$).

10.1.2.2 Brain, thalamic and hippocampal volume differences

Table 10.2 shows the mean-adjusted whole-brain volumes were significantly smaller in the PCA and tAD groups as compared with controls. Mean-adjusted right and left thalamic and hippocampal volumes were significantly smaller in the tAD group compared with controls. Hippocampal and thalamic volumes were intermediate to those in controls and in tAD on the right side in PCA subjects, but the difference did not reach statistical significance. There was no difference in hippocampal or thalamic volume on the left side in PCA subjects.

	Controls	PCA	tAD
Brain volume (ml)	1131 [1107, 1156] .	1036 [1002, 1070] p<0.001	1034 [1012, 1056] p<0.001
Left hippocampal volume (ml)	2.99 [2.84, 3.14] .	2.98 [2.78, 3.18] p=0.9	2.62 [2.49, 2.75] p<0.001
Right hippocampal volume (ml)	3.15 [3.01, 3.29] .	2.95 [2.76, 3.14] p=0.1	2.78 [2.66, 2.90] p<0.001
Left thalamic volume (ml)	6.95 [6.75, 7.15] .	7.00 [6.71, 7.28] p=0.8	6.65 [6.46, 6.83] p=0.03
Right thalamic volume (ml)	6.94 [6.76, 7.12] .	6.77 [6.52, 7.03] p=0.3	6.55 [6.38, 6.71] p=0.003

Table 10.2: Mean adjusted baseline volumes [95% confidence interval] by diagnostic group. Volumes are adjusted for mean age, mean head-size and assume an equal gender split.

10.1.2.3 Thalamic and hippocampal diffusion differences

Table 10.3 shows the results from diffusion imaging reporting the microstructural differences in the three groups. The PCA subjects had significantly higher mean adjusted fractional anisotropy in both the left and right thalami compared with control subjects but there was no difference in the left or right hippocampal FA. Subjects with tAD had a significantly increased mean adjusted FA in the left thalamus but not in the right, and a significantly reduced mean adjusted FA in both the left and right hippocampi as compared with controls. Those with tAD also had significantly increased mean adjusted MD scores in both the left and right hippocampi as compared with controls.

		Controls	PCA	tAD
Mean adjusted MD ($10^{-3}mm^2/s^{-1}$)	Left hippocampus	0.85 [0.83, 0.85] .	0.88 [0.85, 0.91] p=0.140	0.92 [0.90, 0.94] p<0.001
	Right hippocampus	0.86 [0.85, 0.88] .	0.89 [0.87, 0.92] p=0.038	0.90 [0.89, 0.92] p=0.002
	Left thalamus	0.72 [0.71, 0.73] .	0.72 [0.71, 0.74] p=0.999	0.73 [0.72, 0.74] p=0.340
	Right thalamus	0.71 [0.70, 0.72] .	0.70 [0.69, 0.72] p=0.581	0.71 [0.70, 0.72] p=0.725
FA	Left hippocampus	0.111 [0.106, 0.115] .	0.105 [0.098, 0.111] p=0.117	0.104 [0.100, 0.108] p=0.040
	Right hippocampus	0.113 [0.107, 0.118] .	0.114 [0.107, 0.122] p=0.714	0.110 [0.105, 0.115] p=0.505
	Left thalamus	0.334 [0.325, 0.342] .	0.359 [0.347, 0.37] p=0.001	0.346 [0.338, 0.353] p=0.045
	Right thalamus	0.341 [0.332, 0.35] .	0.372 [0.359, 0.384] p<0.001	0.349 [0.341, 0.357] p=0.218

Table 10.3: Mean adjusted [95% confidence interval] FA and MD metrics in the left and right thalamus and hippocampus. Metrics were adjusted for mean age and assume an equal gender split.

10.2 Discussion

In this study, thalamic and hippocampal volumes were found to be significantly smaller in young-onset AD subjects who presented with typical amnestic symptoms compared to controls. These subjects also had evidence of microstructural differences, with significantly higher left thalamic FA, significantly lower right and left hippocampal FA and significantly higher right and left hippocampal MD scores as compared with controls.

In the early-onset AD subjects who presented with PCA, mean-adjusted volumes in the right thalamus and right hippocampus were intermediate to those seen in controls and those who presented with typical AD, but the differences did not reach statistical significance. There was no evidence of volume differences between controls and those with PCA in the left hippocampus or thalamus. Although there was no statistically significant difference in hippocampal or thalamic volumes in those with PCA compared with controls, there was evidence of microstructural diffferences in both the left and right thalami and the right hippocampus. Increased FA was observed in both the right and left thalami and decreased FA in the right hippocampus.

Whilst a number of studies have investigated thalamic volumes in typical late-onset Alzheimer's disease (see chapter 9) only one study to date has specifically investigated thalamic volumes in subjects with young-onset Alzheimer's disease. They found no difference at baseline between those with young-onset AD and controls [Cho et al., 2013a]. Although they did not specify the disease presentation of the participants with young-onset AD in their study, there were no statistically significant differences between the neuropsychological tests scores of those with young-onset AD and those with late-onset AD in their study, suggesting the majority of their young-onset AD cohort had an amnestic presentation. The study had a relatively small number of participants with early-onset AD (n=14) and may have been underpowered to find a difference at baseline. However, they did find that thalamic volumes declined with time in those with young-onset AD.

Although no statistically significant difference in hippocampal volume between PCA and control subjects was found in this study, they were somewhat reduced on the right side. In chapter 6, PCA subjects were found to have significantly lower mean adjusted hippocampal volumes than control subjects on both the right and left sides and hippocampal volumes that were intermediate to those seen in controls and in typical AD. It may be that there was not enough power in this study to find a difference as there were only 14 individu-

als with PCA in this chapter, whilst in chapter 6 there were far more (n=47). In addition, the subjects included in this study had a slightly higher mean MMSE score and slightly lower mean disease duration than those included in the study in chapter 6, so it may be that they were less advanced in their disease course.

The findings in this chapter are also largely in accordance with those in the Alzheimer's disease neuroimaging initiative (ADNI) study reported in chapter 9. The only groups that are truly comparable in a qualitative sense are the control subjects and those with tAD. Similar to the findings in the ADNI subjects reported in chapter 9, when the volume of the thalamus is significantly lower in tAD, then there is some evidence of FA being higher. As with the AD subjects in the ADNI cohort, there was no significant difference in thalamic MD values in the tAD subjects. There are some differences in the hippocampal diffusion findings between the two studies. Hippocampal FA was not found to be significantly lower in the AD group in the ADNI study (who had significantly atrophied hippocampi); in the tAD subjects in this study hippocampal FA was lower on the left side. However, the hippocampal MD values were significantly higher than in controls in both the AD subjects in the ADNI study and in the tAD subjects in this study. The PCA subjects cannot feasibly compared with any of the disease groups in the ADNI study. They had higher thalamic FA values and higher right hippocampal MD values, which does not fit with the patterns seen in any of the disease groups studied in the ADNI cohort. FA and MD values were more similar to those seen in the ADNI AD group than any of the disease groups, despite the lack of thalamic volume differences in the PCA group. In other words - PCA is a unique entity and understanding disease progression in this group will inform us about AD in its variant forms. Further this will aid those with PCA through better monitoring and prognostic information. Longitudinal studies would be required to understand the dynamic diffusion and volumetric changes in PCA.

This is the first study to date to specifically investigate thalamic and

hippocampal diffusion metrics in early-onset sporadic AD and the largest study to date to investigate thalamic volumes in those with young-onset AD.

This study has a number of limitations. First, although MMSE scores and disease durations were not significantly different between those with PCA and tAD, neither are particularly good measures of disease severity in subjects with PCA. Mean-adjusted brain volumes were similar in the PCA and tAD groups however, suggesting similar levels of overall baseline atrophy. In this study, only FA and MD differences were investigated. Participants in this study also had neurite orientation dispersion and density imaging (NODDI) imaging performed and NODDI metrics may provide greater insight into the diffusion differences detected [Zhang et al., 2012]. This study was cross-sectional in nature and longitudinal imaging would be required in order to verify that the differences in volume and diffusion metrics are indeed pathological changes. Finally, as mentioned in chapter 9 both the hippocampus and the thalamus are complex structures making differences in diffusion metrics, particularly FA, hard to interpret.

In summary, this study has shown reduced thalamic and hippocampal volumes in subjects with amnestic young-onset AD as compared to controls and microstructural changes in the thalamus and hippocampus in both amnestic and PCA variants of young-onset AD. These findings give support to the hypothesis that thalamus is affected by Alzheimer's disease, even in the absence of significant thalamic atrophy and may provide both insight into disease processes in differing variants and ultimately biomarkers to track progression.

Chapter 11

General Conclusions

This thesis has investigated, in depth, the effects of Alzheimer's disease on the hippocampus and thalamus in various disease presentations and at different disease stages.

Accelerated T1-weighted scans were found to be equivalent to unaccelerated T1-weighted scans for assessing brain, ventricular and hippocampal volumes and atrophy rates and fewer accelerated scans suffered from motion artefacts than unaccelerated. This is important as it allows for more valuable time in the scanner for other sequences which may provide additional information.

The hippocampus was found to be disproportionately affected by atrophy in apolipoprotein (APOE) $\varepsilon 4$ carriers. This has implications for clinical trials as it means that APOE $\varepsilon 4$ status should be considered when using hippocampal atrophy rates as outcome measures; APOE $\varepsilon 4$ carrier status should be adjusted for in statistical analyses. The hippocampus was also found to be reduced in volume in the posterior cortical atrophy variant of Alzheimer's disease, despite a lack of memory deficits in these individuals. Atrophy was found to be localised in the most posterior portion of the hippocampal tail in these individuals. Hippocampal shape metrics were also found to be valuable for disease classification and were superior to hippocampal volumes alone for this disease group.

A new hippocampal template library, including the full extent of the hip-

pocampal tail, for use in automated segmentation pipelines, was compared to an existing, widely used hippocampal template library. The new template library achieved a good mean dice score in leave-one-out cross validation and was better able to separate diseased subjects from controls than the existing template library to which it was compared.

A set of manually-segmented thalamic regions were generated, using multi-modal imaging, for use in automated segmentation pipelines. The thalami and hippocampi of the Alzheimer's disease neuroimaging initiative (ADNI)-GO/-2 cohort were segmented using both the new thalamic and hippocampal template libraries. Reduced thalamic volumes were observed in those with Alzheimer's disease, late mild cognitive impairment and in those with early mild cognitive impairment (on the left side). This was the first study to date to find lower thalamic volumes at the early mild cognitive impairment stage, suggesting that in addition to the hippocampus, the thalamus is implicated in Alzheimer's disease early on in the disease course. Thalamic volumes in a cohort of young-onset Alzheimer's disease subjects were also investigated. This cohort included individuals with both posterior cortical atrophy and with typical amnestic Alzheimer's disease. Thalamic volumes, in addition to hippocampal volumes, were found to be significantly reduced in those with amnestic Alzheimer's disease. There was no evidence of thalamic volume reductions in those with posterior cortical atrophy (PCA).

Hippocampal and thalamic diffusion metrics were also investigated in the subset of ADNI-GO/-2 subjects that had diffusion imaging available and in the young-onset Alzheimer's disease (AD) cohort. No significant differences in thalamic diffusion metrics were found between the disease groups in the ADNI-GO/-2 cohort and controls, although there was borderline evidence for higher mean diffusivity in the thalamus in early and late mild cognitive impairment (MCI) subjects and borderline evidence for increased thalamic fractional anisotropy (FA) values in subjects with AD. Hippocampal mean diffusivity (MD) values were significantly higher in early MCI, late MCI and AD subjects

on the left side and in AD subjects on the right side, compared with controls. Diffusion metrics were found to significantly improve disease classification in combination with volumetric measures, as compared to volumetric measures alone when classifying early MCI vs controls, late MCI vs controls and AD vs controls. There was also evidence that some diffusion metrics were predictive of subsequent atrophy rates in control subjects and subjects with early mild cognitive impairment (EMCI). In the young-onset AD cohort, the PCA group, who had no evidence of thalamic atrophy, had significantly higher thalamic FA than controls. Those with typical AD had increased FA in the left thalamus, decreased FA in the left hippocampus and increased MD in both the left and right hippocampi. Taken together these findings suggest that thalamic MD may possibly increase before atrophy occurs and decreases as the thalamus atrophies. Further MD in the hippocampus appears to increase with increasing levels of atrophy. Longitudinal studies would be required to confirm these hypotheses however.

In order for accurate early diagnoses, beneficial to both patients and carers, there is a real need for multi-modal markers. Metrics such as those described in this thesis may be useful in identifying those who are likely to progress to a diagnosis of AD, when they are in the early MCI stage and little cognitive function is lost. This may have implications in terms of either inclusion criteria and/or outcome measures in clinical trials of very early disease. The thalamic and hippocampal segmentation libraries may be of use for diagnosis and disease tracking.

In summary, the studies described in this thesis have extended current knowledge of hippocampal and thalamic atrophy patterns and diffusion differences in Alzheimer's disease and it's variants, with important implications for clinical trials of putative Alzheimer's disease therapies.

Chapter 12

Publications

Publications arising from this work:

-

- Manning, E. N., Barnes, J., Cash, D. M., Bartlett, J. W., Leung, K. K., Ourselin, S., Fox, N. C., Alzheimer's Disease NeuroImaging Initiative. (2014), APOE 4 is associated with disproportionate progressive hippocampal atrophy in AD. PloS one, 9(5), e97608.

- Manning, E. N., Macdonald, K. E., Leung, K. K., Young, J., Pepple, T., Lehmann, M., Zuluaga, M. A., Cardoso, M. J., Schott, J. M., Ourselin, S., Crutch, S., Fox, N. C. and Barnes, J. (2015), Differential hippocampal shapes in posterior cortical atrophy patients: A comparison with control and typical AD subjects. Hum. Brain Mapp., 36: 5123–5136. doi:10.1002/hbm.22999

- Manning, E N., Leung, K. K., Nicholas, J. M., Malone, I. B., Cardoso, M. J., Schott, J. M., Fox, N. C. and Barnes, J. A comparison of accelerated and non-accelerated MRI scans for brain volume and volume change measures: evidence from the ADNI dataset. Neuroinformatics, 15: 215-226. doi:10.1007/s12021-017-9326-0

Chapter 13

Appendix: Statistical models used in APOE ε4 study

13.1 Cross-sectional analysis

For the analysis examining the influence of the APOE ε4 allele on baseline hippocampal volumes, the following linear regression model was used:

$$Y_i = \beta_0 + \beta_1 e4carrier_i + \beta_2 age_i + \beta_3 female_i + \beta_4 mmse_i + \beta_0 TIV_i + \varepsilon_i$$

and

$$\varepsilon_i \sim N(0, \sigma^2_{\varepsilon 1})$$

Y_i is the hippocampal volume for subject i, e4carrier is a categorical variable representing the APOE $epsilon4$ carrier status (0 if non-carrier, 1 if carrier), age is the mean centred age, female is a categorical variable for gender (0 if male, 1 if female), mmse is the mean centred MMSE score, brainTIVratio is the mean centred brain to total intracranial volume ratio, TIV is the mean centred total intracranial volume, $\varepsilon_i$is the error term, $\beta_0$is the baseline hippocampal volume in non 4-carriers, $\beta_1$is the difference in volume between non-carriers and carriers and $\beta_2, \beta_3, ..., \beta_6$ are the fixed effects coefficients corresponding to the other covariates.

13.2 Longitudinal analysis

For the analysis examining the influence of the APOE $\varepsilon 4$ allele on rates of hippocampal and whole brain atrophy the following joint mixed model was used:

$$Y_{ij}^{(1)} = (\beta_0^{(1)} + \beta_1^{(1)} e4carrier_i + \beta_2^{(1)} age_i + \beta_3^{(1)} female_i + \beta_4^{(1)} mmse_i \\ + \beta_5^{(1)} brainTIVratio_i + \beta_6^{(1)} TIV_i + b_i^{(1)}) t_{ij} + \varepsilon_{ij}^{(1)}$$

$$Y_{ij}^{(2)} = (\beta_0^{(2)} + \beta_1^{(2)} e4carrier_i + \beta_2^{(2)} age_i + \beta_3^{(2)} female_i + \beta_4^{(2)} mmse_i \\ + \beta_5^{(2)} brainTIVratio_i + \beta_6^{(2)} TIV_i + b_i^{(2)}) t_{ij} + \varepsilon_{ij}^{(2)}$$

Here $Y_{ij}^{(1)}$ denotes the jth measurement of brain loss between baseline and time t_{ij} for subject i, and $Y_{ij}^{(2)}$ the corresponding measure of hippocampal loss. $\beta_0^{(1)}$is the mean adjusted brain atrophy rate in non 4 -carrier males with mean age, mean mmse score, mean brain-to-TIV ratio and mean headsize, $\beta_1^{(1)}$ is the difference in mean adjusted brain atrophy rate between 4 carriers and non-carriers, $\beta_2^{(1)}, \beta_3^{(1)}, ..., \beta_6^{(1)}$ are the fixed effect coefficients for the other covariates (age, gender, MMSE score, brain-to-TIV ratio and TIV) for brain loss and b_i^1 is the random effect slope for subject i, likewise $\beta_0^{(2)}$is the mean adjusted hippocampal atrophy rate in non 4-carrier males with mean age, mean mmse score, mean brain-to-TIV ratio and mean headsize, $\beta_{(1)}^{(2)}$ is the difference in mean adjusted hippocampal atrophy rate between 4 carriers and non-carriers, $\beta_2^{(2)}, \beta_3^{(2)}, ..., \beta_6^{(2)}$ are the fixed effect coefficients for hippocampal loss, and $b_i^{(2)}$ is the random slope for subject i.

We assumed the random slopes satisfy

$$\begin{pmatrix} b_i^{(1)} \\ b_i^{(2)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b1}^{(2)} & \sigma_{b1b2} \\ \sigma_{b1b2} & \sigma_{b2}^{(}2) \end{pmatrix} \right)$$

and

$$\begin{pmatrix} \varepsilon_i^{(1)} \\ \varepsilon_i^{(2)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon 1}^{(2)} & \sigma_{\varepsilon 1 \varepsilon 2} \\ \sigma_{\varepsilon 1 \varepsilon 2} & \sigma_{\varepsilon 2}^{(2)} \end{pmatrix} \right)$$

The covariance $\sigma_{\varepsilon 1 \varepsilon 2}$ allows the residual errors within visits to be correlated for the two measures and the covariance σ_{b1b2} allows the random slopes to be correlated. This enables us to adjust the difference in hippocampal atrophy rates between $\varepsilon 4$ carriers and non-carriers for concurrent whole brain atrophy rate.

Acronyms

AD Alzheimer's disease.

ADNI Alzheimer's disease neuroimaging initiative.

APOE apolipoprotein.

AUC area under the curve.

BSI boundary shift integral.

CI confidence interval.

CN cognitively normal.

CSF cerebrospinal fluid.

CT computer tomography.

DLB dementia with Lewy bodies.

DRC dementia research centre.

DTI diffusion tensor imaging.

DWI diffusion weighted imaging.

EMCI early mild cognitive impairment.

FA fractional anisotropy.

FLAIR Fluid-Attenuated Inversion Recovery.

FOV field of view.

FWE family wise error.

GE General Electric.

GM grey matter.

HMAPS hippocampal multi-atlas propagation and segmentation.

KN-BSI robust boundary shift integral.

LMCI late mild cognitive impairment.

LPA logopenic aphasia.

MAPS Multi-Atlas Propagation and Segmentation.

MCI mild cognitive impairment.

MD mean diffusivity.

MMSE mini-mental state exam.

MNI Montreal Neurological Institute.

MPRAGE Magnetization Prepared RApid Gradient Echo.

MRI magnetic resonance imaging.

NODDI neurite orientation dispersion and density imaging.

PCA posterior cortical atrophy.

PET positron emission tomography.

PNFA progressive nonfluent aphasia.

QC quality control.

ROC receiver operator curve.

ROI region of interest.

SD standard deviation.

SMC subjective memory complaints.

SNR signal-to-noise ratio.

SPHARM SPHerical HARMonics.

SPHARM-PDM SPHerical HARMonics - Point Distribution Models.

STEPS Similarity and Truth Estimation for Propagated Segmentations.

SVM support vector machine.

tAD typical Alzheimer's disease.

TIV total intracranial volume.

WMH white matter hyperintensities.

YOAD young onset Alzheimer's disease.

Bibliography

[Addis et al., 2004] Addis, D. R., Moscovitch, M., Crawley, A. P., and McAndrews, M. P. (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. Hippocampus, 14:752–762.

[Aggleton et al., 2016] Aggleton, J. P., Pralus, A., Nelson, A. J. D., and Hornberger, M. (2016). Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to Papez circuit. Brain : a journal of neurology, 139(Pt 7):1877–90.

[Agosta et al., 2009] 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 4 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 United States of America, 106(6):2018–2022.

[Aljabar et al., 2009] Aljabar, P., Heckemann, R. A., Hammers, A., Hajnal, J. V., and Rueckert, D. (2009). Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. NeuroImage, 46(3):726–738.

[Alladi et al., 2007] Alladi, S., Xuereb, J., Bak, T., Nestor, P., Knibb, J., Patterson, K., and Hodges, J. R. (2007). Focal cortical presentations of Alzheimer's disease. Brain : a journal of neurology, 130(Pt 10):2636–45.

[Atkinson et al., 2001] Atkinson, A. J., Colburn, W. A., DeGruttola, V. G., DeMets, D. L., Downing, G. J., Hoth, D. F., Oates, J. A., Peck, C. C., Schooley, R. T., Spilker, B. A., Woodcock, J., and Zeger, S. L. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework.

[Awate et al., 2006] Awate, S. P., Tasdizen, T., Foster, N., and Whitaker, R. T. (2006). Adaptive Markov modeling for mutual-information-based, unsupervised MRI brain-tissue classification. Medical Image Analysis, 10(5):726–739.

[Awate et al., 2008] Awate, S. P., Zhang, H., Simon, T. J., and Gee, J. C. (2008). Multivariate segmentation of brain tissues by fusion of MRI and DTI data. In 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Proceedings, ISBI, pages 213–216.

[Barker et al., 2002] Barker, W. W., Luis, C. a., Kashuba, A., Luis, M., Harwood, D. G., Loewenstein, D., Waters, C., Jimison, P., Shepherd, E., Sevush, S., Graff-Radford, N., Newland, D., Todd, M., Miller, B., Gold, M., Heilman, K., Doty, L., Goodman, I., Robinson, B., Pearl, G., Dickson, D., and Duara, R. (2002). Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the State of Florida Brain Bank. Alzheimer disease and associated disorders, 16(4):203–212.

[Barnes et al., 2015] Barnes, J., Dickerson, B. C., Frost, C., Jiskoot, L. C., Wolk, D., and Van Der Flier, W. M. (2015). Alzheimer's disease first symptoms are age dependent: Evidence from the NACC dataset. Alzheimer's and Dementia, 11(11):1349–1357.

[Barnes et al., 2008] Barnes, J., Foster, J., Boyes, R. G., Pepple, T., Moore, E. K., Schott, J. M., Frost, C., Scahill, R. I., and Fox, N. C. (2008). A com-

parison of methods for the automated calculation of volumes and atrophy rates in the hippocampus. NeuroImage, 40(4):1655–1671.

[Barnes et al., 2004] Barnes, J., Scahill, R. I., Boyes, R. G., Frost, C., Lewis, E. B., Rossor, C. L., Rossor, M. N., and Fox, N. C. (2004). Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates. NeuroImage, 23(2):574–581.

[Baxter and Warrington, 1994] Baxter, D. M. and Warrington, E. K. (1994). Measuring dysgraphia: A graded-difficulty spelling test. Behavioural Neurology, 7:107–116.

[Behrens et al., 2003] Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., Barker, G. J., Sillery, E. L., Sheehan, K., Ciccarelli, O., Thompson, A. J., Brady, J. M., and Matthews, P. M. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature Neuroscience, 6(7):750–757.

[Benson et al., 1988] Benson, D. F., Davis, R. J., and Snyder, B. D. (1988). Posterior Cortical Atrophy. Archives of Neurology, 45(7):789–793.

[Blennow et al., 2010] Blennow, K., Hampel, H., Weiner, M., and Zetterberg, H. (2010). Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. Nature Reviews Neurology, 6(3):131–144.

[Boccardi et al., 2015a] Boccardi, M., Bocchetta, M., Apostolova, L. G., Barnes, J., Bartzokis, G., Corbetta, G., DeCarli, C., DeToledo-Morrell, L., Firbank, M., Ganzola, R., Gerritsen, L., Henneman, W., Killiany, R. J., Malykhin, N., Pasqualetti, P., Pruessner, J. C., Redolfi, A., Robitaille, N., Soininen, H., Tolomeo, D., Wang, L., Watson, C., Wolf, H., Duvernoy, H., Duchesne, S., Jack, C. R., and Frisoni, G. B. (2015a). Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on mag-

netic resonance. Alzheimer's & dementia : the journal of the Alzheimer's Association, 11(2):126–38.

[Boccardi et al., 2015b] Boccardi, M., Bocchetta, M., Apostolova, L. G., Barnes, J., Bartzokis, G., Corbetta, G., Decarli, C., Detoledo-Morrell, L., Firbank, M., Ganzola, R., Gerritsen, L., Henneman, W., Killiany, R. J., Malykhin, N., Pasqualetti, P., Pruessner, J. C., Redolfi, A., Robitaille, N., Soininen, H., Tolomeo, D., Wang, L., Watson, C., Wolf, H., Duvernoy, H., Duchesne, S., Jack, C. R., and Frisoni, G. B. (2015b). Delphi definition of the EADC-ADNI harmonized protocol for hippocampal segmentation on magnetic resonance. Alzheimer's and Dementia, 11(2):126–138.

[Boyes et al., 2006] Boyes, R. G., Rueckert, D., Aljabar, P., Whitwell, J., Schott, J. M., Hill, D. L. G., and Fox, N. C. (2006). Cerebral atrophy measurements using jacobian integration: Comparison with the boundary shift integral. NeuroImage, 32(1):159–169.

[Braak and Braak, 1991] Braak, H. and Braak, E. (1991). Alzheimer's disease affects limbic nuclei of the thalamus. Acta neuropathologica, 81(3):261–8.

[Braak and Braak, 1995] Braak, H. and Braak, E. (1995). Staging of alzheimer's disease-related neurofibrillary changes. Neurobiology of Aging, 16(3):271–278.

[Buckner et al., 2004] Buckner, R. L., Head, D., Parker, J., Fotenos, A. F., Marcus, D., Morris, J. C., and Snyder, A. Z. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. NeuroImage, 23(2):724–38.

[Busse et al., 2008] Busse, R. F., Brau, A. C. S., Vu, A., Michelich, C. R., Bayram, E., Kijowski, R., Reeder, S. B., and Rowley, H. A. (2008). Effects

of refocusing flip angle modulation and view ordering in 3D fast spin echo. Magnetic Resonance in Medicine, 60(3):640–649.

[Canu et al., 2012] Canu, E., Frisoni, G. B., Agosta, F., Pievani, M., Bonetti, M., and Filippi, M. (2012). Early and late onset Alzheimer's disease patients have distinct patterns of white matter damage. Neurobiology of Aging, 33(6):1023–1033.

[Cardoso et al., 2013] Cardoso, M. J., Leung, K., Modat, M., Keihaninejad, S., Cash, D., Barnes, J., Fox, N. C., and Ourselin, S. (2013). STEPS: Similarity and Truth Estimation for Propagated Segmentations and its application to hippocampal segmentation and brain parcelation. Medical image analysis, 17(6):671–84.

[Cardoso et al., 2015] Cardoso, M. J., Modat, M., Wolz, R., Melbourne, A., Cash, D., Rueckert, D., and Ourselin, S. (2015). Geodesic Information Flows: Spatially-Variant Graphs and Their Application to Segmentation and Fusion. IEEE Transactions on Medical Imaging, 34(9):1976–1988.

[Cash et al., 2015] Cash, D. M., Frost, C., Iheme, L. O., Ünay, D., Kandemir, M., Fripp, J., Salvado, O., Bourgeat, P., Reuter, M., Fischl, B., Lorenzi, M., Frisoni, G. B., Pennec, X., Pierson, R. K., Gunter, J. L., Senjem, M. L., Jack, C. R., Guizard, N., Fonov, V. S., Collins, D. L., Modat, M., Cardoso, M. J., Leung, K. K., Wang, H., Das, S. R., Yushkevich, P. A., Malone, I. B., Fox, N. C., Schott, J. M., and Ourselin, S. (2015). Assessing atrophy measurement techniques in dementia: Results from the MIRIAD atrophy challenge. NeuroImage, 123:149–164.

[Cherubini et al., 2010] Cherubini, A., Péran, P., Spoletini, I., Di Paola, M., Di Iulio, F., Hagberg, G. E., Sancesario, G., Gianni, W., Bossù, P., Caltagirone, C., Sabatini, U., and Spalletta, G. (2010). Combined Volumetry and DTI in Subcortical Structures of Mild Cognitive Impairment and

Alzheimer's Disease Patients. Journal of Alzheimer's Disease, 19(4):1273–1282.

[Chételat et al., 2008] Chételat, G., Fouquet, M., Kalpouzos, G., Denghien, I., De la Sayette, V., Viader, F., Mézenge, F., Landeau, B., Baron, J. C., Eustache, F., and Desgranges, B. (2008). Three-dimensional surface mapping of hippocampal atrophy progression from MCI to AD and over normal aging as assessed using voxel-based morphometry. Neuropsychologia, 46(6):1721–1731.

[Chiang et al., 2010] Chiang, G. C., Insel, P. S., Tosun, D., Schuff, N., Truran-Sacrey, D., Raptentsetsang, S. T., Jack, C. R., Aisen, P. S., Petersen, R. C., and Weiner, M. W. (2010). Hippocampal atrophy rates and CSF biomarkers in elderly APOE2 normal subjects. Neurology, 75(22):1976–1981.

[Chiang et al., 2011] Chiang, G. C., Insel, P. S., Tosun, D., Schuff, N., Truran-Sacrey, D., Raptentsetsang, S. T., Thompson, P. M., Reiman, E. M., Jack, J., Fox, N. C., Jagust, W. J., Harvey, D. J., Beckett, L. A., Gamst, A., Aisen, P. S., Petersen, R. C., and Weiner, M. W. (2011). Impact of apolipoprotein 4-cerebrospinal fluid beta-amyloid interaction on hippocampal volume loss over 1 year in mild cognitive impairment. Alzheimer's and Dementia, 7(5):514–520.

[Child and Benarroch, 2013] Child, N. D. and Benarroch, E. E. (2013). Anterior nucleus of the thalamus: Functional organization and clinical implications. Neurology, 81(21):1869–1876.

[Ching et al., 2015] Ching, C. R. K., Hua, X., Hibar, D. P., Ward, C. P., Gunter, J. L., Bernstein, M. A., Jack, C. R., Weiner, M. W., and Thompson, P. M. (2015). Does MRI scan acceleration affect power to track brain change? Neurobiology of aging, 36 Suppl 1:S167–77.

[Cho et al., 2013a] Cho, H., Seo, S. W., Kim, J.-H., Kim, C., Ye, B. S., Kim, G. H., Noh, Y., Kim, H. J., Yoon, C. W., Seong, J.-K., Kim, C.-H., Kang,

S. J., Chin, J., Kim, S. T., Lee, K.-H., and Na, D. L. (2013a). Changes in subcortical structures in early- versus late-onset Alzheimer's disease. Neurobiology of Aging, 34(7):1740–1747.

[Cho et al., 2013b] Cho, H., Seo, S. W., Kim, J.-H., Suh, M. K., Lee, J.-H., Choe, Y. S., Lee, K.-H., Kim, J. S., Kim, G. H., Noh, Y., Ye, B. S., Kim, H. J., Yoon, C. W., Chin, J., and Na, D. L. (2013b). Amyloid deposition in early onset versus late onset Alzheimer's disease. Journal of Alzheimer's disease : JAD, 35(4):813–21.

[Cohen et al., 2001] Cohen, R. M., Small, C., Lalonde, F., Friz, J., and Sunderland, T. (2001). Effect of apolipoprotein E genotype on hippocampal volume loss in aging healthy women. Neurology, 57(12):2223–2228.

[Collins et al., 1998] Collins, D. L., Zijdenbos, a. P., Kollokian, V., Sled, J. G., Kabani, N. J., Holmes, C. J., and Evans, a. C. (1998). Design and construction of a realistic digital brain phantom. IEEE transactions on medical imaging, 17(3):463–468.

[Corder et al., 1993] Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L., and Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science, 261(0036-8075 (Print)):921–923.

[Cortes and Vapnik, 1995] Cortes, C. and Vapnik, V. (1995). Support-vector networks. Machine Learning, 20:273–297.

[Crivello et al., 2010] Crivello, F., Lematre, H., Dufouil, C., Grassiot, B., Delcroix, N., Tzourio-Mazoyer, N., Tzourio, C., and Mazoyer, B. (2010). Effects of ApoE-4 allele load and age on the rates of grey matter and hippocampal volumes loss in a longitudinal cohort of 1186 healthy elderly persons. Neuroimage, 53(3):1064–1069.

[Crutch et al., 2012] Crutch, S. J., Lehmann, M., Schott, J. M., Rabinovici, G. D., Rossor, M. N., and Fox, N. C. (2012). Posterior cortical atrophy. Lancet neurology, 11(2):170–8.

[Csernansky et al., 2005] Csernansky, J. G., Wang, L., Swank, J., Miller, J. P., Gado, M., McKeel, D., Miller, M. I., and Morris, J. C. (2005). Preclinical detection of Alzheimer's disease: Hippocampal shape and volume predict dementia onset in the elderly. NeuroImage, 25(3):783–792.

[Daga et al., 2013] Daga, P., Modat, M., Winston, G., White, M., Mancini, L., McEvoy, A. W., Thornton, J., Yousry, T., Duncan, J. S., and Ourselin, S. (2013). Susceptibility artefact correction by combining B0 field maps and non-rigid registration using graph cuts. page 86690B.

[De Jong et al., 2008] De Jong, L. W., Van Der Hiele, K., Veer, I. M., Houwing, J. J., Westendorp, R. G. J., Bollen, E. L. E. M., De Bruin, P. W., Middelkoop, H. A. M., Van Buchem, M. A., and Van Der Grond, J. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: An MRI study. Brain, 131(12):3277–3285.

[DeCarli et al., 1999] DeCarli, C., Miller, B. L., Swan, G. E., Reed, T., Wolf, P. A., Garner, J., Jack, L., and Carmelli, D. (1999). Predictors of brain morphology for the men of the NHLBI twin study. Stroke; a journal of cerebral circulation, 30(3):529–36.

[Dickerson et al., 2011] Dickerson, B. C., Stoub, T. R., Shah, R. C., Sperling, R. A., Killiany, R. J., Albert, M. S., Hyman, B. T., Blacker, D., and Detoledo-Morrell, L. (2011). Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. Neurology, 76(16):1395–1402.

[Dickson et al., 1992] Dickson, D. W., Crystal, H. A., Mattiace, L. A., Masur, D. M., Blau, A. D., Davies, P., Yen, S. H., and Aronson, M. K. (1992). Identification of normal and pathological aging in prospectively studied nondemented elderly humans. Neurobiology of Aging, 13(1):179–189.

[Dietterich, 1998] Dietterich, T. G. (1998). Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. Neural Computation, 10:1895–1923.

[Douaud et al., 2011] Douaud, G., Jbabdi, S., Behrens, T. E. J., Menke, R. A., Gass, A., Monsch, A. U., Rao, A., Whitcher, B., Kindlmann, G., Matthews, P. M., and Smith, S. (2011). DTI measures in crossing-fibre areas: Increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. NeuroImage, 55(3):880–890.

[Drzezga et al., 2009] Drzezga, A., Grimmer, T., Henriksen, G., Mhlau, M., Perneczky, R., Miederer, I., Praus, C., Sorg, C., Wohlschlger, A, Riemenschneider, M., Wester, H. J., Foerstl, H., Schwaiger, M., and Kurz, A. (2009). Effect of APOE genotype on amyloid plaque load and gray matter volume in Alzheimer disease. Neurology, 72(17):1487–1494.

[Dubois et al., 2007] Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P. J., and Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet neurology, 6(8):734–46.

[Dubois et al., 2014a] Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., DeKosky, S. T., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G. B., Fox, N. C., Galasko, D., Habert, M.-O., Jicha, G. A., Nordberg, A., Pasquier, F., Rabinovici, G., Robert, P., Rowe, C., Salloway, S., Sarazin, M., Epelbaum, S., de Souza, L. C., Vellas, B., Visser, P. J., Schneider, L., Stern, Y., Scheltens, P., and Cummings, J. L. (2014a). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet neurology, 13(6):614–29.

[Dubois et al., 2014b] Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., Dekosky, S. T., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G. B., Fox, N. C., Galasko, D., Habert, M. O., Jicha, G. A., Nordberg, A., Pasquier, F., Rabinovici, G., Robert, P., Rowe, C., Salloway, S., Sarazin, M., Epelbaum, S., de Souza, L. C., Vellas, B., Visser, P. J., Schneider, L., Stern, Y., Scheltens, P., and Cummings, J. L. (2014b). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria.

[Dubois et al., 2016] Dubois, B., Hampel, H., Feldman, H. H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., Blennow, K., Broich, K., Cavedo, E., Crutch, S., Dartigues, J. F., Duyckaerts, C., Epelbaum, S., Frisoni, G. B., Gauthier, S., Genthon, R., Gouw, A. A., Habert, M. O., Holtzman, D. M., Kivipelto, M., Lista, S., Molinuevo, J. L., O'Bryant, S. E., Rabinovici, G. D., Rowe, C., Salloway, S., Schneider, L. S., Sperling, R., Teichmann, M., Carrillo, M. C., Cummings, J., and Jack, C. R. (2016). Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria.

[Duning et al., 2005] Duning, T., Kloska, S., Steinsträter, O., Kugel, H., Heindel, W., and Knecht, S. (2005). Dehydration confounds the assessment of brain atrophy. Neurology, 64(3):548–550.

[Duvernoy, 2005] Duvernoy, H. M. (2005). The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI. Springer Science & Business Media.

[Eaton-Rosen et al., 2016] Eaton-Rosen, Z., Melbourne, A., Jorge Cardoso, M., Marlow, N., and Ourselin, S. (2016). Beyond the resolution limit: Diffusion parameter estimation in partial volume. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 9902 LNCS, pages 605–612. Springer, Cham.

[Echávarri et al., 2011] Echávarri, C., Caballero, M. C., Aramendía, A., García-Bragado, F., and Tuñón, T. (2011). Multiprotein deposits in neurodegenerative disorders: our experience in the tissue brain bank of Navarra. Anatomical record (Hoboken, N.J. : 2007), 294(7):1191–7.

[Eisenberg et al., 2010] Eisenberg, D. T., Kuzawa, C. W., and Hayes, M. G. (2010). Worldwide allele frequencies of the human apolipoprotein E gene: climate, local adaptations, and evolutionary history. Am.J.Phys.Anthropol., 143(1096-8644 (Electronic)):100–111.

[Eustache et al., 2016] Eustache, P., Nemmi, F., Saint-Aubert, L., Pariente, J., and Péran, P. (2016). Multimodal Magnetic Resonance Imaging in Alzheimer's Disease Patients at Prodromal Stage. Journal of Alzheimer's Disease, 50(4):1035–1050.

[Fischl et al., 2002] Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron, 33(3):341–355.

[Folstein et al., 1975] 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.

[Forman and Scholz, 2010] Forman, G. and Scholz, M. (2010). Apples-to-apples in cross-validation studies.

[Fox et al., 2005] Fox, N. C., Black, R. S., Gilman, S., Rossor, M. N., Griffith, S. G., Jenkins, L., and Koller, M. (2005). Effects of Abeta immunization (AN1792) on MRI measures of cerebral volume in Alzheimer disease. Neurology, 64(9):1563–72.

[Fox et al., 2011] Fox, N. C., Ridgway, G. R., and Schott, J. M. (2011). Algorithms, atrophy and Alzheimer's disease: Cautionary tales for clinical trials.

[Frankó and Joly, 2013] Frankó, E. and Joly, O. (2013). Evaluating Alzheimer's Disease Progression Using Rate of Regional Hippocampal Atrophy. PLoS ONE, 8(8).

[Freeborough et al., 1997] Freeborough, P. A., Fox, N. C., and Kitney, R. I. (1997). Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. Computer Methods and Programs in Biomedicine, 53(1):15–25.

[Frost et al., 2004] Frost, C., Kenward, M. G., and Fox, N. C. (2004). The analysis of repeated 'direct' measures of change illustrated with an application in longitudinal imaging. Statistics in medicine, 23(21):3275–86.

[Gao et al., 2014] Gao, Y., Riklin-Raviv, T., and Bouix, S. (2014). Shape analysis, a field in need of careful validation. Human brain mapping, 00.

[Gerardin et al., 2009] Gerardin, E., Chételat, G., and Chupin, M. (2009). ... classification of hippocampal shape features discriminates Alzheimer's disease and mild .... Neuroimage.

[Geroldi et al., 1999] Geroldi, C., Pihlajamaki, M., Laakso, M. P., DeCarli, C., Beltramello, A., Bianchetti, A., Soininen, H., Trabucchi, M., and Frisoni, G. B. (1999). APOE-epsilon4 is associated with less frontal and more medial temporal lobe atrophy in AD. Neurology, 53(0028-3878 (Print)):1825–1832.

[Glaister et al., 2017] Glaister, J., Carass, A., NessAiver, T., Stough, J. V., Saidha, S., Calabresi, P. A., and Prince, J. L. (2017). Thalamus segmentation using multi-modal feature classification: Validation and pilot study of an age-matched cohort. NeuroImage, 158:430–440.

[Hammers et al., 2003] Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., Mitchell, T. N., Brooks, D. J., and Duncan, J. S.

(2003). Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Human Brain Mapping, 19(4):224–247.

[Hammers et al., 2007] Hammers, A., Chen, C. H., Lemieux, L., Allom, R., Vossos, S., Free, S. L., Myers, R., Brooks, D. J., Duncan, J. S., and Koepp, M. J. (2007). Statistical neuroanatomy of the human inferior frontal gyrus and probabilistic atlas in a standard stereotaxic space. Human Brain Mapping, 28(1):34–48.

[Harvey and Skelton-Robinson, 2003] Harvey, R. J. and Skelton-Robinson, M. (2003). The prevalence and causes of dementia in people under the age of 65 years. J Neurol Neurosurg Psychiatry, 74(9):1206–1209.

[Hashimoto et al., 2009] Hashimoto, R., Hirata, Y., Asada, T., Yamashita, F., Nemoto, K., Mori, T., Moriguchi, Y., Kunugi, H., Arima, K., and Ohnishi, T. (2009). Effect of the brain-derived neurotrophic factor and the apolipoprotein E polymorphisms on disease progression in preclinical Alzheimer's disease. Genes, Brain and Behavior, 8(1):43–52.

[Henneman et al., 2009] Henneman, W. J. P., Sluimer, J. D., Barnes, J., van der Flier, W. M., Sluimer, I. C., Fox, N. C., Scheltens, P., Vrenken, H., and Barkhof, F. (2009). Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures. Neurology, 72(11):999–1007.

[Hill et al., 2014] Hill, D. L. G., Schwarz, A. J., Isaac, M., Pani, L., Vamvakas, S., Hemmings, R., Carrillo, M. C., Yu, P., Sun, J., Beckett, L., Boccardi, M., Brewer, J., Brumfield, M., Cantillon, M., Cole, P. E., Fox, N., Frisoni, G. B., Jack, C., Kelleher, T., Luo, F., Novak, G., Maguire, P., Meibach, R., Patterson, P., Bain, L., Sampaio, C., Raunig, D., Soares, H., Suhy, J., Wang, H., Wolz, R., and Stephenson, D. (2014). Coalition Against Major Diseases/European Medicines Agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's

disease. Alzheimer's & dementia : the journal of the Alzheimer's Association, 10(4):421–9.e3.

[Hirshhorn et al., 2012] Hirshhorn, M., Grady, C., Rosenbaum, R. S., Winocur, G., and Moscovitch, M. (2012). Brain regions involved in the retrieval of spatial and episodic details associated with a familiar environment: An fMRI study. Neuropsychologia, 50:3094–3106.

[Hyman et al., 2012] Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., Dickson, D. W., Duyckaerts, C., Frosch, M. P., Masliah, E., Mirra, S. S., Nelson, P. T., Schneider, J. A., Thal, D. R., Thies, B., Trojanowski, J. Q., Vinters, H. V., and Montine, T. J. (2012). National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association, 8(1):1–13.

[Iosifescu et al., 1997] Iosifescu, D. V., Shenton, M. E., Warfield, S. K., Kikinis, R., Dengler, J., Jolesz, F. A., and McCarley, R. W. (1997). An automated registration algorithm for measuring MRI subcortical brain structures. NeuroImage, 6(1):13–25.

[Jack et al., 2013] Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., Shaw, L. M., Vemuri, P., Wiste, H. J., Weigand, S. D., Lesnick, T. G., Pankratz, V. S., Donohue, M. C., and Trojanowski, J. Q. (2013). Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers.

[Jack et al., 2008] Jack, C. R., Petersen, R. C., Grundman, M., Jin, S., Gamst, A., Ward, C. P., Sencakova, D., Doody, R. S., and Thal, L. J. (2008). Longitudinal MRI findings from the vitamin E and donepezil treatment study for MCI. Neurobiology of aging, 29(9):1285–95.

[Jack et al., 1998] Jack, C. R., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1998). Rate of medial

temporal lobe atrophy in typical aging and Alzheimer's disease. Neurology, 51(4):993–999.

[Jackson and Warrington, 1986] Jackson, M. and Warrington, E. K. (1986). Arithmetic skills in patients with unilateral cerebral lesions. Cortex; a journal devoted to the study of the nervous system and behavior, 22:611–620.

[Jak et al., 2007] Jak, A. J., Houston, W. S., Nagel, B. J., Corey-Bloom, J., and Bondi, M. W. (2007). Differential cross-sectional and longitudinal impact of APOE genotype on hippocampal volumes in nondemented older adults. Dementia and Geriatric Cognitive Disorders, 23(6):382–389.

[Jbabdi et al., 2009] Jbabdi, S., Woolrich, M. W., and Behrens, T. E. J. (2009). Multiple-subjects connectivity-based parcellation using hierarchical Dirichlet process mixture models. NeuroImage, 44(2):373–384.

[Jellinger and Attems, 2015] Jellinger, K. A. and Attems, J. (2015). Challenges of multimorbidity of the aging brain: a critical update.

[Johansen-Berg et al., 2005] Johansen-Berg, H., Behrens, T. E. J., Sillery, E., Ciccarelli, O., Thompson, A. J., Smith, S. M., and Matthews, P. M. (2005). Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. Cerebral Cortex, 15(1):31–39.

[Jovicich et al., 2006] Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., Macfall, J., Fischl, B., and Dale, A. (2006). Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage, 30(2):436–43.

[Kalus et al., 2006] Kalus, P., Slotboom, J., Gallinat, J., Mahlberg, R., Cattapan-Ludewig, K., Wiest, R., Nyffeler, T., Buri, C., Federspiel, A., Kunz, D., Schroth, G., and Kiefer, C. (2006). Examining the gateway to

the limbic system with diffusion tensor imaging: The perforant pathway in dementia. NeuroImage, 30(3):713–720.

[Kantarci et al., 2010] Kantarci, K., Avula, R., Senjem, M. L., Samikoglu, A. R., Zhang, B., Weigand, S. D., Przybelski, S. A., Edmonson, H. A., Vemuri, P., Knopman, D. S., Ferman, T. J., Boeve, B. F., Petersen, R. C., and Jack, C. R. (2010). Dementia with Lewy bodies and Alzheimer disease: Neurodegenerative patterns characterized by DTI. Neurology, 74(22):1814–1821.

[Karas et al., 2004] Karas, G. B., Scheltens, P., Rombouts, S. A. R. B., Visser, P. J., Van Schijndel, R. A., Fox, N. C., and Barkhof, F. (2004). Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. NeuroImage, 23(2):708–716.

[Khan et al., 2013] Khan, W., Giampietro, V., Ginestet, C., Dell'acqua, F., Bouls, D., Newhouse, S., Dobson, R., Banaschewski, T., Barker, G. J., Bokde, A. L., Buchel, C., Conrod, P., Flor, H., Frouin, V., Garavan, H., Gowland, P., Heinz, A., Ittermann, B., Lemaitre, H., Nees, F., Paus, T., Pausova, Z., Rietschel, M., Smolka, M. N., Strohle, A., Gallinat, J., Westman, E., Schumann, G., Lovestone, S., and Simmons, A. (2013). No Differences in Hippocampal Volume between Carriers and Non-Carriers of the ApoE epsilon4 and epsilon2 Alleles in Young Healthy Adolescents. J.Alzheimers.Dis., (1875-8908 (Electronic)).

[Kim et al., 2005] Kim, E. J., Cho, S. S., Jeong, Y., Park, K. C., Kang, S. J., Kang, E., Kim, S. E., Lee, K. H., and Na, D. L. (2005). Glucose metabolism in early onset versus late onset Alzheimer's disease: An SPM analysis of 120 patients. Brain, 128(8):1790–1801.

[Klunk and Mathis, 2008] Klunk, W. E. and Mathis, C. A. (2008). The future of amyloid-beta imaging: a tale of radionuclides and tracer proliferation. Current opinion in neurology, 21(6):683–687.

[Knopman et al., 2012] Knopman, D. S., Jack, C. R., Wiste, H. J., Weigand, S. D., Vemuri, P., Lowe, V., Kantarci, K., Gunter, J. L., Senjem, M. L., Ivnik, R. J., Roberts, R. O., Boeve, B. F., and Petersen, R. C. (2012). Short-term clinical outcomes for stages of NIA-AA preclinical Alzheimer disease. Neurology, 78(20):1576–1582.

[Koedam et al., 2011] Koedam, E. L. G. E., Lehmann, M., Van Der Flier, W. M., Scheltens, P., Pijnenburg, Y. A. L., Fox, N., Barkhof, F., and Wattjes, M. P. (2011). Visual assessment of posterior atrophy development of a MRI rating scale. European Radiology, 21(12):2618–2625.

[Kohavi, 1995] Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. pages 1137–1143.

[Krueger et al., 2012] Krueger, G., Granziera, C., Jack, C. R., Gunter, J. L., Littmann, A., Mortamet, B., Kannengiesser, S., Sorensen, A. G., Ward, C. P., Reyes, D. A., Britson, P. J., Fischer, H., and Bernstein, M. A. (2012). Effects of MRI scan acceleration on brain volume measurement consistency. Journal of magnetic resonance imaging : JMRI, 36(5):1234–40.

[La Joie et al., 2013] La Joie, R., Perrotin, A., de La Sayette, V., Egret, S., Doeuvre, L., Belliard, S., Eustache, F., Desgranges, B., and Chételat, G. (2013). Hippocampal subfield volumetry in mild cognitive impairment, Alzheimer's disease and semantic dementia. NeuroImage. Clinical, 3:155–62.

[Langa et al., 2017] Langa, K. M., Larson, E. B., Crimmins, E. M., Faul, J. D., Levine, D. A., Kabeto, M. U., and Weir, D. R. (2017). A Comparison of the Prevalence of Dementia in the United States in 2000 and 2012. JAMA Internal Medicine, 177(1):51.

[Lehmann et al., 2011] Lehmann, M., Crutch, S. J., Ridgway, G. R., Ridha, B. H., Barnes, J., Warrington, E. K., Rossor, M. N., and Fox, N. C. (2011). Cortical thickness and voxel-based morphometry in posterior cortical atrophy and typical Alzheimer's disease. Neurobiology of aging, 32(8):1466–76.

[Lehtovirta, 1995] Lehtovirta, M. (1995). Volumes of hippocampus, amygdala and frontal lobe in Alzheimer patients with different apolipoprotein E genotypes. Neuroscience, 67(1):65–72.

[Leung et al., 2011] Leung, K. K., Barnes, J., Modat, M., Ridgway, G. R., Bartlett, J. W., Fox, N. C., and Ourselin, S. (2011). Brain MAPS: an automated, accurate and robust brain extraction technique using a template library. Neuroimage., 55(3):1091–1108.

[Leung et al., 2010a] Leung, K. K., Barnes, J., Ridgway, G. R., Bartlett, J. W., Clarkson, M. J., Macdonald, K., Schuff, N., Fox, N. C., and Ourselin, S. (2010a). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. Neuroimage, 51(4):1345–59.

[Leung et al., 2010b] Leung, K. K., Clarkson, M. J., Bartlett, J. W., Clegg, S., Jack, C. R., Weiner, M. W., Fox, N. C., and Ourselin, S. (2010b). Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: Tissue-specific intensity normalization and parameter selection. NeuroImage, 50(2):516–523.

[Li et al., 2007] Li, S., Shi, F., Pu, F., Li, X., Jiang, T., Xie, S., and Wang, Y. (2007). Hippocampal shape analysis of Alzheimer disease based on machine learning methods. AJNR. American journal of neuroradiology, 28(7):1339–45.

[Lindberg et al., 2012] Lindberg, O., Walterfang, M., Looi, J. C. L., Malykhin, N., Ostberg, P., Zandbelt, B., Styner, M., Paniagua, B., Velakoulis, D., Orndahl, E., and Wahlund, L.-O. (2012). Hippocampal shape analysis in Alzheimer's disease and frontotemporal lobar degeneration subtypes. Journal of Alzheimer's disease : JAD, 30(2):355–65.

[Lo et al., 2011] Lo, R. Y., Hubbard, A. E., Shaw, L. M., Trojanowski, J. Q., Petersen, R. C., Aisen, P. S., Weiner, M. W., and Jagust, W. J. (2011). Lon-

gitudinal change of biomarkers in cognitive decline. Archives of Neurology, 68(10):1257–1266.

[Lu et al., 2011] Lu, P. H., Thompson, P. M., Leow, A., Lee, G. J., Lee, A., Yanovsky, I., Parikshak, N., Khoo, T., Wu, S., Geschwind, D., and Bartzokis, G. (2011). Apolipoprotein E genotype is associated with temporal and hippocampal atrophy rates in healthy elderly adults: a tensor-based morphometry study. J Alzheimers Dis, 23(1875-8908 (Electronic)):433–442.

[M. Knapp et al., 2014] M. Knapp, M. G., M. McCrone, P. Prina, M. Comas-Herrera, A. Wittenberg, R., and Adelaja, B. Hu, B. King, D. Rehill, D. S. (2014). Dementia UK: Update. Technical Report 1.

[Mang et al., 2012] Mang, S. C., Busza, A., Reiterer, S., Grodd, W., Klose, and Uwe (2012). Thalamus segmentation based on the local diffusion direction: A group study. Magnetic Resonance in Medicine, 67(1):118–126.

[Marrakchi-Kacem et al., 2010] Marrakchi-Kacem, L., Poupon, C., Mangin, J. F., and Poupon, F. (2010). Multi-contrast deep nuclei segmentation using a probabilistic atlas. In 2010 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2010 - Proceedings, pages 61–64.

[Mayes et al., 2007] Mayes, A. R., Montaldi, D., and Migo, E. (2007). Associative memory and the medial temporal lobes. Trends in Cognitive Sciences, 11:126–135.

[McKhann et al., 1984] McKhann, G. M., Drachman, D., Folstein, M. F., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology, 34:939–44.

[McKhann et al., 2011] 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.

[McNemar, 1947] McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika, 12(2):153–157.

[Melbourne et al., 2016] Melbourne, A., Toussaint, N., Owen, D., Simpson, I., Anthopoulos, T., De Vita, E., Atkinson, D., and Ourselin, S. (2016). NiftyFit: a Software Package for Multi-parametric Model-Fitting of 4D Magnetic Resonance Imaging Data. Neuroinformatics, 14(3):319–337.

[Mendez et al., 2002] Mendez, M. F., Ghajarania, M., and Perryman, K. M. (2002). Posterior cortical atrophy: clinical characteristics and differences compared to Alzheimer's disease. Dementia and geriatric cognitive disorders, 14(1):33–40.

[Mendez et al., 2012] Mendez, M. F., Lee, A. S., Joshi, A., and Shapira, J. S. (2012). Nonamnestic Presentations of Early-Onset Alzheimer's Disease. American Journal of Alzheimer's Disease & Other Dementiasr, 27(6):413–420.

[Modat et al., 2010] Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., Fox, N. C., and Ourselin, S. (2010). Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine, 98(3):278–284.

[Moffat et al., 2000] Moffat, S. D., Szekely, C. A., Zonderman, A. B., Kabani, N. J., and Resnick, S. M. (2000). Longitudinal change in hippocampal volume as a function of apolipoprotein E genotype. Neurology, 55(0028-3878 (Print)):134–136.

[Möller et al., 2015] Möller, C., Dieleman, N., Van Der Flier, W. M., Versteeg, A., Pijnenburg, Y., Scheltens, P., Barkhof, F., and Vrenken, H. (2015). More atrophy of deep gray matter structures in frontotemporal dementia compared to Alzheimer's disease. Journal of Alzheimer's Disease, 44(2):635–647.

[Mori et al., 2002] Mori, E., Lee, K., Yasuda, M., Hashimoto, M., Kazui, H., Hirono, N., and Matsui, M. (2002). Accelerated hippocampal atrophy in Alzheimer's disease with apolipoprotein E 4 allele. Annals of Neurology, 51(2):209–214.

[Morra et al., 2009] Morra, J. H., Tu, Z., Apostolova, L. G., Green, A. E., Avedissian, C., Madsen, S. K., Parikshak, N., Hua, X., Toga, A. W., Jack Jr., C. R., Schuff, N., Weiner, M. W., and Thompson, P. M. (2009). Automated 3D mapping of hippocampal atrophy and its clinical correlates in 400 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. Hum.Brain Mapp., 30(1097-0193 (Electronic)):2766–2788.

[Mueller et al., 2010] Mueller, S. G., Schuff, N., Yaffe, K., Madison, C., Miller, B., and Weiner, M. W. (2010). Hippocampal atrophy patterns in mild cognitive impairment and alzheimer's disease. Human Brain Mapping, 31:1339–1347.

[Müller et al., 2007] Müller, M. J., Greverus, D., Weibrich, C., Dellani, P. R., Scheurich, A., Stoeter, P., and Fellgiebel, A. (2007). Diagnostic utility of hippocampal size and mean diffusivity in amnestic MCI. Neurobiology of Aging, 28(3):398–403.

[Narayana et al., 1988] Narayana, P. A., Brey, W. W., Kulkarni, M. V., and Sievenpiper, C. L. (1988). Compensation for surface coil sensitivity variation in magnetic resonance imaging. Magn Reson.Imaging, 6(0730-725X (Print)):271–274.

[O'Dwyer et al., 2012] O'Dwyer, L., Lamberton, F., Matura, S., Tanner, C., Scheibe, M., Miller, J., Rujescu, D., Prvulovic, D., and Hampel, H. (2012). Reduced hippocampal volume in healthy young ApoE4 carriers: an MRI study. PLoS.One., 7(1932-6203 (Electronic)):e48895.

[Patenaude et al., 2011] 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):907–922.

[Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. The Journal of Machine Learning Research, 12:2825–2830.

[Pievani et al., 2009] Pievani, M., Rasser, P. E., Galluzzi, S., Benussi, L., Ghidoni, R., Sabattoli, F., Bonetti, M., Binetti, G., Thompson, P. M., and Frisoni, G. B. (2009). Mapping the effect of APOE epsilon4 on gray matter loss in Alzheimer's disease in vivo. Neuroimage., 45(1095-9572 (Electronic)):1090–1098.

[Pluta et al., 2012] Pluta, J., Yushkevich, P., Das, S., and Wolk, D. (2012). In vivo analysis of hippocampal subfield atrophy in mild cognitive impairment via semi-automatic segmentation of T2-weighted MRI. Journal of Alzheimer's Disease, 31(1):85–99.

[Poulin et al., 2011] Poulin, S. P., Dautoff, R., Morris, J. C., Barrett, L. F., Dickerson, B. C., and Alzheimer's Disease Neuroimaging Initiative (2011). Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. Psychiatry research, 194(1):7–13.

[Power et al., 2015] Power, B. D., Wilkes, F. A., Hunter-Dickson, M., van Westen, D., Santillo, A. F., Walterfang, M., Nilsson, C., Velakoulis, D., and Looi, J. C. L. (2015). Validation of a protocol for manual segmentation

of the thalamus on magnetic resonance imaging scans. Psychiatry Research - Neuroimaging, 232(1):98–105.

[Prados et al., 2017] Prados, F., Ashburner, J., Blaiotta, C., Brosch, T., Carballido-Gamio, J., Cardoso, M. J., Conrad, B. N., Datta, E., Dávid, G., Leener, B. D., Dupont, S. M., Freund, P., Wheeler-Kingshott, C. A., Grussu, F., Henry, R., Landman, B. A., Ljungberg, E., Lyttle, B., Ourselin, S., Papinutto, N., Saporito, S., Schlaeger, R., Smith, S. A., Summers, P., Tam, R., Yiannakas, M. C., Zhu, A., and Cohen-Adad, J. (2017). Spinal cord grey matter segmentation challenge. NeuroImage, 152:312–329.

[Preboske et al., 2006] Preboske, G. M., Gunter, J. L., Ward, C. P., and Jack, C. R. (2006). Common MRI acquisition non-idealities significantly impact the output of the boundary shift integral method of measuring brain atrophy on serial MRI. NeuroImage, 30(4):1196–202.

[Prince et al., 2015] Prince, M., Wimo, A., Guerchet, M., Gemma-Claire, A., Wu, Y.-T., and Prina, M. (2015). World Alzheimer Report 2015: The Global Impact of Dementia - An analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International, page 84.

[Renner et al., 2004] Renner, J. A., Burns, J. M., Hou, C. E., McKeel, D. W., Storandt, M., and Morris, J. C. (2004). Progressive posterior cortical dysfunction: A clinicopathologic series. Neurology, 63(7):1175–1180.

[Risacher et al., 2010] Risacher, S. L., Shen, L., West, J. D., Kim, S., McDonald, B. C., Beckett, L. A., Harvey, D. J., Jack, C. R., Weiner, M. W., and Saykin, A. J. (2010). Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort. Neurobiology of Aging, 31(8):1401–1418.

[Rose et al., 2006] Rose, S. E., McMahon, K. L., Janke, a. L., O'Dowd, B., de Zubicaray, G., Strudwick, M. W., and Chalk, J. B. (2006). Diffusion indices on magnetic resonance imaging and neuropsychological performance

in amnestic mild cognitive impairment. Journal of neurology, neurosurgery, and psychiatry, 77(10):1122–1128.

[Ryan et al., 2013] Ryan, N. S., Keihaninejad, S., Shakespeare, T. J., Lehmann, M., Crutch, S. J., Malone, I. B., Thornton, J. S., Mancini, L., Hyare, H., Yousry, T., Ridgway, G. R., Zhang, H., Modat, M., Alexander, D. C., Rossor, M. N., Ourselin, S., and Fox, N. C. (2013). Magnetic resonance imaging evidence for presymptomatic change in thalamus and caudate in familial Alzheimer's disease. Brain, 136(5).

[Salloway et al., 2014] Salloway, S., Sperling, R., Fox, N. C., Blennow, K., Klunk, W., Raskind, M., Sabbagh, M., Honig, L. S., Porsteinsson, A. P., Ferris, S., Reichert, M., Ketter, N., Nejadnik, B., Guenzler, V., Miloslavsky, M., Wang, D., Lu, Y., Lull, J., Tudor, I. C., Liu, E., Grundman, M., Yuen, E., Black, R., and Brashear, H. R. (2014). Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. The New England journal of medicine, 370(4):322–33.

[Salloway et al., 2009] Salloway, S., Sperling, R., Gilman, S., Fox, N. C., Blennow, K., Raskind, M., Sabbagh, M., Honig, L. S., Doody, R., van Dyck, C. H., Mulnard, R., Barakos, J., Gregg, K. M., Liu, E., Lieberburg, I., Schenk, D., Black, R., Grundman, M., and Bapineuzumab 201 Clinical Trial Investigators (2009). A phase 2 multiple ascending dose trial of bapineuzumab in mild to moderate Alzheimer disease. Neurology, 73(24):2061–70.

[Scheltens and van de Pol, 2012] Scheltens, P. and van de Pol, L. (2012). Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates: Table 1. Journal of Neurology, Neurosurgery & Psychiatry, 83(11):1038–1040.

[Schneider et al., 2007] Schneider, J. A., Arvanitakis, Z., Bang, W., and Bennett, D. A. (2007). Mixed brain pathologies account for most dementia cases in community-dwelling older persons. Neurology, 69(24):2197–2204.

[Schuff et al., 2009] Schuff, N., Matsumoto, S., Kmiecik, J., Studholme, C., Du, A., Ezekiel, F., Miller, B. L., Kramer, J. H., Jagust, W. J., Chui, H. C., and Weiner, M. W. (2009). Cerebral blood flow in ischemic vascular dementia and Alzheimer's disease, measured by arterial spin-labeling magnetic resonance imaging. Alzheimer's & Dementia, 5(6):454–462.

[Scoville and Milner, 1957] Scoville, W. B. and Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. Journal of Neurology, Neurosurgery & Psychiatry, 20(1):11–21.

[Shattuck et al., 2008] Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., Poldrack, R. A., Bilder, R. M., and Toga, A. W. (2008). Construction of a 3D probabilistic atlas of human cortical structures. NeuroImage, 39(3):1064–1080.

[Shaw et al., 2009] Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., Dean, R., Siemers, E., Potter, W., Lee, V. M.-Y., and Trojanowski, J. Q. (2009). Cerebrospinal fluid biomarker signature in alzheimer's disease neuroimaging initiative subjects. Annals of Neurology, 65(4):403–413.

[Shen et al., 2012] Shen, K.-k., Fripp, J., Mériaudeau, F., Chételat, G., Salvado, O., and Bourgeat, P. (2012). Detecting global and local hippocampal shape changes in Alzheimer's disease using statistical shape models. NeuroImage, 59(3):2155–66.

[Sled et al., 1998] Sled, J. G., Zijdenbos, A. P., and Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE transactions on medical imaging, 17:87–97.

[Smith et al., 2007] Smith, S. M., Rao, A., De Stefano, N., Jenkinson, M., Schott, J. M., Matthews, P. M., and Fox, N. C. (2007). Longitudinal and cross-sectional analysis of atrophy in Alzheimer's disease: Cross-validation of BSI, SIENA and SIENAX. NeuroImage, 36(4):1200–1206.

[Smith et al., 2002] Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P., Federico, A., and De Stefano, N. (2002). Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. NeuroImage, 17(1):479–489.

[Solano-Castiella et al., 2010] Solano-Castiella, E., Anwander, A., Lohmann, G., Weiss, M., Docherty, C., Geyer, S., Reimer, E., Friederici, A. D., and Turner, R. (2010). Diffusion tensor imaging segments the human amygdala in vivo. NeuroImage, 49(4):2958–2965.

[Solano-Castiella et al., 2011] Solano-Castiella, E., Schäfer, A., Reimer, E., Türke, E., Pröger, T., Lohmann, G., Trampel, R., and Turner, R. (2011). Parcellation of human amygdala in vivo using ultra high field structural MRI. NeuroImage, 58(3):741–748.

[Spampinato et al., 2011] Spampinato, M. V., Rumboldt, Z., Hosker, R. J., and Mintzer, J. E. (2011). Apolipoprotein E and gray matter volume loss in patients with mild cognitive impairment and Alzheimer disease. Radiology, 258(3):843–852.

[Squire et al., 2007] Squire, L. R., Wixted, J. T., and Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. Nature reviews. Neuroscience, 8(11):872–83.

[Stěpán-Buksakowska et al., 2013] Stěpán-Buksakowska, I., Szabó, N., Hořínek, D., Tóth, E., Hort, J., Warner, J., Charvát, F., Vécsei, L., Roček, M., and Kincses, Z. T. (2013). Cortical and Subcortical Atrophy in Alzheimer Disease: Parallel Atrophy of Thalamus and Hippocampus. Alzheimer disease and associated disorders, 00(00):1–8.

[Strange et al., 2014] Strange, B. A., Witter, M. P., Lein, E. S., and Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. Nature Reviews Neuroscience, 15(10):655–669.

[Styner et al., 2006] Styner, M., Oguz, I., Xu, S., Brechbühler, C., Pantazis, D., Levitt, J. J., Shenton, M. E., and Gerig, G. (2006). Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM. The insight journal, (1071):242–250.

[Tang et al., 2015] Tang, X., Holland, D., Dale, A. M., Younes, L., and Miller, M. I. (2015). The diffeomorphometry of regional shape change rates and its relevance to cognitive deterioration in mild cognitive impairment and Alzheimer's disease. Human Brain Mapping, 36(6):2093–2117.

[Tang-Wai et al., 2004] Tang-Wai, D. F., Graff-Radford, N. R., Boeve, B. F., Dickson, D. W., Parisi, J. E., Crook, R., Caselli, R. J., Knopman, D. S., and Petersen, R. C. (2004). Clinical, genetic, and neuropathologic characteristics of posterior cortical atrophy. Neurology, 63(7):1168–1174.

[Thompson et al., 2004] Thompson, P. M., Hayashi, K. M., De Zubicaray, G. I., Janke, A. L., Rose, S. E., Semple, J., Hong, M. S., Herman, D. H., Gravano, D., Doddrell, D. M., and Toga, A. W. (2004). Mapping hippocampal and ventricular change in Alzheimer disease. NeuroImage, 22(4):1754–66.

[Tosun et al., 2016] Tosun, D., Schuff, N., Jagust, W., and Weiner, M. W. (2016). Discriminative power of arterial spin labeling magnetic resonance imaging and 18F-fluorodeoxyglucose positron emission tomography changes for amyloid-β-positive subjects in the Alzheimer's disease continuum. Neurodegenerative Diseases, 16(1-2):87–94.

[Unrath et al., 2008] Unrath, A., Klose, U., Grodd, W., Ludolph, A. C., and Kassubek, J. (2008). Directional colour encoding of the human thalamus by diffusion tensor imaging. Neuroscience Letters, 434(3):322–327.

[Van De Pol et al., 2007] Van De Pol, L. A., van der Flier, W. M., Korf, E. S. C., Fox, N. C., Barkhof, F., and Scheltens, P. (2007). Baseline predictors of rates of hippocampal atrophy in mild cognitive impairment. Neurology, 69(15):1491–1497.

[van der Flier and Scheltens, 2009] van der Flier, W. M. and Scheltens, P. (2009). Alzheimer disease: Hippocampal volume loss and Alzheimer disease progression. Nat Rev Neurol, 5(7):361–362.

[van Es and van den Berg, 2009] van Es, M. A. and van den Berg, L. H. (2009). Alzheimer's disease beyond APOE. Nat.Genet., 41(1546-1718 (Electronic)):1047–1048.

[Van Hecke et al., 2016] Van Hecke, W., Emsell, L., and Sunaert, S. (2016). Diffusion Tensor Imaging: A Practical Handbook.

[Vellas et al., 2009] Vellas, B., Black, R., Thal, L., Fox, N., Daniels, M., McLennan, G., Tompkins, C., Leibman, C., Pomfret, M., and Grundman, M. (2009). Long-Term Follow-Up of Patients Immunized with AN1792: Reduced Functional Decline in Antibody Responders. Current Alzheimer Research, 6(2):144–151.

[Vemuri et al., 2015] Vemuri, P., Senjem, M. L., Gunter, J. L., Lundt, E. S., Tosakulwong, N., Weigand, S. D., Borowski, B. J., Bernstein, M. A., Zuk, S. M., Lowe, V. J., Knopman, D. S., Petersen, R. C., Fox, N. C., Thompson, P. M., Weiner, M. W., and Jack, C. R. (2015). Accelerated vs. unaccelerated serial MRI based TBM-SyN measurements for clinical trials in Alzheimer's disease. NeuroImage, 113:61–9.

[Villemagne et al., 2015] Villemagne, V. L., Fodero-Tavoletti, M. T., Masters, C. L., and Rowe, C. C. (2015). Tau imaging: Early progress and future directions.

[Vogt et al., 1990] Vogt, L. J. K., Hyman, B. T., Van Hoesen, G. W., and Damasio, A. R. (1990). Pathological alterations in the amygdala in Alzheimer's disease. Neuroscience, 37(2):377–385.

[Wang et al., 2006] Wang, L., Miller, J. P., Gado, M. H., McKeel, D. W., Rothermich, M., Miller, M. I., Morris, J. C., and Csernansky, J. G. (2006). Abnormalities of hippocampal surface structure in very mild dementia of the Alzheimer type. NeuroImage, 30(1):52–60.

[Wang et al., 2009] Wang, P. N., Liu, H. C., Lirng, J. F., Lin, K. N., and Wu, Z. A. (2009). Accelerated hippocampal atrophy rates in stable and progressive amnestic mild cognitive impairment. Psychiatry Research - Neuroimaging, 171(3):221–231.

[Wang et al., 2012] Wang, Y., West, J. D., Flashman, L. A., Wishart, H. A., Santulli, R. B., Rabin, L. A., Pare, N., Arfanakis, K., and Saykin, A. J. (2012). Selective changes in white matter integrity in MCI and older adults with cognitive complaints. Biochimica et Biophysica Acta - Molecular Basis of Disease, 1822(3):423–430.

[Warfield et al., 2004] Warfield, S. K., Zou, K. H., and Wells, W. M. (2004). Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging, 23(7):903–921.

[Warrington, 1996] Warrington, E. K. (1996). The Camden Memory Tests. Psychology Press, Hove.

[Warrington and James, 1991a] Warrington, E. K. and James, M. (1991a). A new test of object decision: 2D silhouettes featuring a minimal view. Cortex; a journal devoted to the study of the nervous system and behavior, 27:370–383.

[Warrington and James, 1991b] Warrington, E. K. and James, M. (1991b). The Visual Object and Space Perception Battery. Thames Valley Test Company, Bury St Edmunds (UK).

[Whitaker, 2000] Whitaker, R. T. (2000). Reducing Aliasing Artifacts in Iso-Surfaces of Binary Volumes. 2000 IEEE Symposium on Volume Visualization (VV 2000).

[Whitwell et al., 2007] Whitwell, J. L., Jack, C. R., Kantarci, K., Weigand, S. D., Boeve, B. F., Knopman, D. S., Drubach, D. A., Tang-Wai, D. F., Petersen, R. C., and Josephs, K. A. (2007). Imaging correlates of posterior cortical atrophy. Neurobiology of aging, 28(7):1051–61.

[Wiegell et al., 2003a] Wiegell, M. R., Tuch, D. S., Larsson, H. B., and Wedeen, V. J. (2003a). Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. NeuroImage, 19(2):391–401.

[Wiegell et al., 2003b] Wiegell, M. R., Tuch, D. S., Larsson, H. B. W., and Wedeen, V. J. (2003b). Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. NeuroImage, 19:391–401.

[Wilson et al., 2012] Wilson, R. S., Segawa, E., Boyle, P. A., Anagnos, S. E., Hizel, L. P., and Bennett, D. A. (2012). The natural history of cognitive decline in Alzheimer's disease. Psychology and Aging, 27(4):1008–1017.

[Winston, 2012] Winston, G. P. (2012). The physical and biological basis of quantitative parameters derived from diffusion MRI. Quantitative imaging in medicine and surgery, 2(4):254–65.

[Winterburn et al., 2013] Winterburn, J. L., Pruessner, J. C., Chavez, S., Schira, M. M., Lobaugh, N. J., Voineskos, A. N., and Chakravarty, M. M. (2013). A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging. NeuroImage, 74:254–65.

[Wolz et al., 2010] Wolz, R., Heckemann, R. A., Aljabar, P., Hajnal, J. V., Hammers, A., Ltjnen, J., and Rueckert, D. (2010). Measurement of hippocampal atrophy using 4D graph-cut segmentation: Application to ADNI. Neuroimage, 52(1):109–118.

[Worsley et al., 2009] Worsley, K., Taylor, J., Carbonell, F., Chung, M., Duerden, E., Bernhardt, B., Lyttelton, O., Boucher, M., and Evans, A. (2009). SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. NeuroImage, 47:S102.

[Wright et al., 2003] Wright, R. O., Hu, H., Silverman, E. K., Tsaih, S. W., Schwartz, J., Bellinger, D., Palazuelos, E., Weiss, S. T., and Hernandez-Avila, M. (2003). Apolipoprotein E genotype predicts 24-month Bayley scales infant development score. Pediatric Research, 54(6):819–825.

[Wu et al., 2016] Wu, Y.-T., Fratiglioni, L., Matthews, F. E., Lobo, A., Breteler, M. M. B., Skoog, I., and Brayne, C. (2016). Dementia in western Europe: epidemiological evidence and implications for policy making. The Lancet Neurology, 15(1):116–124.

[Yi et al., 2015] Yi, H.-A., Möller, C., Dieleman, N., Bouwman, F. H., Barkhof, F., Scheltens, P., van der Flier, W. M., and Vrenken, H. (2015). Relation between subcortical grey matter atrophy and conversion from mild cognitive impairment to Alzheimer's disease. Journal of neurology, neurosurgery, and psychiatry, pages jnnp–2014–309105.

[Zhang et al., 2012] 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):1000–1016.

[Ziyan et al., 2006] Ziyan, U., Tuch, D., and Westin, C.-F. (2006). Segmentation of thalamic nuclei from DTI using spectral clustering. Medical image

computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, 9(Pt 2):807–814.

[Ziyan and Westin, 2008] Ziyan, U. and Westin, C. F. (2008). Joint segmentation of thalamic nuclei from a population of diffusion tensor MR images. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 5241 LNCS, pages 279–286.