Multimodal Magnetic Resonance Imaging of Frontotemporal Lobar Degeneration

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences

2015

Helen Beaumont

School of Medicine
Contents

List of Figures 10

List of Tables 16

Abstract 19

Acknowledgements 22

About the author 23

Glossary 24

1 Introduction 28

1.1 Overview of this thesis ........................................... 30

1.2 Frontotemporal lobar degeneration .............................. 31

1.2.1 History ...................................................... 33

1.2.2 FTLD terminology ........................................... 33

1.2.3 FTLD diagnosis .............................................. 34

1.2.4 FTLD phenocopy ............................................ 34

1.2.5 Genetics ..................................................... 35

1.2.6 FTLD pathology ............................................. 36

1.2.7 Treatment .................................................. 36
1.2.8 What’s in a name? ........................................... 37
1.2.9 Summary ................................................... 37

2 Background .................................................... 39

2.1 MRI basics .................................................. 39
2.1.1 Magnetic resonance .................................... 39
2.2 The MR signal .............................................. 41
2.2.1 Free induction decay (FID) ........................... 41
2.2.2 Spin-echo, gradient-echo and sequence diagrams 42
2.2.3 Positional information ................................. 43
2.3 MR diffusion ................................................. 46
2.4 MR perfusion ............................................... 51
2.5 Image analysis ............................................. 53
2.5.1 Normalisation ............................................. 54
2.5.2 Voxel-based analysis ................................... 55
2.5.3 Regional analysis ....................................... 56
2.6 Imaging in FTLD ........................................... 56
2.6.1 PET and SPECT ........................................ 57
2.6.1.1 Other ligands ........................................ 60
2.6.2 T1-weighted MRI ........................................ 61
2.6.3 Diffusion MRI ........................................... 63
2.6.4 Perfusion MRI ........................................... 64
2.7 Power calculations ........................................ 65
2.8 Summary ..................................................... 66

3 Mathematical modelling of the ASL signal ............ 70
3.1 Quantifying the pulsed ASL signal .............................................. 70
  3.1.1 Modelling the ASL signal .................................................. 70
    3.1.1.1 The Buxton model .................................................. 71
    3.1.1.2 The single-blood-compartment model .......................... 73
    3.1.1.3 A suggested intermediate model ................................. 74
  3.1.2 Influence of model parameters on ASL subtraction signal ........ 75

3.2 Simulation methods .................................................................. 75
  3.2.1 Comparing results of simulated models - precision and accuracy ............................................................. 79

3.3 Results of fits on simulated data .............................................. 81
  3.3.1 Accuracy of estimates .......................................................... 81
    3.3.1.1 1-parameter fit ......................................................... 81
    3.3.1.2 2-parameter fit .......................................................... 81
    3.3.1.3 3-parameter fit for CBF, arrival time and bolus width .... 84
    3.3.1.4 3-parameter fit for CBF, arrival time and relaxation time ... 87
  3.3.2 Precision of estimates - comparison of goodness of fit .......... 90

3.4 Multi-timepoint dataset ........................................................... 94
  3.4.1 Methods ................................................................. 94
  3.4.2 Results ................................................................. 95
  3.4.3 Measuring SNR .......................................................... 99
    3.4.3.1 Methods .......................................................... 99
    3.4.3.2 Results .......................................................... 99

3.5 Discussion ............................................................................ 100

4 Comparison of gradient-echo and spin-echo MRI sequences in ventral brain regions 102
### Contents

#### 4.1 Introduction ........................................... 102

#### 4.2 Materials and methods ................................. 104

- 4.2.1 MRI acquisition parameters .......................... 104
- 4.2.2 Region definition ...................................... 105
- 4.2.3 Pre-processing .......................................... 105
- 4.2.4 Analysis .................................................. 107

#### 4.3 Results .................................................. 110

- 4.3.1 Pre-processing .......................................... 110
- 4.3.2 Comparison of values ................................... 112

#### 4.4 Discussion .............................................. 123

- 4.4.1 Scanner problems ....................................... 124
- 4.4.2 Conclusion ............................................... 126

#### 5 Regional atrophy in frontotemporal lobar degeneration and relationship to cognitive deficits 127

#### 5.1 Introduction ........................................... 127

#### 5.2 Methods ................................................ 128

- 5.2.1 MRI sequences ......................................... 129
- 5.2.2 Neuropsychological tests ............................... 130
- 5.2.3 Analysis .................................................. 132
  - 5.2.3.1 Voxel-based analysis ................................. 132
  - 5.2.3.2 Regional analysis .................................... 133

#### 5.3 Results .................................................. 136

- 5.3.1 Results of cognitive tests .............................. 136
- 5.3.2 Normalisation .......................................... 137
- 5.3.3 Voxel-based comparison of patient and control groups .... 137
5.3.4 Results of voxel-based regression of patient group with individual cognitive tests ............................................. 140
  5.3.4.1 Naming ............................................. 140
  5.3.4.2 Pyramids and palmtrees ................................. 141
  5.3.4.3 Famous faces recognition ............................... 142
  5.3.4.4 Evaluation of social actions ............................. 142
  5.3.5 Regional analysis ......................................... 144

5.4 Discussion ................................................ 150

6 Cerebral blood flow and arterial arrival time in frontotemporal lobar degeneration ................................................................. 152

6.1 Introduction .................................................. 152

6.2 Methods ....................................................... 154
  6.2.1 ASL sequence and model .................................. 154
  6.2.2 Partial volume correction .................................. 156
  6.2.3 Voxel-based analysis ...................................... 158
  6.2.4 Regional analysis ......................................... 159

6.3 Results ....................................................... 159
  6.3.1 Group differences in CBF and AAT with no partial volume correction ......................................................... 159
  6.3.2 Group differences in CBF calculated from a 1-parameter fit with fixed AAT of 750ms, no partial volume correction ... 165
  6.3.3 Group differences in CBF, AAT corrected for partial volume effects ................................................................. 169
  6.3.4 Regional Analysis .......................................... 174

6.4 Discussion ................................................... 177
  6.4.1 Partial volume correction .................................. 177
  6.4.2 CBF differences ............................................ 177
## 7 Diffusion MRI in frontotemporal lobar degeneration

### 7.1 Introduction

### 7.2 Methods

#### 7.2.1 MRI sequences

#### 7.2.2 Pre-processing

#### 7.2.3 Voxel-based analysis

#### 7.2.4 Regional analysis

### 7.3 Results

#### 7.3.1 Voxel-based group comparison of FA using TBSS

#### 7.3.2 Voxel-based analyses of MD using SPM8

##### 7.3.2.1 Comparison of patient and control groups

##### 7.3.2.2 Regression of naming scores

##### 7.3.2.3 Regression of pyramids and palm trees scores

##### 7.3.2.4 Regression of famous faces recognition scores

##### 7.3.2.5 Evaluation of social actions

#### 7.3.3 Regional analysis

### 7.4 Discussion

## 8 Multimodal modelling in frontotemporal lobar degeneration

### 8.1 Introduction

### 8.2 Methods

### 8.3 Results

#### 8.3.1 Using JHU ROIs thresholded at 20%

##### 8.3.1.1 Classification with all variables
## Contents

8.3.1.2 Using Matlab “pca” (principal component analysis) 207  
8.3.1.3 Classification using only ROI variables seen to be significant in the univariate analysis 207  
8.3.1.4 Classification with varying modalities 208  
8.3.1.5 Classification with varying ROIs 208  
8.3.2 JHU ROIs thresholded at 10% 211  
8.3.3 Classification with all variables 211  
8.3.3.1 Using Matlab “pca” (principal component analysis) 211  
8.3.3.2 Classification using only ROI variables seen to be significant in the univariate analysis 211  
8.3.3.3 Classification with varying modalities 212  
8.3.3.4 Classification with varying ROIs 212  
8.4 Discussion 216  

9 Conclusion and further work 218  
9.1 Summary of findings 218  
9.2 Discussion of findings 222  
9.2.1 FTLD patients in this study 222  
9.2.2 Partial volume corrections 222  
9.2.3 Voxel-based analysis vs regional analysis 223  
9.3 Further work 225  
9.3.1 Further analysis; more data 225  
9.3.2 MRI sequences to test for atrophy 225  
9.3.3 Development of a diagnostic support tool 226  
9.4 Summary 226  

Appendices 228
A  

A Diagnostic criteria for bvFTD, SD and PPA .......................... 228
A.1 bvFTD ................................................................. 228
A.2 Semantic dementia ................................................. 229
A.3 Semantic variant PPA .............................................. 230

B  

B Goodness of fit for 2-parameter fit vs 3-parameter fit for 4 and 8
timepoints .............................................................. 232

C  

C Measuring labelling efficiency ........................................ 234
C.1 Introduction ........................................................ 234
C.2 Methods ............................................................. 234
C.3 Results ............................................................... 234
C.4 Discussion .......................................................... 235

D  

D Day 1 cognitive tests .................................................. 237
D.1 Description of the tests used ...................................... 237
D.2 Results ............................................................... 239

Bibliography  

Word Count 65,000
List of Figures

1.1 Axial images of T₁-weighted images of a patient and a control 32
2.1 The effect of an RF pulse 41
2.2 Free induction decay signal 42
2.3 The evolution of the spin-echo 43
2.4 Simple spin-echo and gradient-echo sequences 44
2.5 Spatial localisation of MR signal 45
2.6 More typical spin-echo and gradient-echo sequences 47
2.7 A gradient-echo EPI sequence 47
2.8 Visualisation of the diffusion tensor 49
2.9 A diffusion-weighted pulse sequence 50
2.10 Building an ASL subtraction image 52
2.11 Planning a pulsed ASL image 54
2.12 Power calculations for sample sizes 67
3.1 Passage of the labelled bolus through the voxel 72
3.2 Influence of model parameters on subtraction signal 76
3.3 Effect on estimated CBF of changing the assumed value of arrival time 77
3.4 Initial model, simulated data and various fits 79
3.5 Estimated CBF from 1-parameter fit with error bars .................. 82
3.6 Percentage error in CBF for 1-parameter fit .......................... 83
3.7 Percentage error in CBF for 2-parameter fit .......................... 85
3.8 Percentage error in arrival time for 2-parameter fit .................... 86
3.9 Percentage error in arrival time as a function of simulated arrival time for 2-parameter fit .................................................... 86
3.10 Percentage error in individual CBF values from 3-parameter fit including bolus width ......................................................... 87
3.11 Percentage error in CBF Values with error bars vs SNR from 3-parameter fit including bolus width ............................................. 88
3.12 Percentage error in CBF values vs SNR from 3-parameter fit including bolus width ................................................................. 89
3.13 Percentage error in individual CBF Values from 3-parameter fit including relaxation time .................................................... 90
3.14 Percentage error in CBF Values with error bars vs SNR from 3-parameter fit including bolus width ............................................. 91
3.15 Percentage error in CBF values vs SNR from 3-parameter fit including relaxation time .................................................... 92
3.16 Comparison of goodness of fit for 1-parameter vs 2-parameter fits . 93
3.17 Comparison of goodness of fit for 16 timepoints for 2-parameter vs 3-parameter fits ............................................................. 94
3.18 Mean whole-brain subtraction signal for 23-timepoint data, and estimated data from multiple fits ............................................... 96
3.19 CBF maps from multi-parameter fits ...................................... 97
3.20 Arrival time maps from multi-parameter fits ............................. 97
3.21 Bolus width and relaxation time maps from multi-parameter fits .. 97
3.22 Comparison of goodness of fit in multi-parameter fits ................. 98
3.23 SNR vs no of averages for ASL data ..................................... 100
4.1 Regions of interest for GE/SE comparison ............................. 106
4.2 Individual subtraction images for slice 14 for subject R12, showing MRI artefacts ........................................ 111
4.3 Images showing where extreme $M_0$ values are to be found ........... 112
4.4 Sagittal images of CBF in GE and SE for both global and voxelwise $m_0^*$ correction ........................................ 114
4.5 Histogram of mean regional CBF values in ml/100ml/min ............. 119
4.6 Histogram of mean regional arrival time values in ms ............... 119
4.7 Comparison of GE/SE regional CBF values in ml/100ml/min for individual subjects ........................................ 120
4.8 Comparison of GE/SE regional arrival time values for individual subjects ........................................ 121
4.9 Comparison of regional signal-to-noise factors between GE and SE images ........................................ 122
4.10 Correlation between GE and SE values for CBF in regions prone to susceptibility distortion and regions not so prone ............... 125

5.1 Sample ROIs overlain on a canonical $T_1$-weighted image .......... 134
5.2 Regions of interest from the aal atlas used in analyses ............... 135
5.3 Regions of interest from the JHU White Matter Tractography Atlas used in analyses ........................................ 135
5.4 Initial and normalised images for an FTLD patient .................. 137
5.5 Results of group comparison between patients and controls: no correction for multiple comparisons, $p<0.001$, minimum cluster size 75 voxels ........................................ 138
5.6 Results of group comparison between patients and controls with cluster-based multiple-comparison correction .................. 139
5.7 Results of regression of patient images against Naming scores: no correction for multiple comparisons, $p<0.001$, minimum cluster size 75 voxels ........................................ 140
5.8 Results of regression of patient images against pyramids and palmtrees scores: no correction for multiple comparisons, $p<0.001$, minimum cluster size 75 voxels ........................................ 141
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.9</td>
<td>Results of regression of patient images against the guilt and anger scores from the evaluation of social action tests: no correction for multiple comparisons, p&lt;0.001, minimum cluster size 75 voxels</td>
<td>143</td>
</tr>
<tr>
<td>6.1</td>
<td>Simulated data for a mixture of grey and white matter</td>
<td>157</td>
</tr>
<tr>
<td>6.2</td>
<td>Axial and coronal images of t-test results where CBF uncorrected for partial volume effects differs between patients and controls (p&lt;0.005, no correction for multiple comparisons)</td>
<td>161</td>
</tr>
<tr>
<td>6.3</td>
<td>Axial and coronal images of t-test results where CBF uncorrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p&lt;0.05).</td>
<td>161</td>
</tr>
<tr>
<td>6.4</td>
<td>Axial and coronal images of t-test results where AAT uncorrected for partial volume effects differs between patients and controls (p&lt;0.005, no correction for multiple comparisons).</td>
<td>162</td>
</tr>
<tr>
<td>6.5</td>
<td>Axial and coronal images of t-test results where AAT uncorrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p&lt;0.05).</td>
<td>162</td>
</tr>
<tr>
<td>6.6</td>
<td>Axial and coronal images of difference in AAT values between patients and controls uncorrected for partial volume effects</td>
<td>163</td>
</tr>
<tr>
<td>6.7</td>
<td>Axial and coronal images of t-test results where CBF calculated with fixed AAT, uncorrected for partial volume effects differs between patients and controls</td>
<td>166</td>
</tr>
<tr>
<td>6.8</td>
<td>Axial and coronal images of t-test results where CBF calculated with fixed AAT, uncorrected for partial volume effects, differs between patients and controls (AFNI 3dClustSim p&lt;0.05)</td>
<td>166</td>
</tr>
<tr>
<td>6.9</td>
<td>Axial images of CBF calculated with fixed AAT, CBF and AAT calculated with 2-parameter fit and T1-weighted image for an FTLD patient</td>
<td>167</td>
</tr>
<tr>
<td>6.10</td>
<td>Axial and coronal images of t-test results where CBF corrected for partial volume effects differs between patients and controls</td>
<td>170</td>
</tr>
<tr>
<td>6.11</td>
<td>Axial and coronal images of t-test results where CBF corrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p&lt;0.05).</td>
<td>171</td>
</tr>
<tr>
<td>6.12</td>
<td>Axial and coronal images of group differences of AAT values corrected for partial volume effects</td>
<td>171</td>
</tr>
</tbody>
</table>
6.13 Axial and coronal images of group differences in AAT values (AFNI 3dClustSim p<0.05) .......................................................... 172
6.14 Axial and coronal images of difference in AAT values between patients and controls corrected for partial volume effects ............ 172
6.15 ROI values from regions where voxel-based analysis shows differing AAT between patients and controls .......................... 175

7.1 Coronal section from T1-weighted images of the left anterior temporal lobe for control and FTLD patient ................................. 180
7.2 Axial and coronal images of the TBSS analysis of a group comparison of FA between patients and controls ........................... 186
7.3 Axial and coronal images of t-test results (p=0.005, no correction for multiple comparisons) where whole-brain mean diffusivity differs between patients and controls ............................................. 188
7.4 Axial and coronal images of t-test results (p < 0.005, no correction for multiple comparisons) where mean diffusivity differs between patients and controls. Images have been improperly masked to exclude voxels containing > 20% CSF .......................................................... 188
7.5 Axial and coronal images of t-test results (p < 0.005, no correction for multiple comparisons) where mean diffusivity differs between patients and controls. Images have been correctly masked to exclude voxels containing > 20% CSF .......................................................... 189
7.6 Axial and coronal images of t-test results (p < 0.005, no correction for multiple comparisons) for regression of patients’ naming scores .......................................................... 190
7.7 Axial and coronal images of t-test results (p < 0.005, no correction for multiple comparisons) for regression of patients’ pyramids and palm trees scores .......................................................... 191
7.8 Axial and coronal images of t-test results (p < 0.005, no correction for multiple comparisons) for regression of patients’ famous faces recognition scores .......................................................... 192

8.1 ROIs for a given modality which give the most accurate classification 213

B.1 Comparison of goodness of fit for 4 timepoints for 2-parameter vs 3-parameter fits .......................................................... 232
B.2 Comparison of goodness of fit for 8 timepoints for 2-parameter vs 3-parameter fits ........................................... 233

C.1 Phantom for labelling efficiency measurement .................. 235

C.2 MR signal for label and control images .......................... 236
List of Tables

3.1 Parameter values used in simulation 78
3.2 Delay times used in simulation 78
3.3 Estimated values from various fits of an ASL dataset with 23 timepoints 95
3.4 Results of comparison of goodness of fit for differing models 96

4.1 Mean regional CBF values for global $m_0^0$ correction 115
4.2 Mean regional CBF values for voxelwise correction 115
4.3 Mean regional CBF values for GE images comparing global and voxelwise correction 116
4.4 Mean regional CBF values for SE images comparing global and voxelwise correction 116
4.5 Mean regional arrival time values for global quantification 117
4.6 Mean regional arrival time values for voxelwise quantification 117
4.7 Mean regional SNR values for global quantification 118
4.8 Mean regional SNR values for voxelwise quantification 118

5.1 Results of novel cognitive tests 136
5.2 Clusters significant after correction for multiple comparisons, $p<0.05$ 138
5.3 Grey matter density values in aal ROIs for patients and controls 146
5.4 White matter density values in aal ROIs for patients and controls 146
5.5 Grey matter density values in JHU ROIs thresholded at probability > 20% for patients and controls ........................................... 147
5.6 Grey matter density values in JHU ROIs thresholded at probability > 10% for patients and controls ................................. 147
5.7 White matter density values in JHU ROIs thresholded at probability > 20% for patients and controls ................................. 148
5.8 White matter density values in JHU ROIs thresholded at probability > 10% for patients and controls ................................. 149

6.1 Cluster statistics showing regions where CBF uncorrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p<0.05) ........................................... 163
6.2 Cluster statistics showing regions where AAT uncorrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p<0.05) ........................................... 164
6.3 Cluster statistics showing regions where CBF calculated with fixed AAT, uncorrected for partial volume effects, differs between patients and controls (AFNI 3dClustSim p<0.05) ........................................... 168
6.4 Cluster statistics showing regions where CBF corrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p<0.05) ........................................... 170
6.5 Cluster statistics showing regions where AAT corrected for partial volume effects differs between patients and controls (AFNI 3dClustSim p<0.05) ........................................... 173
6.6 Regional CBF for patients and controls ................................. 176
6.7 Regional AAT for patients and controls ................................. 176

7.1 Cluster statistics showing regions where mean diffusivity is greater in patients than controls with no masking of CSF, FDR-corrected ........................................... 189
7.2 Cluster statistics showing regions where mean diffusivity is greater in patients than controls. Voxels with >20% CSF have been set to zero ................................. 189
7.3 MD values in aal regions for patients and controls .................. 193
7.4 FA values in aal regions for patients and controls .................. 194
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>MD values in JHU regions for patients and controls in ROIs thresholded at 20% probability</td>
<td>194</td>
</tr>
<tr>
<td>7.6</td>
<td>FA values in JHU regions for patients and controls in ROIs thresholded at 20% probability</td>
<td>195</td>
</tr>
<tr>
<td>7.7</td>
<td>MD values in JHU regions for patients and controls in ROIs thresholded at 10% probability</td>
<td>195</td>
</tr>
<tr>
<td>7.8</td>
<td>FA values in JHU regions for patients and controls in ROIs thresholded at 10% probability</td>
<td>196</td>
</tr>
<tr>
<td>7.9</td>
<td>Published studies of diffusion data in FTLD</td>
<td>199</td>
</tr>
<tr>
<td>8.1</td>
<td>ROIs significant at $p&lt;0.05$, $p&lt;0.01$, $p&lt;0.001$ for all modalities</td>
<td>206</td>
</tr>
<tr>
<td>8.2</td>
<td>Classification excluding/including each modality in turn with JHU ROIs thresholded at 20%</td>
<td>209</td>
</tr>
<tr>
<td>8.3</td>
<td>Classification excluding/including each ROI in turn with JHU ROIs thresholded at 20%</td>
<td>210</td>
</tr>
<tr>
<td>8.4</td>
<td>Classification excluding/including each modality in turn with JHU ROIs thresholded at 10% probability</td>
<td>214</td>
</tr>
<tr>
<td>8.5</td>
<td>Classification excluding/including each ROI in turn with JHU ROIs thresholded at 10% probability</td>
<td>215</td>
</tr>
<tr>
<td>8.6</td>
<td>Modalities and ROIs included for most accurate classification</td>
<td>215</td>
</tr>
<tr>
<td>D.1</td>
<td>Results of standard cognitive tests</td>
<td>240</td>
</tr>
</tbody>
</table>
Abstract

Frontotemporal lobar degeneration (FTLD) is a heterogeneous group of illnesses which can be difficult to diagnose. Modern diagnostic criteria require the presence of imaging abnormalities, but these are not always seen in the early stages of the illness. Hence there is a need to consider the use of more advanced MR techniques. This thesis reports the results of a multimodal MRI study of patients with FTLD, and considers two things: how well data from the different modalities can classify patients, and how well the different modalities can identify affected tissue.

FTLD is thought to involve alterations in cerebral blood flow, but it is possible that microvascular changes will alter additional perfusion parameters, such as the time taken for blood to reach the tissue (the arrival time). Multi-time point arterial spin labelling (ASL) measurements have the ability to extract the relevant parameters. I consider the parameters involved in modelling these data, and report the accuracy of cerebral blood flow (CBF) measurement achievable in a clinically acceptable time. FTLD patients have atrophy in the frontal and temporal lobes, regions problematic for MRI because of susceptibility artefacts caused by adjacent air spaces. I consider two ASL MR read-out sequences (gradient-echo and spin-echo) and show that spin-echo images give higher signal in frontal and temporal regions than gradient-echo.

ASL, T1-weighted and diffusion-weighted images were collected for a group of 17 FTLD patients and 18 controls. I found decreased CBF in highly atrophied regions of cortical grey matter in patients, but this deficit was not seen when corrected for atrophy. An increased arrival time was seen in regions adjacent to the atrophied regions, but a decreased arrival time was seen in the atrophied regions; this is a novel finding. The diffusion metrics of fractional anisotropy (FA) and particularly mean diffusivity (MD) are found to be highly sensitive to differences in FTLD patients. I speculate that this is an increased sensitivity to atrophy because of the increased signal from cerebrospinal fluid.

I combine the regional values of all the modalities in a classification method to distinguish patients from controls, and establish a combination of region and modality that classified 21/22 subjects correctly. This exploratory study is the first time all three modalities have been combined in a study of FTLD patients; it shows that combining MR modalities may lead to improved classification of FTLD patients and better identification of affected tissue.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the Copyright) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the Intellectual Property) and any reproductions of copyright works in the thesis, for example graphs and tables (Reproductions), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

4. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Librар’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
Acknowledgements

I would like to acknowledge the huge part played in this study by my late husband, Clive Beaumont. Even though he died a decade before the study started, he made me realise how necessary the study was and further research is, and gave me the motivation to persevere. So this thesis is dedicated to his memory.

Many many thanks are due to my supervisors, Geoffrey J M Parker, Laura M Parkes and Roland Zahn for their patience and encouragement at all times.

Thanks also to Karl Embleton for modifying his distortion code to fit my data, to Hamied Haroon for the diffusion tensor calculation code, and to Mark Dobbs for showing me how to read Philips ParRec files.

Thanks to many of my colleagues for advice, in particular Rishma Vidyasagar, Dan Cox, Ross Little, Mike Berks, Anita Banerji, Claudia Lindner, Sha Zhao and Michael Goldsmith.

Thanks to the radiographers at the 3T scanner at Salford Royal Infirmary, who remained calm and unflappable throughout.

Thanks to my family for their rather bemused but whole-hearted support.

Above all thank you to the patients and controls who gave up their time to be scanned, and complete the seemingly-endless neuropsychological tests.
About the author

Qualifications
MA Physics   Oxford University
Postgraduate diploma in software engineering   Oxford University
MPhil   Modelling cognitive decline in dementia   Open University
MSc Medical Statistics   Leicester University

Posters arising from this work
Using MRI Arterial Spin Labelling to investigate cerebral perfusion in frontotem-poral dementia  Ecost ASL in dementia Mar2015
Using MRI Arterial Spin Labelling to investigate cerebral perfusion in frontotem-poral dementia - FTD Conference Oct 2014
Using Arterial Spin Labelling MRI to measure perfusion and arrival time in pa-tients with frontotemporal lobar degeneration - ISMRM Apr 2013
Spin echo offers improved ASL signal in ventral brain regions - FTD Conference Oct 2012
Using ASL MRI with Look-Locker readout to measure perfusion and arrival time in patients with Frontotemporal Lobar Degeneration - ISMRM Perfusion Work-shop, October 2012
Quantification of CBF and arrival time using Arterial Spin Labelling: Correction for bolus duration - British Chapter ISMRM Student Symposium, March 2011

Key research experience
Acquiring and analysing MR images
Conducting neuropsychological tests
Software development and computer modelling

Other publications
## Glossary

<table>
<thead>
<tr>
<th>First entry</th>
<th>Second entry</th>
<th>Third entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{18}$F-THK523PET</td>
<td>PET ligand</td>
<td></td>
</tr>
<tr>
<td>AAL/aal</td>
<td>automated anatomical labelling</td>
<td>brain atlas</td>
</tr>
<tr>
<td>AAT</td>
<td>Arterial arrival time</td>
<td></td>
</tr>
<tr>
<td>ACE-R</td>
<td>Addenbrook’s cognitive examination - revised</td>
<td>cognitive test</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
<td></td>
</tr>
<tr>
<td>ALS</td>
<td>Amyotrophic Lateral Sclerosis</td>
<td></td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>Anterior ventromedial pre-frontal cortex</td>
<td>brain region</td>
</tr>
<tr>
<td>ASL</td>
<td>Arterial spin labelling</td>
<td>MR term</td>
</tr>
<tr>
<td>BF</td>
<td>best friend</td>
<td>cognitive test term</td>
</tr>
<tr>
<td>BOLD</td>
<td>Blood oxygenation level dependent</td>
<td>MR term</td>
</tr>
<tr>
<td>bvFTD</td>
<td>behavioural-variant fronto-temporal dementia</td>
<td></td>
</tr>
<tr>
<td>C9ORF72</td>
<td></td>
<td>Gene implicated in FTLD</td>
</tr>
<tr>
<td>CASL</td>
<td>Continuous ASL</td>
<td>MR sequence</td>
</tr>
<tr>
<td>CBD</td>
<td>Corticobasal degeneration</td>
<td></td>
</tr>
<tr>
<td>CBF</td>
<td>cerebral blood flow</td>
<td></td>
</tr>
<tr>
<td>CBS</td>
<td>Corticobasal syndrome</td>
<td></td>
</tr>
<tr>
<td>CCg</td>
<td>the part of the cingulum bundle connected to the cingulate gyrus</td>
<td>white matter tract</td>
</tr>
<tr>
<td>CH</td>
<td>the part of the cingulum bundle connected to the cingulum hippocampus</td>
<td>white matter tract</td>
</tr>
<tr>
<td>CHMP2B</td>
<td></td>
<td>FTLD genotype</td>
</tr>
<tr>
<td>CSF</td>
<td>Cerebrospinal fluid</td>
<td></td>
</tr>
<tr>
<td>DTI</td>
<td>Diffusion tensor imaging</td>
<td>MR term</td>
</tr>
<tr>
<td>EPI</td>
<td>Echo-planar imaging</td>
<td>MR sequence</td>
</tr>
<tr>
<td>EPISTAR</td>
<td>Echo-planar imaging and signal targeting with alternating RF</td>
<td>MR sequence</td>
</tr>
<tr>
<td>FA</td>
<td>Fractional Anisotropy</td>
<td>diffusion metric</td>
</tr>
</tbody>
</table>

*Continued on next page*
Table 1 – Continued from previous page

<table>
<thead>
<tr>
<th>First entry</th>
<th>Second entry</th>
<th>Third entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAS</td>
<td>Letter fluency test with letters F &amp; A &amp; S</td>
<td>cognitive test</td>
</tr>
<tr>
<td>FDG</td>
<td>PET ligand</td>
<td></td>
</tr>
<tr>
<td>FID</td>
<td>Free induction decay</td>
<td>MT term</td>
</tr>
<tr>
<td>FPR</td>
<td>False positive rate</td>
<td></td>
</tr>
<tr>
<td>FTD</td>
<td>Frontotemporal dementia</td>
<td></td>
</tr>
<tr>
<td>FTLD</td>
<td>Frontotemporal lobar degeneration</td>
<td></td>
</tr>
<tr>
<td>FTLD-FUS</td>
<td>FTLD with FUS pathology</td>
<td></td>
</tr>
<tr>
<td>FTLD-tau</td>
<td>FTLD with tau pathology</td>
<td></td>
</tr>
<tr>
<td>FTLD-TDP</td>
<td>FTLD with TDP pathology</td>
<td></td>
</tr>
<tr>
<td>FUS</td>
<td>Fused-in sarcoma</td>
<td>FTLD pathology</td>
</tr>
<tr>
<td>fvFTD</td>
<td>frontal-variant FTD</td>
<td></td>
</tr>
<tr>
<td>FWHM</td>
<td>full width half-height maximum</td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>Gradient-echo</td>
<td>MR sequence</td>
</tr>
<tr>
<td>GENFI</td>
<td>Genetic frontotemporal dementia initiative</td>
<td>Multi-centre organisation</td>
</tr>
<tr>
<td>GM</td>
<td>grey matter</td>
<td></td>
</tr>
<tr>
<td>GRASE</td>
<td>Gradient- and Spin-Echo</td>
<td>MR sequence</td>
</tr>
<tr>
<td>GRN</td>
<td>Progranulin</td>
<td>Gene implicated in FTLD</td>
</tr>
<tr>
<td>HASTE</td>
<td>half-Fourier acquisition single-shot turbo</td>
<td>MR sequence</td>
</tr>
<tr>
<td>HMPAO</td>
<td>99mTc-hexamethylpropyleneamineoxime</td>
<td>PET ligand</td>
</tr>
<tr>
<td>IFOF</td>
<td>inferior frontooccipital fasciculus</td>
<td>white matter tract</td>
</tr>
<tr>
<td>ILF</td>
<td>inferior longitudinal fasciculus</td>
<td>white matter tract</td>
</tr>
<tr>
<td>JHU</td>
<td>John Hopkins University</td>
<td>white matter atlas</td>
</tr>
<tr>
<td>LAntTL</td>
<td>Left anterior temporal lobe</td>
<td>brain region</td>
</tr>
<tr>
<td>LPostTL</td>
<td>Left posterior temporal lobe</td>
<td>brain region</td>
</tr>
<tr>
<td>MAPT</td>
<td>Microtubule associated protein tau</td>
<td>Gene implicated in FTLD</td>
</tr>
<tr>
<td>MD</td>
<td>Mean diffusivity</td>
<td>diffusion metric</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini-mental state exam</td>
<td>cognitive test</td>
</tr>
<tr>
<td>MND</td>
<td>Motor neurone disease</td>
<td></td>
</tr>
<tr>
<td>MNI</td>
<td>Montreal Neurological Institute</td>
<td>brain coordinate system</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
<td></td>
</tr>
<tr>
<td>NaN</td>
<td>Not a number</td>
<td>Matlab abbreviation</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>First entry</th>
<th>Second entry</th>
<th>Third entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>PALPA</td>
<td>Psycholinguistic assessments of language processing in aphasia</td>
<td>cognitive test battery</td>
</tr>
<tr>
<td>PASL</td>
<td>Pulsed arterial spin labelling</td>
<td>MR sequence</td>
</tr>
<tr>
<td>PCASL</td>
<td>Pseudo-continuous arterial spin labelling</td>
<td>MR sequence</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
<td>Imaging system using radioactive ligands</td>
</tr>
<tr>
<td>PiB</td>
<td>Pittsburgh compound B</td>
<td>Pet ligand</td>
</tr>
<tr>
<td>PK11195</td>
<td></td>
<td>Pet ligand</td>
</tr>
<tr>
<td>Post VMPFC</td>
<td>Posterior ventromedial prefrontal cortex</td>
<td>brain region</td>
</tr>
<tr>
<td>PPA</td>
<td>Primary progressive aphasia</td>
<td>illness affecting language</td>
</tr>
<tr>
<td>QUIPSS/QUIPSSII</td>
<td>Quantitative imaging of perfusion using a single subtraction</td>
<td>MR sequence</td>
</tr>
<tr>
<td>RAnTL</td>
<td>Right anterior temporal lobe</td>
<td>brain region</td>
</tr>
<tr>
<td>RF</td>
<td>Radio-frequency</td>
<td>MR pulse</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
<td></td>
</tr>
<tr>
<td>RPostTL</td>
<td>Right posterior temporal lobe</td>
<td>brain region</td>
</tr>
<tr>
<td>SD</td>
<td>Semantic dementia</td>
<td>a type of FTLD</td>
</tr>
<tr>
<td>SE</td>
<td>Spin-echo</td>
<td>MR sequence</td>
</tr>
<tr>
<td>SENSE</td>
<td>Sensitivity encoding</td>
<td>MR sequence</td>
</tr>
<tr>
<td>SLF</td>
<td>superior longitudinal fasciculus</td>
<td>white matter tract</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise</td>
<td></td>
</tr>
<tr>
<td>SPECT</td>
<td>Single photon emission computed tomography</td>
<td>Imaging system using radioactive ligands</td>
</tr>
<tr>
<td>SPM</td>
<td>Statistical parametric mapping</td>
<td>Toolbox for image analysis</td>
</tr>
<tr>
<td>T</td>
<td>tesla</td>
<td>SI Unit of magnetic field strength</td>
</tr>
<tr>
<td>TARDP</td>
<td>TAR-DNA binding protein</td>
<td>Gene implicated in FTLD</td>
</tr>
<tr>
<td>TE</td>
<td>Echo time</td>
<td>MR term</td>
</tr>
<tr>
<td>TI</td>
<td>Inversion time</td>
<td>MR term</td>
</tr>
<tr>
<td>TPR</td>
<td>True positive rate</td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>Repetition time</td>
<td>MR term</td>
</tr>
<tr>
<td>UF</td>
<td>Uncinate fasciculus</td>
<td>white matter tract</td>
</tr>
<tr>
<td>VB</td>
<td>Voxel-based (analysis)</td>
<td>Method for analysing images</td>
</tr>
<tr>
<td>VBM</td>
<td>Voxel-based morphometry</td>
<td>Method for analysing image shape</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>First entry</th>
<th>Second entry</th>
<th>Third entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCP</td>
<td>Valosin containing protein</td>
<td>Gene implicated in FTLD</td>
</tr>
<tr>
<td>WM</td>
<td>white matter</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 – *Continued from previous page*
Chapter 1

Introduction

Alzheimer’s disease (AD) is the best-known of a group of illnesses, the dementias, that cause progressive cognitive deficits. Although often considered as diseases of the elderly, all can occur in people of middle-age. The predominant symptom of AD is loss of recent memory, so a patient can retain childhood memories, but repeatedly ask the same question and forget the answer. Frontotemporal Lobar Degeneration (FTLD) is one of the many names given to a group of non-Alzheimer dementias that are heterogeneous clinically, pathologically and genetically. They all cause selective and progressive cognitive deficits, initially behavioural changes or difficulties with using language, which are accompanied by atrophy of the frontal and/or temporal brain regions [Neary et al., 1998]. They are not common, with prevalence estimates at between 4 and 15 per 100,000 in the European/US working age population [Rabinovici and Miller, 2010]. Collectively they are the second most common cause of dementia in the working age population after AD [Rabinovici and Miller, 2010], and may have similar prevalence, although estimates vary from 1/3 the incidence of AD [Harvey et al., 2003] to similar incidence to AD [Rabinovici and Miller, 2010]. Historically FTLD was called Pick’s disease, and still is often considered a very rare disease. Spatt quotes Onari and Spatz from 1926: We hold the opinion that Pick’s disease does not belong to the extreme rarities, but that it remains often unrecognised by the clinician as well as the anatomist because not enough attention is directed to it [Spatt, 2003]. This comment could have been made today [Warren et al., 2013].

FTLD is difficult to diagnose in the early stages; there can be a delay of sev-
eral years between the first symptoms being noticed by the patient or their family and the patient being referred to a neurologist [Woolley et al., 2011]. The neurologist will carry out a battery of neuropsychological tests, and usually these will be accompanied by a T_1-weighted MRI scan to look for atrophy, and a PET scan to look for hypometabolism; the pattern of any changes can indicate a neurodegenerative disease, and give a differential diagnosis if dementia is suspected (see chapter 1.2). The MRI and PET images will be examined visually by an expert radiologist, but early changes can be subtle and difficult to distinguish from the wide range of normal images [Mendez et al., 2007].

The aim of this study is to test whether other MRI modalities can provide more information than is available in the T_1-weighted MRI image, and hence improve the ability to distinguish between patients with FTLD and controls, with the eventual aim of aiding the diagnosis of FTLD. While PET has already shown its value, it is expensive and invasive, as it involves injecting the subject with a short-lived radioactive tracer. The MRI modalities considered here are specifically Arterial Spin Labelling (ASL), which can measure cerebral blood flow, and diffusion-weighted MRI, which can measure the freedom of water movement and hence is sensitive to brain microstructure. PET studies have shown cerebral hypometabolism in FTLD (see chapter 2.6.1); CBF and glucose metabolism are closely linked [Buxton, 2010], and so it is a reasonable hypothesis that ASL can provide a means of distinguishing FTLD patients from controls. A postmortem study of an FTLD brain found altered diffusion metrics in the frontal lobes [Larsson et al., 2004] (see chapter 2.6.3), and a 2006 in-vivo study also found differences between FTLD patients and controls [Yoshiura et al., 2006].

This study also considers which regions within the brain are most implicated in the illness. It is intended to generate hypotheses which may in future be used to improve the diagnosis of FTLD. A secondary aim is to see if these MRI modalities can give more information about the differences in the tissue vasculature and structure between tissue involved in FTLD and healthy tissue. This can provide clues to the underlying physiological processes that are altered in FTLD.
1.1 Overview of this thesis

In this chapter I go on to give an introduction to the group of illnesses that are covered by the name of frontotemporal lobar degeneration.

In chapter 2 I give the background to the MRI techniques used, consider the use of MR and other types of imaging in FTLD, and look at the power needed to provide a distinction between FTLD patients and controls.

In chapter 3 I describe the modelling needed to generate a quantitative measurement of cerebral blood flow (CBF) and arterial arrival time (AAT) from a perfusion-weighted image, and consider the minimum data required to give a reasonably accurate measurement of CBF and AAT within an acquisition time tolerable to most subjects. I build on published models for quantifying ASL data, but suggest another model, and carry out extensive simulations to observe the effect on the estimation of CBF of errors in other parameters introduced in the models.

In chapter 4 I consider MRI techniques which can be used to better acquire MR images of the ventral frontal and temporal lobes, and describe the results of a study on young healthy subjects comparing two MR image acquisition techniques. As is clear from its name, FLTD causes degeneration and atrophy of the frontal and temporal lobes [Brun, 1987]. MR images of the ventral regions of these lobes can be difficult to acquire because of their proximity to air spaces in the auditory canals and orbital cavities. The susceptibility differences caused by these cavities cause magnetic field inhomogeneities which in turn cause distortion and signal dropout in these regions. This comparison of gradient-echo and spin-echo sequences is novel.

In chapter 5 I move on to describe the first results of the clinical research study in FTLD: a study involving 20 patients showing features of FTLD and 28 controls and using $T_1$-weighted, ASL and diffusion-weighted images. Not all the patients and controls were included in the imaging study for reasons described in the relevant chapters. In this chapter I describe the clinical criteria for including patients and controls, the neuropsychological tests carried out and the $T_1$-weighted MRI images acquired in this study, and describe analyses of these images for 17 patients and 18 controls. I use voxel-based analysis (VBA) and regional analyses; the VBA confirms the results of previous studies, but the ROI analyses are novel.
In chapters 6 and 7 I go on to report and discuss the findings from the ASL and diffusion-weighted images respectively.

In chapter 8 I describe a classification using linear discriminant analysis, and consider which MRI modalities, and which brain regions, are best at classifying this group of patients and controls.

In chapter 9 I discuss the findings of the study, and consider the direction of further research.

1.2 Frontotemporal lobar degeneration

The criteria for FLTD, described by Neary et al originally [Neary et al., 1998] and recently updated by Rascovsky at el [Rascovsky et al., 2011], describe three main clinical variants of FTLD:

- behavioural variant FTD (bvFTD): patients present with changes in personality and behaviour, specifically a decline in social conduct, impairment in regulation of personal conduct, emotional blunting, and loss of insight.

- semantic dementia (SD): patients present with fluent speech that is empty of meaning, loss of understanding of word meaning, impaired recognition of familiar faces or objects yet preserved single-word repetition, ability to read aloud and write to dictation.

- Progressive non-fluent aphasia: patients present with non-fluent spontaneous speech with agrammatism, phonemic paraphasias, and/or anomia.

SD is closely related to the semantic variant of primary progressive aphasia (PPA) [Gorno-Tempini et al., 2011]. The criteria for both are included in appendix A.

People who develop frontotemporal dementia face several hurdles in getting a diagnosis. The disease has a slow and insidious onset, so patients and their family and friends can become accustomed to the changes, and believe them to be normal. When the changes become so marked as to cause the patient or their partners/colleagues to seek medical advice, a neurodegenerative disorder is often not
considered because of the person’s relatively young age. Patients themselves are often unaware of or unconcerned by their symptoms, leading to a tension which can make it difficult for the physician to decide whom to believe. Behavioural changes caused by the illness can be similar to those caused by a non-organic psychiatric disorder, so that patients can be treated for some time (in some cases, several years [Woolley et al., 2011]) before being referred to a neurologist.

Even when seen by an appropriate specialist, the path to a diagnosis is not clear. The early changes can be subtle and difficult to distinguish from “normal” behaviour or use of language. MRI or PET can show atrophy or hypometabolism, and in the later stages of the illness these changes can be very noticeable even to the untrained eye. However, in the early stages of the illness it can be difficult to decide if images are within a normal range or not. Figure 1.1 shows the T₁-weighted images of a patient and a cognitively normal control in this study. FTLD is distinguished by atrophy visible as dilation of the ventricles and diminution of the frontal and temporal lobes, but it can be difficult to determine the cut-off between normal age-related atrophy and atrophy caused by a neurodegenerative process. A recent complication is the description by Kipps et al of a group of patients who meet the clinical criteria for a diagnosis of bvFTD, but whose condition does not deteriorate, or deteriorates much more slowly than in other cases [Kipps et al., 2009].

![Figure 1.1: Axial images of T₁-weighted images of a patient and a control showing the difficulty of diagnosis using only these images. FTLD is distinguished by atrophy visible as dilation of the ventricles and diminution of the frontal and temporal lobes, but it can be difficult to determine the cut-off between normal age-related atrophy and atrophy caused by a neurodegenerative process.](image-url)
1.2.1 History

Arnold Pick was a Czech neurologist and psychiatrist who held the then-contested idea that it was possible for an illness to cause specific cognitive deficits while leaving other cognitive functions intact. From 1892 onwards he described several patients with “progressive loss of vocabulary, marked aphasia... disturbed understanding of speech”, and patients who were “doing bizarre things such as tearing up vegetables in the garden”. (These excerpts are quoted by Kertesz et al [Kertesz et al., 2000]). At postmortem these patients had regions of cerebral atrophy mostly affecting the frontal and temporal lobes. Alois Alzheimer stained and examined the brain tissue, and described “swollen neurons” and “ballooned cells” now called Pick bodies. The illness was given the name of Pick’s disease, characterised by behavioural disturbances and loss of conceptual knowledge. It was later recognised that Pick’s disease was one of a spectrum of disorders which includes primary progressive aphasia, semantic dementia and frontotemporal dementia.

1.2.2 FTLD terminology

Unfortunately, while the clinical symptoms Pick described are not uncommon, the clinical symptoms frequently occur without the pathology (Kertez et al found pathology in around 20% of the clinical cases [Kertesz et al., 2000]). There arose a degree of confusion, with some clinicians using the term Pick’s disease to describe the clinical symptoms, and some reserving it for cases with the pathology, and hence only diagnosable after death. Corticobasal degeneration (CBD), and motor neurone disease (MND) were also recognised by Pick to be part of the same spectrum of diseases, but subsequently often separated. In 1992 criteria for PPA were published [Mesulam and Weintraub, 1992]; they described a language disorder which occurred in isolation for a time, but was often later accompanied by other cognitive deficits. In the late 1990s Neary et al published the Manchester-Lund criteria for FTLD [Neary et al., 1998], including SD and bvFTD. These criteria are given in appendix A. MND is also known as amyotrophic lateral sclerosis (ALS), Lou-Gehrig’s Disease and Charcot Disease; patients diagnosed with MND can show some FTLD symptoms and vice versa. CBS is often clinically distinct from FTLD, but the underlying neuropathology of CBD can be found in some pa-
patients presenting with FTLD.

1.2.3 FTLD diagnosis

The diagnosis of FTLD is well described by Rabinovici and Miller [Rabinovici and Miller, 2010]. bvFTD in particular is difficult to diagnose in the early stages as symptoms suggestive of bvFTD include apathy, disinhibition, poor planning, poor organisation, hyperactive behaviours such as wandering and pacing, and changes in eating, sleeping and sexual behaviours [Josephs et al., 2011]. The clinician must first decide if such behaviour is abnormal, given the wide range of “normal” and the slow onset of symptoms in FTLD, and then decide what might be the cause of such changes. Semantic dementia causes progressive loss of knowledge about words and objects, while other skills are retained. This skill is more amenable to neuropsychological testing, but again there is a large range of “normal”, so subtle changes may be missed.

Once FTLD is suspected, the clinician must first exclude treatable conditions that can cause similar symptoms, such as cancer, metabolic disturbances, nutritional deficiencies, infections of the central nervous system, substance abuse, vascular disease, heavy metal toxicity, psychiatric disturbance... [Rabinovici and Miller, 2010]. The list is long. The diagnosis of bvFTD was further complicated in 2009 by the description by Kipps et al of a bvFTD phenocopy [Kipps et al., 2009]. Nonetheless Rabinovici and Miller go on to say in our experience, it is far more common for patients with degenerative FTLD to be misdiagnosed with a psychiatric disorder than vice versa.

1.2.4 FTLD phenocopy

Kipps et al in 2007 commented that as many as half of patients they had seen who were diagnosed with bvFTD had normal MRI T₁-weighted images, and that these patients did not deteriorate cognitively over the same timescale as patients with the same diagnosis but with abnormal MRI images [Kipps et al., 2007]. This was followed by another study in 2009 [Kipps et al., 2009], which followed 12 controls and 24 patients with bvFTD; 15 with abnormal MRI images and 9 with normal MRI images. They compared MRI regional atrophy, regional hypometabolism
measured by radioactive tracer images, and results of cognitive tests over a 6-year period. The MRI-abnormal group showed more hypometabolism in frontotemporal regions than the MRI-normal group; the MRI-normal group were indistinguishable from controls. Initially cognitive test performance for the MRI-abnormal group was lower than for the MRI-normal group, and for the MRI-normal group was lower than for the controls; but at follow-up scores for the MRI-abnormal group were still lower, while scores for the MRI-normal group were not different from scores for the controls. Duration of illness was not different initially for the two patient groups, but in 2009 8 of the 15 MRI-abnormal group had died: 6 came to postmortem and showed pathological symptoms of FTLD. All of the MRI-normal group were still alive. Kipps et al consequently suggested that a clinical diagnosis of bvFTD did not necessarily indicate a neurodegenerative process, and recommended the addition of imaging abnormalities to the diagnostic criteria. This recommendation was implemented in the revised criteria published by Rascovsky et al [Rascovsky et al., 2011].

1.2.5 Genetics

Although the majority of FTLD cases seem to be sporadic (something which is important to the close relatives of a patient), FTLD differs from Alzheimer’s disease in that a significant proportion of patients have a close relative who also had/has the illness; 30%-50% of cases are thought to be familial [Seelaar et al., 2011]. The pattern of inheritance is autosomal dominant, i.e. a person carrying the gene will eventually develop the illness.

Mutations in the following genes are known at the moment;

- Microtubule associated protein tau (MAPT) [Hutton et al., 1998]
- Progranulin (GRN) [Baker et al., 2006] [Cruts et al., 2006]
- Valosin containing protein (VCP) [Watts et al., 2004]
- Charged multivesicular body protein 2B (CHMP2B) [Skibinski et al., 2005]
- TAR-DNA binding protein (TARDP) [Benajiba et al., 2009]
- C9ORF72 [DeJesus-Hernandez et al., 2011], [Renton et al., 2011]
The first two mutations account for around half of familial cases, while the next three account for fewer than 5%. In all cases the pattern is one of autosomal dominant inheritance [Seltman and Matthews, 2012]. It is likely that more mutations causing FTLD will be found in the future, as large numbers of families are required to provide the necessary statistical power. These numbers will now be available thanks to the creation of the Genetic Frontotemporal Dementia Initiative (GENFI), a group of European and Canadian institutions with expertise in familial FTLD who have agreed to share data.

1.2.6 FTLD pathology

When the brains of FTLD patients are examined at postmortem, localised atrophy, gliosis and neuronal loss are seen in the frontal and temporal lobes. In most cases there is also a build-up of abnormal proteins. These abnormal proteins fall into three groups:

- tau, the microtubule-associated protein - FTLD-TAU
- TDP-43, the transactive response DNA binding protein of 43 kD - FTLD-TDP
- FUS, the tumour associated protein fused in sarcoma - FTLD-FUS

There is currently no definite causality between neuropathology and clinical presentation [Seltman and Matthews, 2012], however there is a strong association of FTLD-TDP with SD, of FTLD-Tau with bvFTD and CBD and of FTLD-FUS with young-onset ALS.

1.2.7 Treatment

Currently there are no approved treatments for FTLD, although at least one is in clinical trials (no peer-reviewed papers as yet, but see http://taurx.com/lmtx-forftd/). Management is therefore directed to controlling symptoms and otherwise helping patients and their caregivers cope with the impact of their illness [Warren et al., 2013].
1.2.8 What’s in a name?

As Shakespeare so famously wrote:

\[
\begin{align*}
\text{What’s in a name? That which we call a rose} \\
\text{By any other name would smell as sweet.}
\end{align*}
\]

But the name given to a condition can have more significance, especially in the days of computerised searches. Grey’s elegy has it that

\[
\begin{align*}
\text{Full many a flower is born to blush unseen,} \\
\text{And waste its sweetness on the desert air}
\end{align*}
\]

A search for “frontotemporal dementia” may not find articles on “fronto-temporal dementia”, and will definitely not find those on “fronto-temporal lobar degeneration”. I find it ironic that Bruno et al in their paper on Alzheimer’s disease and fronto-temporal dementia [Bruno et al., 2012] uses “fronto-temporal” throughout where many others omit the hyphen, changes “fronto-temporal dementia” to “fronto-temporal lobar degeneration” partway through the paper, and concludes with the constant dialogue between clinicians and physicists is producing an increasingly common language, promoting interdisciplinary collaboration, probably the only key to the ultimate defeat of these devastating diseases. The illnesses are intrinsically heterogeneous, but any review of studies into FTLD is made even more difficult by the varying nomenclature.

The distinctions between semantic dementia, behavioural-variant FTD and primary progressive aphasia seem somewhat arbitrary, when the diagnosis seems to depend on the time at which various symptoms appear, and eventually most patients seem to display symptoms typical of all the variants. The distinction between FTD and FTLD in particular seems extremely arbitrary, since many papers, and many talks I have attended, use them interchangeably. I have tried to be consistent in this thesis.

1.2.9 Summary

As shown above, FTLD is a confusing group of illnesses. The illnesses are frequently not recognised in the early stages, because of the slow onset, because of the relatively young age at which they can occur, and because of the ill-defined nature of symptoms. There is currently no marker in CSF which can show the
presence of the abnormal proteins described above. There are now known genetic mutations which can give a good indication of a diagnosis, but the mutations only account for around half of inherited cases, which themselves are fewer than half of all cases. Despite this there are now ongoing clinical trials of disease-modifying drugs; these have been developed from studies of familial cases. There is little hope for more drug development unless it is possible to diagnose cases early, and then monitor the progress of the illness.

One may think that with no easy diagnosis and no available treatment, FTLD is best ignored for as long as possible. However, carers and patients I have spoken to have without exception stressed how important an early diagnosis is for them, and how difficult life can be without a clear diagnosis. Warren states Early, accurate diagnosis and mobilisation of appropriate support services at present offer the best prospect of effective management for patients with FTD [Warren et al., 2013].

PET has proved to be a good diagnostic tool, but the necessary use of a short-lived radioactive tracer limits its availability and frequency of use. MRI is cheaper, more available and involves no exposure to radioactivity for patients or staff. The development of a diagnostic aid from better use of MRI is therefore a worthwhile goal. As can be seen from figure 1.1, in the early stages a T₁-weighted MRI image may not be very diagnostic. There is hope that ASL MRI, which explores the brain microvasculature, and diffusion MRI, which explores tissue microstructure, may be more diagnostic. These methods are explained in the next chapter.
Chapter 2

Background

2.1 MRI basics

2.1.1 Magnetic resonance

The signal that is detected in all the MRI sequences in this study is generated by the atomic nucleus of hydrogen in water; the hydrogen nucleus consists of a single proton. It has a positive charge, and a quantum property of spin, which together give the nucleus a magnetic moment. Without an applied magnetic field, the individual magnetic moments are not aligned, and there is no net magnetic moment. When placed in a magnetic field, the nuclei tend to align along or against the magnetic field, giving the material a net magnetic moment, which is small but measurable. The imbalance between nuclei aligning with the field (the preferred, lower energy direction) and against the field depends on the field strength and temperature. In a 3T field and at the temperature of the human body there are around ten per million more nuclei aligned with the field than against the field.

The protons do not align directly along the field, but precess around it like spinning tops. The rate of precession is the Larmor frequency \( \omega_0 \),

\[
\omega_0 = \gamma B_0, \quad (2.1)
\]

where \( B_0 \) is the magnetic field strength, \( \gamma \) is the gyromagnetic ratio and is specific
to each element. For protons $\gamma/2\pi = 42.58 \text{ MHz/T}$. In the normal state the protons precess at random, and so there is a net longitudinal magnetisation along the field, but no transverse magnetisation.

If a radio-frequency (RF) pulse $B_1$ in a direction orthogonal to $B_0$ is applied to the protons, it exerts a torque which can change the axis of precession. If $B_1$ has the same frequency as the precessing protons, there is a resonance effect which will enhance the effect of the torque, and hence energy is transferred from the RF pulse to the protons. The energy transfer will only occur at when the frequency of $B_1$ matches the Larmor frequency. The individual spins are also brought into phase with the RF pulse, so there is now a net transverse magnetisation. This transverse magnetisation rotating in a magnetic field causes a detectable signal, which can be measured because there will be a current induced in a coil placed with the axis perpendicular to $B_0$. This is the MR signal. The angle through which the spins are tilted is called the flip angle $\alpha$, and the maximum potential signal occurs when $\alpha = 90^0 = \pi/2$. $\alpha$ depends on the amplitude and duration of the $B_1$ pulse:

$$\alpha = \gamma.B_1^{\text{amplitude}}.B_1^{\text{duration}}$$  \hspace{1cm} (2.2)

This is shown schematically in figure 2.1. Figure 2.1(a) shows the system at equilibrium, with a net magnetisation vector along the z direction. In figure 2.1(b) an RF pulse $B_1$ has been applied in the x-direction: the effect is to flip the magnetisation vector through $90^0$ so it lies along the y-direction. In figure 2.1(c) a more intense RF pulse has been applied (either longer duration or greater magnitude) which has flipped the net magnetisation vector through $180^0$, so it now lies along z. Once the RF pulse ends, this signal decays:

- the longitudinal magnetisation, which has been decreased by the applied pulse, gradually returns to equilibrium. The rate at which this occurs depends on the molecular environment around the protons, and is characterised by the parameter $T_1$. This is known as spin-lattice relaxation.

- the transverse magnetisation decays as the individual protons no longer precess in phase. The exact rate at which each proton precesses is due to the local magnetic field, and hence will be slightly different for every proton, as interactions with neighbouring protons will produce local fluctuating fields. The rate at which the transverse magnetisation decays is characterised by the parameter $T_2$. This is known as spin-spin relaxation.
there are additional field inhomogeneities due to factors external to the tissue, e.g. non-linearities in the applied field $B_0$, or susceptibility effects caused by tissue-air interfaces. These effects increase the decay of the transverse magnetisation, and to a good approximation, the decay can be modelled by a single component exponential function which is characterised by the parameter $T_2^*$. Both transverse and longitudinal relaxations occur exponentially, so $T_1$ is the time in which the longitudinal magnetisation recovers to approximately 67% of its prepulse value and $T_2^*$ is the time in which the transverse magnetisation decays to approximately 33% of the initial value. Normally $T_1 >> T_2 > T_2^*$.

**Figure 2.1:** The effect of an RF pulse. Figure 2.1(a) shows the system at equilibrium, with a net magnetisation vector along the z direction. In figure 2.1(b) an RF pulse $B_1$ has been applied in the x-direction: the effect is to flip the magnetisation vector through $90^0$ so it lies along the y-direction. In figure 2.1(c) a more intense RF pulse has been applied (either longer duration or greater magnitude) which has flipped the net magnetisation vector through $180^0$, so it now lies along -z. The magnetisation vector is shown in red throughout, and the $B_1$ pulse in blue.

### 2.2 The MR signal

#### 2.2.1 Free induction decay (FID)

If the MR signal is sampled after the application of an RF pulse at the Larmor frequency, it has the form shown in figure 2.2. The signal oscillates at the Larmor frequency within an exponentially decaying envelope. In an ideal world the en-
velope would relax at $T_2$, as spins dephase due to interactions with nearby spins. However, the real relaxation rate is $T_{2}^{*}$, because the rate of dephasing is increased by field inhomogeneities.

**Figure 2.2:** Free induction decay signal in an ideal and real world. The signal oscillates at the Larmor frequency, and decays in an ideal world with no external field inhomogeneities at rate $T_2$ (extreme outer envelope) and in the real world at rate $T_{2}^{*}$ (bold envelope).

### 2.2.2 Spin-echo, gradient-echo and sequence diagrams

If the applied RF pulse has a flip angle of $90^\circ$, the spins will start to rotate in a plane orthogonal to the field $B_0$, and gradually dephase; this is shown in figure 2.3(a). If after a time $t$ another RF pulse is applied with a flip angle of $180^\circ$, the spins will be flipped, and precess in the opposite direction as shown in figure 2.3(b). At a time equal to $2t$, the spins will be back in phase, as the protons which precess fastest will have furthest to travel after the reversal, and there will be an echo of the original signal. This is a **spin echo**, and the formation of the echo is shown in figure 2.3(c). Another $180^\circ$ pulse $2t$ after the first refocusing pulse will generate another echo, with the maximum signal at each echo diminishing with relaxation time $T_2$. Note that because the spins themselves are flipped, the effect of any static inhomogeneities will be reversed and hence cancelled out. The random field inhomogeneities local to each proton that are caused by its neighbours will of course still be felt, hence the relaxation with time $T_2$. The time $2t$ is the echo time, $TE$. If a constant gradient $G_{GE}$ is applied after the RF pulse which generated the initial signal, the spins will dephase as usual in the presence of this field. If after time $t$ this applied
Figure 2.3: The evolution of the spin-echo. Figure 2.3(a) shows the spins immediately after the application of the $90^\circ$ pulse; all the spins have the same phase. Figure 2.3(b) shows the transverse magnetisation some time after the application of the $90^\circ$ RF pulse; the spins have dephased because of local field inhomogeneities. Figure 2.3(c) shows the spins immediately after the application of the $180^\circ$ refocussing pulse. After time TE the spins have again come back into phase, as shown in figure 2.3(d).

gradient is reversed, the spins will reverse direction and after time $TE = 2t$ there will be an echo. This is a gradient echo. However, this time the refocussing is caused by the reversal of the applied gradient, and any local static inhomogeneities will not be reversed. In the initial phase before the gradient reversal, the local field is the sum of local inhomogeneities $G_{local}$ and $G_{GE}$, i.e. $G_{local}+G_{GE}$. After the gradient reversal, the local field is $G_{local}-G_{GE}$. Hence the refocussing is not as complete as for the spin-echo, and the envelope of the echoes will now have a decay time of $T_2^*$. The advantage of a gradient-echo sequence is that there is only 1 RF pulse, and so less energy is applied to the sample. Typically a lower flip angle than $90^\circ$ is used. The rate at which MRI signals can be acquired depends on the time taken for the previous signal to relax; as spin-echo requires the application of two RF pulses, one of $90^\circ$ and one of $180^\circ$, while gradient-echo requires only one RF pulse, with a flip angle which can be less than $90^\circ$, the repetition rate for a gradient-echo sequence can be much faster than for a spin-echo sequence.

Simplified schematic diagrams for spin-echo and gradient-echo sequences are shown in figure 2.4.

2.2.3 Positional information

The previous section has described the generation and acquisition of the MR signal. For this signal to be used to generate an image, extra data on the location of the signal is needed. Magnetic field gradients can be used to encode positional
Figure 2.4: Simple spin-echo and gradient-echo sequences. Figure 2.4(a) shows a spin-echo sequence: a $90^\circ$ RF pulse is applied, causing the spins to flip orthogonally to $B_0$ and generate a FID signal. After time $TE/2$ a $180^\circ$ pulse is applied. The spins which precessed to be out of phase (hence the decay of the FID) will now precess in the opposite direction, and come back into phase at time $TE$ from the first pulse, generating an echo. This refocusing process is shown in figure 2.3. Figure 2.4(b) shows a gradient-echo sequence: an RF pulse of flip angle $\alpha$ is applied, causing the spins to flip at angle $\alpha$ to $B_0$. At the same time a constant gradient is applied orthogonal to $B_0$. After time $TE/2$, this applied gradient is reversed, again causing the spins to precess in the opposite direction, so that after time $TE$ there will be an echo.
information. If the applied field $B_0$ is constant, all the protons within a given area will have the same Larmor frequency. If a smaller field is applied which varies across the region to be imaged, the Larmor frequency will differ slightly from one position to another. This difference in frequency can be detected, and a Fourier transform of the MR signal can assign signal amplitude to different frequencies and hence give positional information, assuming that there are no other field inhomogeneities. This is demonstrated in figure 2.5, which shows the signal generated by two spatially-separated samples in a constant or a spatially-dependent field. The same principle can be applied to locate the signal in 2 dimensions. In 2-

![Figure 2.5: Spatial localisation of MR signal with a field gradient across the sample, courtesy of Ross Little. Two separate samples are shown, (a) in a constant field, and (b) in a varying field. (1) shows the samples, and (2) the applied gradient. (3) shows the Larmor frequencies of the samples (the difference is greatly exaggerated). (4) shows the result of a Fourier transform of the frequencies in (3), which give spatial localisation.](image)

...dimensional data acquisition, the applied RF pulse, in combination with a slice selection gradient $G_{SS}$, is tuned to excite the protons in a single slice through the image. The position of the slice is determined by the strength of $G_{SS}$, and its thickness by the frequency range of the RF pulse. A shaped RF pulse whose excitation bandwidth acts over a finite bandwidth will excite a finite slice whose thickness is a function of the ratio of the bandwidth and the slice selection gradient. A field gradient $G_{FE}$, the frequency-encoding gradient, applied in a direction orthogonal to the slice selection gradient, will allow the spatial localisation in a single further direction of the protons within the slice, in that the resonant fre-
quency of the protons will depend on their position within the slice, as is shown in figure 2.5(b). A third gradient $G_{PE}$, the phase-encoding gradient, will further modify the rate at which spins dephase and refocus according to the localisation in the third orthogonal direction. The phase-encoding gradient is typically applied at the same time as the first lobe of the frequency-encoding gradient, as is shown in figure 2.6(a). Typically the number of phase-encoding steps is 128 or 256, and the phase-encoding gradient is shown by a series of lines, while the frequency-encoding gradient is shown by a single line. The actual signal that is generated by all these gradients is a complex sum of the signals generated by each individual spin, and is stored as a function of the magnitude of the gradient-encoding and frequency-encoding fields. This complex signal is known as k-space. A 2-dimensional Fourier transform of k-space allows a reconstruction of the image. This acquisition is slow, as limited data can be acquired for each repetition of the sequence. Data for only one frequency-encoding gradient and one phase-encoding gradient can be collected for each RF pulse, after which time must be allowed for the spins to relax.

However, it is not necessary to wait for the signal induced by the original RF pulse to relax completely. It is possible to generate a train of echoes from a single RF pulse for gradient-echo, or a single pair of pulses for spin-echo. This is achieved by a series of alternating frequency-encoding gradients, and a matching series of phase-encoding pulses. This type of sequence is known as echo-planar imaging (EPI), and is shown in figure 2.7 for gradient-echo EPI.

A diagram showing the series of applied pulses over time is known as a sequence diagram. Diagrams for typical spin-echo, gradient-echo and EPI sequences, and the associated signal, are shown in figures 2.6 and 2.7.

2.3 MR diffusion

The molecules in a liquid move about randomly at a rate dependent on temperature and viscosity; the hotter and less viscous the liquid the faster the movement. This can be seen, for example, in the way ink will spread out when dropped into water. This process is known as diffusion. In a liquid the movement of any one molecule is entirely random; it can be seen when particles of a light substance (e.g.
Figure 2.6: More typical spin-echo and gradient-echo sequences. 2.6(a) shows a single spin-echo pulse sequence, with a single value of the phase-encoding gradient. A succession of such sequences is needed to build an image, each sequence with a different value of the phase-encoding gradient, so a more normal representation is shown in figures 2.6(b) for spin-echo, and 2.6(c) for gradient-echo. Here the phase-encoding gradient is represented by a series of lines.

Figure 2.7: A gradient-echo EPI sequence
pollen) are suspended in water they display a random “swarming motion” described by Brown in 1828. In 1855 Fick defined the coefficient of diffusion \( D \), which describes the rate at which one substance will diffuse into another as a function of the concentration gradient. The diffusion coefficient \( D \) can have different values in different directions, and so Fick introduced the concept of the diffusion tensor. The diffusion tensor describes a rotationally-invariant ellipse showing how freely molecules can diffuse within space; this makes it insensitive to the orientation of the subject in the scanner. It is defined by three scalar quantities, the tensor eigenvalues \( \lambda_1, \lambda_2 \) and \( \lambda_3 \), and three orthogonal tensor eigenvectors \( \epsilon_1, \epsilon_2 \) and \( \epsilon_3 \). The diffusion tensor can be visualised as describing an ellipsoid with the longest axis having length \( \lambda_1 \) oriented along \( \epsilon_1 \), as shown in figure 2.8(a). In 1917 Einstein described the bulk movement of a substance by considering the motion of individual molecules as a random walk, where the interesting factor is not the mean displacement of a molecule (which will be zero), but the root mean square displacement. In pure water at 20\(^{\circ}\)C \( D \) is 2.0 \( 10^{-3} \) mm\(^2\)/s.

The brain floats within a bath of fluid, the cerebrospinal fluid (CSF), and the molecules within the CSF also diffuse, but their movement is not random, as it is hindered by the surrounding tissue. Within the ventricles and the subarachnoid space, diffusion is random and isotropic, but within the brain tissue diffusion is hindered by the cells: The diffusion of water in biological tissues occurs inside, outside, around, and through cellular structures [Alexander et al., 2007]. Diffusion hence allows us to investigate cellular microstructure: cellular swelling and increased cellular density will force fluid to diffuse along a more twisted path, whereas necrosis will allow freer diffusion. In fibrous tissues, e.g. as found in white matter, the fluid can take a relatively unhindered path along the direction of the fibres, but movement across the fibres is hindered; this is shown in figure 2.8(b). Basser et al introduced a way of measuring the diffusion tensor using MRI [Basser et al., 1994], and calculating the diffusion tensor such that it is invariant whatever the position of the subject within the scanner.

Diffusion in the brain can be measured with MR by modifying the EPI sequence described above (figure 2.7), adding two diffusion-weighting gradients, as shown in figure 2.9. The first diffusion-weighting gradient is applied before the 180\(^{\circ}\) pulse, and will dephase spins in the imaging voxel. The second diffusion-weighting gradient is similar to the first, but because the spins have been flipped they will experience a gradient in the opposite direction. The effect of the two
Figure 2.8: Visualisation of the diffusion tensor is shown in figure 2.8(a). $\lambda_1$, $\lambda_2$ and $\lambda_3$, are the eigenvalues of the diffusion tensor, and $\epsilon_1$, $\epsilon_2$ and $\epsilon_3$ the eigenvectors. Conventionally $\lambda_1$ is the largest eigenvalue. Figure 2.8(b) represents the movement of fluid within fibres. The direction of greatest diffusivity is parallel to the fibres. The fibres are here shown in red.

Gradients will be zero for spins which have not moved. However, this cancellation of the effect of the two gradients will not be as effective for spins which have moved randomly, so some signal will be lost and the degree of signal loss will depend on how freely the spins can move. The strength of the diffusion-weighting is measured by the $b$-factor

$$b = \gamma^2 G^2 \Delta^2 (\Delta - \frac{\delta}{3})$$

where $\gamma$ is the gyromagnetic ratio, $G$ is the field strength of the diffusion-weighting gradient, $\Delta$ is the time between the two pulses, and $\delta$ is their duration.

A single repetition of the above sequence will give a measure of diffusion along the direction of the applied diffusion gradients (there is no information about which way diffusion occurs - it may be along or against the applied gradient). Changing the direction of the diffusion gradients will give a measure of diffusion in different directions. For the diffusion tensor to be calculated, at least 6 non-collinear diffusion-weighted images need to be taken. Unfortunately echo-planar images are prone to distortions caused by field inhomogeneities and eddy currents. The distortions are particularly severe in the phase-encoding direction, and near tissue-bone-air interfaces such as around the sinuses and auditory canal. Because the diffusion tensor is calculated from data acquired in several images, any substantial movement by the subject can render the images unusable.

As already described, one possible output of a diffusion-weighted image is a 3x3 tensor for each voxel within the image. These data must be simplified to be useful. Two scalar values can be extracted, which give a fair summary of the
Figure 2.9: A diffusion-weighted pulse sequence. The diffusion-weighting pulses are shown in purple, and are equally positioned around the second, 180° RF pulse. Δ is the time from the start of one diffusion-weighting gradient to the start of its matching gradient; δ is the duration of the either of the diffusion-weighting gradients, RF is the radio-frequency pulse.

data: these are the mean diffusivity (MD) and fractional anisotropy (FA). The mean diffusivity gives the average of the different eigenvalues of the diffusion tensor, and thus a measure of how hindered movement is. In a fluid MD will be large; in a substance such as jelly it will be smaller. MD is defined as

\[
MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}
\] (2.4)

MD is always a positive number, normally measured in units of \(10^{-3}\) mm\(^2\)/s; the larger the MD, the more freely fluid can diffuse.

The fractional anisotropy shows how asymmetric the diffusion ellipse is. FA is defined as

\[
FA = \sqrt{\frac{3 \times ((\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2)}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}
\] (2.5)

FA is a positive unit-less number between 0 and 1, and is a measure of how anisotropic diffusion is: the FA of a liquid is 0, and a value close to 1 indicates that diffusion in one direction is very much greater than diffusivity in any orthogonal direction. Note that although FA indicates how unevenly fluid is able to diffuse it gives no directional information, and tissues with equal FA can have very different orientations of \(\epsilon_1\).
2.4 MR perfusion

Arterial Spin Labelling (ASL) uses the MR system itself to change the magnetic properties of blood, and then track this altered magnetisation as the blood perfuses the brain. Alternative methods of measuring perfusion involve the injection of a tracer of some sort, either radioactive or paramagnetic, and are thus limited in availability, acceptability and cost. ASL is completely safe and non-invasive, and can be repeated many times on the same patient without side effects. The technique was first carried out by Detre et al in rats in 1992 [Detre et al., 1992]; he used a specially-designed coil to apply a label with a duration of several seconds.

The earliest ASL measurements were taken with a separate labelling coil applying a continuous labelling pulse with duration of a few seconds to a small volume. This is known as continuous ASL (CASL). It has two disadvantages: most MRI machines do not have the extra coil required, and thus need to be modified, and the long continuous label delivers high RF power to the subject. In 1994 Edelman described a different sequence, the EPISTAR sequence, which used the standard MR coils to apply a labelling pulse of short duration to a larger volume [Edelman et al., 1994]. This is known as Pulsed ASL (PASL). A combination of the two techniques is pseudo-continuous ASL (PCASL) [Dai et al., 2008], which applies a series of short RF pulses. These mimic the labelling of CASL while much reducing the power deposited. PCASL has been introduced since this study started; it has the potential to give a much higher SNR than PASL. The ASL sequence used in this study is the Philips implementation of Edelman’s EPISTAR sequence.

The ASL signal is inherently low, as the labelled water in the blood reaching the voxel is only a small proportion, typically 1%-2%, of the total water in the voxel. The intent is to take an image with the label and another without the label and subtract the two. In an ideal world the signal from the static water would cancel entirely between the two images, leaving only the signal from the inflowing blood. In the real world, two consecutive images will never be identical to the required accuracy, and the act of labelling will induce changes in the image because of magnetisation transfer effects. The Philips STAR sequence uses a single adiabatic pulse to tag inflowing spins in the labelling sequence, and two similar half-power pulses with little effect on the inflowing spins in the control sequence. Adiabatic pulses are specifically shaped to tag the spins to the same flip angle.
within a large range of resonant frequencies, and are thus relatively immune to inhomogeneities in the magnetic field. A pair of images is taken one after the other; in one acquisition a labelling tag is applied and in the other the control tag. In both cases there is a short delay to allow the tagged blood to travel to the imaging volume and then the image is collected, typically an EPI image because of the speed of collection. The subtraction of the two images will give a perfusion-weighted image. A single image is of little use, it is therefore common practice to take many pairs of label/control images and average them to increase the signal-to-noise. The effect of increasing the number of images is shown in figure 2.10.

CBF is defined as the volume of blood in ml which arrives in the capillary bed of 100ml of tissue in one minute.

![Figure 2.10: Building an ASL subtraction image. The top line shows labelled and control (unlabelled) images, and the subtraction image. The bottom line shows how the subtraction image improves as more images are averaged.](image)

The ASL signal will change over time, with a timescale of a few seconds; initially there will be no signal, as the labelled blood has not had time to reach the imaging voxel. The signal will build up, starting at the arterial arrival time (AAT) for that voxel. There are then several competing processes which will modify the signal:

- the rate at which blood arrives in the tissue (the cerebral blood flow or CBF)
- the rate at which blood leaves the tissue
• the rate at which the MRI signal decays (depending on the local $T_1$, $T_2$)
• the temporal width of the bolus label.

This is discussed further in section 3.1.1.

The subtraction signal described above depends on cerebral perfusion among other things, and is known as a perfusion-weighted image. However, as mentioned above the signal depends on more things than CBF. This may be adequate for some clinical needs, where all that is required is an image showing the relative values of CBF within brain regions. To quantify the CBF further work is needed, and there are two principal approaches. One is to take a a series of perfusion-weighted images after different post-labelling delays, and then model the rise and fall of the subtraction signal. This allows quantification of CBF and AAT, and is described in more detail in chapter 3. The other approach is to modify the sequence to minimise the effect of other parameters, and then take an image at a single post-labelling delay. Sequences such as QUIPSS and QUIPSSII described by Wong et al [Wong et al., 1998] apply additional saturation pulses to limit the length of the inflowing labelled bolus, and then wait to collect the image until all inflowing blood has reached the tissue. This is also the case for CASL and PCASL, where the temporal duration of the bolus is known. The subtraction signal now depends principally on $T_1$ and the rates of inflowing and outflowing blood, and hence a quantified CBF image can be collected. This sequence must be used with care; if the subtraction image is collected too soon, not all labelled blood will have reached the tissue, whereas if it is collected too late, there will be very little signal left [Alsop et al., 2015].

2.5 Image analysis

If medical images of brains are to be used in any analysis, maybe of differences between subject groups, of differences between an individual and “the norm”, or of links between specific cognitive skills (or deficits) and particular brain regions, some difficulties need to be overcome. One is that images are pictures, and analysis needs numbers; another is that brains differ from one individual to another as much as do height and body shape. I describe here the two analysis methods...
Figure 2.11: A planning image for PASL showing the labelling and readout volumes. The green hatched area shows the labelling region, and the red box the readout regions used in this study; other methods such as cortical thickness measurements, different ways of defining regions of interest, and differing statistical approaches for comparing differing patient populations, are described by Bruno et al [Bruno et al., 2012] in his review paper.

2.5.1 Normalisation

Group analyses which perform a co-localised assessment across individuals need all the images to be brought into a standard space. Unimodal images of a single individual can be analysed without further processing. If multimodal images from only one individual are to be analysed, and there is no need for comparison with any other individual, the images can be brought into a common space by aligning images one to another, with no need for transformation. This is necessary, for example, in both ASL and diffusion-weighted MRI, where a series of images is taken. However, this approach is not useful if there is a need to compare images from one individual with images from another. The approach here is to normalise the images to standard space, and most image analysis packages have software designed for this. The coordinate system used in this study is known as MNI-152, or just MNI. It was published by the Montreal Neurological Institute (MNI) [Mazziotta et al., 1995b] and adopted by the International Consortium for Brain Mapping as a standard template [Mazziotta et al., 1995a]. It is based on MRI images of 152 normal subjects. The two image analysis packages used in this study, SPM [Ashburner and Friston, 2000] and FSL [Smith et al., 2004] both provide functions
which can normalise a given image to MNI space.

Normalisation is obviously necessary for any automated voxel-based analysis of images, but there are drawbacks, particularly when studying elderly or atrophied brains as here. Regions of the brain will be transformed to fit the standard space, and so atrophied regions are likely to be stretched in a systematic way compared with non-atrophied brains. Care must be taken to examine all the normalised images visually to see if the transformation has been at least reasonable. This is discussed further in chapter 5.3.2.

2.5.2 Voxel-based analysis

Once the images have been transformed to a standard space, statistical analysis is possible. Voxel-based analysis applies standard statistical tests to each individual voxel within an image, assuming that they are all independent of each other. These tests can be of differences between groups of subjects; of differences between individuals, or of regression on such factors as age, gender and particular cognitive scores. Images need to be smoothed with a gaussian filter in order for these tests to work. Each image is likely to consist of several thousand voxels, so if the normal significance test is applied of \( p < 0.05 \) for significance, many differences are likely to be found which are due to random chance. Hence correction for multiple comparisons is necessary, or at the very least a reduction in acceptable \( p \)-value. Clusters of voxels are less likely if the distribution is purely random, so a simple approach is to set a minimum cluster size for acceptable results. There is a balance to be struck here between a stringent significance level, which guarantees that any result is probably genuine (i.e. few false positives) but risks missing a genuine result (i.e. some false negatives). Lieberman and Cunningham recommend using \( p < 0.005 \) and minimum cluster size of 10 voxels of 3.5 x 3.5 x 5 mm\(^3\) [Lieberman and Cunningham, 2009]. The advantage of including data with a less stringent statistical probability is that when published data is of more use for any future meta-analysis. Others recommend a more stringent approach, e.g. [Bennett et al., 2009], [Carp, 2012], [Nichols, 2012]. I have reported both less stringent and more stringent results.

The advantage of voxel-based analysis is that no prior knowledge is necessary; the analysis is completely independent. One disadvantage is that for the stat-
istical technique used to be valid, the images must be smoothed with a gaussian kernel. This smoothing both allows for slight variations in the normalisation, and enforces a gaussian distribution onto the voxels apart from that already assumed by the central limit theorem. The disadvantage of the smoothing is that it may lose important data, and may also introduce artefacts if the image to be smoothed is in fact very structured, as is shown in chapter 7.3.2.1.

2.5.3 Regional analysis

An alternative to voxel-based analysis is to define regions of interest (ROIs), and sum the image parameters within those ROIs. The images can be taken from an automated atlas, or drawn by hand on each image. The second technique has the advantage that it can be used on images before normalisation, this avoiding the distortions introduced by normalising. However, like the visual rating scales mentioned earlier, it needs a skilled operator, can be very time-consuming and is both subjective and difficult of transfer from site to site. If an automated atlas is used, the ROIs are easy to define, and easily transferred from site to site, but have the disadvantage that the images need to be normalised.

ROIs can be defined to cover regions known to be affected in an illness, thus utilising prior knowledge, or can be arbitrary. Using ROIs avoids some of the problems introduced by smoothing images, but there is still a problem with multiple comparisons, although it is much reduced compared with voxel-based analysis.

2.6 Imaging in FTLD

As new imaging methods have been developed they have been applied to patients with FTLD, as an aid to diagnosis and as a research tool to investigate the underlying processes causing the neural degeneration. Glucose metabolism can be measured with radioactive tracers, as described in chapter 2.6.1, and this is in regular clinical use in the UK in the differential diagnosis of dementia. A recent introduction has been a tracer which allows the imaging of amyloid, a neurochemical implicated in Alzheimer’s disease, and this is becoming accepted as a diagnostic tool for AD. A very recent introduction is the use of PET ligands which can meas-
ure neural inflammation (see chapter 2.6.1.1). PET and T\textsubscript{1}-weighted MRI are used routinely now, and diffusion and ASL imaging are gaining acceptance clinically.

### 2.6.1 PET and SPECT

Positron emission tomography (PET) and single-photon emission computerised tomography (SPECT) are similar in that in both a radioactive tracer is injected into the subject, and an image is taken after a time delay to allow the tracer to be taken up within tissue. They differ in the type of tracer used. In SPECT the tracer contains an isotope that decays by emitting a gamma ray; in PET the tracer contains an isotope that decays by emitting a positron. The positron then interacts with a nearby electron, and emits two gamma rays in opposite directions; this allows PET to give more precise localisation than SPECT because of the extra timing information. In both cases the image is collected by a gamma camera. The power of both techniques lies in the biologically-active chemicals that can be radioactively labelled (ligands); the drawbacks are:

- both involve a radiation dose to the subject, thus limiting the frequency with which they can be used.
- ligands can be both expensive and short-lived.
- unless the image is collected from a machine which can also collect CT or MRI images there is no other anatomical data.

Ligands often used in FTLD are

- HMPAO - \textsuperscript{99m}Tc-hexamethylpropyleneamine oxime, used to measure perfusion, half-life 6 hours
- FDG - \textsuperscript{18}F-fluorodeoxyglucose, used to measure glucose metabolism, half-life 110 minutes
- \textsuperscript{11}C-PK11195, a recently-developed ligand measuring inflammation, half-life 20 minutes
- \textsuperscript{18}F-THK523, binds to tau, half-life 110 minutes.
Pittsburgh compound B (PiB) is another ligand which binds to amyloid plaques, found in Alzheimer’s disease but not in FTLD, and so is often used to distinguish between AD and FTLD.

In 1991 Miller et al published results of a SPECT study on 8 patients with “frontal lobe degeneration, in whom the most common presenting symptoms were social withdrawal and behavioural disinhibition: neuropsychological testing showed selective impairment of frontal and memory tasks with relative sparing of attention, language, and visuospatial skills” [Miller et al., 1991]. From the description these patients would probably now have a diagnosis of bvFTD. They found frontal and temporal hypoperfusion with relative sparing of parietal and occipital blood flow. Two years later the same group described a group of 5 patients with progressive right hemisphere degeneration involving frontal and temporal lobes [Miller et al., 1993]. They reported psychosis, compulsions and behavioural disinhibition in these patients, and suggested that the right hemisphere may be primary for the control of social conduct. These are the first published studies using SPECT in FTLD patients, and found a similar pattern of regional hypometabolism as was known from postmortem studies. It is interesting that they suggest that the right hemisphere is responsible for social conduct, whereas now it is considered to be a function of the frontal lobes.

In 1997 Edwards-Lee et al published a study of 47 FTLD patients, although the majority of the paper concerns a group of 10 patients with anterior temporal and orbito-frontal dysfunction with other frontal regions relatively spared [Edwards-Lee et al., 1997]. They used $^{133}$Xe SPECT to provide absolute measurements of regional CBF, and HMPAO ($^{99m}$Tc-hexamethyl-propyleneamineoxime) SPECT for high-resolution qualitative images. 5 of these 10 patients had predominantly right-sided dysfunction, and 5 had predominantly left-sided dysfunction. All performed badly on neuropsychological tests (one had an MMSE score of 1/30), and yet managed the activities of daily living unaided. Those with right-sided dysfunction had behavioural disturbances, while those with left-sided dysfunction had aphasia. In at least half the group symptoms had been present for more than 10 years; in one patient who died during the study post-mortem examination showed typical findings of frontotemporal dementia: gliosis, anterior temporal neuronal loss and tissue spongiosis which in their view agreed with the SPECT findings. (N.B. This extended duration of illness contrasts with the 3.0±0.7 years reported by Kipps et al for patients with definite frontotemporal atrophy [Kipps et al., 2007]. This may indicate
the heterogeneity of FTLD, or changes in diagnosis over time.)

Ishii et al described an FDG-PET study of 21 FTLD patients, 21 AD patients and 21 controls [Ishii et al., 1998]. They report that hypometabolism in frontal and anterior temporal lobes is characteristic of FTLD, although more widespread than reported previously. Glucose metabolism in the left anterior temporal lobe is reported as $4.37 \pm 1.05$ mg/100g/min for FTLD patients, $5.43 \pm 1.18$ mg/100g/min for AD patients and $6.25 \pm 0.72$ mg/100g/min for controls. In the right anterior temporal lobe the glucose metabolism is reported as $4.49 \pm 0.78$ mg/100g/min for FTLD patients, $5.43 \pm 1.27$ mg/100g/min for AD patients and $6.37 \pm 0.68$ mg/100g/min for controls. The pattern of hypometabolism can be used to distinguish between the groups of AD, FTLD and controls.

There followed a succession of papers all of which used FDG-PET to study patients with FTLD. N.B. This is not an exhaustive list.

Salmon et al used voxel-based analysis to study historical FDG-PET images from 3 centres [Salmon et al., 2003], using a voxel-based statistical analysis. They had no associated MRI images, so normalisation must have been carried out on the PET images. Their study included 29 “FTD” patients - no subclassification was given - and they found hypometabolism in the ventromedial frontopolar cortex. In 2006 the same group described a regional analysis of FDG-PET images from 5 centres, with associated MR images [Salmon et al., 2006]. Their study included 70 “frontal-variant FTD” patients, and their analysis found three large clusters of abnormality, the first metabolic cluster included most of the lateral and medial prefrontal cortex, bilaterally and the second two clusters comprised the subcallosal medial frontal region, the temporal pole, medial temporal structures and the striatum, separately in the left and in the right hemisphere.

Foster et al suggested that the distinctive pattern of hypometabolism seen in FDG-PET in FTLD patients might be of use as a biomarker to distinguish FTLD from AD or psychiatric disturbances [Foster, 2003]. A subsequent paper in 2007 compared diagnostic accuracy for AD/FTLD and concluded that FDG-PET added both accuracy and confidence to a diagnosis of FTLD compared with clinical data alone but that it was of less utility for AD [Foster et al., 2007].

Diehl et al reported an FDG-PET study of 25 mild bvFTD patients, 9 mild SD patients and 15 controls, using voxel-based analysis within SPM99 [Diehl et al.,
2004]. bvFTD patients showed symmetrical hypometabolism in frontal lobes; SD patients showed hypometabolism in the whole of the left temporal lobe and in the right temporal pole. They took particular care to include only patients with very mild disease, and with a clear distinction between the groups, and found very different patterns of hypometabolism in the two groups. A follow-up study on 22 of the same patients 19.5±7.5 months later found that hypometabolism had spread to parietal and temporal cortices. [Diehl-Schmid et al., 2007].

Kanda et al compared hypometabolism measured by FGD-PET with grey matter atrophy measured by MRI $T_1$-weighted images in AD and FTLD patients, and found decreased grey matter volume and decreased glucose metabolism in the frontal lobe and anterior temporal lobe in FTLD patients [Kanda et al., 2008]. They also noted that the MRI analysis showed asymmetry which was not seen in the FDG-PET results.

Herholz et al reviewed the use of PET imaging in dementia, and describe the routine use of FDG-PET in dementia research, with characteristic patterns of glucose hypometabolism being used for differential diagnosis [Herholz et al., 2007].

Evers et al used hypometabolism on FDG-PET as diagnostic of FTLD, and reported variable degrees of insight in these patients - her stress was that “loss of insight” is not necessarily always present in FTLD, and should therefore not be part of the core criteria for such a diagnosis. [Evers et al., 2007].

All the PET and SPECT studies report a pattern of hypometabolism in the frontal and anterior temporal lobes that is predictive of FTLD, and differentiates it from AD. They also show the heterogeneity of this group of illnesses, and that while correlated, clinical symptoms can be present without visible hypometabolism.

### 2.6.1.1 Other ligands

Cagnin et al in 2004 used $^{11}$C-PK11195 as a ligand in PET imaging of microglial activation in FTLD: they report that their findings imply the presence of an active glial response reflecting progressive neuronal degeneration [Cagnin et al., 2004]. Varley et al described available PET ligands for imaging neuroinflammation, but also described an MR contrast medium, $1^{-13}$C-acetate, which allows MRS measurement.
of microglial metabolic rate [Varley et al., 2014]. They report increased glial metabolic rates in patients with AD and mild cognitive impairment (thought to be a precursor to AD), but as yet have reported no results on patients with FTLD. This work suggests that imaging of neuroinflammation could be an interesting future avenue for characterisation and understanding of the disease processes in FTLD.

2.6.2 T1-weighted MRI

High-resolution T1-weighted MRI images are now in routine clinical use to aid diagnosis of FTLD. The images can be used to exclude such diagnoses as cancer and cerebrovascular disease, and also the existence of atrophy will indicate a neurodegenerative illness, and the pattern of atrophy can aid in differentiating FTLD from other dementias. The first description of its potential value was in 1993, when Filley and Cullum report a single case of “putative FTD”, where neuropsychological evaluation disclosed evidence of extensive frontal system dysfunction, with lesser problems in memory, language, and visuospatial skills. MRI images were largely nonspecific in the early stage of the illness (17 months after onset), but 16 months later revealed bi-frontal and bi-temporal atrophy with ventricular enlargement [Filley and Cullum, 1993]. They suggest that neuroimaging may be less sensitive in the early stages of the illness compared with a later stage.

Kitagaki et al compared MRI images of 18 FTLD patients, 18 AD patients and 18 controls, and again found a pattern of frontal and anterior temporal atrophy in FTLD compared with AD.

Hundreds of other papers follow these, and MRI was soon in regular clinical use. These published studies differ in their aim; most seek to distinguish between AD patients, FTLD patients and controls, though others look for a link between damage to specific brain regions and specific cognitive deficits. More recently there has been research to identify biomarkers for disease progression, or to identify pathological phenotypes of FTLD rather than clinical phenotypes. Two early papers are described below, and also four recent review papers.

Mummery et al described a study involving 6 SD patients and 14 controls [Mummery et al., 2000]. They used voxel-based analysis, and found the left temporal pole was the most consistently affected. Damage correlated with semantic
memory impairment.

Rosen et al used voxel-based analysis to look at patterns of atrophy in 20 controls, 8 bvFTD patients and 12 SD patients, all with mild-moderate dementia [Rosen et al., 2002]. They reported that the SD patients showed similar patterns of behavioural abnormalities to the bvFTD patients, but also had impairment on tests of language. They found regional atrophy in both groups in bilateral frontal regions, and the left anterior cingulate cortex. The FTLD, but not the SD, group showed atrophy in the right dorsolateral frontal cortex and the left premotor cortex. The SD, but not the FTLD, group showed tissue loss in the anterior temporal cortex and the amygdala/anterior hippocampal region.

Seelaar et al in their review paper looked at the clinical, genetic and pathological heterogeneity of FTLD [Seelaar et al., 2011]. They report that mild bvFTD patients show atrophy in the anterior cingulate cortex and frontal insula as well as medial frontal and orbitofrontal cortices, hippocampus, striatum and thalamus, and that the right hemisphere is more severely affected than the left. With increasing disease severity, there is more diffuse atrophy, although in similar areas, with involvement of more lateral frontal areas and subsequently more posterior temporal and anterior parietal atrophy, although this may depend on the particular pathology. SD patients have damage to the anterior and inferior temporal lobes, usually left but occasionally right, asymmetric and spreading with disease severity.

Rohrer in his review of $\mathrm{T}_1$-weighted brain imaging in frontotemporal dementia describes the clearly differing patterns of damage found in different patient populations [Rohrer, 2011]. However, he warns that while FTLD is usually clinically distinct from typical AD, some FTLD patients have impairment of episodic memory more typical of AD, and some AD patients have symptoms more typical of SD or bvFTD. When comparing imaging studies of FTLD and AD it is therefore important to understand whether the groups being studied are clinical or pathological phenotypes. He also wonders whether particular imaging features described in the different pathologies (or clinical syndromes) at a group level can usefully translate into a way of diagnosing patients on a single case basis.

Whitwell and Josephs review recent advances in the imaging of FTLD [Whitwell and Josephs, 2012]. They report that neuroimaging (including PET) is an invaluable tool for characterising the many different aspects of FTLD, but also helps
distinguish genetic and pathological groups and has the potential to be important biomarkers for disease diagnosis and studies of disease progression.

Degnan and Levy review neuroimaging in the rapidly progressive dementias, among which they include FTLD [Degnan and Levy, 2013]. When discussing “FTD” they also note a pattern of atrophy in the frontal and temporal lobes, but warn that in the early stages of the illness as many as half of cases atrophy may be missed, and imaging is used with the primary purpose of excluding other causes for the symptoms. However, they do not mention SD as a separate condition, and it is not clear whether or not their definition of “FTD” includes SD, or is just bvFTD.

Hu et al - review the progress and challenges of biomarkers for FTLD [Hu et al., 2011], including clinical phenotype/feature characterisation, neuropsychological analysis, CSF analysis, and patterns of brain atrophy detectable on brain imaging. Their final conclusion is that more study is needed on both the accuracy and pathological significance of all the biomarkers they examined, probably in a multi-site study to achieve the necessary numbers. They also warned that neither clinical nor structural imaging-based biomarkers have been accurately correlated with underlying pathology on the individual patient level.

In summary, it is clear once again from these papers that there are many different ways to classify FTLD patients, and patients with a similar diagnosis may have marked differences in both cognitive skills and levels of atrophy at different stages of the illness. Nevertheless there is a clear picture in the later stages of the illness of atrophy in the frontal and temporal lobes, and the pattern of atrophy provides a clear differentiation between FTLD and AD. Atrophy predominantly in the frontal lobes is more linked to behavioural presentations, and atrophy in the left anterior temporal lobe to semantic deficits.

2.6.3 Diffusion MRI

There are as yet far fewer studies of FTLD using diffusion imaging in comparison to basic structural MRI. One of the first studies by Larsson et al was of diffusion tensor imaging of the brain of a frontotemporal dementia patient which had been preserved in formalin [Larsson et al., 2004]. They reported reduced anisotropy in
the frontal lobes compared with other brain regions. Later Yoshiura et al reported on a study comparing 13 FTLD patients to 15 controls, and reported abnormal MD elevation was seen predominantly in the frontal and temporal lobes in FTLD patients using visual rating scales [Yoshiura et al., 2006]. These papers were followed by several more comparing different groups of patients and using different diffusion metrics: they are described more completely in chapter 7.4. They agree that diffusion metrics show greater differences than grey matter atrophy between FTLD patients and healthy controls, and that the pattern of differences can be used to differentiate between FTLD and AD patients.

2.6.4 Perfusion MRI

Studies with PET and SPECT showed hypometabolism in frontal and temporal lobes in FTLD patients (see chapter 2.6.1), and Alsop et al used ASL to show hypoperfusion in Alzheimer’s disease [Alsop et al., 2000]. It was therefore a reasonable assumption that ASL might show hypoperfusion in FTLD, and Du et al published a study that did indeed show that ASL could detect such hypoperfusion [Du et al., 2006]. Their study included 21 patients with “FTD”, 24 patients with AD, and 25 controls, and they used a PASL single timepoint sequence which only covered the superior regions of the brain. They performed a voxel-based analysis with partial volume correction (see chapter 6.2.2), and found hypoperfusion in superior frontal regions in FTLD patients compared with controls.

Hu et al studied a group of 42 FTLD patients, 18 AD patients and 23 controls [Hu et al., 2010]. Their grouping is somewhat perplexing, as the FTLD group consisted of 25 bvFTD, 15 PPA and 2 CBS by clinical diagnosis, while the AD group consisted of 1 bvFTD, 11 PPA and 5 CBS. Their explanation was that their AD group included those subjects with known or predicted AD pathology, whatever the clinical diagnosis. They used a single timepoint CASL sequence, and found significantly reduced CBF in FTLD patients compared with healthy controls in the bilateral dorsolateral prefrontal cortex and the right inferior frontal cortex. They also observed areas of increased CBF in FTLD patients in the medial parietal cortex/precuneus and the posterior cingulate cortex. It is possible that the increased CBF is an artefact of the partial volume correction, as I discuss in chapter 7.3.2.1, as no other group has reported finding increased CBF in any region.
Shimizu et al studied the relationship between brain atrophy measured by T₁-weighted MRI and perfusion measured by ASL [Shimizu et al., 2010]. Their study included 28 “FTD” patients and 29 controls. Again, no further information is given about the patients except that patients with MND were excluded. They used a single timepoint PASL sequence, and found GM atrophy, but not hypoperfusion, in the premotor cortex in FTLD together with GM atrophy and hypoperfusion in the right prefrontal cortex and bilateral medial frontal lobe.

Tosun et al studied a group of 12 bvFTD patients and 12 controls [Tosun et al., 2012]. They used a single timepoint CASL sequence and found that CBF was reduced in patients in the bilateral frontal and temporal cortices as well as in the thalamic nuclei, putamen, caudate, and hippocampal subcortical regions.

Zhang et al published a study combining ASL, T₁-weighted and diffusion images of a group of 20 AD patients, 20 bvFTD patients, and 21 healthy controls [Zhang et al., 2011]. They used a single timepoint cASL sequence, and partial volume correction. They found reduced CBF in bvFTD patients in bilateral frontal lobes, with the most pronounced hypoperfusion in the right inferior frontal gyrus.

### 2.7 Power calculations

A power calculation is required in order to determine the likely size of the varying metrics that can be detected in this study. There are published studies giving the expected variance of CBF, AAT, FA and MD, and these can be used to estimate the number of subjects in the patient group that are required for a given level of difference in the parameter. I have used the sample size calculation `sampsizepwr` in Matlab2009, to calculate this sample size using a 5% significance level as the level at which a positive result is accepted, and 80% as the detection power of the study to reject a negative result. Since it is expected that CBF and FA will be lower in patients, and MD higher, for these parameters I have used a one-tailed test. For AAT I have used a two-tailed test as there is no prior expectation on the value of AAT in this patient group.

The values of whole-brain CBF and AAT in a healthy group have been measured by Thade et al [Thade et al., 2010]: CBF is 47.4±7.5 ml/100ml/min: AAT as 820±120ms. The values of FA and MD have been measured by Heiervang et al...
[Heiervang et al., 2006]: they found whole-brain mean FA is 0.389±0.139; mean MD is 7.18±0.94x10\(^{-4}\) mm\(^2\)/s. Figure 2.12 shows the number of subjects needed to detect a given percentage change in each parameter, assuming that the population means are applicable to the more elderly group on this study. It can be seen that a patient group of around 17 should be sufficient to detect a 10% drop in CBF; for AAT the group size to detect a 10% change in either direction is around 19; to detect a 10% increase in MD the required group size is around 13 and for a 10% decrease in FA the group size is 81. The larger group size necessary to detect a 10% decrease in FA is because of the larger inter-individual variation in FA. All of these group sizes are indicative only, as the studies from which I have taken the data are of young healthy adults, and I have made no correction for any multiple comparisons, e.g. for looking at regional variations in the multiple regions known to be affected in FTLD. As the purpose of this study is to generate hypotheses for further test, I believe this approach is justified.

Reported differences in CBF and metabolism in FTD are mainly > 10%. Ishii et al [Ishii et al., 1998] report regional glucose differences in the left anterior temporal lobe of 13%-25% between patients with FTLD, patients with AD and controls. Lee et al [Lee et al., 2009] report differences of 30% between people with mild cognitive impairment and controls. Smaller CBF differences have been reported in different populations; e.g. Parkes [Parkes, 2002] reported a percentage decrease of CBF with age of 0.5% per year. I accept that smaller changes in CBF may be clinically relevant and the sample size in my study was not large enough to have the power to detect this. This could lead to false negative findings.

2.8 Summary

Previous studies have increasingly shown that FTLD is a heterogeneous group of illnesses, but with some clear neuroimaging findings. In the early stages of the illness MRI and PET images can be considered “normal” for many patients; any abnormality that does exist may be too subtle for existing technology to find. However, as Kipps et al have pointed out with their description of the clinical phenotype [Kipps et al., 2009] a group of patients exist who have clinical symptoms typical of bvFTD, but whose symptoms do not worsen, and who seem to have no evidence of a neurodegenerative process. In the later stages, atrophy is evid-
Figure 2.12: Power calculations for sample sizes to detect between 10% and 50% change in CBF, AAT, MD and FA
ent on T₁-weighted MR images, and hypometabolism on FDG-PET images in the frontal and temporal lobes, and this feature can be used to differentiate between patients with FTLD and patients with AD. ASL images have shown hypoperfusion in frontal and temporal lobes, but coverage of ventral brain regions is very limited. No published studies as yet have looked at the results of a multi-timepoint ASL sequence, so there is no information on whether arterial arrival time is affected in the illness. Diffusion studies are becoming increasingly common, and all suggest that diffusion may be more sensitive than atrophy as an indicator of a neurodegenerative process.

The questions this study addresses are:

1. what precision and accuracy can the PASL sequence used in this study achieve in a clinically acceptable time
2. how many of the parameters the affect the perfusion-weighted MR signal can be estimated in a clinically acceptable time
3. which of the commonly-used 2-dimensional EPI MR readouts give increased signal in the ventral brain regions initially affected in FTLD
4. what degree and pattern of atrophy is seen in patients with FTLD, compared with controls.
5. do the different degrees of atrophy seen in FTLD patients relate to their scores in cognitive tests, in particular the tests of social awareness developed by Zahn et al [Zahn et al., 2009a], [Green, 2011].
6. what degree and pattern of cerebral blood flow and arterial arrival time differences are seen in patients with FTLD, compared with controls.
7. if seen, do the different perfusion parameters in FTLD patients relate to their scores in cognitive tests, in particular the tests of social awareness developed by Zahn et al
8. what degree and pattern changes in the diffusion metrics of fractional anisotropy and mean diffusivity are seen in patients with FTLD, compared with controls.
9. If seen, do the different diffusion metrics in FTLD patients relate to their scores in cognitive tests, in particular the tests of social awareness developed by Zahn et al.

10. What is the optimal combination of MR modalities and regional differences to distinguish between patients with FTLD and controls.

Question 1 is addressed in chapter 3; question 2 in chapter 4; questions 3 and 4 in chapter 5, questions 5 and 6 in chapter 6, questions 7 and 8 in chapter 7 and questions 9 in chapter 8. Chapter 9 summarises the work submitted in this thesis, and considers its future development.
Chapter 3

Mathematical modelling of the ASL signal

In this chapter I investigate the theory and modelling needed to transform the perfusion-weighted image that can be measured with MRI into a quantified measurement of cerebral blood flow. I consider the minimal set of necessary parameters and the accuracy that can be achieved in the quantification. While the models themselves are not new, the extensive use of simulation to determine the effect on the estimate of perfusion that errors in the assumed values of the other parameters involved in the model is novel. In this chapter I also describe the analysis of an ASL dataset with 23 timepoints.

It must be noted that the whole of this chapter applies principally to a PASL sequence similar to that described in chapter 6.2.1. PCASL sequences can have a higher SNR.

3.1 Quantifying the pulsed ASL signal

3.1.1 Modelling the ASL signal

The perfusion-weighted image can be transformed to a quantitative CBF value by modelling the inflow and outflow of blood to the imaging volume. The evolution
of the subtraction signal within an idealised imaging voxel is shown in figure 3.1. Initially there is no labelled blood within the voxel; the label is applied to a bolus outside the brain. While the physical dimension of the bolus is known, the time \( \tau \) which is the time the bolus will take to pass a given point, is known only approximately, as it depends on the physiology of the subject. The passage of the bolus through the voxel is shown in figures 3.1(a), 3.1(b) and figures 3.1(c), and the value of the subtraction signal in figure 3.1(d) according to the single-blood-compartment model [Parkes and Tofts, 2002].

**3.1.1.1 The Buxton model**

In 1998 Buxton et al published a model for the evolution of the ASL subtraction signal over time [Buxton et al., 1998]. The Buxton model gives the time-course of the subtraction signal for PASL in any given voxel as

\[ \Delta M = 0 \quad \text{for} \quad t < t_a \]  
\[ \Delta M = 2f m_0^a \alpha (t - t_a) e^{-\frac{t}{T_{1b}}} q_p(t) \quad \text{for} \quad t_a < t < t_a + \tau \]  
\[ \Delta M = 2f m_0^a \alpha \tau e^{-\frac{\tau}{T_{1b}}} q_p(t) \quad \text{for} \quad t > t_a + \tau \]

where

\[ q_p(t) = \frac{e^{kt}(e^{-kt_a}) - e^{-kt}}{k(t - t_a)} \quad \text{for} \quad t_a < t < t_a \]  
\[ q_p(t) = \frac{e^{kt}(e^{-kt_a}) - e^{-k(\tau + t_a)}}{k(\tau)} \quad \text{for} \quad t > t_a + \tau \]

\[ k = \frac{1}{T_{1b}} - \frac{1}{T_{1}'} \]

\[ \frac{1}{T_{1}'} = \frac{1}{T_{1}} + \frac{f}{\lambda} \]

and

- \( \Delta M \) is the signal difference between label and control images at time \( t \) (arbitrary units) (voxelwise parameter);
- \( m_0^a \) is the equilibrium magnetisation of blood (arbitrary units) (can be a voxelwise or a global parameter);
- \( \alpha \) is the labelling efficiency, i.e. the fraction of the inflowing blood that is
Figure 3.1: Passage of the labelled bolus through the voxel. A-A shows part of the capillary bed with feeding artery and draining vein. B-B shows an artery passing through the voxel, and C-C shows a vein. In figure 3.1(a) the labelled blood (green) is just arriving at the voxel in both the capillary bed and the arterial system. In figure 3.1(b) the blood in the arteries has cleared the voxel, and part of the bolus has entered the capillary bed. In figure 3.1(c) the entire bolus is with the capillary bed. Figure 3.1(d) shows the evolution of the subtraction signal within the voxel according to the single-blood-compartment model [Parkes et al., 2004]. In this last figure, the line shown in blue and marked (a) covers the time before the bolus has arrived at the voxel. The line shown in green and marked (b) covers the time from the arrival of the start of the bolus to the time when half the bolus has arrived. The line shown in red and marked (c) covers the time from the end of (b) until the time when all the bolus has reached the voxel. The line shown in magenta and marked (d) covers the time when all the bolus has arrived in the voxel, and the signal is now decaying as the labelled spins relax.
magnetically labelled (no units) (global parameter);

- $\lambda$ is the blood-brain partition coefficient, i.e. the relative water content of brain and blood. The units of $\lambda$ define the units of CBF; if $\lambda$ is given in ml/g, CBF will be in ml/100g/min; if $\lambda$ is given in ml/ml, CBF will be in ml/100ml/min. (global parameter)

- $\tau$ is the bolus width (ms), i.e. the time taken by the bolus to pass a defined point (voxelwise or global parameter);

- $f$ is the tissue CBF (for units see $\lambda$) (voxelwise parameter);

- $t_a$ is the arrival time of the leading edge of the bolus (ms) (voxelwise or global parameter);

- $T_{1b}$ is the relaxation time of blood (ms) (global parameter);

- $T_{1t}$ is the relaxation time of tissue (ms) (voxelwise or global parameter).

The assumptions made in this model are that

1. the labelled blood arrives at the voxel as a square pulse, with no dispersion from label to voxel;

2. the labelled blood exchanges instantaneously with the tissue as soon as it arrives in the voxel;

3. as a result of [2], by the time blood leaves the voxel, its magnetisation is equal to the average magnetisation of the whole voxel, weighted by $\lambda$, the equilibrium ratio of water in brain compared to blood;

4. the labelled blood is cleared from the tissue at the same rate $f$ as the inflowing blood.

The term $q_p$ includes all factors relating to venous clearance of blood from the tissue, and the differing relaxation rates of blood and tissue.

### 3.1.1.2 The single-blood-compartment model

Parkes et al [Parkes, 2002] suggested a simplification to this model by ignoring venous clearance, and assuming that the labelled inflowing water remains within
the blood rather than moving instantaneously to tissue, thus setting $q_p$ to 1. This assumption simplifies the equations giving the subtraction signal as

\[
\Delta M = 0 \quad \text{for} \quad t < t_a
\] (3.8)

\[
\Delta M = 2f m_0^a \alpha (t - t_a) e^{\left(-\frac{t}{\tau_1}\right)} \quad \text{for} \quad t_a < t < t_a + \tau
\] (3.9)

\[
\Delta M = 2f m_0^a \alpha \tau e^{\left(-\frac{t}{\tau_{intermediate}}\right)} \quad \text{for} \quad t > t_a + \tau
\] (3.10)

The assumptions made in this model are that

- the labelled blood arrives at the voxel as a square pulse, with no dispersion from label to voxel;
- the labelled water remains in the blood within the vasculature, and does not perfuse into the tissue within the time considered by the model, typically less than 3s.

### 3.1.1.3 A suggested intermediate model

The ASL signal is inherently very noisy, and hence does not lend itself well to complex models. The assumptions of both the Buxton and single-blood-compartment model are admitted to be incorrect, but attempt to minimise the number of parameters to be estimated. No models will accurately reflect reality, but these two models have been found to be useful; An intermediate possibility is to allow the relaxation time of the labelled blood to vary between that of blood, and that of tissue; admitting that an unknown fraction of the labelled blood will move to the tissue within the voxel. The outflowing labelled blood is ignored in this intermediate model, which gives equations for the subtraction signal similar to those of the single-blood-compartment model, except that the relaxation time is included as another parameter to be estimated. Hence the equations for the intermediate model are

\[
\Delta M = 0 \quad \text{for} \quad t < t_a
\] (3.11)

\[
\Delta M = 2f m_0^a \alpha (t - t_a) e^{\left(-\frac{t}{\tau_{intermediate}}\right)} \quad \text{for} \quad t_a < t < t_a + \tau
\] (3.12)

\[
\Delta M = 2f m_0^a \alpha \tau e^{\left(-\frac{t}{\tau_{intermediate}}\right)} \quad \text{for} \quad t > t_a + \tau
\] (3.13)
3.1.2 Influence of model parameters on ASL subtraction signal

There are 6 parameters in these equations: 4 that may vary from voxel to voxel, and 2 that are treated within this model as global. The global parameters are $\alpha$, the labelling efficiency, and $m_0^a$, the equilibrium magnetisation of blood. Because they are global parameters they are not considered further here as their effect is only to scale the entire time course up or down; they are considered further in chapter 4.2.4. The effect on the subtraction signal of changing each of the remaining 4 parameters individually is shown in figure 3.2. I have considered variations in the parameters similar to the natural variation in a healthy population. It is clear that the parameter with the biggest effect on the subtraction signal is CBF. Bolus width and assumed relaxation time have a much smaller effect, and in a signal with low SNR they are unlikely to be well-estimated. When calculating the expected subtraction signal from this model, the arrival time has no effect on the subtraction signal once the time at which the signal is measured exceeds (arrival time + bolus width), as is shown in equation 3.13. This is the reason that a single timepoint measurement of CBF is possible; as long as the single measurement is taken beyond the time at which all labelled blood has arrived within the voxel, the measurement is independent of arrival time. However, arrival time is not generally known; if the single measurement is taken too early it will be inaccurate; if too late there will be little signal left to sample. However, when calculating CBF from multi-timepoint data, if only CBF is estimated and all other parameters are assumed, the assumed value of the arrival time can have a marked effect on the estimated CBF, as is shown in figure 3.3.

3.2 Simulation methods

The intermediate model described above was used to generate simulated data with varying levels of noise, and with different values for the four variable parameters. The simulated data were then used as input to fitting routines, and the estimated parameters from the fits compared with the true values from the simulations. The error values from the fitting routines were also used to compare the goodness of fit among the varying fitting routines.

All calculation was done with Matlab 2009a. Values of CBF, arrival time,
Figure 3.2: Influence of model parameters on subtraction signal. The graphs show the effect on the simulated subtraction signal of varying one parameter while the others are held constant. For all graphs, where simulated values are not stated they are CBF 50 ml/100ml/min, arrival time 375 ms, bolus width 900 ms and relaxation time 1350 ms. In the plot for relaxation time the dotted line shows the signal within the label before it enters the voxel, showing the decay in signal from the time of labelling.
bolus width and relaxation time of blood were chosen and used to generate values of the subtraction signal at specified timepoints, using the intermediate model described above. Initial parameter values are described in table 3.1. Varying levels of Gaussian random noise were added using the Matlab function "randn". Noise levels were selected to give an SNR between 2 and 20; the SNR for the ASL data used in this study is calculated to be around 15 (see chapter 3.4.3). The SNR for a given noise level was calculated at a timepoint of 1200ms (close to the maximum of the signal). The mean and standard deviation of the subtraction signal were extracted, and SNR calculated from

$$SNR = \frac{\text{mean}(\text{subtraction signal})}{\text{standard deviation}}$$

(3.14)

Simulated data were calculated for 4, 8 and 16 timepoints, at times between 500 and 3000ms; exact times are shown in table 3.2. The numbers of timepoints were chosen to reflect realistic clinical scanning sequences; the FTD study described elsewhere in this thesis used 4 timepoints. The 8 timepoints dataset was included to investigate the advantages it may offer over either of the others. 100 data sets were simulated for each timepoint/noise level/parameter set. A normalised signal was calculated, so neither $m_0^a$ nor alpha were used. $m_0^a$ and alpha are global

Figure 3.3: Effect on estimated CBF of changing the assumed value of arrival time. The crosses show data simulated from the single-blood-compartment model with a CBF value of 55ml/100ml/min, arrival time of 550ms, bolus width of 900ms, relaxation time of 1350ms, and SNR of 14. CBF is calculated from this data with a 1-parameter fit with varying assumed values of arrival time, an assumed value for the bolus width of 1000ms, and an assumed value for the relaxation time of 1500ms. It can be seen that the calculated value of CBF depends strongly on the assumed value of arrival time.
values, and so were omitted from the calculation; hence the simulated signal is a fraction of the total available.

A CBF value of 50 ml/100ml/min [Parkes et al., 2004] and arrival time of 750 ms [MacIntosh et al., 2010] were chosen as typical for healthy individuals; a bolus width of 1000ms was chosen as typical of the sequence used elsewhere in this study (see section 3.4.2), and relaxation times were chosen between the relaxation time of grey matter (1470±50ms [Ethofer et al., 2003]) and the relaxation time of blood (1550ms [Greenman et al., 2003]). Values for CBF, arrival time and bolus width were chosen as ± 25% and ± 10% of the assumed value. The simulated data were used to estimate 1, 2 or 3 parameters using the Matlab function “fminsearchbnd”, and the value of the calculated parameter was compared with the input parameter. Where parameters were not calculated, an assumed value was used which is given in table 3.1. No attempt was made to fit all 4 parameters; this is obviously not possible with only 4 timepoints, and so was not used in simulations with more timepoints. The process is shown in figure 3.4, where the original model is shown as the black line and the data points generated with an SNR of approximately 14 are shown as black crosses. The 1-parameter fit for CBF is shown as the green line; the 2-parameter fit for fitting CBF and arrival time as the magenta line; the 3-parameter fit for CBF, arrival time and bolus width as the cyan line, and the 3-parameter fit for CBF, arrival time and relaxation time as the red line.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Assumed value</th>
<th>Value(s) for simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBF ( f ) (ml/100ml/min)</td>
<td>-</td>
<td>37.5 45 55 62.5</td>
</tr>
<tr>
<td>arrival time ( t_A ) (ms)</td>
<td>500</td>
<td>375 450 550 625</td>
</tr>
<tr>
<td>bolus width ( \tau ) (ms)</td>
<td>1000</td>
<td>900 950 1050 1100</td>
</tr>
<tr>
<td>relaxation time ( T_1 ) (ms)</td>
<td>1500</td>
<td>1350 1425 1575 1650</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3.2: Delay times used in simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Points</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>16</td>
</tr>
</tbody>
</table>

78
Figure 3.4: Initial model, simulated data and various fits. The black line shows the subtraction signal generated by the intermediate model, and the black crosses simulated data with noise added with an SNR of 14. The 1-parameter fit for CBF is shown as the green line; the 2-parameter fit for fitting CBF and arrival time as the magenta line; the 3-parameter fit for CBF, arrival time and bolus width as the cyan line, and the 3-parameter fit for CBF, arrival time and relaxation time as the red line.

3.2.1 Comparing results of simulated models - precision and accuracy

There are two ways of comparing models such as these, where the ground truth is known. One way is to compare the true value of a parameter with the estimated value from the model. Where multiple parameters are being estimated, different models may give more accurate estimates for different parameters, and it is important to decide which are the principal parameters of interest. In all cases in this chapter, the principal parameter of interest is CBF, followed by arrival time.

Another, and different way of comparing models is to look at the precision of the estimates, by comparing the sum of residuals of each model. Normally the model with the lowest value is to be preferred, allowing for the number of parameters estimated. There may be a conflict between the two ways of comparing models, especially if some parameters are assumed rather than being fitted. Here I am fitting data which has been simulated using the same model with assumed values for all parameters. If I only fit 1 parameter, but use the same assumed values...
for other parameters as were used in the simulation, the error on the assumed parameters is always zero, and hence the total error for a 1-parameter fit is likely to be less than for a fit on more parameters. Hence the $f$-test which shows the model with the lowest sum of residuals, will show the most precise model, but not necessarily the most accurate. As is shown in figure 3.3, for given simulated data, the value of CBF estimated by a 1-parameter model depends on the values assumed for other parameters, particularly arrival time.

Graphs were plotted to show how far the estimated parameter was from the true value, and the effect on this of errors in other assumed parameters.

The error on the fit was calculated from the sum of squared errors returned by “fminsearchbnd”, and an $f$-test was used to compare how well the various models fitted the data (a Goodness of fit test) given the assumptions made. The $f$-test is used to compare the sum of residuals of two nested models with differing numbers of parameters, and returns the probability that one model is better than the other allowing for the differing number of parameters. A p-value of 0.05 was used as the cut-off for which model fits the data better. The models compared were

- a 1-parameter fit for CBF, assuming values of arrival time, bolus width and relaxation time versus a 2-parameter fit for CBF and arrival time, with the same assumed values for bolus width and relaxation time
- a 2-parameter fit as described above versus a 3-parameter fit for CBF, arrival time and bolus width, assuming the same values for relaxation time
- a 2-parameter fit as described above versus a 3-parameter fit for CBF, arrival time and relaxation time, assuming the same values for bolus width.
3.3 Results of fits on simulated data

3.3.1 Accuracy of estimates

3.3.1.1 1-parameter fit

Figure 3.5 shows the estimated CBF for the 4 assumed CBF values for 4, 8 and 16 timepoints at various levels of SNR, with error bars showing the standard deviation of the estimate. The true value used in the simulation is also shown. It can be seen that for all values of CBF, all SNR and all timepoints, the estimated CBF is approximately constant, that the SD decreases as the SNR increases, and that the estimated value is different from the true value. This difference between estimated and true values is because of the errors in the assumed values of the other parameters. Although the error in the estimated CBF value depends on the simulated true value, the percentage error is approximately independent of the SNR for SNR values >10 and given values of the other parameters, as is shown in figure 3.5. This allows simplification of the values for estimated vs true values of parameters, in that the SNR value can be omitted. Figure 3.6 shows the percentage CBF error in the single-parameter fit; there are three figures for the three differing number of timepoints. This figure gives some indication of the effect that each parameter alone has on the accuracy of the estimate of CBF, and the effect of the number of timepoints. The three graphs at top left in figures 3.6(a), 3.6(b) and 3.6(c) are all very similar, as are the graphs at top right, bottom left and bottom right in the three figures; the number of timepoints has little effect on the accuracy of the estimate of CBF. As \(T1_b\) is varied the error in CBF is up to 30%. Errors in arrival time can result in up to 20% error in CBF (compare the 4 graphs in figure 3.6(a)) and errors in the bolus width have a smaller effect of 10% error in CBF (compare the 4 coloured curves in any individual graph in figure 3.6(a)). So errors in the bolus width will have a lesser effect on the accuracy of the estimate of CBF than errors in arrival time, and errors in \(T1_b\) have the greatest effect.

3.3.1.2 2-parameter fit

Figure 3.7 shows percentage errors in CBF for the 2-parameter fit for differing relaxation times and bolus widths, averaged over all values of simulated CBF and
Figure 3.5: Estimated CBF from 1-parameter with error bars. Each graph shows the percentage error in estimated CBF plotted against SNR for a given number of timepoints and a given simulated value of CBF, with bars showing the standard deviation. The numbers underneath each graph show the number of timepoints, and the CBF in ml/100ml/min used in the simulation. All graphs are on the same axes. It can be seen that the value of estimated CBF is close to the true value (i.e. that used to simulate the data.)
Figure 3.6: Percentage error in CBF for 1-parameter fit. The three sets of four graphs show the percentage error in CBF for 4, 8 and 16 timepoints, for 4 assumed values of arrival time (the individual graphs), 4 assumed values of bolus width (points along the x-axis) and 4 assumed values of relaxation time (coloured lines). The number of timepoints does not have much effect.
simulated arrival time. Errors in $T_1b$ of $\pm 10\%$ can cause errors in CBF of up at $\pm 10\%$; errors in bolus width the same. So an underestimation of $T_1b$ by $10\%$, and an underestimation of the bolus width by $10\%$ will cause an overestimation of CBF of $20\%$; underestimation of $T_1b$ or bolus width will cause overestimation of CBF and vice versa. The number of timepoints has little effect on the magnitude of the errors.

Figure 3.8 shows percentage errors in arrival time in the 2-parameter fit is for differing relaxation times and bolus widths, averaged over all values of simulated CBF. Figure 3.9 shows the percentage error in arrival time as a function of simulated value of arrival time for an assumed value of bolus width of 900ms and an assumed value $T_1b$ of 1350ms. Figure 3.9 shows that the percentage error in arrival time depends on the value of arrival time used to simulate data; if the data is simulated with a short arrival time the fitted value can be $25\%$ higher than the true value. However, this error approaches zero when a longer arrival time is used in the simulation. Figure 3.8 shows that an underestimate of either bolus width or $T_1b$ of $10\%$ can give an overestimate of arrival time of $10\%$ (20\% if both parameters are in error), and vice versa.

3.3.1.3 3-parameter fit for CBF, arrival time and bolus width

In this fit, the only parameter not estimated is the relaxation time. CBF, arrival time and bolus width are estimated from simulated data for 4 values of CBF, arrival time, bolus width and relaxation time, so that for each relaxation time there are 64 estimates of CBF, arrival time, and bolus width. This section shows the mean and standard deviation of the percentage errors in CBF for all of these 64 estimates for any given value of relaxation time. Figure 3.10 shows boxplots of the series of individual percentage errors in estimated CBF for all simulated parameters, grouped by the assumed value for relaxation time, at an SNR of 24. It is clear that even at this high SNR the dataset with 4 timepoints gives highly variable results, with CBF errors of up to $35\%$. There is no apparent structure to the size of these errors, e.g. there is no linear change in percentage CBF error with simulated bolus width, CBF or arrival time. The percentage error instead seems to depend in some complex manner on the values of all the other parameters. This is less clear for data with 8 or 16 timepoints; however figure 3.12 shows that this ability to fit three parameters vanishes at lower values of SNR. For an SNR of 10
Figure 3.7: Percentage error in CBF for 2-parameter fit. The three graphs show the percentage error in CBF for 4, 8 and 16 timepoints, for 4 assumed values of bolus width (points along the x axis) and 4 assumed values of relaxation time (coloured lines).
Figure 3.8: Percentage error in arrival time for 2-parameter fit. The three graphs show the percentage error in arrival time for 4, 8 and 16 timepoints, for 4 assumed values of bolus width (points along the x-axis) and 4 assumed values of relaxation time (coloured lines).

Figure 3.9: Percentage error in arrival time as a function of simulated arrival time for 2-parameter fit. The graph shows the percentage error in arrival time for 16 timepoints and an assumed value of $T_1^b_2$ of 1350 ms. for 4 simulated values of arrival time (points along the x-axis) and 4 assumed values of bolus width (coloured lines).
(a realistic estimate of SNR for the ASL scans with 20 averages used elsewhere in this study, see chapter 3.4.3) and 4 timepoints, the percentage error on CBF can approach 50%. The values shown here are the means of all simulated fits with the given relaxation time: the error bars on these means are large, as can be seen in figure 3.11. For brevity only the graphs for CBF estimates for a single relaxation time have been shown: results are similar for the other parameters. It is worth noting that in almost all cases here CBF is overestimated.

**Figure 3.10:** Percentage error in individual CBF values from 3-parameter fit including bolus width. For each of the three sets of timepoints (4, 8 and 16) at an SNR of 24 the boxplots show the variation in percent estimated CBF for all the possible permutations of simulated parameters CBF, arrival time and bolus width.

### 3.3.1.4 3-parameter fit for CBF, arrival time and relaxation time

In this fit, the only parameter not estimated is the bolus width. I have calculated the estimated percentage error in CBF for the maximum SNR of 24 for each of the 4 simulated CBF values, 4 simulated arrival times and 4 simulated bolus widths. There are therefore 64 separate values, and they are plotted along the x-axis. Figure 3.13 shows the a series of individual percentage errors in estimated CBF for all simulated parameters, grouped by the assumed value for relaxation time. It is clear that even at this high SNR the dataset with 4 timepoints gives highly variable results, depending in some complex manner on the values of the other parameters. This is less clear for data with 8 or 16 timepoints; however figure 3.15 shows that this ability to fit three parameters vanishes at lower values of SNR. The values shown here are the means of all simulated fits with the given bolus width: the error bars on these means are large, as can be seen in figure 3.14. For brevity only
Chapter 3. Mathematical modelling of the ASL signal

Figure 3.11: Percentage error in CBF values with error bars vs SNR from 3-parameter fit including bolus width. For each of the three sets of timepoints the graphs show the percentage error in CBF for all 64 simulations with a single value of assumed relaxation time, against SNR, with error bars showing the standard deviation. As expected, the SD decreases as SNR increases, and even for the 16 timepoint dataset the percent error in CBF rises above 10% below an SNR of around 10.
Chapter 3. Mathematical modelling of the ASL signal

Figure 3.12: Percentage error in CBF values vs SNR from 3-parameter fit including bolus width. For the 3-parameter fit including bolus width there are 64 sets of estimated parameters for each assumed value of relaxation time. These graphs show for the three sets of timepoints and the 4 assumed values of relaxation time how the percentage error in CBF changes with SNR. Even for the 16 timepoint dataset the error in CBF rises sharply below and SNR of 10.
the graphs for CBF estimates have been shown: results are similar for the other parameters.

![Boxplots of percentage error in individual CBF Values from 3-parameter fit including relaxation time.](image)

**Figure 3.13:** Percentage error in individual CBF Values from 3-parameter fit including relaxation time. For each of the three sets of timepoints (4, 8, and 16) at an SNR of 24 the boxplots show the variation in percent estimated CBF for all the possible permutations of simulated parameters CBF, arrival time and relaxation time.

### 3.3.2 Precision of estimates - comparison of goodness of fit

For the 4-timepoint data, as figure 3.16(a) shows, for no values of simulated parameters was the 2-parameter fit better than the 1-parameter fit. Even for the highest SNR of 25 and assumed values close to the true values, the probability that the 2-parameter fit was better than the 1-parameter fit was never lower than 0.097.

For the 8 timepoint data, as figure 3.16(b) shows, at SNR levels above 8 - 12 the 2-parameter fit was better than the 1-parameter fit. Even at an SNR level of 25 the 3-parameter fits were never a better fit then the 2-parameter fit.

For the 16 timepoint data, the 2-parameter fit was better than the 1-parameter fit for SNR levels higher than around 7 - 10 (depending on how well the assumed parameters matched the true parameters), but again at the SNR levels considered here the 3-parameter fits were never a better fit than the 2-parameter fit. This is shown in figure 3.17 showing probabilities for 1-parameter vs 2-parameter, and 2-parameter vs both 3-parameter fits.

I have omitted the 2-parameter vs 3-parameter probability plots for the 4 and 8 timepoints for brevity, but they can be found in appendix B.
Chapter 3. Mathematical modelling of the ASL signal

Figure 3.14: Percentage error in CBF values with error bars vs SNR from 3-parameter fit including bolus width. For each of the three sets of timepoints the graphs show the percentage error in CBF for all 64 simulations with a single value of assumed bolus width, against SNR with error bars showing the standard deviation. As expected, the SD decreases as SNR increases, and even for the 16 timepoint dataset the percent error in CBF rises above 10% below an SNR of around 7.
CHAPTER 3. MATHEMATICAL MODELLING OF THE ASL SIGNAL

Figure 3.15: Percentage error in CBF values vs SNR from 3-parameter fit including relaxation time. For each of the 3 sets of timepoints and 4 assumed values of bolus width these graphs show the percentage error in CBF vs SNR. There is a systematic error in CBF depending on the assumed value of bolus width, and even for 16 timepoints the percentage error in CBF rises sharply to above 10% below and SNR of 10.
Figure 3.16: Comparison of goodness of fit for 1-parameter vs 2-parameter fits. These graphs show for each of the 3 sets of timepoints the p-value from an f-test comparing the 1-parameter and 2-parameter fits. For the 2-parameter fit to be "better" (i.e. have relatively lower residuals) the p-value should be less than 0.05. It is clear that for 4 timepoints the 1-parameter fit always has lower residuals. However, the estimated CBF from the 1-parameter fit is not necessarily more accurate than that from the 2-parameter fit, as it will depend on the value chosen for the assumed arrival time.
CHAPTER 3. MATHEMATICAL MODELLING OF THE ASL SIGNAL

(a) 2-p vs 3-p (bolus width)  

Figure 3.17: Comparison of goodness of fit for 16 timepoints for 2-parameter vs 3-parameter fits. The graphs show the p-values from f-tests comparing the 2-parameter fit with the 3-parameter fits, vs SNR. For the residuals from the 3-parameter fits to be relatively less than those from the 2-parameter fit, the p-value should be below 0.05. It is clear that even for the highest SNR this is never the case. 3 parameters cannot be fitted even at an SNR of 15 and with 16 timepoints.

3.4 Multi-timepoint dataset

ASL sequences used for clinical research are of necessity limited in the time available, and hence in the number of averages, and number of timepoints, that can be collected. This section describes the results of a lengthy ASL sequence on a young, healthy volunteer. It was taken to investigate the best sequence of timepoints, and the optimum number of averages, for other clinical scans.

3.4.1 Methods

The standard gradient-echo EPI sequence was used to collect data from a single healthy volunteer (Male, age 29y), but images were taken at 23 timepoints instead of the usual 4. Timepoints chosen were 400ms 600ms 800ms 900ms 950ms 1000ms 1050ms 1100ms 1150ms 1200ms 1250ms 1300ms 1350ms 1400ms 1450ms 1500ms 1550ms 1600ms 1700ms 1800ms 1900ms and 2000ms and 2100ms. A whole-brain mask was calculated by segmenting the $M_0$ image using the 8 “NewSegment” function, and adding the white-matter and grey-matter masks. When this dataset was acquired no corresponding $T_1$-weighted or $T_2$-weighted image was acquired. The mean within-brain subtraction signal was calculated at each timepoint. The fol-
lowing fits were then calculated for the whole-brain subtraction signal, and for the subtraction signal within each individual voxel:

- 1-parameter fit for CBF, assuming arrival time of 750ms, bolus width of 1100ms and relaxation time of blood of 1600ms
- 2-parameter fit for CBF and arrival time, assuming bolus width of 1100ms and relaxation time of blood of 1600ms
- 3-parameter for CBF, arrival time and bolus width, assuming relaxation time of blood as above
- 3-parameter fit for CBF, arrival time and relaxation time of blood, assuming bolus width as above
- 4-parameter for CBF, arrival time, bolus width and relaxation time of blood

For each voxel fit the sum of squared errors was also calculated, and a series of f-tests performed for each voxel giving the probability that the extra parameters included in the fit decreased the sum of squared errors.

### 3.4.2 Results

The results of the fits on the whole-brain subtraction signal, together with the raw data, are shown in figure 3.18. The estimated values from the fits are shown in table 3.3, and the results of the goodness of fit tests in table 3.4. Maps from

<table>
<thead>
<tr>
<th>Fit Type</th>
<th>CBF (ml/100ml/min)</th>
<th>Arrival Time (ms)</th>
<th>Bolus Width (ms)</th>
<th>Relaxation Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-p</td>
<td>34.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2-p</td>
<td>31.6</td>
<td>594</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3-p (bolus width)</td>
<td>30.6</td>
<td>568</td>
<td>1142</td>
<td>-</td>
</tr>
<tr>
<td>3-p (relaxation time)</td>
<td>40.4</td>
<td>637</td>
<td>-</td>
<td>1330</td>
</tr>
<tr>
<td>4-p</td>
<td>27.8</td>
<td>545</td>
<td>1146</td>
<td>1726</td>
</tr>
</tbody>
</table>

The voxelwise fits are shown in figures 3.19, 3.20, 3.21(a) and 3.21(b). The values have not been limited, as has been done in the ASL calculation used in the FTD
CHAPTER 3. MATHEMATICAL MODELLING OF THE ASL SIGNAL

Figure 3.18: Mean whole-brain subtraction signal for 23-timepoint data, and estimated data from multiple fits. The graphs shows the whole-brain subtraction signal (black crosses) and the estimated subtraction signal from multiple fits.

Table 3.4: Results of comparison of goodness of fit for differing models

<table>
<thead>
<tr>
<th>Comparison</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-p to 2-p</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2-p to 3-p (bolus width)</td>
<td>0.68</td>
</tr>
<tr>
<td>2-p to 3-p (rlx. time)</td>
<td>1.0</td>
</tr>
<tr>
<td>3-p (bolus width) to 4-p</td>
<td>0.69</td>
</tr>
<tr>
<td>3-p (rlx. time) to 4-p</td>
<td>0.11</td>
</tr>
</tbody>
</table>

study described elsewhere in this thesis. Despite the large number of timepoints, the 2-parameter fit proves to be the most robust. The CBF maps show that the 3-parameter fit including bolus width gives a consistently lower value of CBF, while the 3-parameter fit including the relaxation time, and the 4-parameter fit give unrealistically high values (e.g. many of the values are $>200$ ml/100ml/min, where typical values are around 40-80 ml/100ml/min [Parkes et al., 2004]). However, there is clear structure within grey matter for the fits with a higher number of parameters, where there is higher signal, and thus better SNR. A more complicated CBF calculation, which uses different models for higher SNR, could perhaps be explored. Figure 3.22 shows the results of the f-test comparison of goodness of fit. The maps have been set up so that voxels where the fit with the greater number of parameters provides a better fit are set to 2; where the fit with fewer parameters is better, the voxel has been set to 1, and voxels outside the brain are set to 0. In
Figure 3.19: CBF maps from multi-parameter fits. The subtraction data has been fitted using 1-parameter, 2-parameter 3-parameter (bolus width), 3-parameter (relaxation time) and 4-parameter fits.

Figure 3.20: Arrival time maps from multi-parameter fits. The subtraction data has been fitted using 2-parameter 3-parameter (bolus width), 3-parameter (relaxation time) and 4-parameter fits. There has been no limiting of the arrival time fit; the apparent holes are where the arrival time is negative.

Figure 3.21: Bolus width and relaxation time maps from multi-parameter fits. These maps show the values from the relevant 3-parameter and 4-parameter fits for bolus width and relaxation time. Although the maximum values for both sets of maps have been set higher than is physiologically sensible, many voxels have these higher values.
14% of the voxels, the 4-parameter fit is better than the 2-parameter fit; in 2% is the 3-parameter fit including bolus width better than the 2-parameter fit; in 18% is the 3-parameter fit including relaxation time better than the 2-parameter fit; in 31% the 4-parameter fit is better than the 3-parameter fit including bolus width and in 5% the 4-parameter fit is better than the 3-parameter fit including relaxation time. This again emphasises that the bolus width should not be included in any fit. The 3-parameter fit including the bolus width is only marginally better than the 2-parameter fit, whereas including the relaxation time does seem to lead to a better fit in some voxels. However, given the extreme values of estimates produced by the 3-parameter fit including relaxation time, and the 4-parameter fit; the 2-parameter fit is to be preferred.

Figure 3.22: Comparison of goodness of fit in multi-parameter fits. These maps show which fit has the lowest residuals for differing regions of the brain. The significance level for a better fit has been set at 0.05. The test above each set of axial maps describes the fits being compared: Figure 3pa shows the 3-parameter fit for bolus width, and Figure 3pb the 3-parameter fit for relaxation time. The maps have been set up so that voxels where the fit with the greater number of parameters provides a better fit are set to 2; where the fit with fewer parameters is better, the voxel has been set to 1, and voxels outside the brain are set to 0.
3.4.3 Measuring SNR

It can be seen above that the SNR of the ASL signal has an effect on the number of parameters that can be estimated, and also on the credibility attached to those measurements. Here I investigate the temporal SNR of a real ASL dataset collected with 23 timepoints.

3.4.3.1 Methods

The standard gradient-echo EPI sequence was used to collect data from a single healthy volunteer (Male, age29.4y), but an extended series of images was taken, with 125 pairs instead of the usual 20. The subtraction signal at a delay time of 1200 ms was calculated; this time was chosen as it is when the signal is close to its maximum. The images were segmented using the $M_0$ image, and a region of interest (ROI) was defined where the grey matter probability was $>0.95$. The value of the ASL signal $\Delta M$ in this ROI was calculated for each pair of images, and these data were sampled with replacement 100 times for each of 5 - 50 averages. The SNR is proportional to the inverse of the variance of these measurements.

Matlab code was written specifically for the analysis. The Matlab routine “fminsearch” was used to fit the model to the data.

3.4.3.2 Results

The SNR from the extended ASL sequence as a function of number of averages is shown in figure 3.23. The number of averages in the sequence used in the FTD study described elsewhere in this thesis is 20, corresponding to an SNR of approximately 10. This sequence has a scan time of just over 8 minutes. This is likely to be the maximum time available in a clinical setting for patients with FLTD, given that other diagnostic scans will be acquired at the same time. In this chapter I have shown that at this SNR it is not possible to fit for more than 2 parameters. It is important to fit for CBF and arrival time, as an incorrect estimate of arrival time can lead to large errors in estimated CBF.
 CHAPTER 3. MATHEMATICAL MODELLING OF THE ASL SIGNAL

**Figure 3.23:** SNR vs no of averages for ASL data. The graph shows the individual data points, and a least-squares fit with a quadratic. The ASL sequence used in the FTD study described elsewhere in this thesis has 20 averages, which correspond to an SNR of approximately 10.

### 3.5 Discussion

Most ASL analyses where a model is used to fit parameters to measured data, use a deterministic algorithm for the fit. The details of the model vary from approach to approach, with most trying to fit extra parameters such as measurement of the blood/brain partition coefficient [Roberts et al., 1996], or endothelial permeability [Parkes, 2002]. But whichever parameters are included in the model, the model itself is fitted using a technique similar to Matlab’s fminsearch function. A different approach is taken by Woolrich [Woolrich et al., 2009], and Chappell et al [Chappell et al., 2009], who use Bayesian modelling to perform the fit. This allows the probability functions for selected parameters to be estimated, and allows the expectation that certain parameters will have certain values to be included a priori. For instance, in a fit using fminsearch, values such as the bolus width, relaxation time can be fitted as free parameters, or given an estimated but exact value. The use of a Bayesian algorithm allows the a priori probability that parameters have certain values to be entered, with the expected value and precision of that value to be given as a prior. This approach also reports a probability function for the estimated parameter, e.g. CBF, arrival time, bolus width, so that the expected errors on these parameters can be estimated independently. This is not possible with the deterministic approach, where the residual is a sum of errors in all the estimated parameters.

As stated above, many papers have proposed modification to the kinetic
model described here, to allow more parameters to be estimated from the ASL signal. However, at the SNRs described (typically 10-20), it is doubtful whether more parameters can be estimated with any degree of accuracy. The principal determinant of the value of the ASL subtraction signal is CBF: if arrival time is not fitted but assumed, the accuracy of the assumed value will have a great effect on the accuracy of the CBF measurement, as shown in figure 3.3. Other variables which are known to affect the signal, such as the movement of labelled water from blood to tissue, or out of the voxel being modelled, are unlikely to be measurable currently in a scan which is clinically acceptable, as shown in figure 3.17. This study confirms the results found by Carr et al [Carr et al., 2007], that using the ASL signal to estimate parameters other than CBF and arrival time quantitatively is unlikely to be successful.

This simulated data shows that even with an SNR of 25, with 4 timepoints a 1-parameter fit is more precise than any other fits (see figure 3.16(a)). However, to get an accurate estimate of CBF it is important that the other assumed values are close to the true values, as shown in figure 3.7. In the absence of an arrival time measured by some other means, the 2-parameter fit is to be preferred as giving a more accurate estimate. The data with 8 timepoints offers a better way of estimating CBF and arrival time, but 3 parameters cannot be fitted with any confidence. The same is true of the data with 16 timepoints: CBF and arrival time can be estimated with reasonable confidence, but not bolus width or relaxation time. The advantage of more data would be offset in the real world by probably a longer scan time (depending on the sequence possibly double the time), with more possibility of the subject moving. Therefore this simulation shows that 8 timepoints are to be preferred.

For the rest of this thesis the ASL model used is the single-blood-compartment model described by equations 3.8, 3.9 and 3.10.
Chapter 4

Comparison of gradient-echo and spin-echo MRI sequences in ventral brain regions

4.1 Introduction

In this chapter I examine the effect of two different read-out sequences on measured cerebral blood flow (CBF) values in ventral brain regions, and how these values are affected by differing ways of quantifying the signal.

Arterial Spin Labelling (ASL) is gaining clinical acceptance, and has the potential to be a cheaper and more convenient alternative to SPECT and PET for measuring cerebral blood flow. ASL uses the water in blood as an endogenous tracer. Inflowing blood is magnetically labelled; a fast read-out is used to take paired images with and without the label, and the cerebral CBF is calculated from the difference between the two images Gradient-echo EPI sequences are frequently used because they typically allow a very fast read-out. However, in regions where there are field inhomogeneities, particularly those caused by tissue/air interfaces, the images obtained with this type of pulse sequence can suffer from distortion and drop-out. The problem increases as the field strength increases; as is mentioned by Wang et al [Wang et al., 2002]. Early studies were carried out at 1.5T, nowadays fields of 3T or more are commonly used because of their greater SNR,
and 7T scanners are also coming into use. Wang et al [Wang et al., 2004] compared a gradient-echo EPI sequence (GE) with a spin-echo EPI sequence (SE) at 1.5T, and found the spin-echo less prone to drop-out in regions of high field inhomogeneity. In his study the inhomogeneity was intentionally induced by placing a small piece of metal close to the left temporal side of the subject’s head; other naturally occurring inhomogeneities (as are found in ventral brain regions) were not mentioned. 3D GRASE sequences [Talagala et al., 2004], [Günther et al., 2005] have also been suggested as offering a faster read-out, but again the papers do not consider ventral brain regions. Chen et al [Chen et al., 1997] suggest a half-Fourier single shot turbo spin-echo (HASTE) which shows clear advantages over gradient-echo EPI in regions close to the frontal sinuses and eyes; unfortunately the read-out is comparatively slow, and does not offer full coverage of the brain. The susceptibility distortion also affects the BOLD signal normally used for functional MRI, and recently Fernandez-Seara et al [Fernández-Seara et al., 2007] and Kemeny et al [Kemeny et al., 2005] have suggested using spin-echo ASL as an alternative to BOLD for functional MRI in susceptible brain regions. Susceptibility artefacts seen in standard EPI readout can limit ASL analysis to upper brain regions: Du et al [Du et al., 2006] and Cha et al [Cha et al., 2013] used only upper brain regions for their statistical analysis to minimise the potential effects of susceptibility artifacts. Regions prone to susceptibility distortion and dropout are now recognised to be involved in language (the anterior and ventral temporal lobes [Hodges et al., 1992], [Sartori and Job, 1988], memory (the medial temporal lobe [Squire and Zola-Morgan, 1991] and executive function (the orbitofrontal lobes [Luria, 1971], [Rolls, 2000]. These are the regions which are affected in the early stages of fronto-temporal lobar degeneration (FTLD) [Neary et al., 2005]. This neurodegenerative disease is difficult to diagnose in the early stages, and hence imaging of these regions could be particularly important to aid early diagnosis. Du et al have shown that CBF can help distinguish FTLD from other dementias [Du et al., 2006], but they did not look at CBF in ventral or orbitofrontal brain regions. As mentioned in chapter 2.2.2, a spin-echo sequence is inherently less prone to distortions caused by field inhomogeneities, although it has the disadvantages of more RF power applied to the subject, and potentially a longer echo time. This study compares CBF values measured using gradient-echo and spin-echo multi-slice 2D EPI readouts in regions of the brain which are likely to suffer from susceptibility artefacts, and those which are not. The hypothesis is that the gradient-echo EPI ASL will suffer from more signal loss in the frontal and temporal regions close to air-tissue inter-
faces than experienced by the spin-echo EPI ASL. Hence we would expect a lower measured value of CBF in these regions compared with other brain regions less prone to susceptibility artefacts, when using a global M₀ correction factor.

The sequences were chosen as being part of the standard set of sequences supplied by most scanner manufacturers, in order to maximise the clinical relevance of the findings.

4.2 Materials and methods

This study was approved by the NHS NW9 Research Ethics Committee. Seventeen healthy right-handed subjects (9 female; mean age 25.4±4.6y) were scanned on a Philips Achieva 3T MRI. The subjects were given identifiers R01 to R17.

4.2.1 MRI acquisition parameters

An 8-channel head coil was used. Voxel size was 3.5x3.5x5.0 mm³, with a 1.0 mm inter-slice gap. 20 slices were acquired. A EPISTAR sequence was used as described by Edelman [Edelman et al., 1994], [Edelman and Chen, 1998], with STAR labelling with a 150mm label thickness and 10mm gap between the top of the labelling slice and the bottom of the read-out volume. 20 pairs of images were acquired with EPI readout at each labelling delay of 800 ms, 1200 ms, 1600 ms and 2000 ms. The images were acquired alternately, with control, label, control, etc. The images for each labelling delay were collected independently, in the order of 2000ms, 1200ms, 800ms, 1600ms. For GE, TE = 21 ms, TR = 3100 ms; for SE TE = 18.5 ms, TR = 3290 ms. Sequence parameters were chosen to make the imaging times comparable. SE images were acquired with SENSE factor of 2 and half-scan factor of 0.6; GE images were acquired without SENSE, and with half-scan turned off. (SENSE is a Philips-specific term describing the use of multiple coils RF and imaging coils. Half-scan describes a reduction in the acquisition of k-space data, with the rest being inferred by interpolation.) Imaging time was 9 min 0 s for GE, and 8 min 40 s for SE. Vascular crushing was enabled to remove large vessel signal, with the signal from blood travelling at a velocity > 5 cm/s being crushed. An additional calibration image with TR=10 s and no labelling was acquired for both
readout types to allow quantification of CBF. This image is referred to as the $M_0$
image, as it collects the baseline proton density of tissue. It was acquired after the
labelling delay images. A high resolution turbo spin-echo T$_2$-weighted image was
also acquired (1x1x1 mm$^3$ voxel size), which was used in the quantification of the
ASL images to provide a head mask. All calculations were done in Matlab 2009
and SPM8.

### 4.2.2 Region definition

Eleven Regions of Interest (ROIs) were selected from the aal library [Tzourio-
Mazoyer et al., 2002]. Four ROIs were chosen as control regions, being unlikely
to suffer from susceptibility distortion. These ROIs were the amygdala, calcar-
ine sulcus (calcarine), cerebellum-6 and fusiform gyrus (fusiform). Other regions
were chosen as likely to suffer from susceptibility distortion; these ROIs were the
inferior frontal pars orbitalis (front-inf-orb), the frontal medial orbital (front-med-
orb), the frontal superior orbital (front-sup-orb), the hippocampus, the rectus, the
inferior temporal (temp-inf) and the medial temporal pole (temp-pole-mid). The
location of these ROIs is shown in figure 4.1. A whole-brain ROI was also used;
defined as the sum of all the AAL regions, and hence excluding CSF. ROIs were
further restricted to grey matter only, using the SPM8 grey matter probability
template, and assuming that a voxel was grey matter if the template probability
was $> 0.75$.

### 4.2.3 Pre-processing

The individual images were aligned using the rigid-body affine “Align” function
in the SPM8 toolbox, aligning all subsequent images to the first available control
image in the 800 ms time delay series. Pairs of images where the linear adjustment
was greater than the voxel size, or the rotational adjustment greater than 30$^0$ were
rejected. Label/control images were acquired in pairs; subtraction images were
calculated by subtracting a labelled image from the following control image (the
label should produce a signal reduction). Intermittent scanner artefacts were ob-
served in some subtraction images, and so quality checks were made. Individual
subtraction images were viewed, and images which showed obvious artefacts (e.g.
those shown in figure 4.2) were rejected visually. In addition, the interquartile range for each subtraction image was plotted, and images with a value where the inter-quartile range was greater than twice the overall mean interquartile range for all images for all subjects were rejected. Subjects were rejected overall if 50% or more of the individual subtraction images at a single time point were rejected.

Data were also rejected if the mean subtraction signal for each delay time for each subject for each ROI was negative. Subjects with three or more negative points in any region were rejected. (Depending on the arrival time at a particular ROI, it is possible with a noisy signal for the first and last points to be negative.)

During the time these images were acquired, the scanner was giving intermittent errors, and on several occasions image acquisition was halted and restarted because of gross visible errors in the images. Unfortunately at the time of acquisition only the averaged perfusion-weighted images were available for inspection, as the code to examine the individual images had not been written.
4.2.4 Analysis

Once the quality of the individual images was assured, average subtraction maps for each delay time were calculated. The $M_0$ image was segmented using the "New Segment" function in SPM8. A whole-brain mask was calculated by adding the grey-matter and white-matter masks (CSF was excluded), and used to extract the average whole-brain $M_0$. CBF and arrival time were calculated for each voxel independently using the single-blood-compartment model developed by Parkes [Parkes and Tofts, 2002] described earlier (see chapter 3). The subtraction signal is modelled by:

$$\Delta M = 0 \quad \text{for } t < t_a$$
$$\Delta M = 2f m_0^0 \alpha (t - t_a) \exp \left( -\frac{t}{T_1b} \right) \quad \text{for } t_a < t < t_a + \tau$$
$$\Delta M = 2f m_0^0 \alpha \tau \exp \left( -\frac{t}{T_1b} \right) \quad \text{for } t > t_a + \tau$$

where

- $T_1b$ - the relaxation time of blood measured in ms
- $\alpha$ - the labelling efficiency - a unitless number
- $m_0^0$ - the equilibrium magnetisation of blood measured in the same units as the subtraction signal itself
- $f$ - the CBF (measured in ml/min/100 ml)
- $t_a$ - the arrival time of the leading edge of the labelled bolus measured in ms
- $\tau$ - the bolus width measured in ms

The equilibrium magnetisation of tissue is measured in the $M_0$ image referred to earlier. The CBF quantification analysis needs the equilibrium magnetisation of blood $m_0^0$. The two are related by $\lambda$ - the blood/brain partition coefficient. $\lambda$ is the blood-brain partition coefficient, i.e. the relative water content of brain and blood. $\lambda$ can be given in ml/g (in which case perfusion will be in ml/100g/min) or in ml/ml (in which case perfusion will be in ml/100ml/min). Here

$$\lambda = \frac{M_0}{m_0^0}$$

107
The model assumes fixed values for $\alpha$, $T_{1_b}$, and $\lambda$ (but see discussion earlier in this thesis in chapter 3.1.1). $\alpha$ has been measured at 0.87 for the this labelling sequence, as described in appendix C). $T_{1_b}$ and $\lambda$ can be taken from the literature; values of 1700ms for $T_{1_b}$ [Greenman et al., 2003]) and 0.9 for $\lambda$ [Roberts et al., 1996]) are used. $\tau$ may vary within the brain as blood takes different paths to arrive at each voxel, but for young healthy volunteers such as used in this study, the variation is likely to be small. A fixed bolus width of 1100ms was used; this was the value measured earlier on the same scanner with the same labelling sequence, where more data were acquired during a longer image (see section 3.4.2). The calculated CBF value is scaled by $M_0$; this factor, and $\alpha$, the labelling efficiency, define the initial signal created within the labelling slice by the RF labelling pulse. As such, it should be a global value for all voxels. However, variations in the coil sensitivity and also variations in the local $B_0$ will also affect both the subtraction signal and the proton density signal. The $M_0$ image measures the tissue proton density, which is related to the blood proton density by

$$m^0_a = G \frac{M_0}{\lambda_v}$$  \hspace{1cm} (4.5)

where $\lambda_v$ is the blood-brain partition coefficient within the voxel and $G$ is a voxel-wise ‘gain’ factor encompassing local variations in $B_0$ and coil sensitivity. Grey matter and white matter have differing values of $\lambda$, so using a voxelwise $M_0$ value will account for variations in coil sensitivity and $B_0$, but not for local changes in $\lambda_v$. Roberts et al [Roberts et al., 1996]) measured $\lambda$ for grey matter as $0.98 \pm 0.04$, and for white matter as $0.84 \pm 0.05$, a difference of 15%. Neither approach is ideal, but using a voxelwise value for $M_0$ will have the effect of compensating for any signal drop-out due to local field variations due to tissue density changes. The ideal sequence would have included another calibration image using the body coil, as used by Macintosh et al [Macintosh et al., 2012], which could have measured variations in both $B_0$ and head-coil sensitivity, but this additional image was not acquired.

This study investigates two possible methods of using the $M_0$ correction factor:

- global correction - use the mean value for the whole brain excluding CSF. In this procedure a mask which includes both grey and white matter, but not cerebro-spinal fluid is calculated from the $T_2$-weighted image. This mask is
co-registered to the first of the ASL control images. The mean of all the $M_0$ voxels within the mask is then calculated, and used to scale the subtraction signal for all voxels. Hence the same scaling factor is used for all voxels.

- voxelwise correction - use the individual voxel value. In this procedure the subtraction signal for each voxel is scaled by the value from the corresponding voxel within the $M_0$ image.

In both cases the scaled subtraction signals are used to calculate CBF and AAT. Values for CBF which are $<0$ are set to 0, and values $>250$ ml/100ml/min are set to 250. Negative AAT values are set to 0.

CBF and arrival times are calculated by fitting the subtraction signal within each voxel using the Matlab function `fmsbnd`, using $M_0$ values as described above. The first ASL control image was normalised to the SPM EPI template, and the same transformation applied to the CBF and arrival time maps. Matlab was used to extract the mean signal within each ROI from the CBF and arrival time maps. Normalised subject images were then averaged so the group mean could be compared for GE and SE readouts using a Student t-test to compare the individual means.

Where susceptibility differences cause signal loss and distortion, a lower or zero signal can be expected. However, the noise within the signal, being a random effect, will be the same in these regions as in others. Consequently it is likely that there will be a lower SNR in regions affected by susceptibility distortions than in other regions. The temporal SNR for the subtraction signal within each ROI was calculated by extracting the mean subtraction signal ($S_n$) at one inversion time for all individual pairs of images for all voxels within the ROI. The inversion time of 1200 ms was used, as that is close to the maximum subtraction signal. The subtraction signal was then divided by the appropriate $M_0$, either global or voxelwise. The SNR was then calculated by

$$SNR_{ROI} = \frac{\text{mean } S_n}{\text{standard deviation}(S_n)}$$

Note that this is the temporal SNR; the CBF value is calculated by fitting a model to data from images that are the average of 20 measurements; these images will have a much higher SNR of the order of $\sqrt{20}$ (i.e. 4.5) times the temporal SNR calculated here.
Chapter 4. Comparison of Gradient-Echo and Spin-Echo MRI Sequences in Ventral Brain Regions

The correlation between CBF values for GE and SE values across individuals for each ROI was also calculated: a weaker correlation is expected in CBF values in regions prone to susceptibility distortion, but not in those regions which are not so prone. This calculation was repeated for arrival times. The correlation between GE and SE would be predicted to be weaker in these affected regions if the hypothesis is correct that GE values are degraded.

4.3 Results

4.3.1 Pre-processing

No subjects were rejected because of excessive movement as found by the alignment.

Two subjects had to be rejected because of image artefacts. One (R12) had excessive artefacts, and one (R08) had no subtraction signal, although the individual label and control images looked to be acceptable. This is possibly because for this individual the labelling pulse failed. Of the other 15 subjects, 8 had some images which were rejected because of excessive noise, but in most cases fewer than 10% of the 80 available images were rejected. 15 subjects (8F; mean age 25.3 ± 5.0 y) are included in the final analysis.

Figure 4.2 shows the individual subtraction images for slice 14 for subject R12 for gradient-echo for TI of 1600ms. This is a Matlab montage, with all images displayed at the same arbitrary scale: it is very clear that these images have been affected by scanner artefacts or by excessive motion. The stripes are possibly caused by a spark or other short-lived spike in the scanner hardware. At the time it was thought that some electrical equipment (e.g. lighting) in the scanner room had an intermittent fault. Eventually the problem was resolved by replacing the scanner gradient coil, but not until after this study had been completed.

Figure 4.3 shows how the $M_0$ signal differs from the median for GE and SE images. Voxels have been allocated numbers 1-5, where 1 means the $M_0$ value within the voxel is lower than the 2.5 percentile; 2 means the voxel $M_0$ value is within the 2.5 percentile to 10.0 percentile; 3 means the value is within 10 per-
Chapter 4. Comparison of gradient-echo and spin-echo MRI sequences in ventral brain regions

Figure 4.2: Individual subtraction images for slice 14 for subject R12, showing MRI artefacts

- 90 percentile, 4 means the value is within 90 percentile - 97.5 percentile and 5 means the value is greater than the 97.5 percentile. The images have been masked by the head mask used for the perfusion calculation. The extreme values are found in the ventral brain regions known to be prone to susceptibility distortion, and around the brain/skull interface. Since the $M_0$ correction factor is an inverse one; particularly low values of $M_0$ can have a disproportionately large effect, and perhaps in future a possible approach might be to use the voxelwise $M_0$ value where it is within the 2.5 percentile to the 97.5 percentile, and to either use the limiting value, or not use the voxels where the value is outside the limits.
CHAPTER 4. COMPARISON OF GRADIENT-ECHO AND SPIN-ECHO MRI SEQUENCES IN VENTRAL BRAIN REGIONS

Figure 4.3: Images for a typical individual showing where extreme $M_0$ values are to be found. Voxels have been allocated numbers 1-5, where 1 means the $M_0$ value within the voxel is lower than the 2.5 percentile; 2 means the voxel $M_0$ value is within the 2.5 percentile to 10.0 percentile; 3 means the value is within 10 percentile - 90 percentile, 4 means the value is within 90 percentile - 97.5 percentile and 5 means the value is greater than the 97.5 percentile.

4.3.2 Comparison of values

Figure 4.4 shows a mid-sagittal CBF map for gradient-echo and spin-echo acquisitions for a typical individual, analysed with both global and voxelwise $M_0$ correction. As is to be expected, the voxelwise corrected images are noisier than the globally corrected images. There are, however, clear differences in the frontal region between GE and SE images for both global and voxelwise quantification.

Figure 4.5 shows the mean CBF for each region averaged across all subjects for global and voxelwise correction. The error bar shown is the standard deviation of the 15 individual subject measurements. A Wilcoxon rank-sum test on the paired GE/SE values for each region shows a significant difference in some regions. This is indicated on the figure by asterisks; a single asterisk indicates a p-value $< 0.05$, and a triple asterisk shows a p-value $< 0.001$. It can be seen that for global correction many of the regions likely to be affected by susceptibility differences show a significant difference between the gradient-echo and spin-echo values; whereas regions unlikely to be affected do not show such a difference. The effect is also shown in figure 4.7, which shows the paired globally-corrected
GE/SE CBF values for each subject for the named regions. This figure shows the large inter-subject variations, which lead to the large error bars in figure 4.5. For each subject the same effect of increased signal with SE compared to GE is seen, which leads to the significant differences found by the Wilcoxon test. There are fewer significant differences for the voxelwise corrected values.

The voxelwise quantification values for the rectus do not follow this trend, with the GE CBF value being significantly higher than the SE. This may be because this is the smallest ROI considered, and hence likely to be the noisiest. The values are dominated by two high GE CBF values of 66.7 ml/100ml/min for R04 and 72.6 ml/100ml/min for R11.

Figures 4.6 and 4.8 show histograms and line plots for regional arrival times. There is less variation in arrival time across the regions, but there are significant differences between GE and SE values for both global and voxel-wise quantification. For perfect data the arrival time would be independent of the scaling factor introduced by the different quantification strategies, however as can be seen, the noise in real data results in differing values for the two quantification methods.

The figure 4.9 shows the mean temporal SNR for each region averaged across all subjects for global and voxelwise correction. It can be seen that the SNR is lower in distorted ROIs for the GE images.

The regional variations found are self-consistent; there is a good correlation between gradient-echo and spin-echo in regions not subject to susceptibility distortion. The correlation between regional gradient-echo and spin-echo measurements of CBF in these regions is 0.72, whereas in regions prone to susceptibility distortion it is 0.44.
Figure 4.4: Sagittal images for a typical individual of CBF in ml/100ml/min in GE and SE images for both global and voxelwise $m_0^0$ correction showing the differences, particularly in ventral and posterior brain regions.
CHAPTER 4. COMPARISON OF GRADIENT-ECCHO AND SPIN-ECCHO MRI SEQUENCES IN VENTRAL BRAIN REGIONS

Table 4.1: Mean regional CBF values in ml/100ml/min for global \( m_s^0 \) correction, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where \( p < 0.05 \) are marked with *; where \( p < 0.01 \) are marked with **; where \( p < 0.001 \) are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Gradient-echo</th>
<th>Spin-echo</th>
<th>Wilcoxon p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>27.4 ±8.4</td>
<td>33.6 ±6.6</td>
<td>0.03*</td>
</tr>
<tr>
<td>Amygdala</td>
<td>23.3 ±8.9</td>
<td>36.7 ±12.9</td>
<td>0.006*</td>
</tr>
<tr>
<td>Calcarine</td>
<td>42.5 ±16.3</td>
<td>41.4 ±13.7</td>
<td>0.80</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>35.6 ±16.1</td>
<td>36.5 ±10.9</td>
<td>0.80</td>
</tr>
<tr>
<td>Fusiform</td>
<td>29.2 ±10.3</td>
<td>35.9 ±11.3</td>
<td>0.09</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>22.8 ±9.1</td>
<td>35.2 ±7.5</td>
<td>0.002*</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>30.6 ±9.9</td>
<td>36.4 ±11.2</td>
<td>0.15</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>16.5 ±10.8</td>
<td>30.8 ±12.4</td>
<td>0.001**</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>25.2 ±9.2</td>
<td>37.6 ±12.3</td>
<td>0.008*</td>
</tr>
<tr>
<td>Rectus</td>
<td>21.1 ±8.5</td>
<td>39.8 ±13.3</td>
<td>0.0002***</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>22.1 ±8.3</td>
<td>29.3 ±10.9</td>
<td>0.06</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>23.8 ±11.6</td>
<td>38.8 ±16.8</td>
<td>0.01*</td>
</tr>
</tbody>
</table>

Table 4.2: Mean regional CBF values in ml/100ml/min for voxelwise correction, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where \( p < 0.05 \) are marked with *; where \( p < 0.01 \) are marked with **; where \( p < 0.001 \) are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Gradient-echo</th>
<th>Spin-echo</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>22.9 ±6.3</td>
<td>24.6 ±4.7</td>
<td>0.5</td>
</tr>
<tr>
<td>Amygdala</td>
<td>20.7 ±7.3</td>
<td>24.5 ±8.8</td>
<td>0.2</td>
</tr>
<tr>
<td>Calcarine</td>
<td>28.3 ±10.3</td>
<td>30.3 ±10.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>25.8 ±10.6</td>
<td>25.8 ±7.7</td>
<td>0.97</td>
</tr>
<tr>
<td>Fusiform</td>
<td>25.2 ±8.3</td>
<td>26.1 ±8.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>17.1 ±7.4</td>
<td>25.1 ±7.4</td>
<td>0.005**</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>27.7 ±11.9</td>
<td>20.6 ±6.6</td>
<td>0.1</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>20.5 ±14.0</td>
<td>25.5 ±12.4</td>
<td>0.2</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>21.9 ±8.1</td>
<td>24.2 ±7.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Rectus</td>
<td>38.3 ±17.2</td>
<td>27.0 ±10.2</td>
<td>0.03*</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>23.9 ±8.7</td>
<td>24.3 ±9.2</td>
<td>1</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>24.1 ±10.8</td>
<td>31.7 ±14.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>
### Table 4.3: Mean regional CBF values in ml/100ml/min for GE images comparing global and voxelwise correction, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where p<0.05 are marked with *; where p<0.01 are marked with **; where p<0.001 are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Global</th>
<th>Voxelwise</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>27.4±8.4</td>
<td>22.9±6.3</td>
<td>0.10</td>
</tr>
<tr>
<td>Amygdala</td>
<td>23.3±8.9</td>
<td>20.7±7.3</td>
<td>0.41</td>
</tr>
<tr>
<td>Calcarine</td>
<td>42.5±16.3</td>
<td>28.3±10.3</td>
<td>0.009**</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>35.6±16.1</td>
<td>25.8±10.6</td>
<td>0.09</td>
</tr>
<tr>
<td>Fusiform</td>
<td>29.2±10.3</td>
<td>25.2±8.3</td>
<td>0.22</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>22.8±9.1</td>
<td>17.1±7.4</td>
<td>0.05*</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>30.6±9.9</td>
<td>27.7±11.9</td>
<td>0.34</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>16.5±10.8</td>
<td>20.5±14.0</td>
<td>0.48</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>25.2±9.2</td>
<td>21.9±8.1</td>
<td>0.34</td>
</tr>
<tr>
<td>Rectus</td>
<td>21.1±8.5</td>
<td>38.3±17.2</td>
<td>0.004**</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>22.1±8.3</td>
<td>23.9±8.7</td>
<td>0.77</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>23.8±11.6</td>
<td>24.1±10.8</td>
<td>0.90</td>
</tr>
</tbody>
</table>

### Table 4.4: Mean regional CBF values in ml/100ml/min for GE images comparing global and voxelwise correction, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where p<0.05 are marked with *; where p<0.01 are marked with **; where p<0.001 are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Global</th>
<th>Voxelwise</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>33.6±6.6</td>
<td>24.6±4.7</td>
<td>0.0009**</td>
</tr>
<tr>
<td>Amygdala</td>
<td>36.7±12.9</td>
<td>24.5±8.8</td>
<td>0.009**</td>
</tr>
<tr>
<td>Calcarine</td>
<td>41.4±13.7</td>
<td>30.3±10.1</td>
<td>0.034*</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>36.5±10.9</td>
<td>25.8±7.7</td>
<td>0.01*</td>
</tr>
<tr>
<td>Fusiform</td>
<td>35.9±11.3</td>
<td>26.1±8.5</td>
<td>0.014</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>35.2±7.5</td>
<td>25.1±7.4</td>
<td>0.002**</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>36.4±11.2</td>
<td>20.6±6.6</td>
<td>0.0009***</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>30.8±12.4</td>
<td>25.5±12.4</td>
<td>0.089</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>37.6±12.3</td>
<td>24.2±7.6</td>
<td>0.0042**</td>
</tr>
<tr>
<td>Rectus</td>
<td>39.8±13.3</td>
<td>27.0±10.2</td>
<td>0.0062**</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>29.3±10.9</td>
<td>24.3±9.2</td>
<td>0.21</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>38.8±16.8</td>
<td>31.7±14.3</td>
<td>0.17</td>
</tr>
</tbody>
</table>
**Table 4.5:** Mean regional arrival time values in ms for global quantification, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where \( p < 0.05 \) are marked with *; where \( p < 0.01 \) are marked with **; where \( p < 0.001 \) are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Gradient-echo</th>
<th>Spin-echo</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>565 ± 87</td>
<td>644 ± 91</td>
<td>0.03 *</td>
</tr>
<tr>
<td>Amygdala</td>
<td>692 ± 206</td>
<td>727 ± 152</td>
<td>0.53</td>
</tr>
<tr>
<td>Calcarine</td>
<td>432 ± 169</td>
<td>572 ± 150</td>
<td>0.014 *</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>546 ± 140</td>
<td>625 ± 103</td>
<td>0.062</td>
</tr>
<tr>
<td>Fusiform</td>
<td>540 ± 92</td>
<td>555 ± 84</td>
<td>0.65</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>404 ± 129</td>
<td>492 ± 161</td>
<td>0.11</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>528 ± 105</td>
<td>563 ± 153</td>
<td>0.41</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>498 ± 147</td>
<td>592 ± 201</td>
<td>0.14</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>650 ± 120</td>
<td>584 ± 155</td>
<td>0.16</td>
</tr>
<tr>
<td>Rectus</td>
<td>678 ± 135</td>
<td>576 ± 133</td>
<td>0.051</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>560 ± 97</td>
<td>648 ± 108</td>
<td>0.015 *</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>612 ± 101</td>
<td>576 ± 145</td>
<td>0.65</td>
</tr>
</tbody>
</table>

**Table 4.6:** Mean regional arrival time values in ms for voxelwise quantification, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where \( p < 0.05 \) are marked with *; where \( p < 0.01 \) are marked with **; where \( p < 0.001 \) are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Gradient-echo</th>
<th>Spin-echo</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>509 ± 89</td>
<td>572 ± 75</td>
<td>0.051</td>
</tr>
<tr>
<td>Amygdala</td>
<td>619 ± 187</td>
<td>688 ± 135</td>
<td>0.26</td>
</tr>
<tr>
<td>Calcarine</td>
<td>434 ± 204</td>
<td>552 ± 140</td>
<td>0.014 *</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>590 ± 178</td>
<td>636 ± 121</td>
<td>0.21</td>
</tr>
<tr>
<td>Fusiform</td>
<td>514 ± 107</td>
<td>574 ± 96</td>
<td>0.16</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>348 ± 124</td>
<td>411 ± 157</td>
<td>0.41</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>379 ± 129</td>
<td>444 ± 129</td>
<td>0.16</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>348 ± 125</td>
<td>449 ± 183</td>
<td>0.14</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>553 ± 138</td>
<td>543 ± 119</td>
<td>0.68</td>
</tr>
<tr>
<td>Rectus</td>
<td>334 ± 125</td>
<td>423 ± 110</td>
<td>0.042 *</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>485 ± 103</td>
<td>605 ± 90</td>
<td>0.0012 **</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>572 ± 107</td>
<td>550 ± 130</td>
<td>0.62</td>
</tr>
</tbody>
</table>
CHAPTER 4. COMPARISON OF GRADIENT-ECHO AND SPIN-ECHO MRI SEQUENCES IN VENTRAL BRAIN REGIONS

Table 4.7: Mean regional SNR values for global quantification, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where $p<0.05$ are marked with *; where $p<0.01$ are marked with **; where $p<0.001$ are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Gradient-echo</th>
<th>Spin-echo</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>1.44 ± 0.88</td>
<td>2.22 ± 1.26</td>
<td>0.05</td>
</tr>
<tr>
<td>Amygdala</td>
<td>0.49 ± 0.34</td>
<td>0.46 ± 0.27</td>
<td>0.90</td>
</tr>
<tr>
<td>Calcarine</td>
<td>1.34 ± 0.97</td>
<td>2.01 ± 1.22</td>
<td>0.07</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>1.46 ± 1.08</td>
<td>1.64 ± 1.00</td>
<td>0.48</td>
</tr>
<tr>
<td>Fusiform</td>
<td>1.38 ± 1/04</td>
<td>1.54 ± 0.95</td>
<td>0.62</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>1.05 ± 1.57</td>
<td>0.75 ± 0.36</td>
<td>0.21</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>0.90 ± 0.38</td>
<td>0.82 ± 0.37</td>
<td>0.51</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>0.56 ± 0.32</td>
<td>0.67 ± 0.44</td>
<td>0.59</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>1.11 ± 0.71</td>
<td>1.19 ± 0.66</td>
<td>0.59</td>
</tr>
<tr>
<td>Rectus</td>
<td>0.56 ± 0.23</td>
<td>0.69 ± 0.25</td>
<td>0.17</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>0.86 ± 0.62</td>
<td>0.76 ± 0.62</td>
<td>0.28</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>0.51 ± 0.39</td>
<td>0.18 ± 0.17</td>
<td>0.0092 *</td>
</tr>
</tbody>
</table>

Table 4.8: Mean regional SNR values for voxelwise quantification, showing the mean and standard deviation for the individual ROI values, and the p-value from a Wilcoxon ranked sum test. Values where $p<0.05$ are marked with *; where $p<0.01$ are marked with **; where $p<0.001$ are marked with ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Gradient-echo</th>
<th>Spin-echo</th>
<th>Wilcoxon p</th>
</tr>
</thead>
<tbody>
<tr>
<td>WholeBrain</td>
<td>1.31 ± 0.81</td>
<td>1.50 ± 1.62</td>
<td>0.48</td>
</tr>
<tr>
<td>Amygdala</td>
<td>0.50 ± 0.34</td>
<td>0.45 ± 0.26</td>
<td>1</td>
</tr>
<tr>
<td>Calcarine</td>
<td>1.34 ± 0.96</td>
<td>2.00 ± 1.22</td>
<td>0.068</td>
</tr>
<tr>
<td>Cerebellum-6</td>
<td>1.43 ± 1.08</td>
<td>1.65 ± 1.02</td>
<td>0.46</td>
</tr>
<tr>
<td>Fusiform</td>
<td>0.91 ± 0.81</td>
<td>1.05 ± 1.06</td>
<td>0.84</td>
</tr>
<tr>
<td>Front-Inf-Orb</td>
<td>0.97 ± 0.52</td>
<td>0.70 ± 0.34</td>
<td>0.17</td>
</tr>
<tr>
<td>Front-Med-Orb</td>
<td>0.85 ± 0.35</td>
<td>0.80 ± 1.06</td>
<td>0.77</td>
</tr>
<tr>
<td>Front-Sup-Orb</td>
<td>0.27 ± 0.26</td>
<td>0.58 ± 0.43</td>
<td>0.016</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>1.05 ± 0.68</td>
<td>1.18 ± 0.63</td>
<td>0.43</td>
</tr>
<tr>
<td>Rectus</td>
<td>0.41 ± 0.27</td>
<td>0.54 ± 0.17</td>
<td>0.062</td>
</tr>
<tr>
<td>Temp-Inf</td>
<td>0.58 ± 0.50</td>
<td>0.51 ± 0.80</td>
<td>0.68</td>
</tr>
<tr>
<td>Temp-Pole-Mid</td>
<td>0.39 ± 0.40</td>
<td>0.13 ± 0.20</td>
<td>0.031 *</td>
</tr>
</tbody>
</table>
Chapter 4. Comparison of gradient-echo and spin-echo MRI sequences in ventral brain regions

Figure 4.5: Histogram of mean regional CBF values in ml/100ml/min

Figure 4.6: Histogram of mean regional arrival time values in ms
CHAPTER 4. COMPARISON OF GRADIENT-ECHO AND SPIN-ECHO MRI SEQUENCES IN VENTRAL BRAIN REGIONS

Figure 4.7: Comparison of GE/SE regional CBF values in ml/100ml/min for individual subjects
CHAPTER 4. COMPARISON OF GRADIENT-ECHO AND SPIN-ECHO MRI SEQUENCES IN VENTRAL BRAIN REGIONS

(a) Global Quantification

(b) Voxelwise Quantification

Figure 4.8: Comparison of GE/SE regional arrival time values for individual subjects
Figure 4.9: Comparison of regional signal-to-noise factors between GE and SE images
4.4 Discussion

The study has considered the difference in CBF measurement between GE and SE sequences in regions of high susceptibility differences. There is a higher CBF value in ventral brain regions with SE sequences for both global and voxelwise $M_0$. This difference is significant for global correction, and less so for some regions for voxelwise correction. It is likely that the voxel-wise correction is removing some of the error due to susceptibility artefacts.

The mean whole-brain CBF of $27\pm8$ ml/100ml/min from the gradient-echo sequence) is lower than is found in other studies. Parkes et al [Parkes et al., 2004]) found values of $53\pm10$ ml/100ml/min in males, and $62\pm7$ ml/100ml/min in females using ASL. In a large multi-centre study Thade et al found mean grey matter CBF of $47.4\pm7.5$ ml/100ml/min [Thade et al., 2010], although they do comment that variation between centres may be due to differences in the labelling efficiency. The lower than normal CBF values found in this study may be due to variable labelling efficiency because of intermittent scanner problems.

This study has also considered two approaches to quantifying the CBF from a perfusion-weighted image, applying a global $M_0$ correction, and a voxelwise $M_0$ correction. Both have advantages and disadvantages, and there is no clear-cut decision as to which is better. The two approaches give different answers for CBF, especially in regions which are prone to susceptibility distortion and dropout, where the difference can be significant, as shown in tables 4.3, 4.4. The term is used in conjunction with $\lambda$, the blood-brain partition coefficient, and as I have used a global value of $\lambda$, using a global value of $M_0$ seems reasonable. However, the measured $M_0$ value also depends on the scanner characteristics, and so a voxelwise value will take account of these regional variations. This approach has its dangers, as where there is dropout both the measured $M_0$ value and the measured subtraction signal may be small. Using a voxelwise $M_0$ correction masks this problem, and hides regions where low SNR may give an inaccurate value for CBF. SNR is not routinely reported on. The white paper recently published by the ISMRM ASL working group recommends using a voxelwise correction, but first smoothing the $M_0$ map, e.g. with a Gaussian filter of 5-8 mm diameter [Alsop et al., 2015]. Whichever approach is chosen, it is important that the choice is documented.
4.4.1 Scanner problems

All the data for this study were acquired over a time-period of two months in between major rebuilds of the scanner. It is clear from the data that had to be rejected that the scanner was not producing consistently acceptable images. The calculated CBF is directly proportional to the labelling efficiency, and if this function was not performed consistently from image to image and subject to subject, a lower and more variable CBF would be found. This would be consistent with one subject having to be rejected entirely because there was no significant difference between label and control images. In view of this, what credence can the variation found between spin-echo and gradient-echo sequences have? Each data set was acquired in a single session lasting around 45 minutes, so the gradient-echo and spin-echo date for each subject will have been affected equally. In addition, the data have been cleaned, so that very noisy data have been rejected and not included in the analysis. The regional variations found are self-consistent; there is a good correlation between gradient-echo and spin-echo in regions not subject to susceptibility distortion, while in regions that are subject to susceptibility distortion the correlation coefficient is much lower. For global correction in regions not subject to susceptibility distortion the correlation between regional gradient-echo and spin-echo measurements of CBF in regions is 0.57, whereas in regions prone to susceptibility distortion it is 0.66. For voxelwise correction the correlation between regional gradient-echo and spin-echo measurements of CBF in regions not subject to susceptibility distortion is 0.67, whereas in regions prone to susceptibility distortion it is 0.31. No scanner malfunction is likely to produce such a consistent result across 15 subjects and 11 ROIs. Scatter plots of these values are shown in figure 4.10. Unfortunately the labelling efficiency provided by the scanner during this time-period was not measured, and cannot now be calculated from these data. A reduction in labelling efficiency of 50% would give the lower values of CBF found in this study.
Figure 4.10: Correlation between GE and SE values for CBF in regions prone to susceptibility distortion and regions not so prone. 4.10(a) shows a scatter plot of GE CBF against SE CBF calculated with global correction for regions not prone to susceptibility distortion. 4.10(b) shows a scatter plot of GE CBF against SE CBF calculated with global correction for regions prone to susceptibility distortion. 4.10(c) shows a scatter plot of GE CBF against SE CBF calculated with voxelwise correction for regions not prone to susceptibility distortion. 4.10(d) shows a scatter plot of GE CBF against SE CBF calculated with voxelwise correction for regions prone to susceptibility distortion.
4.4.2 Conclusion

This study shows that, in ROIs prone to susceptibility distortion, spin-echo gives a higher value of CBF than gradient-echo. In ROIs not prone to susceptibility distortion the measured CBF is comparable. It is reasonable to believe that the lower measured CBF shown by the gradient-echo readout is an artefact due to signal loss, and therefore that the spin-echo readout gives a more accurate measurement. This is of importance when considering regional CBF values in these regions, which are implicated in the early stages of FTLD.
Chapter 5

Regional atrophy in frontotemporal lobar degeneration and relationship to cognitive deficits

5.1 Introduction

Frontotemporal lobar degeneration (FTLD) has already been described in chapter 1.2. In this chapter I describe the results of T₁-weighted MR images of a group of patients with FTLD and age-matched controls. The results of a voxel-based group analysis are described, and compared with the results of a region-of-interest analysis. I describe the results of a voxel-based regression analysis of MR images against some patient neuropsychological test scores. I also describe the neuropsychological test battery used. The voxel-based analyses have been described and conducted before in other studies of FTLD patients (see chapter 2.6.2). The ROI analysis is not new, although less commonly used than voxel-based analyses.

The principal aim of this thesis is to distinguish people with FTD from controls, but it is also interesting to consider the pattern of atrophy, as giving some indication of the differential diagnosis. There is considerable heterogeneity among FTD patients, and some brain regions are abnormal in some patients but not in others. Methods which can pick out these regions may lead to a more accurate differential diagnosis. Mummery et al [Mummery et al., 2000] found greater left
inferior anterior temporal lobe atrophy in patients with lower naming scores, and so I have chosen to investigate the result of this test in this group of patients, to see if the link is also found. Moll et al [Moll et al., 2011] found that patients with greater septal region atrophy exhibited a poorer ability to experience context-appropriate guilt, and so I have chosen to explore this. Zahn et al [Zahn et al., 2004] reports a link between greater left temporal atrophy and poorer non-verbal semantic knowledge. I therefore chose to explore the effect of the pyramids and palm trees test score, which tests this. Snowden et al [Snowden et al., 2004] reported a link between face recognition and atrophy in the right temporal lobe, so I have explored the link in this study of atrophy with the recognition part of the famous faces test.

5.2 Methods

The study is approved by South Manchester NHS Research Ethics Committee, reference 07/H1003/194 South Manchester REC. 20 patients with features of bvFTD or SD were recruited and took part in the first day of diagnostic testing; 3 were not included because they did not meet the study’s diagnostic criteria. 28 controls were recruited but 10 were not included because they did not meet the cognitive criteria for controls or did not have normal-for-age MRI images. 17 patients (mean age 66.0 ± 7.8y, 8 male) and 18 controls (mean age 62.6 ± 10.5y, 12 male) completed the neuropsychological tests and had T1- and T2-weighted images.

This study is part of a larger study into the neural basis of disorders of social knowledge. Earlier results have already been reported by Green [Green et al., 2010], [Green, 2011]. Green initially conducted all neuropsychological testing and supervised MR images of controls; I took over these tasks in 2010. The study has ethical approval from South Manchester NHS Ethics Committee. Patients and controls were excluded if they:

- were acutely suicidal
- had impairments of vision or hearing which could not be corrected during the experiment
- had a history of learning disabilities or developmental disorders or of manic
or hypomanic episodes, of schizophreniform symptoms or schizophrenia, of substance abuse, repeated self injuries, severe obsessive-compulsive disorder, post-traumatic stress disorder, manifest eating disorder, primary anxiety disorder with only secondary depression

- had a history of major medical disorders (significant heart insufficiency, severe chronic obstructive pulmonary disease, uncontrolled hypertension or diabetes, endangiitis obliterans, severe vascular encephalopathy, hypo- or hyperthyroidism, severe liver or kidney disorders, rheumatoid disorders and all other medical conditions affecting brain function, blood flow or metabolism).

Controls were also excluded if they had a history of neurological disorders.

Patients were recruited by Zahn from his memory clinic at the South Manchester University Hospital or by referral from other consultants. One patient was recruited from outside Greater Manchester by self-referral from a website contact. Patients were seen twice; once in their home, where Zahn took a clinical history and conducted a brief neurological examination and Green or myself conducted pencil-and-paper neuropsychological testing. Patients were subsequently seen at Salford Royal Infirmary, where MR images were acquired and computer-based neuropsychological testing was done. Before both visits all patients provided informed written consent if they had the capacity; otherwise a relative, most frequently their spouse, provided a consultee declaration on their behalf. All patients fulfilled Lund-Manchester criteria for SD or bvFTD [Neary et al., 1998].

Controls were recruited from various sources around Greater Manchester, e.g. science clubs, Alzheimer’s Society meetings. Suitable controls provided written informed consent and were recompensed for their travel. They were invited for neuropsychological testing at the University of Manchester. Controls who showed no abnormal cognitive deficits were invited to Salford Royal Infirmary for MRI images and further neuropsychological testing. The MRI images of all controls were examined for abnormalities.

5.2.1 MRI sequences

All images were acquired with a Philips Achieva 3T scanner using an 8-channel head-coil. The $T_1$-weighted image was a high-resolution 3D MPRAGE image, in-
version time 1150ms, 256x164 matrix, 128 slices, voxel size 1mmx1mmx1mm, TE 3.8ms, TR 9.4ms, flip angle $8^\circ$. The $T_2$-weighted image was acquired using a turbo-spin-echo sequence with SENSE factor 2. Matrix size was 512x408, 44 slices, voxel size 0.26mmx0.26mmx3.30mm, TE 80ms, TR 3s, flip angle $90^\circ$. This image was taken to exclude subjects with gross cardiovascular abnormality, but was not used otherwise.

### 5.2.2 Neuropsychological tests

The cognitive tests used in the FTD study were conducted in two sessions. The first session consisted of a battery of pen-and-paper tests used to confirm the diagnosis of patients, or the cognitive status of controls. These tests are described in appendix D. The second battery of tests was run after the MRI scan in a separate room at Salford Royal Infirmary. All of these tests were run on a laptop. For all tasks, the examiner read the information on the screen, and the participant was asked to point to their chosen response on the screen. The examiner then entered the response on the computer. These tests are:

- **Visual discrimination task** [Green, 2011].
  This test is a non-verbal discrimination task, designed to control for the task demands of the subsequent tests. The layout is similar to the pyramids and palmtrees task, in that there is an abstract drawing at the top, and two similar drawings at the same level below. All objects are abstract line drawings; the two lower drawings have been modified from the top; one has been modified more than the other. The participant is asked to select the lower drawing which is more related to the top drawing.

- **Social Concept discrimination Task** [Zahn et al., 2009a].
  This test assesses impairments in conceptual knowledge about social behaviours. The layout is similar to the pyramids and palmtrees test, but the stimulus (top object) is a word that represents a social concept (e.g. "humble") or an animal function concept (e.g. "trainable"). One of the two lower words is related to the stimulus, and the other is a distractor e.g. "(polite", "sympathetic", "ridden", "bites"). The participant is asked to select which of the lower words is more related to the stimulus.
• Consequences of Social Action discrimination Task [Green, 2011].

This test assesses knowledge of the short- and long-term consequences of actions. The layout is similar to the pyramids and palmtrees test, in that there is a stimulus at the top of the screen, and two possible consequences below. The stimulus is a brief description of an action (e.g. “you give a gift to your friend while shopping”) and the possible consequences are more or less related to the stimulus (e.g. “you become close friends”, “you get married”). The examiner reads the stimulus followed by the first consequence, and then the stimulus followed by the second consequence, and the participant is asked to select which of the two consequences is more related to the stimulus. Consequences can be positive or negative, and short- or long-term.

• Social Action evaluation Task - FTD version [Green, 2011].

This test assesses the evaluation of social interactions. The participant is first asked for the name of their best friend (BF). They are then presented with a series of possible interactions between themselves and the friend: some are positive (e.g. “you are loyal to BF”) or negative (e.g. “BF is disagreeable towards you”). The participant is then asked to choose whether this action would make them feel good or bad, and which of the emotions guilt, gratitude, anger or pride they would feel most strongly. One-quarter of the possible actions are positive where the participant is acting (positive self-agency); one-quarter are positive where BF is acting (positive other-agency); one-quarter are negative other-agency, and one-quarter negative self-agency. The expectation is that guilt will be associated with bad feelings and negative self-agency, gratitude with good feelings and positive other-agency, anger with bad feelings and negative other-agency and pride with good feelings and positive self-agency. The sensitivity index $d'$ score is calculated from the proportion of times the target emotion is selected appropriately (true positive rate or TPR), and the number of times the target emotion is selected inappropriately (false positive rate or FPR) [Brophy, 1986]. A higher value of $d'$ indicates a more accurate selection. In cases where the participant scored zero or 100%, the score is either increased or decreased by half a unit to allow a $d'$ to be calculated, e.g. 0/10 is scored as 0.5/10, and 15/15 is scored as 14.5/15 [Durrant et al., 2013].
5.2.3 Analysis

5.2.3.1 Voxel-based analysis

The VBM8 toolbox in SPM8 was used to determine regional volume differences between the patient and control groups by a process of normalisation, grey matter segmentation, and voxel-wise t-test. The FTLD patients were normalised as one group, and the controls as another group to the MNI template within the Dartel toolbox. Modulation was applied, corrected the warped (normalised) tissue intensities so that their regional volume is preserved. The normalised voxel size was 1.5mm x 1.5mm x1.5mm. Default parameters were used except that the option Clean up any partitions was set to “Thorough Cleanup”, as is recommended for atrophied brains. The normalised and segmented grey matter mask for each brain was scrutinised to see if there were any obvious problems with the normalisation, particularly in the case of the very atrophied brains of some of the FTLD patients. The grey matter masks were smoothed with a Gaussian filter and FWHM of 12mm. This smoothing kernel was used throughout the study, and so was used here for compatibility with the analysis of perfusion- and diffusion-weighted scans. With hindsight a smaller kernel of 6 to 8mm might have been better, reflecting the greater resolution of the T1-weighted images. These smoothed images were used as input to the voxel-based modelling analysis. A 2-group t-test was used, with patients as one group and controls as another, with no multi-comparison correction, significance threshold of \( p < 0.001 \) and minimum cluster size 75. The significance level is as recommended by Lieberman and Cunningham [Lieberman and Cunningham, 2009]; their minimum recommended cluster size of 10 voxels with voxel size \( 3.5 \times 3.5 \times 5 \) mm\(^3\) corresponds to a minimum cluster size of 75 voxels where (as here) the voxel size is \( 2 \times 2 \times 2 \) mm\(^3\). Individual analyses were also run, comparing a single patient against the group of controls, with the same analysis factors.

Multiple regression tests were run, with age included as a potential confounder, and results from some of the cognitive tests as covariates of interest. Only one test of interest was included at any time. The aim of these tests was see if correlations found by others were confirmed in these data. Cognitive tests used were naming [Mummery et al., 2000], pyramids and palmtrees [Zahn et al., 2004], famous faces(recognition) [Snowden et al., 2004], and guilt from the evaluation of social actions [Moll et al., 2011]. Controls were not included in these analyses as
they all scored close to the top score. These tests are independent, and so no correction for the number of regression tests was made. Cluster-based correction was made within each regression for the number of voxels tested, as stated within the individual results.

5.2.3.2 Regional analysis

Grey and white matter densities were also compared on a regional basis, by taking the average GM/WM probability density within a given ROI. ROIs were taken from the automated anatomical labeling atlas (aal) [Tzourio-Mazoyer et al., 2002], from the JHU White Matter Tractography Atlas [Mori et al., 2008], and from Zahn et al [Zahn et al., 2009b]. Right(left) posterior and anterior temporal lobe regions were defined by adding the Temporal-Sup-R(L), Temporal-Pole-Sup-R(L), Temporal-Mid-R(L), Temporal-Pole-Mid-R(L), and Temporal-Inf-R(L), and then splitting the regions along the line y=-11. These ROIs are the right/left anterior/posterior temporal lobe (RAntTL, LAntTL, RPostTL, LPostTL). The ventromedial prefrontal cortex described by Zahn et was split along the line y=47 into anterior and posterior (AntVMPFC, PostVMPFC). Control ROIs unlikely to be affected in the illness were taken by adding the aal right and left calcarine (Calc) and cerebellum6 (Cereb6). Regions taken from the JHU White Matter Tractography Atlas were the part of the cingulum bundle connected to the cingulate gyrus (CCg), the part of the cingulum bundle connected to the cingulum-hippocampus (CH), the inferior-fronto-occipital fasciculus (IFOF), the inferior-longitudinal fasciculus (ILF), the superior-longitudinal fasciculus (SLF) and the uncinate fasciculus (UF). For all these regions left and right values were calculated separately. The JHU White Matter Tractography Atlas is a probabilistic atlas, giving probabilities that a certain voxel is included within a particular tract, and so voxels were included within a region at 2 minimum probability levels of 10% and 20%. These two sets of JHU ROIs were used as they give slightly different results, particularly when used in the multimodal analysis (see chapter 8). The expectation is that ROIs defined with the lower probability threshold will include more voxels on the edge of the WM tracts, while those ROIs defined with the higher probability threshold will have voxels within the centre of the tracts. Some voxels will be included in more than one ROI, in particular the inferior-fronto-occipital fasciculus overlaps with the inferior-longitudinal fasciculus and the uncinate fasciculus. Where such over-
lapses occur the voxels have been included in both ROIs.

All of these ROIs are quite large, and include both grey matter, white matter and CSF (see, for example, figure 5.1, which shows the left posterior temporal ROI and the left uncinate fasciculus overlain on the SPM T1 canonical brain), so I have looked at both GM and WM densities in all regions when considering a value which can be used to classify the subjects. I have also considered the voxels within each ROI which have a GM/WM probability of >80%, in order to investigate whether or not there are any differences between patients and controls in the tissue within the ROIs. The aal ROIs used are shown in figure 5.2, and the JHU ROIs thresholded at 20% in figure 5.3. For clarity, the colour coding for JHU ROIs is the same for both left and right regions.

![Figure 5.1: Sample ROIs overlain on a canonical T1-weighted image](image)

Figure 5.1: Sample ROIs overlain on a canonical T1-weighted image. Figure 5.1(a) shows the left posterior temporal ROI, figure 5.1(b) shows the uncinate fasciculus ROI. It can be seen that both ROIs include both grey matter and white matter.
Chapter 5. Regional atrophy in frontotemporal lobar degeneration and relationship to cognitive deficits

Figure 5.2: aal regions of interest from the aal atlas used in analyses

Figure 5.3: Regions of interest taken from the JHU White Matter Tractography Atlas used in analyses
5.3 Results

5.3.1 Results of cognitive tests

The results of the Day 1 cognitive tests are shown in appendix D, and the Day 2 tests in table 5.1. Tests which do not distinguish between patients and controls are the visual discrimination test (an expected result as it was designed to test whether patients understood instructions for the this and other tests), the short-term negative section of the consequences of social action test and the anger and guilt sections of the social action evaluation test. No correction has been made for multiple comparisons as it is expected that the tests are independent, and no hypothesis is being tested here.

<table>
<thead>
<tr>
<th>Test</th>
<th>FTLD patient</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Discrimination</td>
<td>21.6±2.3</td>
<td>21.6±1.5</td>
<td>0.96</td>
</tr>
<tr>
<td>Consequences of Social Action</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Only available for 12 patients)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% correct</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term negative</td>
<td>65.0±17.0</td>
<td>82.0±8.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Long-term positive</td>
<td>76.1±20.6</td>
<td>92.2±2.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Short-term negative</td>
<td>85.0±26.4</td>
<td>98.0±3.8</td>
<td>0.63</td>
</tr>
<tr>
<td>Short-term positive</td>
<td>81.6±26.3</td>
<td>97.6±3.9</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Social Concept Discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal % correct</td>
<td>82.8±12.9</td>
<td>90.0±4.9</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Social % correct</td>
<td>70.1±13.4</td>
<td>93.3±4.6</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Difference % correct</td>
<td>-12.7±9.2</td>
<td>-3.4±6.1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Social Action Evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anger d’ score</td>
<td>1.78±1.00</td>
<td>1.85±0.90</td>
<td>0.83</td>
</tr>
<tr>
<td>guilt d’ score</td>
<td>1.84±1.02</td>
<td>0.09±0.70</td>
<td>0.13</td>
</tr>
<tr>
<td>pride d’ score</td>
<td>1.35±0.83</td>
<td>2.23±1.09</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>gratitude d’ score</td>
<td>1.52±0.85</td>
<td>2.64±0.71</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
5.3.2 Normalisation

The brains of some of the FTLD patients have a high level of regional atrophy and deformed ventricles, leading to doubt over whether the normalisation process used to transform these brains into MNI space would give reliable results, despite the use of DARTEL and a group-specific template. Figure 5.4 show the original T$_1$-weighted brain image, and the normalised modulated grey matter map for one of the most atypical brains. No subjects were rejected because of normalisation problems.

![Initial and normalised images for an FTLD patient.](image)

Figure 5.4: Initial and normalised images for an FTLD patient. Figure 5.4(a) shows the initial T$_1$-weighted image; figure 5.4(b) shows the normalised image. It can be seen that the normalisation is acceptable.

5.3.3 Voxel-based comparison of patient and control groups

The results of a Voxel-based group comparison between 17 FTD patients and 18 controls are shown in figure 5.5 with no correction for multiple comparisons, p<0.001 and minimum cluster size of 75 voxels; several views are shown to display the extent of the grey matter loss. Although it is of no physiological interest, since there are not expected to be any such regions, I have also included the regions where patients have more grey matter than controls, because the comparison has been done at a high p-value, and hence there may be false positives. Regions where patients have less grey matter than controls are the subgenual cingulate and medial orbitofrontal cortex, the right lateral orbitofrontal and right dorsolateral prefrontal cortex, septal region, basal forebrain, midbrain, and bilateral anterior temporal...
lobes spreading into posterior temporal areas, bilateral insula. Regions where patients have less atrophy then controls are small regions around the posterior ventricles. The regions where the comparison survives cluster-based multiple-comparison correction, with cluster-level \( p < 0.05 \) and minimum cluster size of 4 voxels, are shown in figure 5.6; no regions where patients have a greater volume than controls survive the more stringent test. Table 5.2 shows the SPM output for the results of the comparison with cluster-based multiple-comparison correction.

![Image](image_url)

**Figure 5.5**: Results of group comparison between patients and controls: no correction for multiple comparisons, \( p < 0.001 \), minimum cluster size 75 voxels, coronal and axial views, slices 5mm apart.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster(k)</th>
<th>( p ) (cFWE corr)</th>
<th>peak T</th>
<th>( p ) (uncorr)</th>
<th>MNI local maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left hippocampus, fusiform, etc</td>
<td>11423</td>
<td>(&lt;0.001)</td>
<td>9.4</td>
<td>(&lt;0.001)</td>
<td>(-36\ -33\ -6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(&lt;0.001)</td>
<td>7.8</td>
<td>(&lt;0.001)</td>
<td>(-30\ 0\ -17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(&lt;0.001)</td>
<td>7.5</td>
<td>(&lt;0.001)</td>
<td>(-38\ -11\ -41)</td>
</tr>
<tr>
<td>Right hippocampus, fusiform, etc</td>
<td>17211</td>
<td>(&lt;0.001)</td>
<td>9.2</td>
<td>(&lt;0.001)</td>
<td>21   -2  -11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(&lt;0.001)</td>
<td>8.3</td>
<td>(&lt;0.001)</td>
<td>20   -23  -6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(&lt;0.001)</td>
<td>8.0</td>
<td>(&lt;0.001)</td>
<td>36   -33  -2</td>
</tr>
<tr>
<td>R thalamus</td>
<td>139</td>
<td>(&lt;0.05)</td>
<td>6.0</td>
<td>(&lt;0.001)</td>
<td>0   -17  3</td>
</tr>
</tbody>
</table>

**Table 5.2**: Clusters significant after correction for multiple comparisons. \( p \) (FWE-corr) has been cluster-based corrected for multiple comparisons, peak T, \( p \) (uncorrected) and MNI coordinates for all clusters where FTLD patients have less grey matter volume than controls.
Figure 5.6: Results of group comparison between patients and controls: cluster-based multiple-comparison corrected, $p$(FWE-corr)$< 0.05$, coronal and axial views, slices 5mm apart.
5.3.4 Results of voxel-based regression of patient group with individual cognitive tests

5.3.4.1 Naming

This analysis included 15 patients. One patient was not tested as he had recently completed a very similar test as part of another study; one patient did not complete the test because of incorrect paperwork by myself. Figure 5.7 shows the result of a voxel-based regression of patient images against their score in the Cambridge naming test, with no correction for multiple comparisons, \( p < 0.001 \), and minimum cluster size 75 voxels. The regression included age as a nuisance factor and the naming score as the factor of interest. Analysis was carried out within a region defined by the aal regions of left temporal lobe, left temporal mid and left temporal pole mid, and the p-value shown by each cluster is the regional cluster-based corrected value. There is a large cluster of 1350 voxels along the posterior left temporal lobe with peaks at MNI coordinates \([-54 \ -51 \ -8]\) with \( p=0.05 \), \( T=4.1 \), and \([-36 \ -40 \ -17]\) with \( p=0.05 \), \( T=4.1 \), and another small cluster of 140 voxels at \([-41 \ -19 \ -33]\) with \( p=0.71 \), \( T=3.4 \).

![Figure 5.7: Results of regression of patient images against Naming scores: no correction for multiple comparisons, \( p < 0.001 \), minimum cluster size 75 voxels](image)
5.3.4.2 Pyramids and palmtrees

This analysis included 16 patients. One patient did not complete the test because of incorrect paperwork by myself. Figure 5.8 shows the result of a voxel-based regression of patient images against their scores in the pyramids and palmtrees test, with no correction for multiple comparisons, $p < 0.001$, and minimum cluster size 75 voxels. It shows a large cluster along the left temporal lobe, (larger than for Naming), with a cluster of 5900 voxels stretching from $[-35 -39 -15]$ through $[-26 -44 -11]$ to $[-27 -26 -9]$, cluster-based multiple comparison corrected $p$-value=0.019. There are other clusters which do not survive cluster-based correction: one is a cluster of 1970 voxels with peaks at $[-44 -15 -45]$ and $[-57 -2 -41]$ with cluster-based multiple comparison corrected $p$=0.67, and another, smaller cluster in the right temporal lobe at $[-32 -2 9]$ with cluster-based multiple comparison corrected $p$=0.87.

![Figure 5.8: Results of regression of patient images against pyramids and palmtrees scores: no correction for multiple comparisons, $p < 0.001$, minimum cluster size 75 voxels](image)
5.3.4.3 Famous faces recognition

This analysis included 15 patients. One patient was not tested as he had recently completed this test as part of another study; one patient did not complete the test because of incorrect paperwork by myself. No significant clusters were found.

5.3.4.4 Evaluation of social actions

These analyses included 16 patients. No test data was available for 1 patient as the test was not introduced until partway through the study. Figure 5.9 shows the result of a regression of patient images against the guilt and anger scores from the Evaluation of Social Action tests. The two tests are both included as they are closely linked, but the results for guilt treat anger as a confounding variable, and vice versa. If the guilt analysis is carried out for an ROI defined by the aal septal ROI, there is a peak at [5 5 9] with cluster-based corrected p-value=0.10. There are further peaks at [23 48 29] (460 voxels), [0 5 -21] (183 voxels), [9 0 -15] (154 voxels) and [24 -29 -32] (109 voxels). N.B. Moll et al used a septal ROI which was manually drawn [Moll et al., 2011].

The anger analysis shows peaks of 5500 voxels at [-24 -37 -57 stretching to [-39 -73 -38], 800 voxels at [54 -60 39], 450 voxels at [53 44 -5] and 1550 voxels at [6 -85 -36]. None of these results survive correction for multiple comparisons. A similar analysis on the scores for pride and gratitude showed no differences at p< 0.005, minimum cluster size 75 voxels.
CHAPTER 5. REGIONAL ATROPHY IN FRONTOTEMPORAL LOBAR DEGENERATION AND RELATIONSHIP TO COGNITIVE DEFICITS

Figure 5.9: Results of regression of patient images against the guilt and anger scores from the evaluation of social action tests: no correction for multiple comparisons, \( p < 0.001 \), minimum cluster size 75 voxels
5.3.5 Regional analysis

The ROIs used for these analyses have been described in chapter 5.2.3.2. The p-values shown in tables are uncorrected for multiple comparisons; the text indicates which tests are still significant after such correction.

Table 5.3 shows the GM density in the aal ROIs. The leftmost columns show the mean GM density for all voxels within the ROI, for patients and controls, with the result of a Wilcoxon ranked sum test. The rightmost columns show the mean GM density for only voxels where the GM density is > 80%. It can be seen that when all the voxels within the ROI are considered patients have a significantly lower GM density then controls in all the aal ROIs expected to be affected in the illness at a significance level < 0.05, and no difference in the ROIs not expected to be affected; this difference survives a Bonferroni multiple comparison correction (which for these 6 tests would reduce the p-value to < 0.008) for all regions except the anterior ventromedial prefrontal cortex. When only voxels with a GM density > 0.8 are considered, there is still significantly lower GM density in all the temporal ROIs and in the posterior ventromedial prefrontal cortex, although not in the anterior ventromedial prefrontal cortex. There is a significant difference after Bonferroni correction in the right and left posterior temporal lobes, and in the right anterior temporal lobe.

Table 5.4 shows the WM density in the aal ROIs. The leftmost columns show the mean WM density for all voxels within the ROI, for patients and controls, with the result of a Wilcoxon ranked sum test. The rightmost columns show the mean WM density for only voxels where the WM density is > 80%. It can be seen that when all the voxels within the ROI are considered patients have a significantly lower GM density then controls at a significance level < 0.05 in the right posterior temporal lobe, and the left anterior and posterior temporal lobe. These differences are not significant after Bonferroni correction. There is no difference in the ROIs not expected to be affected. When only voxels with a WM density > 0.08 are considered, there are no significant differences between patients and controls in any ROI.

Tables 5.5 and 5.6 show the mean GM density in JHU ROIs which have been defined with a probability > 20% and > 10% respectively. There are slight differences between the tables, but patients have significantly lower GM density in the
right and left inferior fronto-occipital fasciculus, right and left inferior longitudinal fasciculus, the right and left superior longitudinal fasciculus and the right and left uncinate fasciculus. Bonferroni correction with 10 tests would lower the significant probability to 0.004, but all regions except that right and left superior longitudinal fasciculus would still show significantly lower GM density in patients. The GM densities are slightly higher in the ROIs thresholded at a probability level of 10% than they are in the ROIs thresholded at a probability level of 20%. When only voxels with a GM density > 0.8 are considered there are only lower GM densities in patients in the left inferior fronto-occipital fasciculus, left inferior longitudinal fasciculus, and the right superior longitudinal fasciculus; these differences are not significant after Bonferroni correction. When only these voxels are considered, there is little difference between the two sets of ROIs.

Tables 5.7 and 5.8 show the mean WM density in JHU ROIs which have been defined with a probability > 20% and > 10% respectively. There are slight differences between the tables, but patients have significantly lower WM density in the right and left inferior longitudinal fasciculus; these differences do not survive Bonferroni correction. The WM densities are slightly lower in the ROIs thresholded at a probability level of 10% than they are in the ROIs thresholded at a probability level of 20%. When only voxels with a WM density > 0.8 are considered there is only lower WM density in patients in the left inferior longitudinal fasciculus in the ROI thresholded at probability > 20%, and this does not survive Bonferroni correction. When only these voxels are considered, there is little difference between the two sets of ROIs.

Table 5.3 and 5.4 show respectively the GM and WM density in the aal ROIs, which are predominantly cortical.
Table 5.3: Grey matter density values in aal ROIs for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 17 patients and 18 controls. Values for raw grey matter density and grey matter density with where GM probability is > 80% are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw GM density</th>
<th>p</th>
<th>GM &gt; 80%</th>
<th></th>
<th>p</th>
<th>GM &gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>Controls</td>
<td>Patients</td>
<td></td>
<td>Controls</td>
<td></td>
</tr>
<tr>
<td>LPostTL</td>
<td>0.42±0.10</td>
<td>&lt; 10^-4</td>
<td>0.56±0.05</td>
<td>0.91±0.04</td>
<td>&lt; 10^-4</td>
<td>0.96±0.03</td>
</tr>
<tr>
<td>RPostTL</td>
<td>0.45±0.07</td>
<td>&lt; 10^-3</td>
<td>0.54±0.06</td>
<td>0.91±0.03</td>
<td>&lt; 10^-3</td>
<td>0.96±0.03</td>
</tr>
<tr>
<td>LAntTL</td>
<td>0.35±0.15</td>
<td>&lt; 10^-3</td>
<td>0.56±0.05</td>
<td>0.93±0.07</td>
<td>0.02</td>
<td>0.97±0.04</td>
</tr>
<tr>
<td>RAntTL</td>
<td>0.37±0.11</td>
<td>&lt; 10^-4</td>
<td>0.54±0.06</td>
<td>0.91±0.05</td>
<td>&lt; 10^-3</td>
<td>0.98±0.05</td>
</tr>
<tr>
<td>PostVMPFC</td>
<td>0.45±0.08</td>
<td>&lt; 10^-3</td>
<td>0.57±</td>
<td>0.93±0.05</td>
<td>0.01</td>
<td>0.98±0.05</td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>0.44±0.05</td>
<td>0.03</td>
<td>0.48±0.05</td>
<td>0.90±0.04</td>
<td>0.16</td>
<td>0.92±0.04</td>
</tr>
<tr>
<td>Calcarine</td>
<td>0.49±0.05</td>
<td>0.68</td>
<td>0.50±0.08</td>
<td>0.93±0.03</td>
<td>0.96</td>
<td>0.93±0.04</td>
</tr>
<tr>
<td>Cereb6</td>
<td>0.67±0.08</td>
<td>0.10</td>
<td>0.72±0.08</td>
<td>0.96±0.05</td>
<td>0.68</td>
<td>0.96±0.05</td>
</tr>
</tbody>
</table>

Table 5.4: White matter density values in aal ROIs for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 17 patients and 18 controls. Values for raw white matter density and white matter density with WM probability > 80% are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw WM density</th>
<th>p</th>
<th>WM &gt; 80%</th>
<th></th>
<th>p</th>
<th>WM &gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>Controls</td>
<td>Patients</td>
<td></td>
<td>Controls</td>
<td></td>
</tr>
<tr>
<td>LPostTL</td>
<td>0.39±0.05</td>
<td>0.18</td>
<td>0.42±0.03</td>
<td>1.04±0.07</td>
<td>0.40</td>
<td>1.06±0.05</td>
</tr>
<tr>
<td>RPostTL</td>
<td>0.36±0.04</td>
<td>0.02</td>
<td>0.39±0.03</td>
<td>1.01±0.06</td>
<td>0.06</td>
<td>1.05±0.05</td>
</tr>
<tr>
<td>LAntTL</td>
<td>0.30±0.09</td>
<td>0.00</td>
<td>0.38±0.03</td>
<td>1.02±0.12</td>
<td>0.31</td>
<td>1.05±0.07</td>
</tr>
<tr>
<td>RAntTL</td>
<td>0.27±0.07</td>
<td>0.01</td>
<td>0.33±0.04</td>
<td>0.98±0.11</td>
<td>0.14</td>
<td>1.03±0.06</td>
</tr>
<tr>
<td>PostVMPFC</td>
<td>0.53±0.05</td>
<td>0.23</td>
<td>0.55±0.04</td>
<td>1.02±0.07</td>
<td>0.10</td>
<td>1.05±0.07</td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>0.47±0.05</td>
<td>0.73</td>
<td>0.47±0.04</td>
<td>1.06±0.09</td>
<td>0.36</td>
<td>1.08±0.07</td>
</tr>
<tr>
<td>Calcarine</td>
<td>0.43±0.04</td>
<td>0.33</td>
<td>0.42±0.03</td>
<td>1.02±0.07</td>
<td>0.68</td>
<td>1.01±0.05</td>
</tr>
<tr>
<td>Cereb6</td>
<td>0.24±0.03</td>
<td>0.10</td>
<td>0.22±0.02</td>
<td>1.00±0.09</td>
<td>0.13</td>
<td>0.95±0.04</td>
</tr>
</tbody>
</table>
Table 5.5: Grey matter density values in JHU ROIs thresholded at probability > 20% for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 17 patients and 18 controls. Values for raw grey matter density and grey matter density with GM probability > 80% are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Where no value is shown there are no voxels within the ROI with GM probability > 80%. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw GM density</th>
<th>GM &gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LCCg</td>
<td>0.27±0.03</td>
<td>0.48</td>
</tr>
<tr>
<td>RCCg</td>
<td>0.31±0.05</td>
<td>0.52</td>
</tr>
<tr>
<td>LCH</td>
<td>0.15±0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>RCH</td>
<td>0.30±0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>LIFOF</td>
<td>0.26±0.04</td>
<td>&lt; 10^{-5}</td>
</tr>
<tr>
<td>RIFOF</td>
<td>0.22±0.03</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>LILF</td>
<td>0.20±0.03</td>
<td>&lt; 10^{-3}</td>
</tr>
<tr>
<td>RILF</td>
<td>0.16±0.03</td>
<td>&lt; 10^{-3}</td>
</tr>
<tr>
<td>LSLF</td>
<td>0.29±0.04</td>
<td>0.52</td>
</tr>
<tr>
<td>RSLF</td>
<td>0.28±0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>LUF</td>
<td>0.34±0.07</td>
<td>&lt; 10^{-5}</td>
</tr>
<tr>
<td>RUF</td>
<td>0.32±0.08</td>
<td>&lt; 10^{-5}</td>
</tr>
</tbody>
</table>

Table 5.6: Grey matter density values in JHU ROIs thresholded at probability > 10% for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 17 patients and 18 controls. Values for raw grey matter density and grey matter density with GM probability > 80% are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Where no value is shown there are no voxels within the ROI with GM probability > 80%. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw GM density</th>
<th>GM &gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LCCg (10pc)</td>
<td>0.28±0.03</td>
<td>0.80</td>
</tr>
<tr>
<td>RCCg (10pc)</td>
<td>0.36±0.05</td>
<td>0.08</td>
</tr>
<tr>
<td>LCH (10pc)</td>
<td>0.31±0.03</td>
<td>0.48</td>
</tr>
<tr>
<td>RCH (10pc)</td>
<td>0.41±0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>LIFOF (10pc)</td>
<td>0.28±0.03</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>RIFOF (10pc)</td>
<td>0.28±0.03</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>LILF (10pc)</td>
<td>0.26±0.04</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>RILF (10pc)</td>
<td>0.26±0.04</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>LSLF (10pc)</td>
<td>0.34±0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>RSLF (10pc)</td>
<td>0.31±0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>LUF (10pc)</td>
<td>0.33±0.07</td>
<td>&lt; 10^{-5}</td>
</tr>
<tr>
<td>RUF (10pc)</td>
<td>0.33±0.08</td>
<td>&lt; 10^{-5}</td>
</tr>
</tbody>
</table>
Table 5.7: White matter density values in JHU ROIs thresholded at probability > 20% for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 17 patients and 18 controls. Values for raw white matter density and white matter density with WM probability > 80% are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Where no value is shown there are no voxels with WM probability > 80%. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw WM density</th>
<th>WM &gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LCCg</td>
<td>0.79±0.09</td>
<td>0.35</td>
</tr>
<tr>
<td>RCCg</td>
<td>0.68±0.09</td>
<td>0.31</td>
</tr>
<tr>
<td>LCH</td>
<td>0.91±0.19</td>
<td>0.28</td>
</tr>
<tr>
<td>RCH</td>
<td>0.69±0.19</td>
<td>0.40</td>
</tr>
<tr>
<td>LIFOF</td>
<td>0.83±0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>RIFOF</td>
<td>0.82±0.08</td>
<td>0.19</td>
</tr>
<tr>
<td>LILF</td>
<td>0.81±0.12</td>
<td>0.01</td>
</tr>
<tr>
<td>RILF</td>
<td>0.83±0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>LSLF</td>
<td>0.83±0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>RSLF</td>
<td>0.86±0.07</td>
<td>0.42</td>
</tr>
<tr>
<td>LUF</td>
<td>0.64±0.12</td>
<td>0.73</td>
</tr>
<tr>
<td>RUF</td>
<td>0.58±0.16</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Table 5.8: White matter density values in JHU ROIs thresholded at probability > 10% for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 17 patients and 18 controls. Values for raw white matter density and white matter density with WM probability > 80% are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Where no value is shown there are no voxels with WM probability > 80%. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw WM density</th>
<th>WM &gt; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LCCg (10pc)</td>
<td>0.79±0.08</td>
<td>0.14</td>
</tr>
<tr>
<td>RCCg (10pc)</td>
<td>0.62±0.08</td>
<td>0.13</td>
</tr>
<tr>
<td>LCH(10pc)</td>
<td>0.69±0.15</td>
<td>0.27</td>
</tr>
<tr>
<td>RCH(10pc)</td>
<td>0.56±0.15</td>
<td>0.28</td>
</tr>
<tr>
<td>LIFOF(10pc)</td>
<td>0.81±0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>RIFOF(10pc)</td>
<td>0.80±0.08</td>
<td>0.25</td>
</tr>
<tr>
<td>LILF(10pc)</td>
<td>0.75±0.10</td>
<td>0.03</td>
</tr>
<tr>
<td>RILF(10pc)</td>
<td>0.77±0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>LSLF(10pc)</td>
<td>0.77±0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>RSLF(10pc)</td>
<td>0.78±0.06</td>
<td>0.56</td>
</tr>
<tr>
<td>LUF (10pc)</td>
<td>0.74±0.10</td>
<td>0.31</td>
</tr>
<tr>
<td>RUF (10pc)</td>
<td>0.58±0.11</td>
<td>1.00</td>
</tr>
</tbody>
</table>
5.4 Discussion

This study has confirmed the result from postmortem analyses and other MR imaging studies (see chapter 2.6.2) that patients with FTLD have decreased GM density in frontal and temporal lobes when compared with age-matched controls. This is shown by the voxel-based group analysis (see chapter 5.3.3), where large clusters are found in the frontal lobes, and in both left and right temporal lobes. Those clusters which survive cluster-based multiple comparison correction are shown in table 5.3.

The regression tests for specific cognitive test scores in patients alone also show suggestive clusters in agreement with other studies, but not all prove to be significant when cluster-based multiple comparison correction is applied. Like Mummery [Mummery et al., 2000] I found a large cluster in the left temporal lobe which is correlated with the score on the naming test (see chapter 5.3.4.1), and this cluster is significant after cluster-based multiple comparison correction. This cluster has 1350 voxels along the posterior left temporal lobe with peaks at MNI coordinates [-54 -51 -8] with p=0.05, T=4.1, and [-36 -40 -17] with p=0.05, T=4.1. The score from the pyramids and palmtrees test is also correlated with the left temporal lobe (see chapter 5.3.4.2) in agreement with Zahn et al [Zahn et al., 2004]. The cluster is larger than that correlated with the Naming score, with a cluster of 5900 voxels stretching from [-35 -39 -15] through [-26 -44 -11] to [-27 -26 -9], p=0.019. These data provided no evidence for a link between the score on the famous faces recognition task as found by Snowden at al [Snowden et al., 2004], as is shown in chapter 5.3.4.3. There is a suggestion of a link between the guilt score in the evaluation of social actions test and a cluster in the septal region, as found by Moll et al [Moll et al., 2011], but this does not survive correction for multiple comparisons. The cluster is [5 5 9], but the p-value when cluster-based correction for multiple comparisons is applied = 0.1. This may be suggestive given the small numbers in this study.

The group differences are confirmed by the regional analysis (see chapter 5.3.5), where differences are found in GM density in all the aal regions known to be affected. These differences are also found when the tests are restricted to voxels with a high probability of being pure GM. These analyses also allow quantification of the difference. The difference in density is substantial, being around 30% when all the temporal lobe ROIs are considered. There is also decreased GM density in
regions normally considered white matter tracts, with lower GM density in patients in the right and left inferior fronto-occipital fasciculus, right and left inferior longitudinal fasciculus, the right and left superior longitudinal fasciculus and the right and left uncinate fasciculus. This difference is around 20%. WM densities are also lower in patients in the right posterior temporal lobe, and the left anterior and posterior temporal lobe; this difference is also around 20%. WM densities are lower in patients in the right and left inferior longitudinal fasciculus; this difference is around 10%.

This study confirms the existence of a regional pattern of atrophy in FTLD, and the link of certain key regions with specific cognitive test scores. The significance of this link for the development of a possible diagnostic support tool is discussed in chapter 9.3.3.
Chapter 6

Cerebral blood flow and arterial arrival time in frontotemporal lobar degeneration

6.1 Introduction

The aim of ASL imaging in this study is to investigate whether measurement of cerebral blood flow (CBF) and arterial arrival time (AAT) can add information that is not available through study of structural MR images, either for individual diagnosis, or for understanding of the disease processes.

FTLD is described in chapter 1.2, and imaging in FTLD is described in more detail in chapter 2.6, but I summarise briefly here. Early post-mortem studies of the brains of people who showed symptoms of FTLD describe marked regional atrophy in the frontal and temporal regions of the brain [Brun, 1987]. Later SPECT and PET studies showed hypometabolism in the same regions (see chapter 2.6.1), although it is unclear whether this hypometabolism is simply due to atrophy as partial volume correction was not used in these earlier studies. ASL gives us a non-invasive, repeatable method to study cerebral blood flow (CBF) and arterial arrival time (AAT), and, when used in conjunction with anatomical images, a way to investigate CBF per volume of remaining tissue by correcting for atrophy. Unfortunately the regions known to be affected in FTLD are also brain regions which
are difficult to image with MRI because of their proximity to the air spaces within
the sinuses and ears; hence the early studies of FTLD using ASL only describe CBF
in the more superior regions [Du et al., 2006].

This study uses ASL to derive CBF and AAT in the brains of patients di-
agnosed with either behavioural-variant frontotemporal dementia or semantic de-
mentia, and healthy controls. SPM is then used to investigate differences between
the groups. Unfortunately there are features of CBF which make investigating
such differences difficult:

- There is an inherent variability in CBF from individual to individual. Parkes
  et al. [Parkes, 2004] found CBF between individuals of the same age and
gender can vary by up to 100%. There is no reason to believe that this inher-
ent variability will change between patients and controls, but it does make
looking for a group difference in necessarily small groups more difficult

- Grey matter and white matter have different values of CBF and AAT [Law
  et al., 2000]. ASL is an imaging technique with a low SNR, and consequently
large voxels are used to collect data. Each voxel is therefore likely to have a
mixture of grey matter, white matter and CSF, and the measured CBF values
will lie between the values for grey and white matter, assuming that CSF
has no blood flow. Hence the measured CBF in each voxel will depend on
the proportions of each tissue, and will not be directly comparable with CBF
from similar voxels in other subjects

- FTLD patients have localised cerebral atrophy, particularly in grey matter
  [Neary et al., 1998]. Voxels in the same region in patients are likely to have
proportionally less grey matter than similar voxels in controls, and this is
likely to introduce a consistent group difference unless the effect can be ac-
counted for.

- Highly localised atrophy can cause difficulties when normalising patient brains
to MNI space. This has been covered earlier in this thesis (see chapter 5.3.2).

Several recent studies have considered correction methods for PET data, treating
each voxel as a mixture of tissue types, and using linear regression to estimate the
CBF values for each type [Du et al., 2005], [Aston et al., 2002]. Asllani et al. apply
similar methods to continuous ASL data [Asllani et al., 2008], using probabilistic
images from a segmented high-resolution T$_1$-weighted image to weight the CBF values measured within each voxel. Chappell et al also use probabilistic images to weight the signal from each voxel, but they apply the partial volume correction to the raw subtraction signal from multi-timepoint data, and then fit for AAT and CBF for both grey and white matter from the corrected data, which has 10 timepoints [Chappell et al., 2011]. Neither approach can be used for the ASL data in this study. Asllani et al were only collecting CBF data, and had no data on different AATs. Chappell et al had ASL data collected with 10 timepoints, whereas this study only has 4 timepoints, and therefore cannot be used to estimate 4 parameters. The approach I have taken is to use simulation to estimate correction factors for both CBF and AAT within a voxel, depending on the proportions of grey and white matter within each voxel.

6.2 Methods

The study is approved by South Manchester NHS Research Ethics Committee, reference 07/H1003/194 South Manchester REC. 14 patients and 20 healthy controls were scanned as described above. These numbers are different from the numbers for whom there was T$_1$-weighted images because this ASL sequence was introduced part-way through the study. Difficulties with normalisation meant that 1 patient had to be excluded, and MRI abnormalities forced the exclusion of 2 controls; 13 patients (6M, 65.4±8.1y) and 18 controls (12M, 62.6±10.5y) were included in this analysis.

6.2.1 ASL sequence and model

All subjects were scanned in a Philips Achieva 3T MRI scanner with 8-channel head coil. A T$_1$-weighted MPRAGE scan was taken as described earlier (see chapter 5.2.1), and a multi-timepoint ASL sequence: STAR label, label thickness 150mm; label gap 10mm; gradient-echo EPI readout; TE 21ms; TR 3000ms; 4 delay times of 800, 1200, 1600, 2000ms from label to start of readout; FOV 224 x 224mm; matrix size 64 x 64; 20 5mm slices; slice gap 1mm Voxel size 3.5 x 3.5 x 6.0mm; 20 control/label pairs. Vascular crushing was enabled. An additional calibration scan with TR 10000ms no label was taken to allow quantification. The spin-echo EPI
sequence recommended in chapter 4.4.2 was not used in this study because recruitment had started, and the images taken of several subjects, before that work was completed.

All processing was carried out in Matlab2009a and SPM8. The individual control/label images were checked for artefacts as described in chapter 4.2.3. All subsequent images were then aligned to the first image using the SPM “Align function” and checked for excess motion. CBF and AAT images were calculated from the ASL subtraction images using in-house software and the single-blood-compartment model described earlier. The relaxation time of blood was fixed at 1600ms [Greenman et al., 2003]; the labelling efficiency at 0.9; the blood-brain partition coefficient at 0.9 [Roberts et al., 1996] and the bolus width at 1100ms. The labelling efficiency and bolus width were taken from earlier studies of this sequence. If the fitted value for CBF was greater than 250 ml/100ml/min, it was set to 250 ml/100ml/min; AAT was set to zero if the fit returned a negative value, and set to 5000ms if the fit returned a greater value. CBF was also calculated in a single-parameter fit to the subtraction signal, using a fixed arrival time of 750ms.

There is no CBF in the cerebrospinal fluid, and so CSF voxels were masked out during the calculation. This was done by segmenting the T1-weighted image, co-registering the image to the first ASL control image, applying the same transformation to the segmented images, and calculating a mask where grey matter probability + white matter probability is > 0.75. CSF and AAT are only calculated within this mask. SPM8 was used for these calculations.

The CBF and AAT images were normalised to MNI space (voxel size 2x2x2mm³) with the SPM8 “Normalise” function, using the first ASL control image as source, the SPM EPI image as template, and applying the same transformation to all other ASL images. The T1-weighted image was normalised to MNI space using the SPM T1 template, and segmented with the “Segment” function to provide grey-matter and white-matter probabilistic images for partial volume correction. For regional analysis the images were not smoothed; for 75 analysis the images were smoothed with a 12mm FWHM gaussian kernel.
6.2.2 Partial volume correction

Earlier studies have shown that CBF is lower, and AAT longer, in white matter compared with grey matter [van Osch et al., 2009], [van Gelderen et al., 2008]. The ratio of grey matter CBF to white matter CBF has been reported as varying from 2 to 3.6 measured by ASL, and 3.7 to 7.0 for other methods [Law et al., 2000]. Although some DCE studies suggest AAT in deep white matter is of the order of seconds [Law et al., 2000], DCE is measuring the arrival time from the site of injection of the contrast medium to the brain. In an ASL study, van Gelderen et al found that mean WM AAT was 0.65±0.08s longer than grey matter arrival time [van Gelderen et al., 2008]. He also found considerable heterogeneity in both grey matter and WM AAT. Paling et al, also in an ASL study, found average GM AAT of 0.80±0.14s versus WM AAT 0.92±0.15 seconds [Paling et al., 2013]. For this simulation I have assumed that white matter CBF = 0.4 * grey matter CBF [Asllani et al., 2008], and that white matter AAT = 1.5* grey matter AAT [Paling et al., 2013].

A simulated subtraction signal was calculated for grey and white matter with no added noise for times from 500ms to 4000ms, using the single blood compartment model and other values described elsewhere (see chapter 3.1.1). The subtraction signals from grey matter, white matter and a mixture of 50% of each are shown in figure 6.1(a). The signals from grey and white matter were then added for varying percentages of each (assuming no CSF), and values at inversion times of 800ms, 1200ms, 1600ms and 2000ms used to calculate CBF and AAT. Figures 6.1(b) and 6.1(c) show the simulated CBF and AAT as a function of the composition of the mixture. It can be seen that CBF varies linearly with GM/WM proportions, while AAT is dominated by the GM value until GM percentage drops below 60%, below which a correction factor needs to be applied depending on the WM concentration. Any CSF present in the voxel is assumed to have no CBF, and hence no effect on the measured AAT. Let the measured values of CBF and AAT within a voxel be respectively \( f_{vxl} \) \( t_{A_{vxl}} \); for grey matter CBF \( f_{gm} \), for grey matter AAT \( t_{A_{gm}} \), for white matter CBF \( f_{wm} \) and for white matter AAT \( t_{A_{wm}} \). The values for grey and white matter probability from a T1-segmented image are \( prob_{gm} \) \( prob_{wm} \).

**Equation for correcting CBF**  If we assume a fixed ration between CBF for GM and CBF for WM, then the CBF signal within the voxel comes from the sum of the
CHAPTER 6. CEREBRAL BLOOD FLOW AND ARTERIAL ARRIVAL TIME IN FRONTOTEMPORAL LOBAR DEGENERATION

Figure 6.1: Simulated data for a mixture of grey and white matter. 6.1(a) shows the simulated subtraction signal for grey matter, white matter and a 50/50 mix. 6.1(b) shows CBF values calculated from this simulated data for varying proportions of white matter. 6.1(c) shows AAT values calculated from this simulated data for varying proportions of white matter.
GM signal and the WM signal:

\[ f_{\text{vxl}} = \text{prob}_{\text{wm}} \times f_{\text{wm}} + \text{prob}_{\text{gm}} \times f_{\text{gm}} \quad \text{where} \quad f_{\text{wm}} = 0.4 \times f_{\text{gm}} \quad (6.1) \]

and hence

\[ f_{\text{gm}} = \frac{f_{\text{vxl}}}{\text{prob}_{\text{gm}} + 0.4 \times \text{prob}_{\text{wm}}} \quad (6.2) \]

This correction factor approaches infinity as the value of \( \text{prob}_{\text{wm}} \) approaches zero, and so I have only applied it for values of \( \text{prob}_{\text{wm}} > 0.1 \), and also applied a cut-off for CBF of 250ml/100ml/min. This is the same cut-off as is used in the calculation of CBF from the subtraction signal. It is worth noting that \( \text{prob}_{\text{gm}} + \text{prob}_{\text{wm}} \) do not necessarily add up to 1, as there may be CSF within the voxel.

**Equation for correcting AAT**

It is clear from figure 6.1(c) that when the percentage of WM within a voxel is less than 40%, the AAT within the voxel is scarcely affected. It only differs from the AAT for GM when the proportion of WM rises above 40%. I have therefore chosen to apply a correction to the AAT in any given voxel only when the proportion of WM rises above 40%. I have then applied a linear correction, so that at 40% WM the AAT within the voxel is assumed to be that of GM, while at 100% WM the AAT within the voxel is that of WM. I have also assumed that the AAT for WM is 1.5* the AAT for GM. Then:

\[ t_{A_{\text{vxl}}} = t_{A_{\text{gm}}} \quad \text{for} \quad \text{prob}_{\text{wm}} \leq 0.4 \quad (6.3) \]

\[ t_{A_{\text{vxl}}} = t_{A_{\text{gm}}} + (\text{prob}_{\text{wm}} - 0.4) \times \frac{t_{A_{\text{wm}}} - t_{A_{\text{gm}}}}{0.6} \quad \text{for} \quad \text{prob}_{\text{wm}} > 0.4 \quad (6.4) \]

\[ t_{A_{\text{gm}}} = \frac{6 \times t_{A_{\text{vxl}}}}{4 + 5 \times \text{prob}_{\text{wm}}} \quad \text{for} \quad \text{prob}_{\text{wm}} > 0.4 \quad (6.5) \]

Since this correction factor is well-behaved, no cut-off is needed.

### 6.2.3 Voxel-based analysis

Voxel-based analysis in SPM was used to carry out an unpaired 2-tailed t-test group comparison between patients and controls, with no correction for multiple comparison, p-value<0.005 and minimum cluster size of 75; these values correspond to those recommended by Lieberman and Cunningham [Lieberman
and Cunningham, 2009]. Global normalisation was set to on, with a value of 50. This was used because the whole-brain CBF value varies from one individual to another [Parkes et al., 2004], and this overall variation is not of interest when investigating regional differences. Global normalisation will calculate the mean whole-brain value, calculate a factor which is 50 divided by the mean whole-brain value, and then scale all the voxel values according to this factor. Comparisons were carried out for CBF and AAT from the 2-parameter fit both with and without partial volume correction, and for CBF from the 1-parameter fit.

### 6.2.4 Regional analysis

CBF and AAT were also compared on a regional basis, using the aal ROIs described earlier (5.2.3.2) and calculating both the raw and partial volume-corrected values. The ROIs taken from the JHU white matter atlas were not used because of the low SNR of the ASL signal in white matter. An additional ROI (GM-Sig) was used, based on the highly-atrophied region defined by the voxel-based group analysis carried out in the previous chapter (5.6). Voxels were included in this ROI if they had a t-test result > 1 after FWE correction for multiple comparisons.

### 6.3 Results

#### 6.3.1 Group differences in CBF and AAT with no partial volume correction

Figure 6.2 shows axial and coronal images of the result of a 2-tail t-test comparing CBF in FTLD patients with controls, uncorrected for multiple comparisons, uncorrected for partial volume effects, p < 0.005 and minimum cluster size of 75 voxels. The same data is shown in figure 6.3, but corrected to p < 0.05 according to Monte Carlo simulations implemented in AFNIs 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Table 6.1 shows the location in MNI coordinates, the cluster size, the uncorrected voxel p-value and FWE-corrected p-values for the significant clusters shown in 6.3. There is hyperperfusion in patients in the left paracentral region. There is hypoperfusion in the right insula, the right
rolandic operculum, the right mid-frontal region, the left insula and the left putamen.

Figure 6.4 shows axial and coronal images of the result of a 2-tail t-test comparing AAT in FTLD patients with controls, uncorrected for multiple comparisons, uncorrected for partial volume effects, \( p < 0.005 \) and minimum cluster size of 75 voxels. The same data is shown in figure 6.5, but corrected to \( p < 0.05 \) according to Monte Carlo simulations implemented in AFNIs 3dClustSim, which gives a minimum cluster size of 400 voxels for this \( p \)-value. Table 6.2 shows the location in MNI coordinates, the cluster size, the uncorrected voxel \( p \)-value and FWE-corrected \( p \)-values for the significant clusters shown in 6.5. AAT group differences are more extensive and of greater significance than CBF differences. AAT is longer in patients in the right insula, the right precentral region, the pars opercularis of the inferior frontal gyrus, the left calcarine and left lingual. AAT is shorter in patients in the right thalamus, the right superior temporal region, the right caudate, the right superior frontal region, the left fusiform and the left superior temporal region.

Figure 6.6 shows the difference in AAT uncorrected for partial volume effects between FTLD patients and controls: I have calculated an image which is the mean of AAT images for all the FTLD patients, and another image which is the mean of all the control AAT images, and then subtracted one from the other. It can be seen that for the majority of the brain, dark red and dark blue regions show where there is little difference. Red is where the AAT in patients is longer than in controls; blue where the AAT in patients is shorter than in controls. These differences are on the whole less than 50ms. There are however some regions in yellow and green. Yellow regions indicate that here the AAT is considerably longer in patients, by around 300ms. Green regions indicate that here the AAT is shorter in patients, by around 300ms. The average AAT in the brain is around 850ms, so this is a considerable difference.
Figure 6.2: Axial and coronal images of t-test results (p<0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons) where CBF uncorrected for partial volume effects differs between patients and controls. Red shows where patients have greater CBF than controls; blue where patients have lower CBF.

Figure 6.3: Axial and coronal images of t-test results where CBF uncorrected for partial volume effects differs between patients and controls. Statistics were calculated for p<0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Red shows where patients have greater CBF than controls; blue where patients have lower CBF.
**Figure 6.4:** Axial and coronal images of t-test results where AAT uncorrected for partial volume effects differs between patients and controls (p < 0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons). Red shows where patients have longer AAT than controls; blue where patients have shorter AAT.

**Figure 6.5:** Axial and coronal images of t-test results where AAT uncorrected for partial volume effects differs between patients and controls. Statistics were calculated for p < 0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Red shows where patients have longer AAT than controls; blue where patients have shorter AAT.
Table 6.1: Cluster statistics showing regions where CBF uncorrected for partial volume effects differs between patients and controls. Statistics were calculated for p<0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Clusters where FTLD patients have more CBF than controls have a grey background. Clusters marked in red have p(FWE-corrected)<0.05.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster(k)</th>
<th>Peak p (FWE corr)</th>
<th>Peak T</th>
<th>p (uncorr)</th>
<th>MNI local maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L paracentral</td>
<td>444</td>
<td>0.81</td>
<td>4.0</td>
<td>&lt;0.001</td>
<td>-12 -32 56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.93</td>
<td>3.7</td>
<td>&lt;0.001</td>
<td>-8  -30 64</td>
</tr>
<tr>
<td>R insula</td>
<td>514</td>
<td>0.13</td>
<td>5.1</td>
<td>&lt;0.001</td>
<td>46  16  -4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3.0</td>
<td>&lt;0.005</td>
<td>48  -12 10</td>
</tr>
<tr>
<td>R mid-frontal</td>
<td>465</td>
<td>0.87</td>
<td>3.9</td>
<td>&lt;0.001</td>
<td>28  18  34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3.1</td>
<td>&lt;0.005</td>
<td>36  0  44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3.0</td>
<td>&lt;0.005</td>
<td>28  42  30</td>
</tr>
<tr>
<td>L putamen</td>
<td>403</td>
<td>0.95</td>
<td>3.7</td>
<td>&lt;0.001</td>
<td>-34 4  0</td>
</tr>
<tr>
<td>L insula</td>
<td></td>
<td>0.96</td>
<td>3.6</td>
<td>&lt;0.001</td>
<td>-30 10  6</td>
</tr>
</tbody>
</table>

Figure 6.6: Axial and coronal images of difference in AAT values between patients and controls uncorrected for partial volume effects. Positive values are where patients have longer AAT; negative where patients have shorter AAT. Of particular interest are the small regions of yellow and green, where the difference is substantial.
Table 6.2: Cluster statistics showing regions where AAT uncorrected for partial volume effects differs between patients and controls. Statistics were calculated for \( p < 0.05 \) corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this \( p \)-value. Clusters where FTLD patients have longer AAT than controls have a grey background. Clusters with red text have \( p \) (FWE-corrected) < 0.05.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster(k)</th>
<th>Peak p (FWE corr)</th>
<th>Peak T</th>
<th>( p ) (uncorr)</th>
<th>MNI local maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>L calcarine</td>
<td>4852</td>
<td>0.02</td>
<td>6.1</td>
<td>&lt;0.001</td>
<td>-8</td>
</tr>
<tr>
<td>L lingual</td>
<td></td>
<td>0.21</td>
<td>4.9</td>
<td>&lt;0.001</td>
<td>-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.22</td>
<td>4.9</td>
<td>&lt;0.001</td>
<td>-16</td>
</tr>
<tr>
<td>R insula</td>
<td>1034</td>
<td>0.05</td>
<td>5.6</td>
<td>&lt;0.001</td>
<td>40</td>
</tr>
<tr>
<td>R precentral</td>
<td></td>
<td>1.00</td>
<td>3.2</td>
<td>&lt;0.005</td>
<td>56</td>
</tr>
<tr>
<td>R front inf oper</td>
<td></td>
<td>1.00</td>
<td>3.1</td>
<td>&lt;0.005</td>
<td>42</td>
</tr>
<tr>
<td>R sup temp pole</td>
<td>2220</td>
<td>0.01</td>
<td>6.6</td>
<td>&lt;0.001</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3.0</td>
<td>&lt;0.005</td>
<td>56</td>
</tr>
<tr>
<td>L fusiform</td>
<td>1341</td>
<td>0.05</td>
<td>5.6</td>
<td>&lt;0.001</td>
<td>-36</td>
</tr>
<tr>
<td>L sup temp pole</td>
<td></td>
<td>0.82</td>
<td>4.0</td>
<td>&lt;0.001</td>
<td>-46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3.3</td>
<td>&lt;0.001</td>
<td>-24</td>
</tr>
<tr>
<td>R sup frontal</td>
<td>485</td>
<td>0.16</td>
<td>5.1</td>
<td>&lt;0.001</td>
<td>22</td>
</tr>
<tr>
<td>R thalamus</td>
<td>3859</td>
<td>0.20</td>
<td>4.9</td>
<td>&lt;0.001</td>
<td>18</td>
</tr>
<tr>
<td>R caudate</td>
<td></td>
<td>0.26</td>
<td>4.8</td>
<td>&lt;0.001</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.41</td>
<td>4.5</td>
<td>&lt;0.001</td>
<td>6</td>
</tr>
<tr>
<td>L caudate</td>
<td>991</td>
<td>0.63</td>
<td>4.2</td>
<td>&lt;0.001</td>
<td>-34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.98</td>
<td>3.3</td>
<td>&lt;0.001</td>
<td>-30</td>
</tr>
</tbody>
</table>
6.3.2 Group differences in CBF calculated from a 1-parameter fit with fixed AAT of 750ms, no partial volume correction

Figure 6.7 shows axial and coronal images of the result of a 2-tail t-test comparing CBF in FTLD patients with controls, uncorrected for multiple comparisons, uncorrected for partial volume effects, \( p<0.005 \) and minimum cluster size of 75 voxels, where the CBF has been calculated by a 1-parameter fit with a fixed AAT of 750ms. The same data is shown in figure 6.8, but corrected to \( p<0.05 \) according to Monte Carlo simulations implemented in AFNIs 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Table 6.3 shows the location in MNI coordinates, the cluster size, the uncorrected voxel p-value and FWE-corrected p-values for the significant clusters shown in 6.8.

These images have been included for comparison with the CBF images calculated with a 2-parameter fit. The two different approaches have been described earlier in this thesis (see chapter 3). Using this calculation, clusters with higher CBF are found in the left paracentral lobule and the left calcarine; clusters with lower CBF are found in the right and left caudate and the left hippocampus. It can be seen that there is little overlap between clusters found with the different methods of calculation. When only clusters significant at \( p<0.05 \) according to Monte Carlo simulations implemented in AFNIs 3dClustSim are considered, which gives a minimum cluster size of 400 voxels for this p-value., there is one cluster around the caudate head where CBF is lower in patients. There is no indication of lower CBF in regions such as the temporal lobes which are known to be affected in FTLD. This seems an indication that the 2-parameter method if a better way of analysing this data. For further comparison, figure 6.9 shows axial images for an FTLD patient with marked atrophy. Images are shown in ASL native space, and include CBF images calculated by both methods, AAT images and a \( T_1 \)-weighted image.
(a) Axial images  
(b) Coronal images

**Figure 6.7:** Axial and coronal images of t-test results (p<0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons) where CBF calculated with a fixed AAT of 750ms, uncorrected for partial volume effects differs between patients and controls. Red shows where patients have greater CBF than controls; blue where patients have lower CBF.

(a) Axial images  
(b) Coronal images

**Figure 6.8:** Axial and coronal images of t-test results where CBF calculated with a fixed AAT of 750ms, uncorrected for partial volume effects differs between patients and controls. Statistics were calculated for p<0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. There are no regions where patients have greater CBF than controls; regions where patients have lower CBF are shown in blue.
Figure 6.9: Axial images of CBF calculated with fixed AAT, T₁-weighted image and CBF and AAT calculated with 2-parameter fit for an FTLD patient.
**Table 6.3:** Cluster statistics showing regions where CBF calculated with a fixed AAT of 750ms, corrected for partial volume effects, differs between patients and controls. Statistics were calculated for p<0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. There are no clusters where FTLD patients have more CBF, or clusters which reach p(FWE-corrected)<0.05.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster(k)</th>
<th>Peak p (FWE corr)</th>
<th>Peak T</th>
<th>p (uncorr)</th>
<th>MNI local maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caudate head</td>
<td>1700</td>
<td>0.84</td>
<td>3.9</td>
<td>&lt;0.001</td>
<td>6 12 -2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.89</td>
<td>3.8</td>
<td>&lt;0.001</td>
<td>-24 -28 16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.98</td>
<td>3.5</td>
<td>&lt;0.001</td>
<td>-6  -26 10</td>
</tr>
</tbody>
</table>

Chapter 6. Cerebral blood flow and arterial arrival time in Frontotemporal Lobar Degeneration
6.3.3 Group differences in CBF, AAT corrected for partial volume effects

Figure 6.10 shows axial and coronal images of the result of a 2-tail t-test comparing CBF in FTLD patients with controls, uncorrected for multiple comparisons, corrected for partial volume effects, $p<0.005$ and minimum cluster size of 75 voxels. The same data is shown in figure 6.11, but corrected to $p<0.05$ according to Monte Carlo simulations implemented in AFNIs 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Table 6.4 shows the location in MNI coordinates, the cluster size, the uncorrected voxel p-value and FWE-corrected p-values for the significant clusters shown in 6.11. There are fewer group differences between patients and controls in CBF once it has been corrected for partial volume effects. There is hyperperfusion in patients in the left paracentral region and left superior motor area. There is hypoperfusion in right frontal regions.

Figure 6.12 shows axial and coronal images of the result of a 2-tail t-test comparing AAT in FTLD patients with controls, uncorrected for multiple comparisons, uncorrected for partial volume effects, $p<0.005$ and minimum cluster size of 75 voxels. The same data is shown in figure 6.13, but corrected to $p<0.05$ according to Monte Carlo simulations implemented in AFNIs 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Table 6.5 shows the location in MNI coordinates, the cluster size, the uncorrected voxel p-value and FWE-corrected p-values for the significant clusters shown in 6.13.

Figure 6.14 shows the difference in AAT corrected for partial volume effects between FTLD patients and controls: I have calculated an image which is the mean of AAT images for all the FTLD patients, and another image which is the mean of all the control AAT images, and then subtracted one from the other. It can be seen that for the majority of the brain, dark red and dark blue regions show where there is little difference. Red is where the AAT in patients is longer than in controls; blue where the AAT in patients is shorter than in controls. These differences are on the whole less than 50ms. There are however some regions in yellow and green. Yellow regions indicate that here the AAT is considerably longer in patients, by around 300ms. Green regions indicate that here the AAT is shorter in patients, by around 300ms. The average AAT in the brain is around 850ms, so this is a considerable difference.
AAT group differences are more extensive and of greater significance than CBF differences, although again less extensive than AAT differences uncorrected for partial volume effects. AAT is longer in patients in right insula, precentral and frontal regions, and left calcarine and lingual regions. AAT is shorter in patients in right caudate, amygdala and superior temporal region, and left fusiform and superior temporal region.

![Figure 6.10: Axial and coronal images of t-test results) where CBF corrected for partial volume effects differs between patients and controls (p < 0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons. Red shows where patients have greater CBF than controls; blue where patients have lower CBF.]

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster(k)</th>
<th>Peak p (FWE corr)</th>
<th>Peak T</th>
<th>p (uncorr)</th>
<th>MNI local maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>L paracentral</td>
<td>512</td>
<td>0.90</td>
<td>3.8</td>
<td>&lt;0.001</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>3.2</td>
<td>&lt;0.005</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>2.9</td>
<td>&lt;0.005</td>
<td>-10</td>
</tr>
<tr>
<td>R mid frontal</td>
<td>2136</td>
<td>0.14</td>
<td>5.1</td>
<td>&lt;0.001</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.39</td>
<td>4.3</td>
<td>&lt;0.001</td>
<td>28</td>
</tr>
<tr>
<td>R sup frontal</td>
<td></td>
<td>0.59</td>
<td>4.3</td>
<td>&lt;0.001</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 6.4: Cluster statistics showing regions where CBF corrected for partial volume effects differs between patients and controls. Statistics were calculated for p < 0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Clusters where FTLD patients have more CBF than controls have a grey background. There were no clusters with p(FWE-corrected) < 0.05.
Figure 6.11: Axial and coronal images of t-test results where CBF corrected for partial volume effects differs between patients and controls. Statistics were calculated for p<0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Red shows where patients have greater CBF than controls; blue where patients have lower CBF.

Figure 6.12: Axial and coronal images where AAT values corrected for partial volume effects differs between patients and controls. Statistics were calculated for p<0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons. Red shows where patients have longer AAT than controls; blue where patients have shorter AAT.
Figure 6.13: Axial and coronal images of group differences of AAT values between patients and controls corrected for partial volume effects differ between patients and controls. Statistics were calculated for $p<0.05$ corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this $p$-value. Red shows where patients have longer AAT than controls; blue where patients have shorter AAT.

Figure 6.14: Axial and coronal images of difference in AAT values between patients and controls corrected for partial volume effects. Positive values are where patients have longer AAT; negative where patients have shorter AAT. Of particular interest are the small regions of yellow and green, where the difference is substantial.
Table 6.5: Cluster statistics showing regions where AAT corrected for partial volume effects differs between patients and controls. Statistics were calculated for p<0.05 corrected for multiple comparisons according to AFNI 3dClustSim, which gives a minimum cluster size of 400 voxels for this p-value. Clusters where FTLD patients have longer AAT than controls have a grey background. Clusters with red text have p(FWE-corrected)<0.05.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster(k)</th>
<th>Peak p (FWE corr)</th>
<th>Peak T</th>
<th>p (uncorr)</th>
<th>MNI local maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x    y    z</td>
</tr>
<tr>
<td>L calcarine</td>
<td>3980</td>
<td>0.04</td>
<td>5.7</td>
<td>&lt;0.001</td>
<td>-8   -70  8</td>
</tr>
<tr>
<td>L lingual</td>
<td></td>
<td>0.25</td>
<td>4.8</td>
<td>&lt;0.001</td>
<td>-4   -46  8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.32</td>
<td>4.7</td>
<td>&lt;0.001</td>
<td>-16  -50 -2</td>
</tr>
<tr>
<td>R insula</td>
<td>688</td>
<td>0.30</td>
<td>4.7</td>
<td>&lt;0.001</td>
<td>40   -8  12</td>
</tr>
<tr>
<td>R precentral</td>
<td></td>
<td>1.0</td>
<td>3.4</td>
<td>&lt;0.001</td>
<td>42   8   26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>3.1</td>
<td>&lt;0.005</td>
<td>56   4   20</td>
</tr>
<tr>
<td>R sup temporal</td>
<td>1934</td>
<td>0.01</td>
<td>6.2</td>
<td>&lt;0.001</td>
<td>42   -4  -28</td>
</tr>
<tr>
<td>R amygdala</td>
<td></td>
<td>0.09</td>
<td>5.3</td>
<td>&lt;0.001</td>
<td>30   -4  -28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>3.3</td>
<td>&lt;0.001</td>
<td>38   18  -32</td>
</tr>
<tr>
<td>L fusiform</td>
<td>1000</td>
<td>0.06</td>
<td>5.6</td>
<td>&lt;0.001</td>
<td>-36  -12 -24</td>
</tr>
<tr>
<td>L sup temp pole</td>
<td></td>
<td>0.85</td>
<td>3.9</td>
<td>&lt;0.001</td>
<td>-44  8  -14</td>
</tr>
<tr>
<td>R sup frontal</td>
<td>478</td>
<td>0.16</td>
<td>5.1</td>
<td>&lt;0.001</td>
<td>22   66  14</td>
</tr>
<tr>
<td>R thalamus</td>
<td>3138</td>
<td>0.19</td>
<td>5.0</td>
<td>&lt;0.001</td>
<td>18   18  20</td>
</tr>
<tr>
<td>R caudate</td>
<td></td>
<td>0.50</td>
<td>4.4</td>
<td>&lt;0.001</td>
<td>6    16   4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.52</td>
<td>4.4</td>
<td>&lt;0.001</td>
<td>12   -10 16</td>
</tr>
<tr>
<td>L mid temp (extreme edge)</td>
<td>774</td>
<td>0.72</td>
<td>4.1</td>
<td>&lt;0.001</td>
<td>-34  -50 -10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3.1</td>
<td>&lt;0.005</td>
<td>-32  -60 26</td>
</tr>
</tbody>
</table>


6.3.4 Regional Analysis

Tables 6.6 and 6.7 show the regional values for CBF and AAT respectively for patients and controls both with and without partial volume correction. It can be seen that within the region shown to be lower in GM density (see chapter 5.3.3), and in the left anterior temporal lobe, CBF is lower in patients by around 25%; these differences are significant at \( p < 0.05 \) uncorrected for multiple comparisons, but do not survive either partial volume correction or correction for multiple comparisons. AAT is decreased in the left anterior temporal lobe by about 20%, but increased in the calcarine by about 15%. These differences are significant at probability \( < 0.05 \), and survive correction for partial volume effects, but not correction for multiple comparisons.

The differences in AAT shown in the voxel-based analysis (see figures 6.13 for AAT without partial volume correction and 6.5 for AAT with partial volume correction) are confirmed by a regional analysis of these particular regions. I have saved the regions shown as significant with the cluster-based multiple comparison correction, and used these ROIs to extract the mean AAT within these regions for both the partial volume uncorrected and partial volume corrected images. This may seem to be something of a circular analysis; the difference is that the voxel-based analysis is carried out with smoothed images, and the regional analysis with unsmoothed images. Hence I think it is a valid consistency check.

For AAT uncorrected for partial volume effects, where the voxel-based analysis showed that the FTLD patients had a longer AAT then controls, the equivalent ROI values are 685±82ms for patients and 447±127ms for controls, and a 2-tailed t-test for differences in the populations has a p-value \( < 0.000001 \). Where the voxel-based analysis showed that the FTLD patients had a shorter AAT than controls, the ROI values are 478±127ms for patients and 748±87ms for controls, with a similar p-value from a t-test.

For AAT corrected for partial volume effects, where the voxel-based analysis showed that the FTLD patients had a longer AAT than controls, the ROI values are 591±83ms for patients and 376±107ms for controls, with the p-value from a t-test for differences in the population \( < 0.000001 \). Where the voxel-based analysis showed that the FTLD patients had a shorter AAT than controls, the ROI values are 432±107ms for patients and 659±68ms for controls.
These figures are shown graphically in figure 6.15, where I have plotted the individual ROI AAT values for the 17 patients and 18 controls. Figure 6.15(a) shows the values for the ROI where patients have longer AAT than controls for both no partial volume correction and partial volume correction. The leftmost two columns show the values for patients, with the uncorrected figures in the column on the extreme left, and partial volume corrected figures in the second left column. The rightmost two columns show the values for controls, with uncorrected figures on the left, and partial volume corrected on the extreme right. Figure 6.15(b) shows similar data for the ROI where patients have shorter AAT than controls. It can be seen that the ROI analysis supports the conclusions reached by the voxel-based analysis.

**Figure 6.15:** ROI values from regions where voxel-based analysis shows differing AAT between patients (FTLD) and controls (HC). Figure 6.15(a) shows the values for the ROI where patients have longer AAT than controls for both no partial volume correction (NoPVC) and partial volume correction (PVC). The leftmost two columns show the values for patients, with the uncorrected figures in the column on the extreme left, and partial volume corrected figures in the second left column. The rightmost two columns show the values for controls, with uncorrected figures on the left, and partial volume corrected on the extreme right. Figure 6.15(b) shows similar data for the ROI where patients have shorter AAT than controls.
### Table 6.6: Regional CBF for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 13 patients and 18 controls. Values for raw CBF and CBF corrected for partial volume effects are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>CBF (ml/100ml/min)</th>
<th>Corrected for partial volume effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>GM-Sig</td>
<td>23.44±6.62</td>
<td>0.05</td>
</tr>
<tr>
<td>LPostTL</td>
<td>27.20±10.53</td>
<td>0.08</td>
</tr>
<tr>
<td>RPostTL</td>
<td>28.28±9.24</td>
<td>0.14</td>
</tr>
<tr>
<td>LAntTL</td>
<td>22.71±8.78</td>
<td>0.02</td>
</tr>
<tr>
<td>RAntTL</td>
<td>22.46±7.15</td>
<td>0.27</td>
</tr>
<tr>
<td>PostVMPFC</td>
<td>15.22±5.40</td>
<td>0.18</td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>16.83±9.84</td>
<td>0.24</td>
</tr>
<tr>
<td>Calcarine</td>
<td>36.48±15.19</td>
<td>0.37</td>
</tr>
<tr>
<td>Cereb6</td>
<td>37.63±16.58</td>
<td>0.51</td>
</tr>
</tbody>
</table>

### Table 6.7: Regional AAT for patients and controls. Values shown are mean±SD within each ROI for the patient and control group. The analysis includes 13 patients and 18 controls. Values for raw AAT and AAT corrected for partial volume effects are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>AAT (ms)</th>
<th>Corrected for partial volume effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>GM-Sig</td>
<td>543±88</td>
<td>0.84</td>
</tr>
<tr>
<td>LPostTL</td>
<td>591±97</td>
<td>0.76</td>
</tr>
<tr>
<td>RPostTL</td>
<td>537±78</td>
<td>0.98</td>
</tr>
<tr>
<td>LAntTL</td>
<td>375±85</td>
<td>0.02</td>
</tr>
<tr>
<td>RAntTL</td>
<td>355±48</td>
<td>0.16</td>
</tr>
<tr>
<td>PostVMPFC</td>
<td>537±160</td>
<td>0.83</td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>402±149</td>
<td>0.70</td>
</tr>
<tr>
<td>Calcarine</td>
<td>643±116</td>
<td>0.02</td>
</tr>
<tr>
<td>Cereb6</td>
<td>820±167</td>
<td>0.07</td>
</tr>
</tbody>
</table>
6.4 Discussion

6.4.1 Partial volume correction

See chapter 6.2.2 for details of the partial volume corrections used to establish grey matter CBF and grey matter AAT from that measured within each voxel. The partial volume correction for CBF is one used by others [Du et al., 2006]. It relies on the assumptions that the subtraction signal varies linearly with CBF, and that CBF in white matter is always a fixed percentage of the grey matter CBF in the same voxel. The first assumption has been discussed already (see chapter 3) and found to be reasonable for this model; alternatives to the second might be to assume that CBF in white matter always has a fixed value, or to smooth data across nearby voxels, and fit for both GM and WM CBF, which is the approach taken by Asllani et al [Asllani et al., 2008]. I have chosen to use the fixed ratio method. The partial volume correction I have applied to AAT is novel. Again, I have assumed that the ratio of grey matter AAT to white matter AAT is constant, and that the ratio is 1.5. This assumption is based on ASL estimates of grey and white matter AAT by Paling et al [Paling et al., 2013] and MacIntosh [MacIntosh et al., 2010]. Note that AAT estimated by DCE, which is of the order of seconds [Law et al., 2000], is actually measuring a different arrival time as the contrast medium used to provide the signal is injected a considerable distance from the imaging volume. Given the low SNR of ASL in white matter, the non-linearity of the relationship between AAT and subtraction signal, and the overwhelming effect of grey matter AAT, it is questionable whether this correction is necessary. The partial volume correction does make a substantial difference to AAT values, as can be seen from figure 6.15.

6.4.2 CBF differences

This group comparison of 13 FTLD patients and 18 healthy age-matched controls has found some clusters where CBF is lower in patients than in controls using voxel-based analysis (see chapter 6.3.1). When corrected for multiple comparisons using cluster-based correction, there is hyperperfusion in patients in the left paracentral region, and hypoperfusion in regions around the right and left insula. When corrected for partial volume, there is a significant CBF deficit in patients
in the right mid-frontal region; this agrees with the finding of Du et al [Du et al., 2006], although his quoted MNI coordinates are not the same as found here (Du’s coordinates are [22 44 42] and [22 34 40]; in this study they are [28 20 38], [28 4 46] and [36 2 44]). Given the well-reported hypometabolism in frontal and temporal regions in FTLD patients shown by FDG-PET [Hu et al., 2010], [Foster, 2003], [Salmon et al., 2003], [Diehl et al., 2004], [Jeong et al., 2005], it is unexpected that the voxel-based analysis did not find more widespread CBF deficits, although this may be an effect of the smoothing process (see chapter 7.3.2.1), or may be a function of the small and heterogeneous patient group available for this study. The study by Du et al has already been mentioned, and this study found comparable CBF deficits. Hu et al [Hu et al., 2010] reported CBF deficits found with CASL; they used a similar partial volume correction for CBF, and found hypoperfusion in patients in the bilateral frontal lobes, and hyperperfusion in the posterior cingulate and medial parietal/precuneus in a study of 42 patients and 23 controls. The p-values they quote are uncorrected for multiple comparison. The regional comparison also shows significant hypoperfusion in the atrophied area and the left anterior temporal lobe before partial volume correction, and none after (see table 6.6.

### 6.4.3 AAT differences

There are significant AAT clusters when corrected for partial volume effects; arrival times are greater in patients in the left calcarine and right insula; AAT is shorter in patients in the left fusiform, the right mid-temporal lobe, right frontal superior lobe and right caudate. The AAT uncorrected for partial volume effects does reach significance: I have quoted both FWE-corrected and the less stringent AFNI Monte-Carlo probabilities. AAT uncorrected for partial volume effects is greater in patients in two large clusters in the right calcarine/lingual and the right insula/precentral/frontal. AAT uncorrected for partial volume effects is shorter in patients in the right fusiform/temporal pole, the right caudate, the right frontal sup lobe and the left fusiform/temporal pole/insula. The finding of altered AAT in patients compared with controls is novel. Yoshiura et al reported no difference in AAT between patients with Alzheimer’s disease and controls [Yoshiura et al., 2009], but the time they quote is not comparable with the AAT used here: they were measuring the difference in AAT between scans measured with vas-
cular crushing on and off, and hence transit time in the macrovasculature. The regional analysis within regions found to be different in the voxel-based analysis confirms these findings of both increased and decreased regional AAT in patients. Other recent studies have found increased AAT in MS [Paling et al., 2013] and in Parkinson’s disease, [Al-Bachari et al., 2014] using similar methods to this study, but no-one has reported a decreased AAT in a disease process, and it is difficult to imagine what may be the cause. However, the finding is quite robust, and affects the regions known to be implicated in the disease process; perhaps it indicates a reduction in the tortuosity of the microvasculature. It may be suggestive that some familial forms of FTLD are caused by genes implicated in glial cell function [Cooper et al., 1996], [Hu et al., 2011].
Chapter 7

Diffusion MRI in frontotemporal lobar degeneration

7.1 Introduction

The aim of this part of the study is to investigate the white matter differences in FTLD patients using diffusion MRI. Fractional Anisotropy (FA) and mean diffusivity (MD) are used to compare patients with controls.

Figure 7.1: Coronal section from T$_1$-weighted images of the left anterior temporal lobe for control and FTLD patient Figure 7.1(a) shows a coronal section of the whole brain. The red box shows the region which has been enlarged in figures 7.1(b) for a patient and 7.1(c) for a control. The localised atrophy in the FTLD patient is clear visible.

Figure 7.1 shows a coronal slice through the left anterior temporal lobe for a
typical FTLD patient and a control: the very localised atrophy is clear, and also the
loss of white matter; diffusion imaging, by indicating changes in the WM micro-
structure, may provide a more sensitive marker of early white matter change. En-
gleund et al reported white matter differences (consisting of gliosis and loss of myelin)
oberved in post-mortem brains of patients with “frontal lobe dementia of non-
Alzheimer type” [Englund and Brun, 1987] (N.B. this predates the publication by
Neary et al of diagnostic criteria for FTLD [Neary et al., 1998]). In 2004 Larsson
et al reported reduced anisotropy in the frontal lobes of an FTLD patient in a DTI
MRI study of a postmortem formalin-fixed brain[Larsson et al., 2004]. In-vivo and
ex-vivo studies of a mouse brain suggested that tissue architecture revealed by
diffusion imaging is preserved in a formalin-fixed brain [Sun et al., 2003]. Larsson
also reported that the reduced anisotropy differences were more extensive than
differences in a $T_2$-weighted image: (DTI seemed to be more sensitive than MRI for
the detection of white matter abnormalities). The results of a diffusion MRI studies
into differences between FTLD patients, AD patients and controls was first pub-
lished in 2005 [Yoshiura et al., 2005], when they suggested that the observed white
matter MD abnormalities are secondary to damage in the overlying cortex, although in
this study the majority of his patients had AD). In 2006 the same team reported a
study comparing 13 FTLD patients to 15 controls, and reported abnormal MD el-
evation was seen predominantly in the frontal and temporal lobes in FTD patients using
visual rating scales. Borroni et al carried out a voxel-based analysis of white mat-
ter fractional anisotropy in a group of 36 FTLD patients (28 with frontal variant
and 8 with temporal variant) and 23 controls [Borroni et al., 2007]. More recently
Dopper et al studied a group of 75 healthy individuals; 37 known to be carriers of
a familial FTLD allele, and 38 non-carrier siblings. They report reduced fractional
anisotropy and increased radial diffusivity throughout frontotemporal white matter tracts
in presymptomatic carriers that were not apparent in their non-carrier siblings
[Dopper et al., 2013]. The same study found no differences in regional grey matter
volume.
There is thus reason to believe that MRI DTI will reveal differences in white matter
in patients with FTLD that are more extensive, and may precede, obvious damage
to the grey matter.
7.2 Methods

The study has ethical approval from South Manchester NHS Ethics Committee, reference 07/H1003/194 South Manchester REC. 17 patients with FTLD and 15 controls were scanned with the diffusion sequence described above; images from 3 patients and 1 control were unusable because of scanner problems; images from a further 3 patients were unusable because of excess movement during the scan; images from another patient were unusable despite marginally acceptable movement during the scan because of artefacts generated by the distortion correction. 10 patients (6M, 65.0±9.0y) and 14 controls (9M, 61.2±10.9) were included. The distortion correction is particularly sensitive to motion, as it relies on there being no relative motion between the images taken with the different phase-encoding directions, so any differences are only due to the opposing distortions.

7.2.1 MRI sequences

Subjects were scanned in a Philips Achieva 3T scanner with 8-channel head-coil. A pulse triggered DTI spin-echo EPI sequence was used, with matrix size 128x128, FOV 220mm x 220mm x 120mm, 40 axial slices of 3mm thickness acquired in ascending order, voxel size 1.8mm x 1.8mm x 3mm, SENSE factor 2.5, TE 59ms. TR was dependent on timing of subject’s pulse. 16 different sensitizing gradients were applied, with \( b \) value = 1200 sec/mm\(^2\). Two sets of DTI images were taken consecutively, one with R-L phase-encoding, and the other with L-R. There is significant distortion in the phase-encoding direction with and EPI image: the two sets of images were used to attempt a correction for this distortion. The total scan time was of the order of 5-6 minutes, with the R-L encoded images taken approximately 2-3 minutes after the L-R. A high-resolution T\(_1\)-weighted MPRAGE image was also collected as described earlier (see chapter 5.2.1).

7.2.2 Pre-processing

The distortion correction described by Embleton at al [Embleton et al., 2010] was applied to the two sets of diffusion images to provide a composite, distortion-corrected set of images. These were examined for artefacts and excess motion and,
and subjects with movement causing rotational mis-alignment of more than 3°, or translational misalignment of 2mm in any single direction were rejected. The diffusion tensor was calculated for the entire image area by in-house software written by H Haroon (personal communication) from the distortion-corrected images. Fractional anisotropy (FA) and mean diffusivity (MD) were calculated as defined by Alexander et al [Alexander et al., 2007].

7.2.3 Voxel-based analysis

In one analysis the FA images were processed within FSL[Smith et al., 2004], first using the brain extraction tool (BET) to restrict the images to within the brain, and then using tract-based spatial statistics (TBSS) [Smith et al., 2006] for further analysis. TBSS projects FA data from all subjects onto a mean FA tract skeleton, before applying voxelwise cross-subject statistics. The mean tract skeleton was derived from all subjects included in the TBSS analysis. Correction for multiple comparisons was done using the threshold-free cluster enhancement (TFCE) available in FSL [Smith and Nichols, 2009], [Woolrich et al., 2009]. TBSS was used to investigate the FA images, as it uses the directional nature of the FA data to attempt to identify the white matter tracts within the brain. MD, being a scalar quantity, can be analysed with a voxel-based analysis in SPM, as it has similar properties to, e.g. the grey matter probability density images generated from the T₁-weighted image.

Further analyses on the MD images were carried out using SPM8. The b=0 image was normalised to the SPM EPI template, and the same transformation applied to the FA and MD images. The normalised b=0 image for each brain was scrutinised to see if there were any obvious problems with the normalisation, particularly in the case of the very atrophied brains of some of the FTLD patients. The images were smoothed with a Gaussian filter of FWHM of 12mm, and the smoothed images were used as input to the voxel-based analysis. A 2-group t-test was used, with patients as one group and controls as another, with no multi-comparison correction, p of 0.005 and cluster size 75 voxels which match the recommendations of Lieberman and Cunningham [Lieberman and Cunningham, 2009]. Results corrected for multiple comparisons by cluster-based correction were also reported.

These analyses were carried out with the calculated MD images, and also with images masked in an attempt to minimise partial volume effects. The two
analyses, with and without masking, are giving different information: without
masking can perhaps give clearer diagnostic information about the classification
of each subject; the analysis with masking may give information about differ-
ences between disease and normal tissue. This is because the MD of water (or
cerebrospinal fluid (CSF)) is very much higher than the MD of grey or white mat-
ter, and a small difference in the partial volume of CSF in any voxel may make
a very large difference to the overall voxel signal. The $T_1$-weighted image from
each subject was segmented as described earlier (see chapter 5.2.3.1), and a mask
was created for each individual by adding the grey and white matter probability
masks, such that voxels were only included in the analyses if GM probability +
WM probability $>$ 0.8. To exclude the possibility of any voxels excluded from an
individual being included in the group analysis, a mask was created which was
the Boolean AND of all the individual masks where (GM+WM) $>$ 0.8, and the
analysis was masked with this. Voxels outside the mask were set to either 0 or
NaN, a Matlab flag indicating that the value is non-numeric.

Multiple regression tests were run, with age included as a potential con-
founder, and results from some of the cognitive tests as covariates of interest. Only
one test of interest was included at any time. The aim of these tests was see if
correlations found by others were confirmed in these data. Cognitive tests used
were naming [Mummery et al., 2000], pyramids and palmtrees [Zahn et al., 2004],
famous faces (recognition) [Snowden et al., 2004], and guilt from the evaluation of
social actions [Moll et al., 2011]. Controls were not included in these analyses as
they all scored close to the top score. These regressions were only carried out with
the MD images uncorrected for CSF contamination, as these images were expec-
ted to be more sensitive to any differences between patients and controls. These
tests are independent, and so no correction for the number of regression tests was
made. Cluster-based correction was made within each regression for the num-
ber of voxels tested, as stated within the individual results. I hypothesis that, is a
correlation is found, MD will increase as the score decreases.

7.2.4 Regional analysis

A regional analysis was also carried out using the grey and white matter regions
described earlier (5.2.3.2. This analysis was carried out on the unsmoothed im-
ages for both FA and MD, and for images both with and without correction for
b-volume effects. The value within each ROI was calculated with in-house software; the median value was taken for each ROI in order to minimise the effect of a few extreme values, e.g. where the voxel contains a significant amount of CSF. A Wilcoxon rank sum test was used to compare values within each ROI between patients and controls. Although the regional analysis provides a less visual way of comparing the two groups, it does not have the problems of multiple comparisons that can confuse the results from both a TBSS and a voxel-based analysis. It can also provide information on the size of any differences found: all the voxel-based methods supply is an indication of where the voxels differ, and not in the actual values found.

### 7.3 Results

#### 7.3.1 Voxel-based group comparison of FA using TBSS

Two further FTLD patients had to be removed from the analysis because of errors in the skeletonisation, leaving 8 patients and 14 controls in the final analysis. Figure 7.2 shows the results of the TBSS analysis of a group comparison of FA between patients and controls. The regions where the FA of patients is lower than that of controls are shown in red at Threshold-Free Cluster Enhanced \( p < 0.05 \). This is a binary plot; voxels are included or not, and hence there is no scale. The regions have been thickened to make them more visible. These regions are in the right and left inferior fronto-occipital fasciculus, the right forceps minor and the right and left anterior thalamic radiation. There are no regions at this significance level where patients have a higher FA than controls.

#### 7.3.2 Voxel-based analyses of MD using SPM8

##### 7.3.2.1 Comparison of patient and control groups

10 FTLD patients (6M, 65.9±8.9y) and 14 controls (9M, 61.2±10.9y) were included in the following analyses.

Figure 7.3 shows the results of the voxel-based analysis comparing mean
diffusivity between patients and controls, and figures 7.4, 7.5 the results of the masked analyses. In all figures regions where patients have a higher mean diffusivity than controls are shown in yellow/red, and where patients have lower mean diffusivity in blue/green. Tables 7.1 and 7.2 show the location in MNI coordinates, the cluster size, the uncorrected voxel p-value and cluster-based corrected p-values for clusters shown in 7.3, 7.4 that are significant after cluster-based correction. There are no such significant clusters for the analysis shown in 7.5.

The comparison of unmasked MD images (figure 7.3) shows that there are clusters significant at cluster-level p (FDR-corrected) <0.05. Patients have higher MD than controls in the L postcentral, L superior temporal pole and L caudate, and in the R mid-cingulum. There are no regions in unmasked images where patients have lower MD than controls at this significance level.

This comparison of masked MD images, with two different ways of generating the masked images, shows the effect that pre-processing can have on the final result. Figure 7.4 has been generated by creating a binary mask for each individual subject where voxels within the brain which include > 20% probability of CSF are set to 0, and voxels within the brain which have < 20% probability of CSF are set
to 1. The individual MD images are then AND-ed with the mask, such that voxels within the brain tissue are left untouched, and voxels within the CSF are set to 0. The individual images are then smoothed with a Gaussian kernel of 12mm, and the resulting images used in a voxel-based analysis, with the analysis masked to exclude any voxel excluded from any individual. The regions where patients have a higher MD than controls are similar to those in 7.3, but in the regions around the amygdala the difference is reversed, such that patients have a lower MD than controls. Clusters significant at cluster-p (FDR-corrected) < 0.05 are shown in table 7.2.

Figure 7.5 has been generated by again creating a binary mask for each subject as above, but this time the MD voxels within the CSF regions are set to NaN instead of 0. and smoothed with the same Gaussian kernel of 12mm, but preserving the NaN voxels within the CSF. (However, the smoothing algorithm treats NaN as 0 when smoothing into adjacent voxels.) These images are then used in a similar voxel-based analysis to that used to generate both 7.3 and 7.4. In this analysis, the regions where patients have a higher MD than controls are much smaller to those in 7.4 but the regions around the amygdala now have a similar MD to controls. There are no clusters in this analysis where cluster-p (FDR-corrected)< 0.5.

The apparently significant clusters seen in figure 7.4 where the MD of patients is less than that of controls may be thus an artefact of the pre-processing: the value of voxels close to the CSF have been pulled down by voxels within the CSF during the smoothing step. As patients have more atrophy than controls, there is likely to be a confounding difference between the groups. When the masking is done by vigorously excluding these voxels from the analysis the large differences seen in the unmasked analysis disappear. On this data there is no reason to suspect that the MD of the remaining tissue differs between patients and controls. The regions shown in figure 7.5 where the MD of patients is greater than that of controls may indicate such a difference, but they do not approach significance when corrected for multiple comparisons.
Figure 7.3: Axial and coronal images of t-test results (p=0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons) where whole-brain mean diffusivity differs between patients and controls. Red/yellow shows where patients have higher mean diffusivity than controls; blue/green where patients have lower mean diffusivity. These images have no masking to exclude CSF.

Figure 7.4: Axial and coronal images of t-test results (p < 0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons) where mean diffusivity differs between patients and controls. Red/yellow shows where patients have higher mean diffusivity than controls; blue/green where patients have lower mean diffusivity. Images have been masked to exclude voxels with more than 20% probability of CSF, but the excluded voxels have been set to 0. and these voxels seem to affect the analysis.
Table 7.1: Cluster statistics showing regions where mean diffusivity differs between patients and controls after FDR-correction at cluster level for multiple comparisons. Images have not been masked to exclude CSF.

<table>
<thead>
<tr>
<th>cluster</th>
<th>cluster p</th>
<th>MNI coords</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(FDR_corr)</td>
<td>x   y   z</td>
</tr>
<tr>
<td>L postcentral</td>
<td>&lt;0.01</td>
<td>-26 -40 -4</td>
</tr>
<tr>
<td>L sup temp pole</td>
<td>&lt;0.005</td>
<td>-50 6 -22</td>
</tr>
<tr>
<td>R mid cingulum</td>
<td>&lt;0.005</td>
<td>2   -6 30</td>
</tr>
<tr>
<td>L caudate</td>
<td>&lt;0.05</td>
<td>28  20 38</td>
</tr>
</tbody>
</table>

Table 7.2: Cluster statistics showing regions where mean diffusivity differs between patients and controls after FDR-correction at cluster level for multiple comparisons when voxels with >20% CSF have been set to zero.

<table>
<thead>
<tr>
<th>cluster</th>
<th>cluster p</th>
<th>MNI coords</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(FDR_corr)</td>
<td>x   y   z</td>
</tr>
<tr>
<td>R thalamus</td>
<td>&lt;0.01</td>
<td>16 -24 1</td>
</tr>
<tr>
<td>R amygdala</td>
<td>&lt;0.01</td>
<td>26 2 -18</td>
</tr>
</tbody>
</table>

Figure 7.5: Axial and coronal images of t-test results (p < 0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons) where mean diffusivity differs between patients and controls. Red/yellow shows where patients have higher mean diffusivity than controls; blue/green where patients have lower mean diffusivity. Images have been masked to exclude voxels with more than 20% probability of CSF; the excluded voxels have been set to NaN, indicating that they contain a non-numeric value.
7.3.2.2 Regression of naming scores

Figure 7.6 shows the results of the voxel-based analysis of a regression of the naming scores for patients only with the MD, such that MD increases as the score decreases. 9 patients are included in this analysis; one patient did not complete this test. In view of the very small number of patients in this analysis, its results must be considered very tentative. The cluster at [-44 -4 -26] in the left mid-temporal lobe almost reaches significance, with a cluster-based regional FDR-corrected p-value (within the right and left temporal regions) of 0.075.

![Axial and coronal images of t-test results](image)

**Figure 7.6:** Axial and coronal images of t-test results ($p < 0.005$, minimum cluster size 75 voxels, uncorrected for multiple comparisons) for regression of patients’ naming scores. The images show where the MD increases as the patient score decreases.
7.3.2.3 Regression of pyramids and palm trees scores

Figure 7.7 shows the results of the voxel-based analysis of a regression of the pyramids and palmtrees scores for patients only with the MD, such that MD increases as the score decreases. 9 patients are included in this analysis; one patient did not complete this test. In view of the very small number of patients in this analysis, its results must be considered very tentative. Clusters in both the right and left temporal lobes are significant when corrected for multiple comparisons with within the right and left temporal regions. The cluster in the left inferior temporal lobe of 334 voxels centered on [-38 -8 -28] has a cluster-based FDR-corrected p-value of 0.003, and the cluster of 34 voxels in the right superior temporal lobe centered on [52 -40 0] has a cluster-based FDR-corrected p-value of 0.002.

Figure 7.7: Axial and coronal images of t-test results ($p < 0.005$, minimum cluster size 75 voxels, uncorrected for multiple comparisons) for regression of patients’ pyramids and palmtrees scores. The images show where the MD increases as the patient score decreases.
7.3.2.4 Regression of famous faces recognition scores

Figure 7.8 shows the results of the voxel-based analysis of a regression of the famous faces recognition scores for patients only with the MD, such that MD increases as the score decreases. 8 patients are included in this analysis. The cluster in the right temporal lobe is significant when corrected for multiple comparisons with within the right temporal regions. This cluster of 407 voxels centered on [52 -50 0] has a cluster p-value < 0.001.

Figure 7.8: Axial and coronal images of t-test results (p=0.005, minimum cluster size 75 voxels, uncorrected for multiple comparisons) for regression of patients’ famous faces recognition scores. The maps show where the MD increases as the patient score decreases.
7.3.2.5 Evaluation of social actions

No voxels were found in this analysis which had a significance level of < 0.005, minimum cluster size 75 voxels.

7.3.3 Regional analysis

10 patients and 14 controls were included in these analyses. Tables 7.3 and 7.4 show mean diffusivity and fractional anisotropy respectively, of the mean value ± standard deviation within each GM ROI for patients and controls. Tables 7.5 and 7.6 show mean diffusivity and fractional anisotropy respectively, of the mean value ± standard deviation within the WM ROIs.

Table 7.3: MD values in aal regions for patients and controls, Values shown are mean±SD within each ROI for the patients and control group. The analysis includes 10 patients and 14 controls. Values for raw MD and MD with CSF < 20% probability are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p < 0.05; cells in medium green have p < 0.01 and cells in dark green have p < 0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw MD (10⁻³mm²/s)</th>
<th>Excluding CSF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LPostTL</td>
<td>0.903±0.086</td>
<td>&lt; 10⁻³</td>
</tr>
<tr>
<td>RPostTL</td>
<td>0.924±0.124</td>
<td>0.01</td>
</tr>
<tr>
<td>LAntTL</td>
<td>1.226±0.305</td>
<td>&lt; 10⁻³</td>
</tr>
<tr>
<td>RAntTL</td>
<td>1.135±0.243</td>
<td>&lt; 10⁻³</td>
</tr>
<tr>
<td>PostVMPFC</td>
<td>0.902±0.125</td>
<td>0.01</td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>1.027±0.160</td>
<td>0.009</td>
</tr>
<tr>
<td>Calcarine</td>
<td>0.890±0.075</td>
<td>0.07</td>
</tr>
<tr>
<td>Cereb6</td>
<td>0.866±0.085</td>
<td>0.08</td>
</tr>
</tbody>
</table>

In all aal ROIs known to be affected by the illness, both raw and masked MD are higher in patients then in controls. The difference is greater for the raw MD values. In aal regions known not to be affected, there is no significant difference. The only white matter tracts where there is no difference are the R and L cingulum-cingulate gyrus. This may be because of the size of this tract, which includes a large part of the posterior and anterior cingulum bundle. Other studies have shown that damage to the posterior cingulum bundle is a hallmark of Alzheimer’s dis-
Table 7.4: FA values in aal regions for patients and controls. Values shown are mean±SD within each ROI for the patients and control group. The analysis includes 10 patients and 14 controls. Values for raw FA and FA with CSF < 20% probability are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw FA</th>
<th>FA excluding CSF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LPostTL</td>
<td>0.138±0.019</td>
<td>1.00</td>
</tr>
<tr>
<td>RPostTL</td>
<td>0.118±0.016</td>
<td>0.46</td>
</tr>
<tr>
<td>LAntTL</td>
<td>0.113±0.011</td>
<td>0.05</td>
</tr>
<tr>
<td>RAntTL</td>
<td>0.103±0.019</td>
<td>0.08</td>
</tr>
<tr>
<td>PostVMPFC</td>
<td>0.165±0.012</td>
<td>0.62</td>
</tr>
<tr>
<td>AntVMPFC</td>
<td>0.120±0.012</td>
<td>0.23</td>
</tr>
<tr>
<td>Calcarine</td>
<td>0.145±0.019</td>
<td>0.14</td>
</tr>
<tr>
<td>Cereb6</td>
<td>0.139±0.010</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Table 7.5: MD values in JHU regions for patients and controls in ROIs thresholded at 20% probability. Values shown are mean±SD within each ROI for the patients and control group. The analysis includes 10 patients and 14 controls. Values for raw MD and MD with CSF < 20% probability are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw MD (10⁻³mm²/s)</th>
<th>MD excluding CSF (10⁻³mm²/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
</tr>
<tr>
<td>LCCg</td>
<td>0.775±0.093</td>
<td>0.23</td>
</tr>
<tr>
<td>RCCG</td>
<td>0.778±0.084</td>
<td>0.70</td>
</tr>
<tr>
<td>LCH</td>
<td>0.952±0.205</td>
<td>0.14</td>
</tr>
<tr>
<td>RCH</td>
<td>0.946±0.125</td>
<td>0.15</td>
</tr>
<tr>
<td>LIFOF</td>
<td>0.823±0.066</td>
<td>0.01</td>
</tr>
<tr>
<td>RIFOF</td>
<td>0.805±0.058</td>
<td>0.01</td>
</tr>
<tr>
<td>LILF</td>
<td>0.831±0.088</td>
<td>0.02</td>
</tr>
<tr>
<td>RILF</td>
<td>0.798±0.067</td>
<td>0.02</td>
</tr>
<tr>
<td>LSLF</td>
<td>0.745±0.035</td>
<td>0.02</td>
</tr>
<tr>
<td>RSLF</td>
<td>0.698±0.030</td>
<td>0.93</td>
</tr>
<tr>
<td>LUF</td>
<td>0.923±0.130</td>
<td>0.03</td>
</tr>
<tr>
<td>RUF</td>
<td>0.895±0.086</td>
<td>0.03</td>
</tr>
</tbody>
</table>

ease, whereas in FTLD patients the damage is to the anterior part [Boxer et al., 2003]. In all the other WM tracts considered, MD both raw and masked is higher in patients than controls. FA does not show such significant differences, but the anterior temporal lobes show a lower FA in patients than in controls, and the uncinate fasciculus has a lower FA in patients in both the left and the right tracts. As expected, the lower MD seen in patients when the CSF is improperly masked
Table 7.6: FA values in JHU regions for patients and controls in ROIs thresholded at 20% probability. Values shown are mean±SD within each ROI for the patients and control group. The analysis includes 10 patients and 14 controls. Values for raw FA and FA with CSF < 20% probability are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw FA</th>
<th>FA excluding CSF</th>
<th>p</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>0.06</td>
<td>0.194±0.042</td>
<td>0.245±0.065</td>
</tr>
<tr>
<td></td>
<td>RCCG</td>
<td>0.242±0.083</td>
<td>0.252±0.090</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>LCH</td>
<td>0.224±0.025</td>
<td>0.223±0.026</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>RCH</td>
<td>0.207±0.034</td>
<td>0.206±0.033</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>LIFOF</td>
<td>0.265±0.019</td>
<td>0.297±0.027</td>
<td>0.270±0.016</td>
</tr>
<tr>
<td></td>
<td>RIFOF</td>
<td>0.280±0.028</td>
<td>0.293±0.025</td>
<td>0.281±0.030</td>
</tr>
<tr>
<td></td>
<td>LILF</td>
<td>0.278±0.023</td>
<td>0.303±0.027</td>
<td>0.296±0.017</td>
</tr>
<tr>
<td></td>
<td>RILF</td>
<td>0.302±0.029</td>
<td>0.316±0.027</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>LSLF</td>
<td>0.224±0.030</td>
<td>0.248±0.027</td>
<td>0.233±0.025</td>
</tr>
<tr>
<td></td>
<td>RSLF</td>
<td>0.257±0.035</td>
<td>0.252±0.025</td>
<td>0.267±0.036</td>
</tr>
<tr>
<td></td>
<td>LUF</td>
<td>0.198±0.032</td>
<td>0.223±0.042</td>
<td>0.215±0.033</td>
</tr>
<tr>
<td></td>
<td>RUF</td>
<td>0.224±0.043</td>
<td>0.246±0.040</td>
<td>0.248±0.041</td>
</tr>
</tbody>
</table>

Table 7.7: MD values in JHU regions for patients and controls in ROIs thresholded at 10% probability. Values shown are mean±SD within each ROI for the patients and control group. The analysis includes 10 patients and 14 controls. Values for raw MD and MD with CSF < 10% probability are shown. The significance shown is the p-value from a Wilcoxon ranked sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw MD (10⁻³mm²/s)</th>
<th>MD excluding CSF (10⁻³mm²/s)</th>
<th>p</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>0.10</td>
<td>0.730±0.026</td>
<td>0.766±0.063</td>
</tr>
<tr>
<td></td>
<td>RCCG(10pc)</td>
<td>0.791±0.073</td>
<td>0.756±0.033</td>
<td>0.774±0.074</td>
</tr>
<tr>
<td></td>
<td>LCH(10pc)</td>
<td>0.978±0.181</td>
<td>0.844±0.056</td>
<td>0.971±0.177</td>
</tr>
<tr>
<td></td>
<td>RCH(10pc)</td>
<td>0.976±0.141</td>
<td>0.864±0.074</td>
<td>0.966±0.146</td>
</tr>
<tr>
<td></td>
<td>LIFOF(10pc)</td>
<td>0.816±0.061</td>
<td>0.751±0.033</td>
<td>0.807±0.056</td>
</tr>
<tr>
<td></td>
<td>RIFOF(10pc)</td>
<td>0.800±0.052</td>
<td>0.747±0.031</td>
<td>0.794±0.051</td>
</tr>
<tr>
<td></td>
<td>LILF(10pc)</td>
<td>0.834±0.077</td>
<td>0.749±0.028</td>
<td>0.800±0.043</td>
</tr>
<tr>
<td></td>
<td>RILF(10pc)</td>
<td>0.806±0.060</td>
<td>0.740±0.030</td>
<td>0.778±0.048</td>
</tr>
<tr>
<td></td>
<td>LSLF(10pc)</td>
<td>0.774±0.043</td>
<td>0.734±0.040</td>
<td>0.759±0.037</td>
</tr>
<tr>
<td></td>
<td>RSLF(10pc)</td>
<td>0.733±0.028</td>
<td>0.716±0.037</td>
<td>0.718±0.027</td>
</tr>
<tr>
<td></td>
<td>LUF(10pc)</td>
<td>0.872±0.126</td>
<td>0.768±0.042</td>
<td>0.840±0.098</td>
</tr>
<tr>
<td></td>
<td>RUF(10pc)</td>
<td>0.941±0.168</td>
<td>0.799±0.042</td>
<td>0.887±0.133</td>
</tr>
</tbody>
</table>

from the image does not appear in this ROI analysis. Analysis of an additional ROI including only the left and right amygdala (where the lower MD was seen in
Table 7.8: FA values in JHU regions for patients and controls in ROIs thresholded at 10% probability. Values shown are mean±SD within each ROI for the patients and control group. The analysis includes 10 patients and 14 controls. Values for raw FA and FA with CSF < 10% probability are shown. The significance shown is the p-value from a Wilcoxon rank sum test. Cells shown in light green have p<0.05; cells in medium green have p<0.01 and cells in dark green have p<0.001.

<table>
<thead>
<tr>
<th>Region</th>
<th>Raw FA</th>
<th></th>
<th></th>
<th>FA excluding CSF</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>p</td>
<td>Controls</td>
<td>Patients</td>
<td>p</td>
<td>Controls</td>
</tr>
<tr>
<td>LCCg(10pc)</td>
<td>0.239±0.045</td>
<td>0.07</td>
<td>0.211±0.028</td>
<td>0.241±0.046</td>
<td>0.07</td>
<td>0.214±0.028</td>
</tr>
<tr>
<td>RCCG(10pc)</td>
<td>0.214±0.063</td>
<td>0.04</td>
<td>0.164±0.029</td>
<td>0.226±0.077</td>
<td>0.06</td>
<td>0.167±0.031</td>
</tr>
<tr>
<td>LCH(10pc)</td>
<td>0.215±0.025</td>
<td>0.62</td>
<td>0.217±0.031</td>
<td>0.213±0.024</td>
<td>0.43</td>
<td>0.217±0.030</td>
</tr>
<tr>
<td>RCH(10pc)</td>
<td>0.197±0.027</td>
<td>0.50</td>
<td>0.196±0.022</td>
<td>0.195±0.026</td>
<td>0.62</td>
<td>0.193±0.023</td>
</tr>
<tr>
<td>LIFOF(10pc)</td>
<td>0.244±0.014</td>
<td>0.00</td>
<td>0.272±0.019</td>
<td>0.249±0.010</td>
<td>0.003</td>
<td>0.273±0.020</td>
</tr>
<tr>
<td>RIFOF(10pc)</td>
<td>0.255±0.022</td>
<td>0.11</td>
<td>0.271±0.021</td>
<td>0.257±0.024</td>
<td>0.12</td>
<td>0.272±0.021</td>
</tr>
<tr>
<td>LILF(10pc)</td>
<td>0.239±0.018</td>
<td>0.003</td>
<td>0.264±0.019</td>
<td>0.261±0.016</td>
<td>0.12</td>
<td>0.270±0.021</td>
</tr>
<tr>
<td>RILF(10pc)</td>
<td>0.258±0.025</td>
<td>0.25</td>
<td>0.271±0.026</td>
<td>0.276±0.026</td>
<td>0.98</td>
<td>0.274±0.028</td>
</tr>
<tr>
<td>LSLF(10pc)</td>
<td>0.190±0.016</td>
<td>0.04</td>
<td>0.205±0.020</td>
<td>0.205±0.014</td>
<td>0.07</td>
<td>0.219±0.022</td>
</tr>
<tr>
<td>RSLF(10pc)</td>
<td>0.214±0.022</td>
<td>0.58</td>
<td>0.217±0.017</td>
<td>0.231±0.024</td>
<td>0.84</td>
<td>0.232±0.019</td>
</tr>
<tr>
<td>LUF(10pc)</td>
<td>0.196±0.018</td>
<td>0.02</td>
<td>0.225±0.029</td>
<td>0.212±0.017</td>
<td>0.06</td>
<td>0.234±0.030</td>
</tr>
<tr>
<td>RUF(10pc)</td>
<td>0.209±0.028</td>
<td>0.08</td>
<td>0.236±0.029</td>
<td>0.242±0.027</td>
<td>0.46</td>
<td>0.254±0.030</td>
</tr>
</tbody>
</table>

One voxel-based analysis) shows that the raw MD is $1.046±0.214 \times 10^{-3} \text{mm}^2/\text{s}$ in the patient group and $0.831±0.023 \times 10^{-3} \text{mm}^2/\text{s}$ in the control group. The equivalent values when CSF is excluded are $0.910±0.123 \times 10^{-3} \text{mm}^2/\text{s}$ and $0.805±0.019 \times 10^{-3} \text{mm}^2/\text{s}$.

### 7.4 Discussion

Two key facts have emerged from these data:

- Mean diffusivity when not masked to exclude CSF is a very sensitive measure of differences between FTLD patients and controls, being more sensitive than measures of grey matter atrophy.

- Regional analysis is a powerful tool to see whether MD or fractional anisotropy is different in the tissue of controls and the remaining tissue of FTLD patients.

Mean diffusivity is known to be a good measure of oedema [Le Bihan, 2003] because of the very high value of MD in CSF vs GM or WM. From a control in my...
study, the MD in CSF is 1.98±0.08 whereas in a region containing both GM and WM it is 0.64±0.7 i.e. MD of CSF is of the order of 2 to 4 times that of tissue. Hence the two measures of MD considered in this study are measuring different aspects of the illness. An analysis of the whole-brain MD including regions close to the tissue/CSF boundary is likely to be dominated by the voxels with a high partial volume of CSF, and therefore by localised atrophy. An analysis confined to voxels with a low CSF (known from segmentation of the T1-weighted image) would be expected to tell us something about the remaining tissue. The group analysis of raw MD data shows large regions where the MD of patients is greater than the MD of controls, and these are around the regions where from the T1-weighted images there is known to be greater atrophy. The picture presented by the analysis of masked MD is more complicated; the regional analysis shows significant differences in the part of the cingulum bundle connected to the cingulum-hippocampus, the inferior fronto-occipital and longitudinal fasciculus, the superior longitudinal fasciculus and the uncinate fasciculus; most of these differences are bilateral. The voxel-based analysis may be affected by the smoothing necessary for such an analysis, which will spread the higher MD values within the CSF to nearby voxels, but this will not affect the regional analysis, for which there is no smoothing.

Table 7.9 shows papers publishing details of diffusion MRI studies on patients with FTLD. It is a technique that is gaining acceptance, with 9 papers published in 2012 - 2014 compared with one each year from 2007 - 2010. It is also noticeable and frustrating that there are many different ways of describing this heterogeneous group of patients; this is described in more detail earlier (see chapter 1.2). The patients in my study are difficult to classify into bvFTD or SD, as practically all demonstrate symptoms of both, and a classification would rely on a partner’s possibly faulty memory of the first presenting symptom. Semantic-variant PPA (svPPA) is a diagnosis with many similarities to SD, frontal-variant FTD (fvFTD) might now be described as bvFTD, and temporal-variant FTD (tvFTD) as SD. ALS and FTD are now recognised to be on a spectrum of neurodegenerative disorders, with patients with a diagnosis of one disease often exhibiting symptoms of the other.

There is also heterogeneity in the diffusion metrics reported on. All of these papers report differences in fractional anisotropy, but about half also report other diffusion metrics. Acosta-Cabronero et al [Acosta-Cabronero et al., 2011], Galantucci et al [Galantucci et al., 2011] and Santillo at al [Santillo et al., 2013] find greater
differences between patients and controls in MD rather than FA, although Sajjadi et al. [Sajjadi et al., 2013] reports that “Fractional anisotropy was consistently the most sensitive metric although it was followed closely by radial diffusivity”; the study did not report on mean diffusivity. Mahoney et al. [Mahoney et al., 2014] finds radial diffusivity to be the most sensitive metric, closely followed by mean diffusivity. In common with this study, all studies which compare diffusion metrics with grey matter atrophy find the diffusion metrics to give a greater difference between patients and controls [Santillo et al., 2013], [Mahoney et al., 2013], [Mahoney et al., 2014]. All studies find decreased fractional anisotropy and/or increased mean diffusivity in patients. All papers report differences in diffusion metrics in patients in the uncinate fasciculus, and many also report differences in the inferior and superior longitudinal fasciculus.

The raw MD values in patients are 17% higher than in controls in both aal regions (predominantly grey matter, and 10% higher) in JHU regions (predominantly white matter) in all ROIs considered except the part of the cingulum bundle connected to the cingulate gyrus. These differences are likely to be useful as a diagnostic tool. When masked to exclude CSF, MD is still increased in patients compared with controls in the same regions as the uncorrected MD. This may reflect an underlying pathology; the first descriptions of “Pick’s disease” describe “ballooned cells” [Kertesz et al., 2000].

FA in this study is less sensitive than MD, but still reduced in patients compared with controls by approximately 5%, again in images both unmasked and masked to exclude CSF. The most significant change is seen in the left inferior fronto-occipital fasciculus and anterior temporal lobe, but changes are also seen in uncorrected FA in the left inferior-longitudinal-fasciculus, superior-longitudinal-fasciculus and the uncinate-fasciculus. Both of these changes could be caused by swollen cells.
Table 7.9: Published studies of diffusion data in FTLD.

Patient group abbreviations: AD=Alzheimer’s Disease, HC=healthy control, SD=semantic dementia, bv=bevavioural-variant FTD, fvFTD=frontal-variant FTD, tvFTD=temporal variant FTD, PPA=primary progressive aphasia, nvPPA=non-fluent PPA, lvPPA=logopenic variant PPA, svPPA=semantic variant PPA, PNFA=progressive non-fluent aphasia, CBS=corticobasal syndrome, ALS=Amyotrophic Lateral Sclerosis, FTLD-TDP=FTLD with TDP pathology, FTLD-tau=FTLD with tau pathology.

Diffusion metrics abbreviations: FA=fractional anisotropy, AxD=axial diffusivity, MD=mean diffusivity, RD=radial Diffusivity.

<table>
<thead>
<tr>
<th>Lead author</th>
<th>Year</th>
<th>Field (T)</th>
<th>b=0 s/mm²</th>
<th>No of dirs</th>
<th>Patient group</th>
<th>DTI metrics</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borroni</td>
<td>2007</td>
<td>1.5</td>
<td>1000</td>
<td>6</td>
<td>28 fvFTD, 8 tvFTD, 23 HC</td>
<td>FA</td>
<td>VB(SPM2)</td>
</tr>
<tr>
<td>Matsuo</td>
<td>2008</td>
<td>1.5</td>
<td>1000</td>
<td>15</td>
<td>14 bvFTD, 6 SD, 17 HC</td>
<td>FA</td>
<td>ROI</td>
</tr>
<tr>
<td>Zhang</td>
<td>2009</td>
<td>4</td>
<td>800</td>
<td>6</td>
<td>18 FTD, 18 AD, 19 HC</td>
<td>FA</td>
<td>ROI, whole brain</td>
</tr>
<tr>
<td>Whitwell</td>
<td>2010</td>
<td>3</td>
<td>?</td>
<td>21</td>
<td>16 bvFTD, 7 PNFA, 4 SD, 19 HC</td>
<td>FA, MD, RD</td>
<td>ROI</td>
</tr>
<tr>
<td>Acosta-Cabronero</td>
<td>2011</td>
<td>3</td>
<td>1000</td>
<td>63</td>
<td>10 SD, 21 HC</td>
<td>AxD, RD, MD, FA</td>
<td>VB(TBSS)</td>
</tr>
<tr>
<td>Galantucci</td>
<td>2011</td>
<td>3</td>
<td>2000</td>
<td>64</td>
<td>9 svPPA, 9 lvPPA, 21 HC</td>
<td>FA, MD, AxD, RD</td>
<td>ROI</td>
</tr>
<tr>
<td>Agosta</td>
<td>2012</td>
<td>3</td>
<td>1000</td>
<td>32</td>
<td>13 bvFTD, 9 nvPPA, 7 svPPA, 4 lvPPA</td>
<td>AxD, RD, MD, FA</td>
<td>ROI, whole brain</td>
</tr>
<tr>
<td>Lillo</td>
<td>2012</td>
<td>3</td>
<td>1000</td>
<td>32</td>
<td>10 ALS, 10 ALS-FTD, 15 bvFTD, 18 HC</td>
<td>FA</td>
<td>whole brain</td>
</tr>
<tr>
<td>McMillan</td>
<td>2012</td>
<td>3</td>
<td>1000/7</td>
<td>30/12</td>
<td>25 FTLD-TDP, 10 FTLD-tau.</td>
<td>FA</td>
<td>ROI</td>
</tr>
<tr>
<td>Santillo</td>
<td>2013</td>
<td>3</td>
<td>800</td>
<td>48</td>
<td>14 bvFTD, 18 HC</td>
<td>FA, MD, AxD, RD</td>
<td>ROI</td>
</tr>
<tr>
<td>Mahoney</td>
<td>2013</td>
<td>3</td>
<td>1000</td>
<td>64</td>
<td>13 nvPPA, 10 lvPPA, 20 HC</td>
<td>AxD, Rd, FA, tr</td>
<td>VB(TBSS)</td>
</tr>
<tr>
<td>Sajjadi*</td>
<td>2013</td>
<td>3</td>
<td>1000</td>
<td>63</td>
<td>9 SD, 9 nvPPA, 9 AD, 26 HC</td>
<td>AxD, RD, MD, FA</td>
<td>VB(TBSS)</td>
</tr>
<tr>
<td>Schwindt</td>
<td>2013</td>
<td>3</td>
<td>1000</td>
<td>23</td>
<td>9 svPPA, 9 svPPA, 16 HC</td>
<td>FA</td>
<td>VB(TBSS)</td>
</tr>
<tr>
<td>Mahoney</td>
<td>2014</td>
<td>3</td>
<td>1000</td>
<td>64</td>
<td>27 bvFTD, 25 AD, 20 HC</td>
<td>AxD, RD, FA, tr</td>
<td>ROI, whole brain</td>
</tr>
<tr>
<td>TovarMoll</td>
<td>2014</td>
<td>3</td>
<td>1000</td>
<td>32</td>
<td>20 bvFTD, 19 CBS, 15 HC</td>
<td>FA, MD</td>
<td>ROI, VB(TBSS)</td>
</tr>
</tbody>
</table>
Chapter 8

Multimodal modelling in frontotemporal lobar degeneration

8.1 Introduction

So far in this thesis I have considered the different modalities of structural, perfusion and diffusion data in isolation. However, all three modalities might be available to a clinician trying to reach a diagnosis, and in this chapter I consider which modalities are most useful when distinguishing between patients and controls. Such an analysis combines the power of the individual modalities, and could be used to generate a diagnostic support tool to aid the clinician.

I have not included results from the neuropsychological tests in this analysis. For this particular group of patients these tests alone can give a good classification, and so including them might mask data from the MRI analyses. My study is of people who already have a diagnosis, and therefore are already some way through the degenerative process; in the early stages diagnosis is a challenge [Warren et al., 2013]. Once developed such an analysis might also prove to be capable of tracking disease progression; MRI images are almost infinitely repeatable, unlike PET, SPECT or neuropsychological testing. PET and SPECT cannot be used too often because of the radiation exposure they bring to the patient, and the same neuropsychological tests cannot be used frequently even with people with poor memory.
I have only considered MRI values from images without partial-volume correction; these values when considered unimodally give a clearer distinction between the groups. The values corrected for partial volume effects are useful in showing differences in tissue between the groups, perhaps giving an indication of the causes underlying the illness, but for classifying the subjects it matters only that there is a difference, not whether that difference is due to changes in tissue structure or to atrophy.

Other groups have also looked at data from different MRI modalities, although they have not always been interested in using the data for classification.

Santillo et al [Santillo et al., 2013] compared $T_1$-weighted and diffusion images in a group of 14 patients with bvFTD and 22 healthy controls. They looked at the anterior cingulate cortex, using mean, axial, and radial diffusivity, and fractional anisotropy from manually-drawn ROIs in the diffusion images. They used voxel-based analysis and cortical thickness measures to compare $T_1$-weighted images, and found statistically significant differences in all diffusion parameters. In the left hemisphere, cortical thickness correctly classified 78% of cases and voxel-based analysis 83%. In the anterior part of the cingulum bundle, FA correctly classified 84%, MD 90%, axial diffusivity 88%, and radial diffusivity 92%. They considered all parameters individually.

Zhang et al used $T_1$-weighted, perfusion and diffusion MRI to look at a group of 20 AD patients, 20 FTD patients and 21 healthy controls [Zhang et al., 2011]. They considered CBF and GM atrophy in voxels with $>85\%$ GM, and FA in voxels with $>80\%$ WM. CBF values were corrected for partial-volume effects assuming that CBF in WM is 25% of that in GM. They calculated the T-score of each voxel in each modality, as compared to the mean of the voxel in the combined control group; this calculation is described by Signorini et al [Signorini et al., 1999]. They then counted the load, defined as the number of voxels which were beyond a T-score threshold ranging from -2 (mild abnormality) to -6 (severe abnormality). They considered which modality had the highest load for each group. They found that FTD patients had GM loss in bilateral frontal and temporal lobes, decreased FA in frontal and temporal lobes, the anterior corpus callosum and bilateral anterior cingulum, and hypoperfusion in bilateral frontal lobes compared to controls; the GM loss and WM damage were more significant than the hypoperfusion. They did not use the data to classify subjects.
Tosun et al investigated a group of 12 bvFTD patients and 12 controls using T$_1$-weighted and perfusion images [Tosun et al., 2012]. The CBF images were collected with a single timepoint CASL sequence and corrected for partial-volume effects. They then analysed the images from each modality independently with voxel-based analysis, and multimodally with jICA [Calhoun and Adali, 2009]. The unimodal analyses showed widespread brain atrophy and hypoperfusion in the patients, but jICA further revealed two significant joint components of variations between atrophy and hypoperfusion across brain regions in bvFTD patients. The 1st joint component revealed associated brain atrophy and hypoperfusion predominantly in the right brain hemisphere; the 2nd revealed greater atrophy relative to hypoperfusion predominantly in the left hemisphere. They did not try to classify subjects with these data.

Woost et al collected data on cerebral glucose utilisation (as measured by F-18-fluorodeoxyglucose positron emission tomography) and grey matter atrophy (as measured by a T$_1$-weighted image) from patients with AD (21), FTLD (n = 14) or subjective cognitive impairment (n = 13) as a control group. They used the segmentation of the MRI image to perform partial-volume correction on the glucose utilisation images, and then analysed both sets of images independently using voxel-based analysis [Woost et al., 2013]. They reported that atrophy dominated in FTLD, whereas hypometabolism in AD. Their principal aim was to validate the DemTect [Kalbe et al., 2004] as a valid screening device for an early diagnosis of dementia, and so they looked for correlations between regional atrophy and hypometabolism, and DemTect scores.

Only McMillan et al have used multiple MRI modalities in combination to classify subjects, McMillan et al [McMillan et al., 2012], [McMillan et al., 2013]. In the 2012 paper they used diffusion imaging and 3D T$_1$-weighted volumetric imaging to distinguish between patients with AD (34), FTLD (21 PPA, 27 bvFTD, 10 CBS) and healthy controls (38), looking at FA in the anterior corpus callosum and GM density in cortical regions. They considered each modality independently, and the combination of both. They found reduced FA in patients with FTLD in all the tracts they considered (bilateral corticospinal tract, corpus callosum, IFOF, ILF, SLF and UF), with more prominent differences in the anterior portions. GM density was reduced in FTLD patients throughout the frontal and anterior temporal cortex. For classification they used linear regression on the FA value in the anterior corpus callosum and GM density in the precuneus, posterior cingulate
and anterior temporal cortex. They found 79% sensitivity and 59% specificity with WM alone, 87% sensitivity and 66% specificity for GM in the posterior cingulate and 82% specificity and 79% sensitivity in the precuneus. The best classification came from the combination of WM and GM together, with 87% sensitivity and 83% specificity (33 out of 38 FTLD patients and 24 out of 29 AD patients correctly classified). They also commented that individuals with AD and FTLD have significant changes in WM and GM that appear to reflect underlying neuropathological processes. In the 2013 paper they extracted cortical thickness in 90 GM ROIs, and regional FA in 48 WM ROIs, and used the 138 values, together with the non-MRI parameters age, MMSE score, gender, disease duration and APOE status to classify 93 patients with clinically diagnosed FTD into those with AD-type pathology, and those with FTLD-TDP or FTLD-tau pathology. APOE status is a known genetic risk factor for AD [Burke and Roses, 1991]. Patients were classified as FTD if they carried a known genetic allele, had FTD pathology confirmed at postmortem, or if they lacked CSF markers indicating AD. They used a Eigenanatomy [Kandel et al., 2015], a sparse dimensionality reduction method, to identify the most significant volumes of interest (VOI) in an image, and concluded that data-driven VOI analysis using a multimodal combination of GM MRI and DTI achieved the greatest classification accuracy, and classified all of the FTLD-tau cases correctly, and 24 out of 25 of the FTLD-TDP cases.

8.2 Methods

Technologies exist now to allow voxel-based analysis of multimodal data, e.g. Eigenanatomy [Kandel et al., 2015], jICA [Calhoun and Adali, 2009]. However, a voxel-based analysis requires the data to be smoothed, possibly introducing unwanted artefacts (see figures 7.3, 7.4 and 7.5). The addition of multiple modalities also increases the number of comparisons being made, and compounds the problem of correcting for these to allow some certainty that any effect found is real and not a random effect (type I error), while also not rejecting too many real effects as random (type II error).

I have therefore chosen to extract a single ROI value for each subject for each modality and use that data in a linear discriminant analysis. The ROIs chosen were based on prior knowledge of the regions affected in FTLD, and also on their
availability and reproducibility. In this study there are 6 modalities available: GM density, WM density from $T_1$-weighted images, CBF and AAT from ASL images, and FA and MD from diffusion-weighted images. An additional complication is that CBF and AAT are not available in predominantly WM regions because of low SNR. The ROIs used have been described already (5.2.3.2). The aal ROIs known and seen to be unaffected in patients are not included (i.e. exclude calcarine, cerebellum6), and so 6 aal ROIs and 12 JHU ROIs were used. The atlas used to define the JHU ROIs is a probabilistic one, and I have chosen to use two minimum probabilistic values to define the JHU ROIs. One set of JHU ROIs was defined as where the probability for that ROI in the atlas is $> 0.1$, and the other set as where the JHU regional probability is $> 0.2$. The first set can be assumed to contain more voxels with an admixture of GM and/or CSF, and so might hold more information about atrophy than the second set, which may contain more information about the WM integrity. The two sets cannot be included in the same analysis as they break the assumption that variables should be independent of each other. CBF and arrival times within the JHU ROIs were not included.

There are 84 variables for each subject: GM density, WM density, FA and MD for the left posterior temporal lobe (LPostTL), the right posterior temporal lobe (RPostTL), the left anterior temporal lobe (LAntTL), the right anterior temporal lobe (RAntTL), the posterior ventro-medial prefrontal cortex (PostVMPFC), the anterior ventro-medial prefrontal cortex (AntVMPFC), the left part of the cingulum bundle connected to the cingulate gyrus (LCCg), the right part of the cingulum bundle connected to the cingulate gyrus (RCCg), the left part of the cingulum bundle connected to the cingulum-hippocampus (LCH), the right part of the cingulum bundle connected to the cingulum-hippocampus (RCH), the left inferior-fronto-occipital fasciculus (LIFOF), the right inferior-fronto-occipital fasciculus (RIFOF), the left inferior-longitudinal fasciculus (LILF), the right inferior-longitudinal fasciculus (RILF), the left superior-longitudinal fasciculus (LSLF), the right superior-longitudinal fasciculus (RSLF), the left uncinate fasciculus (LUF) and the right uncinate fasciculus (RUF);

CBF and AAT for the left posterior temporal lobe (LPostTL), the right posterior temporal lobe (RPostTL), the left anterior temporal lobe (LAntTL), the right anterior temporal lobe (RAntTL), the posterior ventro-medial prefrontal cortex (PostVMPFC), and the anterior ventro-medial prefrontal cortex (AntVMPFC).

A linear discriminant analysis was used, specifically the Matlab “classify”
function, which fits a multivariate normal density to each group, with a pooled estimate of covariance. (See http://uk.mathworks.com/help/stats/classify.html) This method was chosen as being standard and well-understood. The ROI values used have been reported earlier in the separate chapters. Because of the disparity in potential values, with CBF and arrival time being much larger than all other values, the analysis was carried out on z-standardised values. Neuropsychological variables were not included in the analysis as the interest was in which MRI modalities best separated patients from controls. In this particular subject group, the majority of the neuropsychological values are independently good classifiers (see table D.1).

Only subjects for whom all modalities are available can be included in this analysis; this included 8 FTLD patients and 14 controls. A “leave-one-out” approach was used: all analyses were carried out by iteratively training the classifier on 21 subjects, and then predicting the classification of the remaining subject for each subject in turn.

Types of analysis carried out:

1. classification with all variables
2. using Matlab “pca” (principal component analysis) to reduce the number of variables to 21, (the maximum allowed with 22 subjects; these account for 80% of the variance), and then classifying on the reduced data set
3. classification using only ROI variables seen to be significant at $p<0.05$, $p<0.01$ or $p<0.001$ in the univariate analysis. These items are listed in table 8.1
4. classification leaving out each pair of modalities from a particular image, i.e. GM and WM, CBF and arrival time, FA and MD
5. classification leaving out each modality in turn
6. classification on only 1 modality in turn
7. classification leaving out each ROI in turn
8. classification using only 1 ROI in turn

Test 2 is included as being similar to the analysis carried out by Tosun et al. Test 3 is included to see if there is a modality which does not assist in the classification.
Tests 4 and 5 were included to see which modality contains the most classification information, and tests 6 and 7 to see which ROI contains the most information. As a sanity check on the classification process, I created a dataset with all the subject variables kept together, but a random class (i.e. control, FTLD) allocated to that vector. Since there are two possible classes, this should return a correct classification 50% of the time. I ran this test 1000 times.

Table 8.1: ROIs significant at $p<0.05$, $p<0.01$, $p<0.001$ for all modalities which will be included in a classification scheme; JHU ROIs are thresholded at a probability level of 10%.

<table>
<thead>
<tr>
<th>Modality</th>
<th>ROIs significant at $p&lt;0.05$</th>
<th>ROIs significant at $p&lt;0.01$</th>
<th>ROIs significant at $p&lt;0.001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIFOF, RIFOF, LILF, RILF, RSLE, LUF, RUF</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, PostVMPFC, LIFOF, RIFOF, LIF, RILF, LUE, RUF</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIFOF, RIFOF, LIF, RILF, LUE, RUF</td>
</tr>
<tr>
<td>WM</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIF OF, RILF</td>
<td>LAntTL, RAntTL, LIF, RILF</td>
<td></td>
</tr>
<tr>
<td>CBF</td>
<td>LAntTL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAT</td>
<td>LAntTL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIFOF, RIFOF, LIF, RILF, LSLF, LUF, RUF</td>
<td>LAntTL, RAntTL, LPostTL, AntVMPFC, LIFOF, RIFOF, LIF, RILF, LUE, RUF</td>
<td>LAntTL, RAntTL, LPostTL,</td>
</tr>
<tr>
<td>FA</td>
<td>LAntTL, RCCg, LIFOF, LILF, LSLF, LUF</td>
<td>LIFOF</td>
<td></td>
</tr>
</tbody>
</table>

8.3 Results

There are 22 subjects for whom there is data from all modalities: 8 patients and 14 controls.
8.3.1 Using JHU ROIs thresholded at 20%

Using a random classification 1000 times gives the mean number classified correctly as 10.9±3.1. Assuming that only results greater than 2 standard deviations from the mean are likely non-random, this means that only results where more than 16 subjects are classified correctly are likely to be significant, and I have highlighted these in the tables that follow.

8.3.1.1 Classification with all variables

Using all 84 variables classified 21/22 subjects correctly. This result is unlikely to be the result of chance. The subject mis-classified was FTD-7.

8.3.1.2 Using Matlab “pca” (principal component analysis)

Using pca reduced the number of variables to 21. Classifying on the reduced data set classified 16/22 subjects correctly. This is marginally likely to be the result of chance. The subjects misclassified were:
FTD-7 FTD-9 FTD-14 FTD-16 FTD-19 FTD-21

8.3.1.3 Classification using only ROI variables seen to be significant in the univariate analysis

- at p<0.05 (42 variables) classified 22/22 subjects correctly - probably not due to chance.

- at p<0.01 (26 variables) classified 16/22 subjects correctly - marginal. The subjects classified incorrectly were:
  FTD-7 FTD-9 FTD-19 H3-17 H3-23 H3-24

- at p<0.001 (14 variables) classified 17/22 subjects correctly - probably not due to chance. The subjects classified incorrectly were:
  FTD-7 FTD-14 H3-2 H3-5 H3-24
8.3.1.4 Classification with varying modalities

These results are shown in table 8.2. The first column shows the modality or pair of modalities in question; column two shows the number of subjects classified correctly if that modality is excluded and those subjects classified incorrectly are itemised in column 3. Cells where excluding this modality makes a significant difference are highlighted. Column 4 shows the number of subjects classified correctly if only that modality is included and those subjects classified incorrectly are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted. It can be seen that the diffusion modalities FA and MD stand out in this table; the classification is less accurate when they are excluded, and more accurate when they are included. However the combination of all modalities is better than either only FA, only MD or the two together.

8.3.1.5 Classification with varying ROIs

The results of classifying the subjects both with and without specific ROIs are table 8.3. Column 1 shows the ROI in question; column two shows the number of subjects classified correctly if that ROI is excluded and those subjects are itemised in column 3. Column 4 shows the number of subjects classified correctly if only that ROI is included and those subjects are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted. It can be seen that excluding a single ROI does not make a significant difference to the accuracy of classification. However, several ROIs are adequate classifiers on their own: these are all the aal ROIs (i.e. right and left anterior temporal lobe, right and left posterior temporal lobe, anterior and posterior ventromedial prefrontal cortex), the right and left part of the cingulum bundle connected to the cingulate gyrus, the right and left part of the cingulum bundle connected to the cingulum hippocampus, the right and left inferior longitudinal fasciculus, the right and left superior frontooccipital fasciculus and the right and left uncinate fasciculus.
Table 8.2: Excluding/Including each modality in turn with JHU ROIs thresholded at 20%. The first column shows the modality or pair of modalities in question; column two shows the number of subjects classified correctly if that modality is excluded and those subjects classified incorrectly are itemised in column 3. Cells where excluding this modality makes a significant difference are highlighted. Column 4 shows the number of subjects classified correctly if that modality is included and those subjects classified incorrectly are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted.

<table>
<thead>
<tr>
<th>Modality</th>
<th>#correct if excluded</th>
<th>IDs</th>
<th>#correct if only</th>
<th>IDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>22/22</td>
<td></td>
<td>14/22</td>
<td>FTD-7 FTD-13 FTD-16 H3-4 H3-5 H3-8 H3-24 H3-25</td>
</tr>
<tr>
<td>WM</td>
<td>21/22</td>
<td>FTD-7</td>
<td>13/22</td>
<td>FTD-16 FTD-19 H3-2 H3-3 H3-4 H3-5 H3-17 H3-21 H3-22</td>
</tr>
<tr>
<td>GM &amp; WM</td>
<td>20/22</td>
<td>FTD-7 FTD-13</td>
<td>13/22</td>
<td>FTD-7 FTD-14 FTD-16 H3-4 H3-5 H3-8 H3-19 H3-24 H3-25</td>
</tr>
<tr>
<td>CBF</td>
<td>21/22</td>
<td>FTD-7</td>
<td>11/22</td>
<td>FTD-7 FTD-9 FTD-13 FTD-16 FTD-19 FTD-21 FTD-24 H3-2 H3-21 H3-25 H3-28</td>
</tr>
<tr>
<td>tA</td>
<td>21/22</td>
<td>FTD-7</td>
<td>13/22</td>
<td>FTD-7 FTD-13 FTD-14 FTD-16 H3-3 H3-8 H3-21 H3-22 H3-28</td>
</tr>
<tr>
<td>CBF &amp; tA</td>
<td>21/22</td>
<td>FTD-7</td>
<td>14/22</td>
<td>FTD-9 FTD-13 FTD-14 FTD-16 FTD-19 H3-3 H3-21 H3-25</td>
</tr>
<tr>
<td>MD</td>
<td>19/22</td>
<td>FTD-7 FTD-16 H3-21</td>
<td>18/22</td>
<td>FTD-7 FTD-9 FTD-14 H3-5</td>
</tr>
<tr>
<td>FA</td>
<td>16/22</td>
<td>FTD-7, FTD-9, FTD-14,F TD-16, H3-8, H3-21</td>
<td>15/22</td>
<td>FTD-7 FTD-9 FTD-24 H3-4 H3-21 H3-22 H3-28</td>
</tr>
<tr>
<td>MD &amp; FA</td>
<td>15/22</td>
<td>FTD-7 FTD-14 FTD-16 H3-4 H3-8 H3-19 H3-21</td>
<td>16/22</td>
<td>FTD-7 FTD-13 FTD-24 H3-2 H3-21 H3-28</td>
</tr>
</tbody>
</table>
Excluding/Including each ROI in turn with JHU ROIs thresholded at 20%. The first column shows the ROI in question; column two shows the number of subjects classified correctly if that ROI is excluded and those subjects classified incorrectly are itemised in column 3. Column 4 shows the number of subjects classified correctly if that ROI is included and those subjects classified incorrectly are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted.

<table>
<thead>
<tr>
<th>Region</th>
<th>#correct if excluded</th>
<th>IDs</th>
<th>#correct if only</th>
<th>IDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPostTL</td>
<td>19/22</td>
<td>FTD-7 FTD-16 H3-21</td>
<td>20/22</td>
<td>FTD-7 H3-4</td>
</tr>
<tr>
<td>RPostTL</td>
<td>21/22</td>
<td>FTD-7</td>
<td>18/22</td>
<td>FTD-7 FTD-19 H3-2 H3-4</td>
</tr>
<tr>
<td>LAntTL</td>
<td>21/22</td>
<td>FTD-7</td>
<td>18/22</td>
<td>FTD-7 FTD-16 FTD-19 H3-22</td>
</tr>
<tr>
<td>RAntTL</td>
<td>21/22</td>
<td>FTD-7</td>
<td>18/22</td>
<td>FTD-7 FTD-16 FTD-19 H3-5</td>
</tr>
<tr>
<td>AVMPTC</td>
<td>21/22</td>
<td>FTD-7</td>
<td>18/22</td>
<td>FTD-7 H3-4</td>
</tr>
<tr>
<td>PVMPFC</td>
<td>21/22</td>
<td>FTD-7</td>
<td>16/22</td>
<td>FTD-7 FTD-14 FTD-16 H3-3 H3-4 H3-21</td>
</tr>
<tr>
<td>CCg-L</td>
<td>20/22</td>
<td>FTD-7 FTD-16</td>
<td>17/22</td>
<td>FTD-13 FTD-21 H3-2 H3-21 H3-24</td>
</tr>
<tr>
<td>CCg-R</td>
<td>20/22</td>
<td>FTD-7 FTD-16</td>
<td>17/22</td>
<td>FTD-19 FTD-21 FTD-24 H3-2 H3-21</td>
</tr>
<tr>
<td>CH-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>15/22</td>
<td>FTD-7 FTD-9 FTD-16 FTD-19 FTD-21 H3-17 H3-25</td>
</tr>
<tr>
<td>CH-R</td>
<td>21/22</td>
<td>FTD-7</td>
<td>9/22</td>
<td>FTD-7 FTD-9 FTD-13 FTD-14 FTD-16 FTD-19 FTD-21 FTD-24 H3-2 H3-3 H3-4 H3-25 H3-28</td>
</tr>
<tr>
<td>IFOF-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>19/22</td>
<td>FTD-7 FTD-16 H3-22</td>
</tr>
<tr>
<td>IFOF-R</td>
<td>20/22</td>
<td>FTD-7 FTD-16</td>
<td>20/22</td>
<td>FTD-14 H3-24</td>
</tr>
<tr>
<td>ILF-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>17/22</td>
<td>FTD-7 FTD-16 FTD-21 H3-5 H3-17</td>
</tr>
<tr>
<td>ILF-R</td>
<td>21/22</td>
<td>FTD-7</td>
<td>16/22</td>
<td>FTD-13 FTD-14 FTD-16 H3-2 H3-5 H3-8</td>
</tr>
<tr>
<td>SLF-L</td>
<td>19/22</td>
<td>FTD-7 FTD-16 H3-21</td>
<td>12/11</td>
<td>FTD-7 FTD-9 FTD-16 FTD-19 FTD-21 FTD-24 H3-2 H3-5 H3-17 H3-19</td>
</tr>
<tr>
<td>SLF-R</td>
<td>21/22</td>
<td>FTD-7</td>
<td>13/22</td>
<td>FTD-7 FTD-13 FTD-14 FTD-21 H3-5 H3-8 H3-17 H3-22 H3-23</td>
</tr>
<tr>
<td>UF-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>17/22</td>
<td>FTD-7 FTD-16 FTD-19 H3-5 H3-21</td>
</tr>
<tr>
<td>UF-R</td>
<td>21/22</td>
<td>FTD-7</td>
<td>18/22</td>
<td>FTD-7 FTD-16 H3-8 H3-21</td>
</tr>
</tbody>
</table>
8.3.2 JHU ROIs thresholded at 10%

Using a random classification 1000 times gives the mean number classified correctly as $11.2\pm3.0$; very similar to the result with the JHU ROIs thresholded at 20% probability. Assuming that only results greater than 2 standard deviations from the mean are likely non-random, this means that only results where more than 6 subjects are classified correctly are likely to be significant (i.e. fewer than 6 incorrectly), and I have highlighted these in the tables that follow.

8.3.3 Classification with all variables

The results are identical with those reported in chapter 8.3.1. Using all 84 variables classified 20/22 subjects correctly. This result is unlikely to be the result of chance. The subjects mis-classified were FTD-7 H3-21.

8.3.3.1 Using Matlab “pca” (principal component analysis)

Using pca reduced the number of variables to 21. Classifying on the reduced data set classified 16/22 subjects correctly. This is marginally likely to be the result of chance. The subjects misclassified were:
FTD-7 FTD-9 FTD-14 FTD-16 FTD-19 FTD-21

8.3.3.2 Classification using only ROI variables seen to be significant in the univariate analysis

- at $p<0.05$ (42 variables) classified 22/22 subjects correctly - probably not due to chance.
- at $p<0.01$ (26 variables) classified 16/22 subjects correctly - marginal. The subjects classified incorrectly were:
  FTD-7 FTD-9 FTD-13 FTD-24 H3-2 H3-4
- at $p<0.001$ (14 variables) classified 17/22 subjects correctly - probably not due to chance. The subjects classified incorrectly were:
  FTD-7 FTD-14 FTD-21 H3-2 H3-5
8.3.3.3 Classification with varying modalities

These results are shown in table 8.2. The first column shows the modality or pair of modalities in question; column two shows the number of subjects classified correctly if that modality is excluded and those subjects classified incorrectly are itemised in column 3. Cells where excluding this modality makes a significant difference are highlighted. Column 4 shows the number of subjects classified correctly if only that modality is included and those subjects classified incorrectly are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted. It can be seen that the diffusion modalities FA and MD stand out in this table; the classification is less accurate when they are excluded, and more accurate when they are included. However the combination of all modalities is better than either only FA, only MD or the two together.

8.3.3.4 Classification with varying ROIs

These results are shown in table 8.3. The first column shows the ROI in question; column two shows the number of subjects classified correctly if that ROI is excluded and those subjects are itemised in column 3. Column 4 shows the number of subjects classified correctly if only that ROI is included and those subjects are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted. It can be seen that excluding a single ROI does not make a significant difference to the accuracy of classification. However, several ROIs are adequate classifiers on their own: these are all the aal ROIs, L and R CCg, L and R ILF, L and R IFOF and L and R UF.

The most accurate classification comes from including only the variables which are significant at p<0.05 in a univariate test. The modalities/ROIs are shown in table 8.6 and visually in figure 8.1.
Figure 8.1: ROIs for a given modality which give the most accurate classification. The specific ROIs are given in table 8.6.
Table 8.4: Excluding/Including each modality in turn with JHU ROIs thresholded at 10% probability. The first column shows the modality or pair of modalities in question; column two shows the number of subjects classified correctly if that modality is excluded and those subjects classified incorrectly are itemised in column 3. Cells where excluding this modality makes a significant difference are highlighted. Column 4 shows the number of subjects classified correctly if that modality is included and those subjects classified incorrectly are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted.

<table>
<thead>
<tr>
<th>Modality</th>
<th>#correct if excluded</th>
<th>IDs</th>
<th>#correct if only</th>
<th>IDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>WM</td>
<td>21/22</td>
<td>FTD-7</td>
<td>14/22</td>
<td>FTD-7 FTD-9 FTD-14 FTD-16 FTD-19 H3-4 H3-23 H3-26</td>
</tr>
<tr>
<td>GM &amp; WM</td>
<td>21/22</td>
<td>H3-21</td>
<td>14/22</td>
<td>FTD-7 FTD-13 FTD-16 H3-2 H3-4 H3-5 H3-8 H3-21</td>
</tr>
<tr>
<td>CBF</td>
<td>21/22</td>
<td>FTD-7</td>
<td>11/22</td>
<td>FTD-7 FTD-9 FTD-13 FTD-16 FTD-19 FTD-21 FTD-24 H3-2 H3-21 H3-25 H3-28</td>
</tr>
<tr>
<td>tA</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>13/22</td>
<td>FTD-7 FTD-13 FTD-14 FTD-16 H3-3 H3-8 H3-21 H3-22 H3-28</td>
</tr>
<tr>
<td>CBF &amp; tA</td>
<td>22/22</td>
<td></td>
<td>14/22</td>
<td>FTD-9 FTD-13 FTD-14 FTD-16 FTD-19 H3-3 H3-21 H3-25</td>
</tr>
<tr>
<td>MD</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>12/22</td>
<td>FTD-7 FTD-9 FTD-13 FTD-14 FTD-16 H3-2 H3-5 H3-8 H3-17 H3-25</td>
</tr>
<tr>
<td>FA</td>
<td>19/22</td>
<td>FTD-7 FTD-9 H3-21</td>
<td>19/22</td>
<td>FTD-7 FTD-9 FTD-21</td>
</tr>
<tr>
<td>MD &amp; FA</td>
<td>16/22</td>
<td>FTD-7 FTD-14 FTD-16 H3-4 H3-8 H3-21</td>
<td>19/22</td>
<td>FTD-7 FTD-24 H3-28</td>
</tr>
</tbody>
</table>
Table 8.5: Excluding/Including each ROI in turn with JHU ROIs thresholded at 10% probability. The first column shows the ROI in question; column two shows the number of subjects classified correctly if that ROI is excluded and those subjects classified incorrectly are itemised in column 3. Column 4 shows the number of subjects classified correctly if that ROI is included and those subjects classified incorrectly are itemised in column 6. Cells where 16 or more subjects are classified correctly are highlighted.

<table>
<thead>
<tr>
<th>Region</th>
<th>#correct if excluded</th>
<th>IDs</th>
<th>#correct if only</th>
<th>IDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPostTL</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>20/22</td>
<td>FTD-7 H3-4</td>
</tr>
<tr>
<td>RPostTL</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>18/22</td>
<td>FTD-7 FTD-19 H3-2 H3-4</td>
</tr>
<tr>
<td>LAntTL</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>18/22</td>
<td>FTD-7 FTD-16 FTD-19 H3-22</td>
</tr>
<tr>
<td>RAntTL</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>18/22</td>
<td>FTD-7 FTD-16 FTD-19 H3-5</td>
</tr>
<tr>
<td>AVMPFC</td>
<td>21/22</td>
<td>FTD-7</td>
<td>20/22</td>
<td>FTD-7 H3-4</td>
</tr>
<tr>
<td>PVMPFC</td>
<td>21/22</td>
<td>FTD-7</td>
<td>16/22</td>
<td>FTD-7 FTD-14 FTD-16 H3-3 H3-4 H3-21</td>
</tr>
<tr>
<td>CCg-L</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>16/22</td>
<td>FTD-7 FTD-13 FTD-19 FTD-21 H3-2 H3-21</td>
</tr>
<tr>
<td>CCg-R</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>17/22</td>
<td>FTD-19 FTD-21 FTD-24 H3-2 H3-21</td>
</tr>
<tr>
<td>CH-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>16/22</td>
<td>FTD-7 FTD-16 FTD-19 FTD-21 H3-17 H3-25</td>
</tr>
<tr>
<td>CH-R</td>
<td>2</td>
<td>FTD-7 H3-21</td>
<td>15/22</td>
<td>FTD-7 FTD-14 FTD-16 FTD-19 FTD-21 H3-2 H3-25</td>
</tr>
<tr>
<td>IFOF-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>20/22</td>
<td>FTD-7 H3-5</td>
</tr>
<tr>
<td>IFOF-R</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>19/22</td>
<td>FTD-7 FTD-14 H3-22</td>
</tr>
<tr>
<td>ILF-L</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>17/22</td>
<td>FTD-7 FTD-16 FTD-21 H3-5 H3-17</td>
</tr>
<tr>
<td>ILF-R</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>17/22</td>
<td>FTD-7 FTD-14 FTD-16 H3-5 H3-23</td>
</tr>
<tr>
<td>SLF-L</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>14/22</td>
<td>FTD-7 FTD-16 FTD-21 FTD-24 H3-2 H3-5 H3-17 H3-19</td>
</tr>
<tr>
<td>SLF-R</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>13/22</td>
<td>FTD-9 FTD-13 FTD-14 FTD-16 FTD-21 FTD-24 H3-5 H3-8 H3-17</td>
</tr>
<tr>
<td>UF-L</td>
<td>21/22</td>
<td>FTD-7</td>
<td>18/22</td>
<td>FTD-7 FTD-16 FTD-19 H3-5</td>
</tr>
<tr>
<td>UF-R</td>
<td>20/22</td>
<td>FTD-7 H3-21</td>
<td>17/22</td>
<td>FTD-16 FTD-19 H3-2 H3-8 H3-21</td>
</tr>
</tbody>
</table>

Table 8.6: Modalities and ROIs included for most accurate classification

<table>
<thead>
<tr>
<th>Modality</th>
<th>ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBF &amp; tA</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIFOF, RIFOF, LIF, RIF, RSIF, LUF, RUF</td>
</tr>
<tr>
<td>GM</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIFOF, RIFOF, LIF, RIF, RSIF, LUF, RUF</td>
</tr>
<tr>
<td>WM</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIF, RIF</td>
</tr>
<tr>
<td>MD</td>
<td>LAntTL, RAntTL, LPostTL, RPostTL, AntVMPFC, PostVMPFC, LIFOF, RIFOF, LIF, RIF, LSIF, LUF, RUF</td>
</tr>
<tr>
<td>FA</td>
<td>LAntTL, RCCg, LIFOF, LIF, LSIF, LUF</td>
</tr>
</tbody>
</table>


8.4 Discussion

The number of subjects included in this analysis is very small, and so all results must be considered as tentative, and suggesting interesting areas for further research rather than being conclusive. Nevertheless it is clear that, when considered in isolation, only of the modalities, MD, is better at distinguishing patients from controls than chance, with possibly one more in FA. There is shown to be power in combining modalities. Some ROIs are good classifiers on their own, but the best classification comes from the combination of ROIs and modalities shown in table 8.6.

Data from the ASL images is as good as data from the T$_1$-weighted images. Given the low resolution and low SNR of the ASL images when compared with the T$_1$-weighted images, this is surprising. This study has provided no evidence that the tissue affected in FTLD has lower intrinsic CBF than unaffected tissue (see table 5.3). ASL images seem to be as good a measure of atrophy as T$_1$-weighted images.

It is surprising that the grey matter values in isolation are not particularly good at classifying subjects, even though the Wilcoxon ranked sum test results show that practically all the ROIs are significantly different between patients and controls (see table 5.3). Further investigation discovered that all the subjects who were mis-classified had at least one ROI where the subject value was more than 2 standard deviations from the group mean, which may be something of an explanation.

It is not surprising that excluding GM values makes little difference to the classification results, as the values included in the classification have no partial-volume correction applied. As discussed in chapter 7.4, the MD values included here seem to have a large component that is due to CSF and not brain tissue, and so information on atrophy is included in the MD values.

Although the results from an analysis with JHU ROIs thresholded at 10% probability are broadly similar to those with ROIs thresholded at 20% probability, there are interesting and suggestive differences in the details, particularly when one modality is excluded, or classification performed on only one modality. MD is much more predictive at 20% probability (10 wrong at 10%, 4 wrong at 20%), and
correspondingly FA is better at 10% (3 wrong at 10%, 7 wrong at 20%). This may indicate that the MD difference reflects some inherent change in the WM tissue at the centre of the included tracts, or just be an indication of the inaccuracy of the tract definition. However, FTLD is increasingly being recognised as a disease that affects brain networks [Zhou and Seeley, 2014], and therefore it is logical that white matter should be affected.

It is also interesting that 3 regions in particular seem very predictive of classification; these are the left posterior temporal lobe, the anterior ventromedial prefrontal cortex, and the inferior fronto-occipital fasciculus. These illnesses are known to cause atrophy in the frontal and anterior temporal lobes, so it is surprising that the posterior left temporal lobe. Other studies have found WM damage in the fronto-occipital fasciculus ([Borroni et al., 2007], [Tovar-Moll et al., 2014]). It is surprising that the left posterior temporal lobe is predictive, but the IFOF has projections to both the frontal lobe and posterior temporal lobe [Wakana et al., 2004], so it is possible that there is secondary damage to this region in a manner similar to the effect of Wallerian degeneration.

In all the classification possibilities considered here, one subject in particular has proved difficult to classify: FTD-7. A discussion with the clinician (RZ) who diagnosed and recruited this patient made clear that he was dubious about the diagnosis of FTLD, as this subject had clear behavioural symptoms, but nevertheless well-preserved activities of daily living, and that while he showed symptoms of semantic dementia, some of the semantic symptoms he displayed could also fit the criteria for the logopenic variant of primary progressive aphasia, most often associated with Alzheimer’s disease on post-mortem examination. Another subject, H3-21, is also frequently mis-classified. He is one of the oldest controls recruited, and showed MRI symptoms of cerebrovascular disease which were considered normal for age, but worthy of reporting back to the subject’s GP. In summary, this regional analysis with multimodal imaging data provides a good way of distinguishing between patients and controls in this group.
Chapter 9

Conclusion and further work

9.1 Summary of findings

I have looked at a range of multimodal MR images of patients with FTLD, and healthy age-matched controls, and considered which provide the best data for classifying subjects. The three modalities are

- $T_1$-weighted images, which give information on grey matter (GM) and white matter (WM) densities.
- ASL images, which give information on cerebral blood flow (CBF) and arterial arrival time (AAT).
- diffusion-weighted images, where I have chosen to look at the scalar diffusion metrics of mean diffusivity (MD) and fractional anisotropy (FA).

I have analysed these data using both voxel-based statistical analyses (VBM in SPM8, TBSS in FSL), and with an analysis of regions of interest (ROI).

I have used the VBM analyses to perform group comparisons between patients and controls: for the $T_1$-weighted images this is shown in chapter 5.3.3; for the ASL images in chapters 6.3.1 and 6.3.3; for the diffusion-weighted images in chapters 7.3.1 and 7.3.2.1.
I have used the values from the ROIs for each modality in a classification to distinguish patients from controls. Taken individually, the diffusion-weighted images provide the best means of distinguishing between the groups (see tables 8.2, 8.4), classifying 21/22 subjects correctly. Other modalities are no better than chance. The combination of all modalities proves the best (see chapter 8.4).

I have also considered which ROIs for which modalities are most informative, looking at all modalities in regions from the AAL atlas, and T1-weighted and diffusion-weighted modalities in regions from the JHU white matter atlas. The ROIs I have considered are described in chapter 8.4, and were chosen based on prior knowledge of the ROIs most likely to be affected in this group of illnesses. As the JHU atlas I have used gives probabilities that a given voxel is in a given tract, I have chosen to use two sets of JHU ROIs, one where I have included a voxel within the ROI if the probability is > 10%, and another group of ROIs where the probability is > 20%. The ROIs thresholded at a probability of 10% are of most utility when distinguishing between patients and controls; the ROIs thresholded at a probability of 20% may highlight differences in the underlying tissue.

GM and WM densities are reduced in several ROIs in patients; this data is shown in tables 5.3, 5.5 and 5.6 for GM and 5.4, 5.7 and 5.8 for WM. This difference is found even when the analyses are restricted to voxels with a high percentage of GM or WM (> 80%), it may therefore indicate some intrinsic difference in the tissue, and not be just a reflection of increased atrophy in the patient group. In these images, the signal from CSF is lower than that from GM or WM.

Both ASL images and diffusion-weighted images have relatively large voxels, and thus the actual values of parameters measured within each voxel will depend on the proportion of GM, WM and cerebrospinal fluid (CSF) within the voxel, as well as the value of the parameter in the different tissue types. To determine the value of the parameter for GM it is necessary to carry out some correction for partial volume effects. I have considered the differences between patients and controls for both the uncorrected metrics and those corrected for partial volume effects. Uncorrected values and values corrected for partial volume effects are given in table 6.6 for CBF, in table 6.7 for AAT, in tables 7.4, 7.6 and 7.8 for FA and in tables 7.3, 7.5 and 7.7 for MD.

CBF is decreased in patients in the left anterior temporal lobe, and in an ROI defined by the volume affected by GM atrophy. However, this difference is not
seen once the values are corrected for partial volume effects. This confirms findings by Tosun et al [Tosun et al., 2012] and Woost et al [Woost et al., 2013] that atrophy rather than hypometabolism is the dominant feature of FTLD. AAT is increased in patients in the left calcarine and right insula; this finding survives correction for both multiple comparisons and partial volume effects. These are regions unaffected in the initial stages of the illness; perhaps it reflects a change in regions peripheral to those mostly affected. Such a change could be precipitated by an absence of neuropeptides from the affected regions in a process akin to Wallerian degeneration. Increased AAT has been found in multiple sclerosis and in Parkinson’s disease, as discussed in chapter 6.4.3. Of equal novelty is a decreased AAT in the very atrophied region of the left temporal lobe; perhaps it indicates a reduction in the tortuosity of the microvasculature or a change in the arterial path length.

MD is decreased in patients in all of the aal ROIs known to be affected by the illness, and in the majority of the JHU ROIs thresholded at a probability level of 10%; these differences survive correction for partial volume effects. This is probably a sensitive measure of atrophy, but as the differences are also found in JHU ROIs thresholded at a probability level of 20% it may also indicate a change in the tissue microstructure as well. FA is decreased in patients, although the differences are less marked than those for MD.

The ROIs and modalities of most use when distinguishing between patients and controls are described in table 8.6 and shown in figure 8.1. The left anterior lobe (LAntTL) is the most frequently occurring region, since it is the only region where all the modalities contribute to the classification. It is also the region that is most frequently described as being affected in the disease (see chapter 2.6.2). GM density, WM density and MD contribute in all the aal regions considered here (left and right anterior temporal lobe, left and right posterior temporal lobe, anterior and posterior ventromedial prefrontal cortex). GM density in the JHU regions of the left and right inferior fronto-occipital fasciculus, the left and right inferior longitudinal fasciculus, the left and right uncinate fasciculus and the right superior longitudinal fasciculus contribute to the classification. WM density in JHU regions is only used in the right and left inferior longitudinal fasciculus. MD in the JHU regions of the left and right inferior fronto-occipital fasciculus, the left and right inferior longitudinal fasciculus, the left and right uncinate fasciculus and the left superior longitudinal fasciculus contribute to the classification (note that this is
very similar to the contribution from GM density). FA values from the left anterior lobe, the right part of the cingulum bundle connected to the cingulate, the left inferior longitudinal fasciculus, the left superior and inferior longitudinal fasciculus and the left uncinate fasciculus are lower in patients.

I have confirmed the correlations found by other studies in the regional pattern of grey matter (GM) density in the brains of patients showing symptoms of FTLD, and the link of specific cognitive deficits with specific brain regions (see chapter 5.3.4). These correlations are also found in MD images (see chapter 7.3.2.2, 7.3.2.3 and 7.3.2.4), although not in CBF or AAT. This may be due to the large inter-individual differences in CBF, despite the normalisation to the individual global CBF, which may make a comparatively small regional difference difficult to see.

In this thesis I have carried out extensive simulations on the single blood compartment model proposed by Parkes and Tofts [Parkes and Tofts, 2002] and a multi-timepoint series of images. This model is used to quantify both CBF and AAT. I have explored the parameters used in the model, and the effect on the quantification of CBF and AAT if their values are not accurately known. I have explored the effect of an inaccurate AAT on the estimated value of CBF, and shown that the estimated values of CBF can differ by 30% for different assumed values of AAT if only CBF is fitted (see figure 3.3). I have shown that a 2-parameter fit for CBF and AAT is the most that can be achieved with the ASL signal in a timescale realistic for a clinical scan, and that 8 timepoints provide adequate data for such a fit; there is no value in collecting more timepoints (see chapter 3.5).

I have looked at two differing MRI read-out sequences for use in the sometimes problematic brain regions of the ventral frontal and temporal lobes, and concluded that a spin-echo sequence suffers from less dropout than a gradient-echo sequence. I have also investigated the best way to apply the scaling factor of proton density needed to quantify the CBF, and the advantages and disadvantages of applying this scaling factor either voxelwise or globally (see chapter 4.4).
9.2 Discussion of findings

9.2.1 FTLD patients in this study

The majority of the FTLD patients in this study had a clear diagnosis of FTLD, and had been ill for some time. This is clear from the scores in the neuropsychological tests, where practically all the tests provide a good means of distinguishing patients from controls. One may wonder why MRI is worth bothering with when the neuropsychological tests are so good. There are two answers to that question. The first is that a means is needed of tracking the progression of the illness, and this will become more important when drugs are available to treat the conditions. The second is that, while in the later stages of the illness it is not too difficult to diagnose, in the early stages it can prove very difficult indeed. While in the later stages neuropsychological tests can highlight severe cognitive deficits, this is not so true when the deficits are not so severe. Kipps et al also described a substantial subset of patients with apparent bvFTD according to clinical evaluation and neuropsychological tests, but who show no imaging abnormality and no sign of a neurodegenerative process [Kipps et al., 2009]. A method is needed of distinguishing these patients.

MRI images can be acquired as frequently as the patient will tolerate. This is not true of any other diagnostic test. PET imaging involves a radiation dose, and so the scans cannot be repeated regularly. Neuropsychological tests are not infinitely repeatable. Currently there are no chemical biomarkers in CSF that are diagnostic. MRI has the potential to provide an indication of the spread of the neurodegenerative process. MRI could also be of more use earlier in the diagnostic process, if a clear pattern of MRI differences emerges. This is discussed further in chapter 9.3.3.

9.2.2 Partial volume corrections

When considering whether or not to apply a partial volume correction, it is important to consider what information is wanted from the images. Partial volume correction can be applied to ASL images, as the CBF of GM is markedly different from that of WM, and CSF has no CBF. Partial volume correction can also be con-
sidered for diffusion metrics, where the problem is reversed in that CSF has a much
greater MD, and lower FA, then the surrounding tissue. When comparing groups
such as the FTLD patients and healthy controls studied here, there is a consistent
anatomical difference between the groups, in that patients have a greater degree of
atrophy than controls. Hence if one is looking for a difference in the tissue proper-
ties between the groups, a partial volume correction is essential. However, if one
is interested in classification the partial volume correction is a hindrance, since
it is attempting to correct for group differences. MD (for example), uncorrected
for partial volume effects, is indeed significantly different between FTLD patients
and controls. That difference may be due to alterations in the tissue microstruc-
ture in patients, or it may be due to atrophy. As a means of differentiating between
patients and controls, it does not matter what causes the difference, simply that it
exists. Here I have quoted results both with and without partial volume correction.

There is also a difference in the method of calculating an appropriate correction
for ASL images, and for diffusion-weighted images. I have followed Asllani
et al [Asllani et al., 2008] in assuming a fixed ratio between the CBF of GM and
the CBF of WM, and then used an arithmetical correction, taking GM and WM
densities from the segmentation maps of the $T_1$-weighted images. This is possible
because the CBF of CSF is zero, and hence the CBF of any voxel is assumed to
depend only on the proportions of GM and WM. This method cannot be applied
to MD because the MD of CSF is not zero, but very much greater than the MD of
either GM or WM. There is also no clear relationship between the MD or FA of
GM, WM or CSF. The contribution to the MD of a given voxel thus depends on
the proportions of GM, WM and CSF, and also on the distribution of those tissues
within the voxel. While the FA of CSF is zero, the FA of the voxel overall will
depend very much on the distribution of the various tissues, and not just on their
proportions. Hence I have chosen to try to mask out the CSF by looking only at
voxels with a small proportion of CSF.

### 9.2.3 Voxel-based analysis vs regional analysis

I have shown that voxel-based morphometry is problematic when there are step
changes in values across the brain; for example, from the MD of tissue in one
voxel to a value of a zero in an adjacent voxel, if the image has been masked to
exclude CSF. This is explored in chapter 7.3.2.1. The smoothing necessary for the
technique to work introduces biases, particularly in the temporal lobes because of their location at the edge of the brain. This can be a particular problem if partial volume correction or masking to exclude CSF introduce substantial regions of the image where the value is zero. The majority of the studies I have referenced here have used VBM, and many mention that it is a preferred analysis method as it is not dependent on prior knowledge (e.g. [Bruno et al., 2012]). VBM has been shown to be a sensitive and robust tool when studying FTLD, justifying its inclusion in a multi parametric assessment of the condition. However, it needs to be used with caution when analysing images with large regions of zero values.

It is over 20 years since $T_1$-weighted MRI was first used to study FTLD; a meta-analysis of images from these studies should be able to provide relevant and accurate a-priori regions. Regions can be drawn by hand, defined semi-automatically or defined automatically. The first option is very time-consuming and also subjective and difficult to transfer from one site to another; these limitations apply in a lesser extent to the second option. The third option has the disadvantage that the ROI may be obviously misplaced: this occurs because of incorrect normalisation, and will also occur in a VBM analysis. However, in the VBM analysis the problem will not be so obvious. ROI analysis has the following advantages:

- ROI analysis can use prior knowledge, e.g. that the regions most affected in FTLD are frontal and temporal lobes.
- ROI analysis effectively smooths values over the volume of the region, which may be a more anatomically relevant way of averaging data than the 12mm gaussian kernel used for smoothing here and in other studies.
- ROI analysis gives regional values, whereas VBM only shows where values are increased or decreased, with no absolute values.
- ROI analysis involves fewer comparisons, and hence less stringent multiple comparison corrections.
9.3 Further work

9.3.1 Further analysis; more data

This study is a small study of a heterogeneous group of patients. Nonetheless it has determined the location and sensitivity of T1-weighted, ASL and diffusion-weighted MR imaging changes that accompany FTLD. There is by now a multiplicity of other studies of patients with FTLD, many involving T1-weighted images, and some with ASL images and/or diffusion-weighted images. An organisation now exists, GENFI, which is gathering these data and making them available to researchers in the field. It would be very interesting to assemble images from a larger group of patients, ideally with neuropsychological test scores, and perform a VBM regression of the images against the results of the cognitive tests, to see if it is possible to define more finely regions which are consistently related to specific cognitive tests, and of the expected variation of these values. Such data could be used to track disease progression, and also perhaps to improve current diagnostic tests (see chapter 9.3.3).

Although practically all of the studies considered here have classified patients by diagnosis (the exceptions being studies such as that by McMillan et al [McMillan et al., 2013] which classify by supposed pathology), it may not be a particularly meaningful classifier. This is partly because the diagnosis gives no indication of the level of severity of the illness, but also because the majority of patients display mixed symptoms: patients with SD or PPA are likely eventually to display behavioural symptoms [Rascovsky et al., 2011], etc. It is possible that the scores in a few key cognitive tests may be better classifiers; this needs to be investigated.

9.3.2 MRI sequences to test for atrophy

This study has shown that diffusion metrics, particularly MD, are very sensitive to some difference in tissue between FTLD patients and controls. It is my expectation that this is due to atrophy, and that MD is highly sensitive to the greater proportion of CSF in atrophied tissue. However, it is possible that there is some other tissue property which is causing this difference, and it would be interesting to investigate
further. If indeed the difference seen in MD is due to the greater proportion of CSF, there may be other MRI sequences which are even more sensitive to the presence of CSF (e.g. FLAIR). It would be interesting to investigate these.

### 9.3.3 Development of a diagnostic support tool

At the moment MRI images are described either qualitatively or with a visual rating scale that only looks at a few regions. The scores of neuropsychological tests are used independently. If it is possible to create a more accurate map of which brain regions are linked to which cognitive tests, it would be possible to create an analysis tool which could compare brain atrophy with cognitive test scores. A below-normal cognitive test score would surely be of more significance if combined with a marginally greater than normal level of atrophy in the associated brain region. Combined with an imaging technique that is more sensitive to atrophy, such an analysis tool could be tested in a multi-centre setting to see if it could give a useful prediction of the likely diagnosis. A less ambitious diagnostic support tool could compare the MR images from individual patients against an image derived from a large study of normal age-matched controls, and provide a map showing where differences existed.

### 9.4 Summary

I have looked at ways of differentiating FTLD patients from controls. Many other studies have looked at ways of generating a differential diagnosis of FTLD vs AD, and of differing FTLD subtypes. As drugs become available to treat these illnesses, such a distinction will become more important. Drugs cannot easily be tested without a biomarker to track disease progression, or the lack of it. However, I believe that a differential diagnosis between dementia and psychiatric illness is more important, even in the absence of a medical treatment, especially for the patients and their families and given the importance in social and financial considerations of such a diagnosis, or the lack of it. “We don’t know, come back next year” gives time for the patient to get divorced, be made redundant, lose their house, be arrested for bizarre behaviour....
The late Sir Terry Pratchett said of Alzheimer’s disease: *It is a physical disease, not a mystic curse; therefore it will fall to a physical cure.* Surely the same is true of the frontotemporal dementias.
Appendix A

A Diagnostic criteria for bvFTD, SD and PPA

A.1 bvFTD

[Neary et al., 1998] These criteria have been updated by Rascovsky et al [Rascov-
sky et al., 2011]; I have given the old criteria as the patients in this study were
diagnosed with these criteria, and the majority of the papers referenced here will
have been of patients diagnosed with these criteria.
Character change and disordered social conduct are the dominant features initially
and throughout the disease course. Instrumental functions of perception, spatial
skills, praxis and memory are intact or relatively well preserved.

1. Core diagnostic features
   (a) Insidious onset and gradual progression
   (b) Early decline in social interpersonal conduct
   (c) Early impairment in regulation of personal conduct
   (d) Early emotional blunting
   (e) Early loss of insight

2. Supportive diagnostic features
   (a) Behavioural disorder
      i. Decline in personal hygiene and grooming
      ii. Mental rigidity and inflexibility
      iii. Distractability and impersistence
      iv. Hyperorality and dietary changes
      v. Perseverative and stereotyped behaviour
vi. Utilisation behaviour
(b) Speech and language
   i. Altered speech output
      A. Aspontaneity and economy of speech
      B. Press of speech
      C. Stereotypy of speech
      D. Echolalia
      E. Perseveration
      F. Mutism
(c) Physical signs
   i. Primitive reflexes
   ii. Incontinence
   iii. Akinesia, rigidity and tremor
   iv. Low and labile blood pressure
(d) Investigations
   i. Neuropsychology: significant impairment on frontal lobe tests in the absence of severe amnesia, aphasia or perceptual disorder
   ii. Electroencephalography: normal on conventional EEG despite clinically evident dementia
   iii. Brain imaging (structural and/or functional: predominant frontal and/or temporal abnormality

A.2 Semantic dementia

[Neary et al., 1998] The clinical diagnostic features of semantic aphasia and associative agnosia (SD): Clinical profile
Semantic disorder (impaired understanding of word meaning and/or object identity) is the dominant feature initially and throughout the disease course. Other aspects of cognition, including autobiographic memory, are intact or relatively well preserved.

1. Core diagnostic features
   (a) Insidious onset and gradual progression
   (b) Language Disorder characterized by
      i. Progressive, fluent, empty spontaneous speech
      ii. Loss of word meaning, manifest by impaired naming and comprehension
      iii. Semantic paraphasias and/or
(c) Perceptual disorder characterized by
   i. Prosopagnosia: impaired recognition of identity of familiar faces
   and/or
   ii. Associative agnosia: impaired recognition of object identity
(d) Preserved perceptual matching and drawing reproduction
(e) Preserved single-word repetition
(f) Preserved ability to read aloud and write to dictation orthographically
    regular words

2. Supportive diagnostic features

(a) Speech and language
   i. Press of speech
   ii. Idiosyncratic word usage
   iii. Absence of phonemic paraphasias
   iv. Surface dyslexia and dysgraphia
   v. Preserved calculation
(b) Behaviour
   i. Loss of sympathy and empathy
   ii. Narrowed preoccupations
   iii. Parsimony
(c) Physical signs
   i. Absent or late primitive reflexes
   ii. Akinesia, rigidity, and tremor
(d) Investigations
(e) Neuropsychology
   i. Profound semantic loss, manifest in failure of word comprehension
      and naming and/or face and object recognition
   ii. Preserved phonology and syntax, and elementary perceptual pro-
      cessing, spatial skills, and day-to-day memorizing
(f) Electroencephalography: normal
(g) Brain imaging (structural and/or functional): predominant anterior tem-
    poral abnormality (symmetric or asymmetric)

A.3 Semantic variant PPA

[Mesulam and Weintraub, 1992] These criteria have been updated by Gorno-Tempini
et al [Gorno-Tempini et al., 2011]; I have given the old criteria as the majority of
the papers referenced here will have been of patients diagnosed with these criteria.
1. There is an insidious onset and gradual but progressive impairment of word finding, object naming, syntax, or word comprehension manifested during conversation or assessed with the use of standard neuropsychological tests of language.

2. All major limitations in activities of daily living can be attributed to the language impairment for at least two years after onset.

3. Premorbid language function (except for developmental dyslexia) is known to be intact.

4. Prominent apathy, disinhibition, loss of memory of recent events, visuospatial impairment, visual-recognition deficits, and sensory-motor dysfunction are absent during the initial two years of illness, as indicated by the history, evaluation of activities of daily living, or neuropsychological testing, so that the patient would not fulfill diagnostic criteria for any other dementia syndrome.

5. Acalculia (inability to perform simple mathematical calculations) and ideomotor apraxia (inability to pantomime movement as instructed by an examiner) can be present even in the first two years of illness, and deficits in copying simple drawings and perseveration may also be noted, but neither visuospatial deficits nor behavioural disinhibition substantially limits activities of daily living.

6. Other cognitive functions may be affected after the first two years of illness, but language remains the most impaired function throughout the course of the illness and deteriorates faster than other affected functions.

7. Specific causes of aphasia, such as stroke or tumour, as ascertained by neuroimaging, are absent.
Appendix B

B Goodness of fit for 2-parameter fit vs 3-parameter fit for 4 and 8 timepoints

The goodness of fit graphs for simulated data with 16 timepoints have been included in figure 3.17. Here I show the results for simulated data with 4 and 8 timepoints.

![Graphs of goodness of fit for 2-parameter vs 3-parameter fits](image)

(a) 2-p vs 3-p (bolus width)  
(b) 2-p vs 3-p (relaxation time)

**Figure B.1:** Comparison of goodness of fit for 4 timepoints for 2-parameter vs 3-parameter fits. The graphs show the p-values from f-tests comparing the 2-parameter fit with the 3-parameter fits, vs SNR. For the residuals from the 3-parameter fits to be relatively less than those from the 2-parameter fit, the p-value should be below 0.05. It is clear that even for the highest SNR this is never the case. 3 parameters cannot be fitted with 4 timepoints even at an SNR of 15.
Appendix B.

(a) 2-p vs 3-p (bolus width)  
(b) 2-p vs 3-p (relaxation time)

Figure B.2: Comparison of goodness of fit for 8 timepoints for 2-parameter vs 3-parameter fits. The graphs show the p-values from f-tests comparing the 2-parameter fit with the 3-parameter fits, vs SNR. For the residuals from the 3-parameter fits to be relatively less than those from the 2-parameter fit, the p-value should be below 0.05. It is clear that even for the highest SNR this is never the case. 3 parameters cannot be fitted with 8 timepoints even at an SNR of 15.
Appendix C

C Measuring labelling efficiency

C.1 Introduction

The labelling efficiency of the Phillips STAR sequence had been measured in the past, and given a result of 0.7, i.e. 70% of the available protons had their spins aligned by the sequence. However, an upgrade to the scanner resulted in an overall increase in perfusion measurements, and so it was decided to repeat the efficiency measurement.

C.2 Methods

A phantom containing 0.9 % by volume saline solution was scanned both using the standard STAR labelling sequence, with label and control images interspersed as usual. The position of the label plane is shown in figure C.3. The pre- and post-sequence saturation gradients were turned off to allow the ASL label to persist. A series of images were taken at delay times of 100 - 5000 ms, with 3 averages being taken at each time. A region of interest was defined within the labelling plane; the ROI is shown in figure C.3.

C.3 Results

Ideally there would be no label at all for the control images; however it can be seen from the control signal in figure C.2 that there is a small signal, as the control signal does not remain at zero after the label is applied. The signal for both label
and control pulses decays according to the inversion recovery equation

\[ labelSatRecovM = M^0 \left( 1 - 2\alpha' \left( e^{\frac{-t}{T_1b}} \right) \right) \]  

(C.1)

and the labelling efficiency \( \alpha \) is given by \( \alpha'_{label} - \alpha'_{control} \). The labelling efficiency after the upgrade is therefore 87%.

![Phantom for labelling efficiency measurement](image)

(a) Labelling region  
(b) ROI for Analysis

**Figure C.1:** Phantom for labelling efficiency measurement showing the labelling region, and the ROI for analysis

### C.4 Discussion

The labelling efficiency for the Manchester sequence should in future be set at 0.87. However, this is a fixed parameter in the calculation of perfusion parameters, so while an accurate value is important for estimates of absolute perfusion, it is less important when considering relative patterns within brain regions.
Figure C.2: MR signal for label and control images showing the development in time of the signal within the ROI for control images (circles) and label images (crosses) and the fit for an inversion recovery.
Appendix D

D Day 1 cognitive tests

D.1 Description of the tests used

These tests are pen and paper tests used in the first testing session. The tests were used to confirm the diagnostic status of patients, and the cognitive status of controls. Tests in italics were only used with patients. The standard cognitive tests used are:

- Addenbrookes Cognitive Examination Revised (ACE-R) including the Mini Mental State Exam (MMSE) [Mioshi et al., 2006].
  This is a brief screening test for dementia, and includes tests for attention, orientation, memory, fluency, language, and visuospatial skills.

- Naming and word-Picture matching from the Cambridge Cognitive Brain Sciences Unit Battery [Bozeat et al., 2000].
  This test assesses single-word semantic knowledge and comprehension. The participant is asked to name 64 objects shown as line drawings, and subsequently to select the same object from a set of 10 possibilities.

- Pyramids and palm trees test (3 picture version) [Howard, 1992].
  This test measured non-verbal semantic knowledge. The participant is shown a series of pictures, each with a stimulus object at the top, and two other objects below. The participant is asked to point to the lower object which is more related to the stimulus.

- Trail-making tests A and B [Reitan, 1958].
  This test assesses visual attention and task switching. The A test has a series of numbered dots; the B test has the same number of dots, but half are numbered, and half are labelled alphabetically. The participant is asked to
join the dots, and timed while doing so. Errors are corrected as and when they occur. For the A test the dots are joined in numeric order. For the B test, the numbers and letters are alternated (i.e. A-1-B-2...).

- **Repetition from the Western Aphasia Battery [Kertesz, 1982].**
  This test assesses verbal recall. The participant is asked to repeat a list of progressively more complex words and phrases read out by the examiner.

- **Digit span forwards and backwards from the Wechsler Memory Scale - Revised [Lezak et al., 2004].**
  This test assesses non-verbal recall. The examiner reads a progressively longer series of numbers, and the participant is asked to repeat the numbers, first in the same order as they were read, and subsequently in reverse order.

- **FAS letter fluency [Tombaugh, 1999].**
  This test assesses verbal fluency. The participant is to give as many words as possible beginning with the letters F, A and S, with one minute for each letter.

- **Cookie-theft picture description from the Boston Diagnostic Aphasia Examination [Goodglass and Kaplan, 1983].**
  This test is designed to elicit natural speech by asking the participant to describe a picture. The speech is then scored for information context, and for fluency, grammatical competence and paraphasias.

- **Rey figure copy and delayed recall [Corwin and Bylsma, 1993].**
  This test assesses visuospatial ability and visuospatial memory, as well as planning and attention. The participant is asked first to copy a complex diagram, then to reproduce it immediately afterwards from memory, and finally to reproduce it again after a delay of a few minutes.

- **Benton visual retention test [Benton, 1962].**
  This test assesses visuospatial ability and visuospatial memory, as well as planning and attention. The participant is shown a series of simple diagrams, then asked to reproduce each one from memory after a delay of a few seconds.

- **Raven’s coloured progressive matrices [Lezak et al., 2004].**
  This test assesses reasoning ability: the participant is asked to select which of 6 possibilities will complete a pattern.

- **Verbal memory - story with recall from the Wechsler Memory Scale - Revised [Lezak et al., 2004].**
  This test assesses verbal memory and attention. Two short stories (approx 60 words each) are read to the participant, who is then asked to repeat each story in the same words, both immediately and after a delay of a few minutes.
• Famous Faces recognition and naming from the Face-Place test. [Dudas et al., 2005].
  This test assesses familiarity with faces and associated names. The participant is shown a series of pages, each with 4 faces. 2 of the faces were of famous people, and the other 2 not. The participant was asked to point to the faces which were famous, and then to put a name to the face.

• PALPA Auditory sentence picture comprehension [Kay et al., 1992].
  This tests assesses verbal comprehension. The subject is asked to point to the appropriate picture which agrees with an action read by the examiner.

D.2 Results

No correction has been made to the p-values reported here for multiple comparisons. The tests are descriptive for individual patients; the group comparison is made only to indicate how well the test has performed.
### Table D.1: Results of standard cognitive tests

<table>
<thead>
<tr>
<th>Test</th>
<th>FTLD patient</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE-R (100)</td>
<td>65.7 ± 16.6</td>
<td>96.2 ± 3.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>MMSE (30)</td>
<td>25.9 ± 4.1</td>
<td>29.7 ± 0.7</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>WAB Repetition</td>
<td>90.1 ± 10.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Digit span forward</td>
<td>8.5 ± 1.6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Digit span backward</td>
<td>5.9 ± 2.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Trail-making test A (sec)</td>
<td>54.2 ± 2.3</td>
<td>32.4 ± 11.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Trail-making test B (sec)</td>
<td>101.3 ± 46.8</td>
<td>57.7 ± 28.9</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>FAS letter fluency</td>
<td>21.2 ± 11.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rey figure copy (36)</td>
<td>29.6 ± 11.4</td>
<td>35.7 ± 0.6</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Rey figure immediate recall (36)</td>
<td>7.9 ± 8.3</td>
<td>24.4 ± 4.6</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Rey figure delayed recall(36)</td>
<td>1.6 ± 9.1</td>
<td>23.6 ± 4.8</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Benton test (10)</td>
<td>4.3 ± 2.0</td>
<td>8.1 ± 1.6</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Raven’s matrices</td>
<td>28.2 ± 6.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Naming(64)</td>
<td>46.6 ± 13.9</td>
<td>63.4 ± 0.6</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Word-Picture matching (64)</td>
<td>45.9 ± 14.9</td>
<td>63.6 ± 0.8</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Pyramids and palmtrees (52)</td>
<td>40.6 ± 9.5</td>
<td>50.9 ± 1.0</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Verbal memory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>immediate recall</td>
<td>5.1 ± 3.4</td>
<td>11.8 ±3.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>delayed recall</td>
<td>2.3 ± 3.0</td>
<td>10.1 ±3.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Famous Faces</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition (20)</td>
<td>13.4 ± 5.3</td>
<td>18.1 ±3.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Naming (20)</td>
<td>5.8 ± 6.4</td>
<td>15.9 ±4.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>PALPA (6)</td>
<td>5.9±0.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cookie-theft description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>info content (10)</td>
<td>6.6 ± 2.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>fluency (10)</td>
<td>8.1 ± 1.7</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Bibliography


