THE UNIVERSITY
of EDINBURGH

Quantifying structural changes in the ageing brain from magnetic resonance imaging

Natalie Anne Royle

Doctor of Philosophy

University of Edinburgh

2015

Declaration i

Acknowledgements ii

Abstract iii

Publications arising from this thesis vi

Section 1: Introduction and methods 7

Chapter 1. General Introduction 7

1.1 Ageing and society 7

1.2 Healthy ageing and ageing research 8

1.3 Magnetic Resonance Imaging and its role in studying ageing processes 9

1.4 The ageing brain 12

1.5 Research into ageing 14

1.6 Assessing the ageing brain on MRI 17

Chapter 2. General methods 19

2.1 LBC1936 cohort 19

2.2 Subjects 20

2.3 Cognitive testing 21

2.4 Brain MRI acquisition protocol 23

2.5 Pre-processing steps and Image analysis 24

2.6 Normal-appearing white matter (NAWM) and cerebrospinal fluid (CSF). 25

2.7 Grey matter volume 27

2.8 Whole brain volume 32

2.9 Ventricular volume 33

Section 2: Systematic reviews 38

Chapter 3. Systematic review of automated hippocampal segmentation methods in healthy elderly population 39

3.1 Introduction 39

3.2 Methods 41

3.2.1 Search strategy 41

3.2.2 Inclusion/exclusion criteria 41

3.2.3 Data extraction and synthesis 41

3.3 Results 42

3.3.1 Study search and evaluation 42
3.3.2 Included populations 47
3.3.3 Brief description of automated methods 48
3.3.4 Statistical analysis used to compare methods 50
3.3.5 Reported discrepancies between methods 54
3.4 Discussion **56**

Chapter 4. Systematic review investigating reported protocols for identification of the posterior frontal lobe boundary 60

4.1 Introduction **60**
4.1.1 Anatomical definition 61
4.2 Methods **62**
4.2.1 Search strategy 62
4.2.2 Inclusion/Exclusion criteria 62
4.2.3 Data Extraction and Synthesis 62
4.2.4 Study Selection 62
4.2.5 Study Characteristics 65
4.3 Results **65**
4.4 Discussion **68**

Section 3. Global measures 69

Chapter 5. Intracranial volume as a proxy for brain size in youth 70

5.1 Introduction **70**
5.2 Method **71**
5.3 Results **71**
5.4 Discussion **73**

Chapter 6. Intracranial volume measurement 74

6.1 Introduction **74**
6.2 Method **75**
6.2.1 Subjects 75
6.2.2 Manual method or reference standard 76
6.2.3 Test method 1 [Object Extraction Tool (OET)] 77
6.2.4 Test method 2 [OET plus editing aided by a colour fusion method] 77
6.2.5 Test method 3 [Brain Extraction Tool (BET)] 79
6.2.6 Exclusion of the pituitary fossa 79
6.2.7 Statistical analysis 81

6.3 Results 81
6.4 Discussion 84

Chapter 7. Comparison of ICA with ICV 86
7.1 Introduction 86
7.2 Method 86
7.3 Result 88
7.4 Discussion 89

Chapter 8. Influence of inner table skull thickening on ICV 90
8.1 Introduction 90
8.2 Method 91
8.2.1 Subjects 91
8.2.2 Measurement method 91
8.2.3 Statistical analysis 93
8.3 Results 93
8.4 Discussion 96

Chapter 9. Measuring inner table skull thickening using ICA 99
9.1 Introduction 99
9.2 Method 99
9.2.1 Statistical analysis 101
9.3 Results 101
9.4 Discussion 102

Section 4: Regional measures 103

Chapter 10. Automated vs Manual segmentation of the hippocampus. 103
10.1 Introduction 103
10.2 Methods 105
10.2.1 Reference method 105
10.2.2 Automated methods 107
10.2.3 FSL_FIRST protocol used in the LBC1936 107
10.3 Results 111
10.4 Discussion 116

Chapter 11. Practical application of frontal lobe boundary protocols 119
11.1 Introduction 119

11.2 Methods 121
11.2.1 Participants 121
11.2.2 Image analysis 121
11.2.3 Boundary definitions 121
11.2.4 Cut–plane boundaries 123
11.2.5 Sulcal boundaries 124
11.3 Results 127
11.4 Discussion 132

Section 5: Associations with cognition 134

Chapter 12. Brain size and cognition 134
12.1 Introduction 134
12.2 Methods 136
12.2.1 Participants 136
12.2.2 Statistical Analysis 137
12.2.3 Model Specification 137
12.2.4 Measurement Invariance 138
12.2.5 Model Evaluation 138
12.3 Results 138
12.4 Discussion 145

Chapter 13. Hippocampus and cognition 148
13.1 Introduction 148
13.2 Methods 150
13.2.1 Subjects 150
13.2.2 Image analysis 150
13.2.3 Cognitive Ability Measures 152
13.2.4 Statistical Analysis 152
13.3 Results 153
13.4 Discussion 158

Section 6: Discussion 162

Chapter 14. General discussion 162
14.1 Global measures 164
14.2 Regional measures 165
14.3 Limitations 168
14.4 Future work 169

Appendices 171
Appendix 1: Data extraction frontal lobe systematic review 171
Appendix 2: References for frontal lobe systematic review 191
Appendix 3: Hippocampal segmentation protocol 210
Appendix 4: Bland-Altman plots Chapter 10.2.3 215
Appendix 5: Linear regression models for all subjects 217
Appendix 6: Linear regression models subjects MMSE above 27 220
Appendix 7: Linear regression models for imaging parameters 223
Appendix 8: Bivariate correlation coefficients 224
References 225

Declaration

The University of Edinburgh

Name of Candidate:	Natalie Royle	*UUN*	S097778
University email:	Nat.royle@ed.ac.uk		

Degree Sought:	PhD	*No. of words in the main text of Thesis:*	44,517
Title of Thesis:	Quantifying structural changes in the ageing brain from magnetic resonance imaging		

I certify:

(a) that the thesis has been composed by me, and

(b) either that the work is my own, or, where I have been a member of a research group, that I have made a substantial contribution to the work, such contribution being clearly indicated, and

(c) that the work has not been submitted for any other degree or professional qualification except as specified.

Signature:	

Acknowledgements

I would like to extend an enormous thank you to my supervisors Professor Joanna Wardlaw and Dr. Maria Valdés Hernández for their enduring support and guidance throughout this process. From their many seeds of wisdom, miraculously a researcher grew. Myself and those who collaborated on this work would like to acknowledge the financial support provided by the Medical Research Council and Age UK.

I would like to thank Professor Benjamin Aribisala for conducting the computational processing for the multimodal analysis (chapter 13) and considerable input with the statistical analysis of the results. I would also like to thank Dr Thomas Booth for conducted the statistical analysis of the results presented in chapter 12. Many thanks to those researchers involved with the LBC1936 and running of the Brain Research Imaging Centre. Especially Professor Ian Deary, Dr. Susana Muñoz Maniega and Dr. Simon Cox for their invaluable help, input and advice when producing the papers that have made up several chapters of this thesis. I will be eternally indebted to Moira without whom I would not have got beyond the front door, and to Duncan who actually did get me beyond the front door at BRIC.

The most heartfelt gratitude goes to Dr. Karen Ferguson for her unending patience when teaching me the dark arts of image analysis and to Dr. Mark Bastin for his unwavering optimism and encouragement at all times but especially when it was needed most. Thank you to all the other PhD students, Anna J, David, Andreas, Jehill, Anna H, Islem and Colin for being both drinking and ideas buddies throughout. A special mention to Xin for her wonderful friendship and to Simon for being Schwarzenegger, both of which made me smile.

To my lovely friends, I offer you a lifetime supply of tea, cake and support in the hope that I can pay back the tiniest amount that you have given me over the course of this PhD and beyond. Cheers to my brothers for giving me some perspective when I collapsed in on myself and for making me more resilient than I ever thought I could be. To my parents, I just hope you are proud of me for sticking at it because without you I would never have even tried, so really this achievement was always for you both. And to Rob, thank you for your steadfast support without which there would be no thesis, plus now we can finally go on holiday.

Abstract

Understanding the ageing process is of increasing importance to an ageing society and one aspect of this is investigating what role the brain has in this process. Cognitive ability declines as we age and it is one of the most distressing aspects of getting older. Brain tissue deterioration is a significant contributor to lower cognitive ability in late life but the underlying biological mechanisms in the brain are not yet fully understood. One reason for this is the difficulty in obtaining accurate measures of potential ageing-related brain biomarkers.

The chapters in this thesis explore the difficulties of quantifying brain changes in the ageing brain from Magnetic Resonance Imaging (MRI), and how the changes identified are related to cognition in later life. The data was acquired as part of the second wave of the longitudinal Lothian Birth Cohort 1936 study in which 866 people aged 73 years, returned for cognitive and medical assessment. At this stage of the study 702 underwent MR imaging resulting in 627 complete datasets across all testing. The entire data, a randomly chosen subset of 150 and 416 freely available data were used to investigate global and regional measurement methods in older brains and how the resultant measurements related to cognitive performance. Furthermore the presence of early life cognitive data in the form of a general intelligence test sat at age 11, served as an indicator of cognitive ability prior to the potential influence of the ageing process.

The chapters concerning global measures at first establish, that a measure of intracranial volume (ICV) serves as both a way of correcting for individual differences in brain size between participants and as a proxy premorbid measure of brain size. The analysis, utilising freely available cross-sectional MRI data (http://www.oasis-brains.org) revealed that ICV differed very little between 18-28 year olds and 84-96 year olds where as total brain tissue volume (TBV) differed by 14.1% between the two groups, which was more than twice the standard deviation across the entire age range (18-96 years). Second a validated, reliable method for measuring ICV was investigated using 150 people randomly chosen from the LBC1936 study. Automated and semi-automated methods were validated against reference measurements the results of which showed that common ageing features

make automated and semi-automated methods that do not have an additional manual editing step, ineffective at producing accurate ICV measurements. This analysis also highlighted the need to employ additional spatial overlap assessment to volumetric comparison of measurement methods to reduce the effect of false-positives and false-negatives skewing apparent discrepancies between methods. Using the information gained here ICV and TBV from the entire LBC1936 cohort were analysed in a structural equation model, alongside cognitive ability measures at both age 11 and age 73. We found that TBV was a stronger predictor of later life cognitive ability, after accounting for early life ability, but that a modest association remained between ICV and late life cognition. This suggests that early life factors pay a role in how well we age, though the relationship is complex.

The regional measures chapters look at two brain regions commonly associated with ageing, the hippocampus and the frontal lobes. Measuring either of these brain regions in large samples of healthy older adults is challenging for many reasons. The hippocampus is small and as with all brain regions shows greater variation in older age, this makes employing automated methods that have the advantage of being fast and reproducible difficult. Following the results of our systematic review of automated methods for measuring the hippocampus, the two most commonly used and available automated methods were validated against reference standard measurements. The results indicated that although automated methods present an attractive alternative to laborious manual measurements they still require manual editing to produce accurate measurements in older adults. The modified strategy employed across the LBC1936 was to use an automated method and then manually edit the output; these segmentations were used to investigate the potential of multimodal image analysis in clarifying associations between the hippocampus and cognitive ability in old age. The analysis focused on associations between longitudinal relaxation time (T1), magnetization transfer ratio (MTR), fractional anisotropy (FA) and mean diffusivity (MD) in the hippocampus and general factors of fluid intelligence, cognitive processing speed and memory. The findings show that multi-modal MRI assessments were more sensitive than volumetric measurements at detecting associations with cognitive measures.

The difficulty with producing a relevant frontal lobe measure was made apparent when the result of a large systematic review looking at the manual protocols used revealed 19 methods and 15 different landmarks had been employed. This resulted in an analysis that took the 5 most common boundaries reported and applied them to 10 randomly selected participants from the LBC1936. The results showed significant differences between the resultant volumes, with the smallest measurement when using the genu as the posterior marker representing only 35% of the measurement acquired using the central sulcus. The results from the studies presented in this thesis strongly highlight the need to develop age specific methods when using brain MRI to study ageing. Furthermore the implications of using unstandardised protocols, making assumptions about a methods performance based on validation in younger samples and the need to account for early life factors in this area of research have been made clearer. Studies building on these findings will be beneficial in elucidating the role of the brain in ageing.

Publications arising from this thesis

Royle, N.A., Booth, T., Valdés Hernández, M.C., Penke, L., Murray, C., Gow, A.J., Muñoz Maniega, S., Starr, J., Bastin, M.E., Deary, I.J., Wardlaw, J.M. (2013). Estimated maximal and current brain volume predict cognitive ability in old age. *Neurobiology of Aging*, 34(12), 2726-2733.

Royle, N.A., Valdés Hernández, M.C., Muñoz Maniega, S., Aribisala, B.S., Bastin, M.E., Deary, I.J., Wardlaw, J.M. (2013). Influence of thickening of the inner skull table on intracranial volume measurement in older people. *Magnetic Resonance Imaging*, 31(6), 918-922.

Cox, S.R., Ferguson, K.J., Royle, N.A., Shenkin, S.D., Macpherson, S.E., MacLullich, A.M., Deary, I.J., Wardlaw, J.M. (2014). A systematic review of brain frontal lobe parcellation techniques in magnetic resonance imaging. *Brain Structure and Function*, 219(1), 1-22.

Aribisala, B.S., Royle, N.A., Muñoz Maniega, S., Valdés Hernández, M.C., Murray, C., Penke, L., Gow, A.J., Starr, J., Bastin, M.E., Deary, I.J., Wardlaw, J.M. (2013). Quantitative multi-modal MRI of the Hippocampus and cognitive ability in community-dwelling older subjects. *Cortex*, 53, 34-44.

Aribisala, B.S., Royle, N.A., Valdés Hernández, M.C., Murray, C., Penke, L., Gow, A.J., Muñoz Maniega, S., Starr, J., Bastin, M.E., Deary, I.J., Wardlaw, J.M. (2014). Potential effect of skull thickening on the associations between cognition and brain atrophy in ageing. *Age and Ageing*, *In Press*.

Section 1: Introduction and methods

Chapter 1. General Introduction

1.1 Ageing and society

The Global Age Watch Index (2013) states that 23% of the UK population is over 60 years old but by 2050 that is expected to rise to 29.6%, by which time nearly a third of our population will be over 60. Projections by the UK government Actuary's department suggest that population increase to 2033 will be concentrated in older groups, with the most significant increase in 65-69 year olds. The impact of an ageing population on society has costs and benefits. Financially the impact on the NHS is vast, with the average expenditure on retired households being nearly double that spent on non-retired households. The other financial consideration is the amount spent on pensions, which is the largest expenditure (65%) by the Department of Work and Pensions (Cracknell, 2010). Along with this the recent announcement in the UK government's autumn statement that the age at which we can access our state pension will increase to 70 years of age, should put the discussion of successful ageing at the top of society's agenda (HM Treasury, 2013). The assumption made by the government that we can link things such as pensionable age with life expectancy overlooks a fundamental tenet of this assumption, that we are relatively healthy and able in those extra years.

The prevalence of age-related diseases such as arthritis and dementia further support the idea that we do not all age successfully. The notion that healthy ageing is possible would seem to contradict the negative associations with getting old that are prominent in society, and it is this very idea that has been the catalyst behind much ageing research. The wealth of experience and knowledge individuals possess by the time they reach the latter part of their lives is an invaluable resource, one that is undervalued by the wider community. The huge contribution older people can make to society is often overshadowed by the challenges of ageing and though these seem insurmountable, ageing research aims to tackle these challenges. If we can discover why some people manage to stay independent, relatively physically fit and active, and cognitively able well into later life the benefits would be enormous. We could improve the lives of a substantial portion of the population, reduce the requirements

on resources and help to diminish the more negative aspects of ageing, potentially removing stigma by replacing it with a more positive outlook for everyone.

The natural extension of studying ageing is to also look at age-related diseases such as Parkinson's or Alzheimer's disease, as non-diseased individuals often show similar pathologies, such as ß-amyloid plaques, to dementia sufferers. Dementia is a set of symptoms that include problems with reasoning, communication and memory as well significant mood changes. It is prevalent in older adults with around 800,000 sufferers in the UK, though it is not confined to over 65 year olds with 1 in 1,400 people being aged between 40 and 64 years. The occurrence of dementia rises continuously from 60 to 70 and by the age of 80+ years 1 in 6 people will have a form of the disease (Alzheimer's Society, 2014). With an ageing population the Alzheimer's society estimates that by 2021 there will be a million people with dementia, rising to 1,700,000 by 2051. The cost of dementia to the UK is approximately £23bn and this along with the obvious failings (Dewing and Dijk, 2014) to adequately provide care for people with dementia has resulted in the first G8 Dementia summit (London UK, 11/12/2013) to be held. The major outcome was a pledge to double funding for dementia research in the UK from £66m in 2015 to £122m by 2025, which is still low when compared to funding for other terminal diseases such as cancer that stood at £267m in 2008 (Cancer Research UK, 2014; Siddique, 2013). The increase in funding pledged to dementia research is encouraging but to have the greatest impact all ageing research should receive increased funding.

1.2 Healthy ageing and ageing research

There are many factors that determine how well we age or how vulnerable we are to age-related diseases. Our genes determine some of our biological make up but our environment, lifestyle and social setting across the life span, as well as the interaction between all these factors, affects how robust we are in our later years. The wide variation in healthy ageing makes for an incredibly complex topic for researchers to tackle. A multi-disciplinary approach is essential in achieving real progress and understanding brain ageing is one of the key aspects. The effects of ageing on the brain are widespread, with noticeable changes to the cells, vasculature, anatomy and

performance. Individual differences in these changes are varied but generally the brain reduces in volume, becomes more atypical in appearance and lesions to the white matter become more common (de Leeuw et al, 2011). These factors are related to cognitive changes also common in older age, as well as things such as alterations in mobility (Starr et al, 2003) and personality (Davis et al, 2012). Prior to the application of brain imaging to this area of research, these changes were only detectable post-mortem therefore making it difficult to determine if they were only related to the cause of death rather than other processes (Creasey and Rapoport, 1985). Brain imaging and specifically magnetic resonance imaging (MRI) for the first time permitted researchers to look at the ageing brain in-vivo.

1.3 Magnetic Resonance Imaging and its role in studying ageing processes

There are many imaging techniques used to study the human brain such as Computed Tomography (CT), Positron Emission Tomography (PET) and functional magnetic resonance imaging (fMRI). Where fMRI and PET are used to measure metabolic and blood flow changes in the brain CT and MRI provide structural information about the brain (Filler et al, 2009). CT is most often used in emergency diagnostic of brain injury, stroke or haemorrhage, as it is quick and accurate and widely available in the UK. However with every CT the person being imaged receives a dose of ionizing radiation that means it is used only when necessary. MRI has therefore become more widely used in research settings as there are no reported risks provided contraindication safety measures are followed. MRI works by administering carefully calculated radio frequency pulses, produced by a strong magnetic field, to tissues. These pulses, controlled by gradient coils, excite hydrogen atoms and the time taken for the hydrogen atoms to return to an equilibrium state provides information about tissues. This relaxation time is what is used to distinguish structures from one another in the body as the water content or amount of hydrogen atoms present in a tissue is variable. MRI emerged gradually through several developments in the wider imaging research field but it can be said that the first MRI machine was created by Professor Raymond Damadian in 1972 (Geva et al, 2006). The idea of introducing gradients to the magnetic field was that of Professor Paul Lauterbur who by collecting spatial information about nuclei from radio waves, produced the first 2 and 3 dimensional

images. Professor Peter Mansfield at the University of Nottingham further developed the mathematics of this technique. Professor Mansfield's work allowed researchers to, for the first time, usefully interpret the signals received from MRI. The first clinical application was by a team at the University of Aberdeen, led by Professor John Mallard, where they used an in-house built full body MRI scanner to detect signs of cancer in a patient. The work carried out by Professor Mallard and colleagues encouraged the widespread use of MRI in clinical diagnosis and beyond (Geva et al, 2006). From this point on, MRI has developed rapidly with the invention of ever more sophisticated techniques and equipment resulting in huge breakthroughs in understanding disease states and progression (Chetelet & Baron, 2003), to charting human brain growth (Oishi et al, 2013), and measuring brain activity through changes in blood flow (Bennett & Rypma, 2013).

The use of brain imaging in normative research has contributed to our understanding of how the brain develops and changes over the life course. Brain development is dramatic going from a single cell to an estimated 100 billion neurons in a matter of years. The rapid changes of childhood and adolescence slow somewhat in adulthood but nonetheless continue throughout our lives and neuroimaging has permitted researchers to see these changes in vivo. It is now common for brain imaging studies that are investigating age-related changes to look for associations between cognitive and/or behavioural measures with volume estimates of brain structures. Other neuroimaging techniques such as PET and fMRI have been used to look at metabolic or blood flow activation patterns in relation to cognitive tasks to assess potential correlates between brain structure and function (Bennett & Rypma, 2013). Most commonly cross-sectional studies have been used to look at age-related differences rather than longitudinal studies, which given that brain imaging has only been possible for 30 years, is understandable. Cross-sectional analysis has the benefit of application in a broad age range, making it possible for us to investigate the similarities and differences between 18 and 90 years olds and all in-between. This provides a picture of ageing across the lifespan allowing us to plot the trajectory of age-related differences from maximum capacity to almost end state. However cross-sectional studies struggle with individual differences and cannot assess change over time as they provide data at a single time-point. Longitudinal studies permit researchers to look for age related changes by performing repeat analysis of the same

participants at several time-points, however often only the healthiest people are assessed because of attrition rates due to illness or death.

Quantifying the volume of brain structures from MRI has a number of very useful applications in ageing research as well as in other fields. The most common methods used to assess brain regions are visual rating scales, manual, semi-automated and automated methods. Visual rating scales were the first to emerge as they can be quickly applied in a clinical setting with no extra technical equipment needed. Large numbers of scans can be rated in a relatively small amount of time but they rely heavily on the expertise of the rater. They are predominantly used in clinical assessments and have shown to be comparable to volumetric measurements when estimating cortical atrophy (Galton et al, 2001; Möller et al, 2014). They are also extensively used in assessing white matter changes (Valdés Hernández et al, 2013; Wahlund et al, 2001). Despite their usefulness visual rating scales do not provide precise quantification of structures often sought in a research setting. Manual methods are more effective at supplying detailed volume measurements. They involve the delineation of a target structure by tracing the region on contiguous slices on MR images, with guidance from validated anatomical protocols. Manual methods are labour intensive, time consuming and are more prone to inter-rater errors than automated methods. Advancements in computer technology, allowing utilization of sophisticated computational algorithms, led the way for the development of semi-automated and automated methods. Along with greater reproducibility, automated methods have the advantage of reducing both user input and the time it takes to produce volume segmentations (Bozzali et al, 2008). Despite these huge benefits, automated methods can struggle with less than typical brains as they rely on templates, normative values and common features to guide processing. Normal older adults brains commonly show features such as differential rates of brain tissue atrophy and increased presence of incidental findings, when compared to the rest of the adult population.

Whether performing manual or automated segmentations age-related features in the brain, such as pronounced asymmetry and pathology like white matter hyperintensities all contribute to making measurements difficult. Variation in the literature, discussed below, will in part be due to discrepancies in the use and

application of a variety of different methodologies. Longitudinal MRI studies provide information about brain changes, which can better reflect the ageing process, however multiple time point scans over many years adds up to a large body of data. If further investigation of the potential influence of multiple life factors implicated in ageing is to be performed, large datasets will be needed to accommodate the multivariate analysis necessary to study these potential associations. These two factors as well as the rising proliferation of imaging-genetic studies, are leading to researchers using large datasets to investigate brain ageing.

1.4 The ageing brain

Studies, both cross sectional and longitudinal have shown that there is considerable variation between individuals when looking at brain size and structure, and this is reflected in the way in which our brains differentially age. Post mortem and brain imaging studies have revealed that brain tissue volume starts to decrease once we have reached adulthood and the rate of reduction increases with age, accelerating significantly over age 70 years (Takao, Hayashi and Ohtomo, 2012). Several studies utilising a variety of methodologies have found grey matter volume decline to be associated with age (Giorgio et al, 2010; Peelle, Cusack, Henson, 2012; Sowell et al, 2003;Taki et al 2011; Tisserand et al, 2004). Studies have reported region specific deterioration with some suggesting that deterioration is most prominent in the temporal and prefrontal regions (Fjell et al, 2009; Raz et al, 2010). Grey matter volume decrease is possibly due to neuronal cell death that starts in early adulthood and continues to degenerate throughout our lives. White matter has also been found to deteriorate with age in volume (Ge et al, 2002; Gunning-Dixon et al, 2010; Resnick et al, 2000; Salat et al, 2009) and integrity (Bennett and Madden, 2013; Burgmans et al, 2010; Kochunov et al, 2009). White matter decline is due to degeneration of the myelin sheath that begins later in life than grey matter decline, at around age 40-50 years (Allen et al, 2005; Ikram et al, 2008). Because white and grey matter are physiologically different tissues that develop at different rates it seems clear that they would be differentially affected by ageing. Therefore investigating global brain measures such as total brain tissue volume along with grey and white matter will give a better indicator of the cumulative effect of ageing on the brain. Regional brain

measures are most useful in understanding structure function relationships such as those explored between memory and the hippocampus (den Heijer et al, 2012).

Though overall brain volume decreases with age, tissue loss is not homogenous and varies a great deal between brain regions (Coffey et al, 1992). Studies have reported age-related atrophy in the hippocampus (Raz et al, 2010; Scahill et al, 2003) and the thalamus (Sullivan et al, 2004; Xu et al, 2000) with others finding no changes in either the former (Du et al, 2006; Van Petten et al, 2004) or the latter (Jernigan et al, 2001). The sub regions of the basal ganglia show a lot of variation when associating volume with age. A reduction in caudate volume has been consistently reported with increasing age (Bergfield et al, 2010; Hasan et al, 2008; Raz et al, 2005; Walhovd et al, 2005), the putamen has also been found to have this association (Fjell et al, 2013; Raz et al, 2010) but with some gender effects indicating the association exists in men only (Nunnemann et al, 2007). More gross anatomical regions have been cited as being differentially affected by age-related decline such as the prefrontal/frontal cortex (Lemaitre et al, 2012; Tisserand et al, 2002) and the temporal lobe (Jiang et al, 2014). Allen et al (2005) explained the heterogeneous pattern of deterioration in terms of structures such as the frontal, parietal and occipital lobes showing a negative linear relationship between volume and age, where as the temporal lobes showed a curvilinear (cubic) relationship with volume decreasing little until an accelerated decrease beyond age 70 years. Zeigler et al (2012) also discuss patterns of decline in terms of linear and non-linear trajectories, suggesting that adaptive behaviour in certain areas may be as result of earlier and/or more pronounced decline in other regions. This results in those regions that are potentially engaged in adaptive behaviour delaying decline as a result of general ageing. This explanation sheds some light on why a consensus does not seem to have been found to answer the fundamental questions as to which brain structures are most affected by ageing and when these effects start to become prominent. It may be that along with methodological differences in measurements, discussed in further Sections (2-4) of this thesis, the range of ages in later years are influencing the reported associations. Age differences in the later years may be as nuanced as those in our earlier years, but with a more complex pattern due to cumulative effects acquired during very different life experiences.

Gender differences in brain structure are present throughout brain development and have been reported in adulthood (Ruigrok et al, 2014). Assessing the potential interactions of these differences with brain ageing are important in understanding why factors such as life expectancy and dementia rates differ between men and women.

1.5 Research into ageing

Approximately two-thirds of those people in the UK with a dementia diagnosis are women but this may be because dementia prevalence increases with age and women have a longer life expectancy than men rather than the differences in brain structure. Men have approximately 10% larger intracranial and brain tissue volumes than women, which reflects the generally larger body size of males over females. There is a suggestion in the literature that larger brains can better cope with the cumulative insults accrued during the life course (Staff et al, 2004). The cerebral reserve theory is discussed in context in Chapter 12. Though there are gender differences in brain function the picture is complicated. There is no clear evidence that general intelligence is linked to gender or that the larger male brain equates to better cognitive ability (Beltz, Blakemore & Berenbaum, 2013). Therefore the potential misapprehension that bigger is better is not necessarily the case, how well an organ in the body functions cannot just be attributed to its size. Region-specific differences in brain structure are present between the sexes, with volume differences being reported in regions such as the hippocampus, cingulate, putamen and insula (Ruigrok et al, 2014). This may in part be due to the influence of hormonal factors that contribute to the development of sexual dimorphisms in the brain. Or varied patterns have emerged due to an adaptive response driven by the physiological differences between the sexes (Tobet et al, 2009).

The reporting of gender differences in age-related brain atrophy have been mixed with some suggestions that there are no differences (Greenberg et al, 2008; Takao, Hayashi & Ohtomo, 2012) and others finding that grey and white matter (Allen et al, 2003) cortical thickness (Jiang et al, 2014) and subcortical structures (Takahashi et al, 2011) differed between men and women. A longitudinal study with a broad age range (10 to 85 years) found that the lateral ventricles and sylvian fissure expanded more rapidly in

men over age 60 years than in women of the same age suggesting that general brain volume decline in later life is more pronounced in males than in females (Pfefferbaum et al, 2013). However another 1 year follow-up study found that women showed a faster rate of atrophy than men but only in the group with mild cognitive impairment and not in the healthy control group (Hua et al, 2010). It would seem intuitive that gender-age interactions would be present when looking at brain tissue decline because developmental differences exist, along with differences in behaviour and prevalence of neurodegenerative diseases between males and females (Gur & Gur, 2004). Therefore the way in which volume measurements are adjusted could be one of the reasons why there has been some variation in the reported literature. Adjusting for the fact that women are on average smaller than men is an important consideration when looking at potential gender differences in brain volume. Using measures such as head circumference or skull size are common, and in neuroimaging studies this is most often achieved by using a measure of the cranial vault known as intracranial volume (ICV). Findings that adjusting for ICV effect interactions between age and brain volume, when gender is included, have been reported (Greenberg et al, 2008; Pfefferbaum et al, 2013). Therefore attention should be paid to the practice of accounting for individual difference in head size when looking at brain tissue volumes within and across groups.

Research from healthy ageing significantly contributes to dementia research by not only providing normative values by which a person's cognitive and biological performance can be measured but by discovering early biomarkers of the disease and by improving methods. Studies such as those looking at cognitively normal individuals who go on to develop dementia are vital when looking for potential pre-clinical brain biomarkers. It has been suggested that rapid tissue loss often precedes the onset of dementia (Ikram et al, 2010), making accurate quantification of normal brain deterioration trajectories important for potential identification of those at risk of developing dementia. Dickerson et al (2011) found that of those older adults measured as cognitively normal but who went on to develop Alzheimer's, subtle cortical thinning up to 10 years prior to onset was detectable. Other studies have suggested that atrophy of specific structures such as the hippocampus (Barnes et al, 2009; Costafreda et al, 2011; Ezekiel et al, 2004) or ventricles (Madsen et al, 2013; Thompson et al, 2004) are indicators of a pre-dementia state and that volume

measurements could be used as biomarkers for the disease. The Alzheimer's Disease Neuroimaging Initiative reported that they found significantly more atrophy in those with Alzheimer's than MCI sufferers or healthy controls (Leow et al, 2009). In fact in a review of their findings to date the group reported that overall the MRI measures of change were the most efficient at predicting future decline compared to the other biomarkers assessed (Weiner et al, 2013). Structural changes observed in normal ageing studies may not all be age-related and might be connected to events much earlier in life (Penke et al, 2012). The size and function of the brain in our early years influences the size and functional ability of our brains in old age (Shenkin et al, 2009a). Interactions between factors that span a lifetime such as level of education, exercise, social integration, genetic, environment and nutrition may all lead to events that influence the state of the brain and its capacity to cope with additional insults in later life (Batouli et al, 2014; Brandt, Deindl & Hank, 2012; Desai, Grossberg & Chibnell, 2010).

One of the most apparent and distressing aspects of age-related brain decline is in the diminishment of cognitive abilities as we get older (Deary et al, 2009a). The neurobiological underpinnings of cognitive ageing are not clear and though we know structure and function are related our understanding of the nature of this relationship is continually evolving. Determining pathological from non-pathological cognitive decline is usually done by clinical screening tools, which categorise older adults into discrete groups such as having dementia or being mildly cognitively impaired. Though this is necessary for treatment it is not as useful when trying to understand those who have either been tested in a pre-disease state or those who show some deviation from the group norm but are not classified as cognitively impaired. Also understanding individual differences within a normative population is important if we want to uncover whether a person's cognitive ability has changed from their peak state or is just low relative to their peers. As already mentioned childhood intelligence appears to affect health in later life, possibly due to it resulting in better opportunities and decision making over the lifetime or due to a protective mechanism afforded by a higher starting level from which to decline (Penke et al, 2012). As with structure the question as to whether cognitive ageing is a general process or affects specific cognitive domains more than others is a continual debate in ageing research (Abbott et al, 2012). Studies have found age-related cognitive decline in a number of domains

such as memory (den Heijer et al, 2012), processing speed and executive function (Albinet et al, 2012). As cognition originates in the brain it has been suggested that age-related brain atrophy may be driving the cognitive deficits also associated with ageing. Associations between brain tissue atrophy and cognitive ability are mixed (Ferriera et al, 2014; Taki et al, 2011), with both global and local associations being reported (Cardenas et al, 2001; Kooistra et al, 2014). Rodrigue and Kennedy (2011) provide a comprehensive review of the literature in this area and say that association cortices are more vulnerable than primary sensory areas to the affects of ageing. They go on to suggest that a variety of structural techniques should be applied depending on the cognitive domain being assessed. For example white matter integrity rather than volume seems to be associated with information processing speed therefore using a technique sensitive to detecting this, such as DTI, would be more effective than measuring white matter volume when looking for age-related associations.

1.6 Assessing the ageing brain on MRI

Factors influencing the reliability of brain volume measures are several and range from image quality and scanner type to pre-processing and segmentation protocols (Jovocich et al, 2009). Traditionally to assess the validity of a method, the reference standard measures have been provided by manual tracing of brain structures performed by experienced individuals (Jack Jr et al, 1990). However a common criticism of manual segmentation is that rater error is inevitable and influences the accuracy and reliability of the measurements. Though rater error is a factor, it can, through training and experience, be minimised but this is still dependent on the use of well researched, standardised manual tracing protocols. The variation in manual tracing protocols may be an important factor in explaining why regional brain volume measurements vary in the literature. If each study is working from a slightly different set of anatomical boundaries the ability to compare results becomes difficult (Konrad et al, 2009). This is also a concern for automated methods that use a priori information, necessary for the algorithm to determine anatomical locations, based on manual labeling or segmentations. There have been efforts to produce standardised protocols (Boccardi et al, 2011; Entis et al, 2012) for use in manual segmentations but reviews in this area suggest they have not yet been adopted across the wider field

(Cox et al, 2014).

Other concerns such as the registration method used and atlas choice have implications as to the comparability of results across studies, as well potential accuracy issues especially on less than optimal scans (Dewey et al, 2010). Registering images involves taking two scans and overlapping them so that common features in that image are aligned. This process has many applications such as, looking for differences between scans of the same person at two time points, investigating common features in a group of individuals, or to align a set of images to a common space for image segmentation or voxel based analysis. In the latter example brain atlases are used to improve the accuracy of methods in identifying target brain regions or features. There are numerous registration methods and atlases available, and they are discussed in more detail in Chapters three and ten. Avants et al (2011) discuss the difficulty of finding the most effective registration method stemming from a lack of consistency in how an algorithm has been applied to a dataset. They suggest that parameter settings and optimisation strategies used to achieve the best possible registrations are not always reported in enough detail as to be reproducible. Atlas-based segmentation techniques have been shown to be the most successful of the automated methods (Babalola et al, 2009), with processes that utilise multiple atlases rather than a single atlas showing the highest level of agreement with the reference standard (Aribisala et al, 2013). Though promising, the difficulty of using atlases derived from a young healthy population, rather than an age appropriate population, has been highlighted (Aljabar et al, 2009). In these cases the range of age-related features such as atrophy and white matter lesions are seldom reflected in the atlas, resulting in poorer performance (Cabezas et al, 2011). All methods should be mindful of neuroanatomical information when being used and this becomes more crucial when trying to investigate what could be subtle differences in non-disease groups such as the elderly.

The aims of this thesis are several. Firstly, Chapter 2 outlines the general methods applied in this thesis. Chapters 3 and 4 examine the literature concerning the way in which application of methodologies can affect the accuracy of measurements in the hippocampus and frontal lobes. Secondly, in Chapters 5 to 9 the factors influencing global brain measures in older adults are explored through development of volumetric

neuroimaging methods. Thirdly, factors arising from the practical application of regional measures, with consideration given to the issues that emerged from the systematic reviews, are presented in Chapters 10 and 11. Finally, in Chapters 12 and 13, the resultant measurements from the experimental studies are used to see how neuroimaging variables are related to cognition in later life, and to explore how methods can be used to provide clearer information about these associations.

Chapter 2. General methods

In this chapter the methods employed as part of the brain MRI imaging protocol developed to study the Lothian Birth Cohort 1936 (LBC1936), a group of older adults in whom cognitive data are available in both youth and old age, will be described. This chapter will detail the methods that have not been developed as part of this thesis but contribute to analysis that features in Chapters 10, 12 and 13.

Although a section in Chapter 3 uses imaging and demographic data from a freely available database (http://www.oasis-brains.org), this thesis mainly uses imaging and cognitive data collected from participants of the Lothian Birth Cohort 1936 Study (Deary et al., 2007). This dataset will be described in more detail in context with the analyses appearing in Chapters 12 and 13, but is briefly introduced in the following sections.

2.1 LBC1936 cohort

The Lothian Birth Cohort (1936) was designed to find determining factors of non-pathological cognitive ageing particularly genetic components (Deary et al, 2012). The subjects that make up this cohort were born in 1936 and at the age of 11 they took a version of the Moray House Test No. 12 as part of the Scottish Mental Survey of 1947 (Deary et al, 2012). The Scottish Mental Survey of 1947 was taken in every school in Scotland on the 4th of June, with 70,805 children tested. The Moray House Test (No. 12) is a well-validated test of intelligence, which provides an estimate of a person's verbal reasoning ability prior to life-course influences such as ageing. The Scottish Council for Research in Education who administered the Survey retained the

data. The tests were rediscovered in ledgers by a group of researchers including Professors Lawrence Whalley and Ian Deary, who recognised the importance of the tests in investigating changes in life-course intelligence. The Community Health Index and media advertisements were used for recruitment and 3810 potential participants were identified. Of the 3686 people who were invited, 2318 responded and 1226 people were elgible and interested in taking part. Initially 1091 participants (548 men and 543 women), aged 70 years, retook the Moray House Test in 2006 along with other detailed cognitive and physiological tests. This first wave of testing began in 2006, with subsequent waves between 2007 and 2010 (Wave 2) and between 2011 and 2013 (Wave 3). Of the wave 1 participants, 866 (448 men and 418 women) received testing at wave 2 with the main reasons for attrition being death, chronic incapacity and permenant withdrawal.

The aim of the first wave of testing was to understand the genetic, physiological, social and environmental factors involved in normal ageing. Therefore extensive social, familial, nutritional and medical histories were obtained, as well as quality of life scores, blood samples, genetic analysis and physical activity information (Deary et al, 2007). Three years later the participants were invited to contribute in the second Wave of testing as before but now with additional brain MRI and neurovascular ultrasound imaging. This second wave of testing focused more on the mechanisms that are implicated in white matter damage and how this damage relates to cognitive decline in later life. The inclusion of brain MRI at this stage was intended to directly answer this question by assessing white matter integrity and cognitive function. Along side this analysis whole brain and regional grey matter integrity was also considered of importance, as they influence and are influenced by white matter damage. A third (complete Dec 2013) and fourth wave of testing will further contribute to this impressive dataset as will a brain tissue bank, which is being prospectively recruited for during these testing waves.

2.2 Subjects

Not all participants initially recruited (1091) at wave 1 returned for testing at wave 2 (866) with 700 (368 men and 332 women) of those recruited consenting to have an

MRI scan. A few factors such as claustrophobia, scanner faults and positioning difficulties influenced the participation/inclusion of some of those individuals that entered the scanner. This resulted in 672 (357 men and 315 women) people who completed all sequences necessary to measure brain volume. Full details of the participant demographics accompany each analysis of the data in the Chapters that follow, as they vary due to the inclusion of different variable included..

2.3 Cognitive testing

The cognitive data used in this thesis come from the second wave of cognitive testing, but the first to also include brain and ultrasound imaging. The cognitive domains tested in the LBC1936 Study were distilled into three main factors using Principal Component Analysis (PCA). The three factors are general cognitive ability (*g* factor), general memory (*g* memory) and general speed (*g* speed), (Laude et al, 2013). The subtests that these factors were derived from can be seen in Figures 2.1-2.3.

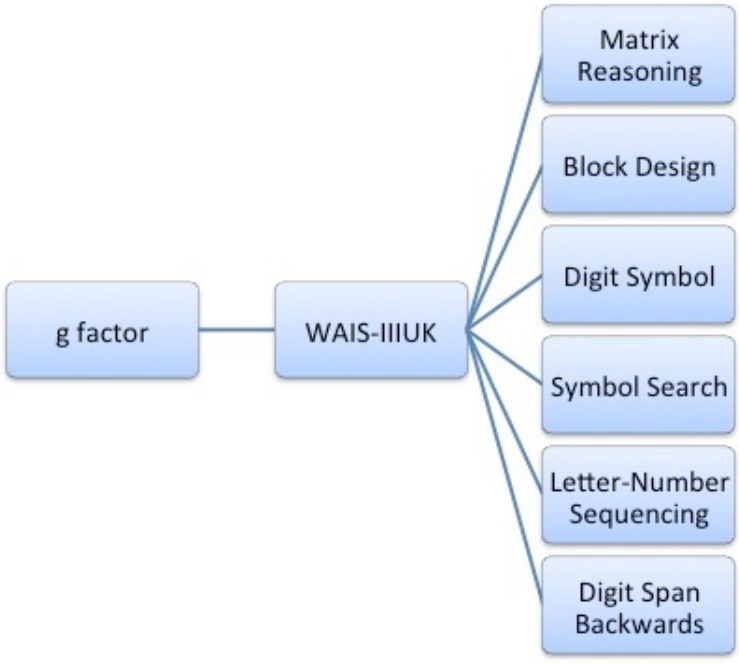


Figure 2.1: Shows the cognitive domains that, using Principle Component Analysis (PCA), make up general cognitive ability (*g* factor) from the entire cognitive test battery.

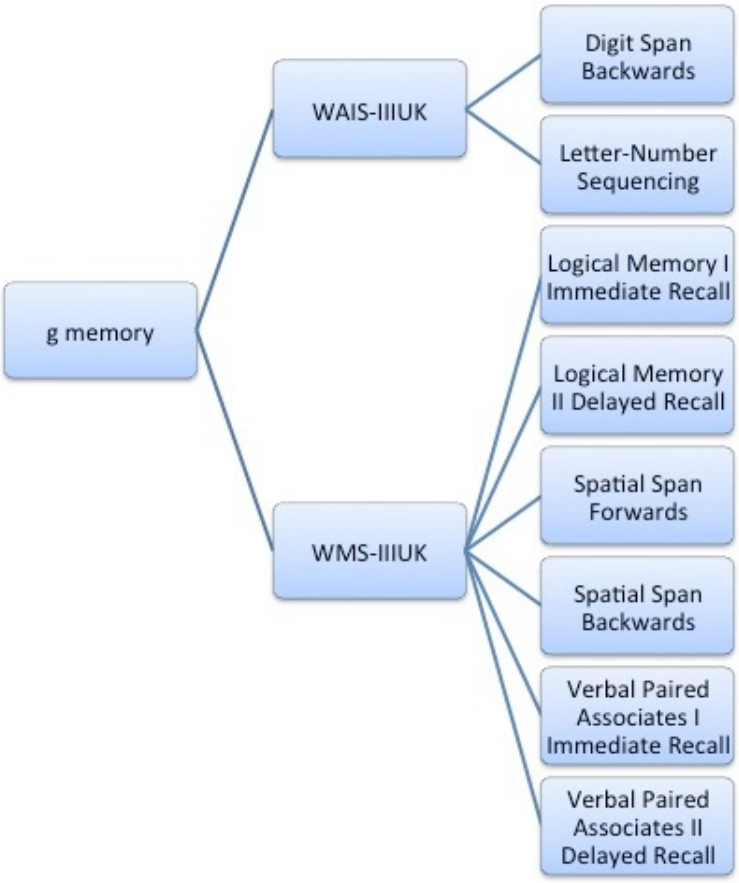


Figure 2.2: Shows the cognitive domains that, using PCA, make up general memory ability (*g*memory) from the entire cognitive test battery.

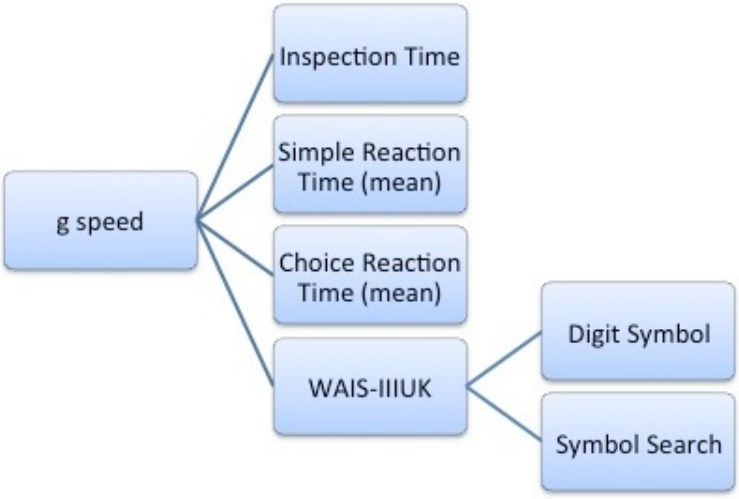


Figure 2.3: Shows the cognitive domains that, using PCA, make up general speed ability (*g* speed) from the entire cognitive test battery.

As part of the dementia screening process each participant completed the Mini-Mental State Examination (MMSE), which is a commonly used assessment tool where a score of 24 or less out of 30 indicates possible dementia (Folstein, Folstein and McHugh, 1975). This assessment was not intended to detect people with dementia but more to try and ensure that the data set comprised of participants free of cognitive impairment related to dementia.

2.4 Brain MRI acquisition protocol

As per Wardlaw et al. (2011), all MRI data were acquired in the Brain Research Imaging Centre, University of Edinburgh (http://www.bric.ed.ac.uk), using a GE Signa Horizon HDx 1.5T clinical scanner (General Electric, Milwaukee, WI, USA) equipped with a self-shielding gradient set (33 mT/m maximum gradient strength) and manufacturer supplied 8-channel phased-array head coil. The examination comprises the following whole brain sequences acquired with contiguous slice locations: T2-weighted (T2W), T2*-weighted (T2*W) and Fluid-Attenuated Inversion Recovery (FLAIR) axial scans. A high-resolution T1-weighted (T1W) volume sequence acquired in the coronal plane, axial T1W fast-spoiled gradient echo (FSPGR) sequences with 2 and 12^{o} flip angles for quantitative T_1-mapping, two standard spin echo sequences acquired with and without a magnetization transfer pulse applied 1 kHz from the water resonance frequency for Magnetisation Transfer-MRI (MT), and finally a Diffusion Tensor-MRI (DTI) protocol consisting of 7 T2W and sets of diffusion-weighted ($b = 1000$ s/mm^2) axial single-shot spin-echo echo-planar (EP) volumes acquired with diffusion gradients applied in 64 non-collinear directions (Jones et al, 2002). The acquisition takes approximately 70 minutes. The acquisition parameters for the all sequences, i.e. field-of-view, imaging matrix, slice thickness and location, were chosen to allow easier co-registration between them so that, for example, $\langle D \rangle$, Fractional Anisotropy (FA), Magnetisation Transfer Ratio (MTR) and T_1 biomarkers could be accurately measured in small structures such as the hippocampus assessed on a structural scan. To allow accurate measurement of intracranial volume (ICV), the structural MRI scans covered the complete intracranial contents, from above the skull vertex to the upper cervical spine below the foramen

magnum. All structural MRI data were examined by a consultant neuroradiologist and any medically relevant incidental findings, which occur in around 3% of individuals and not including age related WMHs (Morris et al, 2009), are identified and notified to a responsible clinician for further action. Finally, each sequence is converted from DICOM (http://dicom.nema.org) to NIfTI-1 (http://nifti.nimh.nih.gov/nifti-1) format for computational image processing.

2.5 Pre-processing steps and Image analysis

Once the data was acquired by the scanner, and the images converted from DICOM to NIfTI format, all sequences were registered to the T2-W sequence for each individual.

The LBC1936 image analysis protocol is extensive, employing various techniques to quantitatively assess structural images. The global volume measurements acquired as part of the image processing protocol were the intracranial contents, total CSF, total brain, ventricles, grey and white matter volumes. Some of the image processing pipeline is stepwise, meaning that as each step of the segmentation is completed it allows the next tissue along the pipeline to be measured. Mostly the process employs a semi-automatic tool that fuses two MRI sequences in red-green colour space to enhance visual differences between tissues (Valdes-Hernandez et al, 2010). The MCMxxxVI uses Minimum Variance Quantization to represent images in a reduced number of clusters mapped in the same colour space. This method is described in detail in (Valdes-Hernandez et al, 2010) but the practical steps used to derive the tissues used in the analysis that follows, will be described here.

To begin with structural scans (i.e. T1W, T2*W and FLAIR) were registered to the T2W sequence using an affine registration tool freely available from the FMRIB software library (Jenkinson et al, 2002).

Following this initial registration step, the intracranial contents were extracted. The research into the development of a standardized and validated protocol for this purpose is presented in detail in Chapter 6.

2.6 Normal-appearing white matter (NAWM) and cerebrospinal fluid (CSF).

We obtained two measures of CSF using the MCMxxxVI method; this is because the outcome of the method differs if different pairs of sequences are fused. For example, the fusion of T1W and T2W makes visible different tissue types than if T2*W and FLAIR sequences are fused. Segmentation of the CSF derived from using the fused T2*W and FLAIR sequences resulted in a measurement that included the veins, sinuses and meningeal tissues that are very close or attached to the brain tissue, together with the CSF, as they all share the same intensity range on each of the two fused images. As this is a step in the white matter lesion segmentation pipeline, it is considered an extra or surrogate measure of CSF. The process of obtaining a CSF measure that does not include the veins, sinuses and meningeal tissues is the same as that described below but the fused T1W_T2W sequences are used rather than the T2*W_FLAIR sequences.

To identify CSF and normal-appearing white matter, the T1W and T2W volumes were fused using Analyze 9.0^{TM}, a licensed image analysis software (http://www.analyzedirect.com/Analyze/). The registered T1W and T2W were loaded into Analyze and selected as input to the registration module of the software. This second registration is necessary to guarantee that both images are well aligned. One image was selected as the base image and mapped in red colour and the other as the match image and mapped in green. If necessary, the images were manually adjusted to overlay the match image on the base image (Figure 2.5) and produce a fused red-green image.

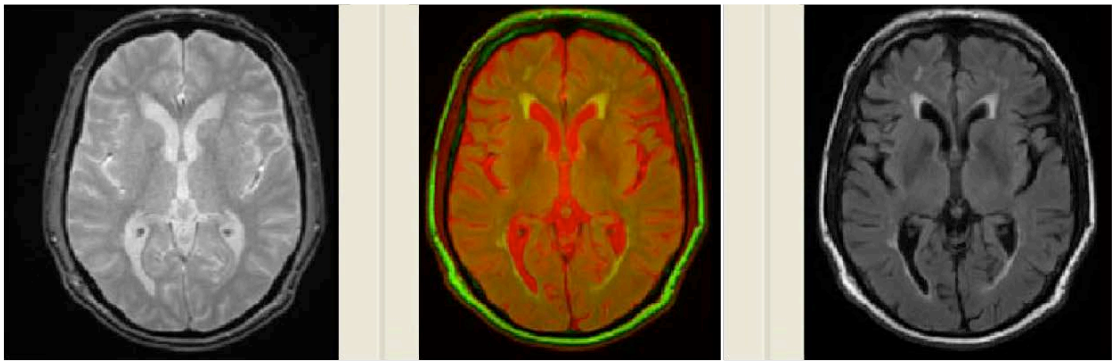


Figure 2.5: Shows a T2*W image fused with a FLAIR image in red-green (left-base, middle-fused, right-match).

MATLAB was then used to load the graphic user interface (GUI) MCMxxxVI. The GUI works by allowing the user to enter the file path of the fused images and the intracranial volume binary mask to segment the specified tissue. Once all the file paths are located and images loaded into the GUI the intensity of the images are adjusted to identify the specified tissue. The tissue included in the segmentation is coloured blue to distinguish it from the red-green background (Figure 2.6).

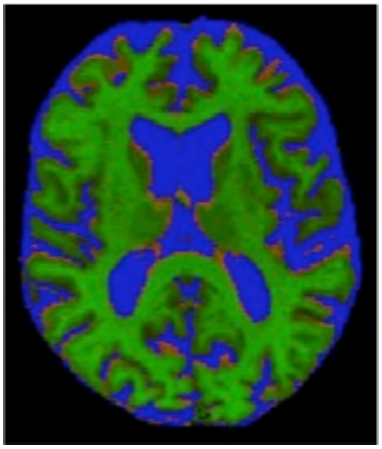


Figure 2.6: A fused image showing the CSF to be segmented in blue.

The flowchart in Figure 2.7 describes the sequence of steps taken to produce the CSF and NAWM masks. Once the CSF is obtained, the next step was to use the binary mask resultant from subtracting the CSF from the ICV and re-input it to the GUI in place of the ICV used in the previous process. Once the intensity values have been adjusted for extracting the NAWM, false positives were removed where erroneous tissue was included.

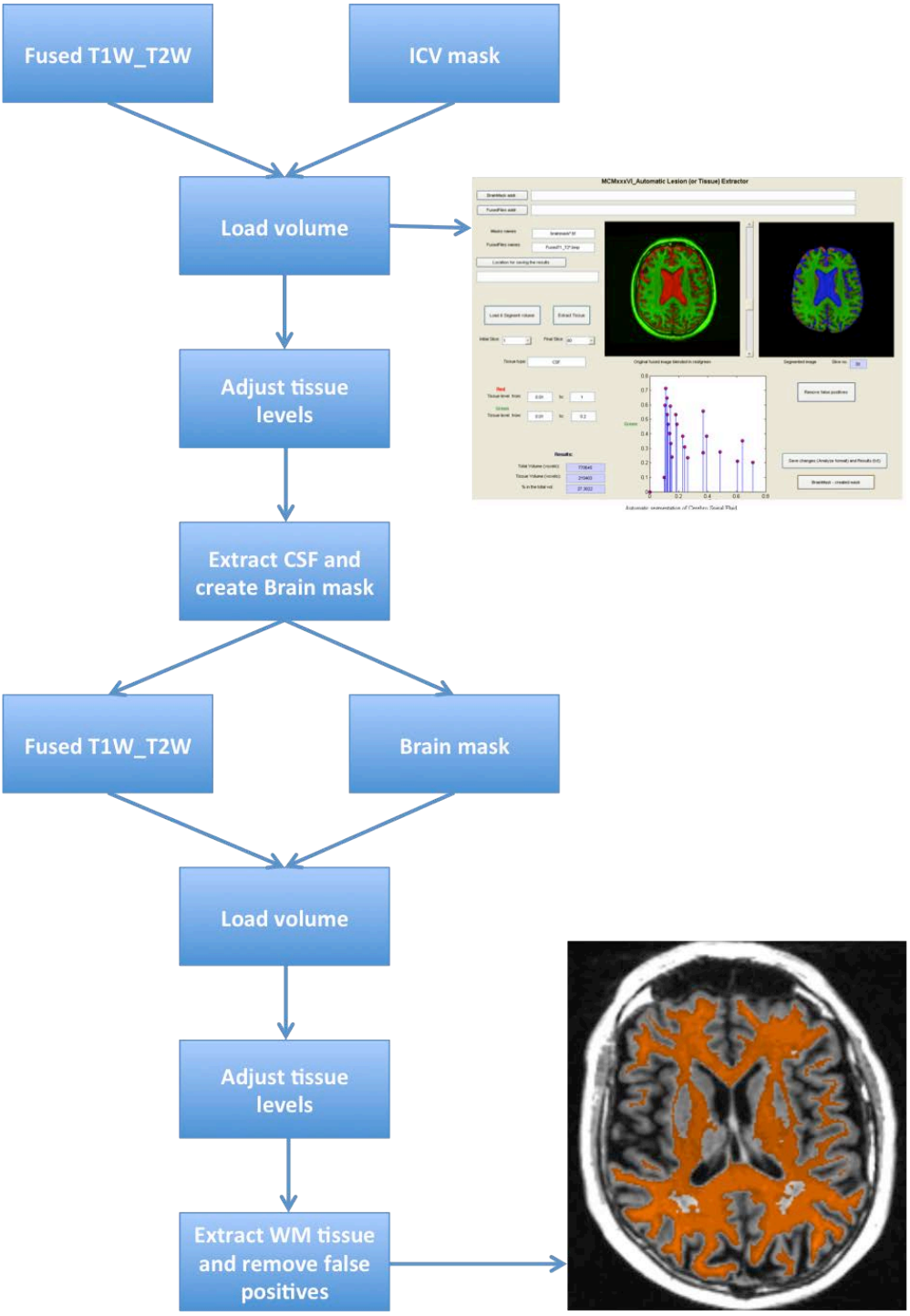


Figure 2.7: Flowchart depicting the steps taken in the MCMxxxVI method to achieve CSF and NAWM segmentations.

2.7 Grey matter volume

The intensity range of grey matter overlaps with that of the white matter hyperintensities (WMH) in almost all sequences. Even in FLAIR sequences, preferred for WMH identification, subtle WMH exhibit the same intensities as some cortical regions. That is why grey matter was derived by subtracting the CSF, the NAWM and WMH masks previously segmented using the MCMxxxVI method from the ICV. Routine visual assessment of the grey matter masks revealed inaccuracies due to inclusion of structures and tissues that are not grey matter and do not belong to any of the other tissue classes previously mentioned, specifically highlighted were the inclusion of the choroid plexus and septum. To investigate the influence in total grey matter volume that inclusion of these erroneous structures represented, manual editing was performed to remove the choroid plexus and septum from the grey matter masks in a randomly selected sample of 40 cases from the LBC1936. Mean grey matter volume (mm^3) was 646501 (± 50170) and after removal of the erroneous structures the mean volume (mm^3) was 644753.4 (± 50181), this only represented a mean difference of 0.27%, with the maximum difference being 0.5% between the two measurements.

Further visual assessment of the MCMxxxVI masks revealed that other areas also were incorrectly classified as grey matter such as the clivus, internal carotid and choroidal arteries. Although the MCMxxxVI method improves the sensitivity by which segmentation of images relying on intensity information can be performed, it does not include shape information in the segmentation process. This aspect of segmentation is dependent on the anatomical knowledge of the user, as is the case with manual methods though less so for this method. The limitation of using only intensity information to segment tissues is that MRI images are limited in the range by which they can display tissues. Distinct structures can appear very similar in intensity, therefore leading to misclassification of tissues when only intensity information is used. It is therefore necessary to manually edit the masks using the ‘Remove false positives’ feature provided in the GUI. This is especially the case when images are affected by motion artifact, tissue abnormalities such as cysts or meningeal hyperintensities. The strength of using a semi-automated technique is that there is some degree of user input, meaning that difficult images such as those described can be corrected where possible.

To establish how comparable the MCMxxxVI method was with other methods of obtaining grey matter segmentation the same 40 subjects were processed using the free software tool FSL_FAST (FMRIB's Automated Segmentation Tool). The flow chart (Figure 2.8) shows the steps in the process FAST uses to produce segmentations.

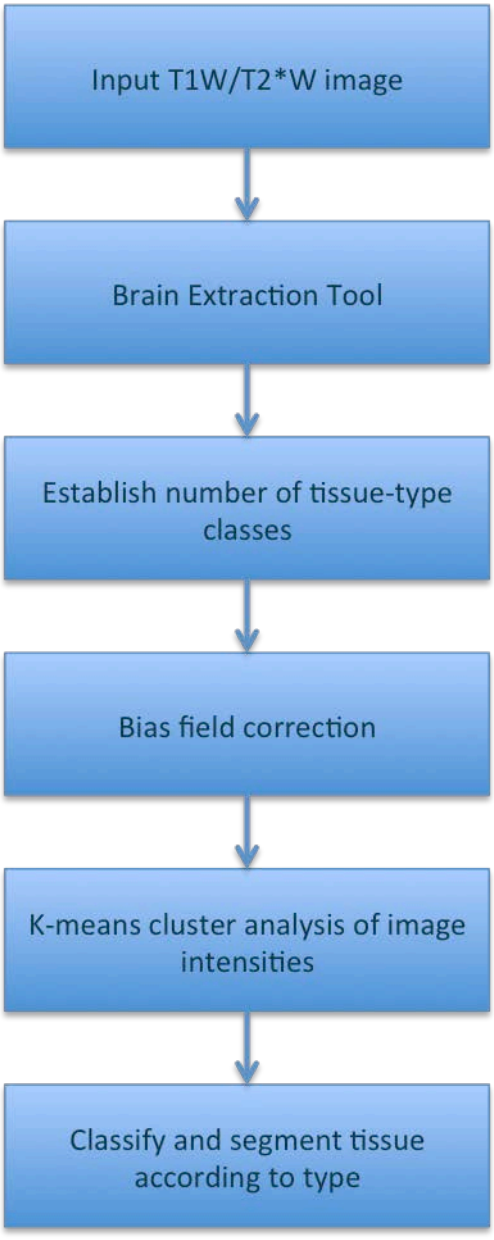


Figure 2.8: Flowchart showing the FSL_FAST processing steps for CSF, grey and white matter segmentation.

Visual assessment of the FSL_FAST results revealed that there were several areas where tissue was misclassified as grey matter. Table 2.1 shows the tissues wrongly included by FSL_FAST and MCMxxxVI. Though no segmentation was marked as a

fail, 38/40 FSL_FAST segmentations had one erroneous structure noted with 23/30 for the MCMxxxVI segmentations.

FSL_FAST	MCMxxxVI
Clivus	Clivus
Internal carotid artery	Internal carotid artery
Trigeminal nerve	-
Periventricular	Periventricular
Optic nerve	-
Septum	Septum
Pons	Pons
Skull	-
Meninges	-
WMLs	-
Mineral deposits	Mineral deposits
Nasal concha	-

Table 2.1: Indicates the visually identified incorrectly segmented tissues for both the FSL_FAST and MCMxxxVI methods.

The FSL_FAST segmentation did not include the choroid plexus in any of the 40 cases and in two cases a small section of the middle frontal gyrus and/or superior frontal sulcus was missed. This also occurred in one of the MCMxxxVI segmentations. MCMxxxVI differed from FSL_FAST in the inclusion of the choroid plexus but generally performed better than FSL_FAST, as fewer tissues were wrongly classified.

Bland Altman analysis performed on the number of voxels included in the grey matter segmentations from FSL_FAST and the MCMxxxVI methods shows a mean difference of -16.45 with a lower confidence limit of 32.01 and upper confidence limit of -0.90 (Figure 2.9).

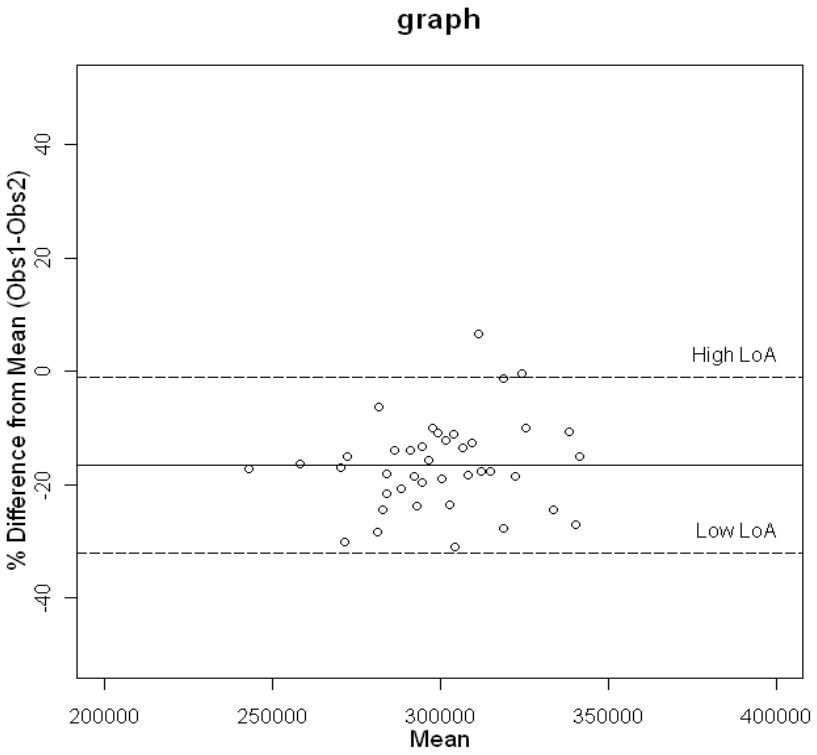


Figure 2.9: Bland-Altman plot comparing percentage mean difference in measurements of grey matter voxels performed by FSL_FAST and MCMxxxVI.

Comparison between FSL_FAST and the MCMxxxVI segmentation, with removal of the septum and choroid plexus, shows a mean difference of -15.74 with a lower confidence limit of 31.68 and an upper confidence limit of -0.19 (Figure 2.10).

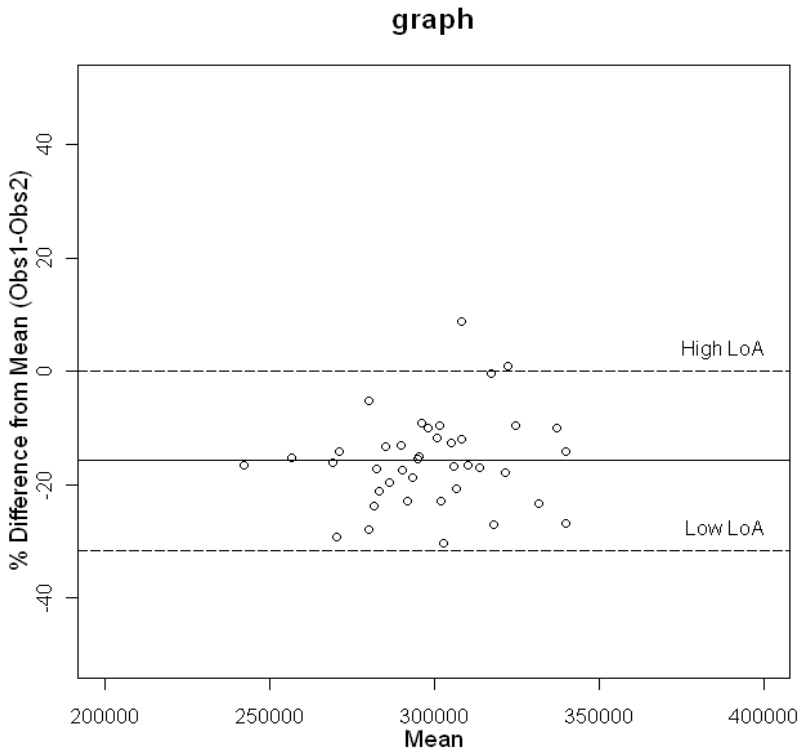


Figure 2.10: Bland-Altman plot comparing percentage mean difference of grey matter voxels by FSL_FAST and MCMxxxVI after manual removal of the choroid plexus and septum.

It was concluded that removal of the choroid plexus from the grey matter masks produced using MCMxxxVI would be an adequate solution to the issue of erroneous tissues. This is achieved by subtracting the contents of the lateral ventricles, where the choroid plexus is located, from the grey matter masks using the masks manually obtained from performing ventricular measurement (described below). Further manual editing where necessary should also be performed.

2.8 Whole brain volume

Subtracting CSF from ICV using the image calculator in Analyze produced whole brain volume. The CSF mask, previously obtained from the MCMxxxVI module by fusing T2*W and FLAIR scans, is subtracted from the ICV mask thus removing all non-brain tissues and CSF within the intracranial vault. This results in a mask that only contains brain tissue, so is used as a measure of whole brain volume.

2.9 Ventricular volume

The ventricles were manually measured using Analyze 9.0 from the original 3D T1W volume acquired in coronal orientation, without being registered to any other sequence. This is to preserve contrast and anatomical information lost when registered images perform interpolation calculations to estimate unmeasured data points while altering the matrix size of the image. The T1W images were thresholded to produce better contrast between grey matter and CSF and to consistently account for partial volume effects across the whole volume. To threshold the image, the ROI module in Analyze was used to place two 3x3mm ROIs, one in the CSF of either lateral ventricle and one in the temporal grey matter (Figure 2.11).

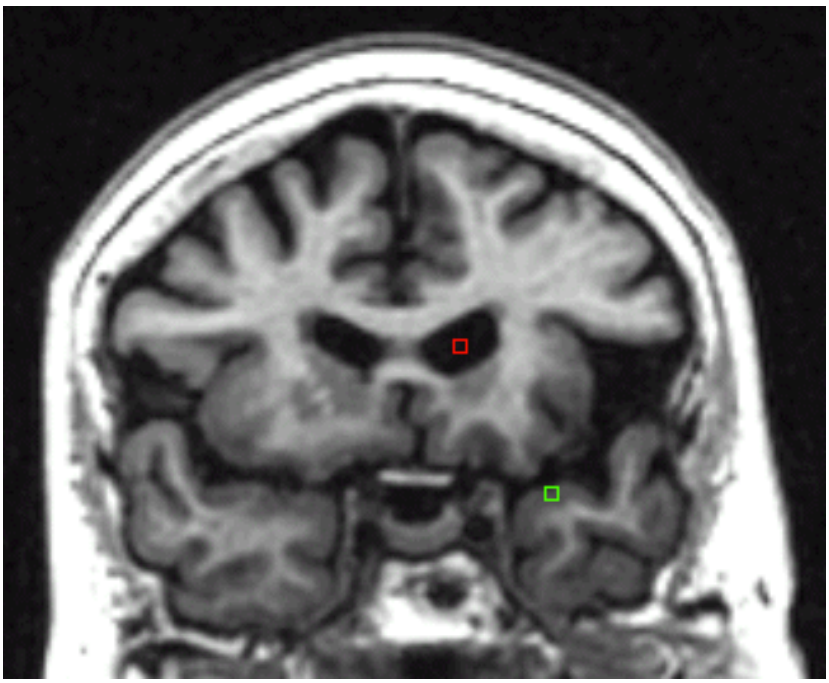


Figure 2.11: Shows the placement of two ROI in the ventricle (red) and the temporal lobe grey matter (green), to obtain a mean intensity value for thresholding the image.

The two ROIs were then copied forward one slice, their position checked and the mean sum of all four ROIs calculated and recorded. This threshold was then applied to the image using the Object Extractor module, which uses an algorithm to extract brain tissue from the skull according to the previously calculated threshold level. The

extracted image was then saved and used to perform the ventricular segmentation.
The thresholded image was loaded into Analyze and the ROI module selected. The first slice of the lateral ventricles was chosen as the anterior slice in which either lateral ventricle appears in deep WM (Figure 2.12); left and right do not always appear at the same time due to variations in head position in the scanner, asymmetry of the hemispheres and differences in ventricular size.

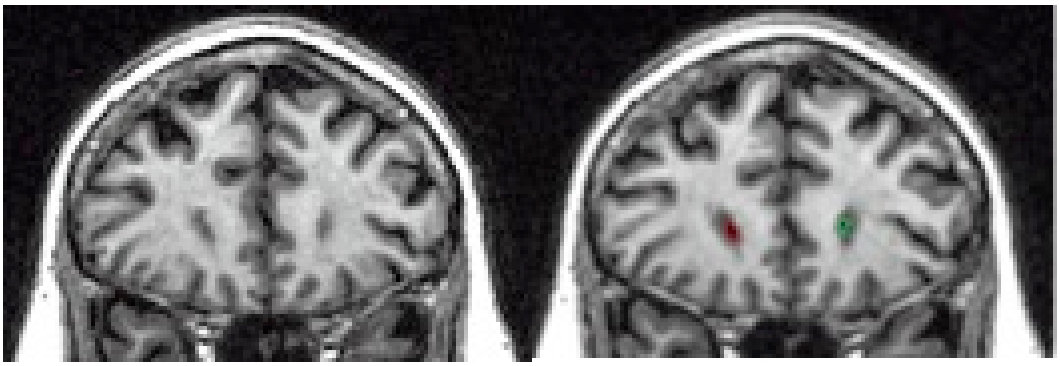


Figure 2.12: Shows the slice preceding and the slice in which the first appearance of the lateral ventricles in the right (red) and left (green) frontal lobes occurs.

A ROI was placed in each lateral ventricle (right and left were separately measured) and the edge of the structure automatically detected. The lateral ventricles were measured in this way working posteriorly through each slice, remembering to identify the temporal horn of the lateral ventricle and include it in the whole lateral ventricle measurement (Figure 2.13).

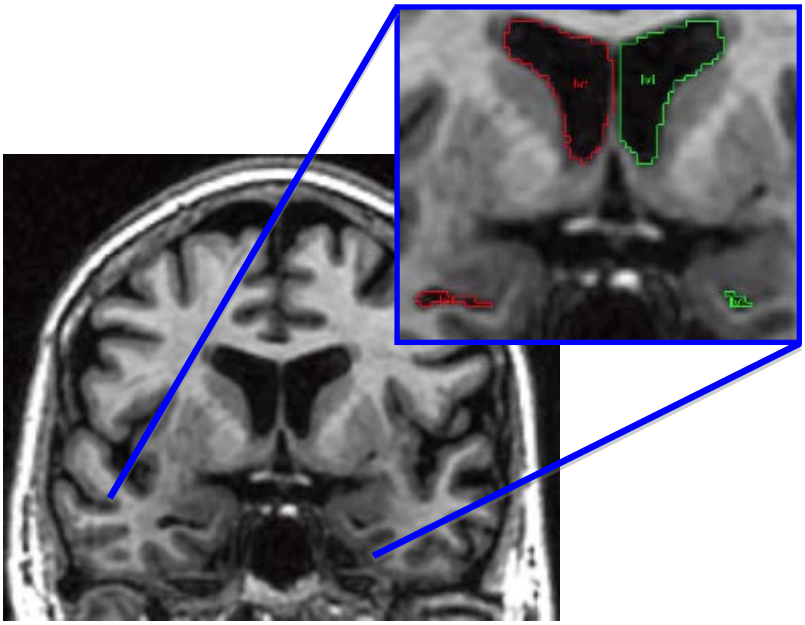


Figure 2.13: Shows the appearance of the horn of the lateral ventricles in the right (red) and left (green) temporal lobes.

The temporal horn of the lateral ventricle is positioned superior to the parahippocampal gyrus and medial to the temporal stem; CSF present superior to the hippocampus is part of the uncal recess of the temporal horn and should therefore be included in this measurement (Figure 2.14).

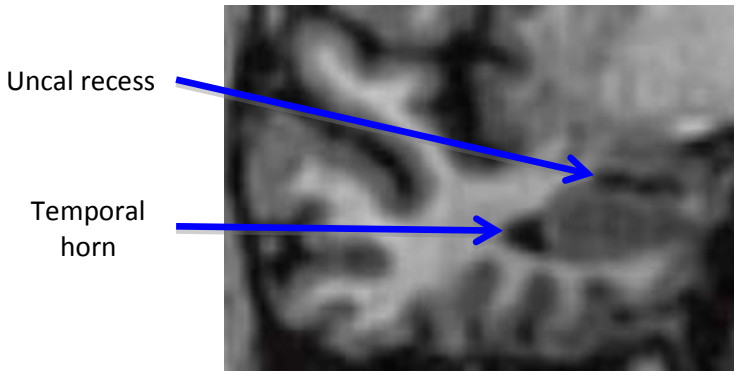


Figure 2.14: Shows the appearance of the uncal recess of the temporal horn superior to the hippocampus.

Occasionally the septum pallucidum appears incomplete and the boundary must be re-established to ensure each ventricle can be measured separately. To do this a straight line is drawn between the two hemispheres (Figure 2.15).

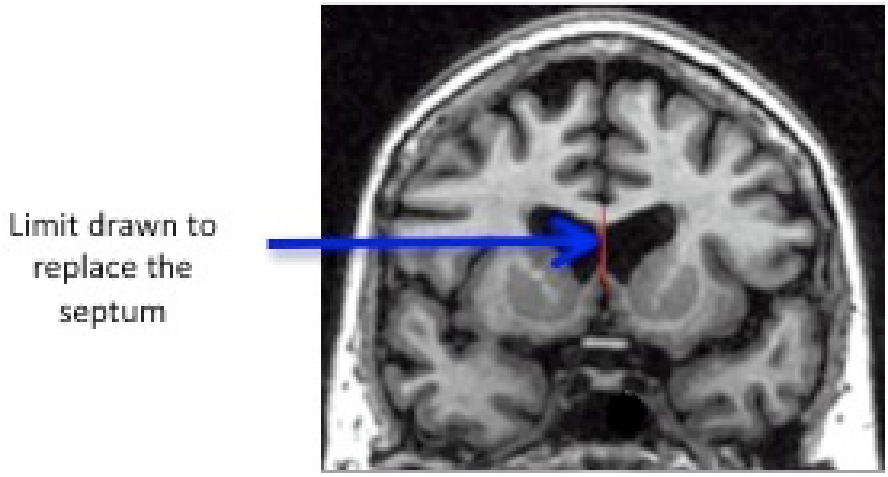


Figure 2.15: Shows the line drawn to re-establish the septum between the left and right lateral ventricles.

The third ventricle appears inferiorly to the lateral ventricles, between the two hemispheres. The first slice in which the third ventricle was measured was the point where the optic tracts converge to form the optic chiasm (Figure 2.16).

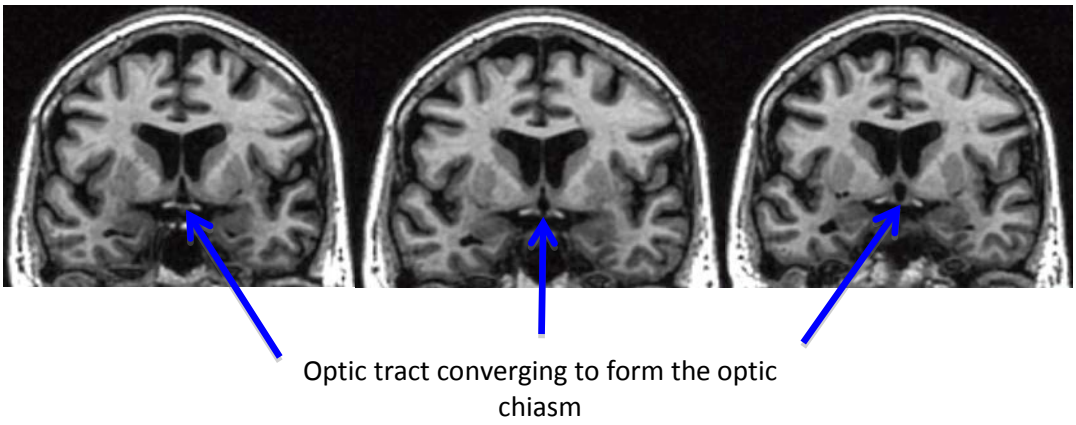


Figure 2.16: Shows the first slice (right hand image) in which the third ventricle is measured once the optic tracts converge to form the optic chiasm.

The interthalamic adhesion divides the third ventricle so it is important to ensure that superior and inferior portions are measured (Figure 2.17). The interpeduncular cistern or the cerebral aqueduct was not included in the third ventricle measurement (Figure 2.17).

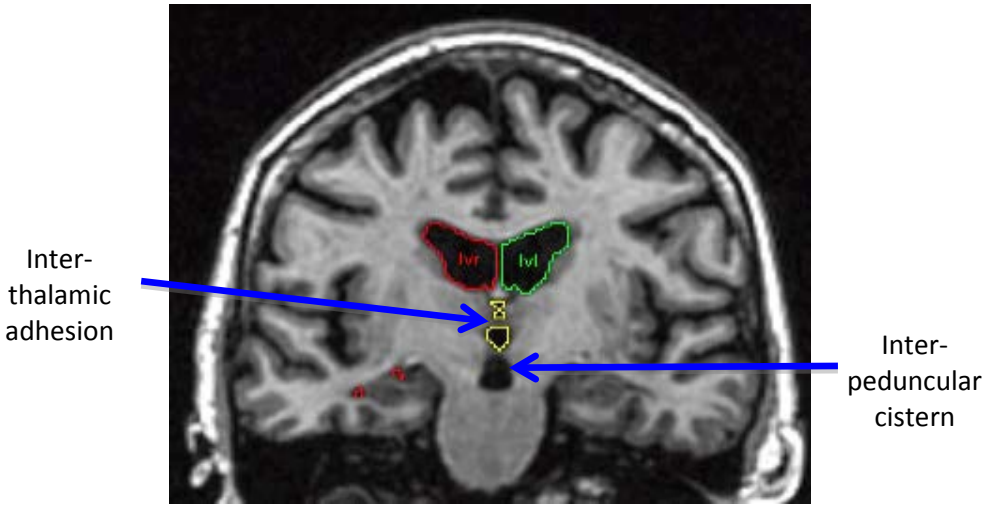


Figure 2.17: Indicates the division of the third ventricle by the interthalamic adhesion and the location of the interpeduncular cistern.

The last slice was that prior to the ventricle opening into the box-shaped cistern of the transverse cerebral fissure (Figure 2.18).

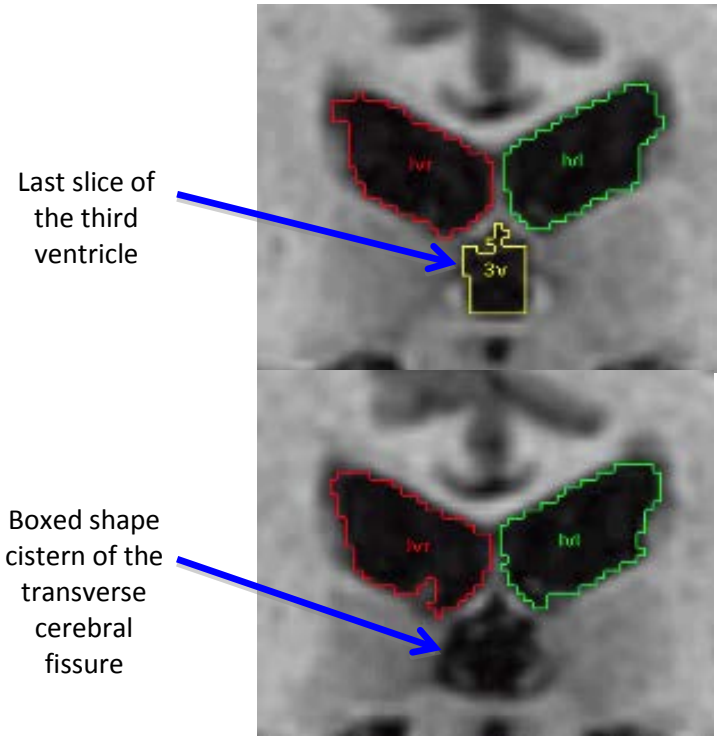


Figure 2.18: Shows the last slice in which the third ventricle is measured before the appearance of the transverse cerebral fissure.

The fourth ventricle is located between the cerebellar hemispheres, superior to the vermis and tends to only appear in a few slices. The first slice of the fourth ventricle appears to form a hexagonal shape and is enclosed by surrounding cerebellar tissue; this structure was measured through the slices until it was no longer visible (Figure 2.19).

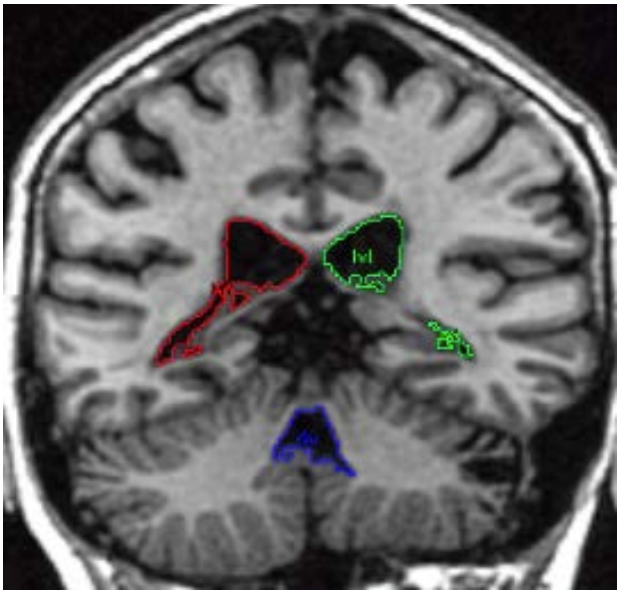


Figure 2.19: Shows the location of the fourth ventricle (blue) between the cerebellar hemispheres.

The complete dataset for this analysis therefore consists of masks, voxel and volume measurements (mm3) for ICV, total CSF, surrogate CSF, whole brain, ventricular, grey and white matter.

Section 2: Systematic reviews

In order to ascertain what methods have already been employed to quantify brain regions of the hippocampus and frontal lobe, reviews of current methods were undertaken. The task of segmenting the hippocampus in a large healthy older cohort (LBC1936) needed to be tackled using automated methods, because the time taken to manually delineate 1400 hippocampi would be unfeasible. Common age-related features such as brain tissue atrophy, asymmetry and prevalence of white matter

hyperintensities can cause difficulties when analysing MR images of elderly participants. A review of automated hippocampal segmentation methods, specifically those that have been applied and adequately validated in healthy older brains, was conducted.

For the frontal lobe review, the main issue being considered was less to do with the application of automated methods, though many of the same concerns that apply to the hippocampus are also pertinent to the frontal lobe, but more to do with anatomical definition of the frontal region. The variety with which researchers assign tissue to the frontal lobe when manually segmenting the brain is troublesome. An agreed upon protocol for exactly what boundaries should be used to accurately delineate the frontal lobe does not as yet exist but would be immensely helpful in reducing variation in findings across studies. In order to understand what the vast literature on the involvement of the frontal lobe in ageing-related decline contributes, it is essential to know what researchers are measuring as frontal lobe tissue. Consequently a review of frontal lobe posterior boundaries was conducted, as part of a larger review of frontal lobe structure function relationship (Cox et al, 2014), to determine how much variation in segmentation protocols exists.

Chapter 3. Systematic review of automated hippocampal segmentation methods in healthy elderly population

3.1 Introduction

The hippocampus is of interest to researchers because it has been shown to be affected by age-related decline and it is implicated in diseases common in old age such as dementia (Baron et al, 2001). It has also been highlighted as a key biomarker in detecting preclinical signs of Alzheimer's disease, due to the relationship between decline in memory performance and smaller hippocampal volume in old age (Backman et al, 1997). Along with ventricular volume (Resnick et al, 2000), it has been suggested that the rate of volume loss in the hippocampus is predictive of the shift from normal cognition, to mild cognitive impairment (MCI) then into dementia (Mungas et al, 2005). Because of these implications there are many studies that

provide evidence to support hippocampal atrophy and poor memory performance in clinical groups (Fischl et al, 2002; Rombouts et al, 2000). However, it is important to understand what the relationship is between the hippocampus and memory function in normal ageing. A meta-analysis looking at associations between hippocampal volume and memory in healthy people found that in older adults results were too variable as to draw a definite conclusion (van Petten, 2004). They did report that a weak positive association was found but that the clearest finding was the considerable degree of variability in this association that seems to occur in later life. It is also the case that null associations between memory performance and hippocampal volume are seldom reported when the healthy older adults form the control group in a patient focussed study. This suggests a potential under reporting of null results therefore making it difficult to determine how much variation in the literature there actually is.

Another possible explanation for why discrepancies exist between published reports could be due to differences that arise from the variety of methods being applied to hippocampal volume measurement. The reference standard method of measuring hippocampal volume is manual delineation; however, it is time consuming, subject to observer variation and not always practical for large studies (Barnes et al, 2004). Consequently, numerous computational automated segmentation algorithms and software programs have been developed to assess subregional brain volumes including the hippocampus. The advantages of using automated techniques are that they should reduce rater variability, perform segmentations with minimum user input and produce precise outputs quickly (Khan et al, 2008). However, assessing subtle changes in the volume of the hippocampus, such as those that occur during healthy ageing, is more difficult due to the degree of sensitivity required. Older individuals brains significantly differ from young people in that they are more asymmetric, suffer from pathology such as atrophy and white matter lesions, and white and grey matter can be difficult to differentiate. Automated methods that utilise shape, contrast and/or intensity information to perform segmentations may struggle to combat these problems leading to errors.

In order to ascertain the effectiveness of automated methods in obtaining whole hippocampal volume measurements, we conducted a review of the published literature

to determine how extensively they have been validated against a reference standard and specifically their application in healthy older people.

3.2 Methods

3.2.1 Search strategy

We searched MEDLINE (1966 to 2012), ISI Web of Knowledge (1966 to Oct/2013) and PubMed electronic databases. We used the subject terms 'Automated', 'Hippocamp$', 'Segmentation', 'Measure$', 'Volume', 'Brain' with the Boolean operator AND in either the title or abstract, and the limits were set for humans only.

3.2.2 Inclusion/exclusion criteria

We included papers that used a fully automatic method to segment the hippocampus on MRI where healthy older adults age 55+ years of age were either the target group being studied or used as controls. We defined 'automated' as those methods that require minimal user input and predominantly use computational processes to derive volume measurements.

Excluded studies were those that were reviews, where only abstracts were available, where insufficient demographic data was provided, duplicate publications using the same dataset and where appropriate validation of the automated method was not stated.

3.2.3 Data extraction and synthesis

We extracted data on total number, gender, age and type of participants; number of participants in each category (i.e. Healthy older adults, Alzheimer's disease, etc.); scanner field strength (Tesla); voxel size and or slice thickness (mm); details of the automated method used; whether the automated method was validated against a reference standard; what statistical analysis or visual assessment was used to assess the segmentation; if the segmentation was corrected for variation in head size between participants (e.g. using intracranial volume); and any information on discrepancies between the hippocampi delineated by automated and reference methods. We

considered the reference standard measurements as those obtained using manual delineation within the same dataset that the automated method was applied to. We tabulated the results and performed descriptive statistics, and intended to perform a meta-analysis, but insufficient data were available.

3.3 Results

3.3.1 Study search and evaluation

The search strategy identified 194 publications. One hundred and eight publications were excluded as they did not fulfil the inclusion criteria, 57 were excluded due to all the participants being under 55 years old. This left 29 papers that described automated segmentation of the hippocampus, which included various patient groups but specifically healthy older adults (Table 3.1). Twelve of the 29 papers did not validate the automated method against a reference standard, so these papers were excluded. The Twelve papers excluded due to lack of validation against a reference standard in the same cohort, used both voxel-based and volumetric methods of segmentation (Table 3.1). Three used SPM99 (Baron et al, 2001; Frisoni et al, 2002; Rombouts et al, 2000), a version of a freely downloadable statistical parametric mapping software. Five papers (Fischl et al, 2002; Messina et al, 2011; Mouiha et al, 2011; Westman et al, 2011; Ystad et al, 2009) used Freesurfer and four others (Erickson et al, 2009; Erickson et al, 2010; Lim et al, 2012b) used FSL_FIRST, both of which are freely available for download. One study developed in-house software to segment the hippocampus automatically (Chupin et al, 2008).

Ref	**N**	**Sample type**	**Mean age (S.D.)**	**Age range**	**Automated method**	**Validated in same data set against reference method**
Baron et al. 2001	19	AD	73(5)	-	SPM99	N
	16	HC	66(8)	-		
Bishop et al. 2011	9	HC	-	65-83	FSL/MatLab	Y
	8	AD	-	65-83	FMASH	
	16	HC	-	20-58		
	16	BPD	-	20-58		
Brewer et al.	20	HC	77	-	NeuroQuant	Y

2009	20	AD	77	-		
Carmichael et al. 2005	20 19 15	AD MCI HC	- - -	- - -	Atlas based segmentation methods	Y
Chupin, et al 2008	25 24 25	AD aMCI HC	73 (6) 74 (8) 64 (8)	-	Minimisation of energy function	N
Csernansky et al. 2000	18 18 15	HOA AD HYA	74.2(5.2) 74(4.8) 30.9(9)	-	High-dimensional bran mapping (HDBM)	Y
Erickson et al. 2009	165	HOA	66.5	59-81	FSL_FIRST	N
Erickson et al. 2010	142	HOA	66.5	59-81	FSL_FIRST	N
Fischl et al. 2002	25 21 71 17	HC AD AD AD	72 74 72 67	-	FreeSurfer	N
Frisoni et al. 2002	26 26	AD HC	74 (9) 69 (8)	-	SPM99	N
Gosche et al. 2001	56	Nun Study		75-106	Knowledge-Guided MRI Analysis Program (KGMAP)	Y
Grabner et al. 2006	118	HOA		-	Atlas based segmentation method	Y
Hsu et al. 2002	20 20 20	HOA MCI AD	74 (6.2)	-	Medtronic Surgical Navigation Technologies	Y
Khan et al. 2011	18 19	HC MCI	- -	72-84 72-84	Freesurfer Large deformation diffeomorphic metric mapping	Y
Leung et al. 2010	10 10 10	HC MCI AD	78.6 (5.4) 75.3 (8.8) 77.2 (6.8)	-	Multiple-atlas propagation and segmentation (MAPS)	Y
Lim et al. 2012a	30 30	HC LLD	72.4 (4.5) 73.7 (6.4)		Freesurfer	Y
Lim et al.	50	HC	70.7 (4.4)		FSL_FIRST	N

2012b	51	AD	73.5 (4.8)			
Messina et al. 2011	46 72 32 15	HC PD PSP MSA-P	66.8 (6.7) 63.8 (9.0) 70.6 (5.3) 64.3 (4.3)		Freesurfer	N
Morra et al. 2010	20 20	AD HC		-	AdaBoost	Y
Mouiha et al. 2011	198 331 154	HC MCI AD	76.1 (5.1) 75.0 (7.2) 75.2 (7.5)		Freesurfer Surgical Navigation Technologies	N
Rombouts et al. 2000	7 7	AD HC	65 57	53-75 51-68	SPM99	N
Sabuncu et al. 2010	10 9 9 11	HC HC HC AD		<30 30-60 60+ -	FreeSurfer	Y
Sánchez-Benavides et al. 2010	41 23 25	HC MCI AD	68.5(8) 73.4 (7) 75.9(6.1)	-	FreeSurfer	Y
Shen et al. 2009	38 39 37 11	HOA EOA MCI AD	70.6(5.2) 72.8(6.1) 72.7(.1) 75.6(6.8)	-	FreeSurfer V4	Y
Shen et al. 2012	103 68 46	HC aMCI AD	70.8 (5.3) 76.8 (5.8) 79.1 (7.1)		Freesurfer Multiple templates	Y
van Der Lijn et al. 2008	518	HOA		60-90	Minimisation of energy function	Y
Westman et al. 2011	36 30	HC AD	76.5 (5.1) 77.3 (5.0)		Freesurfer	Y
Wolz et al. 2010	222 392 182	HC MCI AD	76(5.08) 74.68(7.39) 75.8 (7.63)	-	Learning embeddings for atlas propagation (LEAP)	Y
Ystad et al. 2009	170	HOA	62.2	46-77	FreeSurfer	N

Table 3.1: Total number of participants, population type, mean age, standard deviation, age range and whether the automated methods were validated against a reference standard from the 29 studies that used automatic methods to measure the hippocampus in patients and healthy older adults.

Abbreviations of sample type for tables 1, 2 and 3:
AD – Alzheimer's Disease; HC – Healthy Controls; BPD – Bipolar Disorder; aMCI – amnestic/Mild Cognitive Impairment; HOA – Healthy Older Adults; HYA – Healthy Young Adults; LLD – Late-Life Depression; PD – Parkinson's Disease; PSP – Progressive Supranuclear Palsy; MSA-P – Multiple System Atrophy (Parkinsonian); EOA – Euthymic Older Adults.

This left 17 papers that compared the automated segmentation against manually derived segmentation in the same participants (Table 3.2). Therefore these 17 studies are the focus of this review.

Study reference	**N**	**Sample type**	**Age: mean, (std) range**	**Gender: M/F**	**Scanner field strength: tesla**	**Voxel size/slice thickness (mm)**	**Volumes corrected for ICV**
Bishop et al. 2011	9 8 16 16	MC AD MC BPD	65-83 65-83 20-58 20-58	- - - -	-	0.94x0.94x1.5 0.8x0.8x0.8	No
Brewer et al. 2009	20 20	MC CDR 0.5	77 77	- -	1.5	-	Yes
Carmichael et al. 2005	20 19 15	AD MCI MC	- - -	-	-	1.5mm	No
Csernansky et al. 2000	18 18 15	HOA AD HYA	74.2, (5.2) 74, (4.8) 30.9, (9)	9/9 9/9 11/4	1.5	Isotropic 1mm	Yes
Gosche et al. 2001	56	Nun study	75-106	0/56	1.5	2mm	No
Grabner et al. 2006	118	HOA	-	-	1.5	Isotropic 1mm	No
Hsu et al. 2002	20 20 20	MC CI AD	74, (6.2) 74.2, (1.3) 74.5, (6.2)	10/10 16/4 10/10	1.5	3mm	Yes
Khan et al. 2011	18 19	MC MCI	72-84	17/20	3	Isotropic 1mm	No

Leung et al. 2010	10	MC	78.6, (5.4)	6/4	1.5	1.2mm	No
	10	MCI	75.3, (8.8)	7/3			
	10	AD	77.2, (6.8)	7/3			
Lim et al. 2012a	30	MC	72.4, (4.5)	14/16	3	Isotropic 1mm	Yes
	30	LLD	73.7, (6.4)	15/15			
Morra et al. 2010	20	AD	-	-	-	-	No
	20	MC	-				
Sabuncu et al. 2010	10	MC	<30	-	1.5	Isotropic 1mm	No
	9	MC	30-60				
	9	MC	60+				
	11	AD	-				
Sánchez-Benavides et al. 2010	41	MC	68.5, (8)	46/54%	1.5	1.5mm	No
	23	MCI	73.4, (7)	52/48%			
	25	AD	75.9, (6.1)	32/68%			
Shen et al. 2009	39	MC	-	-	1.5	Isotropic 1mm	Yes
	39	EOA	-				
	37	MCI	-				
	11	AD	-				
Shen et al. 2012	103	MC	70.8, (5.3)	25/78	-	-	Yes
	68	aMCI	76.8, (5.8)	33/35			
	46	AD	79.1, (7.1)	19/27			
van Der Lijn et al. 2008	20	HOA	74.6, (8.2)	55/45%	1.5	1.25mm	Yes
	498	HOA	73.5, (7.9)	50/50%			
Wolz et al. 2010	222	MC	76, (5.1)	106/216	1.5	1.2mm	No
	392	MCI	74.7, (7.4)	138/254			
	182	AD	75.8, (7.6)	91/91			

Table 3.2: Shows the participant demographics; n, population type, age and gender of the 17 included papers as well as the scanner field strength, voxel size, slice thickness and software used for the automated segmentations.

3.3.2 Included populations

The 17 included papers had a total of 2377 participants, aged between 30-106 years. Twelve studies had fewer than 100 participants, whilst five studies had over 100 participants (Table 3.2). Four papers provided the age range of the participants, nine papers gave the mean age and four papers reported that their participants were elderly, but did not provide the age of their participants. The healthy older adults represented 52% of the data from the included studies, 25% had Mild Cognitive Impairment, 16% Alzheimer's disease, with the remaining 7% taken up by other groups (Figure 3.1).

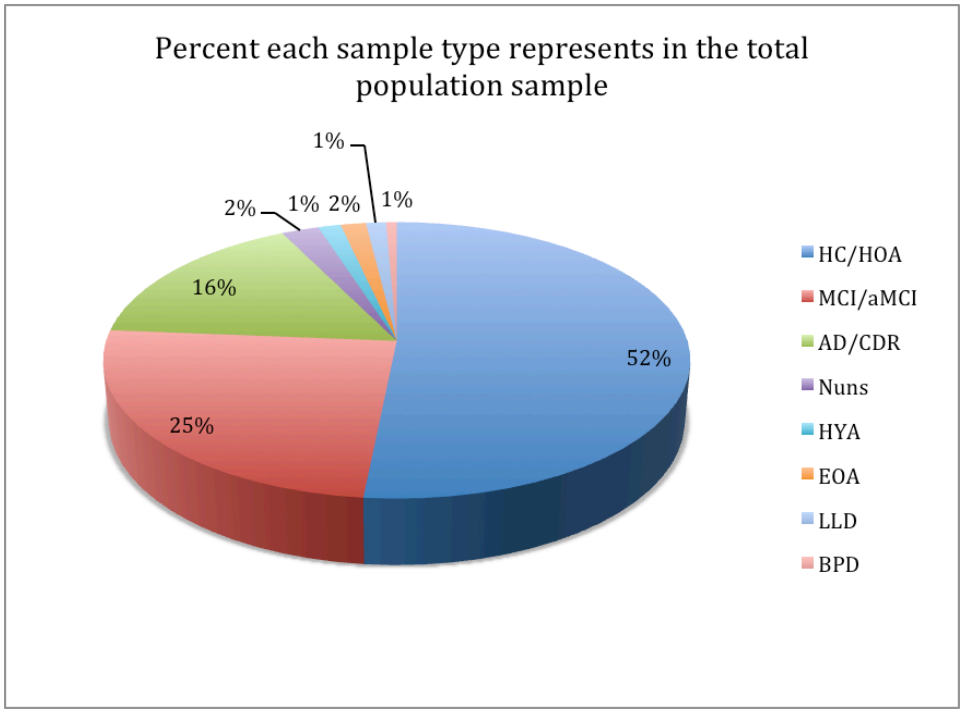


Figure 3.1: Pie chart showing the percent each sample type represents in the whole population sample of included studies.

Of the total number of participants (2377) in the 17 studies, only 1652 participants appeared to be included in the comparison between manual and automated methods. This was due to three papers (Gosche et al, 2001; van Der Lijn, 2008; Wolz et al, 2010) using a subsample of their total data set to compare manual with automated methods. This reduces the percentage represented by healthy older adults, in the total population with manual segmentations (1652) to 43% and those with Alzheimer's disease to 13%. It increases the share of those with Mild Cognitive Impairment to 35%, with the remaining groups to 9% (Figure 3.2).

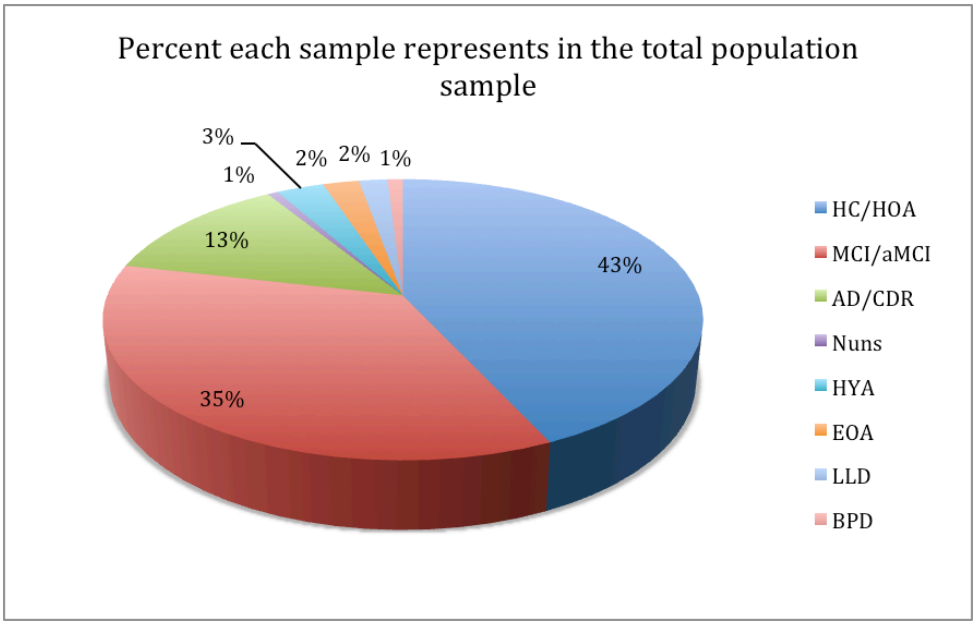


Figure 3.2: Pie chart showing the percent each sample type represents in the total sample involved when validating automated methods in the included studies.

Seven studies did not provide gender information (Bishop et al, 2011; Brewer et al, 2009; Carmichael et al, 2005; Grabner et al, 2006; Morra et al, 2010; Sabuncu et al, 2010; Shen et al, 2009), from the remaining ten there were 842 males and 1172 females. Thirteen out of seventeen papers stated that the images in their study were acquired at scanner strength of either 1.5T (11) or 3T (2). Slice thickness ranged from 0.8mm to 3mm in the fifteen papers that reported this information. Only seven of the seventeen papers stated that they had normalised hippocampal volume by intracranial volume (ICV), of the remaining ten papers no information was provided to suggest that hippocampal volumes had been normalised by any other metric (Table 3.2). ICV should be considered when calculating similarity indices, as it is used to determine how much of the reference standard is contained in the binary mask that identifies the region that has been segmented and how much of the region that should not be included has been correctly left out.

3.3.3 Brief description of automated methods

All 17 included papers performed automatic segmentation of the hippocampus using a registration model that fits a segmented volume from an atlas to that of the subjects in their test sample. The atlas used for segmentation varied from paper to paper, but in

general they were of two main types; standard downloadable atlases such as those from Harvard (Shenton et al, 1995) or the Montreal Neurological Institute (MNI; Evans et al, 1993), and cohort-based atlases. The Harvard atlas uses a white, right-handed, 25 year-old male chosen due to the good quality of the image data. The MNI atlas was developed from a database of 305 participants, 239 males and 66 females with a mean age of 23.4 (4.1) years. Another common atlas used is the Talairach (Lancaster et al, 1997) that is based on post-mortem analysis of the left hemisphere of a 60-year-old female. Cohort atlases are based on one or a set of manually segmented images taken from the population that the automated segmentation will be applied to. The images are chosen at random and used as a reference image with which to train or develop the algorithm on. Standard atlases were used in eight papers (Bishop et al, 2011; Carmichael et al, 2005; Grabner et al, 2006; Leung et al, 2010; Lim et al, 2012a; Morra et al, 2010; Sánchez-Benavides et al, 2010; Shen et al, 2009) and nine papers used cohort atlases (Brewer et al, 2009; Csernansky et al, 2000; van Der Lijn, 2008; Gosche et al, 2001; Hsu et al, 2002; Khan et al, 2008; Sabuncu et al, 2010; Shen et al, 2012 Wolz et al, 2010). Cohort-based atlases appeared to provide better registration (Carmichael et al, 2005; Khan et al, 2011; Shen et al, 2012).

Methods used to register the hippocampus template to the sample ranged from linear affine to fully deformable models, or combined with iterative refinement processes (Wolz et al, 2010; Grabner et al, 2006). There are many different registration algorithms but they can be broadly classified into affine, semi-deformable and fully deformable methods. Affine registration uses rigid/linear transformations to fit the whole chosen atlas to the image, meaning that the relationship between voxels remains constant. Semi-deformable methods spatially align the atlas to the target image smoothing gradually at each iteration, and allowing a greater degree of voxel-to-voxel deformation. Fully deformable registration has no spatial constraints meaning that the geometric position of each voxel can be completely altered to fit the target image. Carmichael et al. (2005) found that more highly deformable geometric transformation models tended to achieve better results than less-deformable models, which was in agreement with earlier studies (Crum et al, 2001; Fischl et al, 2002; Hogan et al, 2000).

3.3.4 Statistical analysis used to compare methods

The 17 included papers used various measures of overlap, correlation, distance and similarity between manual and automated methods to assess the accuracy of the automated segmentation method (Table 3.3)

Study reference	N (included in comparison)	Method of comparison with reference method	Difference between methods		
			Combined	Left	Right
Bishop et al. 2011	17/17	Dice coefficient False-positive rate	Box plots no values given	-	-
Brewer et al. 2009	40/40	ICCs Pearson correlation	0.93 p<0.001 0.88 p<0.001	-	-
Carmichael et al. 2005	6/54	Overlap ratio	Box plots no values given	-	-
Csernansky et al. 2000	51/51	Overlap of contours	80%	-	-
Gosche et al. 2001	11/56	Correlation with two separate raters		0.81 0.79	0.78 0.64
Grabner et al. 2006	118/118	Kappa mean correlation	0.770 (0.027) 0.789 (0.032)	-	-
Hsu et al. 2002	60/60	Correlation	-	0.92	0.91
Khan et al. 2011	37/37		Box plots no values given	-	-

Leung et al. 2010	30/30	Jaccard Index mean (std) MC MCI AD	-	 0.80 (0.03) 0.81 (0.03) 0.79 (0.05)	-
Lim, Hong, et al. 2012	60/60	Dice coefficient	0.7	-	-
Morra et al. 2010	40/40	Correlation with two separate raters	-	0.740 0.694	0.717 0.709
Sabuncu et al. 2010	39/39	Auto – Manual vol relative vol diff	Box plots no values given	-	-
Sanchez-Benavides et al. 2010	89/89	% Overlap % Difference Pearson Correlation	-	79% 10% 0.85	77% 11% 0.84
Shen et al. 2009	125/125	Pearson Correlation	-	0.850	0.832
Shen et al. 2012	217/217	Jaccard Index False-positive ratio False-negative ratio Similarity index 	MNI template Template S Template E MNI template Template S Template E MNI template Template S Template E MNI template Template S Template E	0.51 0.77 0.80 0.42 0.20 0.10 0.56 0.25 0.19 0.13 0.05 0.02	0.35 0.73 0.73 0.54 0.22 0.27 0.75 0.32 0.27 0.19 0.10 0.01
Van der Lijn et al. 2008	20/20	Relative Vol Difference Similarity Index (Mean) Interclass correlation coefficent	-	0.82 0.85 ± 0.04 (0.76-0.90) 0.65	0.80 0.86 ± 0.02 (0.83-0.89) 0.70
	498/498	Dice overlap		0.85	0.85

Wolz et al. 2009	182/796	Dice overlap	Direct	0.78 ± 0.09 (0.47-0.90)	0.79 ± 0.08 (0.44-0.90)
			Direct, GC	0.82 ± 0.06 (0.46-0.90)	0.83 ± 0.07 (0.44-0.90)
			LEAP, N = 300, no GC	0.80 ± 0.05 (0.63-0.90)	0.81 ± 0.05 (0.63-0.90)
			LEAP, N = 1	0.84 ± 0.02 (0.77-0.89)	0.83 ± 0.02 (0.75-0.88)
			LEAP, N = 300	0.85 ± 0.03 (0.68-0.90)	0.85 ± 0.03 (0.73-0.91)

Table 3.3: Seventeen papers that validated the automated method against manual measurements of the same participants; details of the statistical analyses and agreement between methods.

Spatial overlap measures show either the concordance or the differences between two segmentations of the same structure, by overlaying segmentation output masks over one another, usually registering the experimental method to the reference standard mask. The number of voxels that are common and/or differ are calculated, providing a statistical comparison based upon the spatial overlap of the two masks (Figure 3).

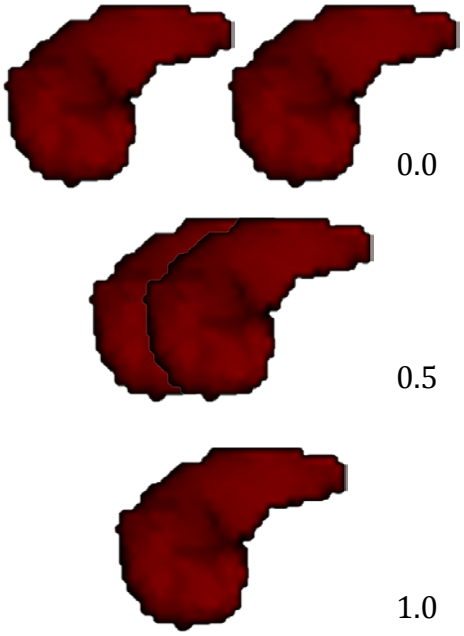


Figure 3.3: Shows an example of a hippocampal segmentation mask at spatial overlap of 0.0, 0.5 and 1.0.

Similarity measures that assess spatial overlap were employed in eleven papers (n=748/1652), (Bishop et al, 2011; Carmichael et al, 2005; Csernansky et al, 2000; Khan et al, 2011; Leung et al, 2010; Lim et al, 2012a; Sabuncu et al, 2010; Sanchez-Benavides et al, 2010; Shen et al, 2012; Van der Lijn et al, 2008; Wolz et al, 2010), eight studies (n=503/1652) used correlation coefficients of the numeric hippocampal volume without accounting for spatial concordance, (Brewer et al, 2009; Gosche et al, 2001; Grabner et al, 2006; Hsu et al, 2002; Morra et al, 2010; Sanchez-Benavides et al, 2010; Shen et al, 2009, Van der Lijn et al, 2008). Four studies (Bishop et al, 2011; Sánchez-Benavides et al, 2010; Shen et al, 2012; van Der Lijn et al, 2008) used more than one type of measure to investigate the differences in hippocampal volume between methods (n=503/1652). From the four papers with over 100 participants used to validate the automated method, two used correlation statistics to assess agreement between methods (Grabner et al, 2006; Shen et al, 2009), one paper employed Jaccard

Index, Similarity index, False-positive and false-negative ratio (Shen et al, 2012), and one paper used the Dice coefficient (Wolz et al, 2010). The range of correlation coefficients reported was from 0.64 to 0.92, and the overlap/similarity measures ranged from 0.77 to 0.87. Seven papers (Bishop et al, 2011; Leung et al, 2010; Morra et al, 2010; Shen et al, 2009; Shen et al, 2012; van Der Lijn et al, 2008; Wolz et al, 2010) visually assessed the segmentations from the automated method, with most reporting under and oversegmentation. All 17 papers reported their agreements as 'acceptable', suggesting that they considered the results to be sufficient in determining hippocampal volume. None of the included papers reported significant differences between methods.

3.3.5 Reported discrepancies between methods

Discrepancies reported between methods included over and under segmentation, registration error, and difficulty identifying the medial boundary of the hippocampus resulting in poor segmentation by the automated method. Both over and underestimation by the automated method were reported, with the direction of difference changing depending on the category of participants. For example, Van der Lijn et al (2008) reported that hippocampal volume correlations with age were significant when measured manually but not when the volumes were derived using the automated method. Others stated that the automated method produced 10% smaller volumes than the manual segmentations in their sample due to the method having difficulty with limited image resolution (Hsu et al, 2002). An optimised, age appropriate template segmentation was reported to have performed better than those using standard templates, however a 9% underestimation of the volume was found compared to manual segmentations of the same sample (Shen et al, 2012). And Bishop et al (2011) reported overestimation by the three automated methods, specifically at the medial inferior boundary and amygdalo-hippocampal border.

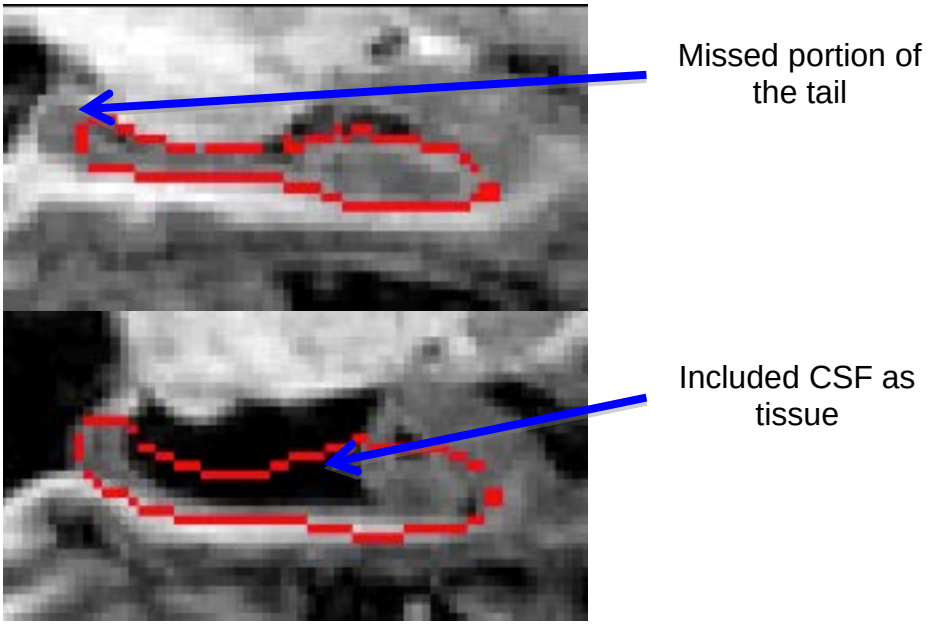


Figure 3.4: Examples of under (top image) and over (bottom image) segmentation by FSL_FIRST.

Differences in manual tracing protocols introduced variability to the reference standard, which may account for some of the differences between automated and manual segmentations. Specifically, the atlases employed by some automated methods were originally derived from manually traced segmentations. If the tissue classification of the hippocampus differs significantly from the manual tracing protocol included for comparison, systematic errors will be introduced (Carmichael et al, 2005). However, we were unable to assess the magnitude of this effect across the studies. Despite reporting no significant difference between the automated and manual methods overall, one paper found that Alzheimer's disease participants showed a significant difference between the size of the hippocampus in the left hemisphere versus the right hemisphere (Sánchez-Benavides et al, 2010) using the manual segmentations, but not using the automated segmentation. The same study found that an association between hippocampal volume and performance on the Mini Mental State Exam that was significant with the automated method, was not significant using the manual volumes (van Der Lijn et al, 2008). Another study that reported different results between groups depending on the method used was Bishop et al (2011). They reported finding that their in-house developed method showed no significant false-positive rate bilateral group effect between the healthy control and

Alzheimer's disease groups, compared to a right hippocampal increase of 23% using FSL_FIRST and 63% using Freesurfer.

3.4 Discussion

Ageing is now a major strategic health priority, and the hippocampus is a vital target for researchers aiming to understand the ageing process. It is implicated in one of the most common dementia types, Alzheimer's disease, but a large degree of variability in hippocampal volume and correlations with cognitive performance is present in normal ageing. Understanding the processes that are affecting non-dementia related hippocampal atrophy is important if we are to discover why some older adults do better in later life than others. Information that can help researchers to determine people most at risk of ageing poorly, would provide an essential pre-disease period in which treatments and even intervention could be more successful. Developing methods to assess hippocampal volume is essential in this process, but to do it successfully can be difficult. The main findings from the review are that automated methods have been successfully applied to healthy older individuals brains. Unfortunately, somewhat limited validation in this population means it is difficult to conclude how reliable these methods are for use in studies of ageing. Those studies that did adequately validate automated methods found good agreement with manual segmentations suggesting that automatic approaches can provide a reliable alternative to manual segmentation of the hippocampus when appropriately setup. Comparison with manual volume measurements in eight of the papers was performed by statistics that simply compare the difference between two volume measurements, not the actual spatial concordance (503/1652 participants representing 30% of total participants in all studies). Hence, where an over-estimation has occurred in a specific area of a target object by one method and underestimation has occurred elsewhere, comparison of volume measurement will not reflect discrepancies. Furthermore correlation analyses are of very limited use in determining agreement or bias between methods.

Despite all the included papers reporting either no significant differences between automated and manual volumes or good overlap agreement, the variability in the strength of associations found between certain factors and automated segmentations

compared to manual segmentations is concerning. The discrepancy may not be large enough to result in a significant difference but it is large enough to effect the outcome of the study findings, as shown for example when an association between a group characteristic and hippocampal size is significant with one method but not another. It is important to note that systematic bias, such as underestimation across a group, could seriously effect the conclusions made from the resultant analysis, therefore all segmentation methods need to be assessed for error rate within the population being studied.

The lack of full demographic information about study participants makes comparison of studies difficult, as does the small sample sizes, varied image analysis methods and statistics used to measure performance. Specifically, the samples used to validate automated methods are, in some cases, considerably smaller than the entire data set. The purpose of employing automated methods is to avoid labour intensive manual segmentations. However, to be confident in the results, a comprehensive comparison with an adequately sized data set should be used. Of the eleven papers that measured agreement between manual and automated methods using statistical analysis that compares the spatial overlap of the two segmentations, only two papers (Shen et al, 2012; Wolz et al, 2010) had relatively large sample sizes of 217 and 182 respectively. The authors reported good agreement with optimised parameters in both studies, significantly the use of templates or atlases that more effectively account for anatomical variation common in older adults produced the best results. This finding is encouraging as it suggests a practical way in which researchers looking to employ automated methods in older subjects can improve the performance of their chosen method. It also highlights the importance of validating methods in large well characterised data sets, something that is acknowledged by Khan et al (2011) who suggests that initial validation of a method may differ from practical scenarios thus application of methods in large, varied datasets is necessary.

Atrophy in the brain occurs in various regions at different rates as we age, making registration using standard landmarks difficult in older brains. For the high degree of anatomical accuracy required to delineate a small structure such as the hippocampus successfully, registration needs to be very good. Templates used in registration have usually been developed from young healthy participants, who do not suffer from

atrophy or high degrees of asymmetry. The differences between the young healthy brains used to produce atlases and those of old healthy brains will introduce error into the registration process, consequently affecting the accuracy of a method's ability to segment a structure. Our findings suggest that a selected cohort atlas, or a manually segmented sample of the target data set from which the algorithm being used is trained, seemed to improve the performance of the automated methods (Carmichael et al, 2005). Also optimising the templates and registration methods for older participants was shown to provide more consistent results from automated software in an older dataset (Shen et al, 2012). Shen et al (2012) acknowledge that it is not generally recognised that age-specific templates are needed, though it seems obvious that templates that better reflect the anatomy of the target dataset would give better segmentation results. This highlights the importance of using age appropriate templates as well as thoroughly validating methods, especially on test data that the method is most likely to be applied to, for example using elderly participants when applying hippocampal segmentation methods.

The automated methods used in these studies appeared to give a good approximation of hippocampal volume considering the minimal amount of user input required to produce segmentations once the program had been setup. The reduction in time taken to perform the segmentations is attractive, especially when large datasets are being analysed. Though good agreement was reported between automated and manual measurements, it was noted that all methods had errors. If the error was systematic across a sample (e.g. all hippocampi measured as slightly larger than true size) then the error may be acceptable; however error that is systematic across a population or patient group within a sample could lead to erroneous associations between variables.

An important consideration when validating automated methods is that of assessing their reliability against manual 'reference' measurements of the same dataset. Definitions of the manual protocols being used needs to be stated when automated methods are being compared to manual measurements, to ensure that the operational definitions of both methods are clear and they are including/excluding the same structures. Though there are potential problems, such as rater error, with assuming manual methods as the reference standard it is necessary to validate methods thoroughly against a recognised and anatomically accepted reference. Efforts are

being made to synthesise variation in manual segmentation protocols of the hippocampus (e.g. http://www.hippocampal-protocol.net). Visual assessment of segmentation output would be an advisable step to ensuring that automated methods are performing well, but this should accompany quantitative validation. Further research is needed to produce better atlases for registration of older brains and better training templates for the hippocampus. Also their need to be more automated hippocampal segmentation methods tested in older participants and validated against manual tracing of the same data set. Automated segmentation techniques provide an alternative to manual segmentation when gross differences are anticipated in hippocampal volume, and bias and error in its measurement are considered to be amplified by the magnitude of biological difference across the study population.

Significant progress in registration algorithms has already started to improve the performance of automated methods (Klein et al, 2009) and continuation of this work with older adult images in mind would be incredibly useful. The advancement in registration techniques feeds into the work showing that methods that employ multi-atlas segmentation methods are more successful than single atlas methods (Aribisala et al, 2013). The advantage of a multi-atlas method is obvious where older adults are involved, as the inclusion of several atlases with more variation will be better able to account for the variability ubiquitous in older adult brains. In turn this leads to the importance of atlases being developed, and incorporated into automated processing pipelines, that are based on large samples of healthy individuals from a wide age range. Development of such atlases are also dependent on successful registration methods, therefore as each area moves forward the effect permeates the other and overall progress in the field significantly advances.

Whilst automated methods are an appealing alternative to manual segmentation, especially in large data sets, development of more standardised techniques by which new automated segmentation methods are validated in relevant datasets, as well as reliable age-appropriate atlases, would be hugely beneficial to the brain imaging research community. The application of two freely available automated methods to a sample of the LBC1936, who also have manual segmentations of the hippocampus for comparison, is presented in Chapter 10. This was undertaken to practically

demonstrate the findings from this review, as well as find a suitable method to use for measuring the hippocampus in the entire LBC1936 cohort.

Chapter 4. Systematic review investigating reported protocols for identification of the posterior frontal lobe boundary

4.1 Introduction

The frontal lobes are implicated in many aspects of cognition such as planning, memory and response inhibition, as well as deficits in this region being linked to various neuropsychiatric and neurodevelopment disorders (Ranta et al, 2009). Age-related decline in cognitive tasks thought to be involved predominantly in the frontal lobe, led researchers to investigate the possibility that this region is adversely affected by brain tissue loss as we age (West, 2000). The 'frontal lobe hypothesis' suggests that functions dependent on the frontal lobes will be adversely affected by ageing whilst functions not reliant on the frontal lobe will be relatively unaffected (West, 2000). A review by Greenwood (2000) examines the efficacy of this hypothesis by looking at the literature regarding the effects of ageing on functions mediated by prefrontal, parietal, temporal and occipitotemporal lobes. The review concluded that age related changes occur both functionally and structurally in the frontal lobe, but not to a greater extent than the other lobes that were examined. It seems likely that frontal regions are involved in age-related decline even if the extent of that involvement is unclear.

One reason for the conflict in how implicated frontal brain regions are in age-related decline may be due to the variation in segmentation protocols applied to measure the frontal lobe. If researchers do not apply the same posterior boundary to segmentations across all studies, reported results may differ due to the fact that the same anatomically defined tissue is not being measured. Neuroanatomical definitions suggest that the central and lateral sulci separate the frontal lobe from both the parietal

and temporal lobe respectively (Duvernoy, 1999). Unlike for subregions of the frontal lobe there seems to be little debate about the definition of the posterior and lateral boundary of the whole frontal lobe (Cox et al, 2014). However, there is much variation in how this anatomical boundary is applied to segmentation of MR images. The fairly robust assumption that brain topology reflects underlying tissue structure can be used to guide segmentation but the practical application of this method from MRI has led to the use of gross geometric landmarks in determining boundaries. The use of cut-planes, where a single anatomical landmark is identified and a straight-line from this point is drawn to determine the frontal lobe posterior boundary, have been employed (Wible et al, 1995). Geometric grids and cut-planes that use structures other than those that have been identified as being anatomically related to the frontal lobe, only serve to introduce variability within the measurement despite having been employed to reduce variation. It was therefore deemed important to look at the variation in reported anatomical boundaries and methods used in segmentation protocols applied to measuring the frontal lobe.

4.1.1 Anatomical definition

At its posterior-lateral edge, the frontal lobe is situated anterior to the central sulcus. Also known as the central fissure of Rolando, this deep sulcus runs from the medial wall, over the lateral convexity until its ventrolateral termination at the sylvian fissure, separating the frontal lobe from parietal tissue. The precentral gyrus itself contains the primary motor cortex (BA4), immediately anterior to the precentral sulcus (PrCS) and supplementary motor area (SMA; BA6; Duvernoy, 1999). The differentiation between frontal and prefrontal lobe is traditionally made on the lateral surface, with the latter excluding both motor and supplementary motor regions (Semendeferi et al., 2001). The frontal lobe is ventrolaterally separated from the temporal lobe by the sylvian fissure, and on the ventral aspect is divided from the insular cortex (considered by some to be an entirely separate lobe; Stephani et al., 2011) by the circular sulcus of the insula.

4.2 Methods

4.2.1 Search strategy

We searched abstracts and article titles using MESH headings in Medline and EMBASE, (covering articles from 1946 to present) on 22nd September 2011 using the search string; (structural OR structure OR volume OR volumetric) AND (parcellate OR parcellated OR parcellation OR measure OR measurement OR estimate OR estimation) AND (frontal OR prefrontal). The references of all screened articles were searched for further relevant papers.

4.2.2 Inclusion/Exclusion criteria

Those studies reporting a method for manual tracing of the human frontal lobe from landmarks or three-dimensional boundaries on MR images; and those papers that were published in English were included. Studies that were excluded were those that involved exclusive use of co-ordinate grid systems; where insufficient detail was given to describe a complete segmentation protocol; if studies repeated the same method using the same participants; if a fully automated method was used or functional MRI was employed. Information was reviewed from both publication and supplementary material where available. Where a protocol was unpublished, the authors were contacted in the first instance, and the study excluded if there was no reply.

4.2.3 Data Extraction and Synthesis

We extracted information on boundary limits for ROI in frontal lobes, study population, sample size, age range, MR sequence used, scanner field strength (Tesla), slice thickness (mm), image pre-processing steps (image alignment), number of regions measured, inter- and intra-rater correlation coefficients (Appendix 1).

4.2.4 Study Selection

There were 1740 records initially identified and once duplicates were removed this became 1544. Of these, 1312 reports did not meet the inclusion criteria mainly due to

using automated structural methods, functional MRI techniques, or animals. Studies applying the same protocol to different cohorts were not excluded as they contribute unique information concerning validity and reproducibility in a range of clinical populations or age groups. Of the 232 remaining studies 24 were excluded due to: lack of boundary information (n=15), regions not intended to be exclusively frontal (n=5), grey literature (n=1), and re-reporting previous results (n=5). This left 208 reviewed publications (Figure 4.1). Due to the number of references, a separate reference list including all reviewed papers can be found in Appendix 2.

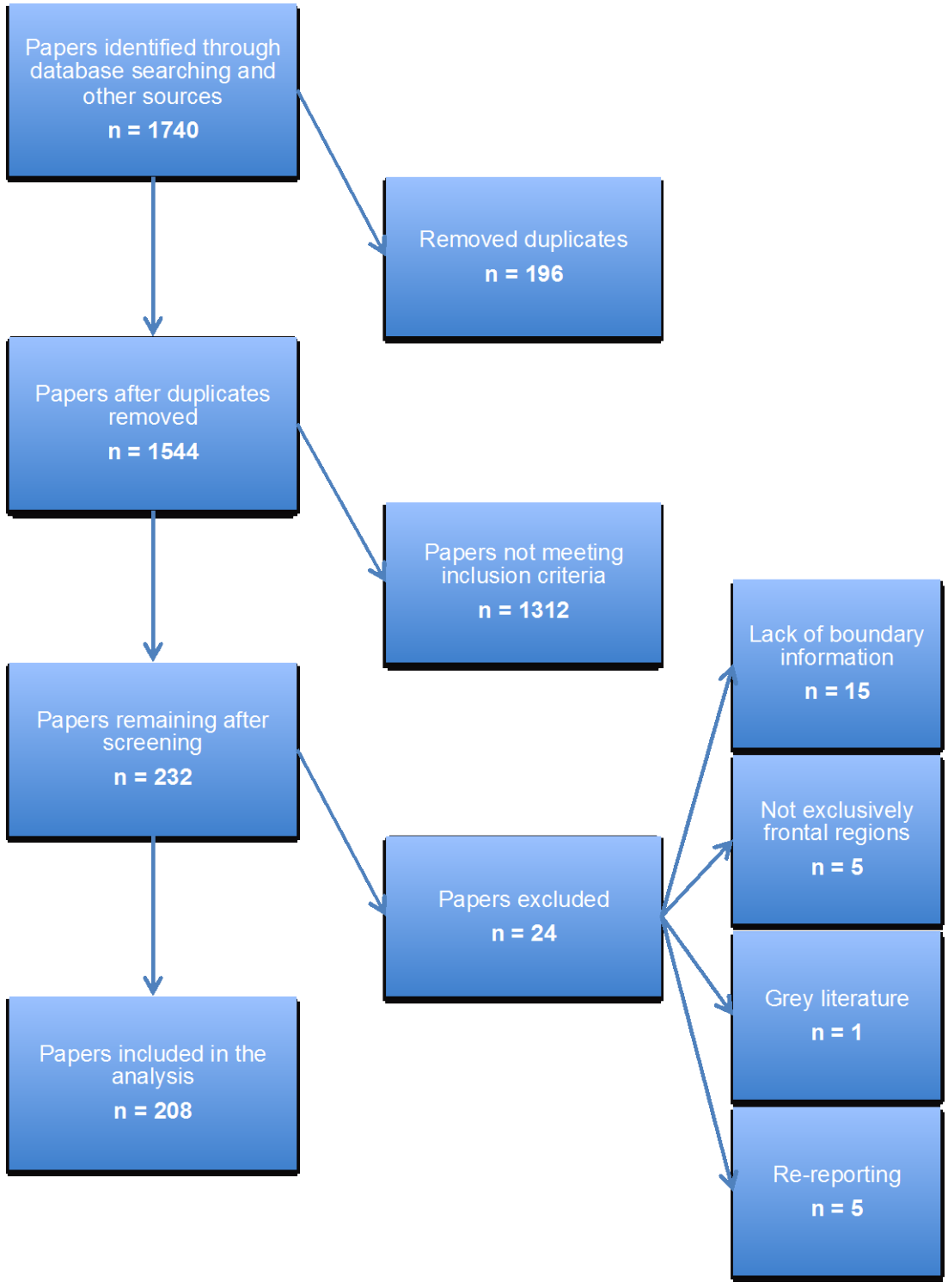


Figure 4.1: Flowchart showing dropout of studies from the initial search to the final included papers.

4.2.5 Study Characteristics

The 208 reviewed papers include 10,903 participants, with a mean of 29 per participant group (median = 22, range = 1-200). The main topics of interest were schizophrenia (25%), affective disorders (unipolar, bipolar, major and minor depressive disorders 13%), dementia (7%), and healthy adults of various ages (26%). Study dates span 1988-2011, and MRI scanners range from 0.1T to 3T in field strength.

4.3 Results

Amongst the reviewed publications, there were a number of variations in the use of lateral, medial and ventral aspects of the posterior frontal lobe boundary. We identified 19 different methods, using 15 different landmarks, for establishing the posterior frontal boundary, which has clear implications for between-study comparison. Though the central sulcus was occasionally adopted as the overall posterior boundary the absence of a clear topographical landmark makes identifying the anterior limit of the supplementary motor area (and therefore the posterior extent of the prefrontal region) problematic. This has led to common use of the precentral sulcus (PrCS) as the most posterior boundary for defining the prefrontal (as opposed to frontal) lobe. Thirty-one papers reported that their measures began anterior to the PrCS. Although use of either central or pre-central sulcus was common, it can be challenging to determine their course when visualising the brain in 2D slices, as reported by several authors (Coffey et al., 1991, Lyoo et al., 1998 and Pantel et al., 1997). Common strategies to overcome this were the use of simultaneous tracing in multiple slice orientations or software that allows ‘painting’ onto 3D renderings to be visualised as a guide during tracing onto standard 2D slices were used (4.2).

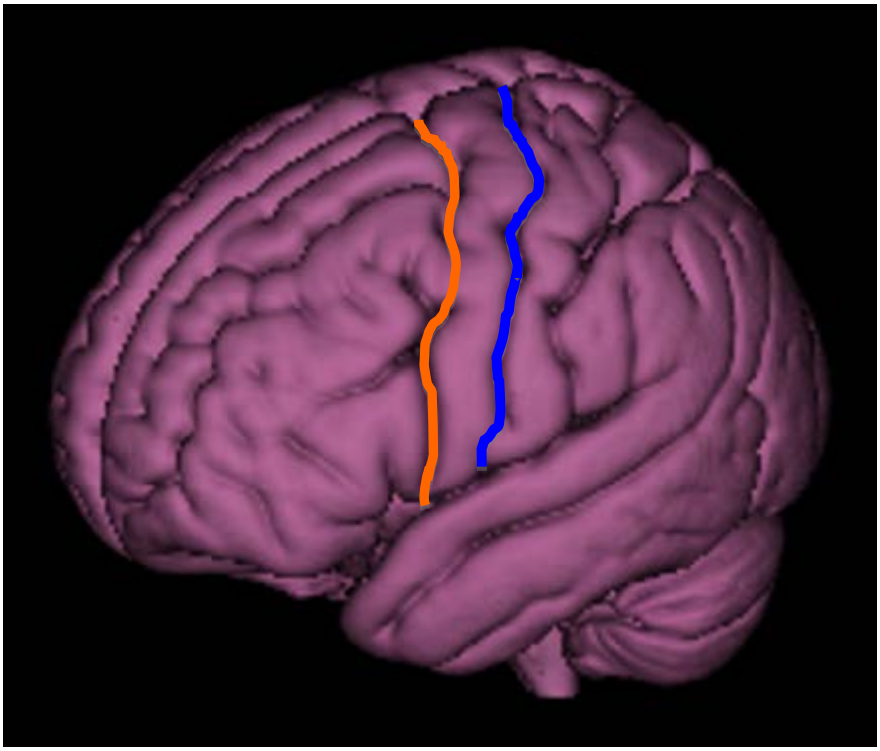


Figure 4.2: Shows the pre-central (orange) and central (blue) sulci marked on the surface of a 3D rendered brain.

Given the difficulty in accurately identifying caudal aspects of the frontal lobe where such methods are unavailable, imposing a coronal cut-plane as the posterior boundary was also found to be a common method. The slice just anterior to, or in which the genu of the corpus callosum appeared was cited by 45 papers as the frontal lobe posterior boundary. The use of two coronal cut planes, one above the body of the corpus callosum where the central sulcus traverses the midsagittal line and one below the genu that intersects the anterior point of the inner surface of the genu, were applied in studies after Crespo-Facorro et al. (1999). Pantel et al. (1997) used the splenium of the corpus callosum but only in the superior slices where it appeared; above the mamillary bodies, a horizontal line from the lateral sulcus (Sylvian fissure) to the midline was used. Other studies used a coronal plane at the midpoint of the corpus callosum (Jernigan et al. 1991; Bartzokis et al. 1993), or a coronal plane a set distance anterior to the most anterior coronal extent of the temporal stem (after Wible et al. 1995). Coronal cut planes have also been employed at the anterior commissure (Bjork et al., 2009; Bremner et al., 2000; Filipek et al. 1997; Nifosi et al., 2010), anterior extent of the lateral ventricles (Coffey et al. 1998), bilateral appearance of the insula (Bäckman et al., 1997; Ginovart et al., 1997), the optic chiasm (Coffey et al.,

1991; Lyoo et al. 1998), the mamillary bodies (Cowell et al., 1994), or 6 mm posterior to the septum pellucidum (Noga et al., 1995) (Figure 4.3). Several papers (Convit et al., 2001; Gold et al., 2005) attempt to distinguish the supplementary motor area from the prefrontal lobe by identifying the coronal plane that equally divides the distance between the cingulate sulcus and the precentral sulcus.

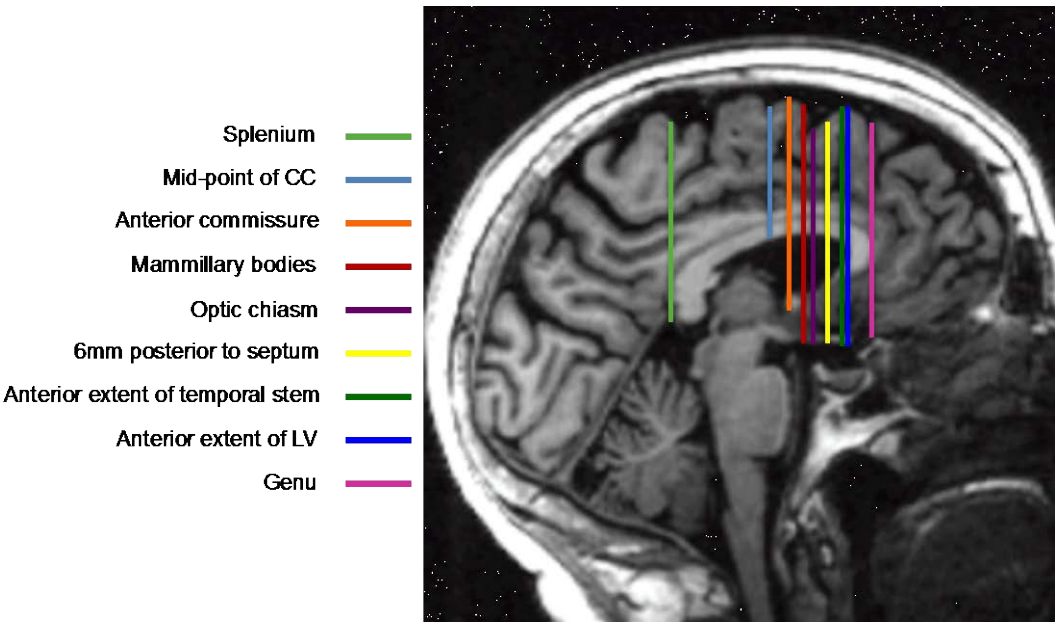


Figure 4.3: Variation of single cut plane boundaries displayed on a sagittal T1W MRI.

Although the majority of cut-plane methods use the selection of distant, sub-cortical landmarks to position cut-planes for the posterior frontal boundaries, explicit attempts to combine cortical topography and cut-planes have also been applied. Kates et al. (2002) selected a coronal slice at the appearance of the precentral gyrus, allowing the exclusion of the supplementary motor area between this plane and the precentral gyrus, based on relevant cortical folding and presumed underlying cytoarchitecture. However, the extent of variation in the angle of the precentral gyrus as it ascends from the dorsal aspect of the brain is unknown. A number of cut planes have also been used to limit the most posterior extent of the ventral frontal lobe. The substantia perforata is a landmark used for many of the papers after Rademacher et al., (1992) to define the posterior boundary of the orbital regions, although Szeszko et al (1999) report difficulties in identifying this. They suggest instead using the most posterior coronal slice in which the olfactory sulcus maintains its characteristic shape, although this, too, may be subject to interpretation. However, the majority of studies that utilise

the central or precentral sulcus to guide frontal lobe segmentation stated the use of the circular sulcus of the insula as the ventral boundary, following traditional anatomical and functional convention. Therefore, whilst the method by Kates et al. (2002) may be a promising approach because it avoids the difficulty in following the central or PrCS in two dimensions and takes account of local topography to some degree, further work would establish whether the area of frontal lobe excluded is equivalent in each individual. The lateral ventricles vary greatly in size within a healthy population, and as such would be a significant determinant of the resultant volumes if using these boundaries (Blatter et al. 1995). Measures obtained using these landmarks has further potential to be influenced by ventricular variation caused by pathology or ageing.

4.4 Discussion

The review of frontal lobe posterior boundaries revealed an overwhelming degree of variation between anatomical definitions and protocols. The consequences of such variability in literature attempting to understand the role of the frontal lobe in ageing is evident in the amount of discrepancies being reported between associations (Zhou et al, 2005). The implications of this variability are wide reaching, as research into frontal lobe decline/deficits straddles many domains and is crucial in our understanding of the potential pattern of brain tissue deterioration with normal ageing. The reasons given for use of geometric or unrelated landmarks, to reduce tracer variability and improve reproducibility of a method, are not sufficient when accuracy is being sacrificed. Protocols that utilise topographical information related to sulcal and gyral patterns alongside software that allows for 3-D visualisation of an image (Schretlen et al, 2000), have proven to be both accurate and reliable with inter-rater correlations of 0.99.

There are important steps that should be taken to reduce variability when employing manual protocols for frontal lobe segmentation, such as ensuring all images are consistently aligned to a chosen axis (e.g. AC-PC), and that decisions concerning individual variation in brain morphology are followed accurately. The importance of following a standardised, well validated protocol cannot be stressed enough, as any other attempts to reduce variability within measurements may be undone by the

failure to identify the correct boundary. Future research should focus on developing a method that combines identification of anatomically clear boundaries with time saving and reproducible techniques, the most effective approach is likely to be semi-automated.

To investigate what implications this variation could mean in actual volumetric measurements, the most commonly applied protocols were applied to a representative sample of the LBC1936. This experiment is described, the results presented and findings discussed in the next section.

Section 3. Global measures

When attempting to quantify age-related brain changes it can be difficult to know where to start, however the most obvious place to begin would be to try to establish a measure of whole brain shrinkage or atrophy. Brain atrophy is one of the most common features of normal ageing and measuring brain volume would provide an acceptable comparison between individuals in cross-sectional studies. The difficulty with comparing volumes between individual's brains is how to account for the variation in head size that naturally occurs in the population. For example a person who is five foot with a petite frame will invariably have a smaller head, and therefore brain, than a six foot large framed person, but it does not necessarily follow that one will suffer more brain atrophy than the other. How then do you establish who has undergone more brain volume deterioration from a single time-point scan without a measure of maximum brain size prior to age-related changes; the simplest way is to use an intracranial volume measurement. The following chapters discuss the reasoning behind using intracranial volume as a proxy measure of maximum brain size, how this measure is made in a large older sample of cross-sectional data and a potential age-related confounder when making an intracranial volume measurement.

Chapter 5. Intracranial volume as a proxy for brain size in youth

5.1 Introduction

It is assumed that once the skull bones have fused and reached maximum size around early adolescence, the size of the cranium remains stable throughout the lifespan (Courchesne et al, 2000). This is unlike brain tissue that degenerates as we age, though the degree and pattern of degeneration varies between individuals (Raz and Rodrigue, 2006). Using cross-sectional MRI we can estimate this change by measuring the contents of the cranium or inner skull table, resulting in a measure of intracranial volume, and seeing what proportion of that volume the brain tissue occupies. The idea is that an ICV measurement provides a proxy measure of the maximum size of the brain in youth from which we can establish the degree to which a person's brain has declined. Development of the skull has been directly linked to brain growth in childhood, but remains stable once peak brain size is reached in late adolescence (Shenkin et al, 2009a; Shenkin et al, 2009b; Shenkin et al, 2009c). However the brain continues to change over the life course, with overall volume steadily decreasing from adulthood and declining more sharply in old age (Courchesne et al, 2000; Tang et al, 2001; Allen et al, 2005).

To establish if using an ICV measurement as a proxy measure of brain volume that is unaffected by age-related decline, longitudinal ICV measurements spanning from childhood to old age would be needed. For obvious reasons there are not yet longitudinal studies of brain changes throughout the entire lifespan but cross sectional data involving large age ranges does exist (Fjell et al, 2009). It is not common for studies of this kind to use ICV measurement as anything other than a way of accounting for differences in head size between individuals. Therefore it can be difficult to assess the trajectories of ICV between ages from published studies. A basic comparison using cross-sectional data, of ICV in young adults with ICV in older adults would provide an indication that ICV is stable at the mean level across adulthood.

Freely available data of healthy adults from the age of 18 to 80 was used to look at the variation of ICV compared to brain tissue volume, to try and establish the relative stability of ICV across the adult lifespan.

5.2 Method

We used data from a freely available online MRI data set (http://www.oasis-brains.org) of 416 (119 males, 197 females) healthy adults with ages ranging from 18 to 96 years. All subjects underwent a clinical assessment to screen for dementia and other neurological disorders.

The scans were T1W magnetisation rapid gradient-echo (MP-RAGE) images acquired on a 1.5 Tesla Siemens scanner in a single session. ICV and total brain volumes were acquired automatically using atlas based segmentation methods (Marcus et al, 2007). We grouped the data roughly by decade to assess change across the age range and compared the mean raw ICV per group; we also compared the mean total brain volume as a percentage of ICV, which was provided online, and compared the mean value for each group. Full details of the image analysis are given in Buckner et al (2004) but briefly ICV is estimated by scaling the manually traced ICV of the atlas by the factor of the affine transform that connects each individual brain to the atlas. FSL_FAST (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) was used to obtain brain tissue volume by assigning voxels as grey, white matter or cerebrospinal fluid. The brain tissue volume is then calculated as all those voxels classified as either grey or white matter within the brain mask. We grouped the data roughly by decade to assess change across the age range and compared the mean raw ICV per group; we also compared the mean total brain volume as a percentage of ICV, which was provided online, and compared the mean value for each group.

5.3 Results

As displayed in Table 5.1, when grouped by decade, mean ICV differs between the youngest (18-28 years old) and oldest group (84-96 years old) by 65cm^3, which is less than the standard deviation of 158cm^3 measured across the whole sample (18-96 years old). The percentage mean TBV, expressed as a percentage of ICV, shows a decrease of 14.1% between the youngest (84.7%) and oldest groups (70.6%), where the standard deviation across the whole group is 6%. In males mean ICV was 181cm^3 larger than in females and there was a 0.01% difference in TBV between men and women.

Group	n	m/f	Age: mean (std)	eTIV (cm3): mean (std)	TBV: (%eTIV) mean (std)
18-28	135	59/76	22.07 (2.58)	1515 (150)	84.7 (1.9)
29-39	19	13/6	32.68 (3.35)	1511 (133)	83.1 (2.0)
40-50	36	12/24	46.19 (2.97)	1446 (164)	82.1 (2.3)
51-61	33	10/23	55.97 (3.11)	1462 (161)	81.0 (2.2)
62-72	63	23/40	68.27 (3.08)	1449 (138)	75.8 (4.2)
73-83	90	29/61	77.68 (3.32)	1478 (182)	73.2 (3.2)
84-96	40	14/26	88.03 (3.01)	1450 (149)	70.6 (3.4)

Table 5.1: Mean estimated intracranial volume (eTIV) and percentage total brain volume of eTIV (TBV) grouped by decade, in the whole dataset and split by gender.

Figure 5.1 plots ICV (Panel A) and TBV (Panel B) by age for the whole sample, males and females. In all cases, ICV remains broadly stable across age, whereas TBV declines with age in an approximately equivalent manner in the whole sample, males and females.

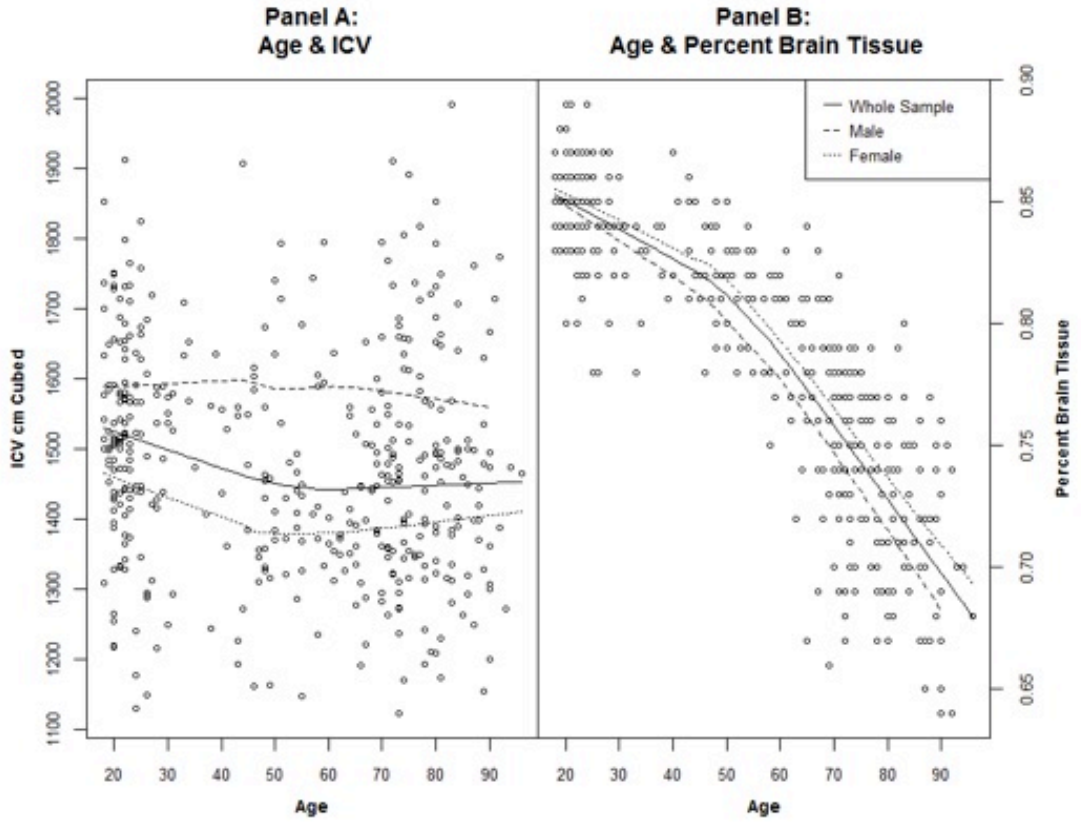


Figure 5.1: Plot depicting the stability of ICV from youth to old age (Panel A) and the decline in the percentage of brain tissue in ICV from youth to old age (Panel B). Data (N=416) were taken from an open source-imaging database (http://www.oasis-brains.org). Data are plotted for the whole sample, and males and females separately.

5.4 Discussion

Overall this simple analysis demonstrates that ICV is a relatively stable variable over the lifespan, with no significant differences between young and old adults. However as expected brain volume declines with age, due to tissue degeneration, with a significant decline in later life (70-80 years old). It is therefore fairly reasonable to assume that ICV can be reliably used as a proxy for maximum brain size, making it an important variable to be used in ageing research.

The brief analysis presented in this chapter supports the hypothesis that intracranial volume can be used as a proxy for brain size in youth, as unlike brain tissue volume, ICV varies little with age. Therefore we can assume that ICV is generally unaffected by age-related changes, providing a stable basis from which to estimate changes in brain tissue volume. Now that the efficacy of ICV measurements in ageing studies has been explored, finding an appropriate method by which to measure ICV is investigated. The specific difficulties associated with a large sample size and the inclusion of older participants in a study requires special attention to be given to the method used when obtaining any volume measurement from MRI. In Chapter 6 the ‘time versus accuracy’ trade-off that is so prevalent in large-scale neuroimaging research is examined in an attempt to find a robust method of measuring ICV. Presented is a comparison of a semi-automated and automated methods of measuring ICV with reference standard manually obtained measurements in the same sample.

Chapter 6. Intracranial volume measurement

6.1 Introduction

Intracranial volume (ICV) represents the maximum brain size in youth and thus provides an important comparator when assessing brain tissue loss in ageing or neurological disorders. ICV is most commonly used to correct for individual differences in head size between participants in a sample, as a way of reducing variation across a group whilst preserving important individual features. Whitwell et al (2001) found that inter-individual variation was significantly reduced when cross-sectional brain volume measures were corrected using total intracranial volume. The process of accounting for individual differences in brain size is especially important in cross-sectional studies where natural variation within the sample population would be expected. This is not the case in longitudinal studies whereby change is accessed within the same person over time thus avoiding confounding differences between individuals.

Traditionally ICV is measured manually or by utilising the boundary between the inner skull and CSF to apply a threshold with which to detect the edge of the intracranial vault. To do this in every slice of a scan is time consuming and would therefore be unfeasible in a large dataset such as the LBC1936. Automated methods of ICV measurement have been compared (Pengas et al, 2009), however despite reports of good performance, overestimation of ICV occurred when compared to manual segmentations of the same scans. Commonly extracranial tissues were found to need to be removed with manual editing; this was also reported by Smith et al (2010) when using a Brain Extraction Tool to obtain a surrogate ICV measurement. One of the problems highlighted by Pengas et al (2009) was that the proximity of the skull and/or dura to the brain tissue sometimes made the edge of the dural margin difficult to distinguish. In older adults the dura and inner skull can thicken resulting in even less distinction in the CSF/skull boundary. Another confound is the presence of image artefacts and incidental findings, common in the ageing brain (Sandeman et al 2013), that mimic brain structures/tissue if a single MR sequence is evaluated (Figure 6.1).

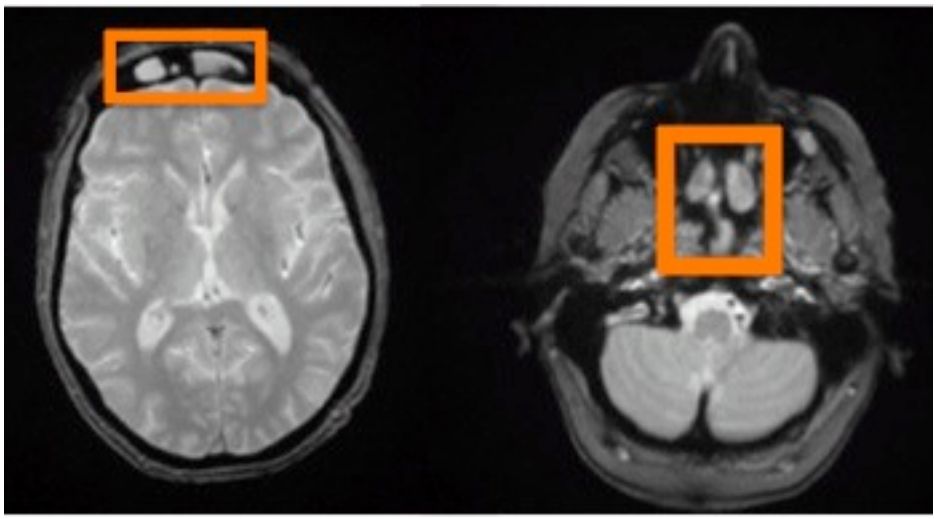


Figure 6.1. Image features on axial T2*W images, which can be mistaken for brain tissue; inflammation and fluid in the frontal sinus (left orange box) and nasal/maxillary polyps (right orange box).

While these anomalies could be easily detected by an experienced neuroradiologist, they could be unnoticed by image analysts, who are the ones that most probably perform the task of computationally extracting ICV. The aim of this research was to determine a fast but accurate method of obtaining an ICV measurement, which could be reproducibly applied to a large number of older adult brains. We explored several computational methods for measuring intracranial volume from magnetic resonance (MR) images and assessed three that represented those explored. Three methods were chosen, two semi-automated and one fully automated, and these were validated against manual measurements, the reference standard, using both volumetric comparison and overlap/agreement statistics.

6.2 Method

6.2.1 Subjects

We randomly selected a sample of 150 individuals using the results of the Moray-House Test taken at age 11 and 70 years, atrophy, brain size and white matter lesion load. The sample is representative of the full range of early and later life cognitive ability in the larger LBC1936 cohort. The sample is made up of 86 males, 64 females and the mean (std) age is 71.9 (0.34) years old.

6.2.2 Manual method or reference standard

For the manual method a T2*W image was used as it shows a clear intensity difference between brain tissue and bone. A T2*W image was loaded into Analyze 9.0, the axial image was then thresholded using the average value intensity from four regions of interest placed in dense white matter (corpus callosum) and bone on two consecutive slices. A seed point set at the calculated average threshold will then detect the edge of the ICV, where the meninges meets the signal void caused by the dense bone of the inner skull table, automatically. Once this seed point is placed, manual editing to remove any erroneous structures such as the optic tracts, orbits and pituitary fossa was performed in those slices where it was necessary (Figure 6.2).

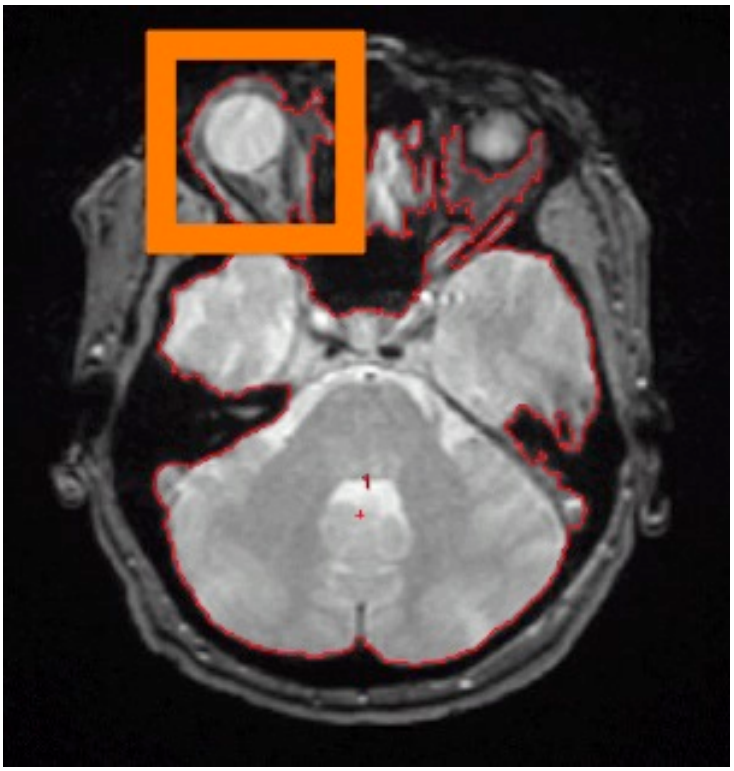


Figure 6.2. Axial T2*W image showing the seed point threshold prior to manual editing to remove erroneous structures such as the orbits and orbital tracts (within orange box).

The inferior boundary of the ICV was determined to be the slice superior to the tip of the odontoid peg at the foramen magnum (Figure 6.3), and excluding the bilateral cavernous and extradural sinuses.

Figure 6.3. Sagittal T2W image showing the inferior boundary of the intracranial cavity at the foramen magnum, indicated by the orange line and boxed inset (top left).

6.2.3 Test method 1 [Object Extraction Tool (OET)]

The Object Extraction Tool in Analyze automatically extracts the contents of the intracranial cavity by applying morphological erosion, dilation, and region growing steps to a threshold predefined. To generate the binary mask, we placed a seed-point in the axial slice where the orbits appear and selected the optimal threshold as the intensity value that separated the optic nerve from the rest of the brain tissue. In the extraction process, morphological dilations were repeated automatically to cover 99% of the voxels in the auto- traced region on the target slice. After extraction, the holes in the extracted object were filled in and a final 6-connected 3D region growing step was performed.

6.2.4 Test method 2 [OET plus editing aided by a colour fusion method]

The colour combination of two or more sequences can help in discerning the boundaries of the intracranial cavity where artefacts and anomalies are present. The idea of using colour to enhance differentiation between tissues and other anatomical features in MR images is not new. Colour MR images were suggested by Holland and Botomley more than thirty years ago (Holland and Botomley, 1977), and the first

clinical colour MR images were published in 1987 by Weiss et al. Soon after, different colour composite techniques were implemented in different ways to improve the information content and enhance conspicuousness of specific tissues and fluids.

The incidental findings of extracranial sinus inflammation previously shown can be easily detected by a non-specialist with the aid of the colour fusion of different sequences (Figure 6.4).

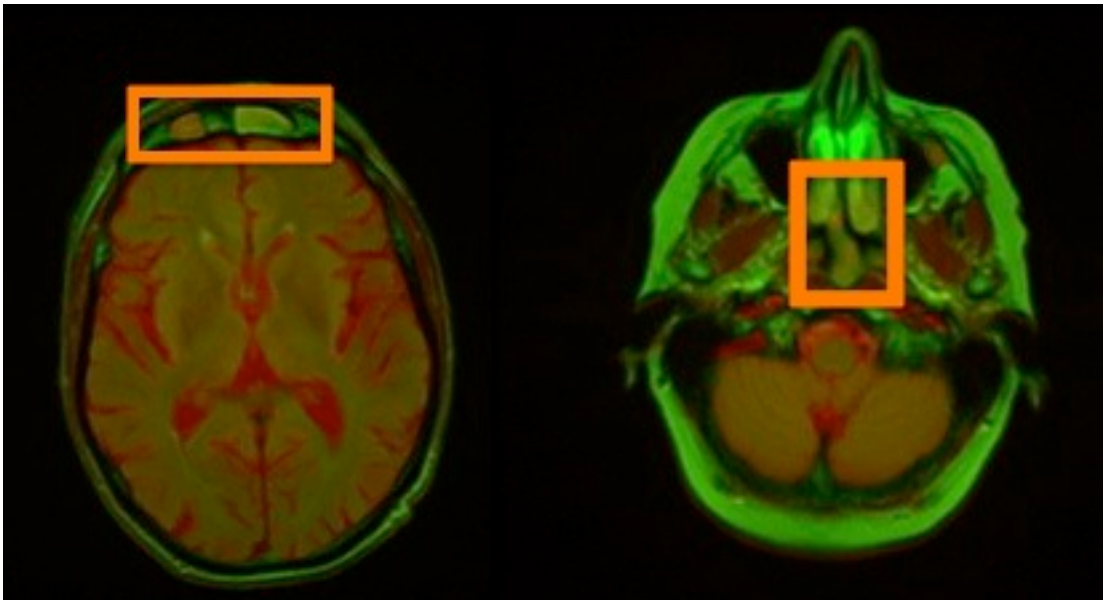


Figure 6.4. Fused axial T2*W and FLAIR images showing the additional visual information used to distinguish non-brain features from brain tissue.

MCMxxxVI has been described in detail in Chapter 2, but briefly it uses colour fusion to increase discrimination between tissue types, thus improving the users ability to accurately identify target tissues for editing. This method differs from the colour composite techniques summarised previously. While utilising a colour scale that mimics the natural appearance of the tissues may be of benefit as an educational tool, it is unlikely that methods that produce seminatural-appearing or virtually realistic appearing tissue tones will serve to highlight pathology. Successfully differentiating otherwise unseen pathology is more likely to be achieved by the simple colour fusion of two or three different MRI sequences and uneducated observer, and will most likely be unrelated to the natural appearance of the tissues.

As an additional step removal of the extracranial tissues following the automatic ICV extraction done by the OET, we used the MCMxxxVI method.

6.2.5 Test method 3 [Brain Extraction Tool (BET)]

BET (Smith et al, 2010) performs a fully automatic brain extraction in three main steps, which is used as a surrogate measure for ICV (Burns et al, 2008; Thoma et al, 2009). Firstly it processes the intensity histogram to find robust lower and upper intensity values for the image, and a rough brain/non-brain threshold. Then, it finds the centre of gravity of the head image, along with the rough size of the head in it. Finally, it performs a triangular tessellation of a sphere's surface inside the intracranial cavity and slowly deforms it, one vertex at a time, following forces that keep the surface well-spaced and smooth, while attempting to move towards the intracranial cavity's edge in an iterative process. This process excludes part of the lower brain stem from the final extracted volume, although does not guarantee that its lower limit coincides with the standard boundary at the foramen magnum.

The BET fractional intensity threshold was optimized by visually inspecting the brain/ICV extractions created with a range of thresholds in a sub-sample of brain volumes. A fractional intensity threshold of 0.6 was found to be the best compromise in our sample and subsequently applied to the full dataset. We did not perform any manual editing of the derived mask obtained by this method.

6.2.6 Exclusion of the pituitary fossa

In all of the methods employed, the pituitary gland was included in the ICV measurement despite it being arguably extracranial (Figure 6.5). The pituitary gland is a small endocrine structure located below the hypothalamus at the base of the brain, it is involved in hormone regulation in the body. Other studies have reported similar exclusion of the pituitary (Keihaninejad et al, 2010; Nordenskjöld et al, 2013). To assess how much the inclusion of this structure would influence the overall ICV, values before and after the exclusion of the pituitary in the manual measurements were compared.

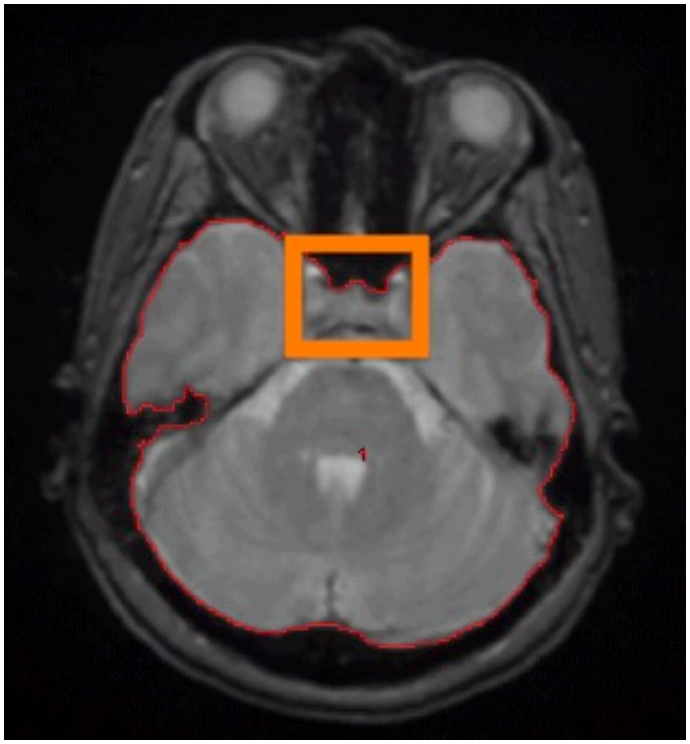


Figure 6.5: Axial T2*W image showing the pituitary gland (within orange box), which has been included in the ICV measurement.

To remove the pituitary gland the line of the sella turcica, the bone cavity in which the gland sits, is followed to establish the boundary of the ICV (Figure 6.6).

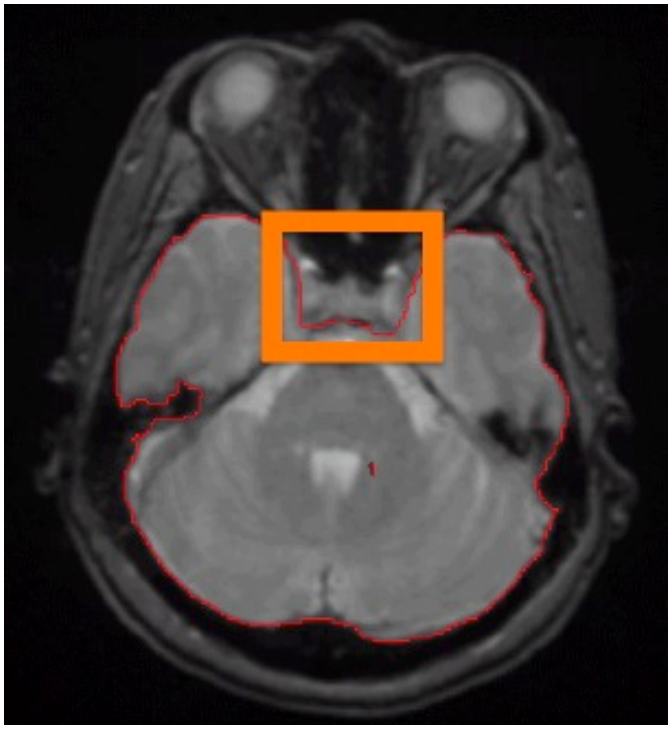


Figure 6.6: Axial T2*W image showing manual exclusion of the pituitary gland (orange box).

6.2.7 Statistical analysis

Agreement between the manual method and the three test methods was assessed using Bland-Altman analysis of absolute and percentage differences, as well as Pearson correlations. The data was normally distributed as assessed by the Kolmogorov-Smirnov test.

6.3 Results

The agreement between the manual method and the 1st) OET, 2nd) OET plus editing with MCMxxxVI and 3rd) BET test methods was assessed using Bland Altman analysis and Pearson correlations. The choice to use Bland Altman analysis in addition to correlation statistics was to try to ascertain the variation between measurements as well as the degree to which they agree. This method is very helpful where under and over segmentation of structures can occur, as it provides positive and negative variance from the mean thus indicting if a method differs from the reference standard due to a volume being larger or smaller.

The reference standard measurement of ICV had a mean of 1502.01cm^3 (sd 142.78), means and standard deviations for all measurements are in Table 6.1.

Methods compared	**Mean ICV ± SD (cm^3)**	**Pearson correlation coefficient (r)**	**Mean difference (cm^3)**	**95% CI of difference (cm^3)**	**Minimum difference (absolute value) (cm^3)**	**Maximum difference (absolute value) (cm^3)**
Reference standard vs. ICV using OET plus editing with MCMxxxVI	1502.01 ± 142.78	0.98	40.27 (2.74%)	105.62 (7.03%)	0.73 (0.05%)	140.58 (9.50%)
Reference standard vs. surrogate	1502.01 ± 142.78	0.97	32.14 (2.18%)	62.51 (8.32%)	0.89 (0.08%)	138.80 (9.37%)

ICV by OET						
Reference standard vs. surrogate ICV by BET	1502.01 ± 142.78	0.98	81.67 (5.38%)	104.16 (6.93%)	2.60 (0.14%)	175.22 (11.75%)
Reference standard including sella vs. Reference standard excluding sella	1501.03 ± 153.53	0.99	8.74 (0.29%)	7.29 (0.49%)	1.61 (0.05%)	18.06 (0.58%)

Legend; SD: standard deviation, CI: confidence interval.

Table 6.1: Comparison of the reference standard measurements with the numeric results obtained by the methods evaluated. Absolute ICV, correlation, differences in volumes with 95% CI of the difference and & mean difference between methods.

The mean difference between the reference standard and the OET method was 2.18%, OET plus MCMxxxVI was 2.74% and BET was 5.38%. The 95% confidence intervals (reported in table 2) are best understood from looking at the Bland Altman plots (Figure 6.7), which show the % difference between two methods plotted against the mean of two methods.

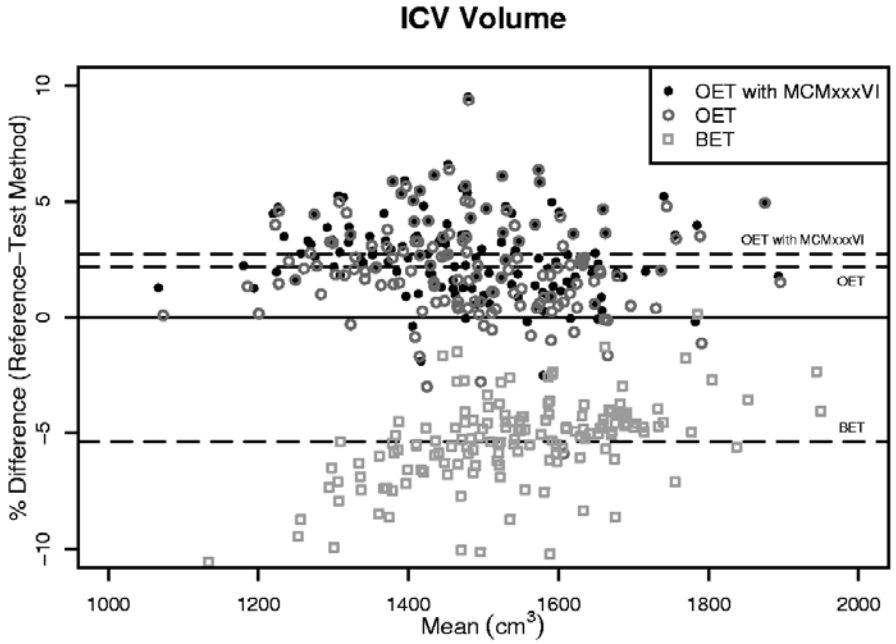


Figure 6.7. Bland-Altman plot comparing the ICV measurements done by thresholding followed by manual editing (reference standard method) and OET (Test method 1), OET with MCMxxxVI (Test method 2) and BET (Test method 3).

The plot shows that although the OET (Test method 1) method may have produced the smallest mean difference (2.18%) compared to the reference standard, it produced a larger variation than OET with MCMxxxVI (Test method 2) whose mean difference was in fact larger (2.74%). As expected this shows that the OET (Test method 1) method alone is less reliable than with the additional step of manual editing, however it performed better than BET (Test method 3) in both instances.

All three methods showed good correlations with the reference standard; OET (0.97r), OET plus MCMxxxVI (0.98r) and BET (0.98r), however in this instance how well two measures are associated does not give the full picture of which performed better.

Bias was found in all three methods and visual inspection of the masks revealed that the main problems were inclusion of structures such as the clivus, orbits and pituitary gland. Also the inferior boundary of the foramen magnum was not consistently applied by the OET or BET methods, causing a small systematic bias.

Jaccard Index (JI) analysis, which provides an indicator of how well segmentation outputs share spatial concordance (Jaccard, 1901), was conducted as part of a separate

investigation and therefore results are only comparing the reference standard, and the OET and OET with manual editing. For the OET method the JI was 0.966 and 0.970 for OET with manual editing, showing that a small advantage can be made by manually editing segmentation output where necessary.

Exclusion of the pituitary gland resulted in an overall mean difference of 0.29% (8.74cm^3), with a maximum difference of 0.58% (18.06cm^3) between the manual measurements that included and that which excluded the structure. This maximum difference of just over half a percent is so small as to be inconsequential in the measurement of ICV, however this is only when volumetric data is being used. If the resultant segmentation masks are to be used in conjunction with other tissue segmentation then the presence of the pituitary gland may result in erroneous measurements. Therefore it is important, where possible, to provide the most accurate segmentations of a structure especially where consistent boundaries can be used to exclude non-target tissues from a segmentation.

6.4 Discussion

Thresholding plus morphological operations followed by manual editing is the most reliable semi-automated measurement method for ICV compared with a reference standard manual method. The two fully automated methods (OET and BET), while fast, both require manual editing to avoid significant errors in ICV measurement and therefore do not necessarily speed up the overall process.

Though the mean differences between the methods and reference standard only ranged from 2.18-5.38%, the variance between measurements was as much as 8.32%. When employing these methods to obtain ICV in studies where subtle differences are expected, these errors could be significant. If age-related differences are one such type of subtle change being investigated, the potential methodological errors highlighted here could hide important differences. Furthermore where ICV is being used in a pipeline, such as for registration of templates or other tissue masks, the inclusion of extracranial structures may influence the effectiveness of the fit thus influencing the success of the registration.

The apparent contradiction of the OET method producing a larger mean value than OET plus editing using MCMxxxVI is due to Bland Altman being unable to account for spatial disagreement. Visual inspection of the OET masks show the exclusion of the dural venous sinuses in some areas due to the automated threshold being higher than the threshold used for the reference standard. Thresholds are calculated using gray- scale intensity levels. When a threshold that corresponds to a lower intensity level is set it will include more voxels in an image than the threshold corresponding to a high intensity level. For the Bland Altman analysis we subtracted the OET (Test method 1), OET with MCMxxxVI (Test method 2) and BET (Test method 3) volumes from the reference standard volumes and gave the percentage difference between methods. OET excludes intracranial structures in some areas but includes extracranial structures below the inferior boundary, when the extra step of removing the extracranial structures is taken the volume decreases causing the percentage difference to increase. These results highlight the importance of visually assessing segmentation output.

The Jaccard Index analysis performed between the OET (Test method 1), OET with MCMxxxVI (Test method 2) and reference standard methods showed that agreement improved when masks were edited where required. This supports the other results reported and adds credence to the explanation that volumetric analysis cannot parse apart false positives from false negatives, thus potentially skewing or hiding discrepancies between segmentation output where both exist.

The use of an older population in this study allowed us to specifically look at the effectiveness of these techniques on scans that inherently cause difficulties for image analysis. Ageing features such as hemispheric asymmetry, brain atrophy and poor tissue contrast between white and grey matter are prevalent in older individuals. This being the case, it is necessary that image analysis methods be able to accommodate ageing features. The added steps of visual assessment and correction by manual editing increase the time it takes to complete a segmentation, but they result in a more accurate and therefore more valid, relevant result. This study has highlighted the steps and considerations necessary to produce a relatively quick yet accurate ICV measurement, using a semi-automated method easily applied to MRI.

The analysis presented in the next chapter discusses the application of a cross-sectional area measurement as a proxy for intracranial volume measurement. The method addresses some of the issues such as time constraints and image acquisition that usually occur in image analysis whilst trying to maintain anatomical sensitivity.

Chapter 7. Comparison of ICA with ICV

7.1 Introduction

Intracranial Cross-sectional Area (ICA) can be used as a surrogate measure for ICV (Ferguson et al 2005; Nandigam et al, 2007). It has the advantage of being fast to apply, as only one slice in the midline sagittal plane needs to be measured rather than making measurements from every slice in an image. As discussed in the previous chapter, it may be unfeasible to manually measure ICV in a large sample but automated methods may not be adequate especially in older adult brains. The possibility of an accurate method that significantly reduces the time taken to produce a measurement is appealing.

ICA utilises previous work that demonstrated the use of linear and cross-sectional measurements as proxy measures of numerous brain structures (Whalley and Wardlaw, 2001). The authors successfully demonstrated that simple linear and cross-sectional measurements could be a fast, accurate and reproducible alternative to detailed volumetric measurements. To investigate if ICA measurements could be used in place of full ICV measurement, 148 existing ICA measurements from the LBC1936 cohort were compared to ICV measurements of the same subsample.

7.2 Method

The sample and ICV measurements were the same as described in the Chapter 6 with the exclusion of two subjects due to an inadequate sagittal localizer being available. The ICV measurements were derived using OET plus manual editing that was found to be the most accurate and efficient method of obtaining an ICV measurement.

Intracranial cross-sectional area measurements were made from the sagittal localizer used at the time of scanning to correctly position a person's head in the scanner. The 2D sagittal T2 localiser imaging parameters were TR = 3520 ms; TE = 102.5 ms, 256 x 256 matrix size and 1 slice with voxel size of 1 x 1 x 5 mm. Measurements were made in the midline sagittal slice, which is located by looking for where the corpus callosum is well defined, the cerebellum is rounded and not yet pyramidal, and structures such as the pons bridge and medulla oblongata are visible (Figure 7.1).

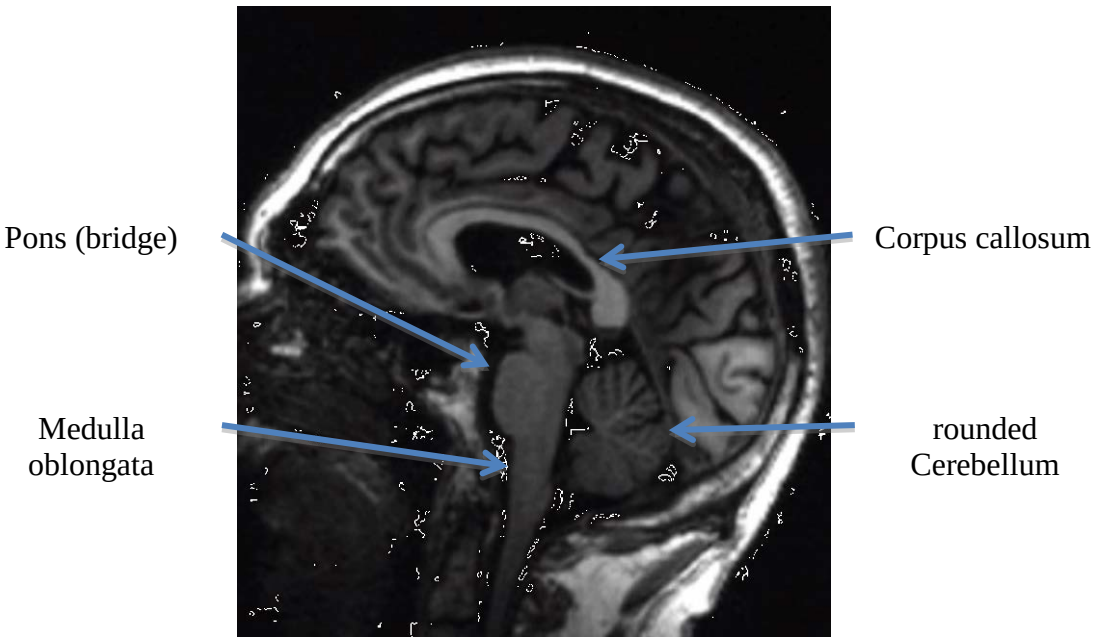


Figure 7.1: T1W midline sagittal slice with structures useful for identification marked.

The ICA is measured using the ROI function in Analyze. The trace tool is used to draw around the inner table of the calvarium from the posterior rim of the foramen magnum to the inferior most point of the clivus, it is continued superiorly over the clivus to the anterior rim of the sella turcica excluding the pituitary gland (Figure 7.2).

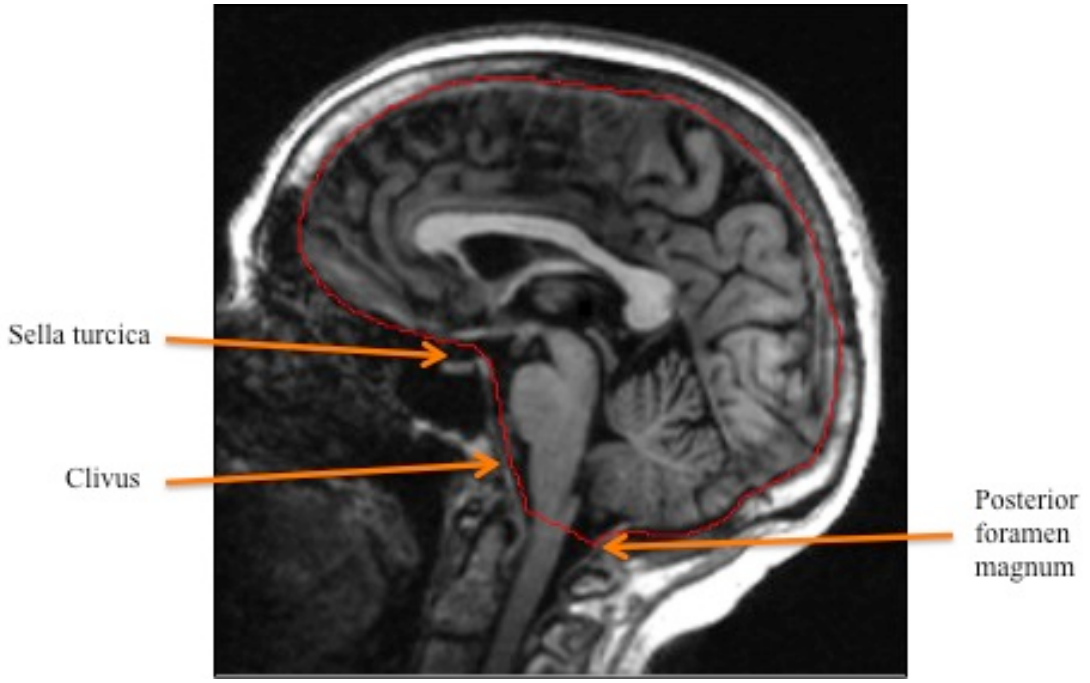


Figure 7.2: Midline sagittal slice identifying the anatomical points of refernce for tracing the ICA.

7.3 Result

The Pearson correlation between ICV and ICA was r= 0.75 (p = 0.01, n = 148). A Bland-Altman plot was generated using Z scores, as the distribution range within each set of raw values is too disparate, this showed a good agreement between the two measurements (Figure 7.3).

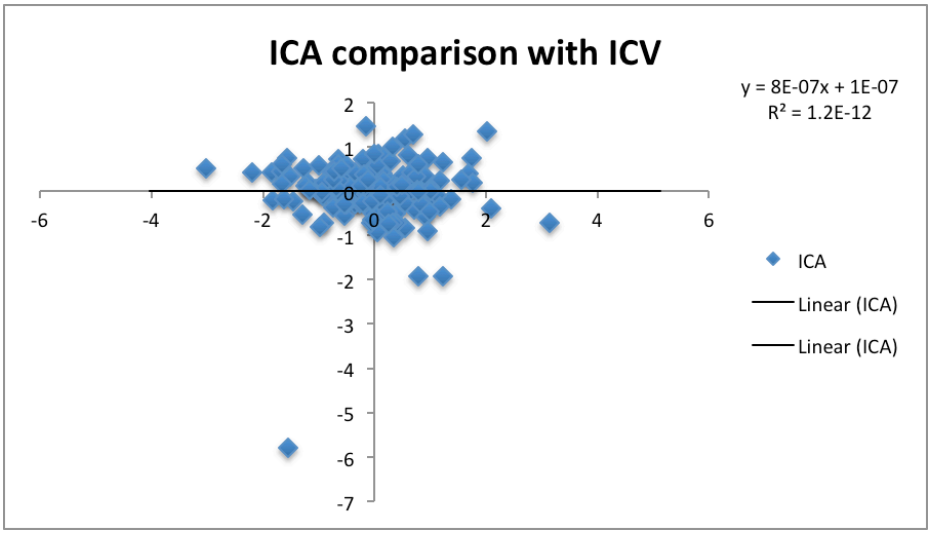


Figure 7.3: Modified Bland-Altman plot of Z scores of intracranial volume and intracranial cross-sectional area.

7.4 Discussion

ICA was reasonably well correlated with ICV suggesting that it is an acceptable proxy measure in the absence of availability of a more thorough measure of ICV. This supports previous work that found measurements made from cross-sectional area are a good estimate of the volume of an entire structure (Wardlaw and Whalley. 2001; Ferguson et al. 2005) when the structure is rounded. The correlation we found was not as large as that previously reported by Ferguson et al (2005) but our sample was larger and comprised of both males and females, therefore gender differences may have be a factor though this would need to be investigated further. The advantages of using ICA are that, due to using a single slice in the sagittal midline, the time taken to produce a measurement is considerably quicker than obtaining an ICV using every slice in an image. This is especially useful where a large sample is being analysed, as manual ICV methods may be too time consuming but automated measurement includes erroneous tissue as discussed in chapter 6. It could also mean that the standard practice of using the midline sagittal slice to position an individual in the scanner could be further utilised by making an ICA measurement from the localizer thus reducing scanning time. Whole brain coverage is not always possible for every individual and this is dependent on the scanner and software being used, accurate ICV measurements require whole brain coverage where ICA measurements do not. Therefore the ICA method would be useful in instances where whole brain coverage is not possible or when a short scanner time is necessary, possibly due to a person being unable to tolerate a long scan.

The limitations of the ICA method are that the plane of view and differences in lateral positioning of the head at the time of acquiring the image, will affect the accuracy of the measurement. Though protocols are followed to reduce differences between individuals, factors such as the size of person's head or their ability to lie in an optimum position for scanning are unavoidable. Corrections can be made using registration tools, though this will increase the amount of processing and therefore time needed to perform a measurement. These are limitations common to all image analysis and not expected to have disproportionately influenced the ICA method. Ideally it would have been more thorough to compare the spatial concordance of these two measures using overlap or false positive – false negative statistical evaluation

methods. However the significantly different shapes of the output masks prohibits this, therefore correlation of volume measuresments was used.

In conclusion the ICA method is an acceptable proxy method of ICV, best used when either scanning acquisition time or time needed for image analysis are at a premium.

Establishing an efficient measure of ICV allows us to correct for head size when looking at individual differences in brain tissue volumes, as well as providing us with a proxy measure of maximal brain size that is unaffected by age-related decline. Both of these uses for ICV measurements are important and they are based on the assumption that our skull remains the same size throughout adulthood. However in the next chapter the influence of the particular ageing feature of inner table skull thickening, common but not ubiquitous amongst older adults, on ICV measurement is investigated. Though not present in all older adults, and not as prevalent as other ageing features such as white matter hyperintensities, inner table skull thickening is routinely identified by neuroradiologists and can be significant in some people. It has thus far been overlooked in neuroimaging studies of ageing, but here the potential influence it has on ICV measurement and estimates of brain atrophy are presented.

Chapter 8. Influence of inner table skull thickening on ICV

8.1 Introduction

Head size is strongly influenced by brain growth in childhood and reaches maximum size by early adulthood (Sahin et al, 2007). It is generally assumed that head size, and therefore intracranial volume (ICV), remains the same from early adulthood to old age. However, age-related skull changes, such as an increase in the thickness of the inner table and overall size of the cranium, have been found (Finby and Kraft, 1972; Israel, 1968). Physiological changes of the skull such as hyperostosis frontalis interna (HFI), thickening of the inner table of the frontal region of the skull, have also long been documented in the medical literature (May et al, 2010). Whereas it is commonly observed by radiologists in older adults, skull thickening is not often mentioned in

ageing research, possibly due to the benign nature of the changes (She and Szakacs, 2004). Although the process is benign, some research suggests that, where the increase is very pronounced, dural irritation and pressure atrophy may occur (Chaljub et al, 1999). Case studies of hydrocephalic children (Griscom and Sang, 1970) and adults with severe brain atrophy (Wolf and Falsetti, 2001) suggest that thickening of the inner skull table may occur in response to the reduction in brain volume caused by atrophy or changes in intracranial pressure. A cause of this sporadic thickening is thought to be hormonal as it is most prominently found in post-menopausal women and some studies have found endocrine abnormalities coincidental with HFI (Harding, 1949; May et al, 2010).

In neuroimaging studies, ICV is used as an estimate of peak prior adult brain volume (Farias et al, 2011; Shen et al, 2009; Staff et al, 2006). Because it is thought that ICV is not influenced by disease or age-related changes, it is therefore often used to estimate brain atrophy. However, the influence that thickening of the inner skull table may have on measures of ICV, and hence on estimates of brain atrophy and its correlations, have yet to be investigated. In this paper, we investigated the potential influence of inner table skull thickening on measurement of ICV and estimates of brain atrophy in a cohort of community-dwelling older adults.

8.2 Method

8.2.1 Subjects

We randomly selected 60 participants from the Lothian Birth Cohort 1936 who had, on visual inspection, a range of inner skull table thickening from significant to little or no thickening. Study participants (31 males and 29 females) were non-demented, community-dwelling older individuals who underwent cognitive tests and brain MRI between 8th November 2007 and 29th June 2010 at 71.1 to 74.3 years of age (mean 72.7, standard deviation (SD) 0.7 years).

8.2.2 Measurement method

All analyses were performed blind to subject details, including gender, on anonymised scans. The scans used were T1W as they allowed for the best contrast between bone and tissue; all were aligned to the anterior-posterior commissure (AC-PC) line to improve reproducibility.

ICV masks were obtained using the method described in Chapter 6. We estimated the original ICV, denoted as 'estimated original' ICV, excluding the effects of inner table skull thickening by editing the current ICV mask slices throughout the skull vault, extending its boundaries to include the inner skull table thickening. The inferior boundary for the measurement of inner table skull thickening was identified as the supraorbital ridge, which is the most prominent point in the midline sagittal view (Hatipoglu et al, 2008). This landmark was used as it is an easily identifiable point and separates the vault where most of the inner table thickening occurs from the frontal sinuses and orbits where there is little thickening and the boundaries that are also more difficult to measure. Using multiplanar display software (MRIcro; www.cabiatl.com/**mricro**/**mricro**/index.html;), the sagittal view was selected and the location of the supraorbital ridge was highlighted to indicate which axial slice was the most inferior limit of the region. Then, for all slices showing ICV and inner table skull thickening superior to the supraorbital ridge, the edge of the current ICV mask was extended by manually tracing along the line where the original skull table was thought to be, as shown in Figure 8.1. The inferior slices remained the same as those in the current ICV mask. The entire mask was re-measured providing an estimate of original ICV measurement without the effects of inner table skull thickening.

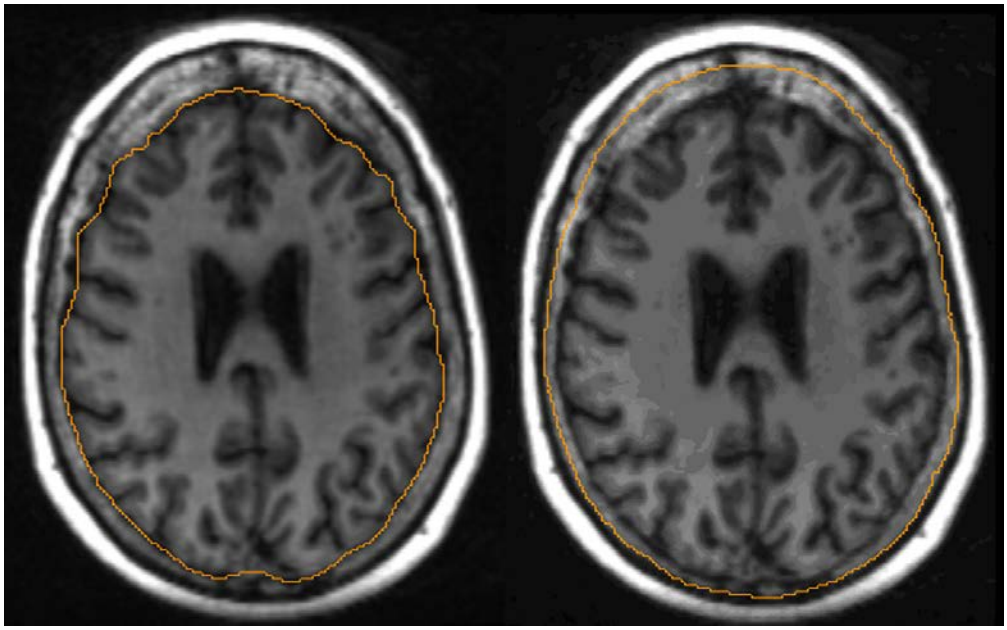


Figure 8.1: Shows the current ICV mask (left) that is manually altered to account for skull thickening resulting in a estimated original ICV mask (right) in an axial view.

Finally, the current brain volume was measured in all subjects and brain atrophy determined by calculating the total brain tissue volume as a percentage of both current and estimated original ICV.

8.2.3 Statistical analysis

The sample was not normally distributed (Shapiro-Wilk normality test) for either the current (W=0.902, p=0.001) or estimated original ICV (W=0.919, p=0.001). The Wilcoxon Signed Ranks test was used to identify differences between current and estimated original ICV across the whole group. The Mann-Whitney U test was applied for differences between men and women. To test the potential effect that thickening of the inner skull table would have on estimates of brain atrophy between youth and old age, the Wilcoxon Signed Ranks test was used to identify the differences between percent brain tissue in current and estimated original ICV in the whole group. The Mann-Whitney U was used to test for gender differences between percentage brain tissue in both measurements.

8.3 Results

The median current ICV for the whole group was 1409.1 ml and the median estimated original ICV was 1480.1 ml. The median difference between the ICV measurements was 108.54 ml representing a percentage median difference of 7.3%, which was significant across the whole group ($z=-6.33$; $p<0.001$). Quartile ranges are presented in Table 8.1.

Measurement	**Group**	**N**	**Median**	**25th percentile**	**75th percentile**
Current ICV (ml)	Whole group	60	1409.10	1229.44	1648.04
	Male	31	1643.46	1480.38	1698.79
	Female	29	1228.54	1199.03	1265.87
Estimated original ICV (ml)	Whole group	60	1480.16	1355.49	1747.41
	Male	31	1741.12	1602.81	1820.38
	Female	29	1354.75	1290.5	1381.23
Absolute difference in ICV (ml)	Whole group	60	108.54	86.74	131.18
	Male	31	101.98	79.42	131.92
	Female	29	114.61	88.31	129.33
% Brain tissue in current ICV	Whole group	60	78.84%	75.04%	81.40%
	Male	31	76.67%	74.29%	79.50%
	Female	29	80.93%	78.05%	82.48%
% Brain tissue in estimated original ICV	Whole group	60	73.10%	70.63%	76.04%
	Male	31	71.71%	70.26%	75.02%
	Female	29	74.08%	71.41%	76.32%
Absolute difference between % brain tissue	Whole group	60	5.29%	4.41%	6.70%
	Male	31	4.74%	3.48%	5.50%

in current and estimated original ICV	Female	29	6.55%	5.07%	7.63%

Table 8.1. Median and interquartile ranges of ICV and brain atrophy measures for the whole group, men and women.

Statistics for Wilcoxon Signed Ranks and Mann-Whitney U tests are shown in Table 8.2.

Difference between current and estimated original ICV (Wilcoxon signed rank test)	Whole group	Z = -6.334; p<0.001
	Males	Z = -4.457; p<0.001
	Females	Z = -4.541; p<0.001
Difference in current and estimated original ICV between males and females (Mann-Whitney U test)	-	Z = -0.718; p>0.05
Difference in current and estimated original ICV between males and females after correcting for head size using current ICV (Mann-Whitney U test)	-	Z = -3.523; p<0.001
Difference in % brain tissue in current and estimated original ICV (Wilcoxon signed rank test)	Whole group	Z = -6.334; p<0.001
Difference in % brain tissue in current ICV between males and females (Mann-Whitney U test)	-	Z = -3.188; p<0.001
Difference in % brain tissue in estimated original ICV between males and females (Mann-Whitney U test)	-	Z = -1.280; p>0.05

Table 8.2: Test score (Z) and p-value for the difference (Mann-Whitney U and Wilcoxon Signed Rank) between measurements within groups and within measurements between groups.

The median current and estimated original ICV in men (current ICV=1643.5 ml and estimated original ICV =1741.1 ml) were larger than in women (current ICV=1228.5 ml and estimated original ICV=1354.8 ml). The absolute difference between current and estimated original ICV was greater for women 114.6 ml (z =-4.541; p<0.001) than for men 101.9 ml (z =-4.457; p<0.001), confirming that women showed a greater

decrease in ICV due to inner table skull thickening (8.3%) than did men (6.2%). However, when considering the difference between current and estimated ICV for men and women without adjusting for head size, the result was not significant ($z = -0.718$; $p>0.05$). Only when individual differences in head size were accounted for by correcting for current ICV prior to comparison, did the difference between current and estimated original ICV between men and women became significant ($z = -3.523$; $p<0.001$).

We assessed the percentage of total brain tissue in current (78.8%) and estimated original ICV (73.1%) as a measure of brain atrophy. The difference between the percentage of total brain tissue in current and estimated original ICV for the whole group (5.3%) was significant ($z = -6.334$; $p<0.001$). The percentage of total brain tissue in current ICV in men (76.7%) and women (80.9%) was significantly different ($z = -3.188$; $p<0.001$) making it appear that women had less brain atrophy than men. However, the percentage of total brain tissue in estimated original ICV between men (71.7%) and women (74.1%) was not significantly different ($z = -1.280$; $p>0.05$) indicating that there were no sex differences in the degree of brain atrophy once the effects of inner table skull thickening had been removed.

The reliability of all measurements was calculated using intra-class and inter-class correlation coefficients. The intra-class correlations for the current ICV and estimated original ICV measurements were 0.98 and 0.98 respectively; and the inter-class correlations for the current ICV and estimated original ICV measurements were 0.96 and 0.98 respectively.

8.4 Discussion

Thickening of the inner skull table occurs with ageing and can significantly affect the measurement of ICV. The reduction in ICV is more pronounced in women, who on average have more inner table skull thickening than men, but the difference in head size between men and women artificially distorts the magnitude of this difference. The finding of significantly more inner table skull thickening in females (8.3%) than males (6.2%) is consistent with the literature on physiological changes of the skull

with age (Harding et al, 1949). The significant difference in current ICV measurements between men and women reflects the fact that men have larger heads than women, but the disappearance of the significant difference when the effect of skull thickness is removed highlights that the extent of skull thickening is greater in women than in men.

The influence that inner table skull thickening can have on estimating brain atrophy is significant and could be a potential problem for estimates of brain tissue loss in ageing. Our findings show that gender differences in atrophy are affected by inner table skull thickening and may be obscuring true differences in brain atrophy with age (Coffey et al, 1998; Xu et al, 2000), or in estimates of original brain size in youth to compare with cognition in old age. This is an important consideration when attempting to approximate whether men or women suffer more age-related brain tissue atrophy. The importance of using ICV as a covariate has been aptly displayed in the study by Scahill et al (2003) which looked at brain volume changes in normal ageing, and found that significant gender effects were lost when ICV was used to correct for differences in head size. The mechanism behind inner table skull thickening may be related to brain tissue loss in old age and although its cause is unclear, it remains an unexplored source of variance in gender comparisons and ageing research.

The strengths of this study are that we used exemplar subjects chosen to represent a range of degrees of inner skull table thickening from a well characterised older cohort. We performed all analyses blind to all other subject information and used well-validated image processing tools and checked the outputs visually and corrected any erroneous tissue inclusion/exclusion manually. We made careful attempts at standardising the measurement and increasing the reproducibility by choosing an easily identifiable inferior boundary, aligning all images to the AC-PC line.

The limitations include the relatively small sample, the subjective decision necessary to delineate the original inner skull table, which was difficult to identify consistently in people with any moderate inner table skull thickening (Figure 8.2), and uncertainty about the generalisability to other populations in which the degree of inner table skull thickening may be even greater. Clearly, having identified the potential scale of the

problem, it is now necessary to direct future research towards finding ways of estimating original ICV reliably.

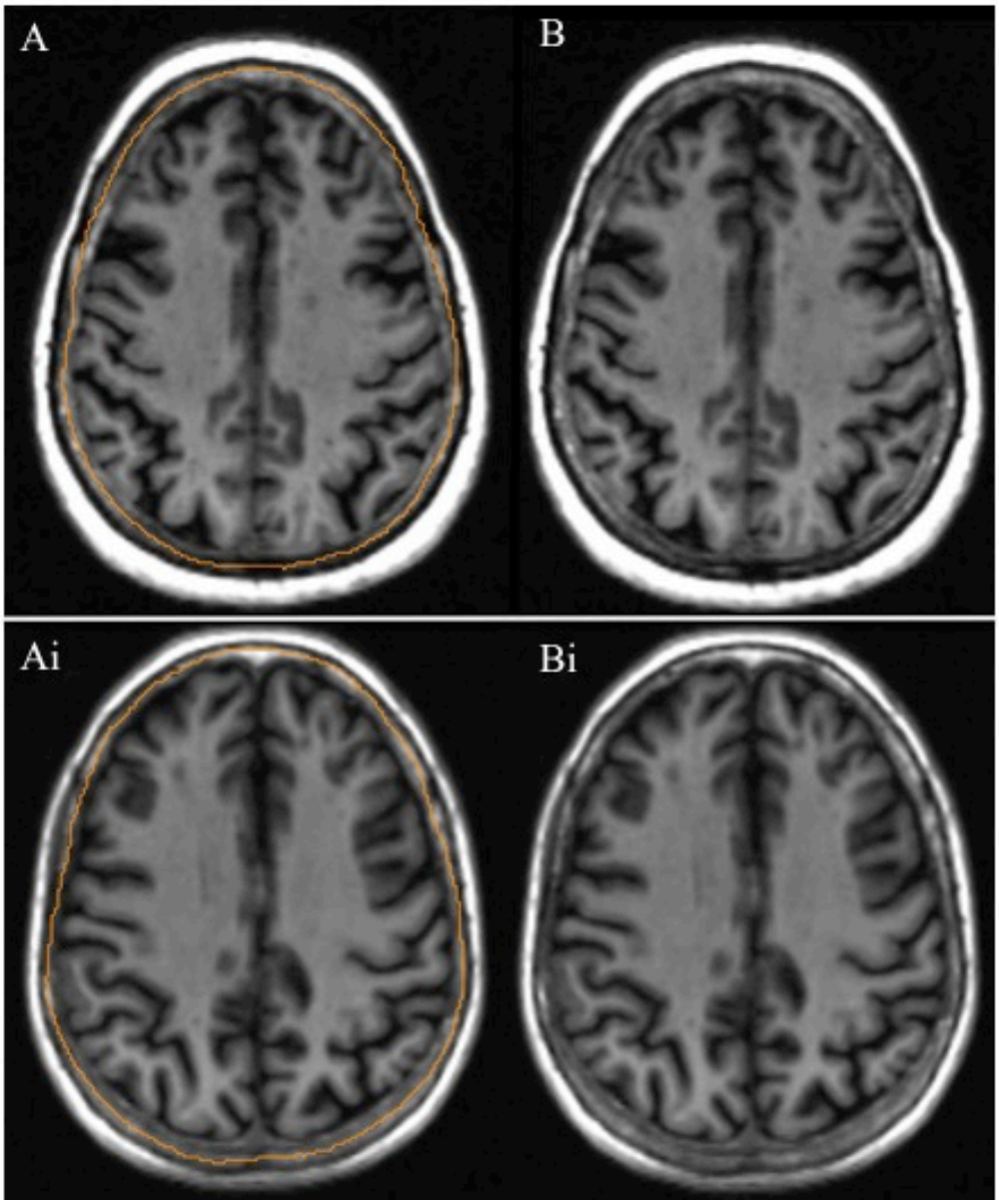


Figure 8.2: Identification of the boundary at which to estimate skull thickening (Orange line in A and Ai) in a person with moderate (B) and little (Bi) inner table skull thickening.

Although the pathological implications of inner skull table thickening may not be of great medical research interest due to its presumed benign nature, the influence this thickening has on estimates of brain atrophy in studies of ageing and potentially in clinical practice should be considered. Inner table skull thickening, while not present in all older people, can have considerable influence on estimates of brain atrophy, especially in woman. These findings call attention to an otherwise overlooked aspect

of research into brain ageing and further informs research concerning gender differences. Though individual differences in inner skull table thickness are difficult to extrapolate from this relatively small sample, the demonstration that large degrees of thickening are present in older people is important. Presentation of results showing that these differences could obscure gender differences in brain atrophy further strengthens the implications of our findings.

In an attempt to find a faster method by which the influence of inner table skull thickening on ICV measurement can be assessed the work presented in Chapter 9 was conducted. The basic mathematical principal of calculating the volume of a sphere using the circumference has been utilised by researchers in trying to obtain a quick brain volume measurement. The successful application of this to measuring the contents of the skull was extended to see if this principal could also be useful when measuring the influence of inner skull thickening.

Chapter 9. Measuring inner table skull thickening using ICA

9.1 Introduction

Intracranial cross sectional area, discussed in Chapter 7, has been shown to be a good proxy measure for obtaining a quick ICV measurement. Therefore it seemed reasonable to investigate if a measurement of ICA could be adjusted to account for estimated skull thickening, as it could be especially useful in large scale studies where time constraints due to sample size, can make applying more detailed but labour intensive methods unfeasible. The difficulty with applying an ICA method to estimating skull thickening is that thickening tends to be less prominent in the midline sagittal region, and it is hard to say whether it is uniform across the skull therefore a single measurement may not reflect the full extent of the thickening. The aim of this research was to investigate a method of accounting for skull thickening within an ICV measurement but by adapting the ICA method to better account for thickening across the skull.

9.2 Method

The subjects were the same as those used in Chapter 8 to assess skull thickening, as such they were all T1W images realigned to the AC-PC line, but with the additional step of resizing the images to make them Isotropic. This was to ensure a clear view of the coronal acquired T1W image in the sagittal plane. This can reduce clarity of structures due to the interpolation, or estimation between slices, necessary to make the image isotropic.

To obtain intracranial area measurements that cover the whole brain, the midline sagittal slice was chosen as a starting point and manually traced according to Ferguson et al (2005). Ten slices from the midline were counted moving laterally and the cranial vault traced on the resulting slice. This was then repeated until 5 measurements from the midline were completed; the same process was performed in the opposite hemisphere to complete 11 measurements covering the whole brain (Figure 9.1).

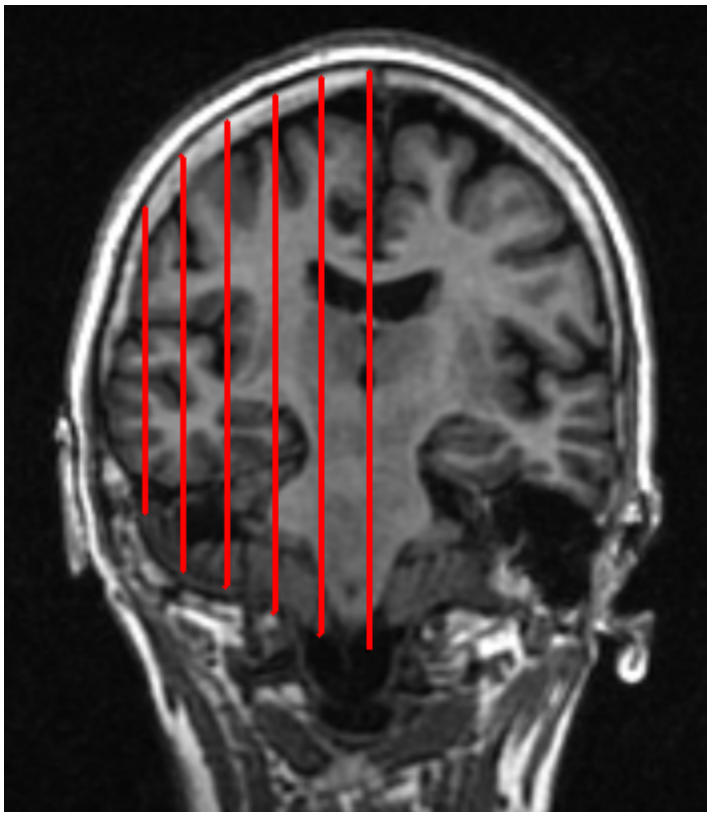


Figure 9.1: Coronal T1W slice showing the points (red vertical lines) at which sagittal ICA measurements were made in the left hemisphere.

The same protocol is followed as for the ICA measurement, however as the inner table is reached above the orbital socket (or glabella) the line is taken into the diploe where the estimated skull increase is visible and continued to the bregma point (Figure 9.2).

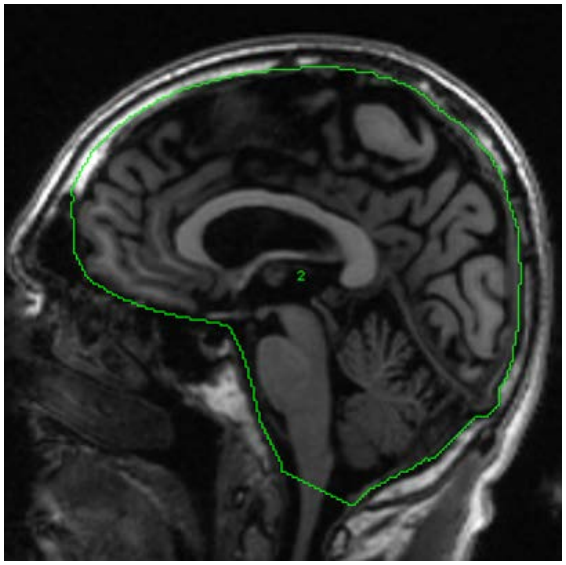


Figure 9.2: ICA measurement on a sagittal T1W image, taking account of inner table skull thickening.

This is completed on 10 evenly spaced slices across the brain, in order to cover both hemispheres and attempt to capture variation in inner table skull thickening across the whole skull.

9.2.1 Statistical analysis

Normality distribution analysis was performed using Shapiro-Wilks. Agreement between the ICV method and ICA method was assessed using Spearman Rho correlation. Percentage mean difference was used to look for differences between ICA and ICA without skull thickening.

9.3 Results

Shapiro-Wilks analysis revealed that neither the ICA (0.37, p = 0.00) or ICA with skull thickening (0.36, 0.00) were normally distributed.

The results obtained from the ICA method were assessed to see if they correlated well with the ICV method of assessing the effects of skull thickening, and Spearman rho correlations showed that the two methods were significantly correlated (r_s = 0.95, p = 0.00). This was also the case for the ICA and ICV with skull thickening, with a significant correlation of r_s = 0.95 (p = 0.00). These results show that the ICA method is a good alternative to ICV when assessing skull thickening, and could be utilised to investigate potential effects in a large dataset.

Comparison of the values for ICA and ICA without skull thickening, show a mean (std) difference of 3779 (1119) mm^3, indicating a change of 2.75% (0.93%). Small gender differences exist between these measurements with a mean difference of 3590 (1367) mm^3 between ICA and ICA with skull thickening in men, and 3918 (901) mm^3 between ICA and ICA with skull thickening in women. This equates to a percentage difference of 0.95% between men and women, which is around the percentage mean variance in the group as a whole. Despite this a Wilcoxon signed rank test revealed a significant difference between the ICA and ICA with skull thickening measurements in the whole group (Z = -5.511, p = 0.00), in men (Z = -3.621, p = 0.00) and in women (Z = -3.621, p = 0.00).

9.4 Discussion

An adapted intracranial area measurement was found to be comparable to a more extensive intracranial volume measurement when trying to assess the potential influence of skull thickening. The results found using this method are not in the same magnitude as those found in the more extensive skull thickening analysis presented in Chapter 8. However significant differences were still found to be present between measures that estimate ICA and measurements that estimate ICA after accounting for the effects of skull thickening.

The difference in results may be due to the difficulty in defining where the original skull table is when trying to adjust for the skull thickening. Although the diploe is relatively clearly defined on a sagittal T1W, the measurement is very dependent on the rater's ability to distinguish the inner skull table prior to age-related skull

thickening. The application of the glabella and bregma limits were designed to reduce discrepancies in judgement, as inferior to these points determining the skull can be problematic. Alongside the slightly reduced resolution and shadow cast by the interhemispheric fissure, a small reduction in volume estimation when making the ICA measurement may have occurred.

The main advantage of this method is the speed at which measurements of the whole brain can be made. Using this method, only 11 slices were used to make the ICA measurement, unlike with the more extensive method where 30+ slices are involved in acquiring a volume measurement. This method also has the potential to be used where whole brain coverage during acquisition is not possible, maybe due to a larger than average skull or where clinical research scans, which due to a shorter acquisition time may not cover the whole skull, need to be obtained. Whole brain coverage is also affected by which scanner is being used, it may not be possible to increase the field of view to increase the number of slices being acquired resulting in incomplete coverage of the skull. Other than speed the advantages of this method over the more extensive measurement method previously described are that identification of the inferior boundary at the foramen magnum is easier in the sagittal plane than in the axial view, resulting in a more consistent measurement of this boundary.

In conclusion this method could be effectively applied to a large number of individuals where inner table skull thickening is considered to be a significant variable in ICV measurement. ICA methods can be used where optimum scan acquisition for volumetric measurements is not possible or where slices are thick, as can be the case in studies involving patients due to the necessity for a short scan time.

Section 4: Regional measures

Chapter 10. Automated vs. Manual segmentation of the hippocampus.

10.1 Introduction

Rising interest in the hippocampus has led to more and more researchers using magnetic resonance imaging (MRI) to study its size and potential associations with memory (Wheeler et al, 2011; Erickson et al, 2011), post-traumatic stress disorder (Karl et al, 2006) and major depressive disorder (Lorenzetti et al, 2009). The connection with development of Alzheimer's disease has sparked the question as to whether early signs of hippocampal deterioration are linked with cognitive decline in older adults (Frisoni et al, 2010). Traditionally the hippocampus has been measured using manual segmentation methods from T1-weighted MRI, but this can be time consuming and requires highly trained raters (Watson et al, 1992). The increasing prevalence of automated software has made it possible for large numbers of participants to be measured in a relatively short amount of time. This allows for epidemiological studies of the hippocampus in normal older adults, especially when looking at associations with genetic data where very large participant numbers are necessary (Stein et al, 2012).

Automated methods rely on computational algorithms to produce segmentations thus reducing rater error by minimising user input (Khan et al, 2008). They utilise shape, contrast and/or intensity information to perform segmentations, relying on standard atlases and registration methods to guide the algorithm. They are widely used in a variety of different fields (Kim et al, 2012; Bergouignan et al, 2009; Joseph et al, 2012), commonly to demonstrate differences in hippocampal volume between control groups and patient groups (Baron et al, 2011; Jafari-Khouzani et al, 2011;) where distinct differences are expected. These factors would seem to suggest that a viable alternative to manual hippocampal segmentation has been found, however some concerning outcomes have been reported such as findings being significant depending on the segmentations in the analysis were from a manual or automated method (Sánchez-Benavides et al, 2010). Pardoe et al (2009) look at the performance of automated methods, compared to manual segmentation, at detecting hippocampal sclerosis in mesial temporal lobe epilepsy sufferers. They concluded that the automated methods (FSL-FIRST, FreeSurfer) were less sensitive than the manual method, furthermore the authors go on to suggest that were the hippocampal volume measurements alone relied upon to classify patients more cases would be missed using automated methods than manual methods.

The results of the systematic review in Chapter 3 informed the choice of methods investigated here. We applied two widely used, and freely available, methods to hippocampal segmentation of a large sample of normal older adults. Though there are many methods that can be used to segment the hippocampus the two methods were chosen specifically because of the ease at which they were available and the amount of comprehensive supporting materials that accompany them. FSL and Freesurfer are both available to download for free, have been cited as being used in many studies (Pardoe et al, 2009; Morey et al, 2009; Lim et al, 2012) and have excellent user guides, providing easy to follow workflows. We looked at validating the output from both these automated methods against manually delineated segmentations in a representative sample of the LBC1936.

10.2 Methods

The sample is the same as that described in Chapter 6 (6.2.1) and the image acquisition parameters as those detailed in Chapter 2 (2.4).

10.2.1 Reference method

Manual segmentations were made from T1W images in the coronal view, using Analyze 9.0. The LBC1936 scan acquisition is along the hippocampal long axis designed to improve visualisation of the hippocampus for segmentation purposes. Each image is individually thresholded to create greater separation between grey matter and CSF, and to make inclusion/exclusion of partial volume less user dependent. This process is described in detail in Appendix 3. The thresholded image is then loaded in the coronal plane and the borders delineated by drawing around the structure to include all tissue that is considered to be hippocampus. As the amygdala recedes the hippocampus begins to appear and the two are separated by the uncal recess of the temporal horn of the lateral ventricle (Figure 10.1).

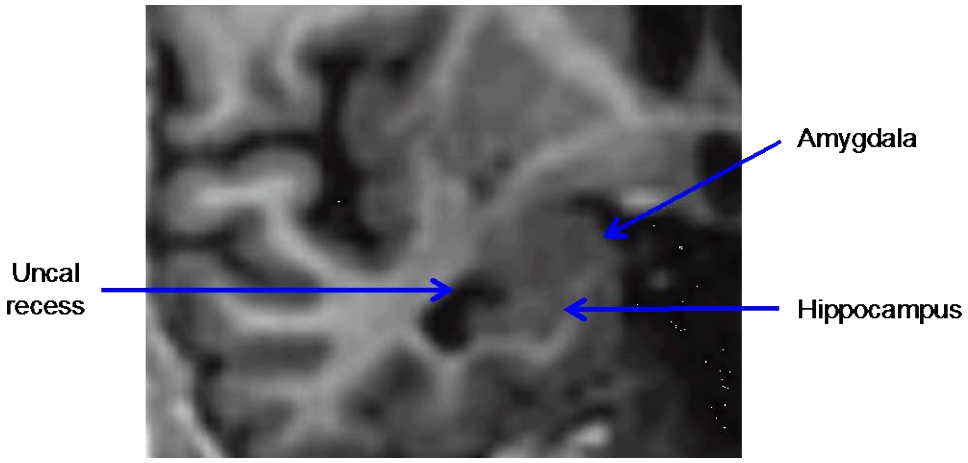


Figure 10.1. Shows the position of the uncal recess, amygdala and hippocampus.

The uncas is included in the hippocampal measurement where present and the body of the hippocampus is followed posteriorly through the slices until the tail is reached. The tail of the hippocampus can be seen as the pulvinar nucleus of the thalamus recedes and the fornix becomes apparent, this informs the choice of the last slice as it is considered to be the slice in which the entire length of the fornix extending superiorly and medially is visible, but has not become continuous with the corpus callosum (Figure 10.2). A full protocol is given in Appendix 3.

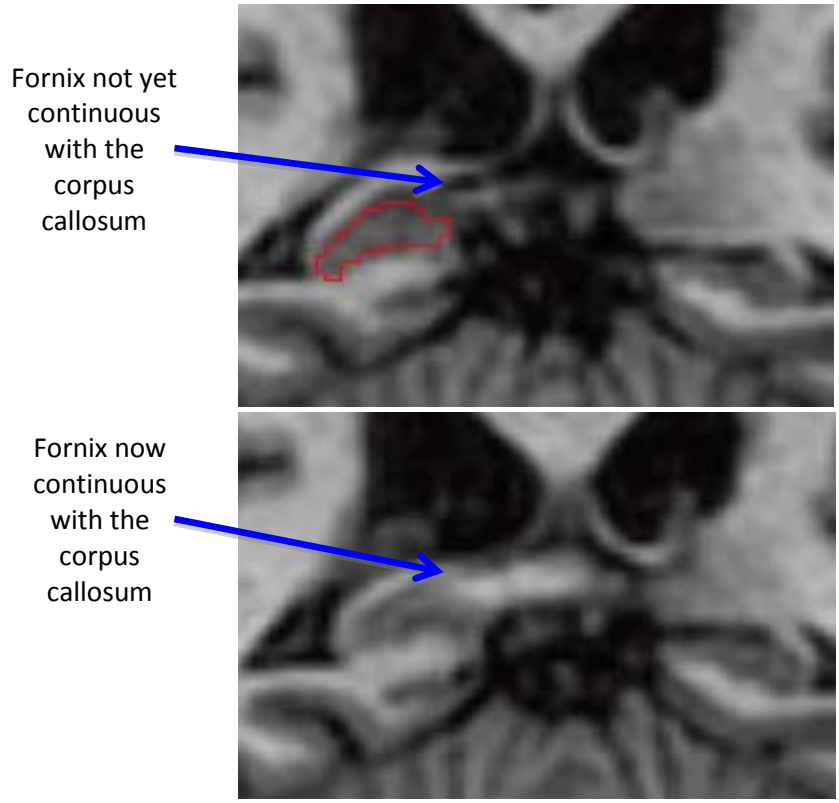


Figure 10.2. Shows the last slice in which the hippocampus is measured in the manual protocol used.

10.2.2 Automated methods

FSL_FIRST (http://www.fmrib.ox.ac.uk/fsl/first/index.html) is a model-based segmentation tool for subcortical structures of the brain such as the hippocampus. The shape model uses previously delineated manual segmentations to produce deformable meshes based on multivariate Gaussian assumptions, and then using these models tries to find the most probable shape from the T1W image. This mesh is then linearly transformed to standard space, as described in Patenaude et al (2011), then left and right hippocampal volumes were calculated.

10.2.3 FSL_FIRST protocol used in the LBC1936

In a preliminary analysis prompted by the opportunity to contribute to a large collaborative neuroimaging and genetics study, 30 randomly selected MR scans from the LBC1936 were segmented using FSL_FIRST. This was designed to determine the efficacy of using FSL_FIRST as an automated method of segmentation in the LBC1936 cohort.

Four versions of the FSL_FIRST protocol were compared; a Standard protocol, the Standard protocol with a template developed from older individuals, the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) protocol for FSL_FIRST and the ENIGMA protocol with prior application of BET following the ENIGMA protocol for obtaining brain and intracranial volumes (Figure 10.3).

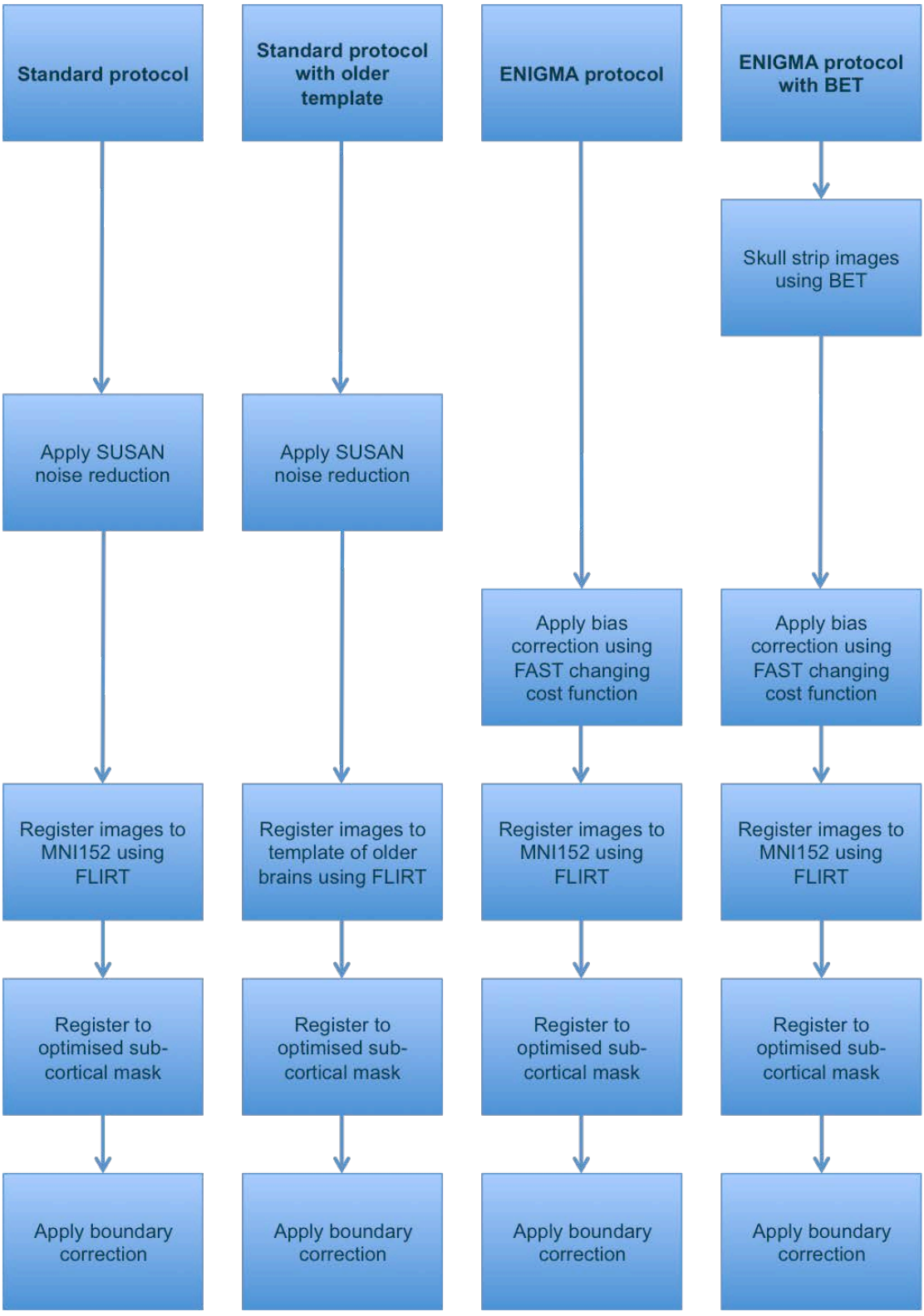


Figure 10.3: Stages performed for automatic segmentation of the Hippocampus using four different protocols of FSL_FIRST.

The Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) suggested protocol for image analysis of the Hippocampus was run on the same data set to assess how it compares to the Standard protocol. 8 of the 30 cases failed at the preliminary registration stage (10.4); therefore brain extraction (BET) according to the ENIGMA ICV protocol was applied prior to FSL_FIRST.

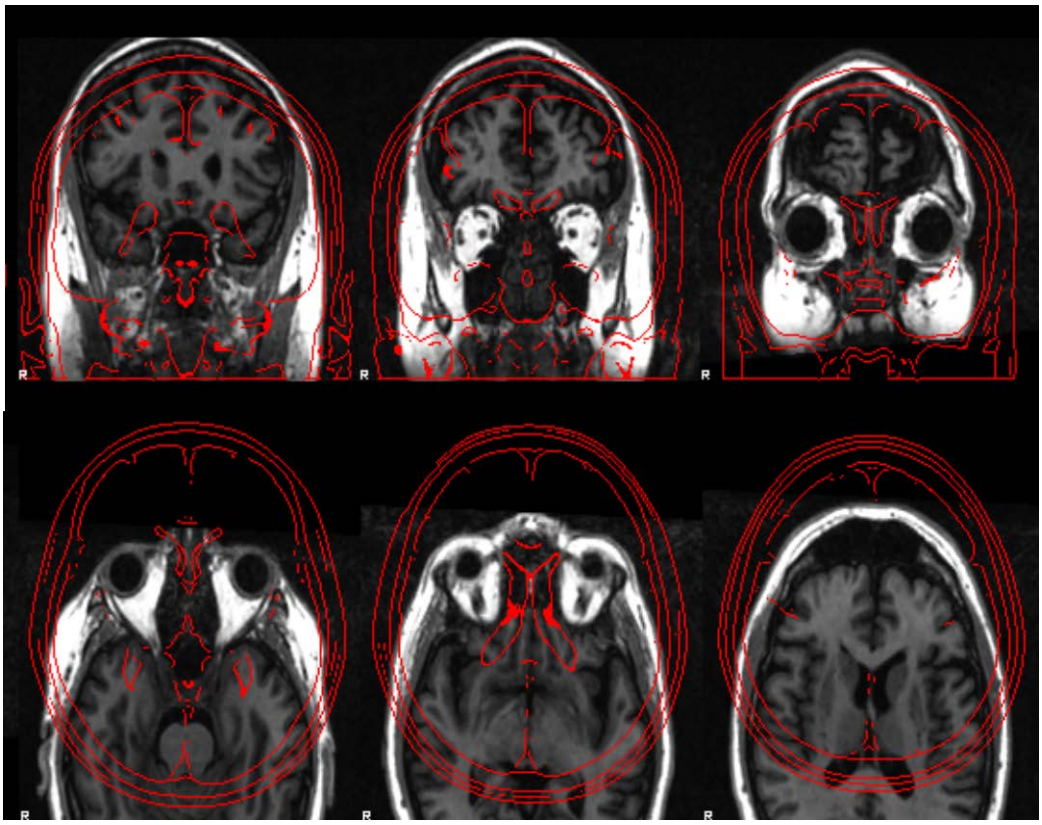


Figure 10.4: An example of a failed registration from the ENIGMA protocol without BET.

This improved the registration success with only 2 failures and the Bland Altman suggests this step also improved the hippocampal segmentation (Table 10.1). The volume measurements from these four protocols were compared with manually derived volume measurements of the same data set. Bland Altman analysis shows that the mean difference between the protocols and manual segmentation is between 16% and 27% (Table 10.1).

Protocol	Left Hippocampus				Right Hippocampus			
	Standard	**SOT**	**ENIGMA**	**EBET**	**Standard**	**SOT**	**ENIGMA**	**EBET**
Mean	20.08	23.23	22.02	27.08	18.54	20.37	15.93	24.20
LCL	-10.41	-4.39	-33.56	-3.67	-1.24	2.33	-32.9	-4.90
UCL	50.84	50.84	77.59	57.83	38.32	38.40	64.76	53.30

Legend: SOT – Standard Older Template; EBET – ENIGMA with BET; LCL – Lower Confidence Limit; UCL – Upper Confidence Limit

Table 10.1: Percentage mean difference with lower and upper confidence limits for each protocol compared to manual measurements.

Overall the Standard with Older template protocol showed the closest agreement for both hemispheres. The intervals for both ENIGMA protocols is wide, suggesting considerable discrepancies between the volume measures they provide and those derived from manual segmentation. As mentioned applying BET prior to the ENIGMA protocol improved the registration success as well as improving the degree of agreement. The mean difference is larger for the ENIGMA with BET protocol than ENIGMA alone however the confidence limits are narrower, showing less variation in the results. Despite this improvement to the ENIGMA protocol both the Standard protocols showed better agreement with manual segmentations. (Bland Altman plots for all protocols can be seen in Appendix 4).

The large degree of variation shown by the sizeable confidence limits reflects the small sample of 30, the following analysis shows the results from a much increased sample size.

Freesurfer (http://surfer.nmr.mgh.harvard.edu) uses a manually labelled data set and Markov random field models to estimate the probability that a voxel belongs to the

structures being targeted for segmentation. A probabilistic atlas is used to aid the algorithm in correctly identifying a target structure in relation to the whole brain and other subcortical brain regions. T1W scans were used to compute the left and right hippocampal volumes.

The output from both automated methods was visually inspected as an initial quality control step and to look for any cases that could be re-run if they had failed at the registration step. Both automated methods failed to successfully segment the hippocampus in all 150 images, therefore the sample size for comparison with the manual measurements differs slightly for the FSL (n=137) and Freesurfer (n=141) methods.

10.3 Results

To assess the performance of the two automated methods they were compared against a manually delineated reference standard of the same sample. The mean and standard deviation of the volume measurements (mm^3) from the three methods for left and right hippocampi can be seen in Table 10.2.

Measurement method	**Mean (std) volume mm^3**
Manual right	3284.96 (442.47)
Manual left	3134.99 (390.85)
FSL right	3800.07 (705.45)
FSL left	3648.60 (580.29)
Freesurfer right	3360.42 (398.60)
Freesurfer left	3337.99 (440.63)

Table 10.2: Mean and standard deviation for manual, FSL and Freesurfer for left and right hippocampus.

On average FSL over estimated the hippocampus by 513.61mm^3 in the left and by 515.11mm^3 in the right, and Freesurfer overestimated by 203mm^3 in the left and 75.46mm^3 in the right. The manual measurements show that the right hippocampus is larger than the left on average by 149.97mm^3, this was also the case for the FSL

measurements by 115.47mm^3 but not to the same extent for the Freesurfer measurements 22.43mm^3.

Bland-Altman analysis was performed to determine the difference between the automated and manual methods, and to give an indication of the measurement variance between methods. The Bland-Altman analysis shows that when comparing the volumes (mm^3) the largest mean difference was between the manual method and FSL in the right hippocampus (-511.64). The smallest mean difference was between the manual method and Freesurfer in the right hippocampus (-71.64) but the largest upper and lower 95% confidence intervals (CI) were also from this comparison. The FSL method and Freesurfer method for the left hippocampus performed comparably when compared to the manual method (Table 10.3).

Measurement method	**BA mean (std)**	**Upper 95% CI**	**Lower 95% CI**	**Pearson correlation**	**Jaccard Index mean (std)**
Manual right vs FSL right	-511.64 (601.79)	667.86	-1.69	r 0.53 p = 0.01	0.67 (0.09)
Manual left vs FSL left	-507.08 (582.10)	633.84	-1.65	r 0.34 p = 0.01	0.65 (0.09)
Manual right vs Freesurfer right	-71.64 (333.55)	582.12	-725.41	r 0.70 p = 0.01	0.56 (0.17)
Manual left vs Freeurfer left	-198.81 (435.38)	654.54	-1.05	r 0.45 p = 0.01	0.54 (0.18)
FSL right vs Freesurfer right	421.17 (1.68)	1.68	-836.97	r 0.46 p = 0.01	0.58 (0.15)
FSL left vs Freesurfer left	298.07 (619.28)	1.51	-915.71	r 0.29 p = 0.01	0.55 (0.17)

Table 10.3: Bland-Altman (BA) mean and standard deviations with upper and lower 95% confidence intervals (CI), Pearson correlation coefficient with significance and Mean (standard deviation) Jaccard Index values between the automated and manual methods, and between the manual and two automated methods for the left and right hippocampi.

The variance between measurements can be seen clearly on the scatter plots (Figure 10.5), which also indicates that the automated methods performed better when segmenting the right rather than the left hippocampus. Overall the Bland-Altman analysis suggests that the Freesurfer method more closely agrees with the manual method than the FSL method.

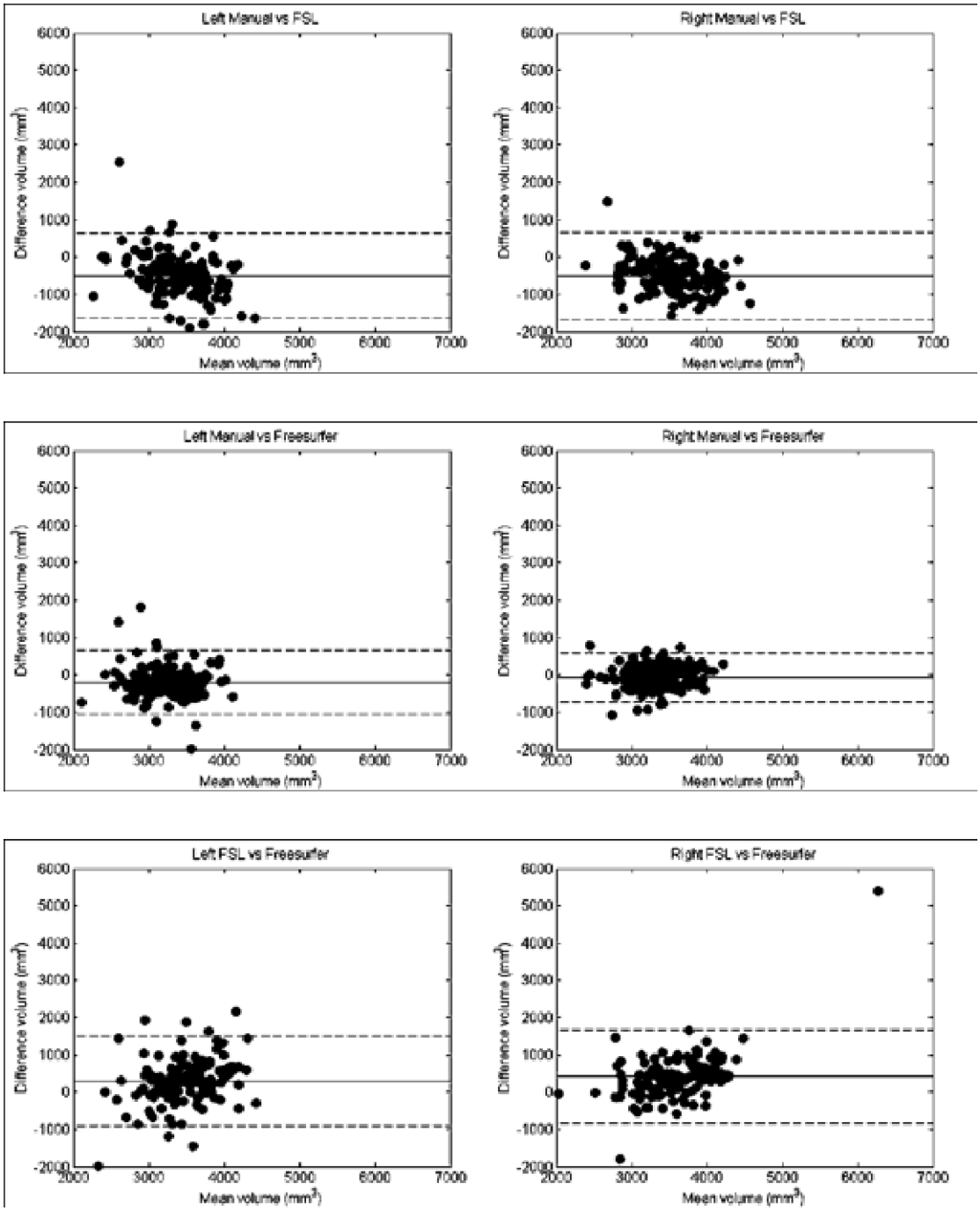


Figure 10.5: Bland-Altman plots showing the values between the automated and manual methods, and between the two automated methods for the left and right hippocampi.

To assess the spatial concordance of the two automated methods with the manual segmentations Jaccard analysis was performed; this provides an index score that indicates the similarity between two segmentations. The segmentation outputs are overlaid and the coefficient is calculated by using the size of the intersection between

two segmentations divided by the size of the union of the segmentations. Jaccard analysis was applied to the two automated methods to see how well they spatially agreed with the manual method and with one another.

The Jaccard Index scores show that the FSL method agrees best with the manual method for both left (0.65) and right hippocampi (0.67), compared to the left (0.54) and right (0.56) output from Freesurfer (Table 10.4). The Jaccard analysis suggests that FSL is more spatially concordant with the manual method than Freesurfer for both the left and right hippocampi.

Methods being compared	**Jaccard Index; mean (std)**	
	Right	**Left**
Manual*FSL	0.67 (0.09)	0.65 (0.09)
Manual*Freesurfer	0.56 (0.17)	0.54 (0.18)
FSL*Freesurfer	0.58 (0.15)	0.55 (0.17)

Table 10.4: Jaccard Index analysis showing agreement between the reference and automated methods and between the two automated methods.

To determine if brain tissue atrophy influenced the accuracy of the automated methods, the manual, FSL and Freesurfer volumes and Jaccard Index scores were correlated with total brain tissue volume (BTV) as a percentage of intracranial volume (%ICV). The inclusion of the Jaccard Index scores is to assess whether overall brain tissue atrophy is associated with the physical overlap of the masks. Pearson correlations revealed that none of the segmentation output volumes were associated with BTV%ICV (Table 10.5).

Method	**Pearson correlation**
Manual right	r = -0.041, p = 0.625
Manual Left	r = 0.115, p = 0.165
FSL Right	r = -0.070, p = 0.418
FSL Left	r = -0.128, p = 0.136
Freesurfer Right	r = 0.147, p = 0.083
Freesurfer Left	r = 0.049, p = 0.564

Table 10.5: Pearson correlation between BTV%ICV and the right and left manual, FSL and Freesurfer hippocampal segmentations.

Spearman rho correlations, chosen because the indices is between 0 and 1, showed that the Freesurfer Jaccard Index scores, with both the manual and FSL output, were significantly correlated with BTV%ICV (Table 10.6).

Method	**Spearman's Rho correlation**
FSL_Manual_Right	$r_s = 0.044$, $p = 0.635$
FSL_Manual_Left	$r_s = 0.022$, $p = 0.810$
Free_Manual_Right	$r_s = 0.292$, $p = 0.001$*
Free_Manual_Left	$r_s = 0.281$, $p = 0.002$*
Free_FSL_Right	$r_s = 0.369$, $p = 0.000$*
Free_FSL_Left	$r_s = 0.339$, $p = 0.000$*

Table 10.6: Spearman rho correlation between BTV%ICV and the right and left Jaccard Index scores for pairings between manual, FSL and Freesurfer. *Significant correlations indicated.

10.4 Discussion

Comparison of two commonly used and freely available automated segmentation methods with a manual reference standard method, in a group of relatively healthy older adults, has revealed that automated methods perform moderately well in this population. Pearson Correlation and Bland-Altman support the use of Freesurfer over FSL, suggesting it has performed marginally better at segmenting the hippocampus in older adults. However the Jaccard analysis, giving spatial similarities between the segmentation output, suggests that FSL has performed better at segmenting the hippocampus in our sample. Our analysis comparing both the volumetric data, as well as assessing the spatial concordance against a reference standard in a good size sample, provides a clear indication that when validating methods statistics that assess overlap or concordance between methods is vital. Analysing only the volumetric agreement does not provide a clear enough picture as to where the discrepancy between methods lie, as it does not adequately reflect concordance in the event that a method has both under and over segmented the same structure. If this occurs the

errors cancel each other out so that when the final numbers are calculated the average measurements seem to be closer to the reference standard than is actually the case.

Our findings disagree with previous publications that looked at these two automated methods for hippocampal segmentation (Morey et al, 2009; Doring et al, 2010) and found that Freesurfer performed better however the samples used in these studies had lower mean ages (37.2 and 35.4 years; 32.9 and 43.8 years respectively) than our sample (71.9 years) therefore it could be that FSL can more easily negotiate the difficulties associated with ageing than Freesurfer.

Both methods over segmented the hippocampus, though Freesurfer less so than FSL, this could have significant consequences if subtle differences or changes in hippocampal volume are expected. The true degree of atrophy would be underestimated, which for example when using volumetric data to detect group classification could lead to clinically significant hippocampal atrophy being missed (Pardoe et al, 2009). The right-larger-than-left asymmetry evident from the manual measurements and previous research (Pedraza, Bowers and Gilmore, 2004) was apparent from the FSL segmentations but not with Freesurfer. This could in part be due to the fact that the Freesurfer method did not include an age appropriate template therefore asymmetry, more pronounced in older groups, may have been lost when transforming data to common space. The potentially significant consequences of losing this information due to measurement method is concerning, especially where volume measurements are being analysed to look for associations with cognitive (Woolard and Heckers, 2012) or clinical measures (Wolf et al, 2001). It is not known whether asymmetry of function in the hippocampus (Burgess, Maguire & O'Keefe, 2002) is associated with the asymmetrical pattern of atrophy found here but it could help to explain why variation in associations are sometimes reported.

The introduction of an age appropriate template to the FSL_FIRST method was taken to try to improve the drawbacks of using the supplied template of MNI 152, as this template was averaged from 150 normal young adults. This improved the segmentation output, resulting in a higher concordance with the manual segmentations. The development of age appropriate atlases or normative image databases are beneficial in providing both data on which algorithms can be tested and

resources such as atlases, which improves existing automated methods (Evans et al, 2012). The results from the review in Chapter 3 suggest that the main causes of discrepancy between manual and automated methods came from registration errors that in turn could have been due to an unrepresentative atlas being applied to the segmentations. However cohort atlases, developed from the study sample, were more successful at dealing with age-related variation in brain images. Therefore it would seem sensible to use cohort over standard atlases, unless large population specific atlases were developed and made available for use in much the same way that the Harvard or MNI atlases are.

The limitations of this study are chiefly in the use of only two automated methods to compare with manual segmentations however apart from being impractical to run several more methods, the two methods chosen represent what is freely available to the average researcher. They are also extensively supported by comprehensive literature, workflows and online tutorials making them relatively easy to use. From the review in Chapter 3 these two methods were the most commonly reportedly, therefore they are representative of those used in the wider community. Another limitation is that of inserting an age relevant atlas into the FSL_FIRST pipeline making comparison with the standard Freesurfer program biased. This is acknowledged but served to investigate the potential effect that using such a template could have on the segmentation outcome. Due to time and processing constraints it was not possible to check the same effect with the Freesurfer program, ideally this study would be repeated with an older adult template being substituted for the standard atlas supplied in both automated methods.

This analysis has revealed that although automated methods represent an attractive alternative to laborious manual segmentations, there use is not without compromise and this may be to an unacceptable level when investigating hippocampal volume size in healthy older adults. Refinements in registration methods and age specific templates can only improve the performance of automated methods when applied to older brains. For the time being manual segmentation remains the most sensitive method of achieving accurate hippocampal segmentations on the ageing brain but an acceptable alternative may be to manually edit all automated segmentations per a standardised manual protocol.

Chapter 11. Practical application of frontal lobe boundary protocols

11.1 Introduction

Unlike disorders whose biological cause can be used to predict potentially vulnerable locations in the brain, such as Parkinson's or Multiple sclerosis, normal age-related decline is not biologically characterised. The pattern of decline is not uniform across all individuals, making it difficult to know which area of the brain to focus on when trying to assess age-related structural brain changes. The frontal lobes have garnered a lot of interest from researchers trying to determine the brain structure that most significantly contributes to decline in old age. Some findings suggest that the frontal lobes are the first to show age-related tissue decline, especially in frontal grey matter (Fjell et al, 2010). Others state that the evidence to support the frontal lobe being differentially affected by ageing is weak and findings often conflict (Greenwood et al, 2000). Generally the evidence for a specific age-related pattern of brain deterioration is inconsistent, with regions such as the hippocampus and factors like cortical thinning being highlighted but not in every study (Raz et al, 2010). There are many explanations for why researchers seem unable to resolve the conflicting evidence regarding the frontal lobe, but a possible reason could be discrepancies between the protocols used to measure frontal lobe volume.

The traditional way of assessing frontal lobe volume, or any regional brain volume, is to make manual measurements using hand tracing. Manual tracing of a structure requires good anatomical knowledge with which to apply a well researched and considered protocol. However a recent review (Cox et al, 2014) highlighted the considerable variation in segmentation protocols in the published literature, finding 19 methods using 15 different landmarks to measure the frontal lobes. These findings suggest that the structure being termed the frontal lobe may not in fact be anatomically the same tissue depending on the protocol applied. The possible implication of variations in measurement protocols is significant, as these methodological differences could confound potential associations between brain

structure and function. Understanding how structural brain changes relate to cognitive function is paramount in trying to understand what underpins ageing trajectories. Poor cognitive function in old age has been correlated with deterioration in brain structure (Raz & Rodrigue, 2006), therefore accurate measurements of both metrics is essential.

Topographical landmarks such as sulcal patterns are the most accurate way of anatomically identifying the frontal lobe but they are difficult, variable and time-consuming to apply. Boundaries based on sulcal landmarks give a better concordance with the underlying cellular architecture of a structure (Zilles & Amunts, 2010), therefore should be used where possible. However, the limitations just mentioned have resulted in researchers finding alternative approaches. The most common alternative approach is the use of geometrical cut planes to measure the frontal lobe. Geometrical cut planes utilise one or several anatomical landmarks to apply straight boundary limits to determine the frontal lobe, for example the use of a coronal plane at the first appearance of the splenium. This type of method has the advantage of being quick to apply when determining a large, anatomically variable region such as the frontal lobe. The natural irregularity of the brain can mean identifying sulcal patterns when defining a region, leading to greater rater error due to subjective assessments of anatomy. Using geometrical cut planes should help to reduce this variability and subjectivity when making measurements (Lacerda et al, 2003). However the counter-point to these benefits is that cut plane methods are less able to reflect inter-individual variations in brain anatomy, especially where the individual has been affected by age-related changes or disease. The resolution of accuracy versus reliability is a constant concern in image analysis but even more so when structural volume measurements are then associated with other metrics such as cognition. The immense variability in boundary definitions used to determine frontal lobe volume only adds to this already complex but important area of research.

As a way of understanding the effects that these discrepancies of boundary definition may have on volumetric measures, a selection were applied to ten randomly chosen images from the larger LBC1936 cohort. As the exact method of segmentation from every paper cannot be applied for practical reasons, a selection of boundaries were chosen as they represented those that most commonly occurred in the papers reviewed. These boundaries were then applied to a randomly chosen subsample of the

LBC1936 cohort, in order to compare the volumetric results of the different boundaries in the same set of brains. The aim of this study was to assess the potential degree of variation between a selection of published frontal lobe boundaries.

11.2 Methods

11.2.1 Participants

Ten T1W scans were chosen at random from the LBC1936 cohort, to which the five boundaries were applied, totalling 100 measurements. The mean (std) age was 72 (0.26) years old. The participants displayed a range of atrophy, white matter lesions and in one case a right sided, fronto-parietal arachnoid cyst. It was decided that the scan with the arachnoid cyst would be included to see how this would influence the application of boundaries in difficult cases. This case can still be considered representative of normal ageing as incidental findings, i.e. asymptomatic intracranial abnormalities, are common where MRI is used in research. A systematic review suggested that prevalence of brain incidental findings is 2.7% but that this increases with age, and that the most common incidental finding reported was an arachnoid cyst followed by meningioma (Morris et al, 2009).

11.2.2 Image analysis

Image acquisition parameters are detailed in Chapter 2 (2.4). T1W scans were chosen as they give the best contrast between grey and white matter, necessary for detecting boundaries between brain structures. As the LBC1936 scans are acquired oriented along the hippocampal long axis the scans were realigned to the anterior-posterior commissure (AC-PC). This is to provide a more consistent view of the frontal region, making identification of landmarks reliable from person to person.

11.2.3 Boundary definitions

The boundaries were chosen as they represented those boundaries reported in the systematic review of frontal lobe parcellation methods, which had been previously conducted (Cox et al, 2014). Three cut plane methods; the genu, the anterior

commissure and the optic chiasm, and two sulcul methods; the central sulcus and the pre-central sulcus were selected. From the 19 methods used in the 208 reviewed papers, forty-five reported using the genu to identify the posterior boundary of the frontal lobe, four reported using the anterior commissure, two the optic chiasm, seventeen the central sulcus and 31 the pre-central sulcus (Figure 11.1).

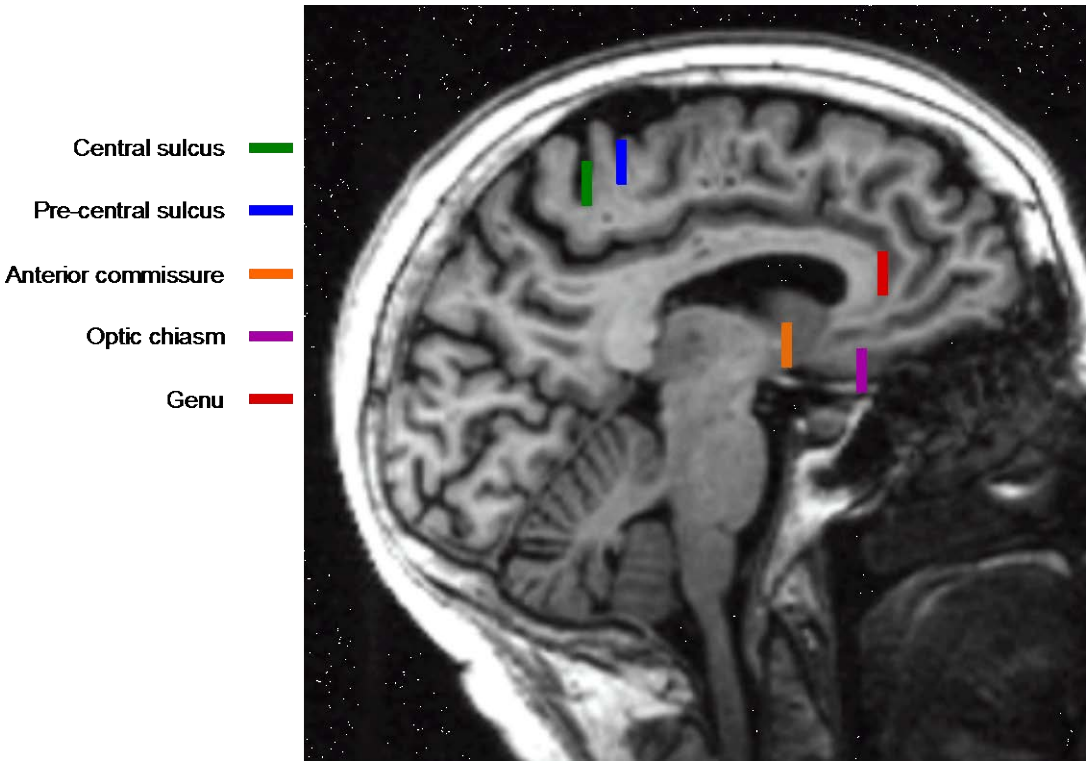


Figure 11.1: Indicates the posterior landmarks given to apply the five frontal lobe boundary protocols on a sagittal T1W image.

All measurements were made using the ROI module in Analyze 11.0, however MRIcro was used to aid boundary identification due to the ability to view all three orthogonal planes simultaneously using this program. The left and right hemispheres of the frontal lobe were measured, allowing for analysis of asymmetry of brain structures to be explored. For the three cut plane boundaries the anatomical landmark was identified in the sagittal plane and the measurements made on the corresponding coronal slice.

For the two sulcul boundaries the appropriate sulcus was identified in the axial view, as this allowed for a top down view of the sulci. Measurements were then made in both the axial and sagittal plane; discussion of each boundary is described below.

11.2.4 Cut–plane boundaries

The genu is the most anterior portion of the corpus collosum and forms a bend that goes down and backwards in front of the septum pellucidum. It is part of the largest bundle of white matter fibres in the brain (corpus collosum), which has contralateral axonal projections, and is distinct from surrounding structures (Figure 11.2). The anterior commissure is a bundle of white matter fibres connecting the two hemispheres across the midline. It is an easily identifiable structure on a T1W image (Figure 11.2). The optic chiasm is the point at which the optic tracts cross; they are located below the hypothalamus. (Figure 11.2).

Figure 11.2: Sagittal T1W image showing the landmarks used to identify the corresponding coronal slice from which to measure the frontal lobe in all anterior slices.

For the anterior commissure and genu boundaries the landmarks were identified in the midline sagittal plane and the corresponding coronal slice was counted as the most posterior boundary for the measurement. The coronal view was used to identify the optic chiasm, which once identified, all brain tissue anterior to and including this slice were measured as the frontal lobe. For all three cut plane measurements a very similar method was used. As per the majority of papers that reported using these boundaries

all brain tissue anterior to and including the slice at the boundary was included in the frontal lobe measurement. The Temporal lobes were disconnected where the temporal stem joins it to the rest of the cortex, by drawing a line from the Sylvian point to the uppermost point of the grey matter of the temporal lobe, this continues medially following this grey matter. To separate the hemispheres a straight line was drawn directly down the midline of the image in every slice, ensuring that the dura was not included in the measurement.

Cerebrospinal Fluid (CSF) is treated as a separate object, therefore it was not counted in the volume calculation for the frontal lobe. The other advantage of this is that the volume of CSF measured within the frontal lobe can be used as an estimated measure of atrophy when calculated as a percentage of ICV. However this is only a measure of subcortical atrophy, as it does not reflect atrophy occurring on the cortical surface.

11.2.5 Sulcal boundaries

The central sulcus is the anatomically correct posterior boundary of the frontal lobe and this is most effectively identified in the axial plane, as this provides the best visualisation of brain topography. The central sulcus was identified by looking for the characteristic omega shaped gyrus, whose corresponding sulcus ran significantly deeper than surrounding sulci (Figure 11.3).

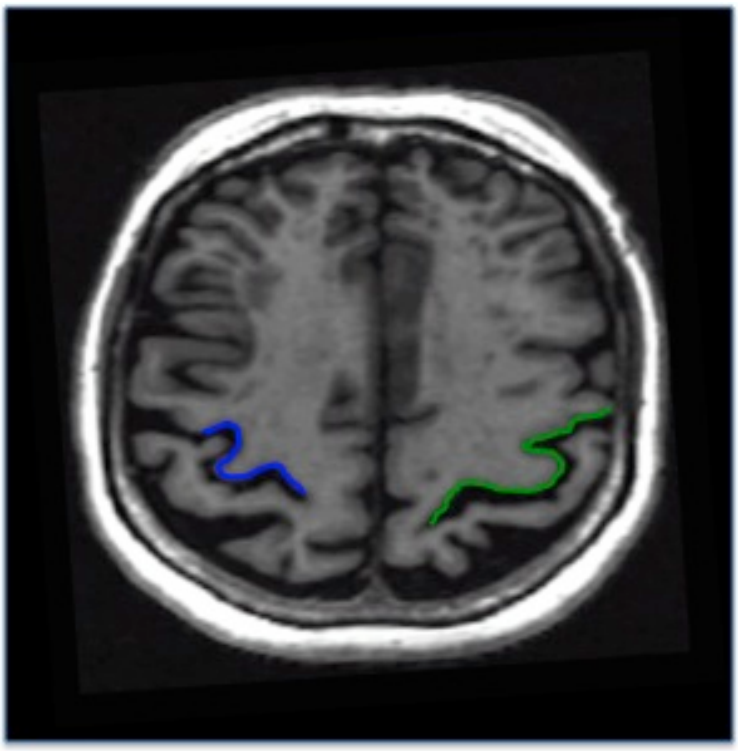


Figure 11.3: Shows the central sulcus (green), the deepest running sulci, and the characteristic omega shape (blue).

This landmark was followed inferiorly until it met the sylvian fissure. Where the Sylvian fissure did not reach the interhemispheric fissure a straight line connecting this point to the interhemispheric midline CSF was drawn (Figure 11.4).

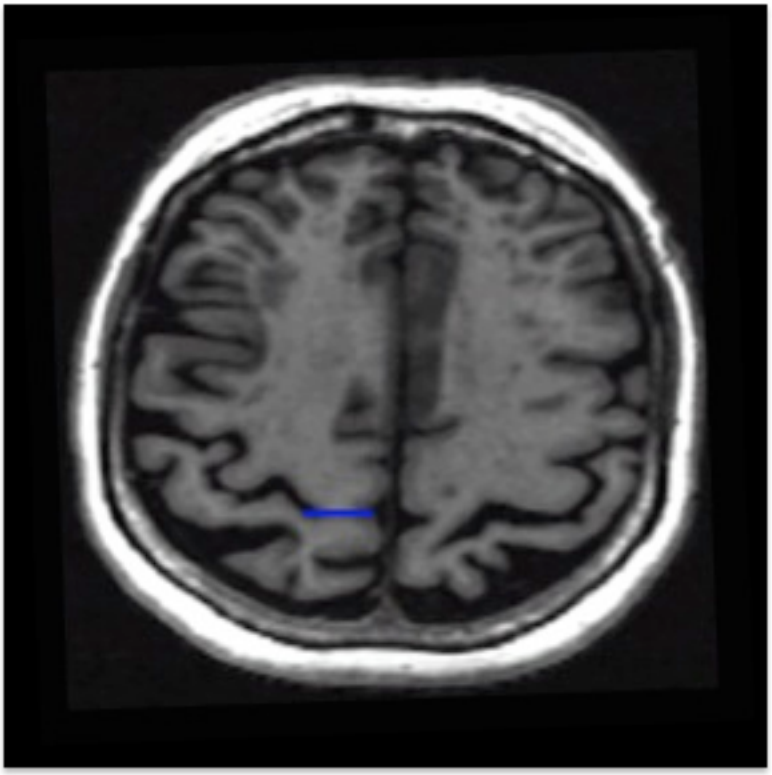


Figure 11.4: Displays the straight line drawn (blue) to continue the central sulcus to the interhemispheric fissure.

The precentral sulcus is often used as an alternative measure of the frontal lobe and this is sometimes termed the prefrontal lobe, consisting of the inferior, middle and superior gyri. Identification of the precentral sulcus was achieved in an axial and sagittal orientation depending on which was clearer, and is the next sulcus anterior to the central sulcus (Figure 11.5).

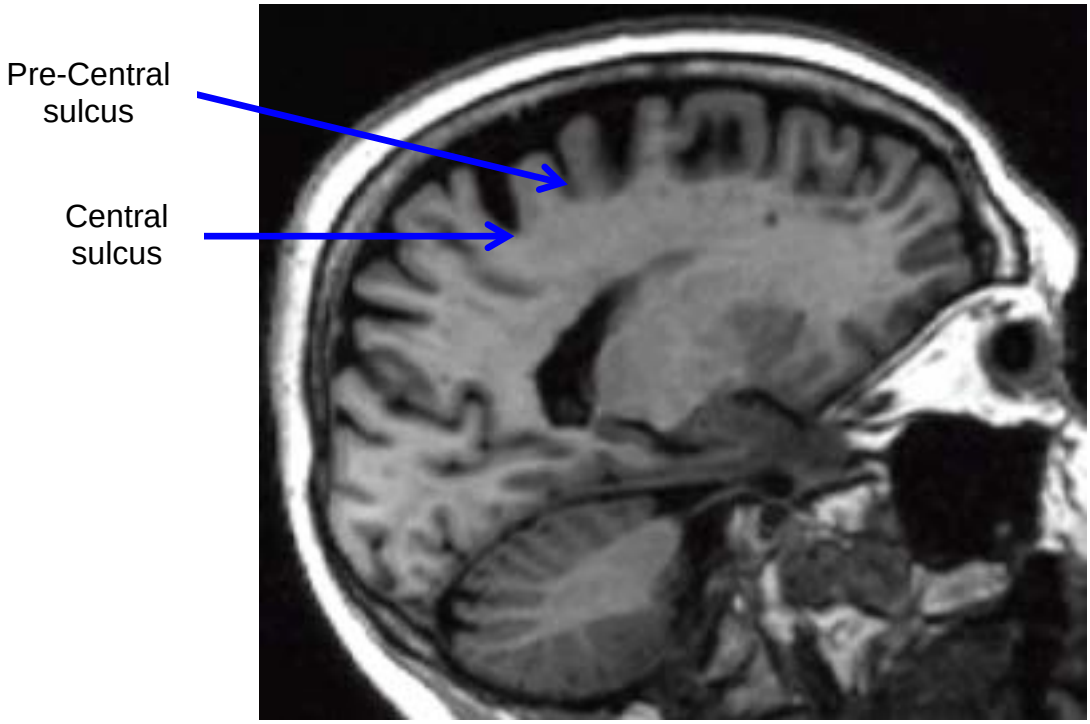


Figure 11.5: Shows the central and pre-central sulci on a T1W sagittal image.

The same procedure as for the central sulcus was adopted. In both cases the cingulate was excluded from the frontal lobe measurement, this was completed on a sagittal plane by following the line of the cingulate sulcus posteriorly until in intersects with the descending central or precentral sulcus. Where the sulci did not meet a straight line was used to define the boundary.

11.3 Results

Shapiro Wilks and Q-Q plots revealed that all of the data was normally distributed, except the total Central Sulcus (S-W= 0.901, (21), p= 0.036) measurement that was not normally distributed (Figure 11.6).

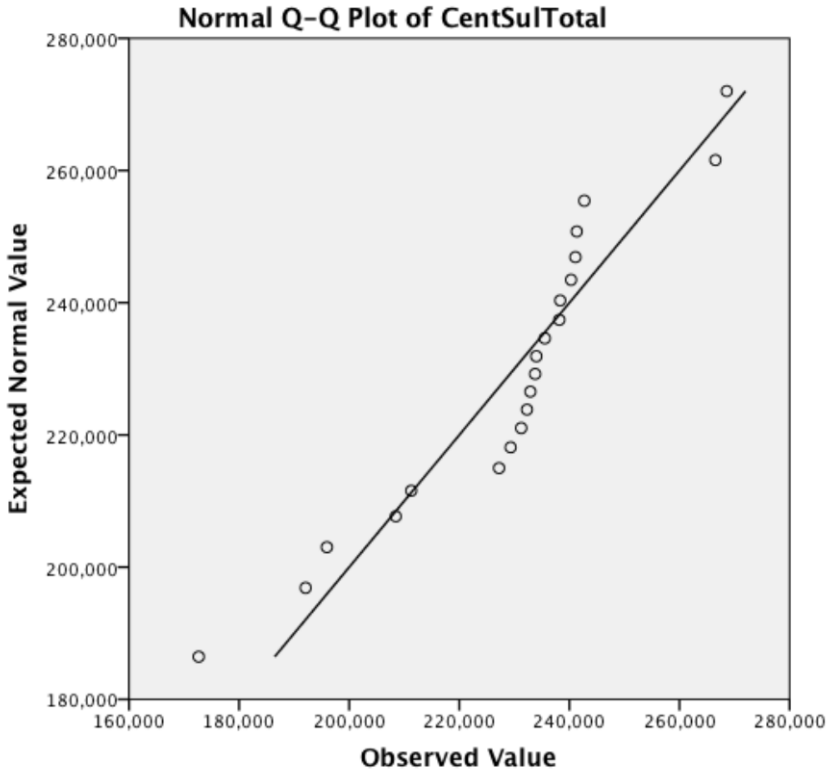


Figure 11.6: Q-Q plot showing that the combined Central sulcus (CentSul) measurement was not normally distributed.

The boundary providing the largest mean/median total frontal lobe volume was the central sulcus (229227/233778 mm^3) and the boundary giving the smallest volume was the genu (81909/72347 mm^3). The results show that between the boundary with the largest measured volume and that with the smallest a mean/median difference of up to 147318/161340mm^3 for total, 72122/78785 mm^3 for right and 75196/81636 mm^3 for left frontal lobe volume was found (Table 11.1).

Boundary	**Total Mean vol mm^3 (Std)**	**Right Mean vol mm^3 (Std)**	**Left Mean vol mm^3 (Std)**
Genu	81909 (23236)	44682 (9415)	37227 (11429)
Anterior Commissure	207808 (32097)	109337 (15842)	98471 (16760)

Optic Chiasm	211501 (37048)	11411 (18905)	100090 (18651)
Central Sulcus	229227 (22639)	116804 (13093)	112423 (11073)
Pre-Central Sulcus	199471 (20317)	101114 (12165)	98357 (9345)

Table 11.1: Mean volume and standard deviation (std) for each frontal lobe boundary measurement.

The large degree of variability between boundaries can be clearly seen in Figure 11.7, which shows each boundary measurement for each participant. The variation between participants reflects individual differences in brain morphology but the variation between measurements for each person, shows the variation introduced by choice of boundary.

All data except the total central sulcus measurement met the assumptions necessary to perform paired sample t-tests, therefore further bootstrap analysis was applied to the pairings including the total central sulcus. Paired sample t-tests were used to look for differences between each measurement, resulting in analysis of every possibly pair of measurements e.g. genu and optic chiasm, genu and central sulcus etc. Furthermore to explore hemispheric asymmetries the measurements were compared for total, right and left frontal lobe differences.

Results show that for the total frontal lobe volumes only the AnteriorCommissure* OpticChiasm (t= -0.640, p= 0.530), AnteriorCommissure*PreCentralSulcus (t= 1.280, p= 0.215) and the OpticChiasm*PreCentralSulcus (t= 1.965, p= 0.064) were not significantly different. For the right frontal lobe volumes only the AnteriorCommissure*OpticChiasm (t= -0.710, p= .486) and OpticChiasm* CentralSulcus (t= -1.499, p= 0.150) were not significantly different from one another. For the left frontal lobe volumes only the AnteriorCommissure* OpticChiasm (t= -0.567, p= 0.577), AnteriorCommissure*PreCentralSulcus (t= 0.036, p= 0.972) and

OpticChiasm*PreCentralSulcus (t= 0.578, p= 0.570) were not significantly different. All values are reported in Table 11.2.

Boundary pairing	Combined right left volume	Right volume	Left volume
Genu * AntComm	t = -22.816 (0.000)	t = -22.840 (0.000)	t = -22.515 (0.000)
Genu * OptChia	t = -13.581 (0.000)	t = -13.562 (0.000)	t = -13.554 (0.000)
Genu * CS	t = 19.702 (0.000)	t = -17.962 (0.000)	t = -20.451 (0.000)
Genu * PreCS	t = -16.524 (0.000)	t = -14.260 (0.000)	t = -18.315 (0.000)
AntComm * OptChia	t = -0.640 (0.530)	t = -0.710 (0.486)	t = -0.567 (0.577)
AntComm * CS	t = 3.180 (0.005)	t = -2.025 (0.056)	t = -4.214 (0.000)
AntComm * PreCS	t = 1.280 (0.215)	t = 2.334 (0.030)	t = 0.036 (0.972)
OptChia * CS	t = 2.905 (0.009)	t = -1.499 (0.150)	t = -4.461 (0.000)
OptChia * PreCS	t = 1.965 (0.064)	t = 3.051 (0.006)	t = 0.578 (0.570)
CS * PreCS	t = 16.983 (0.000)	t = 15.856 (0.000)	t = 13.708 (0.000)

Legend: AntComm – Anterior commissure; OptChia – Optic Chiasm; CS – Central Sulcus; PreCS – PreCentral Sulcus.

Table 11.2: Paired sample T-test for all possible pairings of variables. For those pairings including the combined Central sulcus, bootstrap was applied.

To assess the intra-rater reliability of the methods one scan was selected and each boundary applied as before but blind to the previous segmentation. ICCs were then calculated and revealed that across all measurements the rater reliability was very good at 0.99. With just two measurements per boundary it was not possible to assess ICCs within each boundary measurement but the difference between the first and second volume for each boundary as a percentage of the total group mean value for each measurement was calculated (Table 11.3). This reveals that measurements made using the cut plane boundaries (Genu, 0.37%; AntComm, 0.07%; OptChia, 3.84%) resulted in less percentage mean difference between two volumes in the same brain, than the sulcul boundaries (CS, 5.90%; PreCS, 6.19%). This variation in repeat measurements was more apparent in the left hemispheric volumes where the cut plane methods showed a difference from the mean of between 0.02 – 4.12% and the sulcul measures showed a difference of more than 10% from the mean.

Boundary	**Total % mean vol (mm^3)**	**Right % mean vol (mm^3)**	**Left % mean vol (mm^3)**
Genu	0.37%	0.33%	0.40%
Anterior Commissure	0.07%	0.12%	0.02%
Optic Chiasm	3.84%	3.59%	4.12%
Central Sulcus	6.02%	1.96%	10.24%
Pre-Central Sulcus	6.18%	1.54%	10.94%

Table 11.3: Percentage mean difference of repeat measures for each boundary.

The time taken to perform each boundary measurement was recorded to try to establish if in fact cut plane boundary measurements are faster to perform than following sulcal boundaries. The measurements using the genu as the posterior boundary took on average of 23 minutes to complete, the anterior commissure 27, the optic chiasm 31, central sulcus 108 and the pre-central sulcus 101 minutes. The time taken to perform the sulcal measurements is considerably longer than that taken to perform the cut plane measurements.

11.4 Discussion

This study found that volume measurements made using various boundary definitions of the frontal lobe were significantly different in a group of older community-dwelling adults. The smallest resultant volume (derived using the genu) represented only 35% of the largest total frontal lobe volume measured, meaning that the boundary definition used would eclipse any potential differences found between groups or other brain volume measures. The applied boundaries were selected from published literature, all of which performed some measure of frontal lobe volume in a range of populations (Cox et al, 2014). Our results suggest that comparing findings from studies that apply such considerably different protocols to determining the same structure, will lead to inaccurate conclusions. It undoubtedly hinders researchers efforts to investigate areas of potential vulnerability to age-related decline, as the conflicting findings may in fact be down to measurement differences and not biological variations in samples.

Although the central sulcus is anatomically the correct posterior frontal lobe boundary, the results from this study show it is one of the most variable boundaries to apply in practice. Both sulcal boundaries were found to be the least repeatable of those used, resulting in more variation between repeat measurements. The inability to utilise 3-Dimensional software to apply boundaries, was a hindrance in this study but the software used was representative of the general image analysis packages regularly in use. Difficulty in consistently identifying the central sulcus in inferior slices, especially the anterior limit of the supplementary motor area, has been suggested as a reason for using the precentral sulcus as a posterior marker (Lyoo et al, 1998; Pantel

et al, 1997). However our results show that the precentral sulcus is as variable to apply as the central sulcus measurement, and almost as time consuming. It is no small concern that a method takes a long time to perform; it may simply be unfeasible to employ such a technique especially with a large number of scans. Therefore a cut plane method provides an attractive, and in some instances, the only alternative to acquiring volume measurements of the frontal lobe. Nevertheless producing an anatomically plausible measurement is crucial to understanding structural brain differences between patient groups or associations with cognitive ability, only then can real progress in research be made.

The basic, but crucially important, trade-off between time and repeatability in a measure is one that requires serious consideration. The appearance of software that allows measurements to be made using a 3-D image is a step toward resolving this problem, as better visualisation of boundaries should lead to more consistent application. Automated methods have also been heralded as a possible way to improve both speed and repeatability of segmentations. Most automated methods are either based upon standard atlases or use a small number of manual tracings to predefine boundaries prior to segmentation. However the problem of boundary definition still applies, the atlases or manual segmentations used by the automated methods still require a protocol by which to demarcate the frontal lobe.

The necessity for a standardised protocol to be applied in all studies measuring frontal lobe volume is evident. The results from this study reveal that comparing findings across studies would be incredibly difficult and means that interpretation of findings in the published literature should be undertaken with caution. It is hoped that this study has highlighted a methodological problem in image analysis that may be contributing to the noise found in published literature on the role of the frontal lobe in ageing. Future work to develop a unifying protocol by which measurements can be made, can only help to further progress in this area.

Section 5: Associations with cognition

Chapter 12. Brain size and cognition

12.1 Introduction

As individuals age, some degree of decline is typically observed in the mean test scores of cognitive abilities such as reasoning, memory, processing speed and spatial ability (Salthouse, 2010; Ghisletta et al, 2012). A review by Plassman et al. (2010) found that, although the evidence for the contribution to this decline was in many factors, health problems, negative lifestyle choices (especially smoking), and possession of the *APOE*-ε4 allele were consistently associated with an increased risk of age-related cognitive decline.

Important factors not considered by Plassman et al. (2010) are decline or changes in brain tissue volume. This is surprising as larger brain volume, estimated by magnetic resonance (MR) imaging (MRI), has been associated with higher intelligence, with reported associations ranging from 0.33 to 0.42 across studies (McDaniel, 2005; Rushton and Ankney, 2009; Miller and Penke, 2007). Further, studies have shown, for example, direct links between reduction in brain tissue volume and cognitive decline (e.g. Sluimer et al, 2008); that decline in brain tissue volume and increased cerebrospinal fluid (CSF) longitudinally are associated with lower cognitive performance (e.g. Cardenas et al, 2011); and that tissue loss is accelerated in people suffering from mild cognitive impairment compared to normal decline with ageing (e.g. Driscoll et al, 2009).

However, several factors remain unclear in the associations between brain status and cognitive ability in later life. Firstly, research has been equivocal as to whether prior maximal brain size, as measured by ICV, or current brain status, as measured by current tissue volume, is the stronger predictor of later life cognitive ability. Maximal brain volume is established in late childhood and thereafter is reflected broadly in the internal size of the cranial cavity (Wolf et al, 2003), which can be measured using intracranial volume (ICV) from structural MRI. The volume of the intracranial cavity

is directly related to brain growth in youth and, although brain volume begins to decline in early adulthood, the cranial cavity is considered to remain relatively stable thereafter (Sahin et al, 2007). A number of studies have shown positive associations between intracranial volume (ICV) and cognitive ability in later life. MacLullich et al. (2002) report positive significant correlations, ranging from 0.26 to 0.39, between ICV and of several individual cognitive ability tests. Shenkin et al. (2009a) found, in 107 general healthy older adults aged between 75 and 81 years, that whole brain volume accounted for little (<1%) variance in general cognitive ability, whereas intracranial area (a proxy measure of ICV), explained 6.2% of the variance. More recently Farias et al. (2012) found that ICV and current brain volumes (i.e. total brain, hippocampal and white matter lesions) in older adults were associated with different cognitive domains, but that ICV correlated with cognitive variables after the other brain volume measures were accounted for. These findings suggest that maximal brain size is an important factor in understanding late life cognitive ability against which to determine the effects age-related pathological processes such as brain tissue loss and normal and pathological cognitive decline.

Secondly, despite research evidence suggesting it to be the strongest predictor of later life ability (e.g. Deary et al, 2012), few studies investigating the association between contemporaneous brain volumes and cognitive ability, control for prior ability. Of those studies that have (e.g. Staff et al, 2004; Murray et al, 2011), no support has been found for the prediction of contemporaneous cognitive performance by brain volumes when controlling for prior ability.

Thirdly, brain tissue volume decline in the normal ageing brain is not uniform; grey matter may begin declining in early adulthood and follow a fairly linear pattern thereafter, whereas white matter volume is said to increase until around middle age and then starts to decrease (Ge et al, 2002; Ziegler et al, 2010). Further, research has suggested that the divergence between volumetric changes in grey and white matter tissue volume is reflected in the associations between these tissues and cognitive performance (Raz et al, 2010).

Thus, in order to provide an examination of the associations between brain size and cognitive ability in ageing, in the present study we test whether maximal brain size in youth (as represented by ICV) and current brain tissue volume significantly contribute to cognitive ability in later life. We then test the contributions to both current cognitive ability and change in cognitive ability of white and grey matter volumes separately. Collectively we aimed to provide a thorough assessment of the relative contributions of broad measures of the brain and cognitive performance in a large, age homogenous sample of generally healthy older adults.

12.2 Methods

12.2.1 Participants

Details of the participants, cognitive testing, image acquisition and image analysis methods can be found in Chapter 2. Due to the inclusion of cognitive data in the analysis in this chapter, a further, 41 participants were excluded for incomplete cognitive data (age 11 IQ = 36; individual cognitive assessments at Wave 2 assessment = 5), and 11 participants were excluded as they scored ≤24 on the Mini Mental State Examination. A score of ≤24 the MMSE is a widely used clinical cut-off considered to be indicative of possible pathological cognitive impairment (Folstein, Folstein and McHugh, 1975). The final sample consisted of 620 adults (327 males, 52.7%).

Based on recent studies (O'Bryant et al, 2008; Stephan et al, 2008), and in order to test the robustness of the models described below, we also estimated all models using a more conservative MMSE cut-off of ≥28. This was done to ensure the results were not overly influenced by cases with lower MMSE scores, and thus potentially also those with mild cognitive impairment or early stages of dementia. In this secondary analysis, a further 83 participants (male=57; female=26) were removed resulting in a sample of 537 (270 males, 50.3%).

12.2.2 Statistical Analysis

All models were estimated using multi-group structural equation modelling (MG-SEM). MG-SEM has the advantage of modelling latent constructs, and the associations between them, taking account of measurement error, whilst providing formal tests of their equivalence across groups. As a result, MG-SEM provides highly robust estimates of associations. All models were estimated in Mplus 6.0 (Muthen and Muthen, 2010) using maximum likelihood estimation.

For the current study, we used the multi-group model to estimate parameters for males and females separately. Sex is known to be one of the largest sources of variability in overall head and brain size, and given our large sample; we chose to model this directly rather than simply include sex as a covariate. Input data for all models were standardized residuals of (age 11 IQ), after regressing out, and thus controlling for, variance associated with age. Despite the narrow age range of the current cohort, age still accounted for significant ($p<0.05$) amounts of variance in brain tissue volume, white matter volume, grey matter volume, block design, digit symbol coding, symbol search and matrix reasoning; demonstrating the importance of accounting for the effects of age.

12.2.3 Model Specification

In model 1, we included ICV and TBV as predictors of general cognitive ability (g) to assess the degree of association between measures of maximal brain size, current brain status and current cognitive ability. Next, we included age 11 IQ as a predictor of g in order to assess whether ICV and TBV remain significant predictors of current ability, controlling for past ability. As such, we were asking whether ICV and TBV also predict change in cognitive ability over the life course. In model 2, we follow the same sequence of analyses but replace TBV with white matter (WM) and grey matter (GM) volumes, to explore whether associations with specific tissue types are consistent with g in both males and females. In all analyses, TBV, WM and GM were standardized residuals controlling for age at MRI and ICV.

12.2.4 Measurement Invariance

Prior to testing the equivalence of the regression parameters across males and females, measurement invariance was established for the latent constructs. Measurement invariance ensures that the latent constructs are equivalent across groups, and is required in order to make meaningful interpretation of model parameters across groups (French and Finch, 2006). We established configural invariance (equivalence of the pattern of factor loadings), and metric invariance (degree of factor loadings), for both g and processing speed.

Once measurement invariance has been established, parameters of interest within the models can be fixed to equivalence, and the plausibility of this constraint is tested using the difference in chi-square for the appropriate number of degrees of freedom.

12.2.5 Model Evaluation

In SEM, the degree to which a model conforms to the data is assessed using model fit indices. We adopted cut-off points based on a review (Schermelleh-Engel, Moosbrugger, and Muller, 2003) of ≤ 0.05 for the standardised root mean square residual (SRMR), ≤ 0.06 for the root mean square error of approximation (RMSEA), and ≥ 0.95 for the Tucker-Lewis Index (TLI) and Comparative Fit Index (CFI). In establishing whether the assumptions of measurement invariance for the latent constructs hold, we follow Chen (2007) and suggest changes in CFI of ≤ -0.01 combined with changes in RMSEA ≤ 0.015 support measurement invariance.

12.3 Results

The descriptive statistics are shown in Table 12.1. All variables were approximately normally distributed with no values for skew exceeding ± 0.90, or values for kurtosis exceeding ±1.17 in either the males or females, with the exception of the processing speed indicators. In particular, simple reaction time showed moderate to high levels of kurtosis in both males (3.96) and females (5.19).

	Mean		Standard Deviation		Skew		Kurtosis	
	Male (n=327)	Female (n=293)	Male (n=327)	Female (n=293)	Male (n=327)	Female (n=293)	Male (n=327)	Female (n=293)
Age (Years)	72.47	72.60	0.70	0.73	0.10	-0.06	-0.84	-0.85
MMSE	28.75	29.04	1.27	1.10	-0.93	-1.17	0.25	1.18
Cognitive Ability Age 11 years								
Age 11 IQ	99.72	102.98	16.17	13.39	-0.90	-0.67	1.12	0.49
Cognitive Ability Age 73 years								
Digit Span Backward	7.85	8.04	2.27	2.24	0.34	0.30	-0.03	-0.30
Block Design	35.96	32.93	10.53	8.85	0.28	0.56	-0.28	0.76
Letter-Number Sequencing	11.07	11.07	3.06	2.74	0.36	0.44	0.32	0.51
Matrix Reasoning	14.09	13.00	4.82	4.75	-0.19	-0.03	-0.89	-0.90
Digit Symbol Coding	54.78	58.95	12.17	11.27	0.19	0.22	-0.36	-0.16
Symbol Search	24.74	25.15	6.23	5.69	-0.31	-0.16	0.55	0.93
Estimated Childhood Brain Volume								
ICV (cm^3)	1536.93	1355.39	113.20	101.08	0.25	0.28	-0.15	-0.13
Brain Volume Age 73 years								
Total Brain Tissue Volume (cm^3)	1175.08	1070.08	100.03	83.70	0.16	0.14	-0.07	-0.18

White Matter Volume (cm^3)	522.22	468.04	84.87	68.89	0.46	0.17	0.20	0.31
Grey Matter Volume (cm^3)	521.59	476.49	71.80	62.36	0.14	0.01	1.17	-0.13

Table 12.1: Descriptive Statistics for Cognitive Ability and Brain Imaging Variables in the LBC1936 (n=620).

Table 12.2 contains the uncorrected bivariate correlations. The significant positive correlations between the cognitive ability tests in males (r=0.29 to 0.65) and females (r=0.20 to 0.53), and between the three experimental processing speed tasks in both males (r=0.25 to 0.54) and females (r=0.22 to 0.35), supports the modelling of latent constructs. Total brain tissue volume (male r=0.18 to 0.34; female r=0.10 to 0.25), and white (male r=0.13 to 0.26; female r=0.04 to 0.18) and grey matter (male r=0.04 to 0.26; female r=0.01 to 0.13) volumes show universally positive correlations with cognitive ability tests. Age shows several significant associations with both cognitive and brain volume variables in males (r= -0.18 to 0.13) and females (r= -0.23 to 0.11), supporting its inclusion as a covariate, despite the very narrow range in this cohort.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Age	-	-0.14*	-0.05	-0.03	-0.02	0.05	-0.17**	-0.14*	-0.03	-0.13*	-0.23†	0.11
2. Age 11 IQ	-0.04	-	0.44†	0.40†	0.34†	0.41†	0.35†	0.35†	0.12*	0.11	0.11	-0.01
3. Block Design	-0.18**	0.42†	-	0.53†	0.29†	0.31†	0.38†	0.42†	0.19**	0.25†	0.15**	0.07
4. Matrix Reasoning	-0.12*	0.41†	0.49†	-	0.33†	0.30†	0.28†	0.24†	0.09	0.14*	0.14*	0.01
5. Digit Span Backward	-0.11*	0.33†	0.29†	0.35†	-	0.47†	0.20†	0.18**	0.07	0.11	0.04	0.06
6. Letter-Number Sequencing	-0.11	0.39†	0.33†	0.38†	0.56†	-	0.27†	0.21†	0.07	0.10	0.08	0.04
7. Digit Symbol	-0.16*	0.46†	0.47†	0.43†	0.38†	0.51†	-	0.58†	0.11	0.22†	0.13**	0.13**
8.Symbol Search	-0.13*	0.41†	0.48†	0.42†	0.37†	0.41†	0.65†	-	0.06	0.17**	0.18**	0.04
9. ICV	-0.01	0.28†	0.17**	0.11*	0.14**	0.11	0.21†	0.26†	-	0.83†	0.51†	0.42†
10.Brain Volume	-0.13*	0.26†	0.25†	0.18**	0.23†	0.22†	0.33†	0.34†	0.82†	-	0.58†	0.52†
11.WM Volume	-0.21†	0.18**	0.18**	0.13*	0.19**	0.24†	0.26†	0.19**	0.51†	0.61†	-	-0.07
12. GM Volume	0.13*	0.11*	0.14*	0.04	0.06	0.04	0.18**	0.26†	0.36†	0.48†	-0.09	-

Table 12.2: Pearson's correlations between independent, dependent and covariate variables.

$^{*}p<0.05$; $^{**}p<0.01$; $^{\dagger}p<0.001$

Table 12.2 also provides the associations between cognitive ability and head/brain size from youth and old age. To the extent that ICV may be considered an indicator of maximal brain size in youth, the correlation between ICV and age 11 IQ provides an estimate of brain-cognition associations in youth. In the current sample, these estimates are 0.28 (p<0.001) for males and 0.12 (p<0.05) for females. The correlations between latent g and concurrent brain volume provide similar contemporaneous associations at age 73. In the current sample the g-brain volume correlation is 0.27 (p<0.001) for males and 0.26 (p<0.001) for females; highly comparable to the ICV-age 11 IQ association.

It is also interesting to note that total brain volume and white matter volume have several moderately significant associations with individual cognitive ability subtests in both males and females, whereas the associations with grey matter volume are largely small and non-significant. However, the correlations presented in Table 12.2 are raw correlations, uncorrected for ICV and age, and as such differ to some extent from the estimates presented in the final models.

Finally, Table 12.2 also provides estimates of the associations between ICV and concurrent brain volume, which correlate highly in both males and females (0.82, p<0.001 and 0.83, p<0.001, respectively). Grey matter shows weaker associations with both ICV and brain tissue volume than did white matter in males (ICV: 0.36 versus 0.51; brain volume: 0.48 versus 0.61) and females (ICV: 0.42 versus 0.51; brain volume: 0.52 versus 0.58).

Prior to testing the main models, we first established measurement invariance in the measurement model for g across males and females. The model showed excellent fit to the data (χ^2=22.73(12), p<0.05; CFI=0.99; TLI=0.97; RMSEA=0.054; SRMR=0.023), and the difference in model fit across the configural and metric invariance models fell within the suggested range of model fit (ΔCFI=0.00; ΔRMSEA=-0.004). The measurement model was therefore considered to be invariant across sex at the metric level, with all subsequent models run with the invariance constraints in place.

Models 1 and 2 displayed excellent fit to the data (See Figure 12.1 for final model fits). When all paths in the final models were sequentially constrained to equivalence across males and females, no chi-square differences reached statistical significance ($\Delta\chi^2 \geq 3.84$, p<0.05).

Therefore, in the current sample, there were no substantive differences in the magnitude of parameter estimates in the male and female models.

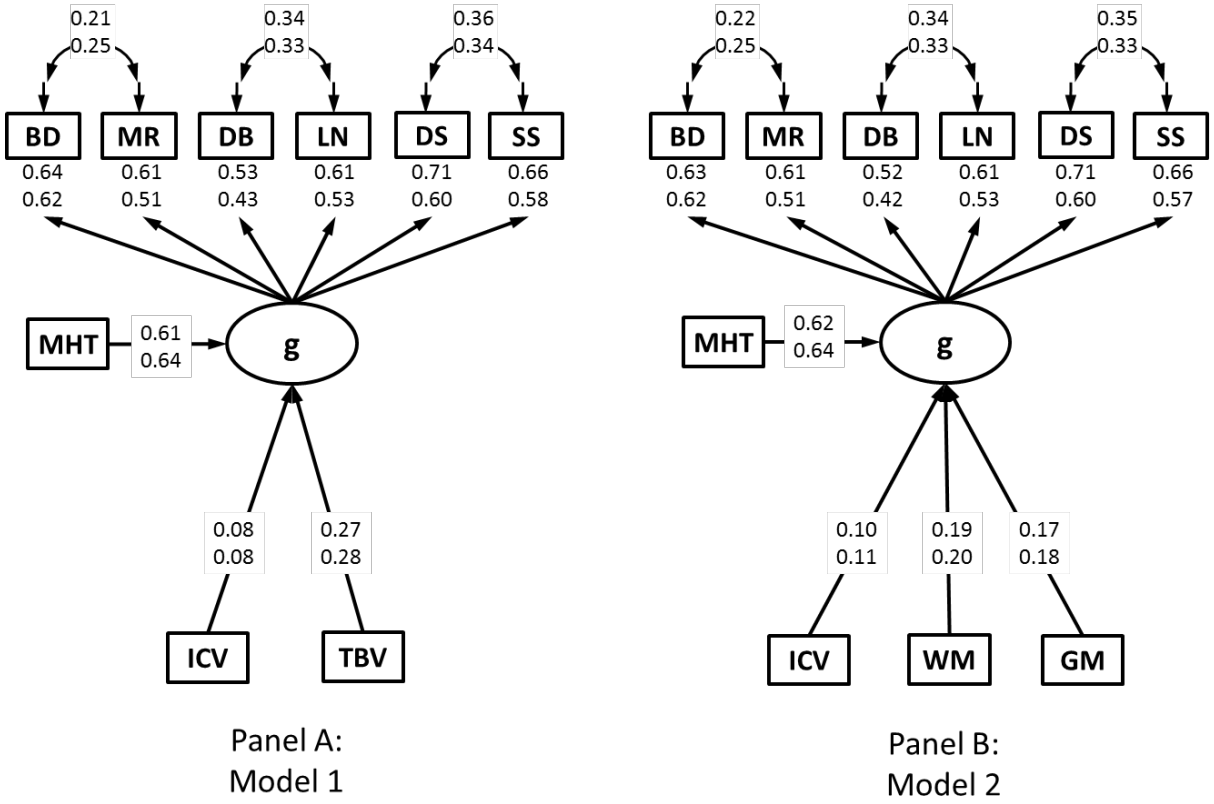


Figure 12.1: Structural diagram for Models 1 and 2. BD=Block Design; MR=Matrix Reasoning; DB=Digit Span Backward; LN=Letter-Number Sequencing; DS=Digit Symbol Coding; SS=Symbol Search; MHT=Age 11 MHT IQ score; ICV=Intracranial volume; TBV=Total Brain Volume; WM=White Matter Volume; GM= Grey Matter Volume. All variables are residuals controlling for age. All values are standardized parameter estimates. Estimates are taken from the final models with parameters constrained across groups. Parameter estimates are presented for males (top) and females (bottom) separately. Note, no differences between male and female models were significant. All values are significant at a minimum of $p<0.05$. Model fit for Model 1 ($\chi^2=59.59(53)$, $p=0.25$; CFI=1.00; TLI=0.99; RMSEA=0.020; SRMR=0.034) and Model 2 ($\chi^2=88.42(64)$, $p<0.05$; CFI=0.98; TLI=0.98; RMSEA=0.035; SRMR=0.036).

For model 1 (Figure 12.1), we first tested a model including only ICV and TBV, but not prior cognitive ability. Both ICV and TBV were significant predictors of current cognitive ability in both males (ICV=0.19, $p<0.001$; TBV=0.30, $p<0.001$) and females (ICV=0.21, $p<0.001$; TBV=0.30, $p<0.001$). Combined ICV and TBV accounted for approximately 14% of the variance in current cognitive ability. Next, we included prior ability (MHT) as a predictor of current ability level (Figure 1, Panel A). Prior ability was the strongest predictor of current ability (Males = 0.61, $p<0.001$; Females = 0.64, $p<0.001$). ICV and TBV remained significant predictors of cognitive ability, however the magnitude of the association with ICV dropped to 0.08 ($p<0.05$). In total, MHT, ICV and TBV accounted for approximately 52% of the variance in later life cognitive ability.

Similarly, for model 2 (Figure 12.1), we first tested a model excluding prior ability. In this model, ICV (Male=0.22, $p<0.001$; Female=0.23, $p<0.001$), WM (Male=0.22, $p<0.001$; Female=0.21, $p<0.001$) and GM (Male=0.18, $p<0.001$; Female=0.18, $p<0.001$) were all significant predictors of current cognitive ability accounting for a combined 11% of the variance in cognitive ability. Including prior ability (Figure 1, Panel B) the magnitude of the association with ICV dropped (Male=0.10, $p<0.05$; Female=0.11, $p<0.05$), but remained significant. The magnitudes of the associations with WM and GM were approximately equal. Overall, model 2 accounted for approximately 48% of the variance in current cognitive ability.

There were no substantive differences in either model 1 or 2 when models were re-estimated using the subsample of participants scoring ≥28 on the MMSE (n=537). The regression paths of ICV to g in model 1 failed to reach significance in the reduced sample, but the parameter estimates were identical at the second decimal place, indicating the lack of significance is most likely due to a reduction in power with sample size.

12.4 Discussion

The current findings suggest that cognitive ability at age 73 years is dependent, in part, on prior cognitive ability, prior or maximal brain size, and current brain tissue

volume. The study also finds that there were generally similar modestly-sized cross-sectional associations between brain size and cognitive ability in childhood and old age, though a proxy measure (ICV) had to be used for brain size in youth.

Across models, current brain tissue volume was a stronger predictor of later life cognitive ability (both with and without controlling for past ability), than ICV, replicating several past studies (Sluimer et al, 2008; Cardenas et al, 2011). However, it is important to note that a small, but significant, effect of ICV on later life cognitive ability remained in both models. The magnitudes of these associations did not differ significantly across males and females. Given the current sample size, we consider these associations to be accurate and robust. As such, the current study confirms only a very modest association of maximal brain size with cognitive ability in old age.

A large body of research has considered whether grey or white matter deterioration has the greater impact on cognitive ageing (e.g. Taki et al, 2011; Ziegler et al, 2010), with mixed results. In the current study, the effects of white and grey matter volume on g were largely equivalent and held in both males and females, suggesting a comparable influence on later life cognitive ability. However, as has been noted previously (Ziegler et al, 2010), grey and white matter deterioration may localise in different areas of the brain, and may therefore have differentiated effects on cognition. Thus, though the associations may be comparable for both grey and white matter whole brain volumes, the functional effect may be differentiated due to their roles in the underlying neural networks.

As previously noted, the simple correlational analyses in the current study provided a number of important estimates for the research literature on both brain size and cognition, and ICV and current brain status in ageing people. Specifically, we found positive associations between brain volume and cognition both in youth (males=0.28; females=0.12), where ICV in old age was used as an indicator of brain size in youth, and later life (males=0.27; females=0.26). These findings are in close agreement with the previous meta-analysis of McDaniel (2005) and the literature reviews of Rushton and Ankney (2009) and Miller and Penke (2007). Our findings contribute significantly to the literature because, in this single study, the sample (n=620 full sample; n=537 MMSE ≥28) is approximately 40-45% of the total sample reported in

these quantitative reviews. Moreover, this sample provides estimates of these effect sizes across 60 years in the life course up to age 73 years in the same subjects.

The study has some strengths. It is rare to have cognitive ability scores from youth in older people. The sample is large and homogenous in terms of the age range of participants. As such, our analyses gain statistical power, and have a natural control for the confounding effect of chronological age (Hofer, Flaherty and Hoffman, 2006). An additional advantage of our large sample was the ability to estimate models reliably in males and females independently, rather than simply including sex as a covariate in statistical analyses. Given the large array of well-validated cognitive tests administered at age ~73 years, and the large sample size, we were able to estimate all models using SEM, thus providing reliable estimates of latent cognitive ability constructs which explicitly account for measurement error, and allowing the simultaneous estimate of all substantive and covariate parameters.

There are some limitations with the current study. The potential influence of age-related changes to the skull, such as thickening of the inner skull table (May et al, 2010), may lead to underestimates of brain changes when ICV is used as a measure of pre-morbid brain volume, especially when inner table skull thickening is known to affect women more than men. However, as of yet there are no reliable methods for estimating ICV that take account of the effect of inner skull thickening. Thus we acknowledge the possibility of bias due to such effects, but are not able to provide any reasonable adjustments to the current findings.

In future research we aim to study the ageing process in more detail, using repeated measures of broad and specific cognitive functions and brain parameters. In the current study, our childhood estimate of cognitive ability was an overall IQ score and, as such, we focussed specifically on general cognitive ability (g) in later life as the principal outcome. However, we recognise that cognitive ability is known to be constituted by several different domains and future research with additional longitudinal data not currently available will be able to consider the associations between these domains and atrophy.

The current study yields a number of important conclusions for the associations between brain status and cognitive ability in older age. Firstly, both prior and current brain size are significant predictors of current cognitive ability, over and above the influence of prior cognitive ability. Secondly, the effects were highly similar for white and grey matter. These conclusions hold in both males and females. In addition, the correlational analyses yields one of largest single study-based contemporaneous associations between brain volume and cognitive ability, and between ICV and current brain tissue volume

Chapter 13. Hippocampus and cognition

13.1 Introduction

The hippocampus is involved in cognitive tasks such as learning, memory, emotional behaviour, stress regulation and spatial navigation (Foerster et al, 2012; Muzzio, Kentros, & Kandel, 2009; Nossin-Manor et al, 2012). Hippocampal volume reduction is associated with the development of Alzheimer's disease and other disorders of memory, with findings showing links between poor cognitive performance and smaller hippocampal volume (Leung et al, 2010; Sabuncu et al, 2010). Reduction in hippocampal volume has been linked to schizophrenia and multiple sclerosis (Adriano, Caltagirone, & Spalletta, 2012; Ceccarelli et al, 2007; Cercignani et al, 2001). It is also thought to be involved in general age-related cognitive decline, though reports are often mixed with some research finding a significant inverse association (Wolz et al, 2010) and others no association (Sanchez-Benavides et al, 2010). Although these inconsistencies might be due to methodological differences, such as image segmentation techniques or the population studied (Adriano et al, 2012), it is important to note that the main focus of these studies was on hippocampal size measured using conventional structural MRI techniques (Nossin-Manor et al, 2012).

Age-related brain tissue loss is most likely to be preceded by cellular changes, such as synaptic loss and neuronal degeneration (Hyman et al, 1984), which may not be detectable by conventional volumetric measurement. Quantitative MRI techniques

such as relaxometry, magnetization transfer (MT-MRI), diffusion tensor (DT-MRI) and perfusion MRI can detect subtle brain tissue changes not identifiable on conventional T1- or T2-weighted MRI (Ceccarelli et al, 2007; Cercignani et al, 2001; Davies et al, 2004; Filippi et al, 2000; Filippi & Rovaris, 2000; Parry et al, 2003; Rovaris & Filippi, 2000; Hugo Vrenken et al, 2006; Vrenken et al, 2006). Some of these techniques have recently been used to uncover associations between brain-wide white matter integrity and cognitive ability in old age (Penke et al, 2012).

T1 is the longitudinal (or spin-lattice) relaxation time and is related to the tissue water content, with increased T1 indicating increased tissue water, e.g. oedema that might, for example, reflect axonal damage (Bastin et al, 2002). MTR measures the efficiency of the magnetization exchange between relatively free water protons and those water protons that are bound to protein macromolecules in cellular membranes. Low MTR values indicate reduced transfer efficiency suggesting axonal damage and demyelination (Bastin et al, 2002; McDonald, Miller, & Barnes, 1992).

DT-MRI is most often used for measuring white matter integrity but it has also been proposed as a measure of grey matter integrity (Bhagat & Beaulieu, 2004; den Heijer et al, 2012; Pal et al, 2011). Fractional anisotropy (FA) and mean diffusivity (MD) are scalar indices obtained from the diffusion tensor, with the former indicating the degree of directionality of the water molecule diffusion when subjected to cellular boundaries within a tissue, and the latter indicating the overall magnitude of water diffusion. When the microstructure of cells break down, water molecules can diffuse further and more uniformly in all directions (Bhagat & Beaulieu, 2004) resulting in increased MD and reduced FA compared with healthy, structurally intact tissue.

It has been reported in several small cohort studies that hippocampal structural changes are detectable using image relaxometry (Kosior et al, 2011; Sumar et al, 2011; Wang et al, 2012) and MTR (Diniz et al, 2011; Margariti et al, 2007; Ropele et al, 2012; van den Bogaard et al, 2012; Vrenken et al, 2007). Increased relaxation time in the hippocampus has been associated with poorer cognitive performance in Alzheimer's disease compared to those with vascular dementia and matched controls (Wang et al, 2004); and MTR has been shown to detect brain changes in medial temporal lobe epilepsy sufferers white and grey matter, in the absence of significant

volume change (Diniz et al, 2011). Additionally, DT-MRI has been reported to be sensitive at detecting hippocampal changes (Carlesimo et al, 2010; Cherubini et al, 2010; den Heijer et al, 2012; Hong et al, 2010; Muller et al, 2005). In view of these previous findings, we anticipate that multivariate analysis of a range of quantitative MRI parameters in a large ageing sample could provide useful information about hippocampal structural changes and their role in cognitive ageing. However, to the best of our knowledge no studies have yet assessed the association between cognition in older people and hippocampal integrity characterised by multiple quantitative MR parameters such as longitudinal relaxation time (T1), magnetization transfer ratio (MTR) and water diffusion tensor parameters.

The aim of the current study was to investigate associations between major, ageing-relevant cognitive ability domains and hippocampal integrity measured using multi-parametric MRI (T1, MTR, FA and MD) in a large sample of community-dwelling older adults. We hypothesized that hippocampal integrity measured using these advanced MRI techniques would be more sensitive at detecting age-related integrity than volumetric measurements alone and hence provide further insights into the role the hippocampus plays in cognitive functioning in old age.

13.2 Methods

13.2.1 Subjects

Amongst the 627 subjects who had complete data for image segmentation, 56 participants did not have complete cognitive ability test scores and 5 were excluded because of segmentation failure, leaving a final sample of 565 (301 men, Table 13.1), aged 71.2 to 74.2 years (mean 72.7, SD 0.7 years). Of these 565 subjects, 483 (245 men) had MMSE scores above 27, and were aged 71.2 to 74.3 years (mean 72.8, SD 0.7 years).

13.2.2 Image analysis

Hippocampal structures were segmented from the high-resolution T1-weighted volume scans using FLIRT-FIRST (Patenaude et al, 2011). All of the generated masks

were visually inspected and, where necessary, corrected by manual editing resulting in a hippocampal mask and volume measurement for each subject. The editing was based on a manual segmentation protocol to reduce rater error and inter-rater reliability ratings were 0.98 based upon a subsample of 103.

T1 and MTR maps were generated on a voxel-by-voxel basis as previously described (Armitage et al, 2007; Wardlaw et al, 2011), and hippocampal regions were extracted from T1 and MTR maps in the following steps. The T1-weighted volumes were first transformed into the native space of the T1 and MTR parametric maps using FLIRT (Jenkinson & Smith, 2001), and the transformation matrices applied to the hippocampal masks. These masks were then applied to the T1 and MTR maps. In order to remove potential partial volume errors due to interpolation and to ensure analysis of pure grey matter tissue within the hippocampal volume, grey matter masks were applied to the T1 and MTR maps, and average T1 and MTR values within hippocampal structures were computed.

DT-MRI data were pre-processed using FSL (http://www.fmrib.ox.ac.uk/fsl), to extract brain (Smith, 2002), remove bulk subject motion and eddy current induced distortions by registering all diffusion-weighted volumes to the first undistorted baseline (b=0 s/mm^2) volume (Jenkinson & Smith, 2001), estimate the water diffusion tensor and calculate parametric maps of FA and MD from its eigenvalues using DTIFIT (Behrens et al, 2003). To extract FA and MD in the hippocampus the high-resolution T1-weighted volume scan was brain extracted using Freesurfer (http://surfer.nmr.mgh.harvard.edu) and then transformed to DT-MRI space using FLIRT (Jenkinson & Smith, 2001). The transformation matrix computed was applied to the hippocampal masks and the resulting masks in DT-MRI space were then applied to the FA and MD parametric maps. The grey matter mask previously segmented was also applied to the FA and MD hippocampal mask producing pure grey matter segmentations, and the average FA and MD values were computed. Finally, the hippocampal masks in the T1, MTR, FA and MD maps were visually checked before computation of average values was performed (Figure 13.1).

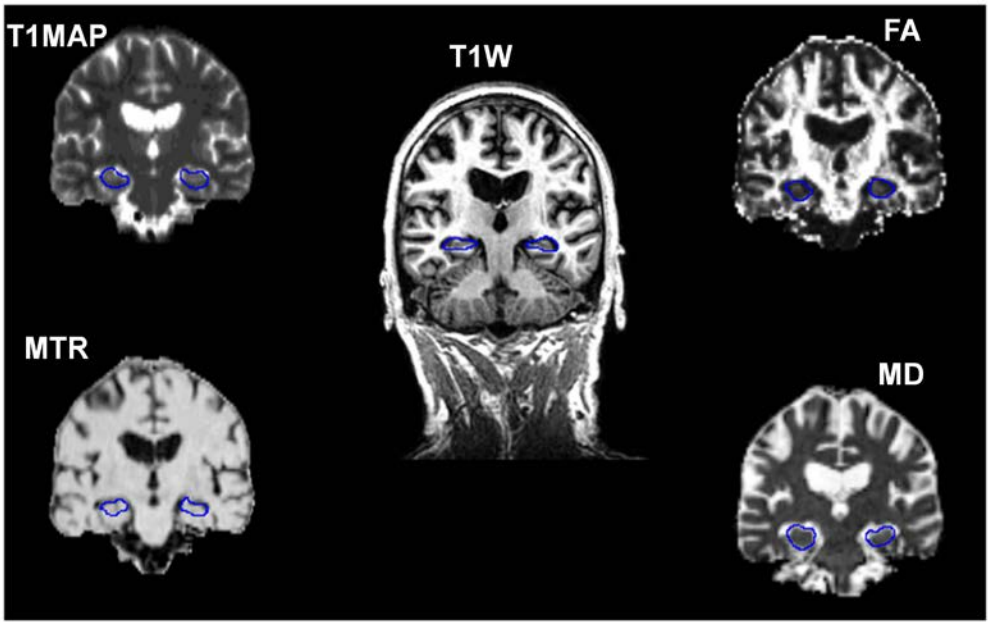


Figure 13.1: Typical images showing the quantitative MR images and the T1-weighted images with the outlines of the left and right hippocampi. T1MAP= T1 relaxation times, MTR = magnetization transfer ratio, MD = mean diffusivity and FA = fractional anisotropy.

13.2.3 Cognitive Ability Measures

Details of the cognitive data used in this analysis are given in Chapter 2 (2.3). Our primary analysis used all subjects, but we also performed secondary analyses using a more conservative MMSE threshold of above 27 to ensure the investigation of those who are free from potential cognitive impairment.

13.2.4 Statistical Analysis

All statistical analyses were performed using SPSS version 18 (SPSS Inc. Chicago III, USA), with all statistical tests being two-tailed, and P values < 0.05 being considered statistically significant. The left and right hippocampal integrity measures were compared using paired t-tests, followed by Bonferroni correction for multiple comparison. Associations between cognitive ability measures and hippocampal integrity measures were examined using multivariate linear regression models. In these models, each cognitive parameter (g, speed, and memory) was the dependent variable and each hippocampal integrity measure (T1, MTR, FA, MD and volume) was the independent variable. All models included gender and age 11 IQ because they

are known to be associated with hippocampus integrity or cognition, while models that assessed associations between cognition and hippocampal volumes included ICV to correct for individual differences in head size. A separate model which predicted cognitive abilities from the combined measures of integrity was used to assess how much variance in cognition in old age is accounted for by multiple measures of hippocampal integrity, age and age 11 IQ. We also assessed association between age 11 IQ and hippocampual integrity, to do this we developed a model that predicted hippocampal integrity from cognitive abilities at age 70 years, age 11 IQ and gender. To assess the effects of including subjects with possible cognitive impairment on any measured associations, analyses were performed for the entire population and for those with MMSE scores above 27. All p values were corrected for multiple testing using the False Discovery Rate approach.

13.3 Results

For the full cohort, left hippocampal volume (mean ± SD 3094.61 ± 444.58 mm^3) was significantly smaller than right (3337.11 ± 439.75 mm^3, $p<0.001$). The mean T1 relaxation time of left hippocampus (1.66 ± 0.16 ms) was significantly shorter than that of the right (1.67 ± 0.16 ms, $p<0.001$). The left hippocampal FA (0.12 ± 0.01) was significantly higher than right (0.11 ± 0.01, $p<0.001$). The left hippocampal MD (942.38 ± 69.44 × 10^{-6} mm^2/s) was significantly smaller than that of the right (966.62 ± 60.68 × 10^{-6} mm^2/s, $p<0.001$). There was no significant difference between left (47.99 ± 2.56 %) and right (48.02 ± 2.49 %, $p=0.60$) hippocampal MTR.

	The whole sample (Mean ± SD)
Ages in years	72.70 ± 0.70
MMSE	28.89 ± 1.35
Logical Memory Total 1st Recall WMS-III	45.92 ± 10.04
Logical Memory 2nd Recall WMS-III	28.97 ± 7.94
Verbal Paired Associates 1st Recall WMS-III	20.92 ± 7.70
Verbal Paired Associates 2nd Recall WMS-III	6.40 ± 2.05
Spatial Span Forward WAIS-IIIUK	7.68 ± 1.65
Spatial Span Backward WAIS-IIIUK	7.12 ± 1.57
Simple Reaction Time Mean Score	0.27 ± 0.05
Choice reaction Time Mean Score	0.64 ± 0.09

Inspection Time Total Correct Responses	111.48 ± 11.73	
Digit Symbol WAIS-III^{UK}	56.43 ± 12.34	
Digit Span Backward WAIS-III^{UK}	7.9 ± 2.30	
Block Design WAIS-III^{UK}	34.16 ± 10.05	
Letter-Number Sequencing WAIS-III^{UK}	10.98 ± 30.00	
Matrix Reasoning WAIS-III^{UK}	13.45 ± 4.87	
Symbol Search WAIS-IIIUK	24.77 ± 6.15	
Brain Tissue volume (mm^3)	1119184 ± 130234	
ICV (mm3)	1451103 ± 140637	
	Right Hippocampus	Left Hippocampus
T1 right (milliseconds)	**1.67 ± 0.17***	1.66 ± 0.17
MTR right (%)	47.93 ± 2.67	47.88 ± 2.74
MD right x 10-6(mm^2/s)	**969.22 ± 69.14***	943.77 ± 75.67
FA right	0.11 ± 0.01	0.12 ± 0.02
Hippocampus volume right (mm3)	**3333 ± 458***	3094 ± 460

*Measure in the left hemisphere significant smaller than that of the right, paired t-test, $p < 0.001$.

Table 13.1: Descriptive statistics of the whole sample, including volumetric measurements and quantitative MRI parameters.

Similar results were obtained when analysis used only those subjects with MMSE scores above 27 (Table 13.2).

	Subjects with MMSE score of 27 and above (Mean ± SD)
Ages in years	72.70 ± 0.70
MMSE	29.31 ± 0.74
Logical Memory Total 1st Recall WMS-III	46.79 ± 9.65
Logical Memory 2nd Recall WMS-III	29.78 ± 7.50
Verbal Paired Associates 1st Recall WMS-III	21.68 ± 7.42
Verbal Paired Associates 2nd Recall WMS-III	6.60± 1.95
Spatial Span Forward WAIS-IIIUK	7.72 ± 1.65
Spatial Span Backward WAIS-IIIUK	7.20 ± 1.59
Simple Reaction Time Mean Score	0.27 ± 0.05
Choice reaction Time Mean Score	0.64 ± 0.08
Inspection Time Total Correct Responses	111.92 ± 11.49
Digit Symbol WAIS-III^{UK}	57.49 ± 12.18
Digit Span Backward WAIS-III^{UK}	8.09 ± 2.26

Block Design WAIS-III^{UK}	35.07 ± 10.05	
Letter-Number Sequencing WAIS-III^{UK}	11.24 ± 2.93	
Matrix Reasoning WAIS-III^{UK}	13.85 ± 4.80	
Symbol Search WAIS-IIIUK	25.28 ± 6.04	
Brain Tissue volume (mm^3)	1119689 ± 132011	
ICV (mm3)	1449383 ± 139779	
	Right Hippocampus	Left Hippocampus
T1 right (milliseconds)	**1.66 ± 0.17***	1.65 ± 0.16
MTR right (%)	47.99 ± 2.60	47.95 ± 2.6
MD right x 10-6(mm^2/s)	**966.92 ± 69.18***	941 ± 67.72
FA right	0.11 ± 0.01	0.12 ± 0.02
Hippocampus volume right (mm3)	**3338 ± 455***	3097 ± 463

*Measure in the left hemisphere significant smaller than that of the right, paired t-test, $p < 0.001$.

Table 13.2: Descriptive statistics of the subjects with MMSE score of 27 and above, including volumetric measurements and quantitative MRI parameters.

In the regression models, after correcting for gender, ICV and age 11 IQ, larger volume of left hippocampus in the entire sample was significantly associated with higher scores of memory ($\beta = 0.11$, $p=0.003$, Table 13.3, Figure 13.2) and larger volume of the right hippocampus was significantly associated with higher scores of g ($\beta = 0.09$, $p=0.023$).

	g factor		*g* speed		*g* memory	
	Right	Left	Right	Left	Right	left
	Hippocampus volume, ICV not corrected for in the model					
Volume	**0.16 (<0.001)**	**0.13 (0.003)**	**0.12 (0.007)**	0.08 (0.068)	**0.11 (0.015)**	**0.16 (<0.001)**
Gender	0.07 (0.106)	0.05 (0.238)	0.08 (0.063)	0.06 (0.148)	**0.13 (0.003)**	**0.14 (0.001)**
	Hippocampus volume, ICV corrected for in the model					
Volume	**0.12 (0.011)**	0.08 (0.06)	0.07 (0.112)	0.03 (0.489)	0.08 (0.081)	**0.14 (0.002)**
Gender	**0.16 (0.003)**	**0.15 (0.005)**	**0.18 (0.001)**	**0.17 (0.001)**	**0.18 (0.001)**	**0.18 (0.001)**
ICV	**0.16 (0.005)**	**0.17 (0.002)**	**0.17 (0.002)**	**0.19 (0.001)**	0.10 (0.094)	0.07 (0.198)
	Hippocampus volume, ICV and Age 11 IQ corrected for in the model					
Volume	**0.09 (0.023)**	0.05 (0.151)	0.05 (0.24)	0.01 (0.828)	0.05 (0.164)	**0.11 (0.003)**
Gender	0.01 (0.82)	0.01 (0.928)	0.08 (0.117)	0.08 (0.137)	0.04 (0.345)	0.04 (0.369)
ICV	0.04 (0.373)	0.06 (0.233)	0.10 (0.069)	**0.11 (0.033)**	-0.02 (0.652)	-0.04 (0.365)
Age 11 IQ	**0.58 (<0.001)**	**0.58 (<0.001)**	**0.38 (<0.001)**	**0.38 (<0.001)**	**0.54 (<0.001)**	**0.54 (<0.001)**
			T1			
T1	**-0.13 (<0.001)**	**-0.14 (<0.001)**	**-0.20 (<0.001)**	**-0.18 (<0.001)**	**-0.11 (0.002)**	**-0.12 (0.001)**
Gender	**-0.07 (0.046)**	**-0.07 (0.036)**	-0.06 (0.144)	-0.06 (0.164)	0.01 (0.873)	0.01 (0.988)
Age 11 IQ	**0.58 (<0.001)**	**0.58 (<0.001)**	**0.39 (<0.001)**	**0.39 (<0.001)**	**0.54 (<0.001)**	**0.54 (<0.001)**
			MTR			
MTR	**0.10 (0.004)**	**0.11 (0.001)**	**0.15 (<0.001)**	**0.14 (<0.001)**	0.06 (0.105)	0.05 (0.157)
Gender	-0.04 (0.252)	-0.04 (0.272)	-0.01 (0.796)	-0.01 (0.865)	0.03 (0.355)	0.04 (0.332)
Age 11 IQ	**0.57 (<0.001)**	**0.57 (<0.001)**	**0.38 (<0.001)**	**0.38 (<0.001)**	**0.54 (<0.001)**	**0.54 (<0.001)**
			MD			
MD	**-0.11 (0.003)**	**-0.13 (<0.001)**	**-0.17 (<0.001)**	**-0.15 (<0.001)**	**-0.10 (0.005)**	**-0.12 (0.001)**
Gender	-0.06 (0.076)	-0.06 (0.092)	-0.03 (0.48)	-0.01 (0.744)	0.02 (0.617)	0.02 (0.621)
Age 11 IQ	**0.58 (<0.001)**	**0.58 (<0.001)**	**0.39 (<0.001)**	**0.39 (<0.001)**	**0.55 (<0.001)**	**0.55 (<0.001)**
			FA			
FA	**0.11 (0.001)**	**0.10 (0.003)**	**0.15 (<0.001)**	**0.12 (0.003)**	0.06 (0.128)	0.06 (0.091)
Gender	-0.05 (0.163)	-0.05 (0.191)	-0.00 (0.953)	0.00 (0.957)	0.04 (0.327)	0.04 (0.324)
Age 11 IQ	**0.57 (<0.001)**	**0.57 (<0.001)**	**0.37 (<0.001)**	**0.38 (<0.001)**	**0.54 (<0.001)**	**0.54 (<0.001)**

Note. Values are the standardized β (and *p* value) for the listed measures of hippocampus integrity predicting measures of cognitive ability. Models used the entire sample. Model: cognition = β1*integrity + β2*Gender + β3*Age 11 IQ. Where integrity represents measures of hippocampus integrity (T1, MTR, FA, MD and hippocampus volume). ICV is included only for hippocampus volume to correct for head size.

Table 13.3: Linear regression models for the association between cognitive abilities and longitudinal relaxation time (T1), magnetization transfer ratio (MTR) and hippocampal volume. N=565.

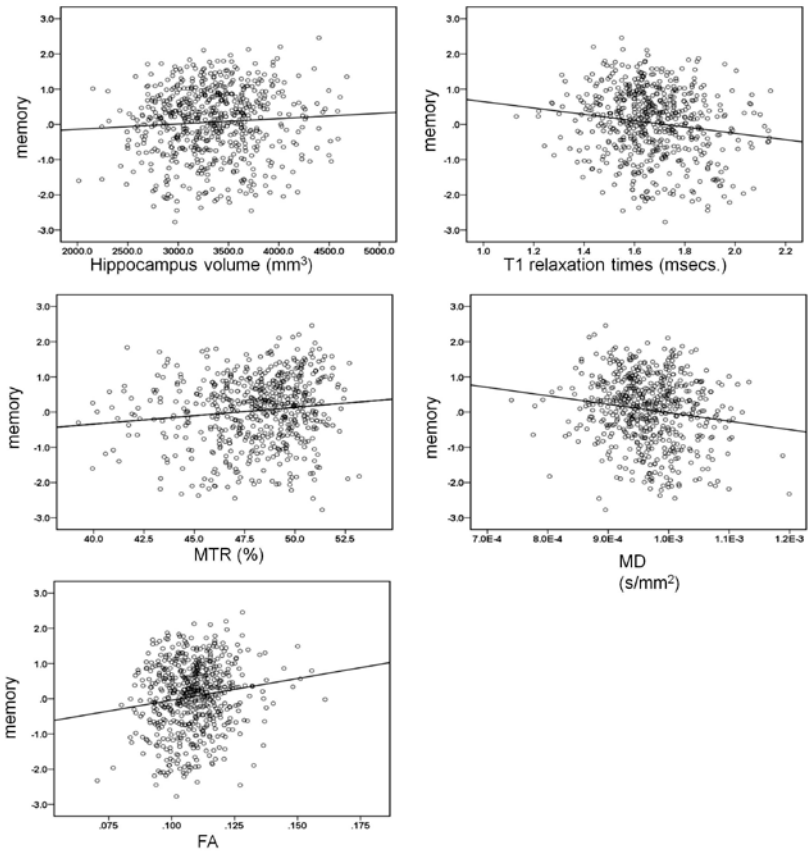


Figure 13.2: Scatter plots with regression lines showing bivariate associations between memory performance and hippocampal volume (a), longitudinal relaxation time (b), magnetization transfer ratio (c), mean diffusivity (d) and fractional anisotropy (e). Plots used only the measures of cognition and measures of integrity without accounting for any covariate.

The model that predicted hippocampal integrity from cognitive ability variables, gender and age 11 IQ showed that there was no association between age 11 IQ and hippocampal integrity (Appendix 5). Associations between cognitive ability variables and hippocampal integrity were similar for those with MMSE scores above 27 (Appendix 6 and 7).

For other measures of hippocampus integrity, after correcting for gender and age 11 IQ, shorter T1 and lower MD values in the hippocampus were significantly associated with higher scores of g, speed and memory (β: right and left, range = -0.10 to -0.20, all $p<0.001$; Appendix 6). Higher MTR and FA values in the hippocampus were

significantly associated with higher scores of g and speed (β: right and left, range = 0.10 to 0.15, all $p<0.001$). Associations were similar when the model was based on subjects with MMSE scores above 27 (Appendix 7). Thus T1 and MD, followed by MTR and FA were significantly associated (in decreasing order of effect size) with cognitive ability after correcting for age 11 IQ, whereas hippocampal volume did not show significant association in most cases. All significant associations between quantitative MRI measures remained after correction for multiple testing using the False Discovery Rate method.

The multivariate model that used the combined T1, MTR, FA and MD showed that, after correcting for age and gender, the combined hippocampus integrity measure explained between 4.8% and 10.2 % of the variance in cognitive ability variables. Age 11 IQ explained between 12.6% and 30.1 % of the variations in cognitive ability variables when entered in the same analyses (Appendix 7). We observed that the measures of hippocampus studied were significantly correlated with each other (Appendix 8). In view of this we investigated whether the correlation could introduce multicollinearity problem by computing the variance inflation factors (VIF) and tolerance. Appendix 9 shows that the models did not suffer from multicollinearity problem as none of the tolerance was less than 0.2 and none of the VIF was greater than 5 (Neter, Wasserman, & Kutner, 1989; Pan & Jackson, 2008). The individual quantitative measures of hippocampus integrity explained between 0.2% and 3.6% of the variance in cognitive ability variables.

13.4 Discussion

In our sample of generally healthy older individuals, we found that: T1 relaxation time and MD in the hippocampus were significantly associated with all cognitive ability variables investigated; hippocampal MTR and FA were associated with general intelligence and speed but not with memory; and only left hippocampal volume was significantly associated with memory, but not speed or intelligence. None of the significant associations was attenuated by the correction for multiple testing. The findings support our hypothesis that hippocampal integrity, measured using

quantitative MRI parameters, is more sensitive at detecting brain tissue structural integrity than volumetric measurements alone. To the best of our knowledge, this is the first study to investigate associations between cognitive ability and hippocampal integrity measured using multi-modal quantitative MRI techniques in a large sample of community-dwelling non-demented older adults.

We performed a separate analysis for participants with MMSE scores above 27. The conventional approach is to set the threshold to 24 that indicates possible cognitive impairment (Filippi et al, 2000), but our choice of a more conservative threshold of 27 allowed us to include those who are unlikely to suffer from cognitive impairment. We found that there was no difference in associations when the analysis included only subjects with MMSE scores above 27 compared with the use of the entire population. This was not a surprise because our participants were generally healthy individuals with no history of cognitive impairment or neuropsychological conditions.

The associations between hippocampal volume and memory are consistent with previous studies (Erickson et al, 2010; van der Lijn et al, 2008; Ystad et al, 2009) supporting the idea that the hippocampus is responsible for encoding and retrieval functions (Muzzio et al, 2009; Tamminga, Stan, & Wagner, 2010) and hence plays a key role in declarative memory (Boyer et al, 2007). Our finding that higher MD values in the hippocampus were associated with poorer cognitive ability is also consistent with previous studies (Carlesimo et al, 2010; den Heijer et al, 2012). We did not find any significant association between hippocampal FA values and memory. This is also in agreement with previous studies (Carlesimo et al, 2010; den Heijer et al, 2012), although both groups measured cognitive ability using only memory performance but in addition to memory, we assessed cognitive ability using both speed of information processing and IQ at older age, and our analysis accounted for age 11 IQ which allowed us to carry out a detailed investigation of the associations between cognitive ability and hippocampal integrity.

The observed associations between poorer performance on the cognitive assessments with increased T1, and increased MD suggest an age-related increase in tissue water, and with reduced MTR supports potential axonal damage as possible mechanism for poorer cognitive ability. This observation is supported by the association between

poorer cognitive ability and lower FA, reflective of further microstructural changes in cellular structure. The associations between quantitative MRI parameters and cognitive measures suggest that subtle changes in hippocampal cellular structure may have begun to affect cognitive processes before changes in volume are detected. The currently ongoing longitudinal MRI of this population will provide an opportunity to study these subtle, but potentially significant changes in cell structure, and allow a better understanding of the interaction between biological age-related changes and their cognitive correlates.

Reuben et al. (Reuben et al, 2011) have suggested that the hippocampus may be involved in logical reasoning, or fluid intelligence which is itself correlated with processing speed (Sheppard & Vernon, 2008). Our finding that MTR was associated with intelligence and processing speed but not memory may reflect this aspect of hippocampal function. We know that information processing speed mediates associations between intelligence and tract integrity (Penke et al, 2010), and that diffusion methods are more sensitive at detecting axonal damage, therefore it would seem that our findings of associations between cognitive ability and FA, and MD reflect changes in the substrates of hippocampal tissue likely to contribute to poorer performance in cognitive measures more associated with neural networks.

Asymmetry in hippocampal volume is common, with a smaller left than right hippocampus being reported in healthy older adults (Woolard & Heckers, 2012) as well as in dementia and dementia subtypes (Eckerstrom et al, 2008). It may be the case that hippocampal degeneration reaches a threshold whereby the volume has reduced significantly enough to affect cognition as maybe the case in Alzheimer's disease, where significant hippocampal atrophy is associated with poor memory when compared to age matched controls (Leung et al, 2010). The association between left hippocampus and memory may indicate that it is differentially affected by the ageing process, though the potential biological underpinnings of this need to be explored in future research.

The differential pattern of associations between cognitive performance and quantitative MRI parameters in the hippocampus, compared to the associations found between hippocampal volume and cognitive measures may indicate that quantitative

MRI biomarkers are sensitive at detecting histopathological changes in the absence of severe neuronal loss. Support for the idea that these measures are more sensitive at detecting microstructural changes comes from studies that have used MD and FA (Hong et al, 2010), and MTR (Hanyu et al, 2005) to differentiate between various patient groups. The successful application of quantitative MRI techniques to distinguish between subtle differences in the underlying pathology of diseases with overlapping characteristics, such as Alzheimer's disease and dementia with Lewy bodies, lends strength to the use of multi-modal MRI in studying age-related structural changes in the hippocampus of normal older adults. To test the pattern of change in multi-modal hippocampal parameters either a longitudinal or large cross-sectional dataset, which included participants with a range of dementia subtypes, mild-cognitive impairment and normal older adults would be helpful. Application of multi-modal MRI in such a dataset would help to elucidate the parameter that is most sensitive to cognitive change, hopefully leading to a clearer understanding of the underlying mechanism that is influencing the cognitive outcome.

The main strength of this study lies in the application of multi-modal MRI to quantify structural integrity in the hippocampus in a large (n=565), well-characterised group of older adults. This study is one of the largest so far to report associations between any measured hippocampal integrity and cognitive ability (Adriano et al, 2012). Where previous studies have successfully applied these techniques to pathological conditions such as brain tumour or multiple sclerosis (Davies et al, 2004; Liang et al, 2012), Alzheimer's disease (Hanyu et al, 2005; Hong et al, 2010; Ropele et al, 2012), dementia with lewy bodies (Hanyu et al, 2005) and cerebrovascular disease (Foerster et al, 2012), we have shown their usefulness in providing more sensitive measures of brain structure than volumetric analysis in detecting subtle associations with cognitive performance. Another strength of the study is the access to early life cognitive data, age 11 IQ, allowing us to control for prior ability when looking at associations between cognitive ability in later life and brain size. We clearly demonstrate, through the assistance of age 11 IQ, that hippocampus integrity is associated with cognitive decline over a lifespan, from youth to later life. Failing to account for earlier life cognition would risk the erroneous assumption that all associations between hippocampus and cognition in later life are the consequence of ageing.

The main limitation of the study is the lack of longitudinal data to assess time dependent changes in the hippocampus and their association with cognitive ability. However, the LBC1936 participants are currently undergoing repeat MRI to provide such longitudinal data.

In conclusion, we found that hippocampal integrity assessed using T1, MTR, MD and FA were significantly associated with nearly all measures of cognitive ability investigated, even after accounting for early life age 11 IQ, whereas volume was less sensitive. Advanced multi-modal MRI measures (obtainable from three MRI sequences) may provide more sensitive measures of age-related changes in hippocampal integrity than volume measurements derived from conventional structural MRI. Furthermore this approach may be more useful in helping us to determine the brain's role in cognitive ageing, specifically individual differences present in the associations between measures of the hippocampus and cognition.

Section 6: Discussion

Chapter 14. General discussion

The work in this thesis is intended to explore the challenges of measuring brain ageing, and to investigate how the resulting data were related to cognition in later life. Global measures of brain status such as intracranial volume and regional measurements of areas often implicated in ageing were studied. The main aims were to explore existing segmentation methods to discover the pros and cons of applying them in a large, older group of healthy adults. By doing so it was hoped that practical solutions to some of those difficulties posed by specific age-related features in the brain would be overcome, or at the very least obvious next steps needed to resolve them would emerge.

The difficulties surrounding defining how well a method has performed are considerable. When trying to determine which method out of several methods available you should choose, there are many factors to consider. Validation of method performance should be appropriate to the use and application of the methods you are evaluating. For example when using a method to classify people into groups the

sensitivity and specificity of that method, at correctly classifying people is most important. Whereas when looking at features within a group, such as brain size, the method being evaluated should assess accuracy of measuring brain size. Therefore an evaluation metric that determines false positive/negative tissue segmentation would be most useful.

Once an evaluation method has been chosen the next difficult step is to define what constitutes the 'best' or most 'accurate' method. When determining the performance of a segmentation method, the ultimate goal is to find a method that obtains a segmentation of the target region that is the most similar to the true value of that region. In brain research the true value can never be truly known as post mortem measurements are the only method by which you can directly measure a structure. Although post mortem studies exist they are not useful when investigating disease evolution or brain changes in healthy populations. Therefore reference standard measurements, obtained via highly trained manual raters, are considered the gold standard. When evaluating a set of methods comparing it to a reference standard is the simplest way of knowing which has performed well. In this instance the 'best' method is the one closest to the reference method. Though a minimum acceptable level of error or accuracy should be defined, otherwise the method that gives the closest result could still be considerably different than the reference standard, just less so than the other methods being evaluated. However the difficulty is in defining what is close enough to the reference standard. The most useful way of defining this is to look at comparable published methods, this way a method can be considered to be the 'best' performing if the magnitude of comparison to the reference standard is better than those published in other studies. This issue becomes further complicated if the method is novel and no comparative research is available. In this instance it is good practice to ensure that rate of error, by whatever validation metric is used, is less than that produced by two rater's measurements.

The variation in the validation strategies used to assess method performance can be seen from the literature review of automated hippocampal methods presented in this thesis. This variability between studies makes it difficult to compare results and though steps have been taken, in the form of publicly available datasets that include

reference measurements, this issue is still a significant problem in the field of computational morphometry.

14.1 Global measures

The studies in section 3 were designed to establish an appropriate method of measuring ICV, to assess if ICV is a stable measure of maximal brain size and to investigate potential confounding age-related factors when measuring ICV. We found that ICV measurements were most effectively made using a semi-automated technique that utilises a combination of image intensity thresholding and manual editing. This analysis specifically highlighted the need to either use spatial concordance statistics or visual assessment when validating methods against a reference standard. The confounding effect of the presence of over and under segmentation in the same image was made apparent when volume comparisons suggested smaller overall differences than were actually present on visual inspection of the segmentations. The method established during this piece of work was adopted as the primary image analysis protocol for ICV segmentation in the wider LBC1936 study. It was deemed to have enough input to allow users to follow well defined inclusion and exclusion criteria of erroneous structures, whilst reducing time investment and increasing reproducibility.

Analysis that built on this looked at whether early life cognitive ability and brain status were better predictors of late life cognitive ability than brain status in old age. The finding that brain tissue volume at age 73 years is the better predictor even after accounting for age 11 IQ and ICV, which functions as a proxy measure for brain size in youth, suggests age-related mechanisms are involved. However the presence of a modest association remaining between ICV and late life cognitive ability is interesting as it suggests some support for the cerebral reserve hypothesis. The main difficulty with making inferences from cross-sectional data is the inability to take account of adaptive processes. Adaptive processes may help older adults to retain abilities by recruiting alternative neural networks to compensate for decline in other areas of the brain (Park and Reuter-Lorenz, 2009). The variation in reported findings between cognitive decline and decreased regional brain volume may result not only from individual factors associated with decline, such as lifestyle and environment but also

from different patterns of compensation over a person's lifetime. It may be that those individuals, whose better cognitive ability was associated with fewer indications of brain tissue atrophy at age 73, began life with a brain better equipped to adapt to changes and insults across the lifespan. Ultimately how well we age, however that is measured, is dependent on a multitude of factors but the state of our brain throughout our lives, not just in old age, is important.

One of the main tenets of using ICV as proxy measure for pre-morbid brain size is that it remains stable overtime, the analysis in Chapter 5 shows that this is the case. As mentioned the only way of being more sure of this would be to obtain MR images from early life to old age but in the absence of this cross-sectional data of a large sample is a good alternative. There is however an age-related process that affects the inner skull table and therefore the intracranial vault, which was identified and investigated. The prevalence of thickening of the inner skull in older adults is not known, partly because other than when it is aggressive it seems to be a benign process and therefore does not foster much attention. The study presented in this thesis revealed that women are more affected than men and in those cases of substantial thickening the ICV measurement was confounded by a significant degree. Though exploratory the results are important in highlighting the need to understand the biological underpinnings of structures being measured in neuroimaging research. Gender differences in age-related brain tissue decline may reflect sexual dimorphisms but they may also in part be influenced by ICV measurement, which in turn is influenced by inner table skull thickening. The process is most prevalent in the frontal region of the skull, what affect this has on frontal brain changes is not known but requires further investigation. This analysis not only flags a potential confounding factor in age-related gender differences and measurement of ICV but underlines the need for researchers using image analysis to take a more holistic approach to quantifying brain status.

14.2 Regional measures

The decision to focus on hippocampal segmentation was initially as a result of how to achieve accurate measurements in a large dataset when manual measurements were

considered too time consuming. Automated methods are commonly used but not always effective in older adults. The review of automated methods (Chapter 3) for hippocampal segmentation suggested that although methods were not often adequately validated in healthy older adults, those that were reported a good agreement with manual segmentations. The difficulty when comparing effectiveness of methods across studies was in the variation in the statistics used to validate methods. Good agreement between a method and a corresponding reference standard measure should be considered in relation to the expected size of the effect being investigated. If an overlap statistic of 80% were reported it would indicate a high degree of agreement between methods. However if a small difference of 5-10% were expected then an 80% agreement leaves a discrepancy of 20% between measurements, which would exceed the expected difference. Another issue that emerged from the review was that of associations between group characteristics and hippocampal volume being present or not depending on the automated method used (Bishop et al, 2011). This has been found in other brain regions that share the same anatomical challenges as that of the hippocampus in that they are small, variable and can have difficult to detect borders with other regions (Tisserand et al, 2002).

A concern that arose from the study comparing automated methods (Chapter 3) was that of the differences between hippocampal asymmetry found with each method. The magnitude of the difference between left and right hippocampal volumes was much smaller for the Freesurfer method than either the manual or FSL_FIRST methods. Automated methods that use atlases not matched to the group being studied risk potential confounding affects such as this. Significantly the papers in the review that reported good agreement between automated and manual methods in large samples both used targeted, age-appropriate templates. All of these factors should lead to the cautious use of automated methods in healthy older populations but not necessarily prohibit it. Small associations and subtle changes require methods that are precise to reduce, as much as possible, the confounding effects of measurement error and bias from the data.

Assessment of two widely used and freely available automated methods was performed to determine if they could be used to provide hippocampal segmentations for the entire LBC1936 dataset. The finding that FSL_FIRST method worked better

due to the introduction of an age appropriate atlas was as expected but manual editing was still required. Segmentation algorithms are becoming increasingly more sophisticated and able to perform to high standards however a standardised procedure for validating methods, automated or otherwise, should be developed to ensure that a certain level of confidence can be placed in the application of a method to providing volumetric analysis of the human brain. The efforts of Shattuck et al (2009), who have developed an online resource containing datasets and tools for validation of segmentation methods, should be recognised as a positive step in addressing this issue. Normal image brain banks and freely available atlases are also starting to become available online (http://www.sinapse.ac.uk/research-resources/brains-project). Progress in image analysis is to the benefit of all those scientific disciplines that utilise its unique contribution, however basic standards of validation, reproducibility and reliability should not be overlooked.

As discussed earlier associations between hippocampal volume and cognitive decline in healthy older adults have been mixed. The analysis in Chapter 13 suggests that these subtle associations may be better detected using multimodal imaging. Segmentations of the hippocampus were still needed to perform this analysis, so despite tissue volume being a less sensitive measure it was important that an accurate mask be obtained. These findings contribute most directly to the debate on the underlying changes occurring in the microstructure of a tissue that may influence cognitive function in the absence of significant tissue deterioration. Smaller left hippocampal volume was associated with memory decline but interestingly the strongest associations with quantitative metrics were not with memory but with general intelligence and speed. It could be interpreted that measures such as MTR, T1, FA and MD detect changes in the hippocampus related to its function in networks other than memory.

One of the most informative findings from this thesis was the extent to which manual protocols used to delineate the frontal lobes vary significantly. The review in Chapter 4 revealed the considerable number of boundaries used to define the frontal lobe and the follow up analysis that applied a selection of these boundaries to segmentation showed the consequences of such substantial variation. Further to the suggestion that validation protocols need to be developed, manual tracing protocols that are

standardised and widely applied are vital. The huge implications in our comprehension of the vast amount of literature discussing the role of the frontal lobe in healthy ageing is being held back by the inability to compare findings across studies. If each research study has a different understanding of frontal lobe boundaries it becomes impossible to usefully interpret findings in the context of what already exists. Crucially the application of protocols analysis (Chapter 11) brought out the delicate trade off between accuracy and reproducibility that is pervasive in image analysis. Though it is concerning that anatomically, accurate boundaries, such as the central sulcus, show significant variation in application and reproducibility the alternative of using a cut plane is inadequate.

14.3 Limitations

It must be noted that there are methodological limitations that apply to all of the studies in this thesis. Though the LBC1936 cohort is part of a longitudinal study, the data used in this thesis is cross-sectional, as it was performed prior to completion of the second wave of testing. This means that any significant associations reported could be due to factors accrued at any point in the lives of the participants and not necessarily as a result of ageing. Follow up data at a second time point is currently being acquired and subsequent waves are planned with detailed analysis.

Due to time, data storage and computational capacity constraints it was not possible to perform a like-for-like comparison of automated hippocampal methods. Additional analysis which included segmentations derived from Freesurfer with an age appropriate template in place of the standard atlas would have given a fairer indication of the methods capabilities. However, the purpose of this study was not to determine which of the two methods was best, but to assess the individual performance of each method against manual segmentation of the same subjects. Therefore it is sensible to interpret the results in two ways. First, as giving an indication of how a freely available automated method performs in a dataset that it has not been developed in. Second, of how the optimisation of such methods can improve overall performance in an older group.

14.4 Future work

The finding that brain status in old age is a good predictor of ability in late life is not surprising but the residual, modest association remaining between ICV and cognitive ability in old age is interesting. Support for early life factors predicting late life function could be very useful in helping researchers identify groups who are more at risk of ageing poorly. These groups could be given extra support and early treatments to try and reduce some of the other contributing, and more easily targeted, factors such as lifestyle choices that influence cognitive functioning in later life. Longitudinal analysis, ongoing for the LBC1936, will help to further explore this idea by recording the ageing trajectories of the cohort over a decade in late life and beyond. Though a more definitive answer to this question cannot be reached without following a group from birth to old age, or even in utero to death, future work on the subsequent waves of data collection from this group may help to indicate which measures are better predictors as the effects of ageing become more pronounced. The measurement methods explored in this thesis provide an age sensitive set of methods that can be used to assess further waves, as well as contributing to the baseline data.

The methodological limitations highlighted, in both manual boundary protocols and automated methods, in this thesis can be viewed as positive steps towards addressing some of the contributing factors to variation in the reported literature. It is hugely important that research groups consider the image analysis methods they are using in the context of the group they are studying, even if that population is considered normal. The occurrence of normal brain changes in old age is common knowledge and yet methods developed for either younger groups or clinical populations are applied to, what should be considered as, an unique group. Future work should aim to take advantage of the rapidly developing online resources, to aid researchers in gaining access to high quality data for use in validation of methods in diverse samples. Knowing how well a method performs in a particular population would enable researchers to not only choose the most effective segmentation method for their study but will encourage improvements. Small but significant improvements in accuracy of segmentation methods can be significant when looking at subtle differences. Therefore the development of more age-appropriate templates in

improving segmentations in older adults, as well as other groups (Fonov et al, 2011), would be encouraged.

The major issue in image analysis of 'time versus accuracy' was not resolved in this work but significant steps towards emphasising those aspects in a process that can be standardised and improved was. It seems the most effective way of resolving this problem would be to refine an automated or semi-automated method that reduces rater variation but maintains quality control through expert user input (Aribisala et al, 2013). This would be reliant, either through the use of atlases or manual segmentations to train the algorithm, on a well researched, anatomically sensitive protocol being developed prior to any future methods work. In the case of the work presented here valuable further analysis, which would clarify the implications of comparing studies on the role of the frontal lobes in healthy ageing, would involve including cognitive data. Repeat analysis looking at the potential differences in strength of associations between the various boundary measurements, applied in Chapter 11, and cognitive performance would be useful. This would provide a clear indication as to the influence that having such a wide variety of protocols in the literature has upon our understanding of the frontal lobes involvement in normal ageing.

The finding that a multi-modal image analysis approach was more sensitive at detecting associations between cognition and the hippocampus should be further developed. Combining several brain image measures that contribute unique information about several tissue properties, rather than just volume for example, may be the next key step in understanding subtle brain ageing. The stronger associations with speed and general intelligence over memory suggest the role of the hippocampus in ageing is complex. Future work in this area should concentrate on assessing quantitative metrics in the component structures of networks implicated in a specific cognitive domain. The connective white matter tract integrity involved in the network also needs to be assessed. This level of investigation, involving both connectivity and region integrity, would provide the necessary degree of information required to fully understand how brain tissue decline is associated with cognitive decline in ageing.

Appendices

Appendix 1: Data extraction frontal lobe systematic review

Author & Year	Sample type	*n*	Mean Age (std)	Seq	Tesla	Slice Thickness mm	Pre-processing Aligned	Pre-processing Seg	Posterior Boundary
Almeida et al (2003)	EOD	24	72.8	T1	1	1	?	X	Most anterior slice on which the corpus callosum can be seen.
	LOD	27	75.5						
	HC	39	72.9						
Kohler et al (2010)	Depression	25	60+	T1	1	1	✓	X	Most anterior slice on which the corpus callosum can be seen.
	HC	29	60+						
Aylward *et al.*, (1997)	HA	10	47.3	T1	1.5	1.5	✓	X	Central sulcus.
Schretlen *et al.*, (2000)	HA	112	54 (19)	T1	1.5	1.5	✓	X	Central Sulcus
Baaré *et al.*, (1999)	Schizophrenia	14	28.5 (5.7)	T1	0.5	1.2	✓	✓	Precentral Sulcus
	HC	14	26.9(5.9)					Manual	
Staal et al., (2000)	Schizophrenia	16	40.6(8.2)	T1	1.5	1.2	✓	✓	PreCentral Sulcus
	Healthy Siblings	16	40.9(8.6)						
	HC	32	40.3(9.3)						
Ballmaier *et al.* (2004)	Depressed	24	66	T1	1.5	1.4	✓	✓	Precentral sulcus (but not for ACC).
	HC	19	66						
Blanton et al	Children	46	11	T1	1.5		✓	✓	PreCentral Sulcus

(2004)									
Elderkin-Thompson *et al.*, (2008)	HA	23	61-88	T1	1.5	1.4	✓	✓	PreCentral Sulcus
Elderkin-Thompson *et al.*, (2009)	Depressed	26	70 (7.7)	T1	1.5	1.4	✓	✓	PreCentral Sulcus
	HC	23	71 (7.9)						
Bartzokis et al (1993)	Healthy males	70	38.6 (1.5.6)	T2	1.5	3	✓	X	Midpoint of the CC
Berryhill et al (1995)	Severe CHI children	14	5-15	T2	0.5-1.5	5-8	X	✓	Last locator line on the midsaggital image passing through the rostrum of the CC
	Mild CHI children	14	5-15						
Beyer et al (2009)	Bipolar	56	60.5		1.5	3	✓	X	Coronal plane containing the CC genu.
	HC	43	58.1						
Bjork et al (2009)	HA	29	37.4(11)	T1	1.5	2	✓	✓	Coronal plane containing AC
Bokde et al (2002)	HCs	5		T1	1.5	1.2	✓	X	Central sulcus
	AD	3					☐	☐	
	Vascular Dementia	1					☐	☐	
	FTD	1					☐	☐	
Bokde et al (2005)	HA	10	60.3(8.8)	T1	1.5	1.2	✓	X	Central Sulcus
Convit *et al.*, (2001)	Schizophrenia	9	30-52	T1	1.5	2	✓	X	*"Coronal slice that bisects the distance between the cingulate sulcus and the PrCS in two equal parts."*
	HC - middle aged	9	30-52						
	HC - old	9	58-76						
Gold *et al.*, (2005)	Hypertensive	27	66.5(8.9)	T1		1.2	X	X	*"Coronal slice that bisects the distance between the cingulate sulcus and the PrCS in two equal parts."*
	HC	27	59.8(8.5)						
Bremner *et*	HA	11	18-65	T1	1.5	3	?	?	Line from inferior frontal sulcus to sylvian

***al.*, (1998)**									fissure
Bremner et al (2000)	MDD	16	43(8)	?	?	?	?	?	All coronal slices anterior to the AC
	HC	16	45(10)						
Bremner *et al.*, (2002)	MDD	15	43(8)	T1	1.5	3	?	?	All coronal slices anterior to the AC
	HC	20	45(11)						
Carper & Courchesne (2000)	Autism	42	5.4(1.7)	T2	1.5	3	X	✓	Central sulcus
	HC	29	6(1.8)				☐	☐	
Carper & Courchesne (2005)	Autism	25	2-9	T1	1.5	3	X	✓	Central sulcus
	HC	18	2-9						
Castellanos et al (1996)	ADHD	55	5-18	T1	1.5	1.5-2	✓	X	Anterior to the CC genu coronal slice
	HC	57	5-18						
Coffey et al (1991)	Depressed	35	55.7(17)	T1	1.5	5	✓	X	Optic chiasm - most posterior tissue not included
Coffey et al (1998)	HA	330	74.98 (5.09)	T1	1.5	5	X	X	Most anterior aspect of lateral ventricles
Colchester et al 2001	Korsakoff	11	39.1 (13.5)	T1/T2*	1.5	1.2-2.3	X	X	Central Sulcus
	Herpes E	9							
	Focal frontal lesion	6							
	HC	10	45.9 (17.3)						
Kopelman et al (2001)	Amnestic	40		T1	1.5	1.2	X	X	Central Sulcus
	HC	10							
Crespo-Facorro *et al.*, (1999)	HC	NR	NR	T1	1.5	1.5	✓	✓	A coronal plane passing through the most anterior tip of the inner surface of the genu of the CC.
	Schizophrenia	NR	NR				☐	☐	
	Schizophreniform	NR	NR				☐	☐	
	Autism	NR	NR						
	Schizoaffective	NR	NR						

Crespo-Facorro 2000	HC	34	25.2(6.2)	T1	1.5	Resampled to 1	✓	✓	As Crespo-Facorro *et al.*, (1999)
	Schizophrenia	26	26.4(6.7)						
Chemerinski et al (2002)	Schizophrenia	45	30(8.9)	T1	1.5	1.5	✓	✓	As Crespo-Facorro *et al.*, (1999)
	HC	45	30.5(8.5)						
Coryell et al (2005)	MDD	10	21.9(4.9)	T1	1.5	1.5	✓	✓	As Crespo-Facorro *et al.*, (1999)
	Schizophrenia	10	22.2(4.2)						
	HCs	10	22.1(6.0)						
Antonucci et al (2006)	Psychiatric (various)	15	39(8.7)	T1	1.5	5	✓	✓	As Crespo-Facorro *et al.*, (1999)
Gansler et al (2009)	Psychiatric (various)	41	40.12 (8.3)	T1	1.5	1.5	✓	✓	As Crespo-Facorro *et al.*, (1999)
	HC	19	40.94 (7.5)						
Szendi et al (2006)	Schizophrenia	13	25.9(5.4)	T1	1	1.5	X	X	As Crespo-Facorro *et al.*, (1999)
	HC	13	29.3(4.7)						
Boes et al (2007)	Cleft Palate	30	7-12	T1	1.5	1.5	✓	✓	As Crespo-Facorro *et al.*, (1999)
	HC	43	7-12						
Wood et al (2007) - adults	HA	60	18-50	T1	1.5	1.5	✓	✓	As Crespo-Facorro *et al.*, (1999)
Wood et al (2008) children	HA	74	7-17	T1	1.5	1.5	✓	✓	As Crespo-Facorro *et al.*, (1999)
Lindberg et al (2009)	FTD	12	42-72	T1	1.5	2.5	X	X	As Crespo Facorro but Fornito protocol for the ACC….
	SD	13	52-77				☐	☐	
	PNFA	9	57-78				☐	☐	
	HC	27	53-78				☐	☐	
Croxson et al (2005)	HA	10	24-35	T1	1.5	1	?	?	Medial - coronal plane 6mm posterior to the AC. Orbital - a line drawn between the posterior extent of the medial and lateral orbital sulci.

									Lateral - coronal plane 10mm anterior to the meeting of the inferior branch of the SFS and PrCS.
De Bellis et al (2005)	Alcohol Use Disorder	14	17(2.1)		1.5	1.5	✓	X	Coronal plane at CC genu
	HC	28	16.9(2.3)						
Drevets et al (1997)	Bipolar	21	35 (8.2)	T1	1.5	2.5	X	✓	Anterior most coronal slice where the CC no longer divides the striatum
	Unipolar	17	35 (9.4)					☐	
	HC	21	34 (8.2)					☐	
Hirayasu et al (1999)	Schizophrenia	17	27.2(7.4)	T1	1.5	1.5	X	X	As Drevets (Axial tracing)
	Affective disorder	24	23.7(5.1)						
	HC	20	24(4.3)						
Hastings et al (2004)	Depression	18	38.9 (11.4)	T1	1.5	1.5	X	Manual	As Drevets (Axial tracing)
	HC	18	34.8 (13.6)						
Nifosi et al (2010)	FPDD	15	43.9(11)	T1	1	1.5	X	X	Coronal plane at the AC
	HC	15	37.8(11)						
Botterton et al (2002)	Depression (young)	30	20.2(1.6)	T1	1.5	0.5	✓	X	As Drevets.
	Depression (old)	18	35.8(8.1)				☐		
	HC (young)	8	20.2(1.6)				☐		
	HC (old)	9	35.8(8.1)				☐		
Kegeles et al (2003)	MDD/Bipolar	19	36(11)	T1	NR	1.5	X	X	As Drevets.
	HC	10	39(19)				☐		
Brambilla et al (2002)	Unipolar	18	42(10)	T1	1.5	1.5	X	Manual	As Drevets.
	Bipolar	27	35(11)						
	HC	38	37(10)						
Egan et al (1994)	Schizophrenia	16	19-39	?	0.5	10	X	X	CC genu
	HC	16	19-43						

Exner et al (2002)	Basal Ganglia lesion	20	53(11)	T1	1.5	1	X	X	Precentral sulcus
	HC	20	52(9)						
Exner et al (2006)	Schizophrenia	15	29(9)	T1	1.5	1	X	X	Precentral sulcus
	HC	15	28(8)						
Filipek et al (1997)	ADHD	15	12.4(3.4)	T1	1.5	3	✓	✓	Coronal plane at AC.
	HC	15	14.4(3.4)						
Flashman et al (2001)	Schizophrenia	15	31.9(11)	T1	1.5	1.5	X	X	Precentral gyrus, insula,
Fornito *et al.*, (2006)	HA	24	25-29	T1	1.5	1.5	✓	✓	The first coronal slice posterior to the AC.
Foundas et al (2001)	HA	12	29.8(14)	T1	1.5	1.5	✓	X	Sylvian fissure
Knaus et al (2006)	HA	48	29.2(7.57)	T1	1.5	1.5	✓	X	Precentral sulcus (sagittal tracing)
Knaus et al (2007)	HA	60	30.15(8.97)	T1	1.5	1.5	✓	X	Precentral sulcus
Knaus et al (2009)	ASD	40	12.22(3.3)	T1	3	1	✓	X	Precentral sulcus
	HC	40	12.24(3.25)						
Fukui et al (2000)	PPA	17		T1/T2	1.5	5.5-6.5	X	X	CS - line connecting anterior sylvian fissure and frontal horn
	FTD	11							
	AD	24							
Giedd et al (1996)	HA	104	11.6(3.5)	T1/T2	1.5	1.5	✓	X	Anterior to coronal plane at CC genu
Casey et al (1997)	ADHD	26	9.69(1.99)	T1	1.5	1.5	X	X	Prefrontal - anterior most point of CC
	HC	26	9.8(1.7)						
Kumra et al (2000)	Schizophrenia	44	14(2.3)	T1	1.5	1.5-2	✓	X	Prefrontal - anterior most point of CC
	HC	64	13.7(2.2)						

ACC only

	Psychotic Disorder	27	12.3(2.9)						
	HC	42	11.7(3.3)						
Geroldi et al 1999	AD	28	53-86	T1	1.5	2	✓	X	Anterior to the coronal slice where the CC is present
	HC	30	53-86						
Ginovart et al (1997)	Huntington's Disease	5	49.4(7.6)	?	1.5	1.5	X	X	Drawn from the imaging planes in which the insula was visible bilaterally to the planes where the putamen disappeared.
	HC	5	48(7.8)						
Backman et al (1997)	Huntington's Disease	5	49.4(7.6)	?	1.5	1.5	X	X	As above
	HC	5	48(7.8)						
Gur et al (2000)	Schizophrenia	70	18-45	T1	1.5	1	✓	✓	Dorsal regions: coronal plane at the anterior tip of the CC.
	HC	81							
Cowell et al (1994)	HA - mid	96	18-40	T1	1.5	5	✓	X	Coronal appearance of the mamillary bodies
	- old	34	41-80						
Turetsky et al (1995)	Schizophrenia	**71**	29.7(3.2)	T2	1.5	5	✓	X	As Cowell
	HC	77	28(7.3)						
Kumar et al (1997)	Minor Depression	18	70.94 (8.69)	T2?	1.5	5?	X	X	As Cowell
	HC	31	69.7 (6.18)						
Kumar et al (2000)	MDD	51	74.3 (6.56)	T2?	1.5	5?	X	X	As Cowell
	HC	30	69.43 (6.09)						
Matsui et al (2000)	18-42yrs(59)			T1	1.5	5	✓	✓	As Cowell
Gur et al (2002)	HA	116	18-49yrs	T1	1.5	1	✓	✓	As Cowell
Gur et al (2004)	Schizophrenia	31	<50 yrs	T1	1.5	5	✓	✓	As Gur 2000
	HC	80							

Mohlman et al (2009)	GAD	15	60+yrs	T1	1.5	1.5	✓	X	As Gur 2000
	HC	15	60+yrs						
Hanninen et al (1997)	AAMI	43	69(5.4)	T1	1.5	1.5-2	✓	X	Posterior: coronal slice at AC. A straight line was drawn from the bottom of the lateral fissure (ventral insular sulcus) to the medially located choroidal fissure in order to separate the temporal lobe from the frontal lobe.
	HC	47	71.1(4)				□		
Harris et al (1994)	HA	57	31.5(7.9)	T2+pw	1.5	5	X	✓	As above (axial)
Schlaepfer et al (1994)	Schizophrenia	46	31.8(7.8)	T2+pw	1.5	5	X	✓	As above (axial)
	HC	60	31.6(8)						
	Bipolar	27	34.9(8.6)						
Schlaepfer et al (1995)	HA	60	30-43	T2+pw	1.5	5	X	✓	As above (axial)
Hasan et al (2011)	HA MZ Twins	12	19-36	T1	1.5	3	✓	X	Coronal slice at CC genu
	DZ Twins	12	24-29						
Haznedaar et al (1997)	ASD	7	24.3 (10.7)	T1	1.5	1.2	✓	X	ACC only
	HC	7	26.4(9.2)						
Haznedaar et al (2004)	Schizophrenia	27	38.3 (14.3)	T1	1.5	1.2	✓	X	ACC only
	Schizotypy	13	43.3 (13.6)						
	HC	32	41.8 (12.1)						
Hill et al (2003)	ADHD	23	9.35 (1.82)	T1	1.5	3	✓	X	Just dorsal and ventral PFC
	HC	24	9.36 (1.64)						
Howard et al (1995)	Delusional Disorder	19	79-86	T1	1.5	5	✓	X	Coronal plane at the CC genu
	Schizophrenia	31	75-80						

	HC	35	76-82						
Iordanova *et al.*, (2006)	LBD	8	74.6(9.2)	T1	1.5	1.5	X	X	Precentral sulcus for SFG & MFG, Circular sulcus of the insula or Sylvian fissure for the IFG.
	AD	8	78(5.7)						
	HC	9	62(9.2)						
James et al., (2004)	Schizophrenia	16		T1	1.5	3	X	✓	Posterior part of the genu of the CC in (coronal)
	HC	16							
Jernigan et al (1991)	HA	55	53.8 (14.1)	T1	1.5	5	X	X	A coronal plane bisecting the midpoint between CC genu and splenium
Jernigan et al (2001)	HA	78	62.9(18)	T1	?	1.5	X	X	As above
Sowell et al (2002)	HA	35	7-16	T1	?	1.5	✓	✓	Central sulcus
John *et al.*, (2006)	Schizophrenia	5	22.7(4.7)	?	?	1.25	✓	X	Precentral Sulcus & insular
frontal gyri only	HC	5	20.6(3.1)						
John *et al.*, (2007)	HA	20	22-40	T1	1.5	1.25	✓	✓	Frontal Pole only
John et al., (2009)	Schizophrenia	23	30.13(5.79)	T1	1.5	0.5	✓	✓	Frontal Pole only
	HC	23	30.13(5.05)						
Kates et al (2002)	Tourettes Syndrome	13	9.9(1.1)	T1	1.5	1.5	✓	X	Anteriormost coronal plane at which the precentral gyrus can be visualised.
	ADHD	13	9.4(1.2)					□	
	HC	13	10(1.5)						
Kelsoe et al (1988)	Schizophrenia	27	29(1)	T1	0.1	10	✓	X	Coronal plane at the CC genu
	HC	14	(31(1)						
Lacerda et al (2003) OFC only	HA	20		T1	1.5	1.5	✓	✓	Anterior perforated substance

Riffkin *et al.*, (2005)	OCD	18	36.1(12.99)	T1	1.5	1.5	✓	✓	ACC, OFC only
	Schizophrenia	18	35.9(11.97)						
	HC	18	34.6(11.82)						
Najt et al (2007)	Bipolar	14	15.5(3.2)	T1	1.5	1.5	✓	✓	Tip of the genu of the CC
	HCs	20	16.9(3.8)						
Girgis et al (2007)	ASD	11	8-12	T1	1.5	1.5	✓	✓	Tip of the genu of the CC
	HC	18	8-12						
Monkul et al (2007)	Unipolar	17		T1	1.5	1.5	✓	✓	Tip of the genu of the CC
	HC	17							
Chanen et al (2007) OFC only	BPD	20	17.3(1.1)	T1	1.5	1.5	✓	✓	As Riffkin 2005
	HC	20	19(2.2)						
Nery et al (2009)	Bipolar	28	34(11.9)	T1	1.5	1.5	✓	✓	As Lacerda
	HC	28	32.5(8.5)						
Atmaca et al (2010)	Body dysmorphic	12	29.6(4.8)	T1	1.5	2.4			As Lacerda, but no OFC sub-division
	HC	12	26.44 (5.7)						
Lindberg et al (2012)	AD	20	62	T1	1.5	1	X	X	DLPFC ACC OFC
	HC	30	62						
	FTD	12	59						
	SD	12	64						
Lai et al (2000)	Depressed	20	66.65(5.65)	T2	1.5	3	X	✓	OFC only
	HC	20	71.79(4.44)						

Lyoo et al (1998)	BPD	25	26.2(3.6)	T2	1.5	5	X	X	Optic chiasm - posteriormost tissue not included
	HC	25	24.9(4.1)						
MacLullich et al (2002)	HA - males	100	65-70	T1	1.9	1.5	X	X	coronal plane at CC genu
MacLullich et al (2006)	HA	20	65-70	T1	1.9	1.5	X	X	Coronal appearance of mamillary bodies
McCormick *et al.*, (2006)	Random selection of Schiz & HC	14		T1	1.5	1.5	✓	✓	ACC only
Asami et al (2008)	Panic Disorder	26	37.7 (10.1)	T1	1.5	1.5	✓	✓	As McCormick
	HC	26	38.2(9.7)						
McLaughlin et al (2009)	Affective disorders	20	40.2(7.6)	T1	1.5	1.5	✓	✓	Most anterior appearance of the frontal operculum, or when the rostrum and body of the CC could be clearly differentiated coronally.
Mueller et al (1998)	HA - young-old	11	70.36 (2.43)	T1	1.5	4	CC?	X	Central sulcus
	HA - middle-old	15	81.07 (2.81)						
	HA - old-old	20	86.96 (2.23)						
Murphy et al (1993)	Turner's syndrome	18	30(7)	?	0.5+1.5	5-7mm	X	X	Supratemporal structures anterior to the aqueduct of Sylvius
	HC	19	27(8)						
Murphy et al (1996)	HA - M	35	44(23)	?	0.5+1.5	5-7mm	X	X	As above
	- F	34	50(21)						
van Amelsvoort et	VCFS	10	32(9)	?	1.5	1.5	✓	X	As above
	HC	13	37(10)						

al (2001)									
McAlonan et al (2002)	Asperger	21	32(10)		1.5	1..5	✓	X	As above
	HC	24	33(7)				☐	☐	
Nakamura et al (2008)	Schizophrenia	24	39.1 (10.3)	T1	1.5	0.94	✓	X	Circular sulcus of the insula
	HC	25	41.1(9.1)						
Nagel et al (2006)	Healthy teens	65	14.95(1.9)	T1	1.5	1.3	✓	✓	Coronal plane at the A-C.
Medina et al (2008)	Alcohol Use Disorder	14	15-18	T1	1.5	1.3	✓	X	As above
	HC	17	15-18						
Medina et al (2009)	Marijuana Users	16	16-18	T1	1.5	1.3	✓	X	As above
	HC	16	16-18						
Noga et al (1995) ACC only	Schizophrenia	14	31(4.7)	T1	1.5	3	X	X	6mm posterior to the coronal slice containing a clear view of the septum pellucidum.
	HC	14	32.2(6)				Acquired at AC-PC		
Kaur et al (2005)	Bipolar	16	15.5(3.4)	T1	1.5	1.5	X	X	Coronal anterior appearance of the A-C
	HC	21	16.9(3.8)				Unclear		
Nolan et al (2002)	MDD	22	9-17yrs	?	1.5	1.5	✓	X	CC genu
	HC	22	9-17yrs						
Pantel et al (1997)	Depressed	19	72.4(8.8)	T1	1.5	1.25	X	X	CC - splenium posterior aspects excluded
	AD	27	71.9(8)						
	HC	13	68.2(5.3)						
Paus et al (1996)	HA	105	25.2(7.7)	T1	1.5	0.86	✓	✓	1cm posterior to A-C (coronal)

Rademacher *et al.*, (1992)	THIS PROTOCOL WAS NOT APPLIED AT THIS STAGE.						Not spec'd - schema only		Medial:Coronal plane at AC Lateral: PrCS Measures central gyrus separately. Ventrolateral: insular cortex
Caviness *et al.*, (1996)	HA	15	NR	T1	1.5	3	✓	✓	Operationalized the Rademacher schema
Grachev et al (1997)	Trichotillomania (10)	10	31(10)	T1	1.5	3	✓	✓	As Caviness
	HCs (10)	10	28.5 (11.2)				☐		
Tzourio *et al.*, (1997)	No measurements made in this study??								Precentral Sulcus
Kennedy et al (1998)	HA M	10	27.4(5)	T1	1.5	3	✓	✓	As Caviness
	F	10	26.9(5.3)						
Goldstein et al (1999)	Schizophrenia	29	44.8(10.5)	T1	1.5	3.1	✓	✓	As Caviness
	HC	26	39.8(11.5)						
Szeszko et al (1999a)	Schizophrenia	19		T1	1	3.1	✓	X	MODIFIED CAVINESS
	HC	26							
Szeszko et al (1999b)	OCD	26		T1	1	1.5	✓	X	MODIFIED CAVINESS
	HC	26							
Szeszko et al (2000)	Schizophrenia	35		T1	1	1.5	✓	X	MODIFIED CAVINESS
Allen et al (2002)	HA M	23	32.1(8.8)	?	1.5	1.5-1.6	✓	X	As Caviness
	F	23	32.6(7.5)						
Rauch et al (2003)	PTSD	9		T1	1.5	1.5	✓	✓	As Caviness
	Controls	9							

Takeoka et al (2003)	Ramsussen Encaphalitis	1	5	T1	1.5	1.5	✓	✓	As Caviness
Fossé et al (2004)	ALI	16	6-14	T1	1.5	1.5	✓	✓	Only measures IFG
	ALN	6	6-14						
	SLI	9	6-14						
	HC	18	6-14						
Allen et al (2005)	HA M	43	49.4 (20.8)	T1	1.5	1.5-1.6	✓	X	As Caviness
	F	44	47(16.7)				☐		
Frazier et al (2005)	Bipolar Disorder	32	11.2(2.8)	T1	1.5	1.5	✓	X	As Caviness
	HC	15	11.2(3)						
Rupp et al (2005)	Schizophrenia	33	25.4(4.7)	T1	1.5	1-1.2	✓	✓	As Szeszko
	HC	40	26.2(4.7)				☐		
Allen et al (2006)	Anoxic	13	51.6	?	1.5	1.5-1.6	✓	X	As Caviness
	HC M	43	49.4 (20.8)				☐		
	F	44	47(16.7)						
Seidman et al (2006)	ADHD	24	18-59	T1	1.5	1.5	✓	✓	As Caviness
	HC	18	18-59						
Rosso et al (2010)	FHR Schizophrenia	27	19(4.2)	T1	1.5	1.33	✓	✓	As Caviness
	HC	48	17.7(3.7)						
Betjemann et al (2010)	HA MZ twins	41	15.4 (2.69)	T1	1.5	3	✓	✓	As Caviness
	DZ twins	30	15.4 (2.69)						
Raine et al (1991)	HA	17	33.9 (11.8)	T1	0.15	10	X	X	Coronal plane at CC genu
Rankin et al	FTD	27	64.9(9.8)	T1	NR	NR	X	✓	First slice anterior to the optic chiasm

(2004)									
Ranta et al., (2009)	ADHD	15	8-12	T1	1.5	1.2	✓	✓	Cut plane where the ascending ramus of the Sylvian fissure and the fissure both meet. In medial regions, the posterior boundary is defined by the anterior extent of the lateral ventricle.
	HC	15	8-12						
Ratnanather et al (2001)	HC	5	NR	T1	1.5	1	X	✓	5mm anterior to the CC coronally
Raz *et al.*, (1995)	Down's Syndrome	13	22-50yrs	T1	1.5	0.86	✓	X	Rostral to the genu of the CC.
	HA	12	23-49yrs						
Raz et al (1997)	HA	148	18-77yrs	T1	1.5	0.86	✓	X	As Raz *et al.*, (1995)
Head et al (2002)	HA	68	49(16.65)	T1	1.5	0.86	✓	X	As Raz *et al.*, (1995)
Hesslinger et al (2002)	ADHD	8	31.4(4.4)	T1	2	0.98	✓	X	As Raz *et al.*, (1995)
	HC	17	30.2(7.9)						
Raz et al (2003)	Hypertensive	40	61.56 (11.68)	T1	1.5	1.3	✓	X	As Raz *et al.*, (1995)
	HC	40	61.63 (11.25)						
van Elst et al (2003)	BPD	8	33.5(6.3)	T1	2	0.98	✓	X	As Raz *et al.*, (1995)
	HC	8	30.5(5.1)						
Raz et al (2004)	HC	200	29-63	T1	1.5	1.3	✓	X	As Raz *et al.*, (1995)
Raz et al (2005)	HA	72	52.49	T1	1.5	1.3	✓	X	As Raz *et al.*, (1995)
Raz et al (2007)	Vascular Risk	23	47-77	T1	1.5	1.3	✓	X	As Raz *et al.*, (1995)
	HC	23	45-75						
Raz et al (2010)	HA	40	49-85	T1	1.5	1.5	✓	X	As Raz *et al.*, (1995)

Rosen et al (2002)	tvFTD	9	66(8.3)	T1	1.5	1.5	✓	✓	First coronal slice anterior to the optic chiasm
	HC	10	60.3(8.1)				☐	☐	
Salat et al (1999a)	AD	22	69.8	T2	1.5	4	✓	✓	Coronal slice at CC genu
	HC - old	22	88.9						
	HC - young	26	71						
Salat et al (1999b)	AD	30	64-75	T1	1.5	4	✓	✓	Coronal slice at CC genu
	HC	17	66-77						
Salat et al (2001)	AD	22	69.8	T2	1.5	4	✓	✓	Coronal slice at CC genu
	HC - old	22	88.9						
	HC - young	26	71						
Salat et al (2002)	HA - young	20	21-43	T2	1.5	4	✓	✓	Coronal slice at CC genu
	- old	31	72-94						
Sanches et al., (2009) DLPFC only	HA	10	26.8(6.5)	T1	1.5	1	✓	✓	The slice immediately anterior to the most rostral slice where the CC can be viewed as a bridge in coronal view
Sanfilipo et al (2000)	Schizophrenia	53	38.7(5.5)	T1	1.5	2.8	✓	✓	Slice anterior to the CC genu
	HC	29	35.8(8.7)						
Sanfilipo et al (2002)	Schizophrenia	62	38.8(5.3)	T1	1.5	2.8	✓	✓	Slice anterior to the CC genu
	HC	27	13.1(1.6)						
Seidman et al (1994)	Schizophrenia	19	33.6(8.4)	T1	1.5	5-6	X	✓	Most posterior coronal slice passing through the CC genu
Gilbert et al (2001)	Schizophrenia	16	26.56 (7.3)	T1	1.5	1.5	✓	✓	Most posterior coronal slice passing through the CC genu
	HC	25	23.6 (4.66)						
Prasad et al (2005)	Schizophrenia	25/19 *		T1	1.5	1.5	✓	✓	Most posterior coronal slice passing through the CC genu

Semendeferi et al (1997)	HA	4		T1		1.2-1.6	X	X	Laterally, the Central Sulcus
Semendeferi & Damasio (2000)	HA	10		T1	1.5	1.6	X	X	Central sulcus
Semendeferi et al (2002)	HA	10		T1	1.5	1.6	X	X	Central sulcus
Schenker et al (2005)	Sub-regional	10	32.4 (8.55)	T1	1.5	1.6	X	X	Central sulcus
Sherwood et al (2011)	HA	87	22-88	T1	1.5+3	1.5-1.6	✓	X	Central Sulcus
Soininen et al (1995)	AAMI	16	68(7)	T1	1.5	1.5-1.8	X	X	Coronal slice at AC
	HC	16	70(5)						
Laakso et al (1995)	AD	32	69(8)	T1	1.5	1.5-1.8			Coronal slice at AC Excluded temporal lobe in posterior slices by drawing a line from the lateral fissure to the choiroid fissure, then above the optic tract to the midline.
	HC	16	70(5)			but traced at 5mm			
Suga et al (2010)	Schizophrenia	29	30.9(6.4)	T1	1.5	1.5	X	X	BA 44 & 45 after Tomaiuolo 1999 and Knaus 2006
	HC	29	28.9(3.9)						
Yamasaki et al (2010)	hfASD	13	28.5(10.2)	T1	1.5	1.5	X	X	BA 44 & 45 after Tomaiuolo 1999 and Knaus 2006
	HC	11	29.1(3.1)						
Suzuki *et al.*, (2005)	Schizotypal disorder	25	25(5.7)	T1	1.5	1	✓	✓	Pre-central Sulcus
	Schizophrenia	53	25.3(5)				☐	☐	
	HC	59	24.3(5.3)						
Matsui et al (2008)	Schizotypal disorder	25	24.9(4.5)	T1	1.5	1	✓	✓	Pre-central Sulcus
	Schizophrenia	35	23.8(5.2)						
	HC	19	27.1(6.7)						

Takahashi *et al.*, (2002a) ACC only	Schizophrenia	40	26.1(5)	T1	1.5	1	✓	✓	Anterior Commisure
	HC	40	25.31 (5.8)						
Takahashi et al., (2002b)	Schizotypal disorder	24	22.7(4.5)	T1	1.5	1	✓	✓	Anterior Commisure
	Schizophrenia	40	26.1(5)						
	HC	48	24.2(5.9)						
Takahashi *et al.*, (2003)	Schizophrenia	40	26.1(5)	T1	1.5	1	✓	✓	Anterior Commisure
	HC	40	25.31 (5.8)						
Tisserand *et al.*, (2002)	HA	57	55.7(16.2)	T1	1.5	1.5	✓	✓	Precentral sulcus for DL 3 slices posterior to AC for ACC
Jones et al (2006)	AD	10	51.8(7.4)	T1	1.5	1.5			Precentral sulcus
	HC	10	51.0(8.0)						
Burgmans et al (2009)	HC - no decline	35	69.1(7.7)	T1	1.5	1.5	✓	✓	Precentral sulcus for the DL and IFG regions.
	HC - decline	30	69.2(8.1)						Coronal plane at inner CC for OFC
	Dementia	9	73.8(4.3)						
Uylings et al (2010)	Post-mortem	32	23-86	T1	1.5	1.17	✓	X	OFC only. As Tisserand - a coronal plane at the inner curvature of the CC.
Tomaiuolo et al (1999)	HA	50	25(10.2) 25.6(7.2)	T1	1.5	0.86	✓	✓	Precentral sulcus
Van Petten et al (2004)	HA	48	65-85	T1 Axial	1.5	4 Summed 4x1mm axial slices.		Manual	Precentral sulcus Frontal gyri only (no ACC or OFC).
Wible et al (1995)	Schizophrenia	15		T1	1.5	1.5	X	✓	3 slices anterior to the 1st slice of temporal stem
	HC	15							
Wible *et al.*, (1997)	HA	15	20-55	T1	1.5	1.5	X	✓	As above
Rosenberg et al	OCD	19	12(3)		1.5	1.5	✓	✓	Coronal plane at genu of CC

(1997)	HC	19	12(3)						
Hirayasu et al (2001)	Shizophrenia	17	22.8(3.6)	T1	1.5	1.5	?	✓	As Wible 1995
	Affective disorder	17	22.6(3.8)						
	HC	17	22.2(3.8)						
Wible et al (2001)	Schizophrenia	17	44	T1	1.5	1.5	?	✓	As Wible 1995
	HC	17	40						
Yamasue *et al.*, (2004)	Schizophrenia	27	30.4(7.9)	T1	1.5	0.9375	✓	✓	Same PFC boundary, but introduced a more posterior ACC boundary - coronal appearance of the mammilary bodies
	HC	27	30(5.6)						
Wilde et al (2005)	TBI	16	12.9(2.5)	T1	1.5	1	✓	✓	Coronal slice just anterior to the CC genu.
	HC	16	12.8(2.4)				☐	☐	
Woods et al., (1996) ONLY AREA FOR FL ROIs	Schizophrenia	19	35	T1	1.5	5-6	X	✓	Coronal plane passing through the CC genu. They recognise this is a PFC measure.
	HC	19	35						
Maher et al (1998)	Schizophrenia	16	33.75(8.44)	?	1.5	5-6	X	X	As above
Woodward et al (2005)	Gulf War Veterans	36	38	T1	1.5	1.5-1.7	✓	✓	TT sector E1. ACC only
	Vietnam Veterans	63	56						
Yucel et al (2008)	MDD	65	28.8(10.3)	T1	1.5/3	1.2			vACC only.
	HC	93	28.4(10.7)						
Zipursky et al (1992)	Schizophrenia	22	34.1(5.5)	?	1.5	5	✓	✓	Coronal plane at the CC genu
	HC	20	36.2(7)					"cortical ring"	
Sullivan et al (1996)	Schizophrenia	34	36.9(7.8)			5	✓	✓	Coronal plane at the CC genu
	HC	47	37.9(9.2)						

Pfefferbaum et al (1997)	HC - young	65	32.9(6.5)		1.5	5	✓	✓	Coronal plane at the CC genu
	HC - old	27	53.2(6)						
	Alcoholic - young	33	37.5(4.5)						
	Alcoholic - old	29	52.7(6)						
Fama et al (1997)	AD	50	51-87		1.5	5	✓	✓	Coronal plane at the CC genu
	HC	136	20-84						
Mathalon et al (2001)	Schizophrenia	24	39.4(6.4)		1.5	5	✓	✓	Coronal plane at the CC genu
	HC	25	40.7(8.5)						
Fama et al (2004)	AD	50	71.4(7.4)	?	1.5	5	X	X	Coronal plane at the CC genu
Greenwood et al (2005)	Neurofibramatosis I Relatives	36	9.3(2.3)	?	1.5	3	✓	X	Coronal plane at the CC genu
	Relatives	36	9.5(2.5)						
Zhou et al (2005)	Schizophrenia	59	25	T1	1.5	1	✓	X	Central sulcus
	HC	58	25						
Zuffante et al (2001)	Schizophrenia	23	46.5(4.2)	T1	1.5	1	✓	X	TT y=32
	HC	23	43.3(9.6)						

Appendix 2: References for frontal lobe systematic review

Allen JS, Bruss J, Brown CK and Damasio H. Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. *Neurobiology of Aging*, 26(9): 1245-1260, 2005.

Allen JS, Damasio H and Grabowski TJ. Normal neuroanatomical variation in the human brain: an MRI-volumetric study. *American Journal of Physical Anthropology*, 118(4), 341-58, 2002.

Allen JS, Tranel D, Bruss J and Damasio H. Correlations between regional brain volumes and memory performance in anoxia. *Journal of Clinical and Experimental Neuropsychology*, 28(4), 457-76, 2006.

Almeida OP, Burton EK, Ferrier N, McKeith IG and O'Brien JT. Depression with late onset is associated with right frontal lobe atrophy. *Psychological Medicine*, 33(4): 675-681, 2003.

Antonucci AS, Gansler DA, Tan S, Bhadelia R, Patz S and Fulweiler C. Orbitofrontal correlates of aggression and impulsivity in psychiatric patients. *Psychiatry Research: Neuroimaging*, 147:213-220, 2006.

Asami T, Hayano F, Nakamura M, Yamasue H, Uehara K, Otsuka T, Roppongi T, Nihashi N, Inoue T and Hirayasu Y. Anterior cingulate cortex volume reduction in patients with panic disorder. *Psychiatry and Clinical Neurosciences*, 62(3): 322-330, 2008.

Atmaca M, Bingol I, Aydin A, Yildirim H, Okur I, Yildirim MA, Mermi O and Gurok MG. Brain morphology of patients with body dysmorphic disorder. *Journal of Affective Disorders*, 123(1-3): 258-263, 2010.

Aylward, Elizabeth H, Augustine, A., Li, Q., & Barta, P. E. Measurement of frontal lobe volume on magnetic resonance imaging scans. *Psychiatry research: Neuroimaging*, *75*, 23-30, 1997.

Baaré WF, Hulshoff PHE, Hijman R, Mali WP, Viergever MA, and Kahn RS. Volumetric analysis of frontal lobe regions in schizophrenia: relation to cognitive function and symptomatology. *Biological Psychiatry*, 45(12): 1597-605, 1999.

Backman L, Robins-Wahlin T-B, Lundin A, Ginovart N and Farde L. Cognitive deficits in Huntington's disease are predicted by dopaminergic PET markers and brain volumes. *Brain*, 120:2207-2217, 1997.

Ballmaier M, Toga A, Blanton R, Sowell ER, Lavretsky H, Peterson BS, Pham D and Kumar A.Anterior cingulate, gyrus rectus, and orbitofrontal abnormalities in elderly depressed patients: an MRI-based parcellation of the prefrontal cortex. *American Journal of Psychiatry*, 161: 99-108, 2004.

Bartzokis G, Mintz J, Marx P, Osborn D, Gutkind D, Chiang F, Phelan CK and Marder SR. Reliability of in vivo volume measures of hippocampus and other brain structures using MRI. *Magnetic resonance imaging,* 11: 993-1006, 1993.

Berryhill P, Lilly MA. Levin HS, Hillman GR, Mendelsohn D, Brunder DG, Fletcher JM, Kufera J, Kent TA, Yeakley J, Bruce D and Eisenberg HM.. Frontal Lobe Changes after Severe Diffuse Closed Head Injury in Children: A Volumetric Study of Magnetic Resonance Imaging. *Neurosurgery,* 37(3): 392–400, 1995.

Betjemann RS, Johnson EP, Barnard H, Boada R, Filley CM, Filipek PA, Willcutt EG, DeFries JC and Pennington BF. Genetic covariation between brain volumes and IQ, reading performance, and processing speed. *Behavior Genetics*, *40*(2), 135-45, 2010.

Beyer JL, Kuchibhatla M, Payne ME, Macfall J, Cassidy F and Krishnan KRR. Gray and white matter brain volumes in older adults with bipolar disorder. *International Journal of Geriatric Psychiatry*, 24:1445-1452, 2009.

Bjork JM, Momenan R and Hommer DW. Delay discounting correlates with proportional lateral frontal cortex volumes. *Biological Psychiatry*, 65(8): 710-713, 2009.

Blanton RE, Levitt JG, Peterson JR, Fadale D, Sporty ML, Lee M, To D, Mormino EC, Thompson PM, McCracken JT and Toga AW. Gender differences in the left inferior frontal gyrus in normal children. *Neruoimage,* 22:626-636, 2004.

Boes AD, Murko V, Wood JL, Langbehn DR, Canady J, Richman L and Nopoulos P. Social function in boys with cleft lip and palate: Relationship to ventral frontal cortex morphology. *Behavioural Brain Research,* 181:224-231, 2007.

Bokde AL, Teipel SJ, Zebuhr Y, Leinsinger G, Gootjes L, Schwarz R, Buerger K, Scheltens P, Moeller HJ and Hampel H. A new rapid landmark-based regional MRI segmentation method of the brain. *Journal of Neurological Science,* 194: 35-40, 2002.

Bokde ALW, Teipel SJ, Schwarz R, Leinsinger G, Buerger K, Moeller T, Möller H-J and Hampel H. Reliable manual segmentation of the frontal, parietal, temporal, and occipital lobes on magnetic resonance images of healthy subjects. *Brain Research Protocols*, 14(3), 135-145, 2005.

Botteron KN, Raichle ME, Drevets WC, Heath AC and Todd RD. Volumetric reduction in left subgenual prefrontal cortex in early onset depression. *Biological Psychiatry*, 51(4): 342-344, 2002.

Brambilla P, Nicoletti MA, Harenski K, Sassi RB, Mallinger AG, Frank E, Kupfer DJ, Keshavan MS and Soares JC. Anatomical MRI study of subgenual prefrontal cortex in bipolar and unipolar subjects. *Neuropsychopharmacology,* 27(5): 792-799, 2002.

Bremner JD, Bronen RA, Erasquin GD, Vermetten E, Staib LH, Ng CK, Soufer R, Charney DS and Innis RB. Development and reliability of a method for using

magnetic resonance imaging for the definition of regions of interest for Positron Emission Tomography. *Clinical Positron Imaging,* 1(3): 145-159, 1998.

Bremner JD, Narayan M, Anderson ER, Staib LH, Miller HL and Charney DS. Hippocampal volume reduction in major depression. *American Journal of Psychiatry,* 157:115-117, 2000.

Bremner JD, Vythilingam M, Vermetten E, Nazeer A, Adil J, Khan S, Staib LH and Charney DS. Reduced volume of orbitofrontal cortex in major depression. *Biological Psychiatry*, 51(4): 273-279, 2002.

Burgmans S, van Boxtel MPJ, Smeets F, Vuurman EFPM, Gronenschild EHBM, Verhey FRJ, Uylings HBM and Jolles J. Prefrontal cortex atrophy predicts dementia of a six-year period. *Neurobiology of Aging,* 30(9): 1413-1419.

Carper RA and Courchesne E. Inverse correlation between frontal lobe and cerebellum sizes in children with autism. *Brain,* 123(4): 836-44, 2000.

Carper RA and Courchesne E. Localized enlargement of the frontal cortex in early autism. *Biological Psychiatry*, 57(2): 126-133, 2005.

Casey BJ, Castellanos FX, Giedd JN, Marsh WL, Hamburger SD, Schubert AB, Vauss YC, Vaituzis AC, Dickstein DP, Sarfatti SE and Rapoport JL. Implication of right frontostriatal circuitry in response inhibition and attention deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry,* 36(3):374-383, 1997.

Castellanos FX, Giedd JN, Marsh WL, Hamburger SD, Vaituzis AC, Dickstein DP, Sarfatti SE, Vauss YC, Snell JW, Rajapakse JC and Rapoport JL. Quantitative brain mangnetic resonance imaging in attention-deficit hyperactivity disorder. *Archives of General Psychiatry*, 53: 607-616, 1996.

Caviness VS, Meyer J, Makris N and Kennedy DN. MRI-Based Topographic Parcellation of Human Neocortex: An Anatomically Specified Method with Estimate of Reliability. *Journal of Cognitive Neuroscience*, 8(6): 566-587, 1996.

Chanen AM, Velakoulis D, Carison K, Gaunson K, Wood SJ, Yuen HP, Yucel M, Jackson HJ, McGorry PD and Pantelis C. Orbitofrontal, amygdala and hippocampal volumes in teenagers with first-presentation borderline personality disorder. *Psychiatry Research: Neuroimaging,* 163:116-125, 2008.

Chemerinski E, nopoulos PC, Crespo-Facorro B, Andreasen NC and Magnotta V. Morphology of the ventral frontal cortex in schizophrenia: Relationship with social dysfunction. *Biological Psychiatry,* 52:1-8, 2002.

Coffey CE, Weiner RD, Djang W, Figiel G, Soady S, Patterson L, Holt PD, Spritzer CE and Wilkinson WE. Brain anatomic effects of electroconvulsive therapy. *Archives of General Psychiatry*, 48: 1013-1021, 1991.

Coffey CE, Lucke JF, Saxton JA, Ratcliff G, Unitas LJ, Billig B and Bryan RN. Sex differences in brain aging: a quantitative magnetic resonance imaging study. *Archives of Neurology*, 55(2): 169-79, 1998.

Colchester A, Kingsley D, Lasserson D, Kendall B, Bello F, Rush C, Stevens TG, Goodman G, Heilpern G, Stanhope N and Kopelman MD. Structural MRI volumetric analysis in patients with organic amnesia, 1: methods and comparative findings across diagnostic groups. *Journal of Neurology, Neurosurgery, and Psychiatry*, 71(1): 13-22, 2001.

Convit A, Wolf OT, de Leon MJ, Patalinjug M, Kandil E, Caraos C, Scherer A, Saint Louis LA and Cancro, R. Volumetric analysis of the pre-frontal regions: findings in aging and schizophrenia. *Psychiatry Research: Neuroimaging*, 107(2): 61-73, 2001.

Coryell W, Nopoulos P, Drevets W, Wilson T and Andreasen NC. Subgenual prefrontal cortex volumes in major depressive disorder and schizophrenia: diagnostic specificity and prognostic implications. *The American Journal of Psychiatry*, 162(9): 1706–1712, 2005.

Cowell PE, Turetsky BI, Gur RC, Grossman RI, Shtasel DL and Gur RE. Sex differences in aging of the human frontal and temporal lobes. *The Journal of Neuroscience*, 14(8): 4748-55, 1994.

Crespo-Facorro B, Kim JJ, Andreasen NC, O'Leary DS, Wiser AK, Bailey JM, Harris, G and Magnotta VA. Human frontal cortex: an MRI-based parcellation method. *NeuroImage*, 10(5): 500-519, 1999.

Crespo-Facorro B, Kim J, Andreasen NC, O'Leary DS and Magnotta V. Regional frontal abnormalities in schizophrenia: a quantitative gray matter volume and cortical surface size study. *Biological Psychiatry*, 48(2): 110-119, 2000a.

Croxson PL, Johansen-Berg H, Behrens TEJ, Robson MD, Pinsk MA, Gross CG, Richter W, Kastner S and Rushworth MFS. Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *The Journal of Neuroscience*, 25(39): 8854-8866, 2005.

De Bellis MD, Narasimhan A, Thatcher DL, Keshevan MS, Soloff P and Clark DB. Prefrontal cortex, thalamus, and cerebellar volumes in adolescents and young adults with adolescent-onset alcohol use disorders and comorbid mental disorders. *Alcoholism: Clincal and Experimental Research*, 29(9):1590-1600, 2005.

Drevets WC, Price J, Simpson J and Todd R. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature*, 386: 824-827, 1997.

Egan MF, Duncan CC, Suddath RL, Kirch DG, Mirsky AF and Wyatt RJ. Event-related potential abnormalities correlate with structural brain alterations and clinical features in patients with chronic schizophrenia. *Schizophrenia Research*, 11:259-271, 1994.

Elderkin-Thompson V, Ballmaier M, Hellemann G, Pham D and Kumar A. Executive function and MRI prefrontal volumes among healthy older adults. *Neuropsychology*, 22(5): 626-637, 2008.

Elderkin-Thompson V, Hellemann G, Pham D and Kumar A. Prefrontal brain morphology and executive function in healthy and depressed elderly. *International Journal of Geriatric Psychiatry*, 24(5): 459-468, 2009.

Exner C, Koshack J and Irle E. The differential role of premotor frontal cortex and basal ganglia in motor sequence learning: Evidence from focal basal ganglia lesions. *Learning and Memory*, 9:376-386, 2002.

Exner C, Weniger G, Schmidt-Samoa C and Irle E. Reduced size of the pre-supplementary motor cortex and impaired motor sequence learning in first-episode schizophrenia. *Schizophrenia Research*, 84(2-3): 386-96, 2006.

Fama R, Sullivan EV, Shear PK, Marsh L, Yesavage JA, Tinklenberg JR, Lim KO and Pfefferbaum A.. Selective cortical and hippocampal volume correlates of Mattis Dementia Rating Scale in Alzheimer disease. *Archives of Neurology, 54*(6): 719-28, 1997.

Fama R, Marsh L and Sullivan EV. Dissociation of remote and anterograde memory impairment and neural correlates in alcoholic Korsakoff syndrome. *Journal of the International Neuropsychological Society*, 10(3): 427-41, 2004.

Filipek PA, Semrud-Clikeman M, Steingard RJ, Renshaw PF, Kennedy DN and Biederman J. Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology*, 48(3): 589-601, 1997.

Flashman LA, McAllister TW, Johnson SC, Rick JH, Green RL and Saykin AJ. Specific frontal lobe subregions correlated with unawareness of illness in schizophrenia: a preliminary study. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 13(2): 255-257, 2001.

Fornito A, Whittle S, Wood SJ, Velakoulis D, Pantelis C and Yücel M. The influence of sulcal variability on morphometry of the human anterior cingulate and paracingulate cortex. *NeuroImage*, 33(3): 843-854, 2006.

Fossé LD, Hodge SM, Makris N, Kennedy DN, Caviness VS, McGrath L, Steele S, Ziegler DA, Herbert MR, Frazier JA, Tager-Flusberg H and Harris GJ. Language-association cortex asymmetry in autism and specific language impairment. *Annals of Neurology*, 56:757-766, 2004.

Foundas AL, Weisberg A, Browning CA and Weinberger DR. Morphology of the frontal operculum: A volumetric magnetic resonance imaging study of the pars triangularis. *Journal of Neuroimaging*, 11:153-159, 2001.

Frazier J, Breeze J, Makris N, Giuliano A, Herbert M, Seidman L, Biederman J, Hodge SM, Dieterich ME, Gerstein ED, Kennedy DN, Rauch SL, Cohen BM and Caviness VS. Cortical gray matter differences identified by structural magnetic

resonance imaging in pediatric bipolar disorder. *Bipolar Disorders*, 7(6): 555-569, 2005.

Fukui T. Volumetric study of lobar atrophy in Pick complex and Alzheimer's disease. *Journal of the Neurological Sciences*, 174(2): 111-121, 2000.

Gansler DA, McLaughlin NCR, Iguchi L, Jerram M, Moore DW, Bhadelia R and Fulwiler C. A multivariate approach to aggression and the orbital frontal cortex in psychiatric patients. *Psychiatry Research*, 171(3): 145-154, 2009.

Geroldi C, Pihlajamaki M, Laakso MP, DeCarli C, Beltramell A, Bianchetti A, Soininen H, Trabicchi M and Frisoni GB. APOE-4 is associated with less frontal and more temporal lobe atrophy in Alzheimer's Disease. *Neurology*, 53(8): 1825-1832, 1999.

Gilbert AR. Thalamic Volumes in Patients With First-Episode Schizophrenia. *American Journal of Psychiatry*, 158(4), 618-624, 2001.

Ginovart N, Lundin A, Farde L, Halldin C, Backman L, Swahn CG, Pauli S and Sedvall G. PET study of the pre- and post-synaptic dopaminergic markers for the neurodegenerative process in Huntington's disease. *Brain*, 120: 503-514, 1997.

Girgis R, Munshew NJ, Melhem NM, Nutche JJ, Keshavan MS and Hardan AY. Volumetric alterations of the orbitofrontal cortex in autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 31:41-45, 2007.

Gold SM, Dziobek I, Rogers K, Bayoumy A, McHugh PF and Convit A. Hypertension and hypothalamo-pituitary-adrenal axis hyperactivity affect frontal lobe integrity. *The Journal of Clinical Endocrinology and Metabolism*, 90(6): 3262-3267, 2005.

Goldstein JM, Goodman JM, Seidman LJ, Kennedy DN, Makris N, Lee H, Tourville J, Caviness VS, Faraone SV and Tsuang MT. Cortical abnormalities in schizophrenia identified by structural magnetic resonance imaging. *Archives of General Psychiatry*, 56(6): 537-547, 1999.

Grachev ID. MRI-based morphometric topographic parcellation of human neocortex in trichotillomania. *Psychiatry and Clinical Neurosciences*, 51(5): 315–321, 1997.

Greenwood RS, Tupler LA, Whitt JK, Buu A, Dombeck CB, Harp AG, Payne ME Eastwood JD, Krishnan KRR and MacFall JR.. Brain morphometry, T2-weighted hyperintensities, and IQ in children with neurofibromatosis type 1. *Archives of Neurology*, 62(12): 1904-8, 2005.

Gur RE, Cowell PE, Latshaw A, Turetsky BI, Grossman RI, Arnold SE, Bilker WBand Gur RC. Reduced dorsal and orbital prefrontal gray matter volumes in schizophrenia. *Archives of General Psychiatry*, 57(8): 761-768, 2000.

Gur RE, Kohler C, Turetsky BI, Siegel SJ, Kanes SJ, Bilker WB, Brennan AR and Gur RC. A sexually dimorphic ratio of orbitofrontal to amygdala volume is altered in schizophrenia. *Biological Psychiatry*, 55(5): 512-517, 2004.

Gur, RC, Gunning-Dixon F, Bilker WB and Gur, RE. Sex differences in temporo-limbic and frontal brain volumes of healthy adults. *Cerebral Cortex,* 12(9), 998-1003, 2002.

Hänninen T, Hallikainen M, Koivisto K, Partanen K, Laakso MP, Riekkinen P and Soininen H. Decline of frontal lobe functions in subjects with age-associated memory impairment. *Neurology*, 48(1): 148-53, 1997.

Harris, GJ, Barta PE, Peng LW, Lee S, Brettschneider PD, Shah A, Henderer JD, Schlaepfer TE and Pearlson GD. MR gray and white matter segmentation using manual thresholding: Dependence on image brightness. *American Journal of Neuroradiology,* 15:225-230, 1994.

Hasan A, McIntosh AM, Droese U-A, Schneider-Axmann T, Lawrie SM, Moorhead TW, Tepest R,Maier W, Falkai R and Wobrock T. Prefrontal cortex gyrification index in twins: an MRI study. *European Archives of Psychiatry and Clinical Neuroscience*, 261(7): 459-65, 2011.

Hastings RS, Parsey RV, Oquendo MA, Arango V and Mann JJ. Volumetric analysis of the prefrontal cortex, amygdala, and hippocampus in major depression. *Neuropsychopharmacology,* 29(5): 952-959, 2004.

Haznedar MM, Buchsbaum MS, Metzger M, Solimando A, Spiegel-Cohen J and Hollander E. Anterior cingulate gyrus volume and glucose metabolism in autistic disorder. *American Journal of Psychiatry*, 154:1047-1050, 1997.

Haznedar MM, Buchsbaum MS, Hazlett EA, Shihabuddin L, New A and Siever LJ. Cingulate gyrus volume and metabolism in the schizophrenia spectrum. *Schizophrenia Research,* 71:249-262, 2004.

Head D, Raz N, Gunning-Dixon F, Williamson A and Acker JD. Age-related differences in the course of cognitive skill acquisition: The role of regional cortical shrinkage and cognitive resources. *Psychology and Aging*, 17(1): 72-84, 2002.

Hesslinger B, van Elst LT, Thiel T, Haegele K, Hennig J and Ebert D. Frontoorbital volume reductions in adult patients with attention deficit hyperactivity disorder. *Neuroscience Letters,* 328:319-321, 2002.

Hill DE, Yeo RA, Campbell RA, Hart B, Vigil J and Brooks W. Magnetic resonance imaging correlates of attention-deficit/hyperactivity disorder in children. *Neuropsychology,* 17(3): 496-506, 2003.

Hirayasu Y, Shenton ME, Salisbury DF, Kwon JS, Wible CG, Fischer IA, Yurgelun-Todd D, Zarate C, Kikinis R, Jolesz FA and McCarley RW. Subgenual cingulate cortex volume in first-episode psychosis. *American Journal of Psychiatry,* 156: 1091-1093, 1999.

Hirayasu Y, Tanaka S, Shenton ME, Salisbury DF, DeSantis MA, Levitt JJ, Wible C, Yurgelun, Todd D, Kikinis R, Jolesz FA and McCarley RW. Prefrontal gray matter volume reduction in first episode schizophrenia. *Cerebral Cortex,* 11(4): 374-381, 2001.

Howard R, Mellers J, Petty R, Bonner D, Menon R, Almeida O, Graves M, Renshaw C and Levy R. Magnetic resonance imaging volumetric measurements of the superior temporal gyrus, hippocampus, parahippocampal gyrus, frontal and temporal lobes in late paraphrenia. *Psychological medicine*, 25(3): 495-503, 1995.

Iordanova B, Rosenbaum D, Norman D, Weiner M and Studholme C. MR imaging anatomy in neurodegeneration: A robust volumetric parcellations method of frontal lobe gyri with quantitative validation in patients with dementia. *American Journal of Neuroradiology*, 27:1747-1754, 2006.

James A, James S, Smith D and Javaloyes A. Cerebellar, prefrontal cortex, and thalamic volumes over two time points in adolescent-onset schizophrenia. *American Journal of Psychiatry*, 161: 1023-1029, 2004.

Jernigan TL, Archibald SL, Berhow MT, Sowell ER, Foster DS and Hesselink JR. Cerebral structure on MRI, part1: Localization of age-related changes. *Biological Psychiatry*, 29:55-67, 1991.

Jernigan TL, Archibald SL, Fennema-Notestine C, Gamst AC, Stout JC, Bonner J and Hesselink JR. Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiology of Aging*, 22:581-594, 2001.

John JP, Wang L, Moffitt AJ, Singh HK, Gado MH and Csernansky JG. Inter-rater reliability of manual segmentation of the superior, inferior and middle frontal gyri. *Psychiatry Research*, 148(2-3): 151-163, 2006.

John JP, Yashavantha BS, Gado M, Veena R, Jain S, Ravishankar and Csernansky JG. A proposal for MRI-based parcellations of the frontal pole. *Brain Structure and Function,* 212:245-253, 2007.

John JP, Burgess PW, Yashavantha BS, SHakeel MK, Halahalli HN and Jain S. Differential relationship of frontal pole and whole brain volumetric measures with age in neuroleptic-naïve schizophrenia and healthy subjects. *Schizophrenia Research*, 109:148-158, 2009.

Jones BF, Barnes J, Uylings HBM, Fox NC, Frost, C, Witter MP and Scheltens P. Differential regional atrophy of the cingulate gyrus in Alzheimer disease: A volumetric MRI study. *Cerebral Cortex,* 16(12): 1701-1708, 2006.

Kates WR, Frederikse M, Mostofsky SH, Folley BS, Cooper K, Mazur-Hopkins P, Kofman O, Singer HS, Denckla MB, Pearlson GD and Kaufmann WE. MRI parcellation of the frontal lobe in boys with attention deficit hyperactivity disorder or Tourette syndrome. *Psychiatry Research, 116*(1-2), 63-81, 2002.

Kaur S, Sassi RB, Axelson D, Nicoletti M, Brambilla P, Monkul ES, Hatch JP, Keshevan MS, Ryan N, Birmaher B and Soares JC. Cingulate cortex anatomical abnormalities in children and adolescents with bipolar disorder. *American Journal of Psychiatry*, 162(9): 1637-1643, 2005.

Kegeles LS, Malone KM, Slifstein M, Ellis SP, Xanthopoulos E, Keilp JG, Campbell C, Oquendo M, van Heertum RL and Mann JJ. Response of cortical metabolic deficits to serotonergic challenge in familial mood disorders. *Psychiatry: Interpersonal and Biological Processes*, 160(1): 76-82, 2003.

Kelsoe JR, Cadet JL, Pickar D and Weinberger DR. Quantitative neuroanatomy in schizophrenia. A controlled magnetic resonance imaging study. *Archives of General Psychiatry*, 45(6): 533-41, 1988.

Kennedy DN, Lange N, Makris N, Bates J, Meyer J and Caviness VS. Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cerebral Cortex*, 8(4): 372-384, 1998.

Knaus TA, Bollich AM, Corey DM, Lemen LC and Foundas AL. Variability in perisylvian brain anatomy in healthy adults. *Brain and Language*, 97:219-232, 2006.

Knaus TA, Corey DM, Bollich AM, Lemen LC and Foundas AL. Anatomical asymmetries of anterior perisylvian speech-language regions. *Cortex*, 43:499-510, 2007.

Knaus TA, Silver AM, Dominick KC, Schuring MD, Schaffer N, Lindgren KA, Joseph RM and Tager-Flusberg H. Age-related changes in the anatomy of language regions in autism spectrum disorder. *Brain Imaging Behaviour*, 3(1):51-63, 2009.

Köhler S, Thomas AJ, Lloyd A, Barber R, Almeida OP and O'Brien JT. White matter hyperintensities, cortisol levels, brain atrophy and continuing cognitive deficits in late-life depression. *The British Journal of Psychiatry*, 196(2): 143-9, 2010.

Kopelman MD, Lasserson D, Kingsley D, Bello F, Rush C, Stanhope N, Stevens T, Goodman G, Heilpern G, Kendall B and Colchester A. Structural MRI volumetric analysis in patients with organic amnesia, 2: correlations with anterograde memory and executive tests in 40 patients. *Journal of Neurology, Neurosurgery, and Psychiatry*, 71(1): 23-8, 2001.

Kumar A, Bilker W, Jin Z and Udupa J. Atrophy and high intensity lesions: complementary neurobiological mechanisms in late-life major depression. *Neuropsychopharmacology*, 22(3): 264-74, 2000.

Kumar A, Schweizer E, Zhisong J, Miller D, Bilker W, Swan LL and Gottleib G. Neuroanatomical substrates of late life minor depression: A quantitative magnetic resonance imaging study. *Archives of Neurology*, 54:613-617, 1997.

Kumra S, Giedd JN, Vaituzis AC, Jacobsen LK, McKenna K, Bedwell J, Hamburger S, Nelson JE, Lenane M and Rapoport JL. Childhood-onset psychotic disorders:

magnetic resonance imaging of volumetric differences in brain structure. *The American Journal of Psychiatry*, 157(9): 1467-74, 2000.

Laakso MP, Soininen H, Partanen K, Helkala E-L, Hartikainen P, Vainio P, Hallikainen M, Hanninen T and Riekkinen Sr PJ. Volumes of hippocampus, amygdala and frontal lobes in the MRI-based diagnosis of early Alzheimer's disease: correlation with memory functions. *Journal of Neural Transmission: Parkinson's Disease and Dementia Section,* 9:73-86, 1995.

Lacerda AL, Hardan AY, Yorbik O and Keshavan MS. Measurement of the orbitofrontal cortex: a validation study of a new method. *NeuroImage*, 19(3): 665-673, 2003.

Lai TJ, Payne ME, Byrum CE, Steffens DC and Krishnan KRR. Reduction of orbital frontal cortex volume in geriatric depression. *Biological Psychiatry*, 48(10): 971–975, 2000.

Lindberg O, Ostberg P, Zandbelt BB, Oberg J, Zhang Y, Andersen C, Looi JCL, Bogdanovic and Wahlund L-O. Cortical morphometric subclassification of frontotemporal lobar degeneration. *American Journal of Neuroradiology*, 30(6): 1233-1239, 2009.

Lindberg O, Manzouri A, Westmas E and Wahlund L-O. A comparison between volumetric data generated by voxel-based morphometry and manual parcellation of multimodal regions of the frontal lobe. *American Journal of Neuroradiology,* 33(10): 1957-1963, 2012.

Lyoo IK, Han MH and Cho DY. A brain MRI study in subjects with borderline personality disorder. *Journal of Affective Disorders*, 50(2-3): 235-43, 1998.

MacLullich AMJ, Ferguson KJ, Deary IJ, Seckl JR, Starr JM and Wardlaw JM. Intracranial capacity and brain volumes are associated with cognition in healthy elderly men. *Neurology*, 59(2): 169-74, 2002.

MacLullich AMJ, Ferguson KJ, Wardlaw JM, Starr JM, Deary IJ and Seckl JR. Smaller left anterior cingulate cortex volumes are associated with impaired hypothalamic-pituitary-adrenal axis regulation in healthy elderly men. *The Journal of Clinical Endocrinology and Metabolism*, 91(4): 1591-1594, 2006.

Maher BA, Manschreck TC, Yurgelun-Todd DA and Tsuang MT. Hemispheric asymmetry of frontal and temporal gray matter and age of onset in schizophrenia. *Biological Psychiatry*, 44(6): 413-7, 1998.

Mathalon DH, Sullivan EV, Lim KO and Pfefferbaum A. Progressive brain volume changes and the clinical course of schizophrenia in men: a longitudinal magnetic resonance imaging study. *Archives of General Psychiatry*, 58(2): 148-57, 2001.

Matsui M, Gur RC, Turetsky BI, Yan MXH and Gur RE. The relation between tendency for psychopathology and reduced frontal brain volume in healthy people. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, 13(3): 155-162, 2000.

Matsui M, Suzuki M, Zhou S-Y, Takahashi T, Kawasaki Y, Yuuki H, Kato K and Kurachi M. The relationship between prefrontal brain volume and characteristics of memory strategy in schizophrenia spectrum disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 32(8): 1854-1862, 2008.

McAlonan GM, Daly E, Kumari V, Critchley HD, Amelsvoort TV, Suckling J, Simmons A, Sigmundsson T, Greenwood K, Russell A, Schmitz N, Happe F, Howlin P and Murphy DGM. Brain anatomy and sensorimotor gating in Asperger's syndrome. *Brain*, 127: 1594-1606 2002.

McCormick LM, Ziebell S, Nopoulos P, Cassell M, Andreasen NC and Brumm M. Anterior cingulate cortex: an MRI-based parcellation method. *NeuroImage, 32*(3): 1167-1175, 2006.

McLaughlin NCR, Moore DW, Fulwiler C, Bhadelia R and Gansler DA. Differential Contributions of Lateral Prefrontal Cortex Regions to Visual Memory Processes. *Brain Imaging and Behavior*, 3(2): 202-211, 2009.

Medina KL, McQueeny T, Nagel BJ, Hanson KL, Schweinsburg AD and Tapert SF. Prefrontal cortex volumes in adolescents with alcohol use disorders: unique gender effects. *Alcoholism, Clinical and Experimental Research*, 32(3): 386-394, 2008.

Medina K, McQueeny T, Nagel B, Hanson KL, Yang T and Tapert SF. Prefrontal cortex morphometry in abstinent adolescent marijuana users: subtle gender effects. *Addiction Biology*, 14(4): 457-468, 2009.

Mohlman J, Price RB, Eldreth DA, Chazin D, Glover DM and Kates WR. The relation of worry to prefrontal cortex volume in older adults with and without generalized anxiety disorder. *Psychiatry Research*, 173(2): 121-127, 2009.

Monkul ES, Hatch JP, Nicoletti MA, Spence S, Brambilla P, Lacerda ALT, Sassi RB, Mallinger AG, Keshevan MS and Soares JC. Fronto-limbic brain structures in suicidal and non-suicidal female patients with major depressive disorder. *Molecular Psychiatry*, 12(4): 360-366, 2007.

Mueller EA, Moore MM, Kerr DC, Sexton G, Camicioli RM, Howieson DB, Quinn JF and Kaye JA. Brain volume preserved in healthy elderly through the eleventh decade. *Neurology*, 51(6): 1555-1562, 1998.

Murphy DG, DeCarli CD, Daly E, Gillette JA, McIntosh AR, Haxby JV, Teichberg D, Schapiro MB, Rapoport SI and Horwitz B. Volumetric magnetic resonance imaging in men with dementia of the Alzheimer type: correlations with disease severity. *Biological Psychiatry*, 34(9): 612-21, 1993.

Murphy DGM, DeCarli C, McIntosh AR, Daly E, Mentis MJ, Pietrini P, Szczepanik J, Schapiro MB, Grady CL, Horwitz B and Rapoport SI. Sex differences in human brain morphometry and metabolism: An in vivo quantitative magnetic resonance imaging and positron emission tomography study on the effect of aging. *Archives of General Psychiatry*, 53: 585-594, 1996.

Nagel, B, Medina K, Yoshii J, Schweinsburg AD, Moadab I and Tapert SF. Age-related changes in prefrontal white matter volume across adolescense. *Neuroreport*, 17(13): 1427-1431, 2006.

Najt P, Nicoletti M, Chen HH, Hatch JP, Caetano SC, Sassi RB, Axelson D, Brmabilla P, Keshevan MS, Ryan ND, Birmaher and Soares JC. Anatomical measurements of the orbitofrontal cortex in child and adolescent patients with bipolar disorder. *Neuroscience*, 413(3): 183-186, 2007.

Nakamura M, Nestor PG, Levitt JJ, Cohen AS, Kawashima T, Shenton ME and McCarley RW. Orbitofrontal volume deficit in schizophrenia and thought disorder. *Brain,* 131(1), 180-195, 2008.

Nery FG, Chen H-H, Hatch JP, Nicoletti MA, Brambilla P, Sassi RB, Mallinger AG, Keshavan MS and Soares JC. Orbitofrontal cortex gray matter volumes in bipolar disorder patients: a region-of-interest MRI study. *Bipolar Disorders,* 11:145-153, 2009.

Nifosì F, Toffanin T, Follador H, Zonta F, Padovan G, Pigato G, Carollo C, Ermani M, Amista P and Perini GI. Reduced right posterior hippocampal volume in women with recurrent familial pure depressive disorder. *Psychiatry Research,* 184(1): 23-28, 2010.

Noga JT, Aylward E, Barta PE and Pearlson GD. Cingulate gyrus in schizophrenic patients and normal volunteers. *Psychiatry Research,* 61(4): 201-208, 1995.

Nolan CL, Moore GJ, Madden R, Farchione T, Bartoi M, Lorch E, Stewart CM and Rosenberg DR. Prefrontal cortical volume in childhood-onset major depression: preliminary findings. *Archives of General Psychiatry*, 59(2): 173-9, 2002.

Pantel J, Schroder J, Essig M, Popp D, Dech H, Knopp MV, Schad LR, Eysenbach K, Backenstrass M and Friedlinger M. Quantitative magnetic resonance imaging in geriatric depression and primary degenerative dementia. *Journal of Affective Disorders*, 42(1): 69-83, 1997.

Paus T, Otaky N, Caramanos Z, MacDonald D, Zijdenbos A, D'Avirro D, Gutmans D, Holmes C, Tomiauolo F and Evans AC. In vivo morphometry of the intrasulcal gray matter in the human cingulate, paracingulate, and superior-rostral sulci: hemispheric asymmetries, gender differences and probability maps. *The Journal of Comparative Neurology*, 376(4): 664-673, 1996.

Pfefferbaum A, Sullivan EV, Mathalon DH and Lim KO. Frontal lobe volume loss observed with magnetic resonance imaging in older chronic alcoholics. *Alcoholism, Clinical and Experimental Research,* 21(3): 521-529, 1997

Prasad KMR, Sahni SD, Rohm BR and Keshavan MS. Dorsolateral prefrontal cortex morphology and short-term outcome in first-episode schizophrenia. *Psychiatry Research,* 140(2): 147-155, 2005.

Rademacher J, Galaburda AM, Kennedy DN, Filipek PA and Caviness VS. Human Cerebral Cortex: Localization, Parcellation, and Morphometry with Magnetic Resonance Imaging. *Journal of Cognitive Neuroscience*, 4(4): 352-374, 1992.

Raine A, Reynolds G and Sheard C. Neuroanatomical correlates of skin conductance orienting in normal humans: a magnetic resonance imaging study. *Psychophysiology*, 28(5), 548-558, 1991.

Rankin KP, Rosen HJ, Kramer JH, Chaier GF, Weiner MW, Schuff N and Miller BL. Right and left medial orbitofrontal volumes shown an opposite relationship to agreeableness if FTD. *Dementia and Geriatric Cognitive Disorders*, 17(4): 328-332, 2004.

Ranta ME, Crocetti D, Clauss JA, Kraut MA, Mostofsky SH and Kaufmann WE. Manual MRI parcellation of the frontal lobe. *Psychiatry Research*, 172(2), 147-154, 2009.

Ratnanather JT, Botteron KN, Nishino T, Massie AB, Lal RM, Patel SG, Peddi S, Todd RD and Miller MI. Validating cortical surface analysis of medial prefrontal cortex. *NeuroImage*, 14(5): 1058-1069, 2001.

Rauch SL, Shin LM, Segal E, Pitman RK, Carson MA, McMullin K, Whalen PJ and Makris N. Selectively reduced regional cortical volumes in post-traumatic stress disorder. *Neuroreport*, 14(7): 913-916, 2003.

Raz N Gunning FM, Head D, Dupuis JH, McQuain J, Briggs SD, Loken WJ, Thornton AE and Acker JD. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral Cortex,* 7(3): 268-282, 1997.

Raz N, Rodrigue KM and Acker JD. Hypertension and the brain: vulnerability of the prefrontal regions and executive functions. *Behavioral Neuroscience*, 117(6): 1169-80, 2003.

Raz N, Torres IJ, Briggs SD, Spencer WD, Thornton AE, Loken WJ, Gunning FM, McQuain JD, Driesen NR and Acker JD. Selective neuroanatomic abnormalities in Down's syndrome and their cognitive correlates: Evidence from MRI morphometry. *Neurology*, 45*:* 356-366, 1995.

Raz N, Ghisletta P, Rodrigue KM, Kennedy KM and Lindenberger U. Trajectories of brain aging in middle-aged and older adults: regional and individual differences. *NeuroImage*, 51(2): 501-511, 2010.

Raz N, Gunning-Dixon F, Head D, Rodrigue KM, Williamson A and Acker JD. Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiology of Aging*, 25(3): 377-396, 2004.

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D and Acker JD. Regional brain changes in aging healthy adults: general

trends, individual differences and modifiers. *Cerebral Cortex*, 15(11): 1676-1689, 2005.

Raz N, Rodrigue KM, Kennedy KM and Acker JD. Vascular health and longitudinal changes in brain and cognition in middle-aged and older adults. *Neuropsychology*, 21(2): 149-157, 2007.

Riffkin J, Yücel M, Maruff P, Wood SJ, Soulsby B, Olver J, Kyrios M, Velakoulis D and Pantelis C. A manual and automated MRI study of anterior cingulate and orbito-frontal cortices, and caudate nucleus in obsessive-compulsive disorder: comparison with healthy controls and patients with schizophrenia. *Psychiatry Research*, 138(2): 99-113, 2005.

Rosen HJ, Perry RJ, Murphy J, Kramer JH, Mychack P, Schuff N, Weiner M, Levenson RW and Miller BL. Emotion comprehension in the temporal variant of frontotemporal dementia. *Brain*, 125:2286-2295, 2002.

Rosenberg DR, Keshavan MS, O'Hearn KM, Dick EL, Bagwell WW, Seymour AB, Montrose DM, Pierri JN and Birmaher B. Frontostriatal measurement in treatment-naive children with obsessive-compulsive disorder. *Archives of General Psychiatry*, 54(9): 824-830, 1997.

Rosso IM, Makris N, Thermenos HW, Hodge SM, Brown A, Kennedy D, Caviness VS, Faraone SV, Tsuang MT and Seiman LJ. Regional prefrontal cortex gray matter volumes in youth at familial risk for schizophrenia from the Harvard Adolescent High Risk Study. *Schizophrenia Research*, 123(1): 15-21, 2010.

Rupp CI, Fleischhacker WW, Kemmler G, Oberbauer H, Scholtz AW, Wanko C and Hinterhuber H. Various bilateral olfactory deficits in male patients with schizophrenia. *Schizophrenia bulletin*, *31*(1), 155-165, 2005.

Salat DH, Kaye JA and Janowsky JS. Prefrontal gray and white matter volumes in healthy aging and Alzheimer disease. *Archives of Neurology*, 56(3): 338-44, 1999a.

Salat DH, Kaye JA and Janowsky JS. Selective preservation and degeneration within the prefrontal cortex in aging and Alzheimer disease. *Archives of Neurology*, *58*(9), 1403-1408, 2001.

Salat, DH, Kaye JA and Janowsky JS. Greater orbital prefrontal volume selectively predicts worse working memory performance in older adults. *Cerebral Cortex*, *12*(5), 494-505, 2002.

Salat D, Stangl P, Kaye J and Janowsky J. Sex differences in prefrontal volume with aging and Alzheimer's disease. *Neurobiology of Aging*, 20(6): 591-596, 1999b.

Sanches M, Caetano S, Nicoletti M, Monkul ES, Chen HH, Hatch JP, Yeh P-H, Mullis RL, Keshevan MS, Rajkowska G and Soares JC. An MRI-based approach for the measurement of the dorsolateral prefrontal cortex in humans. *Psychiatry Research*, 173(2): 150-154, 2009.

Sanfilipo M, Lafargue T, Rusinek H, Arena L, Loneragan C, Lautin A, Feiner D, Rotrosen J and Wolkin A. Volumetric measure of the frontal and temporal lobe regions in schizophrenia: relationship to negative symptoms. *Archives of General Psychiatry*, 57(5): 471-480, 2000.

Sanfilipo M, Lafargue T, Rusinek H, Arena L, Loneragan C, Lautin A, Rotrosen J and Wokin A. Cognitive performance in Schizophrenia: relationship to regional brain volumes and psychiatric symptoms. *Psychiatry Research Neuroimaging* 116: 1-23, 2002.

Schenker NM, Desgouttes A-M and Semendeferi K. Neural connectivity and cortical substrates of cognition in hominoids. *Journal of Human Evolution*, 49(5): 547-569, 2005.

Schlaepfer TE, Harris GJ, Tien AY, Peng L, Lee S and Pearlson GD. Structural differences in the cerebral cortex of healthy female and male subjects: a magnetic resonance imaging study. *Psychiatry Research*, 61(3): 129-135, 1995.

Schlaepfer TE, Harris GJ, Tien AY, Peng LW, Lee S, Federman EB, Chase GA, Barta PE and Pearlson GD. Decreased regional cortical gray matter volume in schizophrenia. *American Journal of Psychiatry*, 151: 842-848, 1994.

Schretlen D, Pearlson GD, Anthony JC, Aylward EH, Augustine AM, Davis A and Barta P. Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society*, 6(1): 52-61, 2000.

Seidman LJ, Yurgelun-Todd D, Kremen WS, Woods BT, Goldstein JM, Faraone SV and Tsuang MT. Relationship of prefrontal and temporal lobe MRI measures to neuropsychological performance in chronic schizophrenia. *Biological Psychiatry*, 35(4): 235-246, 1994.

Seidman LJ, Valera EM., Makris N, Monuteaux MC, Boriel DL, Kelkar K, Kennedy DN, Caviness VS, Bush G, Aleardi M, Faraone SV and Biederman J. Dorsolateral prefrontal and anterior cingulate cortex volumetric abnormalities in adults with attention-deficit/hyperactivity disorder identified by magnetic resonance imaging. *Biological Psychiatry*, 60(10): 1071-1080, 2006.

Semendeferi K and Damasio H. The brain and its main anatomical subdivisions in living hominoids using magnetic resonance imaging. *Journal of Human Evolution*, 38(2): 317-32, 2000.

Semendeferi K, Damasio H, Frank R and Van Hoesen GW. The evolution of the frontal lobes: a volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution*, 32(4): 375-388, 1997.

Semendeferi K, Lu A, Schenker N and Damasio H. Humans and great apes share a large frontal cortex. *Nature Neuroscience*, 5(3): 272-6, 2002.

Sherwood CC, Gordon AD, Allen JS, Phillips KA, Erwin JM, Hof PR and Hopkins WD. Aging of the cerebral cortex differs between humans and chimpanzees. *Proceedings of the National Academy of Sciences*, 108(32): 13029-13034, 2011.

Soininen HS, Karhu J, Partanen J, Pääkkönen A, Jousmäki V, Hänninen T, Hallikainen M, Partanen K, Laakso MP, Koivisto K and Reikkinen PJ. Habituation of auditory N100 correlates with amygdaloid volumes and frontal functions in age-associated memory impairment. *Physiology & Behavior*, 57(5): 927-35, 1995.

Sowell ER, Trauner DA, Gamst A and Jernigan TL. Development of cortical and subcortical brain structures in childhood and adolescence: a structural MRI study. *Developmental Medicine and Child Neurology*, 44(1): 4-16, 2002.

Staal W, Hulshoff Pol H, Schnack HG, Hoogendoorn MLC, Jellema K and Kahn RS. Structural brain abnormalities in patients with schizophrenia and their healthy siblings. *American Journal of Psychiatry*, 157: 416-421, 2000.

Sullivan EV, Shear PK, Lim KO, Zipursky RB and Pfefferbaum A. Cognitive and motor impairments are related to gray matter volume deficits in schizophrenia. *Biological Psychiatry*, 39(4): 234-240, 1996.

Suga M, Yamasue H, Abe O, Yamasaki S, Yamada H, Inoue H, Takei K, Aoki S and Kasai K. Reduced gray matter volume of Brodmann's Area 45 is associated with severe psychotic symptoms in patients with schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, 260:465-473, 2010.

Suzuki M, Zhou S-Y, Takahashi T, Hagino H, Kawasaki Y, Niu L, Matsui M,, Seto H and Kurachi M. Differential contributions of prefrontal and temporolimbic pathology to mechanisms of psychosis. *Brain*, 128(9): 2109-2122, 2005.

Szendi I, Kiss M, Racsmany M, Boda K, Cimmer C, Voros E, Kovacs ZA, Szekeres G, Galso G, Pleh C, Csernay L and Janka Z. Correlations between clinical symptoms, working memory functions and structural brain abnormalities in men with schizophrenia. *Psychiatry Research: Neuroimaging*, 147:47-55, 2006.

Szeszko PR, Bilder RM, Lencz T, Ashtari M, Goldman RS, Reiter G, Wu H and Lieberman JA. Reduced anterior cingulate gyrus volume correlates with executive dysfunction in men with first-episode schizophrenia. *Schizophrenia Research*, 43(2-3): 97-108, 2000.

Szeszko PR, Bilder RM, Lencz T, Pollack S, Alvir JM, Ashtari M, Wu H and Lieberman JA. Investigation of frontal lobe subregions in first-episode schizophrenia. *Psychiatry Research*, 90(1): 1-15, 1999a.

Szeszko PR, Robinson D, Alvir JM, Bilder RM, Lencz T, Ashtari M, Wu H and Bogerts B. Orbital frontal and amygdala volume reductions in obsessive-compulsive disorder. *Archives of General Psychiatry*, 56(10): 913-919, 1999b.

Takahashi T, Kawasaki Y, Kurokawa K, Hagino H, Nohara S, Yamashita I, Nakamura K, Murata M, Matsui M, Suzuki M, Seto H and Kurachi M. Lack of

normal structural asymmetry of the anterior cingulate gyrus in female patients with schizophrenia: a volumetric magnetic resonance imaging study. *Schizophrenia Research*, 55(1-2): 69-81, 2002a.

Takahashi T, Suzuki M, Kawasaki Y, Hagino H, Yamashita I, Nohara S, Nakamura K, Seto H and Kurachi M. Perigenual Cingulate Gyrus Volume in Patients with Schizophrenia: A Magnetic Resonance Imaging Study. *Biological Psychiatry*, *53*: 593-600, 2003.

Takahashi, T, Suzuki M, Kawasaki Y, Kurokawa K, Hagino H, Yamashita I, Zhou S-Y, Nohara S, Nakamura K, Seto H and Kurachi M. Volumetric magnetic resonance imaging study of the anterior cingulate gyrus in schizotypal disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 252(6): 268-277, 2002b.

Takeoka M, Kim F, Caviness VS, Kennedy DN, Makris N and Holmes GL. MRI volumetric analysis in rasmussen encephalitis: a longitudinal study. *Epilepsia*, 44(2): 247-251, 2003.

Tisserand DJ, Pruessner JC, Arigita EJS, Boxtel MPJV, Evans AC, Jolles J and Uylings HBM. Regional Frontal Cortical Volumes Decrease Differentially in Aging: An MRI Study to Compare Volumetric Approaches and Voxel-Based Morphometry. *NeuroImage*, 17:657- 669, 2002.

Tomaiuolo F, MacDonald JD, Caramanos Z, Posner G, Chiavaras M, Evans AC and Petrides M. Morphology morphometry and probability mapping of the pars opercularis of the inferior frontal gyrus: an *in vivo* MRI analysis. *European Journal of Neuroscience*, 11:3033-3046, 1999.

Turetsky B, Cowell PE, Gur RC, Grossman RI, Shtasel DL and Gur RE. Frontal and temporal lobe brain volumes in schizophrenia: relationship to symptoms and clinical subtype. *Archives of General Psychiatry*, 52(12): 1061-1070, 1995

Tzourio N, Petit L, Mellet E, Orssaud C, Crivello F, Benali K, Salamon G and Mazoyer B. Use of anatomical parcellation to catalog and study structure-function relationships in the human brain. *Human Brain Mapping*, 5(4): 228-232, 1997.

Uylings HBM, Sanz-Arigita EJ, de Vos K, Pool CW, Evers P and Rajkowska G. 3-D cytoarchitectonic parcellation of human orbitofrontal cortex correlation with postmortem MRI. *Psychiatry Research*, 183(1): 1-20, 2010.

Van Amelsvoort T. Structural brain abnormalities associated with deletion at chromosome 22q11: Quantitative neuroimaging study of adults with velo-cardio-facial syndrome. *The British Journal of Psychiatry*, 178(5), 412-419, 2001.

Van Elst, LTV, Hesslinger B, Thiel T, Geiger E, Haegele K, Lemieux L, Lieb K, Bohus M, Hennig J and Ebert D. Frontolimbic brain abnormalities in patients with boderline personality disorder: A volumetric magnetic resonance imaging study. *Biological Psychiatry*, 54(2): 163-171, 2003.

Van Petten C, Plante E, Davidson PSR, Kuo TY, Bajuscak L and Glisky EL. Memory and executive function in older adults: Relationships with temporal and prefrontal gray matter volumes and white matter hyperintensities. *Neuropsycholgia,* 42:1313-1335, 2004.

Wible CG, Shenton ME, Fischer IA, Allard JE, Kikinis R, Jolesz FA, Iosifescu DV and McCarley RW. Parcellation of the human prefrontal cortex using MRI. *Psychiatry Research: Neuroimaging Section,* 76:29-40, 1997.

Wible CG, Anderson J, Shenton ME, Kricun A, Hirayasu Y, Tanaka S, Levitt JJ, O'Donnell BF, Kikinis R, Jolesz FA and McCarley RW. Prefrontal cortex, negative symptoms, and schizophrenia: an MRI study. *Psychiatry Research*, 108(2): 65-78, 2001.

Wible CG, Shenton ME, Hokama H, Kikinis R, Jolesz FA, Metcalf D and Mccarley RW. Prefrontal cortex and schizophrenia. *Archives of General Psychiatry*, 52: 279-288, 1995.

Wilde EAA, Hunter JV, Newsome MR, Schiebel RS, Bigler ED, Johnson JL, Fearing MA, Cleavinger HB, Li X, Swank PR, Pedroza C, Roberson GS, Bachevalier J and Levin HS. Frontal and temporal morphometric findings on MRI in children after moderate to severe traumatic brain injury. *Journal of Neurotrauma,* 22(3):333-344, 2005.

Wood JL, Heitmiller D, Andreasen NC and Nopoulos P. Morphology of the ventral frontal cortex: Relationship to femininity and social cognition. *Cerebral Cortex,* 18:534-540, 2007.

Wood JL, Murko V and Nopoulos P. Ventral frontal cortex in children: Morphology, social cognition and femininity/masculinity. *Scan,* 3:168-176, 2008.

Woods BT, Yurgelun-Todd D, Goldstein JM, Seidman LJ and Tsuang MT. MRI brain abnormalities in chronic schizophrenia: one process or more? *Biological Psychiatry*, 40(7): 585-596, 1996.

Woodward SH, Kaloupek DG, Streeter CC, Martinez C, Schaer M and Eliez S. Decreased anterior cingulate volume in combat-related PTSD. *Biological Psychiatry*, 59(7): 582-587, 2006.

Yamasaki S, Yamasue H, Abe O, Suga M, Yamada H, Inoue H, Kuwabara H, Kawakubo Y, Yahata N, Aoki S, Kano Y, Kato N and Kasai K. Reduced gray matter volume of pars opercularis is associated with impaired social communication in high-functioning autism spectrum disorders. *Biological Psychiatry* 68(12): 1141-1147, 2010.

Yamasue H, Iwanami A, Hirayasu Y, Yamada H, Abe O, Kuroki N, Fukuda R, Tsujii K, Aoki S, Ohtomo K, Kato N and Kasai K. Localized volume reduction in prefrontal, temporolimbic, and paralimbic regions in schizophrenia: an MRI parcellation study. *Psychiatry Research*, 131(3): 195-207, 2004.

Yücel M, McKinnon MC, Chahal R, Taylor VH, Macdonald K, Joffe R and MacQuenn GM. Anterior cingulate volumes in never-treated patients with major depressive disorder. *Neuropsychopharmacology,* 33:3157-3163, 2008.

Zhou S-Y, Suzuki M, Hagino H, Takahashi T, Kawasaki Y, Matsui M, Seto H ad Kurachi M. Volumetric analysis of sulci/gyri-defined in vivo frontal lobe regions in schizophrenia: Precentral gyrus, cingulate gyrus, and prefrontal region. *Psychiatry Research*, 139(2): 127-39, 2005.

Zipursky RB, Lim KO, Sullivan EV, Brown BW and Pfefferbaum A. Widespread cerebral gray matter volume deficits in schizophrenia. *Archives of General Psychiatry*, 49(3): 195-205, 1992.

Zuffante P, Leonard CM, Kuldau JM, Bauer RM, Doty EG and Bilder RM. Working memory deficits in schizophrenia are not necessarily specific of associated with MRI-based estimates of area 46 volumes. *Psychiatry Research: Neuroimaging Section*, 108:187-209, 2001.

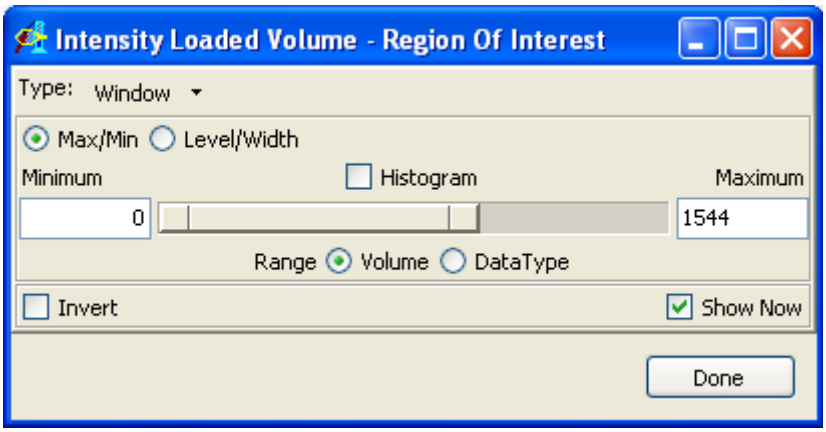
Appendix 3: Hippocampal segmentation protocol

In the Analyze 11.0 window, File_Load, load the thresholded image (see relevant SOP).

With the image icon highlighted, Click on Regions of Interest 

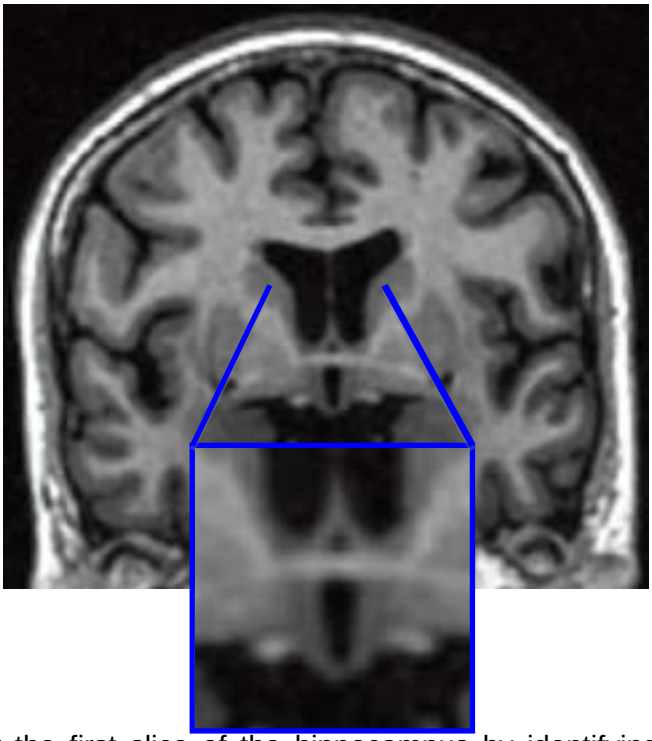
To ensure the image is in the correct orientation select Generate > Orientation > Coronal

If the image is dark select View > Intensities > Loaded Volume

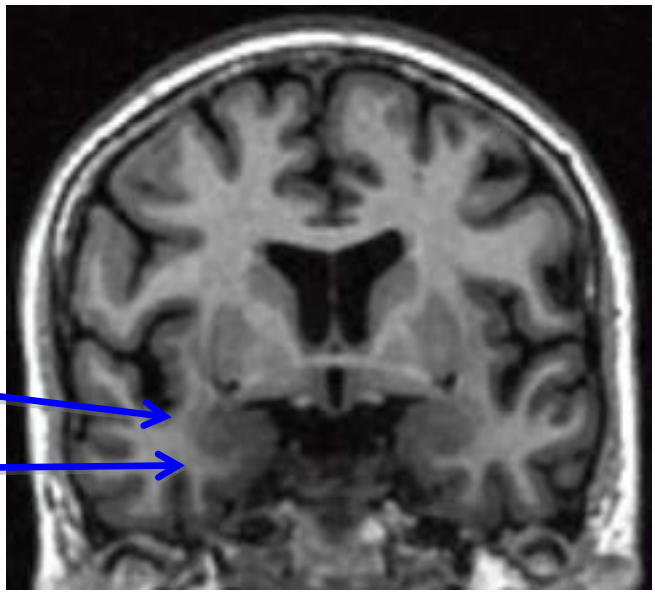


Move the slider down to adjust the Maximum value, brightening the image. Try to select a consistent range within a dataset so that the brightness of the image does not influence identification of tissue.

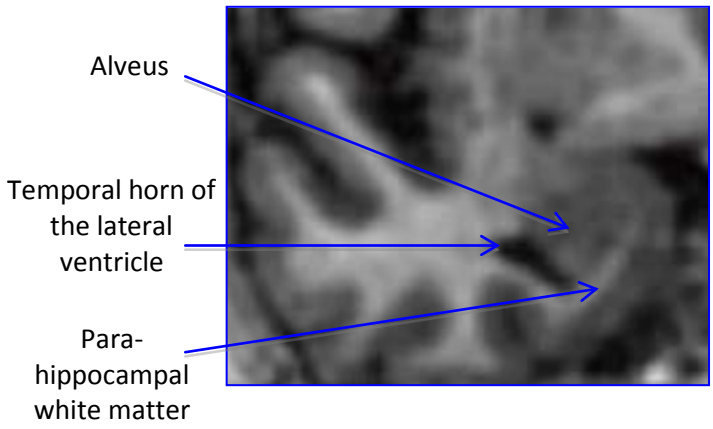
The hippocampus is located in the midbrain so use the Slice slider beneath the image and scroll to the slice in which the third ventricle is fully visible.



Begin looking for the first slice of the hippocampus by identifying an 'L' shape of white matter in the right hand temporal lobe (inverted 'L' on the left), made up of parahippocampal gyral white matter and the temporal stem.



The hippocampus will appear medially and superior to the 'L' shaped white matter, with the inferior boundary being the white matter of the parahippocampal gyrus and the superior boundary being the alveus. Often the temporal horn of the lateral ventricles will appear lateral to the hippocampus.

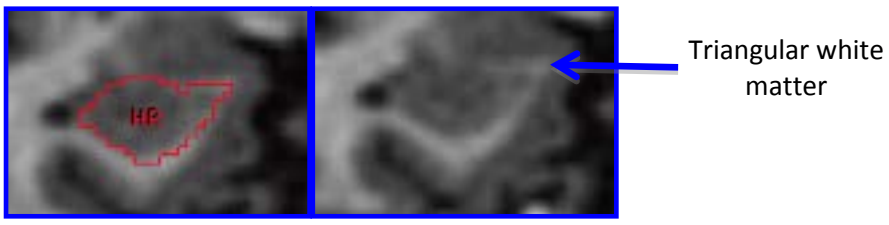


Use the 

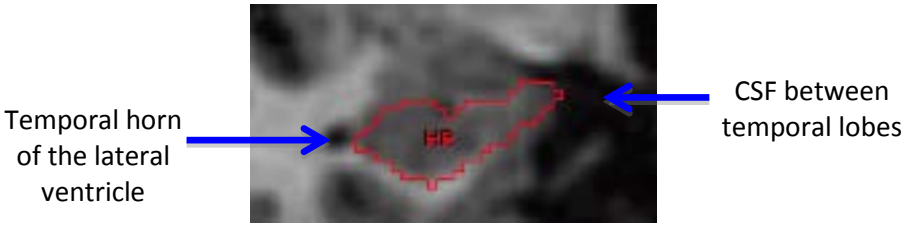
function to trace around the hippocampus, then select the auto-trace  function and click within the traced area. Adjust the intensity range so that the maximum is set to full and the minimum is at the mean threshold previously calculated.

Repeat this process for the contralateral side ensuring that Object to Define is set to **New**. Be aware that the hippocampus will appear asymmetrically due to the differing sizes of the right and left hippocampi and the position of the head during scanning.
When two objects have been established as right and left hippocampus, they can be renamed by going to View > Objects and overwriting the name beside “Name” to HR (Hippocampus Right) for “1.Object”. The return key must be pressed before progressing to name “2.Object” as HL (Hippocampus Left), and after this process or the name will not be changed.

Continue to trace around the hippocampus using the parahippocampal white matter as the inferior boundary, the temporal stem or temporal horn of the lateral ventricle as the lateral boundary and the alveus as the superior boundary. Where visible the triangle of white (see below) matter should be used as the medial boundary.

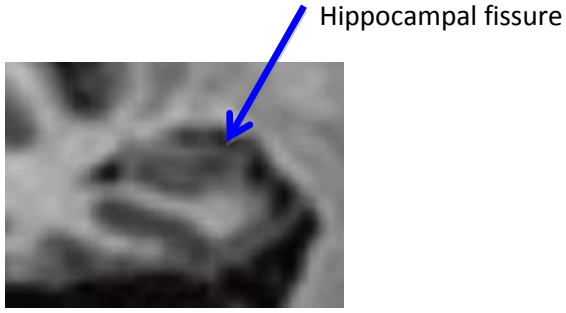


The hippocampus will change shape as you move through the coronal slices, always trace up to the ventricle and, as the uncus is included, to the cerebrospinal fluid (CSF) between the temporal lobes.

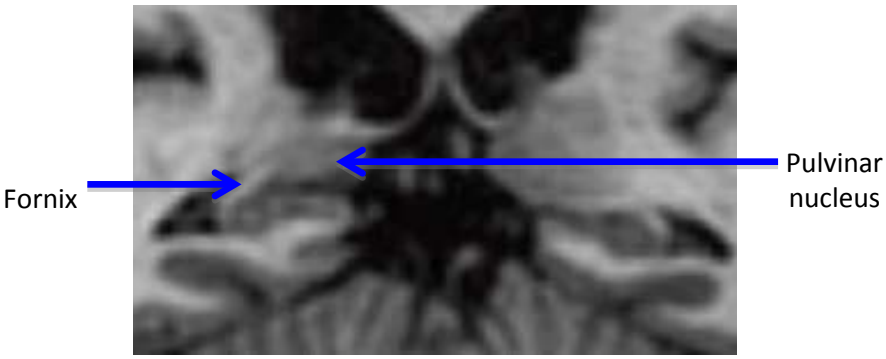


Where the inferior white matter boundary does not extend to the meet the CSF, continue with a straight line from the white matter extending out to the CSF.
Where possible try to follow the digitations or pes of the hippocampi; these are the rounded undulations that can be clear.

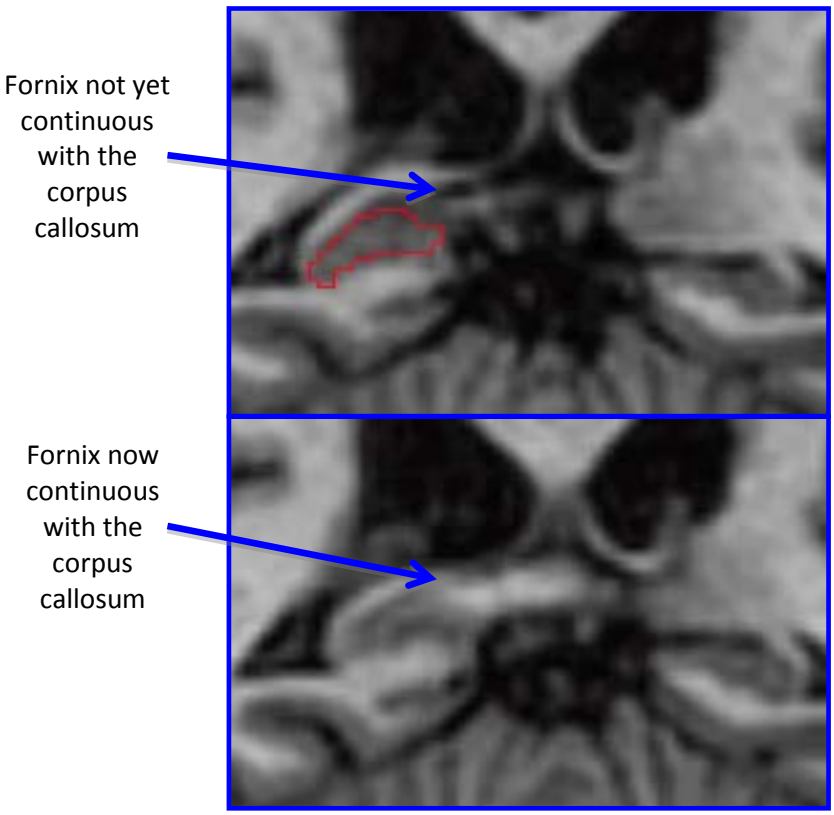
The hippocampal fissure is included with the main body of the hippocampus.



As the tail of the hippocampus is reached the structure becomes smaller and the pulvinar nucleus of the thalamus recedes (superior to hippocampus) with the emergence of the fornix.



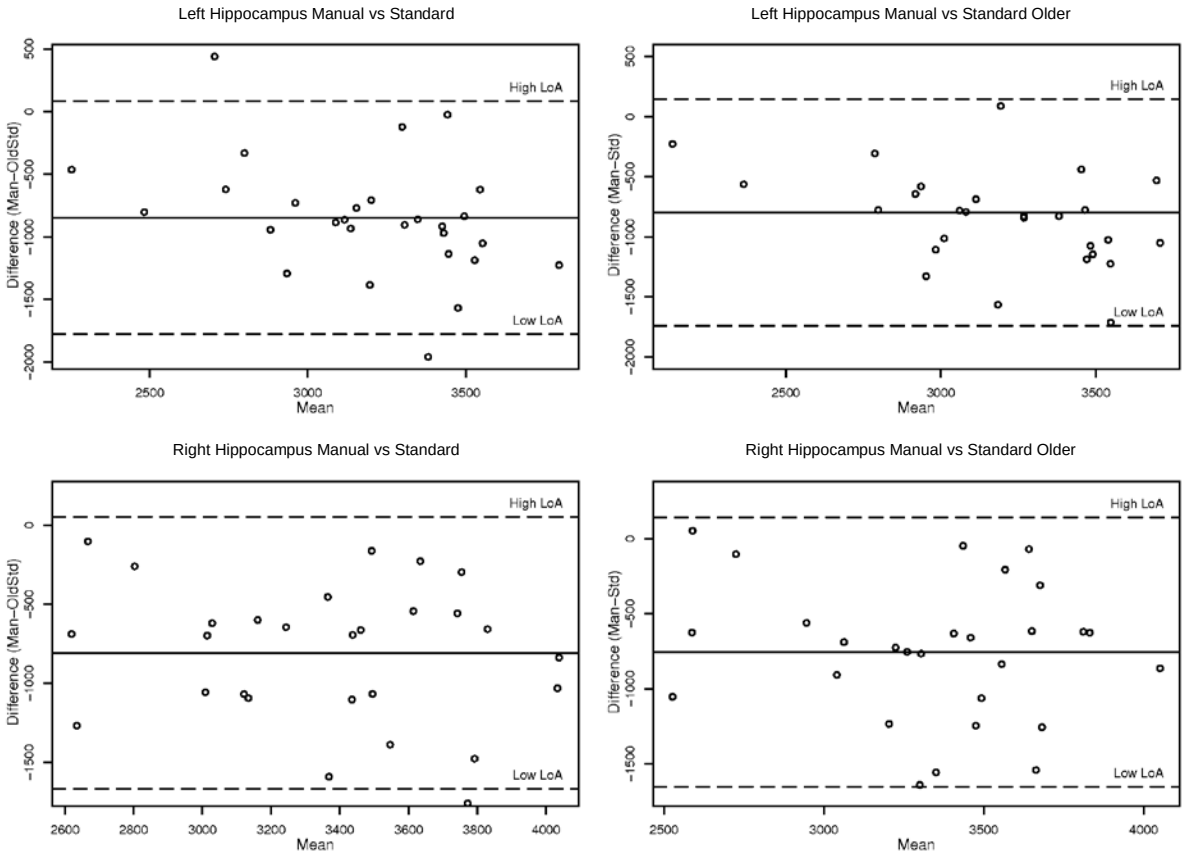
The last slice of the hippocampus is that in which the entire length of the fornix is visible but not yet continuous with the splenium of the corpus callosum.

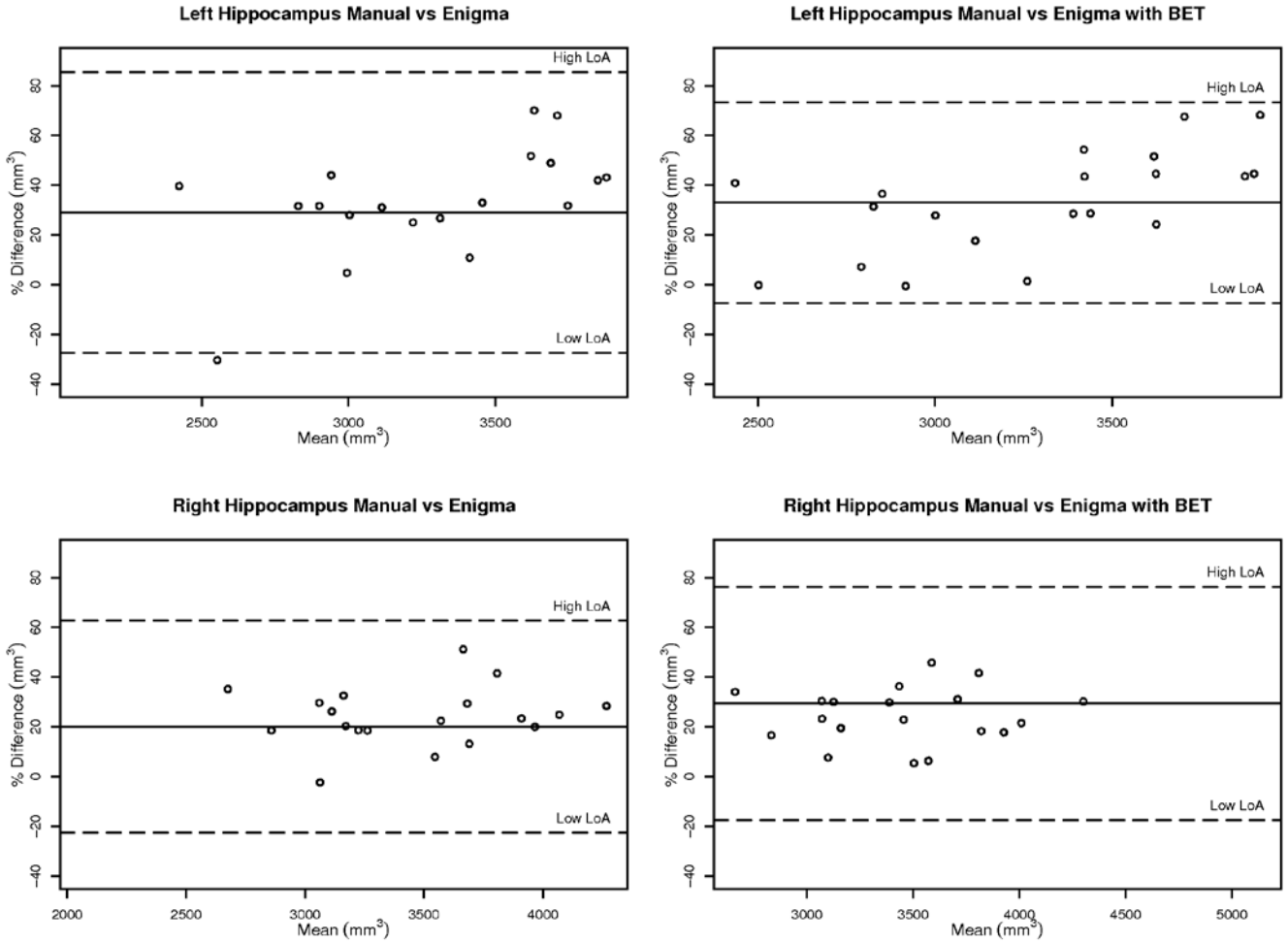


CSF spaces are contained within the hippocampus, to avoid inclusion of these spaces from the calculated volume, set the Object to Define to **New** and use autotrace  on one area of CSF then select Apply. Change the Object to Define to “3.Object” before clicking on the other areas of CSF. By clicking Apply, all the areas of CSF will be assigned to 3.Object; this object can be renamed by going to View > Objects as before.
It may be easier to define the each hippocampus in their entirety and then go back through the slices again to define the CSF spaces. This avoids having to change the intensity range of the autotrace for every slice as well as affording the opportunity to review the segmentation to ensure mistakes have not been made.

Select File > Save object map ensuring the target folder where the object is to be saved is correct and name the file appropriately.

Appendix 4: Bland-Altman plots Chapter 10.2.3




Left Hippocampus Manual vs Enigma
High LoA
Low LoA
% Difference (mm3)
Mean (mm3)
Left Hippocampus Manual vs Enigma with BET
High LoA
Low LoA
% Difference (mm3)
Mean (mm3)
Right Hippocampus Manual vs Enigma
High LoA
Low LoA
% Difference (mm3)
Mean (mm3)
Right Hippocampus Manual vs Enigma with BET
High LoA
Low LoA
% Difference (mm3)
Mean (mm3)

Appendix 5: Linear regression models for all subjects

		Model 1		Model 2		
				T1		
		Cog	Gender	Cog	Gender	Age11IQ
T1 Right	g	**-0.15 (0<0.001)**	**-0.27 (0<0.001)**	**-0.18 (0<0.001)**	**-0.27 (0<0.001)**	0.07 (0.184)
T1 Left	g	**-0.15 (0<0.001)**	**-0.28 (0<0.001)**	**-0.20 (0<0.001)**	**-0.28 (0<0.001)**	0.07 (0.137)
T1 Right	speed	**-0.20 (0<0.001)**	**-0.26 (0<0.001)**	**-0.22 (0<0.001)**	**-0.27 (0<0.001)**	0.04 (0.315)
T1 Left	speed	**-0.18 (0<0.001)**	**-0.28 (0<0.001)**	**-0.20 (0<0.001)**	**-0.28 (0<0.001)**	0.04 (0.391)
T1 Right	memory	**-0.13 (0.002)**	**-0.26 (0<0.001)**	**-0.15 (0.002)**	**-0.26 (0<0.001)**	0.05 (0.312)
T1 Left	memory	**-0.13 (0.001)**	**-0.28 (0<0.001)**	**-0.16 (0.001)**	**-0.28 (0<0.001)**	0.06 (0.247)
				MTR		
		Cog	Gender	Cog	Gender	Age11IQ
MTR Right	g	**0.17 (0<0.001)**	0.04 (0.343)	**0.15 (0.004)**	0.036 (0.393)	0.04 (0.408)
MTR Left	g	**0.18 (0<0.001)**	0.02 (0.601)	**0.17 (0.001)**	0.02 (0.627)	0.02 (0.745)
MTR Right	speed	**0.20 (0<0.001)**	0.03 (0.432)	**0.17 (0<0.001)**	0.027 (0.516)	0.06 (0.169)
MTR Left	speed	**0.19 (0<0.001)**	0.02 (0.738)	**0.17 (0<0.001)**	0.01 (0.82)	0.05 (0.27)
MTR Right	memory	**0.13 (0.002)**	0.03 (0.441)	0.08 (0.105)	0.03 (0.508)	0.09 (0.082)
MTR Left	memory	**0.12 (0.006)**	0.03 (0.52)	0.07 (0.157)	0.02 (0.586)	0.08 (0.1)
				MD		
		Cog	Gender	Cog	Gender	Age11IQ
MD Right	g	**-0.14**	**-0.21 (0<0.001)**	**-0.15 (0.003)**	**-0.21 (0<0.001)**	0.02 (0.738)

		(0<0.001)						
MD Left	g	**-0.16 (0<0.001)**	**-0.14 (0.001)**	**-0.19 (0<0.001)**	**-0.14 (0.001)**	0.06 (0.276)		
MD Right	speed	**-0.20 (0<0.001)**	**-0.20 (0<0.001)**	**-0.20 (0<0.001)**	**-0.20 (0<0.001)**	0.01 (0.947)		
MD Left	speed	**-0.17 (0<0.001)**	**-0.13 (0.002)**	**-0.18 (0<0.001)**	**-0.13 (0.002)**	0.01 (0.814)		
MD Right	memory	**-0.13 (0.001)**	**-0.19 (0<0.001)**	**-0.14 (0.005)**	**-0.19 (0<0.001)**	0.01 (0.832)		
MD Left	memory	**-0.15 (0.001)**	**-0.17 (0<0.001)**	**-0.16 (0.001)**	**-0.17 (0<0.001)**	0.03 (0.544)		
				FA				
		Cog	Gender	Cog	Gender	Age11IQ		
FA Right	g	**0.21 (0<0.001)**	0.08 (0.069)	**0.17 (0.001)**	0.07 (0.094)	0.07 (0.178)		
FA Left	g	**0.18 (0<0.001)**	0.05 (0.219)	**0.15 (0.003)**	0.05 (0.26)	0.05 (0.355)		
FA Right	speed	**0.21 (0<0.001)**	0.07 (0.1)	**0.17 (0<0.001)**	0.06 (0.145)	0.10 (0.032)		
FA Left	speed	**0.17 (0<0.001)**	0.05 (0.26)	**0.14 (0.003)**	0.04 (0.331)	0.08 (0.079)		
FA Right	memory	**0.15 (0.001)**	0.06 (0.153)	0.08 (0.128)	0.06 (0.193)	0.13 (0.014)		
FA Left	memory	**0.13 (0.002)**	0.05 (0.233)	0.09 (0.091)	0.05 (0.27)	0.09 (0.087)		
			Hippocampus Volume					
		Model 3			Model 4			
		Cog	Gender	ICV	cog	Gender	Age11IQ	ICV
Hippo volume Right	g	**0.09 (0.011)**	-0.08 (0.082)	**0.39 (<0.001)**	**0.10 (0.023)**	-0.08 (0.097)	-0.02 (0.72)	**0.39 (<0.001)**
Hippo volume Left	g	0.07 (0.06)	-0.01 (0.797)	**0.38 (<0.001)**	0.07 (0.151)	-0.01 (0.777)	0.01 (0.853)	**0.38 (<0.001)**
Hippo volume Right	speed	0.06 (0.112)	-0.09 (0.073)	**0.39 (<0.001)**	0.05 (0.24)	-0.09 (0.058)	0.03 (0.427)	**0.39 (<0.001)**
Hippo volume Left	speed	0.03 (0.489)	-0.01 (0.852)	**0.39 (<0.001)**	0.01 (0.828)	-0.02 (0.717)	0.05 (0.257)	**0.38 (<0.001)**

Hippo volume Right	memory	0.07 (0.081)	-0.08 (0.089)	**0.40 (<0.001)**	0.06 (0.164)	-0.08 (0.088)	0.01 (0.878)	**0.40 (<0.001)**
Hippo volume Left	memory	**0.12 (0.002)**	-0.03 (0.585)	**0.39 (<0.001)**	0.13 (0.003)	-0.02 (0.651)	-0.03 (0.552)	**0.39 (<0.001)**

* represents associations that became non-significant at p<0.05 after correction for multiple testing
Model 1: integrity = β1*cog + β2*Gender
Model 2: integrity = β1*cog + β2*Gender + β3*Age 11 IQ
Model 3: hippo volume = β1*cog + β2*Gender + β3*ICV
Model 4: hippo volume = β1*cog + β2*Gender + β3*ICV + β4*Age 11 IQ

Linear regression models predicting measures of integrity from cognition and age 11 IQ. Covariates were Gender and ICV. Models used the entire population. Values are standardized beta (and p). N=565

Appendix 6: Linear regression models subjects MMSE above 27

		model 1		model 2		
		T1				
		cog	Gender	cog	Gender	age11IQ
T1 Right	g	**-0.13 (0.003)**	**-0.29 (0<0.001)**	**-0.16 (0.002)**	**-0.29 (0<0.001)**	0.06 (0.21)
T1 Left	g	**-0.13 (0.003)**	**-0.31 (0<0.001)**	**-0.18 (0<0.001)**	**-0.31 (0<0.001)**	0.10 (0.05)
T1 Right	speed	**-0.17 (0<0.001)**	**-0.29 (0<0.001)**	**-0.19 (0<0.001)**	**-0.29 (0<0.001)**	0.05 (0.321)
T1 Left	speed	**-0.16 (0<0.001)**	**-0.31 (0<0.001)**	**-0.19 (0<0.001)**	**-0.31 (0<0.001)**	0.07 (0.129)
T1 Right	memory	**-0.11 (0.016)**	**-0.27 (0<0.001)**	**-0.13 (0.009)**	**-0.27 (0<0.001)**	0.05 (0.282)
T1 Left	memory	**-0.10 (0.028)**	**-0.29 (0<0.001)**	**-0.13 (0.007)**	**-0.29 (0<0.001)**	0.08 (0.111)
		MTR				
		Cog	Gender	Cog	Gender	age11IQ
MTR Right	g	**0.19 (0<0.001)**	0.04 (0.377)	**0.16 (0.002)**	0.04 (0.405)	0.04 (0.426)
MTR Left	g	**0.19 (0<0.001)**	0.02 (0.721)	**0.18 (0.001)**	0.02 (0.742)	0.03 (0.631)
MTR Right	speed	**0.22 (0<0.001)**	0.03 (0.469)	**0.20 (0<0.001)**	0.03 (0.509)	0.06 (0.195)
MTR Left	speed	**0.22 (0<0.001)**	0.01 (0.893)	**0.20 (0<0.001)**	0.01 (0.931)	0.05 (0.294)
MTR Right	memory	**0.14 (0.003)**	0.03 (0.479)	0.09 (0.073)	0.03 (0.501)	0.09 (0.089)
MTR Left	memory	**0.12 (0.009)**	0.01 (0.809)	0.08 (0.147)	0.01 (0.835)	0.09 (0.079)
		MD				
		Cog	Gender	Cog	Gender	age11IQ
MD Right	g	**-0.15 (0.001)**	**-0.23 (0<0.001)**	**-0.16 (0.002)**	**-0.23 (0<0.001)**	0.02 (0.679)

MD Left	g	**-0.19 (0<0.001)**	**-0.20 (0<0.001)**	**-0.22 (0<0.001)**	**-0.21 (0<0.001)**	0.06 (0.26)		
MD Right	speed	**-0.21 (0<0.001)**	**-0.22 (0<0.001)**	**-0.21 (0<0.001)**	**-0.22 (0<0.001)**	0.01 (0.851)		
MD Left	speed	**-0.23 (0<0.001)**	**-0.19 (0<0.001)**	**-0.24 (0<0.001)**	**-0.19 (0<0.001)**	0.02 (0.62)		
MD Right	memory	**-0.14 (0.001)**	**-0.21 (0<0.001)**	**-0.15 (0.003)**	**-0.21 (0<0.001)**	0.02 (0.699)		
MD Left	memory	**-0.17 (0<0.001)**	**-0.19 (0<0.001)**	**-0.19 (0<0.001)**	**-0.19 (0<0.001)**	0.04 (0.479)		
		FA						
		Cog	Gender	Cog	Gender	age11IQ		
FA Right	g	**0.22 (0<0.001)**	**0.107 (0.016)**	**0.18 (0.001)**	**0.10 (0.02)**	0.09 (0.105)		
FA Left	g	**0.22 (0<0.001)**	0.071 (0.112)	**0.19 (0<0.001)**	0.07 (0.127)	0.06 (0.292)		
FA Right	speed	**0.23 (0<0.001)**	0.096 (0.031)	**0.19 (0<0.001)**	**0.09 (0.037)**	0.11 (0.025)		
FA Left	speed	**0.21 (0<0.001)**	0.062 (0.168)	**0.18 (0<0.001)**	0.06 (0.189)	0.09 (0.06)		
FA Right	memory	**0.16 (0<0.001)**	0.087 (0.053)	0.10 (0.058)	0.09 (0.055)	**0.13 (0.011)**		
FA Left	memory	**0.17 (0<0.001)**	0.06 (0.18)	**0.13 (0.014)**	0.06 (0.185)	0.09 (0.072)		
		Hippo volume						
		Model 1			Model 2			
		Cog	Gender	ICV	cog	Gender	age11IQ	ICV
Hippo volume Right	g	**0.08 (0.035)***	-0.08 (0.12)	**0.40 (<0.001)**	0.09 (0.057)	-0.08 (0.13)	-0.01 (0.824)	**0.40 (<0.001)**
Hippo volume Left	g	0.07 (0.095)	-0.01 (0.954)	**0.41 (<0.001)**	0.07 (0.149)	-0.00 (0.959)	-0.00 (0.961)	**0.41 (<0.001)**
Hippo volume Right	speed	0.03 (0.398)	-0.08 (0.105)	**0.41 (<0.001)**	0.02 (0.6)	-0.09 (0.088)	0.04 (0.407)	**0.40 (<0.001)**
Hippo volume Left	speed	0.01 (0.78)	-0.01 (0.992)	**0.43 (<0.001)**	0.00 (0.992)	-0.01 (0.911)	0.04 (0.406)	**0.42 (<0.001)**
Hippo volume	memory	0.05 (0.189)	-0.08 (0.112)	**0.41 (<0.001)**	0.05 (0.296)	-0.08	0.01 (0.819)	**0.41 (<0.001)**

Right						(0.109)		
Hippo volume Left	memory	**0.11 (0.009)***	-0.01 (0.863)	**0.42 (<0.001)**	0.12 (0.009)	-0.01 (0.925)	-0.03 (0.514)	**0.42 (<0.001)**

Note. See Appendix 5
* represents associations that became non-significant at p<0.05 after correction for multiple testing

Linear regression models predicting measures of integrity from cognition and age 11 IQ. Covariates were Gender and ICV. Models used subjects with MMSE 27 and above. Values are standardized beta (and p). N=483

Appendix 7: Linear regression models for imaging parameters

	g		speed		memory	
	Right	Left	Right	Left	Right	left
			Hippocampus volume			
Volume	0.08 (0.057)*	0.06 (0.149)	0.03 (0.6)	0.00 (0.992)	0.05 (0.296)	**0.11 (0.009)***
Gender	0.01 (0.838)	0.00 (0.935)	0.06 (0.253)	0.06 (0.268)	0.03 (0.603)	0.023 (0.648)
ICV	0.08 (0.146)	0.08 (0.109)	0.12 (0.032)	0.13 (0.021)	-0.00 (0.947)	-0.033 (0.549)
Age 11 IQ	**0.53 (<0.001)**	**0.53 (<0.001)**	**0.33 (<0.001)**	**0.33 (<0.001)**	**0.48 (<0.001)**	**0.48 (<0.001)**
			T1			
T1	**-0.13 (0.002)**	**-0.14 (<0.001)**	**-0.18 (<0.001)**	**-0.18 (<0.001)**	**-0.11 (0.009)**	**-0.11 (0.007)**
Gender	**-0.09 (0.022)**	**-0.10 (0.013)**	-0.09 (0.052)	**-0.09 (0.045)**	**-0.02 (0.574)**	-0.03 (0.529)
Age 11 IQ	**0.54 (<0.001)**	**0.54 (<0.001)**	**0.35 (<0.001)**	**0.36 (<0.001)**	**0.49 (<0.001)**	**0.49 (<0.001)**
			MTR			
MTR	**0.12 (0.002)**	**0.13 (0.001)**	**0.17 (<0.001)**	**0.18 (<0.001)**	0.07 (0.073)	0.06 (0.147)
Gender	-0.06 (0.118)	-0.06 (0.138)	-0.040 (0.358)	-0.03 (0.429)	0.00 (0.94)	0.01 (0.881)
Age 11 IQ	**0.53 (<0.001)**	**0.53 (<0.001)**	**0.33 (<0.001)**	**0.33 (<0.001)**	**0.48 (<0.001)**	**0.48 (<0.001)**
			MD			
MD	**-0.12 (0.002)**	**-0.16 (<0.001)**	**-0.20 (<0.001)**	**-0.22 (<0.001)**	**-0.12 (0.003)**	**-0.15 (<0.001)**
Gender	**-0.09 (0.025)**	**-0.09 (0.017)**	-0.06 (0.139)	-0.06 (0.148)	-0.02 (0.68)	-0.02 (0.627)
Age 11 IQ	**0.54 (<0.001)**	**0.54 (<0.001)**	**0.34 (<0.001)**	**0.34 (<0.001)**	**0.48 (<0.001)**	**0.48 (<0.001)**
			FA			
FA	**0.13 (0.001)**	**0.14 (<0.001)**	**0.18 (<0.001)**	**0.16 (<0.001)**	**0.08 (0.058)**	**0.10 (0.014)**
Gender	-0.07 (0.056)	-0.07 (0.07)	-0.04 (0.388)	-0.03 (0.482)	0.00 (0.945)	0.00 (0.932)
Age 11 IQ	**0.52 (<0.001)**	**0.52 (<0.001)**	**0.32 (<0.001)**	**0.33 (<0.001)**	**0.48 (<0.001)**	**0.48 (<0.001)**

Note: Values are the standardized β (and p value) for the listed measures of hippocampus integrity predicting measures of cognitive ability.
Model: cognition = β1*integrity + β2*Gender + β3*Age 11 IQ
Where integrity represents measures of hippocampus integrity (T1, MTR, FA, MD and hippocampus volume). ICV is included only for hippocampus volume to correct for head size.

* represents associations that became non-significant at p<0.05 after correction for multiple testing

Linear regression models for the association between cognitive ability and longitudinal relaxation time (T1), magnetization transfer ratio (MTR) and hippocampal volume. Models used subjects with MMSE above 27, N=483. Values are standardized beta (and p).

Appendix 8: Bivariate correlation coefficients

		Volume		T1		MTR		MD		FA	Variance Inflation Factor		Tolerance	
		Right	left	right	left	right	left	right	left	right	range	**mean (std)**	range	**mean (std)**
Volume	left	0.77 (<0.001)												
T1	right	-0.05 (0.198)	-0.04 (0.299)								1.19 to 1.20	1.20(0.003)	0.83 to 0.84	0.83(0.002)
	left	-0.03 (0.515)	-0.06 (0.135)	0.92 (<0.001)							1.24 to 1.27	1.26(0.017)	0.79 to 0.81	0.80(0.011)
MTR	right	**0.17 (<0.001)**	**0.16 (<0.001)**	0.04 (0.331)	0.05 (0.213)						1.76 to 1.77	1.77(0.015)	0.56 to 0.57	0.56(0.005)
	left	**0.16 (<0.001)**	**0.15 (<0.001)**	0.06 (0.121)	0.00 (0.916)	0.87 (<0.001)					1.69 to 1.83	1.78(0.074)	0.55 to 0.59	0.56(0.024)
MD	right	-0.01 (0.785)	-0.04 (0.343)	**0.21 (<0.001)**	**0.18 (<0.001)**	**-0.48 (<0.001)**	**-0.35 (<0.001)**				1.75 to 1.78	1.76(0.019)	0.56 to 0.57	0.57(0.006)
	left	-0.03 (0.425)	0 (0.906)	**0.17 (<0.001)**	**0.26 (<0.001)**	**-0.4 (<0.001)**	**-0.48 (<0.001)**	**0.68 (<0.001)**			1.52 to 1.66	1.61(0.076)	0.60 to 0.66	0.62(0.030)
FA	right	**0.18 (<0.001)**	**0.17 (<0.001)**	**-0.29 (<0.001)**	**-0.22 (<0.001)**	**0.29 (<0.001)**	**0.23 (<0.001)**	**-0.51 (<0.001)**	**-0.32 (<0.001)**		1.48 to 1.49	1.49(0.004)	0.67 to 0.67	0.67(0.002)
	left	**0.16 (<0.001)**	**0.23 (<0.001)**	**-0.3 (<0.001)**	**-0.32 (<0.001)**	**0.21 (<0.001)**	**0.25 (<0.001)**	**-0.37 (<0.001)**	**-0.41 (<0.001)**	**0.65 (<0.001)**	1.34 to 1.36	1.35(0.009)	0.74 to 0.75	0.74(0.005)

Variance Inflation Factor and tolerance were computed as measures of collinearity from models that included the combined hippocampus integrity measures (T1, MTR, FA, MD).
All significant associations remained after correction for multiple comparisons using Bonferroni correction

Bivariate correlation coefficient (and p) showing associations between measures of hippocampus integrity using Pearson correlation. Models used the entire population. N=565

References

Abbott, A. (2012). Cognition: the brain's decline. *Nature*, vol. 492, S4-S5.

Adriano, F., Caltagirone, C., & Spalletta, G. (2012). Hippocampal Volume Reduction in First-Episode and Chronic Schizophrenia: A Review and Meta-Analysis. *Neuroscientist*, vol. 18, no. 2, pp. 180-200.

Albinet, C.T., Boucard, G., Bouquet, C.A., Audiffren, M. (2012). Processing speed and executive functions in cognitive aging: how to disentangle their mutual relationship? *Brain and Cognition*, vol. 79, no. 1, pp. 1-11.

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D. (2009). Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage*, vol. 46, no. 3, pp. 726-738.

Allen, J.S., Bruss, J., Brown, K.C., Damasio, H. (2005). Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. *Neurobiology of Aging*, vol. 26, no. 9, pp.1245-1260.

Allen, J.S., Damasio, H., Grabowski, T.J., Bruss, J., Zhang, W. (2003). Sexual dimorphism and asymmetries in the gray-white composition of the human cerebrum. *NeuroImage*, vol. 18, no. 4, pp. 880-894.

Alzheimer's Society [Online], Available: http://www.alzheimers.org.uk/statistics. Date accessed 02/04/2014.

Aribisala, B.S., Cox, S.R., Ferguson, K.J., MacPherson, S.E., MacLullich, A.M.J., Royle, N.A., Valdés Hernández, M.C., Bastin, M.E., Deary, I.J., Wardlaw, J.M. (2013). Assessing the performance of atlas-based prefrontal brain parcellation in an ageing cohort. *Journal of Computer Assisted Tomography*, vol. 37, no. 2, pp. 257-264.

Armitage, P. A., Schwindack, C., Bastin, M. E., & Whittle, I. R. (2007). Quantitative assessment of intracranial tumor response to dexamethasone using diffusion, perfusion and permeability magnetic resonance imaging. *Magnetic Resonance Imaging*, vol. 25, no. 3, pp. 303-310.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, vol. 54, no. 3, pp. 2033-2044.

Babalola, K.O., Parenaude, B., Aljabar, P., Schnabel, J., Kennedy, D., Crum, W., Smith, S., Cootes, T., Jenkinson, M., Rueckert, D. (2009). An evaluation of four automatic methods of segmenting the subcortical structures in the brain. *NeuroImage*, vol. 47, no. 4, pp. 1435-1447.

Bäckman, L., Robins-Wahlin, T.B., Lundin, A., Ginovart, N., Farde, L. (1997). Cognitive deficits in Huntington's disease are predicted by dopaminergic PET

markers and brain volumes. *Brain*, vol. 120, no. 12, pp. 2207-2217.

Barnes J., Bartlett J.W., van de Pol L.A., Loy C.T., Scahill R.I., Frost C., Thompson, P., ox, N.C. (2009). A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. *Neurobiology of Aging*, vol. 30, no. 11, pp. 1711–23.

Barnes, J., Scahill, R.I., Boyes, R.G., Frost, C., Lewis, E.B., Rossor, C.L., Rossor, M.N., Fox, N.C. (2004). Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates. *NeuroImage* vol. 23, no. 2, pp. 574-81.

Baron, J.C., Chételat, G., Desgranges, B., Perchey, G., Landeau, B., de la Sayette, V., Eustache, F. (2001). In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. *NeuroImage*, vol. 14, no. 2, pp. 298-309.

Bartzokis, G., Mintz, J., Marx, P., Osborn, D., Gutkind, D., Chiang, F., Phelan, C.K. Marder, S.R. (1993). Reliability of in vivo volume measures of hippocampus and other brain structures using MRI. *Magnetic Resonance Imaging*, vol. 11, no. 7, pp. 993-1006.

Bastin, M. E., Sinha, S., Whittle, A. R., & Wardlaw, J. M. (2002). Measurements of water diffusion and T1 values in peritumoural oedematous brain. *Neuroreport*, vol. 13, no. 10, pp. 1335-1340.

Batouli, S.A.H., Trollor, J.N., Wen, W., Sachdev, P.S. (2014). The heritability of volumes of brain structures and its relationship to age: a review of twin and family studies. *Ageing Research Reviews*, vol. 13, pp. 1-9.

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, M., Nunes, R. G., Clare, S., et al. (2003). Characterization and propegation of uncertainty in diffusion weighted MR images. *Magnetic Resonance Medicine*, vol. 50, no. 5, pp. 1077-1088.

Beltz, A.M., Blakemore, J.E.O., Berenbaum, S.A. (2013). Sex differences in brain and behavioral development. *Neural circuit development and function in the brain: Comprehensive developmental neuroscience*. Oxford: Elsevier.

Bennett, I.J., Madden, D.J. (2013). Disconnected aging: cerebral white matter integrity and age-related differences in cognition. *Neuroscience*, doi: 10.1016/j.neuroscience.2013.11.026. [Epub ahead of print].

Bennett, I.J., Rypma, B. (2013). Advances in functional neuroanatomy: a review of combined DTI and fMRI studies in healthy younger and older adults. *Neuroscience & Biobehavioral Reviews*, vol. 37, no. 7, pp. 1201-1210.

Bergfield, K.L., Hanson, K.D., Chen, K., Teipel, S.J., Hampel, H., Rapoport, S.I., Moeller, J.R., Alexander, G.E. (2010). Age-related networks of regional covariance in MRI gray matter: reproducible multivariate patterns in healthy aging. *NeuroImage*, vol. 49, no. 2, pp. 1750-1759.

Bergouignan, L., Chupin, M., Czechowska, Y., Kinkingnehun, S., Lemogne, C., Le Bastard, G., Lepage, M., Garnero, L., Colliot, O., Fossati, P. (2009). Can voxel based

morphometry, manual segmentation and automated segmentation equally detect hippocampal volume differences in acute depression? *NeuroImage*, vol. 45, no. 1, pp. 29-37.

Bhagat, Y. A., & Beaulieu, C. (2004). Diffusion anisotropy in subcortical white matter and cortical gray matter: Changes with aging and the role of CSF-suppression. *Journal of Magnetic Resonance Imaging*, vol. 20, no. 2, pp. 216-227.

Bishop, C.A., Jenkinson, M., Andersson, J., Declerck, J., Mehof, D. (2011). Novel fast marching for automated segmentation of the hippocampus (FMASH): Method and validation on clinical data. *NeuroImage*, vol. 55, no. 3, pp. 1009-1019.

Bjork, J.M., Momenan, R., Hommer, D.W. (2009). Delay discounting correlates with proportional lateral frontal cortex volumes. *Biological Psychiatry*, vol. 65, no. 8, pp. 710-713.

Blatter, D.D., Bigler, E.D., Gale, S.D., Johnson, S.C., Anderson, C., Burnett, B.M., Parker, N., Kurth, S., Horn, S. (1995). Quantitative volumetric analysis of brain MRI: normative database spanning five decades of life. *American Journal of Neuroradiology*, vol. 16, no. 2, pp. 241–245.

Boccardi, M., Ganzola, R., Bocchetta, M., Pievani, M., Redolfi, A., Bartzokis, G., Camicioli, R., Csernansky, J.G., de Leon, M.J., de Toledo-Morrell, L., Killiany, R.J., Lehéricy, S., Pantel, J., Pruessner, J.C., Soininen, H., Watson, C., Duchesne, S., Jack jr, C.R., Frisoni, G.B. (2011). Survey of protocols for the manual segmentation of the hippocampus: preparatory steps towards a joint EADC-ADNI harmonized protocol. *Journal of Alzheimer's disease*, vol. 26, no. 3, pp. 61-75.

van den Bogaard, S. J. A., Dumas, E. M., Milles, J., Reilmann, R., Stout, J. C., Craufurd, D., et al. (2012). Magnetization Transfer Imaging in Premanifest and Manifest Huntington Disease. *American Journal of Neuroradiology*, vol. 33, no. 5, pp. 884-889.

Boyer, P., Phillips, J. L., Rousseau, F. L., & Ilivitsky, S. (2007). Hippocampal abnormalities and memory deficits: New evidence of a strong pathophysiological link in schizophrenia. Brain Research Reviews, vol. 54, no. 1, pp. 92-112.

Bozzali, M., Cercignani, M., Caltagirone, C. (2008). Brain volumetrics to investigate aging and the principal forms of degenerative decline: a brief review. *Magnetic Resonance Imaging*, vol. 26, no. 7, pp. 1065-1070.

Brandt, M., Deindl, C., Hank, K. (2012). Tracing the origins of successful aging: the role of childhood conditions and social inequality in explaining later life health. *Social Science & Medicine*, vol. 74, no. 9, pp. 1418-1425.

Bremner, J.D., Narayan, M., Anderson, E.R., Staib, L.H., Miller, H.L., Charney, D.S. (2000). Hippocampal volume reduction in major depression. *American Journal of Psychiatry*, vol. 157, no. 1, pp. 115-117.

Brewer, J.B., Magda, S., Airriess, C., Smith, M.E. (2009). Fully-automated quantification of regional brain volumes for improved detection of focal atrophy in Alzheimer disease. *American Journal of Neuroradiology*, vol. 30, no. 3, pp. 578-580.

Buckner, R.L., Head, D., Parker, J., Fotenos, A.F., Marcus, D., Morris, J.C., Snyder, A.Z. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *NeuroImage*, vol. 23, no. 2, pp. 724-738.

Burgess, N., Maguire, E.A., O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, vol. 35, no. 4, pp. 625-641.

Burgmans, S., van Boxtel, M.P.J., Gronenschild, E.H.B.M., Vuurman, E.F.P.M., Hofman, P., Uylings, H.B.M., Jolles, J., Raz, N. (2010). Multiple indicators of age-related differences in cerebral white matter and the modifying effects of hypertension. *NeuroImage*, vol. 49, no. 3, pp. 2083-2093.

Burns, J.M., Cronk, B.B., Anderson, H.S., Donelli, J.E., Thomas, G.P., Harsha, A., Brooks, W.M., Swerdlow, R.H. (2008). Cardiorespiratory fitness and brain atrophy in early Alzheimer's disease. *Neurology*, vol. 71, no. 3, pp. 210-216.

Cabezas, M., Oliver, A., Llado, X., Freixenet, J., Cuadra, M.B. (2011). A review of atlas-based segmentation for magnetic resonance brain images. *Computer methods and programs in biomedicine*, vol. 104, no 3, e158-e177.

Cancer Research UK [Online, 20/03/2014], Available: http://www.cancerresearchuk.org/cancer-info/cancerstats/. Date accessed 02/04/2014.

Cardenas, V.A., Chao, L.L., Studholme, C., Yaffe, K., Miller, B.L., Madison, C., buckley, S.T., Mungas, D., Schuff, N., Weiner, M.W. (2011). Brain atrophy associated with baseline and longitudinal measures of cognition. *Neurobiology of Aging*, vol. 32, no. 4, pp. 572-580.

Carmichael, O.T., Aizenstein, H.A., Davis, S.W., Becker, J.T., Thompson, P.M., Meltzer, C.C., Liu, Y. (2005). Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *NeuroImage*, vol. 27, no. 4, pp. 979-90.

Carlesimo, G. A., Cherubini, A., Caltagirone, C., & Spalletta, G. (2010). Hippocampal mean diffusivity and memory in healthy elderly individuals A cross-sectional study. *Neurology*, vol. 74, no. 3, pp. 194-200.

Ceccarelli, A., Rocca, M. A., Falini, A., Tortorella, P., Pagani, E., Rodegher, M., et al. (2007). Normal-appearing white and grey matter damage in MS - A volumetric and diffusion tensor MRI study at 3.0 Tesla. *Journal of Neurology*, vol. 254, no. 4, pp. 513-518.

Cercignani, M., Bozzali, M., Iannucci, G., Comi, G., & Filippi, M. (2001). Magnetisation transfer ratio and mean diffusivity of normal appearing white and grey matter from patients with multiple sclerosis. *Journal of Neurology Neurosurgery and*

Psychiatry, vol. 70, no. 3, pp. 311-317.

Chaljub G, Johnson III R.F., Johnson Jr. R.F., Sitton C.W. (1999). Unusually exuberant hyperostosis frontalis interna: MRI. *Neuroradiology*, vol. 41, no. 1, pp. 44-45.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation ModelingModelling: A Multidisciplinary Journal*, vol. 14, no. 3, pp. 464-504.

Cherubini, A., Peran, P., Spoletini, I., Di Paola, M., Di Iulio, F., Hagberg, G. E., et al. (2010). Combined Volumetry and DTI in Subcortical Structures of Mild Cognitive Impairment and Alzheimer's Disease Patients. *Journal of Alzheimers Disease*, vol. 19, no. 4, pp. 1273-1282.

Chetelat, G., Baron, J.C. (2003). Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. *NeuroImage*, vol. 18, no. 2, pp. 525-541.

Christensen, K. (2008). Human biodemography: some challenges and possibilities. *Demographic research*, vol. 19, no. 43, pp. 1575-1586.

Chupin, M., Chételat, G., Lemieux, L., Dubois, B., Garnero, l., Benali, H., Eustache, F., Lehéricy, S., Desgranges, B., Colliot, O. (2008). Fully automatic hippocampus segmentation discriminates between early Alzheimer's disease and normal aging. IEEE, doi: 10.1109/ISBI.2008.4540941.

Coffey, C.E., Lucke, J.F., Saxton, J.A., Ratcliff, G., Unitas, L.J., Billig, B., Bryan, R.N. (1998). Sex differences in brain aging: a quantitative magnetic resonance imaging study. *Archives of Neurology*, vol. 55, no. 2, pp. 169-79.

Coffey, C.E., Weiner, R.D., Djang, W., Figiel, G., Soady, S., Patterson, L., Holt, P.D., Spritzer, C.E., Wilkinson, W.E. (1991). Brain anatomic effects of electroconvulsive therapy. *Archives of General Psychiatry*, vol. 48, no. 11, pp. 1013-1021.

Coffey, C.E., Wilkinson, W.E., Parashos, I.A., Soady, S.A., Sullivan, R.J., Patterson, L.J., Figiel, G.S., Webb, M.C., Spritzer, C.E., Djang, W.T. (1992). Quantitative cerebral anatomy of the aging human brain: a cross-sectional study using magnetic resonance imaging. *Neurology*, vol. 42, no. 3, part 1, pp. 527–536.

Convit, A., Wolf, O.T., de Leon, M.J., Patalinjug, M., Kandil, E., Caraos, C., Scherer, A., Saint Louis, L.A., Cancro, R. (2001). Volumetric analysis of the pre-frontal regions: findings in aging and schizophrenia. *Psychiatry Research: Neuroimaging*, vol. 107, no. 2, pp. 61-73.

Corley, J., Gow, A. J., Starr, J. M., & Deary, I. J. (2010). Is Body Mass Index in Old Age Related to Cognitive Abilities? The Lothian Birth Cohort 1936 Study. *Psychology and Aging*, vol. 25, no. 4, pp. 867-875.

Corley, J., Jia, X., Brett, C. E., Gow, A. J., Starr, J. M., Kyle, J. A., McNeill, G., Deary, I.J. (2011). Alcohol Intake and Cognitive Abilities in Old Age: The Lothian

Birth Cohort 1936 Study. *Neuropsychology*, vol. 25, no. 2, pp. 166-175.

Costafreda, S.G., Dinov, I.D., Tu, Z., Shi, Y., Liu, C-Y., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Wahlund, L-O., Spenger, C., Toga, A.W., Lovestone, S., Simmons, A. (2011). Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. *NeuroImage*, vol. 56, no. 1, pp. 212-219.

Courchesne, E., Chisum, H.J., Townsend, J., Cowles, A., Covington, J., Egaas, B., Harwood, M., Hids, S., Press, G.A. (2000). Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology*, vol. 216, no. 3, pp. 672-682.

Cowell, P.E., Turetsky, B.I., Gur, R.C., Grossman, R.I., Shtasel, D.L., Gur, R.E. (1994). Sex differences in aging of the human frontal and temporal lobes. *The Journal of Neuroscience*, vol. 14, no. 8, pp. 4748-55.

Cox, S.R., Ferguson, K.J., Royle, N.A., Shenkin, S.D., MacPherson, S.E., MacLullich, A.M., Deary, I.J., Wardlaw, J.M. (2014). A systematic review of brain frontal lobe parcellation techniques in magnetic resonance imaging. *Brain structure and function*, vol. 19, no. 1, pp. 1-22.

Cracknell, R. [Online, 2010], Avaliable: http://www.parliament.uk/documents/commons/lib/research/key_issues/Key%20Issues%20The%20ageing%20population2007.pdf. Date Accessed: 02/04/2014.

Creasey, H., Rapoport, S.I. (1985). The aging human brain. *Annals of neurology*, vol. 17, no. 1, pp. 2-10.

Crespo-Facorro, B., Kim, J.J., Andreasen, N.C., O'Leary, D.S., Wiser, A.K., Bailey, J.M., Harris, G., Magnotta, V.A. (1999). Human frontal cortex: an MRI-based parcellation method. *NeuroImage*, vol. 10, no. 5, pp. 500-519.

Crum, W.R., Scahill, R.I., Fox, N.C. (2001). Automated hippocampal segmentation by regional fluid registration of serial MRI: Validation and application in Alzheimer's disease. *NeuroImage*, vol. 13, no. 5, pp. 847-855.

Csernansky, J.G., Wang, L., Joshi, S., (2000). Early DAT is distinguished from aging by high-dimensional mapping of the hippocampus. *Neurology*, vol. 55, no. 11, pp. 1636-43.

Davies, G. R., Ramio-Torrenta, L., Hadjiprocopis, A., Chard, D. T., Griffin, C. M. B., Rashid, W., et al. (2004). Evidence for grey matter MTR abnormality in minimally disabled patients with early relapsing-remitting multiple sclerosis. *Journal of Neurology Neurosurgery and Psychiatry*, vol. 75, no. 7, pp. 998-1002.

Davis, J.C., Nagamatsu, L.S., Hsu, C.L., Beattie, B.L., Liu-Ambrose, T. (2012). Self-efficacy is independently associated with brain volume in older women. *Age and Ageing*, vol. 41, no. 4, pp. 495-501.

Deary, I. J., Bastin, M. E., Pattie, A., Clayden, J. D., Whalley, L. J., Starr, J. M., Wardlaw, J. M., 2006. White matter integrity and cognition in childhood and old age. *Neurology*. vol. 66, no. 4, pp. 505-512.

Deary, I.J., Corley, J., Gow, A.J., Harris, S.E., Houlihan, L.M., Marioni, R.E., Penke, L., Rafnsson, S.B., Starr, J.M. (2009a). Age-associated cognitive decline. *British Medical Bulletin*, vol. 92, no. 1, pp. 135-152.

Deary, I. J., Der, G., & Ford, G., 2001. Reaction times and intelligence differences; a population-based cohort study. *Intelligence*, vol. 29, no.5, pp. 389–399.

Deary I.J., Gow, A.J., Pattie, A., Starr, J.M. (2012). Cohort profile: The Lothian birth cohorts of 1921 and 1936. *International Journal of Epidemiology*, vol. 41, no. 6, pp. 1576-1584.

Deary I.J., Gow A.J., Taylor M.D., Corley J., Brett C., Wilson V., Campbell H., Whalley L.J., Visscher P.M., Porteous D.J., Starr J.M. (2007). The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatric*, vol. 7:28.

Deary IJ, Johnson W. (2010). Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology*, *vol.* 39, no. 5, pp. 1362-9.

Deary IJ, Johnson W, Houlihan LM. (2009b). Genetic foundations of human intelligence. *Human Genetics*, vol.126, no. 1, pp. 215-32.

Deary, I. J., Penke, L., Johnson, W., 2010. The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, vol. 11, no. 3, pp. 201-211.

Deary, I.J., Yang, J., Davies, G., Harris, S.E., Tenesa, A., Liewald, D., Luciano, M., Lopez, L.M., Gow, A.J., Corley, J., Redmond, P., Fox, H.C., Rowe, S.J., Haggerty, P., McNeill, G., Goddard, M.E., Porteous, D.J., Whalley, L.J., Starr, J.M., Vizzcher, P.M., 2012. Genetic contributions to stability and change in intelligence from childhood to old age. *Nature*, vol. 482, pp. 212-215.

Desai, A.K., Grossberg, G.T., Chibnall, J.T. (2010). Healthy brain aging: a road map. *Clinics in Geriatric Medicine*, vol. 26, no. 1, pp.1-16.

Dewey, J., Hana, G., Russell, T., Price, J., McCaffrey, D., Harezlak, J., Sem, E., Anyanwu, J.C., Guttmann, C.R., Navia, B., Cohen, R., Tate, D.F. and the HIV Neuroimaging Consortium. (2010). Reliability and validity of MRI-based automated volumetry software relative to auto-assisted manual measurement of subcortical structures in HIV-infected patients from multisite study. *NeuroImage*, vol. 51, no. 4, pp. 1334-1344.

Dewing, J., Dijk, S. (2014). What is the current state of care for older people with dementia in general hospitals? A literature review. Dementia (London), [Epub ahead of print].

Dickerson, B.C., Stoub, T.R., Shah, R.C., Sperling, R.A., Kiliany, R.J., Albert, M.S., Hyman, B.T., Blacker, D., de Toledo-Morrell, L. (2011). Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. *Neurology*, vol. 76, no. 16, pp. 1395-1402.

Diniz, P. R. B., Velasco, T. R., Salmon, C. E. G., Sakamoto, A. C., Leite, J. P., & Santos, A. C. (2011). Extratemporal Damage in Temporal Lobe Epilepsy: Magnetization Transfer Adds Information to Volumetric MR Imaging. *American Journal of Neuroradiology*, vol. 32, no. 10, pp. 1857-1861.

Doring, T.M., Kubo, T.T.A., Domingues, R.C., Gasparetto, E.L. (2010). Evaluation of hippocampal volume based on MRI applying manual and automatic segmentation techniques. *Revista Brasileira de fisica medica*, vol. 4, no. 1, pp. 89-91.

Driscoll, I., Davatzikos, C., An, Y., Wu, X., Shen, D., Kraut, M., Resnick, S.M., 2009. Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. *Neurology*, vol. 72, no. 22, pp. 1906-1913.

Du, A.T., Schuff, N., Chao, L.L., Kornak, J., Jagust, W.J., Kramer, J.H., Reed, B.R., Miller, B.L., Norman, D., Chui, H.C., Weiner, M.W. (2006). Age effects on atrophy rates of entorhinal cortex and hippocampus. *Neurobiology of aging*, vol. 27, no. 5, pp. 733-740.

Duvernoy, H.M. (1999). The human brain: surface, blood supply, and three-dimensional sectional anatomy, 2nd edition. Wien, Springer.

Eckerstrom, C., Olsson, E., Borga, M., Ekholm, S., Ribbelin, S., Rolstad, S., et al. (2008). Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: The Goteborg MCI study. *Journal of the Neurological Sciences*, vol. 272, no. 1-2, pp. 48-59.

Entis, J.J., Doerga, P., Barrett, L.F., Dickerson, B.C. (2012). A reliable protocol for the manual segmentation of the human amygdala and its subregions using ultra-high resolution MRI. *Neuroimage*, vol. 60, no. 2, pp. 1226-1235.

Erickson, K.I., Prakash, R.S., Voss, M.W., Chaddock, L., Hu, L., Morris, K.S., White, S.M., Wójcicki, T.R., McAuley, E., Kramer, A.F. (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus*, vol. 19, no. 10, pp. 1030-1039.

Erickson, K.I., Prakash, R.S., Voss, M.W., Chaddock, L., Heo, S., McLaren, M., Pence, B.D., Martin, S.A., Vieira, V.J., Woods, J.A., McAuley, E., Kramer, A.F. (2010). Brain-derived neurotrophic factor is associated with age-related decline in hippocampal volume. *The Journal of Neuroscience*, vol. 30, no. 15, pp. 5368-5375.

Erickson, K.I., Voss, M.W., Prakash, R.S., Basak, C., Szabo, A., Chaddock, L., Kim, J.S., Heo, S., Alves, H., White, S.M., Wojcicki, T.R., Mailey, E., Vieira, V.J., Martin, S.A., Pence, B.D., Woods, J.A., McAuley, E., Kramer, A.F. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences*, vol. 108, no. 7, pp. 3017-3022.

Evans, A.C., Collins, D.L., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *IEEE Nuclear Science Symposium Medical Imaging*, vol. 3, no. 1-3, pp. 1813-1817.

Evans, A.C., Janke, A.L., Collins, D.L., (2012). Baillet, S. Brain templates and atlases. *Neuroimage*, vol. 62, no. 2, pp. 911-922.

Ezekiel F., Chao L., Kornak J., Du A-T., Cardenas V., Truran D., Jagust, W., Chui, H., Miller, B., Yaffe, K., Schuff, N., Weiner, M. (2004). Comparisons between global and focal brain atrophy rates in normal aging and Alzheimer disease. Boundary shift integral versus tracing of the entorhinal cortex and hippocampus. *Alzheimer Disease & Associated Disorders*, vol. 18, no. 41, pp. 96–201.

Farias, S.T., Mungas, D., Reed, B., Carmichael, O., Beckett, L., Harvey, D., Olichney, J., Simmons, A., DeCarli, C., 2012. Maximal brain size remains an important predictor of cognition in old age, independent of current brain pathology. *Neurobiology of Aging*, vol. 33, no. 8, pp. 1758-68.

Ferguson, K.J., Wardlaw, J.M., Edmond, C.L., Deary, I.J., MacLullich, A.M.J. (2005). Intracranial Area: A validated method of estimating intracranial volume. *Journal of Neuroimaging*, vol, 15, no. 1, pp. 76-78.

Ferreira, D., Molina, Y., Machado, A., Westman, E., Wahlund, L-O., Nieto, A., Correia, R., Junque', C., Diaz-Flores, L., Barroso, J. (2014). Cognitive decline is mediated by gray matter changes during middle age. *Neurobiology of aging*, vol. 35, no. 5, pp. 1086-1094.

Filler, A. (2009). The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: CT, MRI and DTI. *Nature precedings*, doi:10.1038/npre.2009.3267.5.

Filipek, P.A., Semrud-Clikeman, M., Steingard, R.J., Renshaw, P.F., Kennedy, D.N., Biederman, J. (1997). Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology*, vol. 48, no. 3, pp. 589-601.

Filippi, M., Cercignani, M., Bozzali, M., Iannucci, G., & Comi, G. (2000). MTR and mean diffusivity of normal appearing white and grey matter from patients with MS. *European Journal of Neurology*, vol. 7, no. 3, pp. 311-317.

Filippi, M., & Rovaris, M. (2000). Magnetisation transfer imaging in multiple sclerosis. *Journal of NeuroVirology*, vol. 6, no.2, pp. S115-S120.

Finby, N., Kraft, E. (1972). The aging skull: comparative Roentgen study 25 to 34 year interval. *Clinical Radiology*, vol. 23, no. 4, pp. 410-414.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ven der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M. (2002). Whole brain segmentation: automated labeling of

neuroanatomical structures in the human brain. *Neuron,* vol. 33, no. 3, pp. 341-55.

Fjell, A.M. and Walhovd, K.B. (2010). Structural brain changes in aging: courses, causes and cognitive consequences. *Reviews in the Neurosciences*, vol. 21, no.3, pp. 187-221.

Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M. (2009). One-year brain atrophy evident in healthy aging. *The journal of Neuroscience*, vol. 29, no. 48, pp. 15223-15231.

Fjell A.M, Westlye L.T., Amlien I., Espeseth T., Reinvang I., Raz N., Agartz I., Salat D.H., Greve D.N., Fischl B., Dale A.M., Walhovd K.B. (2009). Minute effects of sex on the aging brain: a multisample magnetic resonance imaging study of healthy aging and Alzheimer's disease. *Journal of Neuroscience*, vol. 29, no. 27, pp. 8774–8783.

Fjell, A.M., Westlye, L.T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Holland, D., Brewer, J.B., Dale, A.M., Walhovd, K.B., and Alzheimer Disease Neuroimaging Initiative. (2013). Critical ages in the life course of the adult brain: nonlinear subcortical aging. *Neurobiology of aging*, vol. 34, no. 10, pp. 2239-2247.

Foerster, A., Griebe, M., Gass, A., Kern, R., Hennerici, M. G., & Szabo, K. (2012). Diffusion-Weighted Imaging for the Differential Diagnosis of Disorders Affecting the Hippocampus. *Cerebrovascular Diseases*, vol. 33, no. 2, pp. 104-115.

Folstein, M.F., Folstein, S.E., McHugh, P.R. (1975). Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, vol. 12, no.3, pp. 189-198.

Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., Brain development Cooperative Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage,* vol. 54, no. 1, pp. 313-327.

French, B.F., Finch, W.H., 2006. Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, vol.13, no. 3, pp. 378-402.

Frisoni, G.B., Fox, N.C., Jack Jr, C.R., Scheltens, P., Thompson, P.M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nature Review: Neuroscience,* vol. 6, no. 2, pp. 67-77.

Frisoni, G.B., Testa, C., Zorzan, A., Sabattoli, F., Beltramello, A., Soininen, H., Laakso, M.P. (2002). Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. *Journal of Neurology, Neurosurgery & Psychiatry,* vol. 73, no. 6, pp. 657-664.

Galton, C.J., Gomez-Anson, B., Antoun, N., Scheltens, P., Patterson, K., Graves, M., Sahakian, B.J., Hodges, J.R. (2001). Temporal lobe rating scale: application to Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery and Psychiatry,* vol. 70, no. 2, pp. 165-173.

Ge, Y., Grossman, R.I., Babb, J.S., Rabin, M.L., Mannon, L.J., Kolson, D.L. (2002). Age-related total gray matter and white matter changes in normal adult brain. Part 1: volumetric MR imaging analysis. *American Journal of Neuroradiology,* vol. 23, no. 8m pp. 1327-1333.

Geva, T. (2006). Magnetic resonance imaging: historical perspective. *Journal of cardiovascular magnetic resonance*, vol. 8, no. 4, pp. 573-580.

Ghisletta, P, Rabbitt, P., Lunn, M., Lindenberger, U., 2012. Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. *Intelligence,* vol. 40, no. 3, pp. 260-269.

Ginovart, N., Lundin, A., Farde, L., Halldin, C., Backman, L., Swahn, C.G., Pauli, S., Sedvall, G. (1997). PET study of the pre- and post-synaptic dopaminergic markers for the neurodegenerative process in Huntington's disease. *Brain,* vol. 120, no. 3, pp. 503-514.

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage,* vol. 51, no. 3, pp. 943-951.

Global Age Watch Index 2013 [Online], Available: http://www.helpage.org/global-agewatch/. Date accessed 02/04/2014.

Gold, S.M., Dziobek, I., Rogers, K., Bayoumy, A., McHugh, P.F., Convit, A. (2005). Hypertension and hypothalamo-pituitary-adrenal axis hyperactivity affect frontal lobe integrity. *The Journal of Clinical Endocrinology and Metabolism*, vol. 90, no. 6, pp. 3262-3267.

Gosche, K.M., Mortimer, J.A., Smith, C.D., Markesbery, W.R., Snowdon, D.A. 2001. An automated technique for measuring hippocampal volumes from MR imaging studies. *American Journal of Neuroradiology,* vol. 22, no. 9, pp. 1686-9.

Grabner, G., Janke, A.L., Budge, M.M., Smith, D., Pruessner, J., Collins, D.L. (2006). Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention,* vol. 9, no. 2, pp. 58-66

Greenberg, D.L., Messer, D.F., Payne, M.E., MacFall, J.R., Provenzale, J.M., Steffens, D.C., Krishnan, R.R. (2008). Aging, gender, and the elderly adult brain: an examination of analytical strategies. *Neurobiology of Aging*, vol. 29, no. 2, pp.. 290-302.

Greenwood, P.M. (2000). The frontal aging hypothesis evaluated. *Journal of the International Neuropsychological Society,* vol. 6, pp. 705-726.

Griscom N.T., Sang Oh K. (1970). The contracting skull: Inward growth of the inner table as a physiological response to diminution of intracranial content in children. *American Journal of Roentgenology,* vol. 110 , no. 1, pp. 106-110.

Gunning-Dixon, F.M., Brickman, A.M., Cheng, J.C., Alexopoulos, G.S. (2009). Aging of cerebral white matter: a review of MRI findings. *International journal of geriatric psychiatry*, vol. 24, no. 2, pp.109-117.

Gur, R.C., Gur, R.E. (2004). Gender differences in the functional organization of the brain. *Principles of gender-specific medicine*. San Diego: Elsevier.

Hanyu, H., Shimizu, S., Tanaka, Y., Kanetaka, H., Iwamoto, T., & Abe, K. (2005). Differences in magnetization transfer ratios of the hippocampus between dementia with Lewy bodies and Alzheimer's disease. *Neuroscience Letters*, vol. 380, no. 1-2, pp. 166-169.

Harding F.E. (1949). Endocrinopathies associated with Hyperostosis frontalis interna. *The American Journal of Medicine*, vol. 6, no. 3, pp. 329-335.

Hasan, K.M., Halphen, C., Boska, M.D., Narayana, P.A. (2008). Diffusion tensor metrics, T(2) relaxation and volumetry of the naturally aging human caudate nuclei in healthy young and middle-aged adults: possible implications for the neurobiology of human brain aging and disease. *Magnetic Resonance Medicine*, vol. 59, no. 1, pp. 7-13.

Hatipoglu H.G., Ozcan H.N., Hatipoglu U.S., Yuksel E. (2008). Age, sex and body mass index in relation to calvarial diploe thickness and craniometric data on MRI. *Forensic Science International*, vol. 182, no. 1-3, pp. 46-51.

den Heijer, T., van der Lijn, F., Vernooij, M.W., de Groot, M., Koudstaal, P.J., van der Lugt, A., Krestin, G.P., Hofman, A., Niessen, W.J., Breteler, M.M.B. (2012). Structural and diffusion MRI measures of the hippocampus and memory performance. *Neuroimage*, vol. 63, no. 4, pp. 1782-1789.

HM Treasury [December, 2013], https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263942/35062_Autumn_Statement_2013.pdf. Date accessed: 02/04/2014.

Hofer, S.M., Flaherty, B.P., Hoffman, L., 2006. Cross-sectional analysis of time-dependent data: Mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohort samples. *Multivariate Behavioural Research*, vol. 41, no. 2, pp. 165-187.

Hogan, R.E., Mark, K.E., Wang, L., Joshi, S., Miller, M.I., Bucholz, R.D. (2000). Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. *Radiology*, vol. 216, no.1, pp. 291-297.

Holland, G.N., Botomley, P.A. (1977). A color display technique for NMR imaging. *Journal of Physics E*, vol. 10, no.7, pp. 14-16.

Hong, Y. J., Yoon, B., Shim, Y. S., Cho, A. H., Lim, S. C., Ahn, K. J., et al. (2010). Differences in Microstructural Alterations of the Hippocampus in Alzheimer Disease and Idiopathic Normal Pressure Hydrocephalus: A Diffusion Tensor Imaging Study.

American Journal of Neuroradiology, vol. 31, no. 10, pp. 1867-1872.

Houlihan LM, Harris SE, Luciano M, Gow AJ, Starr JM, Visscher PM, Deary IJ. (2009). Replication study of candidate genes for cognitive abilities: the Lothian Birth Cohort 1936. *Genes, Brain and Behavior*, vol. 8, no. 2, pp. 238-47.

Hsu, Y-Y., Schuff, N., Du, A-T., Mark, K., Zhu, X., Hardin, D., Weiner, M.W. (2002). Comparison of automated and manual MRI volumetry of hippocampus in normal ageing and dementia. *Journal of Magnetic Resonance Imaging*, vol. 16, no. 3, pp. 305-310.

Hua, X., Hibar, D.P., Lee, S., Toga, A.W., Jack Jr, C.R., Weiner, M.W., Thompson, P.M., the Alzheimer's Disease Neuroimaging Initiative. (2010). Sex and age differences in atrophic rates: an ADNI study with *n*= 1368 MRI scans. *Neurobiology of Aging*, vol. 31, no. 8, pp. 1463-1480.

Hyman, B. T., Vanhoesen, G. W., Damasio, A. R., & Barnes, C. L. (1984). Alzheimer-Disease cell specific pathology isolates the hippocampal-formulation. Science, 225(4667), 1168-1170.

Ikram MA, Vrooman HA, Vernooij MW, den Heijer T, Hofman A, Niessen WJ, van der Lugt A, Koudstaal PJ, Breteler MM. (2010). Brain tissue volumes in relation to cognitive function and risk of dementia. Neurobiology of Aging, vol. 31, no. 3, pp. 378-86.

Israel, H. (1968). Continuing growth in the human cranial skeleton. *Archives of Oral Biology*, vol. 13, no.1, pp. 133-137.

Jaccard, P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, vol. 37, pp. 547-579.

Jafari-Khouzani, K., Elisevich, K.V., Patel, S., Soltanian-Zadeh, H. (2011). Dataset of magnetic resonance images of nonepileptic subjects and temporal lobe epilepsy patients for validation of hippocampal segmentation techniques. *Neuroinformatics*, vol. 9, no.4, pp. 335-346.

Jack, C.R. Jr., Bentley, M.D., Twomey, C.K., Zinsmeister, A.R. (1990). MR imaging-based volume measurements of the hippocampal formation and anterior temporal lobe: validation studies. *Radiology*, vol. 176, no.1, pp. 205-209.

Jenkinson, M., Bannister, P., Brady, J. M. and Smith, S. M. (2002). Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, vol. 17, no. 2, pp. 825-841.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, vol. 5, no. 2, pp. 143-156.

Jernigan, T.L., Archibald, S.L., Berhow, M.T., Sowell, E.R., Foster, D.S. Hesselink, J.R. (1991). Cerebral structure on MRI, part1: Localization of age-related changes.

Biological Psychiatry, vol. 29, pp. 55-67.

Jernigan, T.L., Archibald, S.L., Fennema-Notestine, C., Gamst, A.C., Stout, J.C., Bonner, J., Hesselink, J.R. (2001). Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiology of Aging*, vol. 22, no. 4, pp. 581-594.

Jiang, J., Sachdev, P., Lipnicki, D.M., Zhang, H., Liu, T., Zhu, W., Suo, C., Zhuang, L., Crawford, J., Reppermund, S., Trollor, J., Brodaty, H., Wen, W. (2014). A longitudinal study of brain atrophy over two years in community-dwelling older individuals. *NeuroImage*, vol. 86, pp. 203-211.

Jones, D.K., Williams, S.C., Gasston, D., Horsfield, M.A., Simmons, A., Howard, R. (2002). Isotropic resolution diffusion tensor imaging with whole brain acquisition in a clinically acceptable time. *Human Brain Mapping*, vol. 15, no. 4, pp. 216-230.

Joseph, J., Warton, C., Jacobson, S.W., Jacobson, J.L., Molteno, C.D., Eicher, A., Marais, P., Phillips, O.R., Narr, K.L., Meintjes, E.M. (2012). Three-dimensional surface deformation-based shape analysis of hippocampus and caudate nucleus in children with fetal alcohol spectrum disorders. *Human brain mapping*, vol. 35, no. 2, pp. 659-672.

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., Dale, A. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, vol. 30, no. 2, pp. 436-443.

Karl, A., Schaefer, M., Malta, L.S., Dorfel, D., Rohleder, N., Werner, A. (2006). A meta-analysis of structural brain abnormalities in PTSD. *Neuroscience and Biobehavioural Reviews*, vol. 30, no.7, pp. 1004-1031.

Kates, W.R., Frederikse, M., Mostofsky, S.H., Folley, B.S., Cooper, K., Mazur-Hopkins, P., Kofman, O., Singer, H.S., Denckla, M.B., Pearlson, G.D., Kaufmann, W.E. (2002). MRI parcellation of the frontal lobe in boys with attention deficit hyperactivity disorder or Tourette syndrome. *Psychiatry Research*, vol. 116, no. 1-2, pp. 63-81.

Keihaninejad, S., Heckemann, R.A., Fagiolo, G., Symms, M.R., Hajnal, J.V., Hammers, A., The Alzheimer's Disease Neuroimaging Initiative. (2010). A robust method to estimate the intracranial volume across MRI field strengths (1.5T and 3T). *Neuroimage*, vol. 50, no. 4, pp.1427-1437.

Khan, A.R., Wang, L., Beg, M.F. (2008). FreeSurfer-initiated fully-automated subcortical brain segmentation in MRI using Large Deformation Diffeomorphic Metric Mapping. *NeuroImage*, vol. 41, no. 3, pp. 735-46.

Khan, A.R., Cherbuin, N., Wen, W., Antsey, K.J., Sachdev, P., Beg, M.F. (2011). Optimal weights for local multi-atlas fusion using supervised learning and dynamic information (SuperDyn): Validation on hippocampus segmentation. *NeuroImage*, vol. 56, no. 1, pp. 126-139.

Kim, H., Mansi, T., Bernasconi, N., Bernasconi, A. (2012). Surface-based multi-template automated hippocampal segmentation: application to temporal lobe epilepsy. *Medical Image Analysis,* vol. 16, no.7, pp. 1445-1455.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage,* vol. 46, no. 3, pp. 786-802.

Kochunov, P., Ramage, A.E., Lancaster, J.L., Robin, D.A., Narayana, S., Coyle, T., Royall, Fox, P. (2009). Loss of cerebral white matter structural integrity tracks the gray matter metabolic decline in normal aging. *Neuroimage,* vol. 45, no. 1, pp. 17-28.

Konrad, C., Ukas, T., Nebel, C., Arolt, V., Toga, A.W., Narr, K.L. (2009). Defining the human hippocampus in cerebral magnetic resonance images – an overview of current segmentation protocols. *Neuroimage,* vol. 47, no. 4, pp. 1185-1195.

Kooistra, M., Geerlings, M.I., van der Graaf, Y., Mali, W.P., Vincken, K.L., Kappelle, L.J., Muller, M., Biessels, G.J., SMART-MR study group. (2014). Vascular brain lesions, brain atrophy and cognitive decline. The Second Manifestations of ARTerial disease – Magnetic Resonance (SMART-MR) study. *Neurobiology of Aging,* vol. 35, no. 1, pp. 35-41.

Kosior, R. K., Lauzon, M. L., Federico, P., & Frayne, R. (2011). Algebraic T2 estimation improves detection of right temporal lobe epilepsy by MR T2 relaxometry. *NeuroImage,* vol. 58, no. 1, pp. 189-197.

Lacerda, A.L.T, Hardan, A.Y., Yorbik, O., Keshavan, M.S. (2003). Measurement of the orbitofrontal cortex: a validation study of a new method. *Neuroimage,* vol. 19, no. 3, pp. 665-673.

Lancaster, J.L., Rainey, L.H., Summerlin, J.L., Freitas, C.S., Fox, P.T., Evans, A.C., Toga, A.W., Mazziotta, J.C. (1997). Automated labeling of the human brain. *Human Brain Mapping,* vol. 5, no. 4, pp. 238-242.

Laude, A., Lascaratos, G., Henderson, R.D., Starr, J.M., Deary, I.J., Dhillon, B. (2013). Retinal nerve fiber layer thickness and cognitive ability in older people: the Lothian Birth Cohort 1936 study. *BMC Opthalmology*, 13:28.

Lawlor DA, Batty GD, Morton SM, Deary IJ, Macintyre S, Ronalds G, Leon DA. (2005). Early life predictors of childhood intelligence: evidence from the Aberdeen children of the 1950s study. *Journal of Epidemiology & Community Health,* vol. 59, no.8, pp. 656-63.

de Leeuw. F.E., de Groot, J.C., Achten, E., Oudkerk, M., Ramos, L.M.P., Heijboer, R., Hofman, A., Jolles, J., van Gijn, J., Breteler, M.M.B. (2001). Prevalence of cerebral white matter lesions in elderly people: a population based magnetic resonance imaging study. The Rotterdam Scan study. *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 70, no. 1, pp. 9-14.

Lemaitre, H., Goldman, A.L., Sambataro, F., Verchinski, B.A., Meyer-Lindenberg, A., Weinberger, D.R., Mattay, V.S. (2012). Normal age-related brain morphometric changes: nonuniformity across cortical thickness, surface area and gray matter volume? *Neurobiology of Aging*, vol. 33, no.3, 617.e1-617.e9.

Leow, A.D,. Yanovsky, I., Parikshak, N., Hua, X., Lee, S., Toga, A.W., Jack Jr, C.R., Bernstein, M.A., Britson, P.J, Gunter, J.L., Ward, C.P., Borowski, B., Shaw, L.M., Trojanowski, J.Q., Fleisher, A.S., Harvey, D., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W. Thompson, P.M. and The Alzheimer's Disease Neuroimaging Initiative. (2009). Alzheimer's Disease Neuroimaging Initiative: A one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. *Neuroimage*, vol. 45, no.3, pp. 645-655.

Leung, K. K., Barnes, J., Ridgway, G. R., Bartlett, J. W., Clarkson, M. J., Macdonald, K., et al. (2010). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *NeuroImage*, vol. 51, no. 4, pp. 1345-1359.

Liang, A. L. W., Vavasour, I. M., Madler, B., Traboulsee, A. L., Lang, D. J., Li, D. K. B., et al. (2012). Short-term stability of T (1) and T (2) relaxation measures in multiple sclerosis normal appearing white matter. *Journal of Neurology*, vol. 259, no. 6, pp. 1151-1158.

van der Lijn, F., den Heijer, T., Breteler, M.M.B., Niessen, W.J. (2008). Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. *NeuroImage*, vol. 43, no. 4, pp. 708-20.

Lim, H.K., Hong, S.C., Jung, W.S., Ahn, K.J., Won, W.Y., Hahn, C., Kim, I., Lee, C.U. (2012). Automated hippocampal subfields segmentation in late life depression. *Journal of Affective Disorders*, vol. 143, no. 1-3, pp. 253-256.

Lim, H.K., Jung, W.S., Ahn, K.J., Won, W.Y., Hahn, C., Lee, S.Y., Kim, I., Lee, C.U. (2012). Relationships between hippocampal shape and cognitive performances in drug-naive patients with Alzheimer's disease. *Neuroscience Letters*, vol. 516, no.1, pp. 124-129.

Lorenzetti, V., Allen, N.B., Fornito, A., Yucel, M. (2009). Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. *Journal of Affective Disorders*, vol. 117, no. 1-2, pp. 1-17.

Luciano, M., Wright, M. J., Smith, G. A., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. *Behavior Genetics*, vol. 31, no. 6, pp. 581-592.

Lyoo, I.K., Han, M.H., Cho, D.Y. (1998). A brain MRI study in subjects with borderline personality disorder. *Journal of Affective Disorders*, vol. 50, no. 2-3, pp. 235-43.

Madsen, S.K., Gutman, B.A., Joshi, S.H., Toga, A.W., Jack Jr, C.R., Weiner, M.W. Thompson, P.M. (2013). Mapping dynamic changes in ventricular volume onto baseline cortical surfaces in normal aging, MCI, and Alzheimer's Disease. *MBIA: Lecture notes in computer science*, vol. 8159, pp. 84-94.

Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L. (2007). Open access series of imaging studies (OASIS), cross-sectional MRI data in young, middle aged, nondemented and demented older adults. *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498-1507.

Margariti, P. N., Blekas, K., Katzioti, F. G., Zikou, A. K., Tzoufi, M., & Argyropoulou, M. I. (2007). Magnetization transfer ratio and volumetric analysis of the brain in macrocephalic patients with neurofibromatosis type 1. European Radiology, vol. 17, no. 2, pp. 433-438.

May H., Peled N., Dar G., Abbas J., Medlej B., Masharawi Y., Hershkovitz I. (2010). Hyperostosis frontalis interna and androgen suppression. *The Anatomical Record*, vol. 239, no. 8, pp. 1333-1336.

McDaniel, M.A., 2005. Big-brained people are smarter: a meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence*, vol. 33, pp. 337-346.

McDonald, W. I., Miller, D. H., & Barnes, D. (1992). The pathological evolution of multi-sclerosis. Neuropathology and Applied Neurobiology, vol. 18, no. 4, pp. 319-334.

Messina, D., Cerasa, A., Condino, F., Arabia, G., Novellino, F., Nicoletti, G., Salsone, M., Morelli, M., Lanza, P.L., Quattrone, A. (2011). Patterns of brain atrophy in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Parkinsonism and Related Disorders*, vol. 17, no. 3, pp. 172-176.

Miller, G. F., Penke, L., 2007. The evolution of human intelligence and the coefficient of additive genetic variance in human brain size. *Intelligence*, vol. 35, no. 2, pp. 97-114.

Möller, C., van der Flier, W.M., Versteeg, A., Benedictus, M.R., Wattjes, M.P., Koedam, E.L.G.M., Scheltens, P., Barkhof, F., Vrenken, H. (2014). Quantitative regional validation of the visual rating scale for posterior cortical atrophy. *European Radiology*, vol. 24, no.2, pp. 397-404.

Morey, R.A., Petty, C.M., Xu, Y., Hayes, J.P., Wagner, R.H., Lewis, D.V., LaBar, K.S., Styner, M., McCarthy, G. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage*, vol. 45, no. 3, pp. 855-866.

Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Toga, A.W., Thompson, P.M. (2010). Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. I*EEE*

Transactions on Medical Imaging, vol. 29, no. 1, pp. 30-43.

Morris, Z., Whiteley, W.N., Longstreth, W.T., Weber, F., Lee, Y.C., Tsushima, Y., Alphs, H., Ladd, S.C., Warlow, C., Wardlaw, J.M., Al-Shahi Salman, R. (2009). Incidental findings on brain magnetic resonance imaging: systematic review and meta-analysis. *British Medical Journal,* 339, b3016.

Mouiha, A., Duchesne, S., The Alzheimer's Disease Neuroimaging Initiative. (2011). Hippocampal atrophy rates in Alzheimer's disease: Automated segmentation variability analysis. *Neuroscience Letters,* vol. 495, no.1, pp. 6-10.

Muller, M. J., Greverus, D., Dellani, P. R., Weibrich, C., Wille, P. R., Scheurich, A., et al. (2005). Functional implications of hippocampal volume and diffusivity in mild cognitive impairment. *NeuroImage,* vol. 28, no. 4, pp. 1033 - 1042.

Mungas, D., Harvey, D., Reed, B.R., Jagust, W.J., DeCarli, C., Beckett, L., Mack, W.J., Kramer, J.H., Weiner, M.W., Schuff, N., Chui, H.C. (2005). Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology*, vol. 65, no. 4, pp. 565-571.

Murray, A.D., Staff, R.T., McNeil, C.J., Salarirad, S., Ahearn, T.S., Mustafa, N., Whalley, L.J., 2011. The balance between cognitive reserve and brain imaging biomarkers of cerebrovascular and Alzheimer's diseases. *Brain*, vol. 134, no. 12, pp. 3687-3696

Mustafa, N., Ahearn, T.S., Waiter, G.D., Murray, A.D., Whalley, L.J., Staff, R.T. (2012). Brain structural complexity and life course cognitive change. *Neuroimage,* vol. 61, no.3, pp. 694-701.

Muthen, L.K., Muthen, B.O., 2010. Mplus User's Guide, Sixth Edition, Los Angeles, CA: Muthen & Muthen.

Muzzio, I. A., Kentros, C., & Kandel, E. (2009). What is remembered? Role of attention on the encoding and retrieval of hippocampal representations. *Journal of Physiology - London,* vol. 587, no. 12, pp. 2837-2854.

Nandigam, R.N., Chen, Y.W., Gurol, M.E., Rosand, J., Greenberg, S.M., Smith, E.E. (2007). Validation of intracranial area as a surrogate measure of intracranial volume when using clinical MRI. *Journal of Neuroimaging*, vol. 17, no. 1, pp. 74-77.

Neter, J., Wasserman, W., & Kutner, M. H. (1989). Applied Linear Regression Models. Homewood, IL: Irwin.

Nifosì, F., Toffanin, T., Follador, H., Zonta, F., Padovan, G., Pigato, G., Carollo, C., Ermani, M., Amista, P., Perini, G.I. (2010). Reduced right posterior hippocampal volume in women with recurrent familial pure depressive disorder. *Psychiatry Research,* vol. 184, no. 1, pp. 23-28.

Noga, J.T., Aylward, E., Barta, P.E., Pearlson, G.D. (1995). Cingulate gyrus in schizophrenic patients and normal volunteers. *Psychiatry Research*, vol. 61, no. 4, pp. 201-208.

Nordenskjöld, R., Malmberg, F., Larsson, E-M., Simmons, A., Brooks, S.J., Lind, L., Ahlström, H., Johansson, L., Kullberg, J. (2013). Intracranial volume estimated with commonly used methods could introduce bias in studies including brain volume measurements. *NeuroImage*, vol. 83, pp. 355-360.

Nossin-Manor, R., Chung, A. D., Whyte, H. E. A., Shroff, M. M., Taylor, M. J., & Sled, J. G. (2012). Deep Gray Matter Maturation in Very Preterm Neonates: Regional Variations and Pathology-related Age-dependent Changes in Magnetization Transfer Ratio. *Radiology*, vol. 263, no. 2, pp. 510-517.

Nunnemann, S., Wohlschlager, A.m., Ilg, R., Gaser, C., Etgen, T., Conrad, B., Zimmer, C., Muhlau, M. (2009). Accelerated aging of the putamen in men but not in women. *Neurobiology of Aging*, vol. 30, no.1, pp. 147-151.

Oishi, K., Faria, A.V., Yoshida, S., Chang, L., Mori, S. (2013). Quantitative evaluation of brain development using anatomical MRI and diffusion tensor imaging. *International Journal of Developmental Neuroscience*, vol. 31, no. 7, pp. 512-524.

O'Bryant, S.E., Humphreys, J.D., Smith, G.E., Ivnik, R.J., Graff-Radford, N.R., Peterson, R.C., Lucas, J.A. 2008. Detecting Dementia with the Mini-Mental State Examination in Highly Educated Individuals. *Archives of Neurology*, vol. 65, no. 7, pp. 963-967.

Pal, D., Trivedi, R., Saksena, S., Yadav, A., Kumar, M., Pandey, C. M., et al. (2011). Quantification of age- and gender-related changes in diffusion tensor imaging indices in deep grey matter of the normal human brain. *Journal of Clinical Neuroscience*, vol. 18, no. 2, pp.193-196.

Pan, Y., & Jackson, R. T. (2008). Ethnic difference in the relationship between acute inflammation and serum ferritin in US adult males. *Epidemiology and Infection*, vol.136, no. 3, pp. 421-431.

Pantel, J., Schroder, J., Essig, M., Popp, D., Dech, H., Knopp, M.V., Schad, L.R., Eysenbach, K., Backenstrass, M., Friedlinger, M. (1997). Quantitative magnetic resonance imaging in geriatric depression and primary degenerative dementia. *Journal of Affective Disorders*, vol. 42, no. 1, pp. 69-83.

Pardoe, H.R., Pell, G.S., Abbott, D.F., Jackson, G.D. (2009). Hippocampal volume assessment in temporal lobe epilepsy: how good is automated segmentation? *Epilepsia*, vol. 50, no. 12, pp. 2586-2592.

Park, D.C., Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology*, vol. 60, pp. 173-196.

Parry, A., Clare, S., Jenkinson, M., Smith, S., Palace, J., & Matthews, P. M. (2003). MRI brain T1 relaxation time changes in MS patients increase over time in both the

white matter and the cortex. *Journal of Neuroimaging*, vol. 13, no. 3, pp. 234-239.

Patenaude, B., Smith, S.M., Kennedy, D., and Jenkinson M. (2011). A Bayesian Model of Shape and Appearance for Subcortical Brain. *Neuroimage,* vol. 56, no. 3, pp. 907-922.

Pedraza, O., Bowers, D., Gilmore, R. (2004). Asymmetry of the hippocampus in MRI volumetric measurements of normal adults. *Journal of the international neuropsychological society*, vol. 10, no. 5, pp. 664-678.

Peelle, J.E., Cusack, R., Henson, R.N.A. (2012). Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging. *Neuroimage,* vol. 60, no. 2, pp. 1503-1516.

Pengas, G., Pereira, J.M.S., Williams, G.B., Nestor, P.J. (2009).Comparative Reliability of Total Intracranial Volume Estimation Methods and the Influence of Atrophy in a Longitudinal Semantic Dementia Cohort. *Journal of Neuroimaging*, vol. 19, no. 1, pp. 37-46.

Penke, L., Muñoz Maniega, S., Bastin, M., Valdés Hernández, M.C., Murray, C., Royle, N.A., Starr, J.M., Wardlaw, J.M., Deary, I.J., (2012). Brain white matter tract integrity as a neural foundation for general intelligence. *Molecular Psychiatry*. doi: 10.1038/mp.2012.66.

Penke, L., Muñoz Maniega, S., Murray, C., Gow, A. J., Valdés Hernández, M. C., Clayden, J. D., et al. (2010). A General Factor of Brain White Matter Integrity Predicts Information Processing Speed in Healthy Older People. Journal of Neuroscience, vol. 30, no. 22, pp. 7569-7574.

van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia*, vol. 42, no. 10, pp. 1394-1413.

Pfefferbaum, A., Rohlfing, T., Rosenbloom, M.J., Chu, W., Colrain, I.M., Sullivan, E.V. (2013). Variation in longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85 years) measured with atlas-based parcellation of MRI. *Neuroimage*, vol. 65, pp. 176-193.

Plassman, B.L., Williams, J.W., Burke, J.R., Holsinger, T., Benjamin, S., (2010). Systematic Review: Factors associated with risk for and possible prevention of cognitive decline in later life. *Annals of Internal Medicine*, vol. 153, no. 3, pp. 182-193.

Rademacher, J., Galaburda, A.M., Kennedy, D.N., Filipek, P.A., Caviness, V.S. (1992). Human Cerebral Cortex: Localization, Parcellation, and Morphometry with Magnetic Resonance Imaging. *Journal of Cognitive Neuroscience*, vol. 4, no. 4, pp. 352-374.

Ranta, M.E., Crocetti, D., Clauss, J.A., Kraut, M.A., Mostofsky, S.H., Kaufmann, W.E. (2009). Manual MRI parcellation of the frontal lobe. *Psychiatry Research:*

Neuroimaging, vol. 172, no. 2, pp. 147-154.

Raz, N., Ghisletta, P., Rodrigue, K.M., Kennedy, K.M., Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. *NeuroImage*, vol. 51, no. 2, pp. 501-511.

Raz, N., Lindenberger, U., Rodrigue, K.M., Kennedy, K.M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., Acker, J.D. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebral cortex*, vol. 15, no. 11, pp. 1676-1689.

Raz, N., Rodrigue, K.M. (2006). Differential aging of the brain: patterns, cognitive correlates and modifiers. *Neuroscience and Biobehavioural Reviews*, vol. 30, no. 6, pp. 730-748.

Resnick, S.M., Goldszal, A.F., Davatzikos, C., Golski, S., Kraut, M.A., Metter, E.J., Bryan, R.N., Zonderman, A.B. (2000). One-year age changes in MRI brain volumes in older adults. *Cerebral Cortex*, vol. 10, no. 5, pp. 464-472.

Reuben, A., Brickman, A. M., Muraskin, J., Steffener, J., & Stern, Y. (2011). Hippocampal Atrophy Relates to Fluid Intelligence Decline in the Elderly. Journal of the International Neuropsychological Society, vol. 17, no. 1, pp. 56-61.

Richards M, Deary IJ. (2005). A life course approach to cognitive reserve: a model for cognitive aging and development? *Annals of Neurology*, vol. 58, no.4, pp. 617-22.
Rodrigue, K.M., Kennedy, K.M. (2011). The cognitive consequences of structural changes in the aging brain. *Handbook of the psychology of aging*, seventh edition, Elsevier.

Rombouts, S.A.R.B., Barkhof, F., Witter, M.P., Scheltens, P. (2000). Unbiased whole-brain analysis of gray matter loss in Alzheimer's disease. *Neuroscience Letters*, vol. 285, no. 3, pp. 231-233.

Ropele, S., Schmidt, R., Enzinger, C., Windisch, M., Martinez, N. P., & Fazekas, F. (2012). Longitudinal Magnetization Transfer Imaging in Mild to Severe Alzheimer Disease. *American Journal of Neuroradiology*, vol. 33, no. 3, pp. 570-575.

Rovaris, M., & Filippi, M. (2000). MRI correlates of cognitive dysfunction in multiple sclerosis patients. *Journal of Neurovirology*, vol. 6, S172-S175.

Ruigrok, A.N.V., Salini-Khorshidi, G., Lai, M-C., Baron-Cohen, S., Lombardo, M.V., Tait, R.J., Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioural Reviews*, vol. 39, pp. 34-50.

Rushton, J.P., Ankney, C.D. (2009). Whole brain size and general mental ability: A review. *International Journal of Neuroscience*, vol.119, no. 5, pp. 691-731.

Sabuncu, M. R., Yeo, B. T. T., Van Leemput, K., Fischl, B., & Golland, P. (2010). A Generative Model for Image Segmentation Based on Label Fusion. *IEEE Transactions on Medical Imaging*, vol. 29, no. 10, pp. 1714-1729.

Sahin B., Acer N., Sonmez O.F., Emirzeoglu M., Basaloglu H., Uzun A., Bilgic S. (2007). Comparison of four methods for the estimation of intracranial volume: a gold standard study. *Clinical Anatomy,* vol. 20, no. 7, pp. 766-773.

Salat, D.H., Greve, D.N., Pacheco, J.L., Quinn, B.T., Helmer, K.G., Buckner, R.L., Fischl, B. (2009). Regional white matter volume differences in nondemented aging and Alzheimer's disease. *NeuroImage*, vol. 44, no. 4, pp.1247-1258.

Salthouse, T.A., 1996. The processing-speed theory of adult age differences in cognition. *Psychology Review*, vol. 103, no. 3, pp. 403-428.

Salthouse, T.A., 2010. Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, vol.16, no. 5, pp. 754-760.

Sánchez-Benavides, G., Gómez-Ansón, B., Sainz, A., Vives, Y., Delfino, M., Peña-Casanova, J. (2010). Manual validation of FreeSurfer's automated hippocampal segmentation in normal aging, mild cognitive impairment, and Alzheimer Disease subjects. *Psychiatry Research*, vol. 181, no. 3, pp. 219-25.

Sandeman EM, Valdés Hernández MC, Bastin ME, Murray C, Gow AJ, Corley J, Henderson R, Deary IJ, Starr JM, Wardlaw JM. (2013). Incidental Findings on Brain MR Imaging in Older Community-dwelling Subjects are Common but Serious Medical Consequences are Rare. A Cohort Study. *PLoS ONE*, vol. 8, no. 8.

Scahill, R.I., Frost, C., Jenkins, R., Whitwell, J.L., Rossor, M.N., Fox, N.C. (2003). A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Archives of Neurology*, vol. 60, no. 7, pp. 989-994.

Schermelleh-Engel, K., Moosbrugger, H., Muller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research*, vol. 8, no. 2, pp. 23-74.

Schretlen, D., Perlson, G.D., Anthony, J.C., Aylward, E.H., Augustine, A.M., Davis, A., Barta, P. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society*, vol. 6, no.1, pp. 52-61.

Scottish Council for Research in Education. (1949). The Trend of Scottish Cognitive Ability: A Comparison of the 1947 and 1932 Surveys of the Cognitive Ability of Eleven-Year-Old Pupils. University Publishing Group, London.

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., Van Hoesen, G.W. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. *American Journal of Physical Anthropology*, vol. 114, no. 3, pp. 224–241.

Shattuck, D.W., Prasad, G., Mirza, M., Narr, K.L., Toga, A.W. (2009) Online resource for validation of brain segmentation methods. *NeuroImage*, vol. 45, no. 2, pp. 431-9.

She R., Szakacs J. (2004). Hyperostosis frontalis interna: case report and review of literature. *Annals of Clinical & Laboratory Science*, vol. 34, no. 2, pp. 206-208.

Shen, L., Firpi, H.A., Saykin, A.J., West, J.D. (2009). Parametric surface modeling and registration for comparison of manual and automated segmentation of the hippocampus. *Hippocampus*, vol. 19, no.6, pp. 588-95.

Shen, Q., Zhao, W., Loewenstein, D.A., Potter, E., Greig, M.T., Raj, A., Barker, W., Potter, H., Duara, R. (2012). Comparing new templates and atlas-based segmentations in the volumetric analysis of brain magnetic resonance images for diagnosing Alzheimer's disease. *Alzheimer's & Dementia*, vol. 8, no. 5, pp. 399-406.

Shenkin S.D., Bastin M.E., Macgillivray T.J., Deary I.J., Starr J.M., Wardlaw J.M. (2009c). Birth parameters are associated with late-life white matter integrity in community-dwelling older people. *Stroke*, vol. 40, no. 4, pp.1225-8.

Shenkin S.D., Deary I.J., Starr J.M. (2009b). Birth parameters and cognitive ability in older age: a follow- up study of people born 1921-1926. *Gerontology*, vol. 55, no.1, pp. 92-8.

Shenkin, D.S., Rivers, C.S., Deary, I.J., Starr, J.M., Wardlaw, J.M. (2009a). Maximum (prior) brain size, not atrophy, correlates with cognition in community-dwelling older people: a cross-sectional neuroimaging study. *BMC Geriatrics*, 9, 12.

Shenton, M. E., Kikinis, R., McCarley, R. W., Saiviroonporn, P., Hokama, H.H., Robatino, A., Metcalf, D., Wible, C.G., Portas, C.M., Iosifescu, D.V., Donnino, R., Goldstein, J.M. & Jolesz, F.A. (1995). Harvard brain atlas: A teaching and visualisation tool. *Proceedings of the IEEE Biomedical Image Analysis*, pp. 10-17.

Siddique, H. [Online, 11/12/2013], http://www.theguardian.com/society/2013/dec/11/dementia-research-doubled-david-cameron-alzheimers-nhs. Date accessed 02/04/2014.

Singhal A, Cole TJ, Lucas A. (2001). Early nutrition in preterm infants and later blood pressure: two cohorts after randomised trials. *Lancet*, vol. 357, pp. 413-9.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. Personality and Individual Differences, vol.44, no. 3, pp. 535-551.

Sluimer, J.D., van der Flier, W.M., Karas, G.B., Fox, N.C., Scheltens, P., Barkhof, F., Vrenken, H., 2008. Whole-brain atrophy rate and cognitive decline: longitudinal MR study of memory clinic patients. *Radiology*, vol. 248, no.2, pp. 590-598.

Smith, S. M. (2010). Fast Robust Automated Brain Extraction, *Human Brain Mapping*, vol. 17, no. 3, pp. 143- 155.

Sowell, E.R., Peterson, B.S. Thompson, P.M., Welcome, S.E., Henkenius, A.L., Toga, A.W. (2003). Mapping cortical change across the human life span. *Nature*

Neuroscience, vol. 6, no.3, pp. 309-315.

Staff, R.T., Murray, A.D., Deary, I.J., Whalley, L.J. (2004). What provides cerebral reserve? *Brain*, vol. 127, no. 5, pp.1191-1199.
Staff R.T., Murray A.D., Deary I.J., Whalley L.J. (2006). Generality and specificity in cognitive aging: a volumetric brain analysis. *Neuroimage*, vol. 30, no. 4, pp. 1433-1440.

Starr, J.M., Leaper, S.A., Murray, A.D., Lemmon, H.A., Staff, R.T., Deary, I.J., Whalley, L.J. (2003). Brain white matter lesions detected by magnetic resonance imaging are associated with balance and gait speed. *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 74, no. 1, pp. 94-98.

Stein, J.L., Medland, S.E., Vasquez, A.A., Hibar, D.P., Senstad, R.E., Winkler, A.M., et al. (2012). Identification of common variants associated with human hippocampal and intracranial volumes. *Nature Genetics*, vol. 44, no. 5, pp. 552-561.

Stephan, B.C.M., Savva, G.M., Brayne, C., Bond, J., McKeith, I.G., Matthews, F.E., MRC CFAS. 2008. Optimizing Mild Cognitive Impairment for Discriminating Dementia Risk in the General Older Population. *The American Journal of Geriatric Psychiatry*, vol. 18, no. 8, pp. 662-673.

Stephani, C., Fernandez-Baca, Vaca. G., Maciunas, R., Koubeissi, M., Luders, H.O. (2011). Functional neuroanatomy of the insular lobe. *Brain structure and function*, vol. 216, no. 2, pp. 137-149.

Sullivan, E.V., Rosenbloom, M., Serventi, K.L., Pfefferbaum, A. (2004). Effects of age and sex on volumes of the thalamus, pons and cortex. *Neurobiology of Aging*, vol. 25, no. 2, pp. 185-192.

Sumar, I., Kosior, R. K., Frayne, R., & Federico, P. (2011). Hippocampal T2 abnormalities in healthy adults. *Epilepsy Research*, vol. 95, no. 3, pp. 273-276.

Szeszko, P.R., Bilder, R.M., Lencz, T., Pollack, S., Alvir, J.M., Ashtari, M., Wu, H., Lieberman, J.A. (1999) Investigation of frontal lobe subregions in first-episode schizophrenia. *Psychiatry Research*, vol. 90, no. 1, pp. 1-15.

Takahashi, R., Ishii, K., Kakigi, T., Yokoyama, K. (2011). Gender and age differences in normal adult human brain: voxel-based morphometric study. *Human brain mapping*, vol. 32, no. 7, pp.1050-1058.

Takao, H., Hayashi, N., Ohtomo, K. (2012). A longitudinal study of brain volume changes in normal aging. *European Journal of Radiology*, vol. 81, no.10, pp. 2801-2804.

Taki, Y., Kinomura, S., Sato, K., Goto, R., Wu, K., Kawashima, R., Fukuda, H., 2011. Correlation between gray/white matter volume and cognition in healthy elderly people. *Brain and Cognition*, vol. 75, no. 2, pp. 170-176.

Taki, Y., Thyreau, B., Kinomura, S., Sato, K., Goto, R., Kawashima, R., Fukuda, H. (2011). Correlations among brain gray matter volumes, age, gender and hemisphere in healthy individuals. *PLoS ONE*, vol. 6, no. 7, e22734.

Tamminga, C. A., Stan, A. D., & Wagner, A. D. (2010). The Hippocampal Formation in Schizophrenia. *American Journal of Psychiatry*, vol. 167, no. 10, pp. 1178-1193.

Tang, Y., Whitman, G.T., Lopez, I., Baloh, R.W. (2001). Brain volume changes on longitudinal magnetic resonance imaging in normal older people. *Journal of Neuroimaging*, vol. 11, no. 4, pp. 393-400.

Thoma, R.J., Monnig, M., Hanlon, F.M., Miller, G.A., Petropoulos, H., Mayer, A.R., Yeo, R., Euler, M., Lysne, P., Moses, S.N., Canive, J.M. (2009). Hippocampus Volume and Episodic Memory in Schizophrenia. *Journal of International Neuropsychological Society*, vol. 15, no. 2, pp. 182-195.

Thompson P.M., Hayashi K.M., de Zubicaray G.I., Janke A.L., Rose S.E., Semple J., Hong, M.S., Herman, D.H., Gravano, D., Doddrell, D.M., Toga, A.W. (2004). Mapping hippocampal and ventricular change in Alzheimer disease. *Neuroimage,* vol. 22, no. 4, pp. 1754–66.

Tisserand, D.J., Pruessner, J.C., Sanz Arigita, E.J., van Boxtel, M.P., Evans, A.C., Jolles, J., Uylings, H.B. (2002). Regional frontal cortical volumes decrease differentially in aging: an MRI study to compare volumetric approaches and voxel-based morphometry. *NeuroImage,* vol. 17, no. 2, pp. 657–669.

Tisserand, D.J., van Boxtel, M.P.J., Pruessner, P., Hofman, P., Evans, A.C., Jolles, J. (2004). A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. *Cerebral Cortex*, vol.14, no. 9, pp. 966-973.

Tobet, S., Knoll, J.G., Hartshon, C., Aurand, E., Sratton, M., Kuma, P., Searcy, B., McClellan, K. (2009). Brain sex differences and hormone influences: A moving experience? *Journal of Neuroendocrinology*, vol. 21, no. 4, pp. 387-392.

Valdes Hernandez, M.C., Ferguson, K.J., Chappell, F.M., Wardlaw, J.M. (2010). New multispectral MRI data fusion technique for white matter lesion segmentation: method and comparison with thresholding in FLAIR images. *European Radiology*, vol. 20, no. 7, pp.1684-1691.

Valdes Hernandez, M.C., Morris, Z., Dickie, D.A., Royle, N.A., Munoz Maniega, S., Aribisala, B.S., Bastin, M.E., Deary, I.J., Wardlaw, J.M. (2013). Close correlation between quantitative and qualitative assessments of white matter lesions. *Neuroepidemiology*, vol. 40, no. 1, pp. 13-22.

Vrenken, H., Geurts, J. J. G., Knol, D. L., van Dijk, L. N., Dattola, V., Jasperse, B., et al. (2006). Whole-brain T1 mapping in multiple sclerosis: Global changes of normal-appearing gray and white matter. *Radiology*, vol. 240, no. 3, pp. 811-820.

Vrenken, H., Pouwels, P. J. W., Ropele, S., Knol, D. L., Geurts, J. J. G., Polman, C.

H., et al. (2007). Magnetization transfer ratio measurement in multiple sclerosis normal-appearing brain tissue: limited differences with controls but relationships with clinical and MR measures of disease. *Multiple Sclerosis*, vol. 13, no. 6, pp. 708-716.

Vrenken, H., Rombouts, S., Pouwels, P. J. W., & Barkhof, F. (2006). Voxel-based analysis of quantitative T1 maps demonstrates that multiple sclerosis acts throughout the normal-appearing white matter. American Journal of Neuroradiology, 27(4), 868-874.

Wahlund, L.O., Barkhof, F., Fazekas, F., Bronge, L., Augustin, M., Sjögren, M., Wallin, A., Ader, H., Leys, D., Pantoni, L., Pasquier, F., Erkinjuntti, T., Scheltens, P. (2001). A new rating scale for age-related white matter changes applicable to MRI and CT. *Stroke*, vol. 32, no. 6, pp. 1318-1322.

Walhovd, K.B., Fjell, A.M., Reinvang, I., Lundervold, A., Dale, A.M., Eilertsen, D.E., Quinn, B.T., Salat, D., Makris, N., Fischl, B. (2005). Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiology of Aging*, vol. 26, no. 9, pp. 1261-1270.

Wang, H. L., Yuan, H. S., Shu, L., Xie, J. X., & Zhang, D. (2004). Prolongation of T-2 relaxation times of hippocampus and amygdala in Alzheimer's disease. *Neuroscience Letters*, vol. 363, no. 2, pp. 150-153.

Wang, J. L., Shaffer, M. L., Eslinger, P. J., Sun, X. Y., Weitekamp, C. W., Patel, M. M., et al. (2012). Maturational and Aging Effects on Human Brain Apparent Transverse Relaxation. *Plos One*, vol. 7, no. 2, 11.

Wardlaw, J.M., Bastin, M.E., Valdes Hernandez, M.C., Munoz Maniega, S., Royle, N.A., Morris, Z., Clayden, J.D., Sandeman, E.M., Eadie, E., Murray, C., Starr, J.M., Deary, I.J. (2011). Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. *International Journal of Stroke*, vol. 6, no. 6, pp. 547-559.

Watson, C. Andermann, F., Gloor, P., Jones-Gotman, M., Peters, T., Evans, A., Olivier, A., Melanson, D., Leroux, G. (1992). Anatomic basis of amygdaloid and hippocampal volume measurement by magnetic resonance imaging. *Neurology*, vol. 42, no. 9, pp. 1723-1750.

Wechsler D. WAIS-III^{UK} administration and scoring manual. 1998. London: Psychological Corporation.

Weiner, M.W., Veitch, D.P., Aisen, P.S., Beckett, L.A., Cairns, N.J., Green, R.C., Harvey, D., Jack, C.R., Jagust, W., Liu, E., Morris, J.C., Petersen, R.C., Saykin, A.J., Schmidt, M.E., Shaw, L., Shen, L., Siuciak, J.A., Soares, H., Toga, A.W., Trojanowski, J.Q., Alzheimer's Disease Neuroimaging Initiative. (2013). The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. *Alzheimer's & Dementia*, vol. 8, e111-e194.

Weiss, K.L., Stiving, S.O., Herderick, E.E., Cornhill, J.F., Chakeres, D.W. (1987). Hybrid color MR imaging display. *American Journal of Roentgenology*, vol. 149, no.

4, pp. 825-829.

West, R. (2000). In defense of the frontal lobe hypothesis of cognitive aging. *Journal of the International Neuropsychological Society*, vol. 6, no. 6, pp. 727-729.

Westman, E., Wahlund, L-O., Foy, C., Poppe, M., Cooper, A., Murphy, D., Spenger, C., Lovestone, S., Simmons, A. (2011). Magnetic resonance imaging and magnetic resonance spectroscopy for detection of early Alzheimer's disease. *Journal of Alzheimer's Disease*, 26, 307-319.

Whalley, H.C. and Wardlaw, J.M. (2001). Accuracy and reproducibility of simple cross-sectional linear and area measurements of brain structures and their comparison with volume measurements. *Neuroradiology*, vol. 43, no. 4, pp. 263-271.

Wheeler, S.M., Willoughby, K.A., McAndrews, M.P., Rovet, J.F. (2011). Hippocampal size and memory functioning in children and adolescents with congenital hypothyroidism. *Journal of Clinical Endocrinology Metabolism*, vol. 96, no. 9, pp. 1427-1434.

Whitwell, J.L., Crum, W.R., Watt, H.C., Fox, N.C. (2001). Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. *American journal of neuroradiology*, vol. 22, no. 8, pp. 1483-1489.

Wible, C.G., Shenton, M.E., Hokama, H., Kikinis, R., Jolesz, F.A., Metcalf, D., McCarley, R.W. (1995). Prefrontal cortex and schizophrenia. *Archives of General Psychiatry*, vol. 52, no. 4, pp. 279-288.

Wolf D.A., Falsetti A.B. (2001). Hyperostosis cranii ex vacuo in adults: a consequence of brain atrophy from diverse causes. *Journal of forensic science*, vol. 46, no. 2, pp. 370-373.

Wolf, H., Grunwald, M., Kruggel, F., Riedel-Heller, S.G., Angerhofer, S., Hojjatoleslami, A., Hensel, A., Arendt, T., Gertz, H.J. (2001). Hippocampal volume discriminates between normal cognition; questionable and mild dementia in the elderly. *Neurobiology of Aging*, vol. 22, no. 2, pp. 177-186.

Wolf, H., Kruggel, F., Hensel, A., Wahlund, L-O., Arendt, T., Gertz, H-J. (2003). The relationship between head size and intracranial volume in elderly subjects. *Brain Research*, vol. 973, no. 1, pp. 74-80.

Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D. and Alzheimer's Disease Neuroimaging Initiative. (2010). LEAP: learning embeddings for atlas propagation. *NeuroImage*, vol. 49, no. 2, pp.1316-25.

Woolard, A.A., Hexkers, S. (2012). Anatomical and functional correlates of human hippocampal volume asymmetry. *Psychiatry Research: Neuroimaging*, vol. 201, no. 1, pp. 48-53.

Xu, J., Kobayashi, S., Yamaguchi, S., Iijima, K., Okada, K., Yamashita, K. (2000). Gender effects on age-related changes in brain structure. *American journal of Neuroradiology*, vol. 21, no. 1, pp. 112-118.

Ystad, M.A., Lundervold, A.J., Wehling, E., Espeseth, T., Rootwelt, H., Westlye, L.T., Andersson, M., Adolfsdottir, S., Geitung, J.T., Fjell, A.M., Reinvang, I., Lundervold, A. (2009). Hippocampal volumes are important predictors for memory function in elderly women. *BMC Medical Imaging* 9, 17.

Zeigler, G., Dahnke, R., Jäncke, L., Yotter, R.A., May, A., Gaser, C. (2012). Brain structural trajectories over the adult lifespan. *Human Brain Mapping*, vol. 33, no. 10, pp. 2377-2389.

Zhou, S-Y., Suzuki, M., Hagino, H., Takahashi, T., Kawasaki, Y., Matsui, M., Seto, H., Kurachi, M. (2005). Volumetric analysis of sulci/gyri-defined in vivo frontal lobe regions in schizophrenia: Precentral gyrus, cingulate gyrus, and prefrontal region. *Psychiatry Research: Neuroimaging*, vol. 139, no. 2, pp. 127-139.

Ziegler, D.A., Piguet, O., Salat, D.H., Prince, K., Connally, E., Corkin, S. (2010). Cognition in healthy aging is related to regional white matter integrity, but not cortical thickness. *Neurobiology of Aging*, vol. 31, no. 11, pp. 1912-1926.

Zilles, K. and Amunts, K. Centenary of Brodmann's map – conception and fate. *Nature reviews: Neuroscience*, vol. 11, no. 2, pp. 139-145.