

In-vivo Analysis of Cortical Thickness using Magnetic Resonance Images

Jason Lerch

Department of Neurology and Neurosurgery

McGill University, Montreal, Canada

Monday, August 1, 2005

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

© Jason Lerch, 2005

Table of Contents

Table of Contents.....	2
List of Figures.....	7
List of Tables.....	8
Acknowledgements.....	9
Contributions of Authors.....	11
Other Related Publications.....	13
Abstract.....	16
Résumé.....	18
Original Contribution.....	21
1. Introduction.....	23
2. Background.....	28
2.1. Structure of the cerebral cortex.....	28
2.1.1. The cortex is composed of varying neuronal types.....	30
2.1.2. The cortex is horizontally structured in layers.....	32
2.1.3. The cortex is vertically organized into columns.....	32
2.1.4. The cortex is organized into cytoarchitectonic areas.....	34
2.1.5. Functional systems are hierarchically organized.....	36
2.1.6. The thickness of the cortex.....	38
2.2. The cortex in health and disease.....	43
2.2.1. The developing cortex.....	43
2.2.2. The ageing cortex.....	46
2.2.3. Degenerative diseases.....	47
Alzheimer's Disease.....	47
Huntington's and Parkinson's Disease.....	49
Multiple Sclerosis.....	50

2.2.4. The cortex in psychiatric disorders.....	50
2.3. Structure-function relationships.....	52
2.4. Principles of structural MRI.....	54
2.5. Image Processing.....	56
2.5.1. MNI Image Processing Pipeline.....	59
2.5.2. Other MRI based techniques.....	68
Manual or semi-automatic methods.....	70
Automated volumetric methods.....	71
Mixed mesh-volumetric methods.....	71
Mesh-based methods.....	73
2.5.3. Statistical Analysis.....	74
3. Validation Study.....	80
3.1. Preface.....	80
3.2. Abstract.....	83
3.3. Introduction.....	83
3.4. Methodology.....	87
3.4.1. Measuring Cortical Thickness.....	87
3.4.2. Statistical Analysis.....	92
3.4.3. Analyzing the Variance.....	93
3.4.4. Population Simulation.....	94
3.5. Results.....	96
3.5.1. Variability.....	96
Variability across different cortical thickness metrics.....	97
Variability is not uniform across the cortex.....	97
Variability declines with increased blurring–up to a point.....	100
Power calculations.....	100
3.5.2. Population Simulation.....	103
30 millimeters is the optimal kernel FWHM.....	104
tlink is the most sensitive method.....	105
Controlling for multiple comparisons.....	108
3.6. Discussion.....	109
3.6.1. Comparing the different metrics.....	110
3.6.2. Varying variability.....	112
3.6.3. Effects of blurring.....	112
3.6.4. Thresholding statistical maps.....	113
3.7. Conclusions.....	114
3.8. Acknowledgements.....	115

4. Alzheimer's Disease.....	116
4.1. Preface.....	116
4.2. Abstract.....	119
4.3. Materials and Methods.....	121
4.4. Results.....	127
4.4.1. Group Differences (Normals vs. AD).....	127
4.4.2. Regression of Cortical Thickness versus MMSE.....	133
4.4.3. Regression of Cortical Thickness versus progression..	134
4.5. Discussion.....	135
5. AD prediction.....	140
5.1. Preface.....	140
5.2. Abstract.....	143
5.3. Introduction.....	143
5.4. Methods.....	146
5.4.1. Participants.....	146
5.4.2. Image acquisition.....	147
5.4.3. Image processing.....	147
5.4.4. Discriminant analysis.....	150
5.5. Results.....	153
5.5.1. Controlling for multiple comparisons.....	153
5.5.2. Separating AD patients from normal elderly controls.	153
5.5.3. Comparison of discriminant techniques.....	157
5.6. Discussion.....	157
5.6.1. Discriminant ability of cortical thickness.....	157
5.6.2. On the use of multivariate discriminants.....	158
5.6.3. On the use of different types of discriminants.....	159
5.7. Conclusions.....	159
5.8. Acknowledgements.....	160
6. MACACC.....	161
6.1. Preface.....	161
6.2. Abstract.....	164
6.3. Introduction.....	165
6.4. Methods.....	168
6.4.1. Extraction of Morphometric Data.....	168
6.4.2. Statistical Techniques.....	169
MACACC methods - correlations.....	169
Assessing group differences.....	171

MACACC-slope.....	171
MACACC-variance.....	173
Multiple Comparisons.....	175
Using Residuals.....	176
6.5. Application to normal development.....	176
6.5.1. Correlation Strength.....	177
6.5.2. Seed region: BA 44.....	178
6.5.3. Developmental MACACC differences.....	179
6.5.4. Variance tests.....	181
6.5.5. IQ MACACC differences.....	182
6.6. Discussion.....	184
6.7. Acknowledgments.....	187
6.8. Appendix.....	188
7. Conclusions.....	189
7.1. Summary and Implications.....	189
7.2. Future Work.....	191
Bibliography.....	195

List of Figures

Figure 1.1: Cortical slice.....	25
Figure 1.2: Thickness publications by year.....	26
Figure 2.1: Example cortex.....	29
Figure 2.2: Cortical Lamination.....	33
Figure 2.3: Columnar pattern.....	34
Figure 2.4: Cytoarchitectonic Maps.....	35
Figure 2.5: Economo's Thickness Map.....	39
Figure 2.6: Thickness across evolution.....	42
Figure 2.7: Cortical ontogenesis.....	44
Figure 2.8: DLPFC Thickness vs Age.....	45
Figure 2.9: The ageing cortex.....	46
Figure 2.10: Disease specific maps.....	52
Figure 2.11: Example MRIs.....	55
Figure 2.12: Cortical thickness image processing pipeline.....	61
Figure 2.13: Image preprocessing.....	62
Figure 2.14: CSF skeletonization.....	66
Figure 2.15: CLASP.....	67
Figure 2.16: Surface coordinate system.....	75
Figure 3.1: Blurring kernel illustration.....	91
Figure 3.2: Induced thinning.....	95
Figure 3.3: Thickness variability.....	99
Figure 3.4: Std. across blurring kernels.....	100
Figure 3.5: Power of t-link.....	102
Figure 3.6: Effect of kernel on threshold.....	103
Figure 3.7: Effect of blurring kernel.....	104
Figure 3.8: Result of metrics.....	106

Figure 3.9: Sensitivity of metrics.....	107
Figure 3.10: Thresholding.....	109
Figure 4.1: Methods.....	123
Figure 4.2: AD results: back, front, bottom.....	128
Figure 4.3: AD results, left.....	130
Figure 4.4: AD results, right.....	131
Figure 4.5: AD results, PHG.....	132
Figure 4.6: MMSE results.....	134
Figure 4.7: AD progression.....	135
Figure 5.1: Methods.....	148
Figure 5.2: Anatomical segmentation.....	150
Figure 5.3: QDA results.....	155
Figure 5.4: HC vs IFG.....	156
Figure 6.1: Example cortical thickness map.....	169
Figure 6.2: Group differences:.....	172
Figure 6.3: MACACC-strength.....	177
Figure 6.4: MACACC of BA 44.....	179
Figure 6.5: MACACC age difference.....	181
Figure 6.6: MACACC-variance tests.....	182
Figure 6.7: IQ differences.....	183

List of Tables

Table 2.1: Neuron types.....	31
Table 2.2: Relative thicknesses of cortical layers.....	40
Table 2.3: Thickness publications.....	70
Table 3.1: Thickness metrics.....	89
Table 3.2: Epidemiology statistics.....	96
Table 3.3: Power analysis.....	97
Table 4.1: Demographics.....	121
Table 4.2: AD results.	129
Table 5.1: Results of QDA by cortical structure.....	154
Table 6.1: Population Characteristics.....	176
Table 6.2: Definition of terms.....	188

Acknowledgements

For his guidance, brilliant insight, support, encouragement and good humour that made this project not only possible but also an immensely enjoyable endeavour, I thank my supervisor Dr. Alan C. Evans.

To paraphrase Bernard of Chartres, I am like a dwarf standing on the shoulders of giants: this thesis would not have been possible without the advice, support and mentorship from the Brain Imaging Centre's wise men: Alex Zijdenbos, Louis Collins, Peter Neelin, Jens Pruessner, Andrew Janke, Steve Robbins and John Sled.

Further thanks must go to various superb scientists located both at McGill and elsewhere who through lengthy discussions have both enhanced this thesis and deepened my love of the wonders of neuroscience: Kate Watkins, Tomas Paus, Keith Worsley, Jean-Francois Mangin, Katrin Amunts, Hartmut Mohlberg, Karl Zilles, Robert Zatorre, Virginia Penhune, Michael Petrides, Noor Kabani, Veronique Bohbot, Francesco Tomaiuolo, Hiroshi Fukuda, Shigeo Kinomura, Kazunori Sato, Harald Hampel, Stefan Teipel, Jong-Min Lee, Bruce Pike, Alain Dagher, Andrea Bernasconi, Neda Bernasconi, Jay Giedd, Judith Rapoport, Philip Shaw, Dede Greenstein, Jonathan Blumenthal, Mike Weiner, Howard Chertkow, David MacDonald, and Heidi Johansen-Berg. I am further eternally grateful to my fellow students and co-workers who have helped me with this project over the years: Vivek Singh, Yasser Ad-Dab'bagh, June-sic Kim, Claude Lepage, Arnaud Charil, Scott Mackey, Bojana Stevanovic, Najmeh Khalili,

Penelope Kostopoulos, Mallar Chakravarty, Simon Duchesne, Marie-Helene Grosbras, Oleg Ivanov, Oliver Lyttelton, Jonathan Harlap, Denise Milovan, and Juergen Germann.

I also would like to thank the administrative staff of the MNI for lending a helping hand whenever needed: Zia Merchant, Michele L'Allier-Davies, Jennifer Chew, and Monique Lederman. And further humble appreciation for maintaining an excellent environment to work within goes towards the system administrators: Jean-Francois Malouin, Dale Einarson, and Sylvain Milot.

Lastly, a gracious thank-you to the K.M. Hunter/Canadian Institutes of Health Research for funding my doctoral studies.

Contributions of Authors

I am the first author of all four manuscripts included in this thesis and have performed all of the data analysis. This included all stages of image processing, quality control, statistical analysis, visualization, and interpretation of the data. Software development was undertaken where needed, particularly for the statistics and visualization. The contributions of the co-authors included supervision of the projects, data collection, and advice on the style and content of the manuscripts.

The following list summarizes the contributions of the co-authors situated at the Montreal Neurological Institute:

1. **Alan C. Evans.** Overall supervision of this PhD project.
2. **Jens C. Pruessner.** Headed the collaboration with the Ludwig Maximilian University, Munich, Germany, responsible for the Alzheimer's Disease data used in this thesis.
3. **Alex P. Zijdenbos.** Adjunct supervision of this PhD project.
4. **D. Louis Collins.** Aided in the design of the patient from control discrimination study described in chapter 5.
5. **Keith J. Worsley.** Provided advice on the statistical methods used in chapter 6.

A list of co-authors situated at other institutes who collaborated on various components of this thesis:

1. **Ludwig Maximilian University, Munich Germany.** This collaboration brought access to the Alzheimer's Disease data used in chapter 4 and 5, and is composed of the following two co-authors:
 - a. **Harald Hampel.**
 - b. **Stefan Teipel.**

2. **National Institute of Mental Health, Bethesda, Maryland.** This collaboration brought access to the normal brain development data used in chapter 6, and is composed of the following co-authors:
 - a. **W. Philip Shaw.**
 - b. **Dede Greenstein.**
 - c. **Rhoshel K. Lenroot.**
 - d. **Jay Giedd.**

Other Related Publications

Related publications can be divided into three components: 1) conference abstracts; 2) invited presentations, and 3) related co-authored articles.

Conference Abstracts

Ad-Dabbagh Y, Singh V, Robbins S, Lerch JP, Lyttelton O, Fombonne E, Evans AC. Native-Space Cortical Thickness Measurement and the Absence of Correlation to Cerebral Volume. *Human Brain Mapping* 2005; **11**

Hyde KL, Peretz I, Zatorre RJ, Griffiths TD, Lerch JP. The Brain Morphometry of Congenital Amusic Individuals. *Human Brain Mapping* 2005; **11**

Kinomura S, Lerch JP, Singh V, Sato K, Taki Y, Goto R, Inoue K, Uchida S, Ito H, Okada K, Tsuji I, Kawashima R, Evans AC, Fukuda H. Subcortical ischemic lesions detected with tissue classification technique from 3D magnetic resonance images: An aging effect. *Human Brain Mapping* 2005; **11**

Charil A, Lerch JP, Zijdenbos A, Dagher A, Evans AC. Cortical Atrophy in Multiple Sclerosis. *Human Brain Mapping* 2004;

Kinomura S, Lerch JP, Zijdenbos A, Sato K, Taki Y, Goto R, Inoue K, Uchida S, Tsuji I, Kawashima R, Evans AC, Fukuda H. Automatic detection of subcortical ischemic lesion from 3-D MRI data in the aged brain. *Human Brain Mapping* 2004;

Lerch JP, Pruessner J, Zijdenbos A, Collins DL, Teipel SJ, Hampel H, Evans AC. Predicting Alzheimer's using Cortical Thickness. *Conference on Alzheimer's Disease and Related Disorders* 2004;

Singh V, Lerch JP, Khalili N, Evans AC. Measuring Consistency of Brain Tissue Classification Methods. *Human Brain Mapping* 2004;

Tomaiuolo F, Worsley K, Lerch JP, Di Paola M, Carlesimo A, Bonanni R, Caltagirone C, Paus T. Long-term consequences of severe non-missile traumatic brain injury in white matter: An in vivo T1 weighted MRI computational analysis study. *Human Brain Mapping* 2004;

Ad-Dabbagh Y, Hulshoff Pol HE, Lerch JP, Schnack HG, Neeltje E, van Haren M, Fombonne E, Evans AC, Kahn RS. Case-control comparison of cortical thickness maps in schizophrenia: An anatomical MRI study. *Human Brain Mapping* 2003; **9**

Lerch JP, Pruessner J, Zijdenbos A, Evans AC. A validation study of cortical thickness analysis. *Human Brain Mapping* 2003;

Lerch JP, Zijdenbos A, Goto R, Taki Y, Sato K, Kawashima R, Fukuda H, Evans AC. Cortical Thinning in Aging. *Human Brain Mapping* 2003;

Lerch JP, Bernasconi N, Bernasconi A, Evans AC. Cortical Thickness in Left Temporal Lobe Epilepsy. *Human Brain Mapping* 2003;

Mohlberg H, Lerch JP, Amunts K, Evans AC, Zilles K. Probabilistic cytoarchitectonic maps transformed into MNI space. *Human Brain Mapping* 2003;

Lerch JP, Pruessner J, Zijdenbos A, Teipel SJ, Buerger K, Hampel H, Evans AC. Changes in Cortical Integrity in Alzheimer's Disease. *Conference on Alzheimer's Disease and Related Disorders* 2002;

Watkins K, Paus T, Mangin JF, Zijdenbos A, Collins DL, Lerch JP, Worsley K, Blumenthal J, Giedd J, Rapoport J. Maturation of white matter tracts during adolescence: a longitudinal MRI study. *Human Brain Mapping* 2001;

Invited Presentations

Lerch J. 2005. Structural connectivity through thickness correlations. *Human Brain Mapping*. Toronto, Ontario, Canada.

Lerch J. 2005. Predicting Alzheimer's Disease using Cortical Thickness. *International Conference on Prevention of Dementia: Early diagnosis and intervention*. Washington, DC, USA.

Lerch J. 2004. Cortical thickness from MRI: With a focus on Alzheimer's Disease. *Statistics Society of Canada*. Montreal, Quebec, Canada.

Lerch J. 2004. Cortical thickness analysis in MRI: Applications in studies of neurodevelopment and disease. *Brain Imaging Group Research Seminar*. Douglas Hospital, Montreal, Quebec, Canada.

Lerch J. 2003. Structural Brain Imaging from MRI. *Symposium on Advanced Technologies in Functional Biomedical Imaging*. Sendai, Japan.

Lerch J. 2003. Cortical Thickness Analysis: Methods and Applications. *NIH Neuroscience Interest Group*. National Institutes of Health, Bethesda, Maryland, USA.

Lerch J. 2002. Epidemiology and structural MRI. *Departments of Radiology and Epidemiology Research Seminar*. Erasmus Medical Centre, University of Rotterdam, The Netherlands.

Related co-authored articles

Worsley KJ, Taylor JE, Tomaiuolo F, Lerch J. Unified univariate and multivariate random field theory. *Neuroimage* 2004; **23 Suppl 1**:S189-95.

Worsley K, Chen J-I, Lerch J, Evanc AC. Comparing connectivity via thresholding correlations and SVD. *Philosophical Transactions of the Royal Society* 2005; **in Press**

Shaw WP, Lerch J, Greenstein D, Lenroot R, Evanc AC, Giedd J, Rapoport JL. Handedness and cortical development. 2005; **in preparation**

Charil A, Dagher A, Lerch J, Zijdenbos A, Worsley K, Evanc AC. Focal cortical atrophy in multiple sclerosis: relation to lesion load and disability. *Annals of Neurology* 2005; **submitted**

Shaw WP, Lerch J, Greenstein D, Sharp WS, Clasen LS, Evanc AC, Giedd J, Castellanos FX, Rapoport JL. Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention deficit/hyperactivity disorder. *Arch Gen Psychiatry* 2005; **under review**

Shaw WP, Lerch J, Greenstein D, Clasen LS, Lenroot R, Evans A, Rapoport JL, Giedd J. Intellectual ability and cortical development in children and adolescents. *Nature* 2005; **under review**

Ad-Dab'bagh Y, Hulshoff Pol HE, Lerch J, Robbins S, Schnack HG, van Haren NEM, Fombonne E, Evanc AC, Kahn RS. Case-control comparison of cortical thickness maps in schizophrenia: a vertex-based corticometry study. *Arch Gen Psychiatry* 2005; **in preparation**

Bobb AJ, Addington AM, Sidransky E, Gornick MC, Lerch JP, Greenstein DK, Clasen LS, Sharp WS, Inoff-Germain G, Wavrant-De Vrieze F, Arcos-Burgos M, Straub RE, Hardy JA, Castellanos FX, Rapoport JL. Support for association between ADHD and two candidate genes: NET1 and DRD1. *Am J Med Genet B Neuropsychiatr Genet* 2005;

Nagano-Saito A, Washimi Y, Arahata Y, Kachi T, Lerch JP, Evans AC, Dagher A, Ito K. Cerebral atrophy and its relation to cognitive impairment in Parkinson disease. *Neurology* 2005; **64**:224-229.

Kim JS, Singh V, Lee JK, Lerch J, Ad-Dab'bagh Y, Macdonald D, Lee JM, Kim SI, Evans AC. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage* 2005

Abstract

This dissertation describes a series of investigations designed to establish in-vivo analysis of cortical thickness using magnetic resonance images (MRI) as a viable technique in human brain imaging. Chapters 1 and 2 provide an introduction and overview of the relevant background. Chapters 3-6 take the form of manuscripts either published or submitted and describe (i) the validation of the methodology, (ii) its application in pathophysiology research, (iii) its application in clinical diagnosis and (iv) methodology and research application of the concept of "correlated" changes in cortical thickness. The final chapter summarizes the findings and considers the future development of this methodology.

The main contributions of this thesis, chapters 3-6, are summarized as follows:

Chapter 3: The highest precision in a population simulation was reached at a sensitivity of 0.93 with 100% specificity. Smoothing cortical thickness maps along the surface was established as clearly necessary, ideally in the 20-35mm blurring kernel size range. Different multiple comparisons correction techniques provided varying sensitivity to both false positives and false negatives, with Random Field Theory providing the best guarantee against false positives, False Discovery Rate the best optimization of true positives over false positives and false negatives. A power analysis found that, given two groups of 25 subjects each, a 0.6mm (15%) change in cortical thickness can be found.

Chapter 4: The pattern of cortical atrophy in Alzheimer's Disease found was in line with predictions based on the known advance of histopathology in the disease. The amount of thinning was quantified, reaching a maximum of 1.25mm in the parahippocampal gyrus. Correlations of cortical thickness with disease progression and Mini Mental State Exam scores showed increasing atrophy in the parahippocampal gyrus, lateral temporal and frontal lobes.

Chapter 5: Given mean thickness measures over the parahippocampal gyrus and one of six other structures, cortical thickness was found to be capable of perfectly reproducing the initial clinical diagnosis of probable Alzheimer's Disease. Maps showing the diagnostic accuracy of every single vertex were provided. Different discriminant techniques were compared and found to be roughly comparable in performance.

Chapter 6: Maps of anatomical correlations across the cerebral cortex were computed in a normal brain development population. The correlation map of Brodmann Area 44, the chosen seed region, bore striking resemblance to diffusion tensor fibre maps of the same area. This relationship tightened with age, showing a developmental pattern of cross-cortical correlations. It also changed with IQ, the high IQ sample having strong correlations with the anterior cingulate and lateral parietal lobes, something absent in the low IQ group. Lastly, a map of correlation strength of each vertex was shown, which featured the association cortices as the most correlated areas.

The findings shown in this thesis establish fully automated cortical thickness analysis from in-vivo MRI as a precise technique capable of providing quantitative measures of cortical morphology in both clinical and neurologically normal populations.

Résumé

Cette thèse décrit une série d'investigations visant à utiliser la résonance magnétique (IRM) comme technique viable en imagerie du cerveau humain pour établir le rôle de l'analyse in-vivo de l'épaisseur du cortex cérébral. Les deux premiers chapitres présentent une introduction des travaux, suivie d'un survol des connaissances de base en imagerie. Les chapitres 3 à 6 sont contient de manuscrits, soit déjà publiés ou soumis pour publication, ayant pour but de démontrer (i) la validation de la méthodologie, (ii) l'application de cette dernière en recherche en physiopathologie, (iii) l'usage de la méthodologie vis-à-vis le diagnostic clinique et (iv) l'application de la méthodologie dans des études de corrélation des variations de l'épaisseur du cortex cérébral. Le dernier chapitre résume les résultats obtenus et propose des améliorations à apporter dans le cadre de développements futurs.

Les chapitres 3 à 6 représentent la contribution principale de cette thèse:

Chapitre 3: Dans une simulation sur une population on a atteint une haute précision avec une sensibilité de 0.93 avec une spécificité de 100%. On a clairement établi la nécessité de lisser les données de l'épaisseur de la surface corticale, idéalement avec un noyau de 10 à 35mm. Différentes techniques de correction pour les comparaisons multiples ont démontré des sensibilités variées quant aux faux positifs et aux faux négatifs. La théorie du champ aléatoire (Random Field Theory) a démontré la meilleure assurance contre les faux positifs et le taux de fausses

découvertes (False Discovery Rate) la meilleure optimisation des vrais positifs vis-à-vis les faux positifs et faux négatifs. Une analyse de puissance posée sur deux groupes de 25 sujets a démontré qu'un changement de 0.6 mm (15%) de l'épaisseur du cortex peut être perceptible.

Chapitre 4: Le profil d'atrophie corticale détecté chez des sujets atteints de la maladie d'Alzheimer correspond aux prédictions basées sur l'évolution répertoriée de l'histopathologie de la maladie. Le degré de l'amincissement a été quantifié, atteignant un maximum de 1.25 mm dans le gyrus para-hippocampal. Des corrélations de l'épaisseur du cortex avec l'évolution de la maladie et les résultats du MMSE (Mini Mental State Exam) démontrent une atrophie croissante dans le gyrus para-hippocampal et les lobes latéro-temporaux et frontaux.

Chapitre 5: En utilisant la distribution moyenne de l'épaisseur du cortex au niveau du gyrus para-hippocampal et celle d'une deuxième structure parmi six autres, on a trouvé que l'épaisseur du cortex peut correctement reproduire le diagnostic clinique initial dans les cas probables de maladie d'Alzheimer. On a inclus des données montrant l'exactitude de chaque point cortical à prédire dans le contexte du diagnostic de la maladie de l'Alzheimer. Différentes techniques d'analyse discriminante ont été comparées, et ils ont indiqué des performances équivalentes.

Chapitre 6: Des corrélations anatomiques du cortex cérébral ont été calculés à partir d'une population de sujets normaux. Les corrélations de la région d'intérêt, l'aire 44 de Brodmann, démontrent une ressemblance frappante avec le réseau des fibres de tenseur de diffusion de la même région. Cette corrélation est accentuée avec l'âge, et est associée avec des corrélations inter-corticales suivant le développement adulte. Les corrélations varient aussi avec le quotient intellectuel: l'échantillon de sujets avec un quotient intellectuel élevé a démontré une corrélation forte entre

le cingulum antérieur et les lobes latéro-pariétaux. Cette corrélation est absente chez un groupe avec un quotient intellectuel faible. Finalement, on a établi des données sur l'indices de corrélations pour tous les points du cerveau, indiquant que les régions du cerveau d'association sont les plus fortement corrélées.

Les résultats de cette thèse démontrent que l'analyse automatisée de l'épaisseur du cortex à partir d'imagerie par résonance magnétique in-vivo est une méthode précise capable de fournir des mesures quantitatives sur la morphologie corticale des cas cliniques ainsi que des populations neurologiquement normales.

Original Contribution

The original contributions to this thesis are:

1. Creation of a population simulation based validation framework for cortical thickness analysis.
2. Finding that the t_{link} method is the most precise of six cortical thickness metrics tested.
3. Determining that blurring along the cortical sheet is necessary for adequate precision, and that the optimal blurring kernel is around 30-35mm within the simulation framework used.
4. Ascertaining the relative properties of multiple comparison correction techniques and finding that the False Discovery Rate is optimal at retaining true positives, Random Field Theory at eliminating false positives.
5. Describing the distribution of cortical atrophy in Alzheimer's Disease with its focus on the medial temporal lobes.
6. Showing correlated cortical thickness loss with decline in MMSE scores in the left lateral temporal lobes as well as the medial temporal lobes.
7. Finding increased atrophy with disease duration in Alzheimer's Disease in the frontal and temporal lobes.
8. Showing that cortical thickness can accurately separate

Alzheimer's Disease patients from normal elderly controls.

9. Describing the distribution of AD patient classification across the cortex and finding that the parahippocampal gyrus along with one of six other structures provides the best discrimination.
10. Showing that the cross-cortical correlations of Brodmann Area 44 bear striking similarity to diffusion tensor maps of the same region.
11. Describing maturation changes in BA 44 correlations with the lateral temporal lobes.
12. Finding that a sub-population with high IQs had greater correlations between BA 44 and the cingulate and lateral parietal lobes than children with low IQs.
13. Creating a pipeline environment to run the image processing in a rapid and repeatable fashion across multiple datasets.
14. Creating a statistical toolbox for running linear models, mixed models, discriminant functions, and cross-cortical correlations for cortical data.
15. Creating visualization tools to allow for interactive analysis of results of cortical thickness statistics.

Chapter 1: Introduction

One is allowed to expect answers to a series of questions of utmost importance through the study of the structure of the cerebral cortex: questions relating the anatomical underpinnings of physiological events to psychological phenomena and the correlation of specific sensory experiences with the structure of the brain. Within these questions one finds the beginnings of the ultimate problem, the problem of the soul, while also addressing general questions of comparative anatomy and specific queries on the relationship between psychopathology and pathological anatomy.

Konstantin von Economo and Georg Koskinas, 1925.

The quote above, from one of the giants of neuro-anatomy, illustrates the fascination with the structure of the cerebral cortex that underpins this thesis. Early examples of traumatic injury to the brain illustrated how profoundly structure and function are related; witness the case of Phineas Gage, who recovered fully from a metal bolt through his frontal lobe but returned a changed person, unable to plan for the future or follow social rules he had previously effortlessly operated within (Magoun et al., 1998; Kandel et al., 2000). Or Pierre Paul Broca's patient who could understand language but not speak, whilst still being able to sing or utter isolated

words. He suffered from damage to a precise area of the left frontal lobes, now known as Broca's area (Magoun et al., 1998).

The ability to understand the structure of the cortex and infer its relationship to function has always been limited by technology. The case of Phineas Gage and Broca's work mentioned above, for example, were dependent on accidents injuring specific parts of the brain. Precise investigations of cortical anatomy, while owing much to the work of early anatomists such as Willis, Sherrington, and Gall, began at the turn of the century with the research of Campbell, Brodmann, von Economo, Flechsig, and Ramon y Cajal, among others. The underpinnings were stained slices of post-mortem tissue allowing inferences about the micro-structure of the brain (Magoun et al., 1998). This research continues to this day (Zilles et al., 1988; Zilles, 1992; Sexton et al., 1994; Morris et al., 2000), shedding new light on, for example, cerebral asymmetries (Geschwind and Levitsky, 1968), cortical development (Rakic, 1974; Rakic, 1976), and the variability of micro-structure within populations (Amunts et al., 1999; Zilles et al., 2004).

Detailed in-vivo studies of human cortical anatomy were initially limited to traumatic injuries or, beginning with Wilder Penfield, stimulations during surgery. The advent of Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET) changed the field considerably. MRI in particular, with its excellent soft tissue contrast capabilities, allowed in-vivo neuroanatomy to flourish.

Higher resolution images furthermore brought image processing and computer science to bear on neuroanatomical questions. Manual segmentation of brain structures is now complemented by automated voxel-wise tests of tissue density (Wright et al., 1995; Paus et al., 1999; Ashburner and Friston, 2000), segmentation of cortical and sub-cortical

structures (Collins et al., 1995; Fischl et al., 2002; Mangin et al., 2004), and detection of boundary shifts based on deformations between acquisitions or from individual subjects to a common template (Fox et al., 2000; Resnick et al., 2003; Liu et al., 2004).

The thickness of the cerebral cortex has been investigated using post-mortem preparations in multiple types of mammals. One of the earliest studies was that of von Economo, providing detailed maps of not just cortical thickness but also the thickness of the individual cortical layers (Economo and Koskinas, 1925). An example slice showing this laminar pattern of the cortex can be seen in figure 1.1.

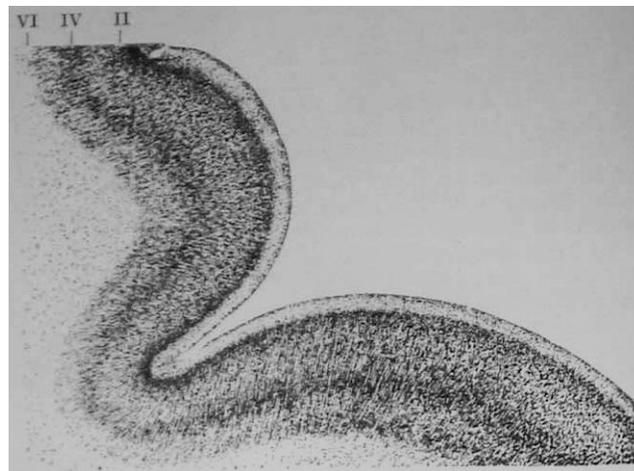


Figure 1.1: An example histology slice showing the lamination pattern of the cortex. Taken from (Economo and Koskinas, 1925).

The main processing unit of the cortex is the column, running from the white/grey matter junction to the pial surface (Buxhoeveden et al., 1996; Mountcastle, 1998). Changes in the length of the column through pathological processes such as neural or neuropil loss results in a concurrent loss of cortical thickness, and a likely corresponding alteration of function. However, direct study of cortical thickness in multiple normal or pathological population groupings had to wait for the advent of in-vivo

neuroanatomy

MR imaging made in-vivo and repeated studies of cortical thickness possible. Here, for the first time, came the ability to infer - in a quantitative and biologically meaningful fashion - a measure of structure at every point of the cortex in living humans. Extracting the boundaries of the cortex is a difficult task due to the tightly folded nature of the human brain and the limited resolution of MR. At the start of this thesis project only a small set of publications on thickness measurements from MRI existed, as shown in figure 1.2 (and detailed in table 2.3).

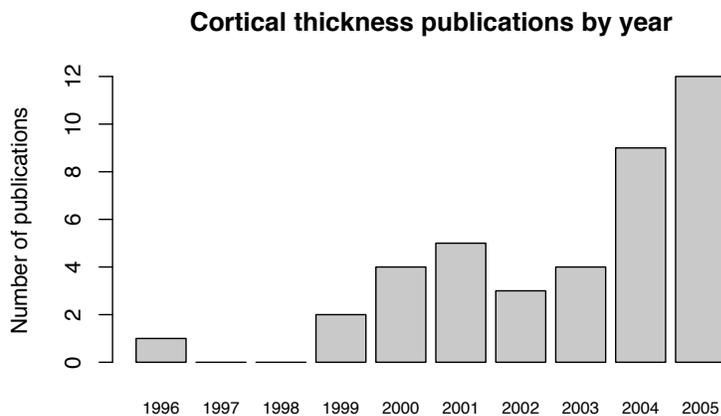


Figure 1.2: The number of cortical thickness publications by year, as revealed by the a PubMed search on The PubMed search used to generate this list was "cortical thickness" AND MRI NOT bone NOT mice NOT canine NOT rat NOT kidney. Only references in English language journals were used. The number for the year 2005 is estimated based on the 5 publications that existed by May of that year.

The overall aim of this thesis was thus to establish cortical thickness measurements from MRI as a viable technique in human brain imaging and to, in the process, further our understanding of the human cortex in health and disease. In particular, the goals of this thesis were to:

1. Validate cortical thickness measurements using simulation studies.
2. Apply the cortical thickness algorithm to a clinical population where cortical pathology is implicated.
3. Assess the potential utility of cortical thickness analysis in a clinical setting.
4. Investigate the potential of using the correlated thickness between different cortical areas as a research tool.

The thesis is organized into four separate papers preceded by a background section. The literature review in the background section sets the scene in terms of neuroanatomy as well as providing an overview of the image processing methodology used throughout this thesis and elsewhere. The first paper, published in *NeuroImage*, contains a validation study designed to assess the precision of cortical thickness analysis from MRI as well as to find optimal parameters for several settings in the image processing pipeline. The second article, published in *Cerebral Cortex*, applies cortical thickness analysis to an Alzheimer's Disease population, the goal being to quantify cortical atrophy within that disease. The third article, under review in *Neurobiology of Aging*, stays within the same population sample but addresses the question of how accurately cortical thickness can match clinical diagnosis. The last article, under review in *NeuroImage*, moves to a normal brain development population and attempts to measure how the thickness of different cortical areas are related in a cross-sectional study. Here too the question of whether the cross-cortical relationships change with different population grouping is addressed.

Chapter 2: Background

This chapter consists of five sections, beginning with a description of the structure of the cortex with a particular focus on cortical thickness. The emphasis here will be on the anatomical knowledge gained about the cortex over the last hundred years. The changing nature of the cortex in health and disease is then discussed, with particular attention paid to normal development, ageing, Alzheimer's disease and Schizophrenia. A brief overview of the principles of structural magnetic resonance imaging is then provided, followed by a detailed discussion of the image processing methods used to extract cortical thickness and analyse thickness data.

2.1. Structure of the cerebral cortex

The cerebral cortex is the highly folded sheet of grey matter encasing the brain, home to most higher cognitive functions. Its thickness varies significantly by cortical region, changes across the human life-span, and is affected in multiple disease and disorders.

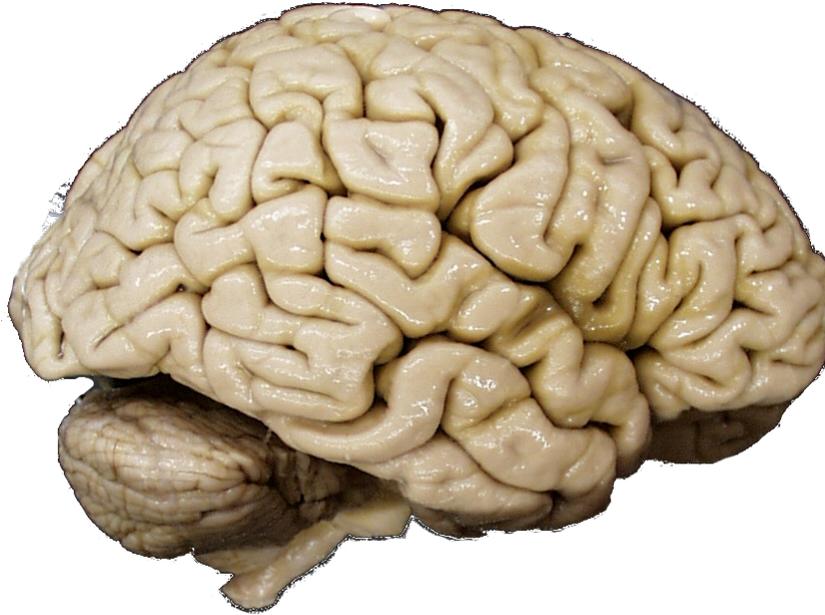


Figure 2.1: A picture of the outer cortical surface, cerebellum, and parts of the brain-stem of a human post-mortem brain.

There are five organizing principles behind the cerebral cortex. They are listed below and will be addressed in turn:

1. The cortex is composed of varying neuronal types.
2. The cortex is horizontally structured in layers.
3. The cortex is vertically organized into columns.
4. The cortex is organized into cytoarchitectonically unique regions.
5. Functional systems in the cortex are hierarchically organized, whereby cortical regions interact, either through direct connections or via sub-cortical mediation, to achieve higher-order mechanisms of sensori-motor and cognitive processing.

The fundamental concern of this thesis is the thickness of the cerebral cortex. The final sub-section will thus look at cortical thickness in more detail, with a special focus on how the five elements listed above affect it.

It should be noted that, due to the emphasis on structure, a detailed study of receptor types and chemical properties of the cortex are absent in this overview. While these are of course the primary mechanisms of neuronal action, the consequences of these mechanisms can be observed in the permanent structural changes which are induced by their action.

2.1.1. The cortex is composed of varying neuronal types

There are two main categories of neurons in the cerebral cortex: projection neurons and local inter-neurons (Kandel et al., 2000). Cortical neurons can be further classified into different types based on several criteria, such as cell form, usual position within the cortical laminae, axonal targets, type of synaptic terminals, direction of trans-synaptic action, and type of transmitter molecules. The seven main types of neurons are outlined in table 2.1.

Two main types of synapses occur within the neocortex. The first are asymmetrical, excitatory synapses which make up about 75-80% of the total. The remaining synapses are symmetric and inhibitory (Mountcastle, 1998).

Spiny pyramidal cells make up 70-80% of cortical neurons in mammals. They are located in different densities in all cortical layers except for layer I. Pyramidal cells in different layers have different targets: layer II targets ipsilateral cortical areas; layer III contralateral cortical areas; layer IV generally has no extrinsic targets (with an exception in the visual cortex); layer V targets spinal cord, pons, medulla, midbrain, thalamic nuclei, and basal ganglia; layer VI principally targets the thalamus and the claustrum. Pyramidal cells receive inputs from intrinsic cortical neurons, collaterals of other pyramidal cells, axonal terminals of pyramidal cells in other cortical areas, thalamic nuclei, brainstem monoaminergic systems, and from the

claustrum (Mountcastle, 1998). Non-pyramidal excitatory neurons make up another 2-3% of the cortex (Mountcastle, 1998). The remaining 20% of neurons in the cortex are non-pyramidal inhibitory neurons whose connections remain intrinsic to the cortex. Their primary transmitter is gamma-aminobutyric acid (GABA). Many of them co-localize neuropeptides (Mountcastle, 1998).

Cell Type	Transmitter	Terminal Type	Targets
Spiny pyramidal	Glutamate	Asymmetric Excitatory	Extracortical structures; Intrinsic
Spiny non-pyramidal	Glutamate	Asymmetric Excitatory	Intrinsic
Basket Cells	GABA	Symmetric Inhibitory	Intrinsic
Double bouquet cells	GABA	Symmetric Inhibitory	Intrinsic
Chandelier cells	GABA	Symmetric Inhibitory	Intrinsic
Peptide cells	GABA	Symmetric Inhibitory	Intrinsic
Neurogliaform	GABA	Symmetric Inhibitory	Intrinsic

Table 2.1: Neuron types of the mammalian neocortex. Adapted from (Mountcastle, 1998).

Neuron structure varies across the cortex, identified by three features: cell size, branching pattern, and number/distribution of inputs. These regional variation in structure underlie fundamental differences in cortical circuitry, as illustrated by the example of visual processing in the cortex discussed

later in this chapter (Elston, 2003).

2.1.2. The cortex is horizontally structured in layers

The neocortex is organized into six layers (Brodmann, 1909; Economo and Koskinas, 1925). Layers are differentiated by cell composition and density. Layer I, the outermost layer, is called the *molecular or plexiform layer* and is generally acellular, being occupied by dendrites of cells located deeper in the cortex. Layer II, the *external granule cell layer*, contains tightly packed small granule cells. Layer III, the *external pyramidal cell layer*, is comprised of pyramidal cells which increase in size towards the depth of the cortex. Layer IV, the *internal granule cell layer*, is similar to layer II, also mostly containing small granule cells. Layer V, the *internal pyramidal cell layer*, is similar to layer III, containing large pyramidal cells. Finally, layer VI, the *polymorphic, fusiform or multiform layer*, contains spindleform cells and gradually blends into the white matter. The cortical layers are illustrated in figure 2.2.

The layering of neurons provides means of organizing inputs and outputs of the cortex. Layer IVC α of the visual cortex, for example, receives input from the magnocellular layers of the Lateral Geniculate Nucleus of the Thalamus, sends projection interneurons to layer IVB, from there to layers II and III which in turn project to extrastriatal areas (Kandel et al., 2000).

2.1.3. The cortex is vertically organized into columns

Along with the horizontal laminar pattern cortical neurons are also vertically or radially structured into columns. Columnar organization varies throughout the cortex, being prominent in the parietal lobes, less visible in parts of the frontal lobe (Economo and Koskinas, 1925). Neurons within a particular column have similar response properties, likely being part of the

same local processing network.

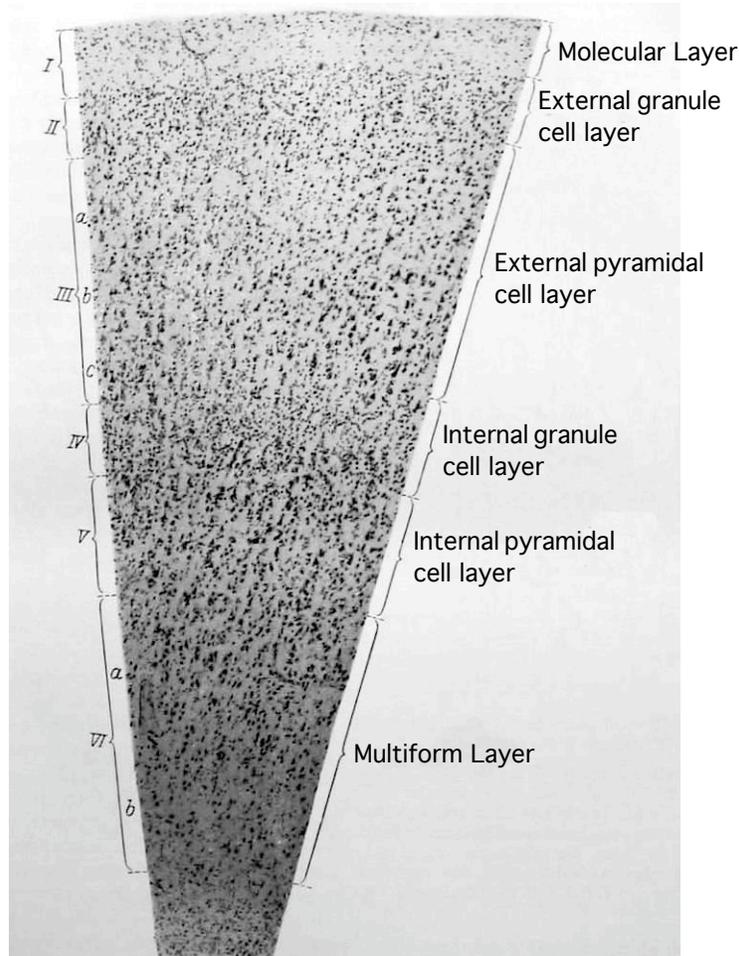


Figure 2.2: The six layers of the cortex illustrated on a single silver stained slice. Adapted from (Economo and Koskinas, 1925).

Evidence for the functional organisation comes from microelectrode penetrations into the cortex along with transection and nerve regeneration studies. The primary visual cortex responds selectively to differentially oriented lines. Electrode penetration studies of the visual cortex found that perpendicular penetrations showed constant responses to a stimulus of a single orientation, whereas penetrations made nearly parallel to the surface found a consistent change in sensitivity to differently

oriented stimuli (Obermayer and Blasdel, 1993; Mountcastle, 1998). Figure 2.3 provides a schematic illustration of these results.

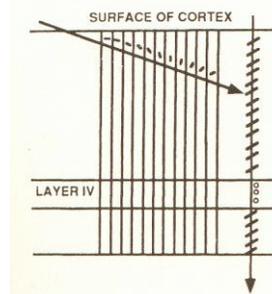


Figure 2.3: This figure shows orientation sensitivity of columns in the striate cortex. Vertical penetrations (rightmost arrow) show a single orientation, whereas parallel penetrations (diagonal arrow) show multiple orientations. The small lines above the diagonal penetration show the orientation of the line in the visual field which that particular column responds to. From (Mountcastle, 1998).

Similarly, studies in the somatic sensory cortex found columns coding for both modality (i.e. α fibers from Meissner receptors) as well as location (point on finger). Evidence for columnar existence in the sensory cortex includes nerve regeneration experiments in monkeys. It was found that neurons in adjacent columns are related to adjoining and overlapping peripheral receptive fields. Section and resuture of the contralateral medial nerve caused the regenerating nerve fibres to be misdirected, revealing cortical columns separated by 50-60 μm (Mountcastle, 1998).

2.1.4. The cortex is organized into cytoarchitectonic areas

Differential composition of neurons, along with varying lamination, subdivides the cortex into cytoarchitectonic areas, as mapped in figure 2.4.

An example to illustrate the difference between two cytoarchitectonic areas is that of Brodmann Area (BA) 44 and 45, both part of Broca's area and implicated in language function. In BA 44 layer IV is dysgranular and

barely recognizable, often invaded by pyramidal cells from layers III and V. BA 45 features a pronounced layer IV, though still less clear than its rostrally adjoining areas 46 and 10 (Economo and Koskinas, 1925; Amunts et al., 1999).

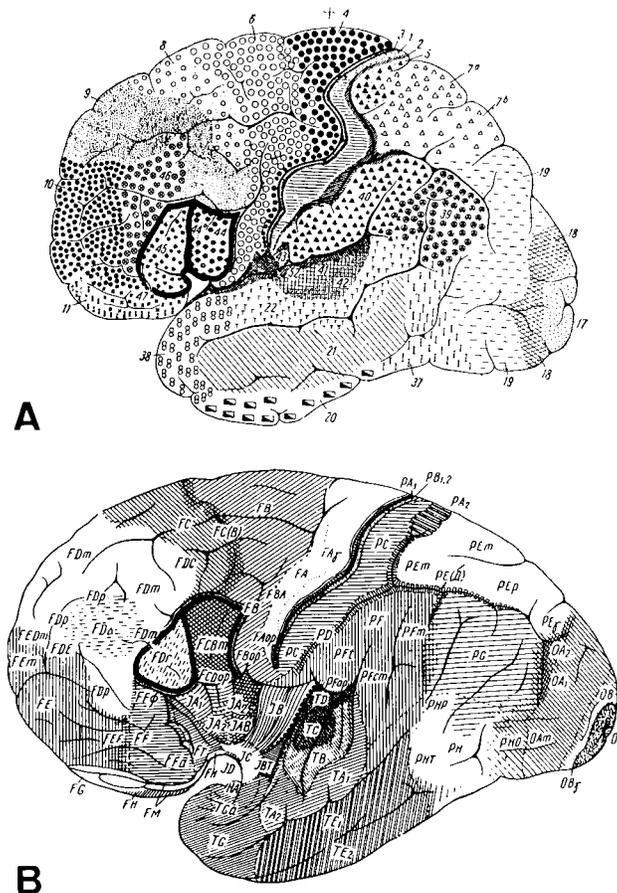


Figure 2.4: Two cytoarchitectonic maps dividing the cortex into morphologically similar areas are shown here. Part A shows the map from Brodmann, part B from Economo. Reproduced from (Amunts et al., 1999). The boundaries for areas 44 and 45, discussed in more detail below, are in bold.

Cytoarchitectonic areas encompass the functional subdivision of the cortex. To return to the previous example, Broca's area is composed of BA 44 and 45, and is involved in language production (Jurgens et al., 1982;

Zatorre et al., 1992; Chertkow and Murtha, 1997; Amunts et al., 2004; Dronkers et al., 2004; Fiebach et al., 2005). The two areas perform slightly different roles, area 45 is more involved with semantic processing, area 44 with higher level language production (Amunts et al., 2004). The cellular, laminar, and columnar structures that are reflected in cytoarchitectonic boundaries thus clearly contribute to the functional specialization of different cortical areas.

2.1.5. Functional systems are hierarchically organized

Each cytoarchitectonic area is responsible for a set of functions; entire functional modalities, such as language or vision, need multiple cortical areas operating within a connected network. The classic example of such a cortical network comes from the vision literature. Two processing streams, the "what" and "where" streams, are thought to exist: the dorsal pathway, running from the primary visual cortex to the parietal lobes, is responsible for motion, depth, and spatial information. The ventral pathway, also originating in the primary visual cortex in the occipital lobes and extending to the temporal lobes, analyses form and colour (Kandel et al., 2000).

It is worth examining cortical processing of visual information in some more detail as it provides an excellent overview of some general properties of the cortex. Input from the retina is passed via the optic chiasm to the lateral geniculate nucleus (LGN). Information at this point is already separated into the parvocellular (P) pathway, responding to colour, and the magnocellular (M) pathway, which processes luminance (Kandel et al., 2000).

Output from the LGN is sent to the primary visual (striate) cortex - BA 17. Afferent fibres from P cells terminate on layer IV β , from M cells on layer

IVC α . Both M and P cells also send collaterals to layer VI, and intralaminar regions of the LGN (neither M nor P) terminate on layers II and III. Neurons in layer IVC of the cortex project to layer IVB or layers II and III, which in turn project to the pyramidal cells of layer V. Layer V neurons in turn project to layer VI as well as back to layers II and III in a feedback loop (Kandel et al., 2000).

The intra-cortical connections illustrated above form the basic processing unit of the neocortex, the cortical column. Within the primary visual cortex, cortical columns code primarily for the orientation of the visual input, the eye from whence the input came (ocular dominance), as well as the colour. Every part of the visual field is represented by multiple columns ultimately coding for all possible orientations, colours, and ocular inputs; the sum of these representations is often referred to as a *hypercolumn* (Kandel et al., 2000).

Layers II and III of area 17 project to extra-striate cortical areas. As alluded to earlier, the ventral visual pathway, running from the primary visual cortex to the inferior temporal cortices, is responsible for form and colour - the details of objects in the visual field. The dorsal visual pathway, also originating in the striate cortex and progressing to the parietal lobes, analyses the location and movement. Neurons in the higher processing centres of either pathway have larger receptive fields than those in the lower cortical areas. Moreover, they respond to increasingly complex stimuli. Neurons in the inferior temporal cortex, for example, respond to a given shape at any position in the visual field, whereas neurons in the primary visual cortex only respond to a particular edge in a precise position within the visual field. Higher visual processing centres are also more dependent on selective attention than cortical areas lower in the pathway (Kandel et al., 2000).

The structure of pyramidal neurons changes with cytoarchitectonic area, increasing their number of spines in areas subserving higher order visual processing tasks. This increase in spines raises the number of putative excitatory inputs along the dendrites. The pattern of connectivity of individual neurons changes as well, with higher order areas featuring neurons capable of sampling from larger areas of the visual field (Elston, 2003).

The example of visual processing in the cortex highlights the significance of the structural elements of the cortex discussed above to its function. Laminar structure underlies local input processing and output redirection; cytoarchitectonic boundaries subserve different functions within the data processing network, and the structure of individual neurons is important for their participation in different tasks within the network.

2.1.6. The thickness of the cortex

The thickness of the cortex is defined by the distance between the inner and outer extents of cortical grey matter, i.e. the grey-white intersection as the inside boundary and the pial surface as the outside boundary. The earliest detailed study of the thickness of the cortex was performed by Konstantin von Economo and Georg Koskinas (Economo and Koskinas, 1925) in 1925, and is still used as the classic reference in the field. Using human brains fixed in formalin and embedded in paraffin they found the thickness of the cortex to vary between 1.5 and 4.5 millimetres.

As seen in figure 2.5, cortical thickness varies significantly by region, being thinnest on the post-central gyrus and the occipital lobes, thickest in the anterior temporal lobes and the pre-central gyrus.



Figure 2.5: The 1925 thickness map by von Economo and Koskinas. Different shading represents changes in cortical thickness, black being highest at 4.5mm, white lowest at less than 2mm. Taken from (Economo and Koskinas, 1925).

Economo and Koskinas also found that cortical thickness varies within cortical folds, being thickest on the crowns of gyri and thinnest in the fundi (Economo and Koskinas, 1925), a finding reproduced in (Fischl and Dale, 2000). This variation is consistent throughout the cortex. The variation in thickness from crown to fundus is not, however, reflected by an equal

thinning across all layers. Layer I actually increases in thickness in the fundus, as does layer II. Layer III declines slightly, layers V and VI considerably. Layer IV remains relatively constant across cortical folds (Economo and Koskinas, 1925).

The six layers of the cortex are not equally thick. As can be expected from the discription of cytoarchitectonic boundaries offered in section 2.1.4, the relative thicknesses of each layer vary with these boundaries. Example proportions for two areas are listed in table 2.2.

Layer	Mean	Central Sulcus	Calcarine Sulcus
I	9%	6%	10%
II	7%	0%	7%
III	33%	43%	14%
IV	9%	0%	42%
V	20%	23%	14%
VI	22%	29%	13%
Total	100%	100%	100%

Table 2.2: illustrated from the crowns of the central sulcus, calcarine sulcus, as well as the mean proportion across the entire cortex. Taken from (Economo and Koskinas, 1925).

One view is that the number of neurons remains steady across the cerebral cortex, even as the thickness varies (Rockel et al., 1980; Mountcastle, 1998). Even more startling is the suggestion that this is true across different mammalian species. For example, the number of neurons within a $25 \times 35 \mu\text{m}$ patch in the motor cortex is 73.4 ± 4.5 in the mouse,

72.7 ± 3.9 in the rat, 69.8 ± 5.1 in the cat, 74.1 ± 6.3 in the monkey, and 68.7 ± 6.5 in the human (Rockel et al., 1980; Mountcastle, 1998). The only exception appears to be the visual cortex in primates, featuring neuron counts around 174-180, whereas in the mouse the count is 75, 72 in the rat, and 74 in the cat (Mountcastle, 1998). The visual cortex moreover is one of the thinnest cortices in primates (Economo and Koskinas, 1925). This theory of uniformity of neuron numbers within the cortical column has been challenged (Skoglund et al., 1996). Even in this data, however, there is no relationship between thickness of the cortex and neuron numbers in the vertical unit of the cortex (Skoglund et al., 1996). While not as well studied, it is thought that overall densities of glial and neuronal cells are about the same, though there is a greater variation across different cortical areas for glial cells (Mountcastle, 1998). Differences in the thickness of the cortex - in non-pathological conditions - are thus not due to neuron numbers, but rather appears to change with the density of the cell packing in various layers along with different innervation patterns, likely supported by differential glial numbers.

The thickness of the cortex, as shown in figure 2.6, is surprisingly constant across the mammalian order varying by a ratio of 4-1 between mouse and human. The surface area of the cortex, on the other hand, varies dramatically - by 3 orders of magnitude (Mountcastle, 1998). Taken together with the fact that neuron numbers are relatively constant within a patch of cortex it would suggest that the vertical organization of the cortex is relatively conserved throughout evolution.

Thickness measures, running perpendicular to the cortical layers, approximate the cortical column, though, unlike columns, they extend across layer I to the pial surface. Changes in the thickness of the cortex thus reflect a change in the length of the underlying columns.

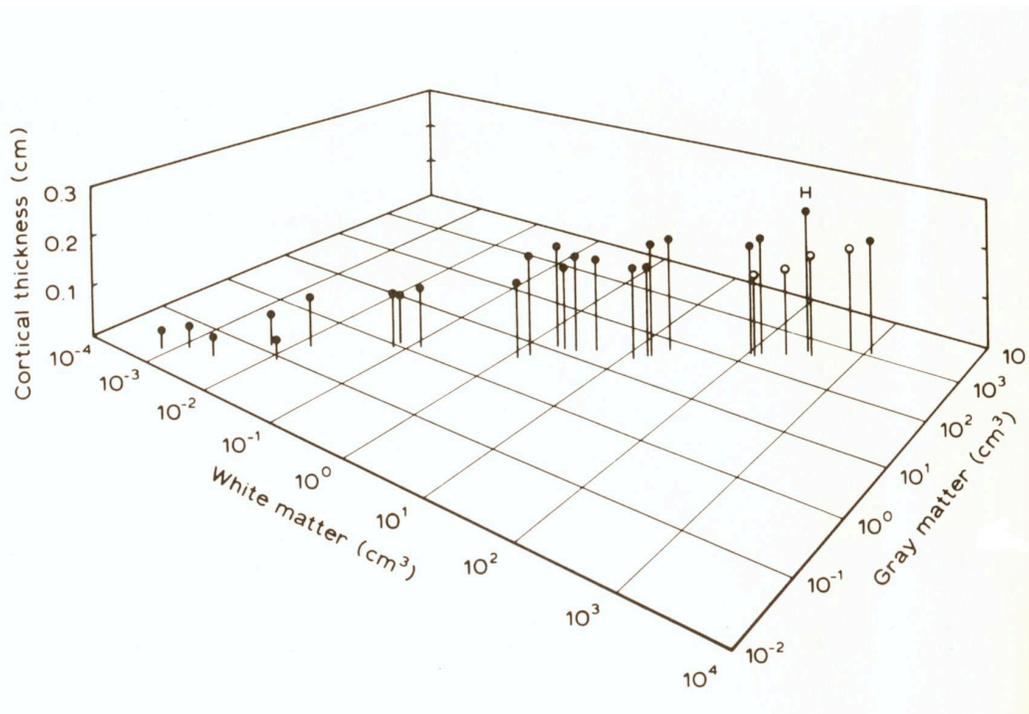


Figure 2.6: The relationships of cortical thickness, cortical grey matter and white matter across evolution. Each point represents a different species. Homo Sapiens are represented by the H. Whales and dolphins, which have a low cortical thickness relative to cortical volume, are represented by hollow points. One can tell from this graph that there is a linear relationship between white matter and grey matter volumes, but that cortical thickness changes at a much shallower rate, and is virtually identical for all species with a total volume above 3cm^3 . Graph taken from (Mountcastle, 1998).

The thickness measures described in this section have all come from post-mortem studies. The great advantage to the use of silver stained slice for this purpose is that it provides cellular level resolution, and thus can infer information about the underlying cellular patterns, be they layers, columns, or otherwise. There are, however, considerable problems with single slice estimation of cortical thickness as well. By far the most important is that the measurement of thickness is, in the convoluted cortex in the human, strongly dependent on the cutting angle used to extract that slice. Even

slight shifts in the angle can create considerable differences in the final measurement. A second problem is related to the use of silver-stained preparations, where the boundary between grey and white matter is diffuse and difficult to delineate precisely. According to Economo and Koskinas the blurred nature of the boundary caused differences of up to 0.5mm and greater in the estimation of thickness at the same location by different raters (Economo and Koskinas, 1925).

2.2. The cortex in health and disease

Studies of the human cerebral cortex were, until the advent of MRI, rare due to the labour intensive nature of working with post-mortem samples. Modern brain-imaging techniques have brought about an increasing literature on how the thickness of the cortex changes in health and disease. This section will discuss how the cortex changes in health and disease, the next section will provide a detailed analysis of the methods used to derive cortical thickness measures from MRI.

2.2.1. The developing cortex

The complex three-dimensional cortical structure of layers and columns is generated from a simple two-dimensional array of progenitor cells which line the lateral ventricular and sub-ventricular zones early in embryonic life. Neurons migrate from their place of origin to their final position in the cortex; the cortex is settled in an inside-out fashion, where neurons generated early in development finish their migration in the lower layers of the cortex (close to the white matter), cells generated later terminate closer to the pial surface. Neurons differentiate further after migration, become arranged in cortical layers, and begin to form efferent and receive

afferent connections. They also then develop trans-laminar intrinsic connections that form the basis of cortical columns (Rakic, 1995; Mountcastle, 1998). Cortical ontogenesis is illustrated further in figure 2.7.

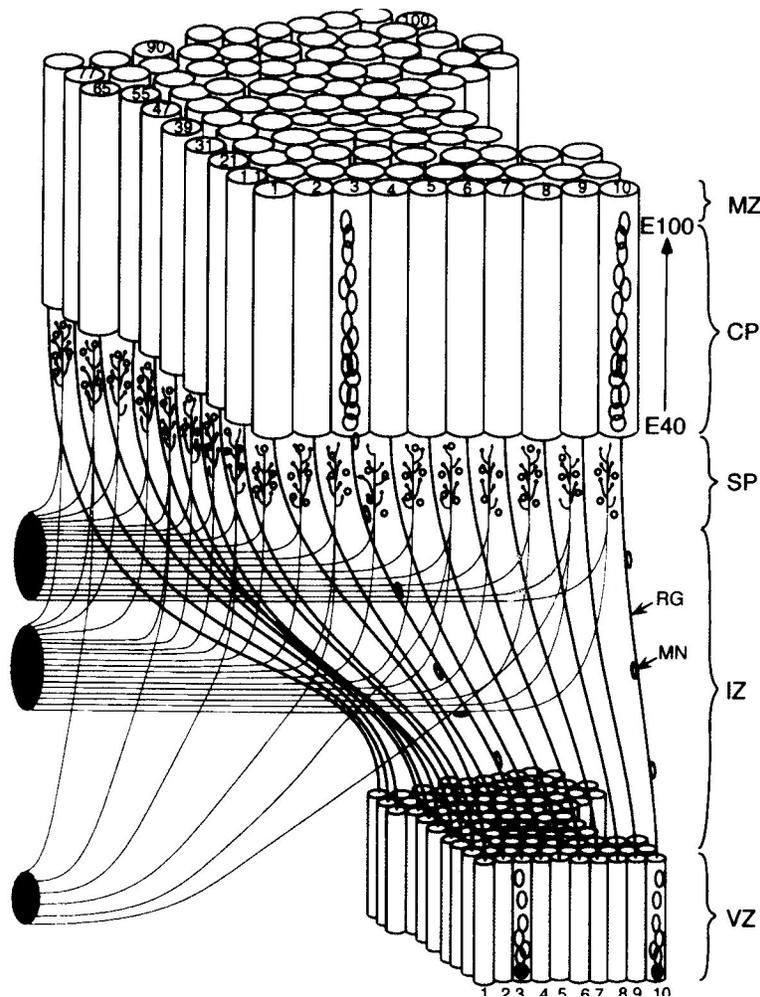


Figure 2.7: An illustration of cortical ontogenesis in the monkey. Neurons are generated in the ventricular zone (VZ), from whence they migrate (MN = migrating neuron) along radial glia (RG) through the intermediate zone (IZ) and the sub plate (SP). Neurons generated early in development, beginning at E40, settle in the deep cortical layers, subsequent neurons bypass them and settle closer to the marginal zone (MZ, the future cortical layer I). Taken from (Rakic, 1995; Mountcastle, 1998).

In the human the full complement of cortical neurons is formed by the third trimester. Between 15-30% of these neurons are then lost during maturation, mostly through competitive interactions between axons for synaptic space on target cells. The total number of synapses decreases by an even larger margin by the time puberty is reached, about 60-70% of the maximum (Mountcastle, 1998).

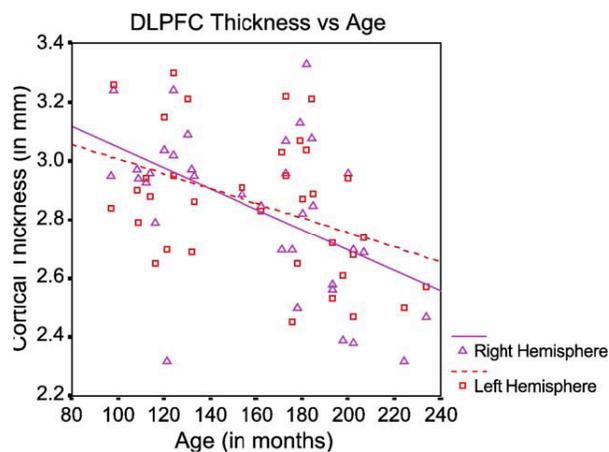


Figure 2.8: A plot of cortical thickness against age in the Dorsolateral prefrontal cortex (DLPFC). There is a linear decline within the population sample. Taken from (O'Donnell et al., 2005).

The cortex increases in size during the first three to five years of life, then begins to decline. The trajectory of cortical grey matter volume as a function of age is often described as an "inverted U," with the initial increase in size, likely caused by increasing arborization of neurons, turning into decreases later on in development, probably due to extensive pruning of synapses (Giedd, 2004). Between the ages of 6-20, the thickness of the cortex declines linearly and rapidly in most cortical regions (O'Donnell et al., 2005), while still growing in others, particularly temporal lobe regions (Sowell et al., 2004). An example plot is shown in figure 2.8. There is likely to be a difference in slopes of grey matter volume versus age when compared to cortical thickness versus age, as an expansion in

volume is potentially attributed to increasing surface area with no change in cortical thickness.

2.2.2. The ageing cortex

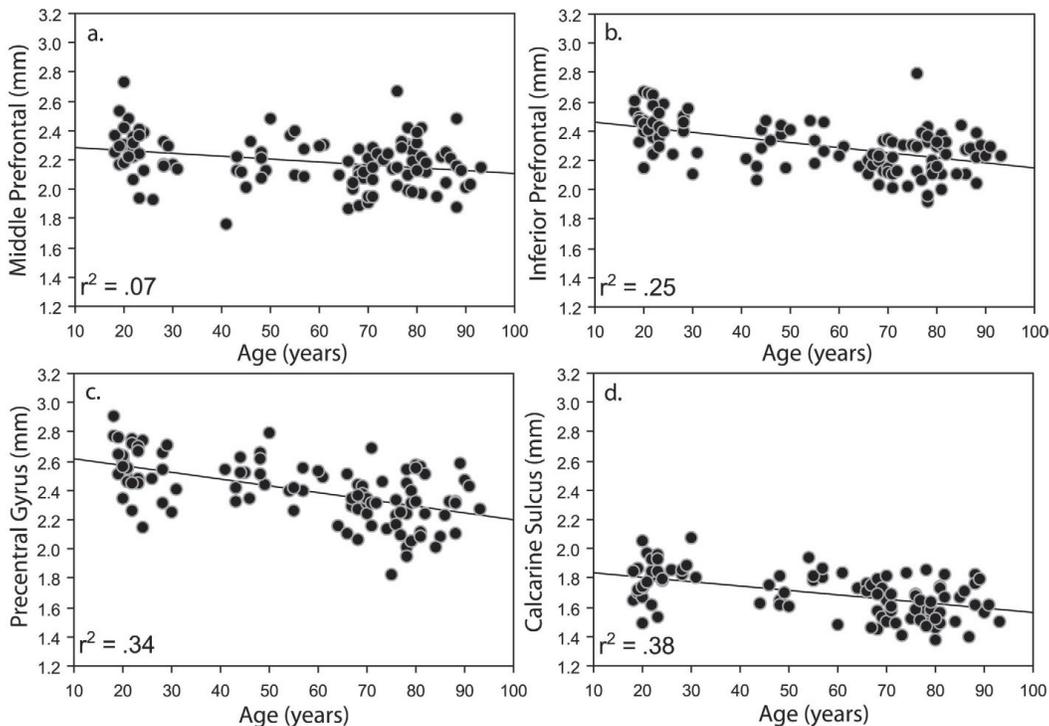


Figure 2.9: Four graphs of cortical thickness against age, each graph taken from a different cortical region as labelled on the y axis. One can see a linear decline with age which varies in slope by region, being highest in the precentral gyrus. Taken from (Salat et al., 2004).

Studies of normal ageing in the brain have shown an overall loss of brain volume and weight with increasing age, with the hippocampus and frontal lobes showing the greatest loss (Jernigan et al., 2001; Anderton, 2002; Resnick et al., 2003). The cause of this weight and volume loss is still disputed, and there is controversy over whether neuronal loss occurs as part of normal ageing (Anderton, 2002). The general opinion holds that, with age, cells shrink, and synaptic loss and neuronal loss are possible in

a few cortical regions (Anderton, 2002).

A recent study explored the role of the cortex in normal ageing (Salat et al., 2004). The cortex was found to thin with age, the greatest effect noticeable in the prefrontal cortex, primary motor cortex, and primary visual cortex, with relative sparing of the temporal and parahippocampal cortices, as shown in figure 2.9. The general pattern was the same for men and women. The decline was, however, quite small, averaging 0.016mm per decade, with a maximum of 0.07mm/decade in the primary motor cortex (Salat et al., 2004). Along with a decline in cortical thickness the gyral structure of the cortex also changes, with gyri becoming more sharply curved and sulci widening (Magnotta et al., 1999).

2.2.3. Degenerative diseases

Alzheimer's Disease

The normal development and ageing of the cortex can be altered by pathological processes. One pathology explored extensively in this thesis is Alzheimer's Disease (AD). Subtle impairment of memory is the earliest noticeable feature in AD patients. With progression of the disease further cognitive changes accumulate, including personality changes, language impairment, and ultimately even motor problems (Braak and Braak, 1991).

The gradual progression of the disease is accompanied by the accumulation of cellular pathologies: amyloid β ($A\beta$) deposition and neurofibrillary tangles (NFT) along with concomitant neuronal loss (Braak and Braak, 1991; Braak and Braak, 1995; Braak and Braak, 1996; Delacourte et al., 1999). $A\beta$ is a normal product of the cleavage of the $A\beta$ precursor protein. Depending on the secretase at the cleavage site either $A\beta_{1-40}$ or $A\beta_{1-42}$ are deposited, the former by far the most common. $A\beta_{1-42}$,

however, is far more fibrillogenic and is the major A β species present in senile plaques. A β deposition leads to increased oxidative stress, though the precise mechanism is unknown. There is some correlation between A β deposition and neuronal loss and cognitive impairment, though it is relatively weak (Rottkamp et al., 2002). NFTs are formed from the aggregation of pathological tau proteins (Delacourte et al., 1999), and are most commonly found in pyramidal neurons (Braak and Braak, 1995). In a study of the superior temporal sulcus it was found that a) neuron numbers are stable throughout the adult lifespan; b) are lost at a dramatic rate in AD; c) neuronal loss correlates highly with NFT presence; and d) neuronal loss does not correlate with senile plaques (Gomez-Isla et al., 1997). The columnar structure of the cortex, visible in the superior temporal sulcus throughout normal ageing, collapses with the presence of NFTs (Buldyrev et al., 2000). Moreover, increased duration of NFT presence is associated with increased atrophy (Grignon et al., 1998).

The progression of AD pathology is stereotyped: it begins in the entorhinal cortex, spreads across the trans-entorhinal cortex to the hippocampus, the rest of the limbic system, the associated cortices, and finally the motor and visual regions. This spread is characterized into either six (Braak and Braak, 1991; Braak and Braak, 1995; Braak and Braak, 1996) or ten (Delacourte et al., 1999) stages. Progression along the stages of AD is closely associated with clinical and cognitive deterioration, reflecting the degeneration of the cortical areas associated with these functions.

The association between neuronal loss and NFTs and the stereotypical progression of the disease leads to the hypothesis that, in patients diagnosed with probable AD, cortical areas involved earlier in the disease progression will show greater atrophy than areas involved later (Gomez-Isla et al., 1997). Imaging research has thus tried to find biomarkers that can be used in place of post-mortem analyses. MRI-based research has

focused on developing precise segmentation protocols for the early affected temporal lobe structures such as the entorhinal cortex (EC), or the hippocampus (HC) (Van Hoesen, 1995; de Leon et al., 1997; Mori et al., 1997; Krasuski et al., 1998; De Toledo-Morrell et al., 2000; Du et al., 2004; Ezekiel et al., 2004). A few studies have used global search algorithms to try to localize the spread of the disease (Fox et al., 2000; Thompson et al., 2001; Good et al., 2002; Scahill et al., 2002; Scahill et al., 2003), generally finding the greatest atrophy in areas implicated in the earlier stages of AD.

One of the key roles that imaging can play in dementia research as well as clinical practice is to aid the early diagnosis process (Scheltens et al., 2002; Chetelat and Baron, 2003; Zakzanis et al., 2003). The crucial task is to identify image-based metrics that can accurately differentiate patients from normal elderly controls. Ultimately, the goal will be to discover a metric that can detect dementia at an earlier stage than a standard neurological diagnosis. Discriminating normal elderly controls from patients diagnosed with probable AD using MR imaging techniques achieves accuracies ranging from 58% to 100%, with the perfect accuracy results relying on small samples and incomplete statistical controls for over-training (Chetelat and Baron, 2003).

Huntington's and Parkinson's Disease

Huntington's Disease (HD) is another degenerative disease recently explored using advanced imaging techniques. The main neuroanatomical underpinning to HD is striatal degeneration, but sensorimotor cortical areas are also affected (Rosas et al., 2002). Parkinson's Disease (PD) is similarly most closely associated with subcortical structures, though here too the cortex thins with age, particularly in the limbic and prefrontal cortex (Nagano-Saito et al., 2005). Even greater areas of the cortex are

implicated if PD is accompanied by dementia (Nagano-Saito et al., 2005).

Multiple Sclerosis

Multiple sclerosis, a progressive auto-immune disorder usually associated with white matter lesions, has also recently been shown to feature cortical grey matter atrophy, particularly in frontal and temporal regions (Sailer et al., 2003; Chen et al., 2004). The extent of cortical thinning increases with disease progression, featuring further loss of thickness in the motor cortex, and dissociated stable patients from patients with progressive disability (Sailer et al., 2003; Chen et al., 2004).

2.2.4. The cortex in psychiatric disorders

Schizophrenia is a disorder that includes symptoms such as altered preception of external events, cognitive and emotional states, along with the inability to identify the boundaries of the self (Innocenti et al., 2003). It is thought to be a progressive developmental disorder involving abnormal connectivity between different cortical centres (Innocenti et al., 2003).

The cortex in schizophrenia appears to contain similar cell numbers to normal controls, but greater cell density. The abnormalities lie in the neuropil - a reduction in size and numbers of neuron processes (Selemon and Goldman-Rakic, 1999). This ties in well with evidence of abnormal synaptic pruning during development and altered connectivity between cortical areas (Innocenti et al., 2003).

MRI based studies of cortical thickness in schizophrenia found thinning in the orbito-frontal, inferior frontal, inferior temporal, and occipito-temporal regions (Kuperberg et al., 2003). The gyri appear to be somewhat flattened compared to normal controls as well (White et al., 2003). The amount of cortical thinning has been related to the length of time that

schizophrenic processes have been active (Wiegand et al., 2004).

Specific phobias, particularly animal phobia, have also recently shown cortical thickness changes compared to normal controls. In this case, the cortex increases in thickness (Rauch et al., 2004). The areas showing focal increases were the bilateral insula, bilateral anterior cingulate, bilateral posterior cingulate as well as the left visual cortical regions (Rauch et al., 2004). The mechanisms underlying these cortical changes are unknown (Cannistraro and Rauch, 2003).

Autism shows a rapid increase in brain-size above and beyond normal controls in very early childhood (Redcay and Courchesne, 2005). Early evidence indicates that this is a global increase in size rather than the more focal nature of the disorders described above (Palmen et al., 2005), though some focus has been placed on the amygdala and the frontal lobes (Santangelo and Tsatsanis, 2005). The thickness of the cortex showed some focal thickening in primarily orbito-frontal and temporal regions (Chung et al., 2005). Attention Deficit/Hyperactivity Disorder (ADHD) shows relatively minor structural abnormalities (Bobb et al., 2005), these being focused on the fronto-striatal network (Willis and Weiler, 2005), along with associated white matter changes (Ashtari et al., 2005).

The overview of the cortex in degenerative and psychiatric disorders provided above illustrates that different diseases do not all affect the cortex in an identical fashion. Differences exist both in the type of change, including thickening (specific phobia, Autism) as well as thinning (AD, HD, and others). The underlying microscopic changes vary as well, ranging from neuronal loss (AD) to neuropil shrinkage (schizophrenia). The locus of cortical thickness differences varies as shown in figure 2.10. These patterns suggest the possibility of differentially diagnosing disorders that affect the cortex based on their cortical thickness.

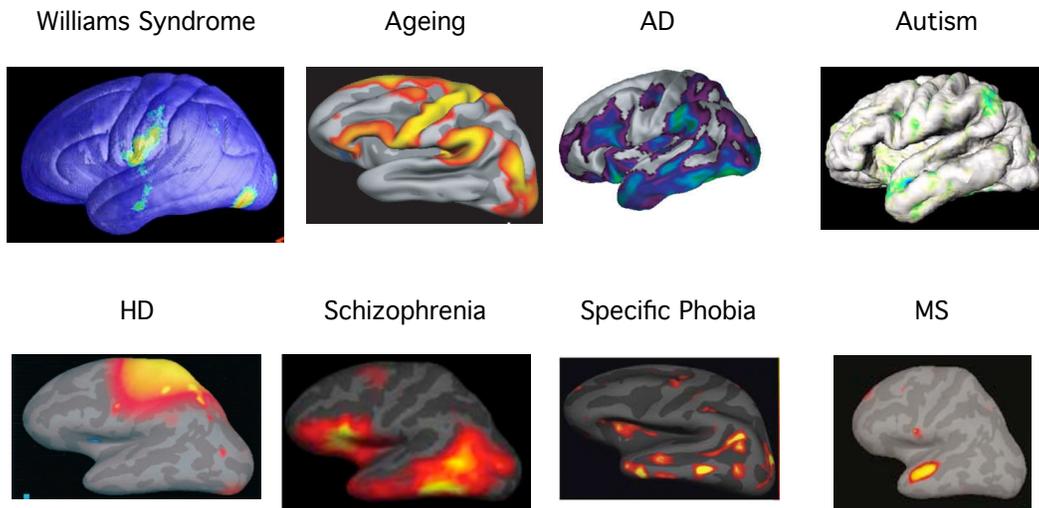


Figure 2.10: Statistics maps of cortical thickness change in the left hemisphere in eight different conditions are shown. The methods used to extract cortical thickness, control for covariates, and threshold for multiple comparisons differ between these eight maps, as does the population used to generate the analyses, so the comparison should not be overinterpreted. Nevertheless, one does notice different patterns of cortical thickness change in these eight populations. For instance, HD features strong correlations in the motor and sensori-motor regions, which are relatively spared in AD. Maps take from (Rosas et al., 2002; Kuperberg et al., 2003; Sailer et al., 2003; Lerch et al., 2004; Rauch et al., 2004; Salat et al., 2004; Chung et al., 2005; Thompson et al., 2005).

2.3. Structure-function relationships

The cerebral cortex does not only follow a stereotyped pattern of development that can be altered through disease processes, it can also be changed with experience. Such plasticity can be due to recovery from injury (Cohen et al., 1991) as well as training within a functional paradigm. It is the latter of these two that can teach us the most about the structure-function relationship within the cortex.

Pascual-Leone and colleagues showed that five days of piano finger exercises were sufficient to enlarge the cortical representation of the long finger flexor and extensor muscles (Pascual-Leone et al., 1995). Early acquisition of motor skills in string players also featured larger finger representations in the sensory-motor cortex (Elbert et al., 1995). These types of changes in cortical representations are likely caused by a combination of formation of new synapses, recruitment of cortical tissue for a task that was previously not used for that purpose, as well as a strengthening of existing synapses (Schlaug, 2001).

Behavioural training can cause cortical changes significant enough to be visible on MRIs. One of the best models for studying such cortical adaptations is through the comparison of musicians to non-musicians (Schlaug, 2001). In one such study searching across the entire cerebral grey matter professional musicians were found to have larger motor, auditory, and visuo-spatial regions when compared to amateur musicians (Gaser and Schlaug, 2003a; Gaser and Schlaug, 2003b).

Similarly, the hippocampus, a structure in part responsible for spatial navigation, of Taxi drivers was found enlarged when compared to controls (Maguire et al., 2000; Maguire et al., 2003). The hippocampus is special in the sense that we know neuro-genesis to occur there even late into the life-span. A similar pattern was found in correlating brain structure with the ability to learn novel sounds. Here greater white matter concentrations were found in the parietal lobes (Golestani et al., 2002). These types of changes can occur over a relatively short time period as shown in a study in which a group of participants were taught to juggle, this showed an increase in parietal lobe grey matter within 3 months of skill acquisition (Draganski et al., 2004).

These studies cited above tell us that the brain is adaptable to change,

and that task-specific specializations have their structural correlates. The microstructural component of these changes is not precisely known, though likely contains components of generations of new neurons (in the hippocampus at least), strengthening of existing synapses as well as the formation of new synapses (Gaser and Schlaug, 2003a; Gaser and Schlaug, 2003b). Changes in morphology based on functional adaptations suggest that the cortical thickness of areas subserving related functions might change in a correlated fashion.

2.4. Principles of structural MRI

Magnetic Resonance Imaging (MRI) provides among the best in-vivo capabilities of identifying soft tissue structure. MR signal is essentially produced by protons. Each proton rotates around its own axis at random. When a strong magnetic field and a radio frequency (RF) pulse are applied at right angles to each other, the protons as a group begin to precess about their own axis in synchrony with each other. When the RF pulse is turned off, the system returns to normal and in doing so create small local magnetic fields which in turn gives rise to small electric currents in receiving coils. It is this current that is ultimately measured in MRI. Specifically, the current collected during the time period that the protons return to the rest state is captured in two time constants, T_1 and T_2 . T_1 images emphasize the righting of the protons as they realign with the magnetic field, T_2 the dephasing of the protons. T_1 and T_2 time constants vary depending on whether the protons belong to fat, CSF, or white matter, leading to the excellent soft tissue contrast provided by MRI. Proton Density (PD) images, a third type of commonly used MR modality, measures the number of protons per unit of tissue. Spatial resolution in MRI is provided by varying the strength of the magnetic field along three

axes - magnetic gradients (Kandel et al., 2000). Example images are shown in figure 2.11.

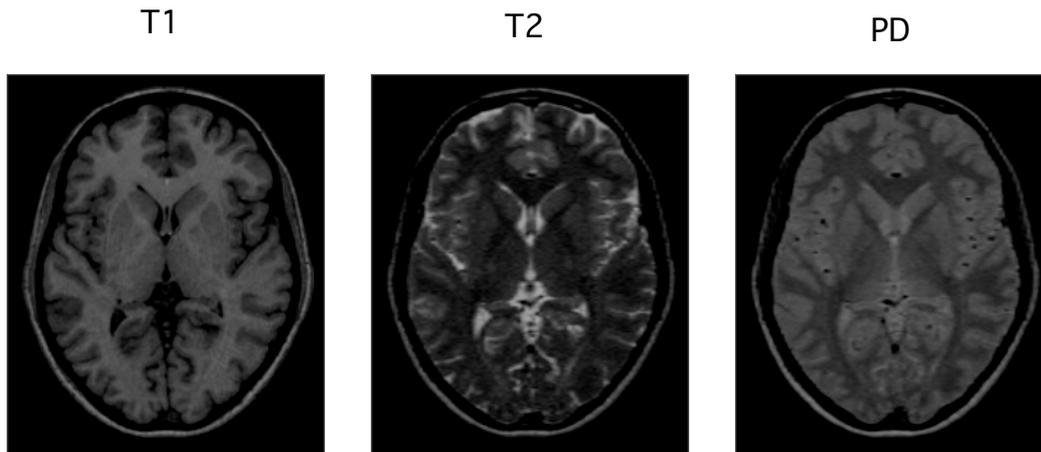


Figure 2.11: Three MRIs of the same subject are shown. These illustrate the difference in signal by three acquisition types (T_1 , T_2 , PD) as well as the level of resolution obtainable with standard MRIs (1mm^3).

The resolution of in-vivo MRI is most commonly 1mm^3 for T_1 images, 2 or 3mm^3 for T_2 and PD. At this level of detail one can distinguish between different tissue types. There is, using these standard modalities, no ability to be able to measure the cellular events that might be causing a change in tissue volume or cortical thickness as measured from MRI.

Two of the distinct diseases that affect the cortex outlined above - AD and schizophrenia - illustrate both the power of MRI as well as its limitations. MR was shown to be effective in localizing cortical changes in the disease, which in both cases involved thinning of the cortical ribbon. Yet the cellular changes that underly the reduction in cortical thickness appear to be quite different. AD features a loss of neurons associated with the spread of neurofibrillary tangles (Gomez-Isla et al., 1997; Buldyrev et al., 2000). Schizophrenia, on the other hand, shows no loss of neurons, instead neuron density is higher and the neuropil is altered (Selemon and

Goldman-Rakic, 1999; Innocenti et al., 2003). The message from these two cases is clearly that, at this point in time, in-vivo MRI can solely point to macroscopic changes occurring in the cortex, and cannot infer the microscopic mechanism underlying these changes without additional, non-MR based analyses.

2.5. Image Processing

The raw data obtained from the MR scanner requires multiple operations to derive information for neuro-anatomical studies. The types of data generated include global measures, such as total cerebral volume, localised information about different brain structures, or even data at the smallest sampling unit of MRIs, the voxel.

MRI processing is usually arranged in a series of steps that combine to produce the final data to be analysed. These steps can be either fully automatic, semi-automatic, or manual. The advantages to fully automatic processing are extensibility to large datasets as well as reproducibility of results. Manual or semi-automatic techniques are labour intensive and subject to inter- and intra-rater variability. There are nevertheless several domains, foremost amongst them segmentation of the brain into its component structures, where automatic techniques only recently began approximating manual structure delineations in accuracy (Collins et al., 1995; Fischl et al., 2002; Rettmann et al., 2002; Liu et al., 2004; Mangin et al., 2004).

Volumes of interest (VOI) as defined by structural delineation are primarily used to assess differences in volumes across populations. Such studies include the use of lobar volumes to understand the growth of the brain during normal development (Giedd, 2004; Gogtay et al., 2004) and studies

of the entorhinal cortex or hippocampus in early AD (Juottonen et al., 1998; Krasuski et al., 1998; Bobinski et al., 1999; Juottonen et al., 1999; Killiany et al., 2000; Xu et al., 2000; Pruessner et al., 2002).

Using volumes of interest can, however, obscure information. In early AD, for example, the parahippocampal gyrus is atrophied, whereas the lateral temporal lobes are still relatively intact (Braak and Braak, 1991; Killiany et al., 1993; Braak and Braak, 1995; Krasuski et al., 1998; Bobinski et al., 1999; Chetelat and Baron, 2003). If the VOI consists of the temporal lobes as a whole, then the change occurring solely within the parahippocampal gyrus might be missed. The alternative is to use techniques which search the whole brain simultaneously. The dominant such algorithm is Voxel Based Morphometry (VBM). VBM relies on the segmentation of the brain into its component tissue types: white matter (WM), grey matter (GM), and cerebro-spinal fluid (CSF). The desired tissue is then blurred with a Gaussian kernel to produce a tissue density estimate at each voxel, which represents the relative amount of tissue within the blurring kernel. Statistical models can then be performed at every voxel across a population whose scans have all been aligned to the same stereotaxic space (Wright et al., 1995; Paus et al., 1999; Ashburner and Friston, 2000; Good et al., 2002).

An alternate approach is to extract the best possible non-linear warping between datasets and then perform analysis on the resulting deformation field (Janke et al., 2001; Liu et al., 2004; Robbins et al., 2004; Shen and Davatzikos, 2004). Such studies can show local displacements in areas of brain growth or shrinkage. These methods can also be combined with VBM to produce results partly contained in the deformation field, partly within the image, referred to as voxel compression maps or the boundary shift integral (Fox et al., 2000; Scahill et al., 2002; Scahill et al., 2003;

Ezekiel et al., 2004).

As discussed earlier, the structure of the cortex approximates a sheet with varying thickness. Processing of information occurs first within the vertical unit of the cortex, the columnar ensemble, before being distributed to other cortical regions. Volumetric information has several drawbacks for the analysis of cortical data. The most important of these is that topology is not respected: two points that might be very close in x-y-z coordinate space might be far apart with respect to the surface of the cortex. It is for this reason that data about the cortex is best represented on a surface mesh rather than at every voxel. Surface based representation allows for much better analysis of cortical information processing and has even been used for cross-species registration (Van Essen et al., 1998).

The focus of this dissertation is the analysis of cortical thickness, whose methodology will be addressed in detail. There are, however, other surface based analyses available. One can study the placement and variability of different cortical sulci, analyzing, for example, whether handedness affects the size and shape of the central sulcus (Le Goualher et al., 1999; Le Goualher et al., 2000; Davatzikos and Bryan, 2002; Mangin et al., 2004). Another oft employed technique is to create deformation fields mapping one cortex to another (Drury et al., 1996; Van Essen et al., 1998; Fischl et al., 1999; Resnick et al., 2003; Hurdal and Stephenson, 2004; Liu et al., 2004; Robbins et al., 2004; Van Essen, 2004). These warps can then be used to study, for example, age associated changes across a longitudinal time series (Resnick et al., 2003).

The next section will describe the methods used to extract cortical thickness. Once the MRI(s) for a subject has been acquired, it has to be processed in order to extract quantitative data about the shape of the

brain. The first part will hence discuss the various steps necessary to measure cortical thickness, starting with an in depth description of the methods used at the Montreal Neurological Institute which were employed to generate all the data for the studies later in this thesis. This will be followed by an overview of the image processing literature to examine similarities and differences with other methods available.

2.5.1. MNI Image Processing Pipeline

The Montreal Neurological Image Processing Pipeline is based upon the MINC (Medical Imaging NetCDF) file format and toolkit. The pipeline consists of multiple independent programs that are tied together through an intricate set of dependencies. The pipeline is outlined in figure 2.12 and described in more detail below.

The basic processing flow is the following: the native MRI is registered into stereotaxic space and corrected for intensity non-uniformity artifacts, then classified into grey matter, white matter, and cerebro-spinal fluid. Non cerebrum tissue is removed and the partial volume content for each tissue class estimated. The white matter surface is then extracted, followed by an expansion towards the pial surface boundary. Cortical thickness is estimated by taking the distance between the white and grey surfaces, and thickness maps are then blurred along the surface. The surfaces can optionally be aligned towards a standard model using non-linear deformations. Automatic segmentations techniques can also be added to the data processing in order to obtain regional thickness estimates.

One of the first steps in the pipeline is the removal of intensity inhomogeneity artefacts. The magnetic signal measured from the brain within the MR scanner is not perfectly homogeneous - a requirement for tissue classification. The intensity of the signal instead varies gradually

across the image due to radio frequency coil non-uniformity, gradient driven eddy currents, and the fact that the subject being scanned is not entirely within the scanner's field of view (Sled et al., 1998). The image non-uniformity is removed by finding the smooth, slowly varying field that maximizes frequency content of the intensity distribution of the image (Sled et al., 1998).

At the same time that the non-uniformity artefact is being removed the image is registered into stereotaxic space. The basis of MNI stereotaxic space is the human brain atlas created by Talairach and Tournoux in 1988 (Talairach and Tournoux, 1988). This atlas was, however, based on a single subject. In order to remove the idiosyncracies associated with any single brain a set of 305 scans was semi-automatically registered towards Talairach space, then all registered MRIs were averaged to create an average brain that could be used as a registration target (Collins et al., 1994; Mazziotta et al., 2001). New images are linearly aligned towards the average brain by maximizing the image cross-correlation, resulting in a transformation with independent translations, rotations, and scale factors for each of the three cardinal directions (Collins et al., 1994). If multiple MR sequences were acquired then each subject's acquisition are aligned to each other using mutual information as the objective function (Collins et al., 1994).

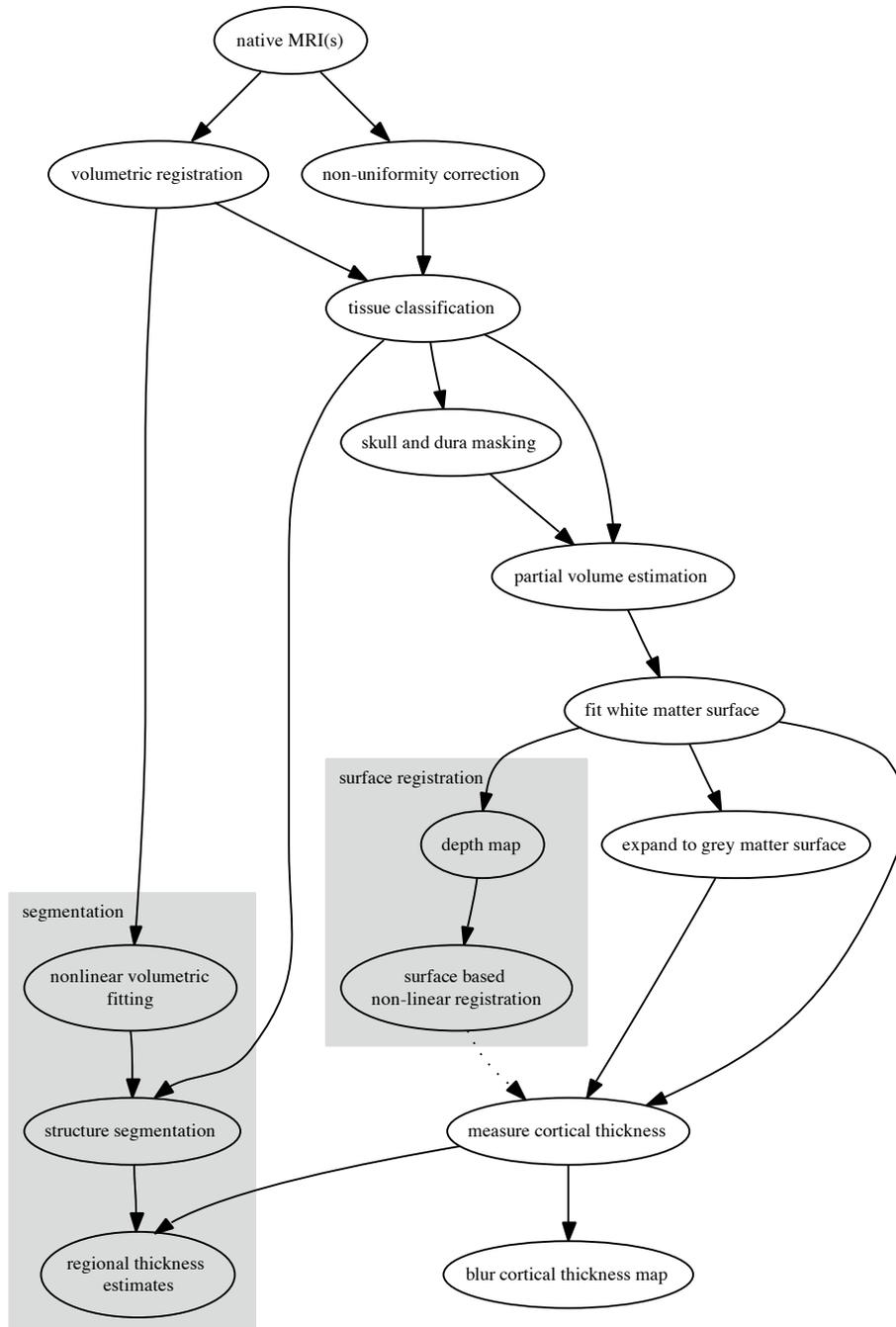


Figure 2.12: all the major image processing steps needed to generate cortical thickness maps from native MRIs are shown here. Greyed out areas are optional. See the text for a detailed description of all the steps.

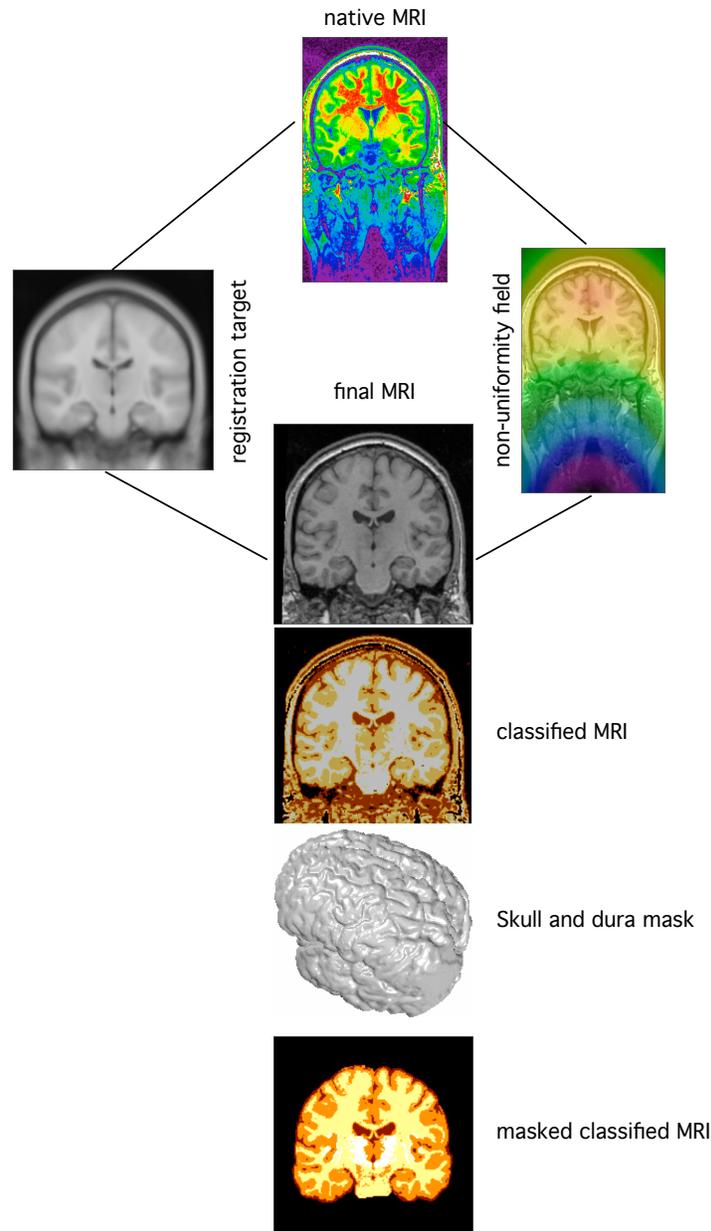


Figure 2.13: the steps needed to classify the MRI in stereotaxic space are shown. The native MRI is registered towards an average template while simultaneously the non-uniformity field is removed. The resulting "final" MRI is then classified into WM, GM, and CSF, a mask of the skull and dura computed, and the mask applied to the classification. This image shows a separate tissue class for subcortical GM as well.

One of the advantages of registering images to a standard space is in the ability to provide models for further image processing. An example comes from the tissue classification step. The goal is to associate each voxel in the MRI with one of four classes: white matter, containing the myelinated fibre tracts in the brain; grey matter, containing both cortical and sub-cortical neurons; cerebro-spinal fluid (CSF); and background. This classification has to be tuned for each individual subject, as different scanners and scanning protocols can result in varying image intensities for each tissue class. The way tissue classification is carried out is thus the following: a series of images in stereotaxic space are semi-automatically classified, and probability maps of each tissue class created. 1000 points with a 90% or greater chance of coming from each class are randomly sampled and used to train an artificial neural network classifier based on the signal intensity characteristics of each new MRI to be classified (Kollokian, 1996; Zijdenbos et al., 2002). Data processing up to this point in the pipeline is illustrated in figure 2.13.

The next step in the pipeline is the removal of skull, dura, brain-stem, and cerebellum. This is performed by fitting a coarse cortical surface to the classified image and removing any voxels lying outside of the surface. The fundamental concepts are the same as for the white and grey matter surface fitting and will thus be discussed in more detail later.

The classification described above is discrete - every voxel in the brain is labelled as one tissue class. Voxels are of a defined size, usually 1mm^3 , and the actual composition of that voxel contains more than one tissue type. This happens at the boundary of the cortex, for example, and is called the partial volume effect. By estimating the relative contribution of each tissue type at any given voxel one can more accurately extract fine

cortical features that are obscured if one solely relies on the discrete labelling of voxels on a one millimetre grid. We use a trimmed minimum covariance determinant method for estimating the partial volume effect (PVE) (Tohka et al., 2004). This method takes the discrete classification as well as the stereotaxic space MRIs as inputs, then uses a set of morphometric operators to remove all voxels which could possibly be contaminated by partial volume. The thus eroded classified image is used to estimate the mean and covariance for grey matter, white matter, and CSF, and each voxel then has a value between 0 and 1 associated with it for each tissue type, representing the proportion of that voxel containing that tissue type (Tohka et al., 2004; Kim et al., 2005). A medial surface image and a curvature image are also derived from the T1 input image in order to improve the detectability of sulcal CSF (Kim et al., 2005).

The inner and outer cortical surfaces are then extracted using the Anatomic Segmentation using Proximity (ASP) (MacDonald et al., 2000) or Constrained Laplacian ASP (CLASP) (Kim et al., 2005) algorithm. Since CLASP is an improved version of ASP, discussion will focus upon it. The essence of CLASP is the creation of simple (non self-intersecting) surfaces with spherical topologies using deformable models.

The process begins with the deformation towards the white matter surface. The discrete tissue classification volume is taken as input and a polyhedron is iteratively deformed towards its white matter boundary. Along with the image information, $T_{boundary-dist}$, several model terms are used to constrain the fit, and self-intersection is explicitly prohibited. The model terms are: $T_{stretch}$, constraining distances between neighbouring vertices; T_{bend} constraining deviation from model shape; and $T_{self-proximity}$, constraining the proximity of pairs of non-adjacent polygons (MacDonald et al., 2000; Kim et al., 2005).

Extraction of the pial surfaces from MRI images is a difficult task due to the complex folded nature of the cortex combined with the limited resolution of modern anatomical MRI. One of the key problems encountered is that of fused gyri, where the sulcus separating two gyri is too small to be easily discerned on MRI and the discrete classification thus merges the grey matter of the two gyri into one structure. The approach taken in CLASP is to use the partial volume information to compute a skeleton of cortical CSF. Since PVE estimation is capable of detecting sulcal CSF in locations where the voxel would otherwise be classified as grey matter this skeleton can correctly reproduce the topology of sulci, as shown in figure 2.14.

The pial surface is obtained by creating a laplacian volume which contains the cortical grey matter minus the skeletonized CSF. The white matter surface is taken as the starting point and the algorithm then expands towards the outer cortical boundary using the same constraints as listed for the white matter surface along with $T_{surface-surface}$, preventing the two surfaces from coming within a certain distance of each other, and $T_{laplace}$, which follows the laplacian map. The entire algorithm is illustrated in figure 2.15.

Once the inner and outer and cortical surfaces have been extracted cortical thickness can be measured as the distance between the two surfaces. This is the topic of the first article in this thesis and will thus be left out of the background literature review. See chapter 3 for the full discussion of this topic.

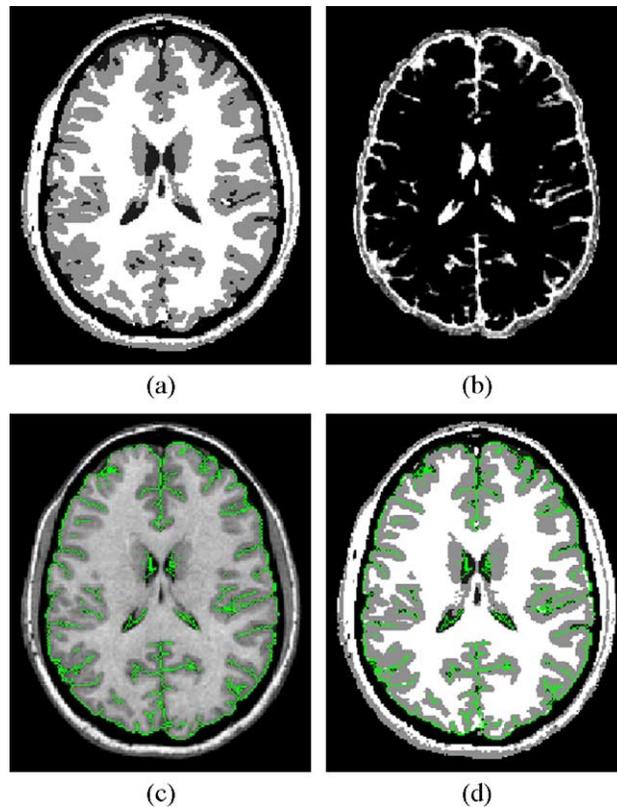


Figure 2.14: part (a) of the figure shows the discrete tissue classification; part (b) the partial volume CSF map; part (c) the T1 overlaid with the CSF skeleton in green, and part (d) the classification with the same CSF skeleton. The skeletonization of cortical CSF leads to full sulcal penetration. Reproduced from (Kim et al., 2005).

Due to both image processing artefacts and, more importantly, the limited resolution of structural MRI, the cortical thickness data obtained is inherently noisy. In order to smooth the thickness maps and thereby also increase comparability of vertices across subjects the data is blurred along the surface using a diffusion smoothing kernel (Chung et al., 2002; Chung et al., 2003). The basic principle behind smoothing is to use a Gaussian filter which gives decreasing weight to signal further away from the point being smoothed up to a distance that can be specified by the user (20-35mm for most of the data used in this thesis). Due to the non-

Euclidian nature of cortical geometry the weights for the diffusion operator has to be estimated, in this case using the Laplace-Beltrami operator - an extension of Gaussian kernel smoothing to arbitrary manifolds (Chung et al., 2003). Diffusion smoothing on the surface is similar to the volumetric blurring employed for voxel based morphometry, the major exception being that it follows surface topology.

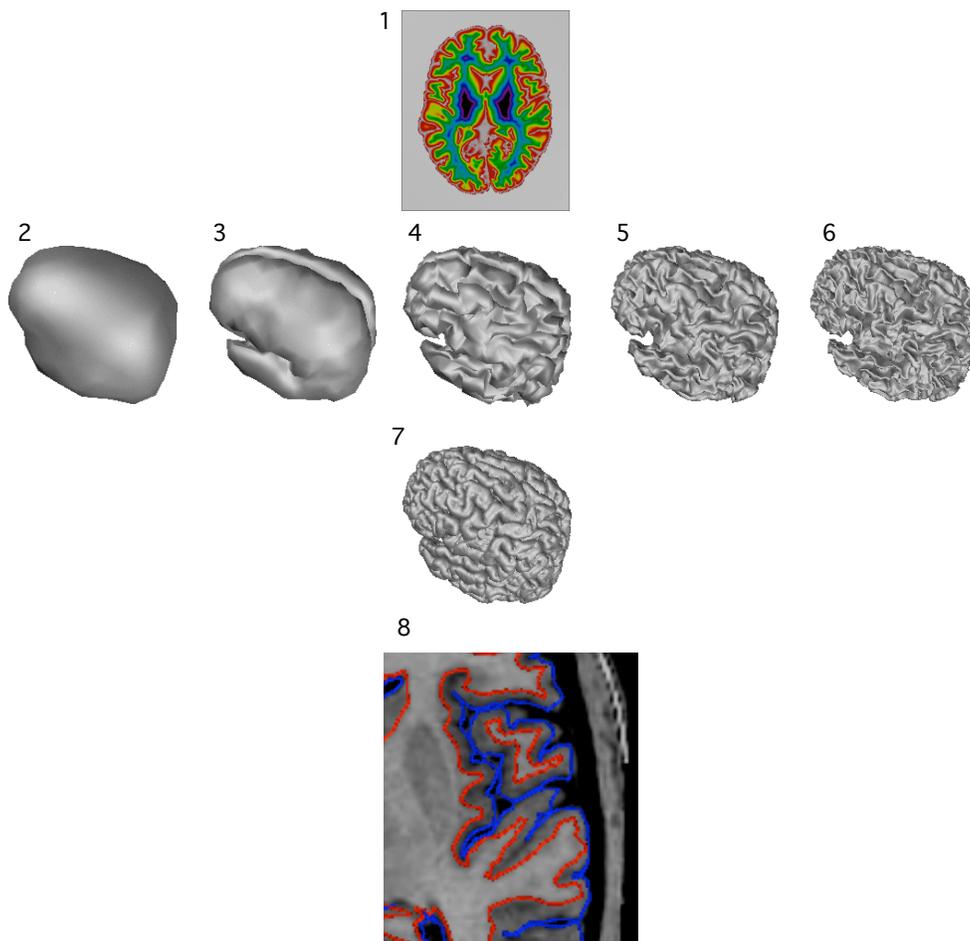


Figure 2.15: An illustration of the CLASP surface fitting procedure. A laplacian map is generated (1) and the white matter surface is fit through an iterative procedure with increasing resolution (2-6), upon whose completion the algorithm expands towards the grey matter surface (7). The intersection with the volume is shown in (8) in order to illustrate the accuracy of the fit.

There are two optional elements to the image processing pipeline: automatic region segmentation of the cortex and non-linear surface based registration. These steps add respectively the ability to have regional definitions of cortical thickness as well as using cortical anatomy to bring vertices into closer correspondence than can be done solely through linear registration or even volumetric non-linear warping (Robbins et al., 2004).

Automatic segmentation of the cortex is accomplished through the ANIMAL algorithm (Collins et al., 1995). A subject's MRI is non-linearly aligned with a template. This template has been segmented into all major cortical and subcortical structures, such as, for example, the superior temporal gyrus or the parietal lobule. The non-linear transformation is inverted and the segmentation from the template is transferred to the volume. The tissue classification is used to ensure that grey matter structure labels are confined to voxels labelled as grey matter (Collins et al., 1995).

The surface-based registration procedure uses depth maps to drive the non-linear warping. Depth maps measure the distance of any point on the cortical surface to a rigid sheet encasing the entire cortex. Points on the crowns of gyri thus have the least depth, points in the fundi of sulci have the greatest depth. By aligning based on depth maps one maximizes the chance that corresponding vertices will be in similar positions within a cortical fold (Robbins et al., 2004).

2.5.2. Other MRI based techniques

The use of automatically extracted cortical thickness information is still rare, though it is becoming increasingly popular, as shown in figure 1.1. A literature search on cortical thickness reveals 37 references, relying on a much smaller number of methodological implementations - see table 2.3.

Reference	Topic
(Chung et al., 2005)	Methods: diffusion smoothing
(Kim et al., 2005)	Methods: surface definition
(Lerch and Evans, 2005)	Methods: thickness validation
(Narr et al., 2005)	Applications: schizophrenia
(Barta et al., 2005)	Methods: laminar structure
(O'Donnell et al., 2005)	Applications: normal development
(Thompson et al., 2005)	Applications: William's Syndrome
(Preul et al., 2005)	Applications: Cerebral Microangiopathy
(Sowell et al., 2004)	Applications: normal development
(Lerch et al., 2004)	Applications: Alzheimer's Disease
(Rauch et al., 2004)	Applications: Animal Phobia
(Sbarbati et al., 2004)	Applications: foetal development
(Salat et al., 2004)	Applications: normal ageing
(Wiegand et al., 2004)	Applications: schizophrenia
(Casanova et al., 2004)	Applications: dyslexia
(Worsley et al., 2004)	Methods: statistical thresholding
(Tosun et al., 2004)	Methods: surface definition
(Kuperberg et al., 2003)	Applications: schizophrenia
(Zhang et al., 2003)	Applications: cortical dysplasia
(White et al., 2003)	Applications: schizophrenia
(Sailer et al., 2003)	Applications: multiple sclerosis
(Rosas et al., 2002)	Applications: Huntington's Disease
(Hund-Georgiadis et al., 2002)	Applications: Arachnoid cysts
(Antel et al., 2002)	Applications: cortical dysplasia
(Ross et al., 2001)	Applications: lissencephaly

(Woo et al., 2001)	Applications: cortical dysplasia
(Ratnanather et al., 2001)	Methods: surface definition validation
(Sowell et al., 2001)	Applications: normal development
(Kabani et al., 2001)	Methods: thickness validation
(Fischl and Dale, 2000)	Methods: thickness measurement
(Jones et al., 2000)	Methods: thickness measurement
(MacDonald et al., 2000)	Methods: surface definition
(Miller et al., 2000)	Methods: thickness measurement
(van der Valk et al., 1999)	Applications: heterotopia
(Zeng et al., 1999)	Methods: surface definition
(Magnotta et al., 1999)	Applications: normal ageing
(Meyer et al., 1996)	Methods: thickness measurement

Table 2.3: a comprehensive list of all the publications relating cortical thickness to MRI in the human, sorted by date of publication. Methods papers are in grey. The PubMed search used to generate this list was "cortical thickness" AND MRI NOT bone NOT mice NOT canine NOT rat NOT kidney. Only references in English language journals were used.

There are four basic types of thickness measures: manual or semi-automatic methods, automated volume based methods, automated mixed mesh-volumetric methods, and automated mesh-based methods. These will be discussed in turn below.

Manual or semi-automatic methods

Manual measurements of cortical thickness from MRI involves selecting a slice perpendicular to the piece of cortex under investigation and using the equivalent of a jeweler's eye-piece to estimate the distance of the inside surface to the outside surface (Meyer et al., 1996). In cases of gross

morphological disorders such as those evident in malformations of cortical development a simple visual assessment of cortical thickness can be made (Ross et al., 2001).

Automated volumetric methods

Volumetric methods define cortical thickness entirely on the basis of a preprocessed MR volume. The best example of a pure volumetric technique comes from Jones and colleagues (Jones et al., 2000), also employed in (Antel et al., 2002) and tested in (Lerch and Evans, 2005). Here the thickness of the cortex is solved through a partial differential equation. White matter voxels are defined as the inside boundary, CSF or background voxels as the outside boundary. Laplace's equation is then iteratively solved across the voxels of the cortical grey matter. Upon completion of this algorithm every cortical voxel has an associated thickness value (Jones et al., 2000). This approach is extended in (Yezzi and Prince, 2003) where the need to explicitly construct correspondence trajectories is eliminated. Lohmann et al. (Lohmann et al., 2003; Preul et al., 2005) measure the 3D Euclidian distance of any GM/CSF boundary voxel to the white matter surface. Similarly, (Wiegand et al., 2004) identify both the GM/CSF and the GM/WM boundaries in the volume, and estimate thickness by taking the distance between closest points from one boundary to the other.

Mixed mesh-volumetric methods

Mixed mesh-volumetric methods define cortical thickness through both the use of one or two cortical meshes as well as volumetric information. The recent work of Paul Thompson's group is a good example of such a mixed technique. Here a polyhedral mesh is deformed towards the outer cortical

surface using an older version of the method described in (MacDonald et al., 2000; Kabani et al., 2001; Kim et al., 2005). A single surface is, by definition, not adequate for defining the thickness of the cortex. Instead cortical thickness is solved volumetrically by measuring the distance from each cortical grey matter voxel to the cortical mesh. The next step is to then use a volumetric blurring kernel across the thickness volume at each point on the cortical mesh to define the thickness of the cortex at that point (Thompson et al., 2005).

A similar method is employed in (Miller et al., 2000; Ratnanather et al., 2001). Here the tissue classified MR volume is used as an input to generate the outer cortical surface through an isosurface technique similar in principle to the marching cubes algorithm. This algorithm generates triangles at the outer boundary of a set image threshold. Thickness is then measured at each voxel by computing its distance from the cortical surface (Miller et al., 2000; Ratnanather et al., 2001). This is related to the method used in (Magnotta et al., 1999), where the cortical surface represented by pure GM voxels (no CSF partial volume contamination) is reconstructed using a marching cubes algorithm and the distance to the closest point on the GM/WM boundary is determined at every vertex. This measure is referred to as cortical depth rather than thickness since it does not truly span the entire cortex, seeing how the cortical surface does not go all the way to the pial boundary (Magnotta et al., 1999).

One of the key issues faced by volumetric or mixed mesh-volumetric techniques is accurate estimation of cortical thickness in the presence of buried sulci. Three solutions to this problem are proposed: a pre-processing step which emphasizes sulcal CSF, either by distance based operators (Lohmann et al., 2003; Preul et al., 2005) or through analysis of the MR signal (Jones et al., 2000). The second solution is to compute cortical thickness by iteratively propagating the thickness values from the

GM/WM intersection (Thompson et al., 2005). This ensures that in the presence of a buried sulcus the maximal thickness value will be equidistant from the two opposing WM banks, and, in combination with the blurring step, negates the problem (Thompson et al., 2005). The third solution is manual editing of the input volume to ensure that sulci are fully represented (Wiegand et al., 2004).

Mesh-based methods

Mesh-based methods construct polygonal models for both the inside and the outside cortical surface. Cortical thickness is then measured as the distance between the two surfaces (though multiple types of distances can be defined). One of the most used methods for estimating cortical thickness is implemented in the freely available FreeSurfer program and is mesh based (Dale et al., 1999; Fischl et al., 1999; Fischl and Dale, 2000). This algorithm is similar in principle to the deformable model approach described in the section on the MNI image processing pipeline (MacDonald, 1997; MacDonald et al., 2000; Kabani et al., 2001; Kim et al., 2005). White matter voxels are identified, and a surface reconstructed using a marching cubes like algorithm. A deformable model is then used to fit the tessellated WM surface, followed by an expansion to the pial boundary. Self intersection is prevented throughout. Some topological errors might remain; these can be corrected by hand. Cortical thickness is measured as the closest distance between the white and pial surfaces, computed twice (once from each surface) and averaged (Dale et al., 1999; Fischl and Dale, 2000).

A similar method is introduced by Tosun et al. (Tosun et al., 2004) wherein the WM surface is constructed through a modified version of the marching cubes algorithm, followed by a deformable model expansion to an intermediate cortical surface and the pial surface. Thickness is measured

at each voxel of the cortex by taking the distance to the closest inner and outer surface vertex (Tosun et al., 2004).

Zeng and colleagues create the two surfaces of the cortical sheet through the use of a coupled-surface level set algorithm (Zeng et al., 1999). Two embedded surfaces are evolved simultaneously towards two different image-derived targets (the WM/GM boundary and the CSF/GM boundary) while constraining the distance that the two surfaces can be apart. The surfaces themselves are constructed using a marching cubes algorithm at each iteration of the level set. Cortical thickness is measured by taking the distance between the two surfaces at every point, correspondence being kept through the coupled surface propagation model (Zeng et al., 1999).

2.5.3. **Statistical Analysis**

Once the cortical thickness maps are extracted the data can be analyzed. There are three issues to contend with:

1. Identifying corresponding points in the study population.
2. Using the appropriate statistical test.
3. Controlling for multiple comparisons.

There are two ways to provide corresponding points for statistical analysis: alignment of surface-meshes or segmentation through regional definitions. Regional segmentation takes the average cortical thickness over a specified area, and compares these averages across subjects. This type of analysis is attainable regardless of the thickness definition used.

Point-by-point comparisons of cortical thickness, on the other hand, requires that cortical thickness be measured on a mesh. The main reason is that human cortical features are quite variable from person to person, and volumetric alignment cannot guarantee that one is comparing cortical

voxel to cortical voxel. This especially the case when one considers matching one subjects 2mm thick cortex to another's 4mm thick cortex on the same gyrus: it is inevitable that at least 2 voxels in one subject will have no match in the other. By transferring data to a mesh this problem disappears, since these meshes are infinitely thin and will thus have just one thickness measure per vertical strip of cortex.

Cortical meshes can be aligned in one of two fashions: through linear alignment, usually performed volumetrically, or through spherical alignment based directly on cortical features. The linear alignment approach is used by default in the MNI image processing pipeline, and relies on volumetric registration of the input MR volumes. The cortical surfaces are then generated from the same starting model. The resulting surfaces will thus have individual vertices in rough correspondence across subjects as shown in figure 2.16.

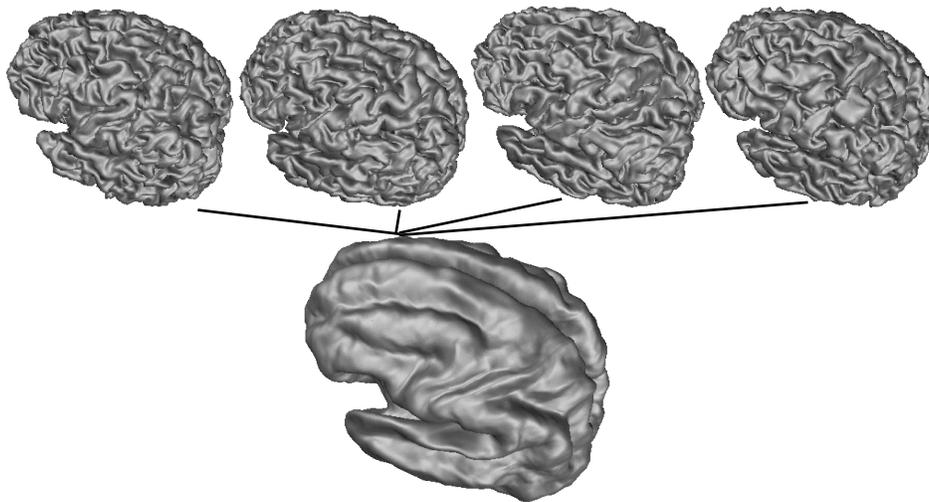


Figure 2.16: The top row shows four individual cortices, the bottom an average cortex created by geometrically averaging all individual cortices in a study sample. One can see the retention of the major sulci such as the central sulcus, superior frontal and superior temporal sulci, while the more variable sulci are washed out on the average.

The second method, non-linear alignment uses cortical features to bring individual vertices into closer correspondence. This technique is described in (Robbins et al., 2004), and is also implemented in the thickness methods described in (Fischl et al., 1999) and (Thompson et al., 2005). The first two use cortical depth or convexity on the sphere as the feature whose cross-correlation is to be maximized, whereas (Thompson et al., 2005) uses manually identified sulcal ribbons to drive the deformation.

Once vertices are brought into correspondence a statistical test can be performed at each node. By far the most commonly used is the general linear model, which tests the effect of a series of explanatory variables on the dependent variable - cortical thickness. One is not limited to linear models, however; non-linear tests, discriminants, regressions, and more are all available.

Hypothesis tests in statistics depend on a confidence region for accepting or rejecting the null hypothesis. When testing cortical thickness maps these tests are conducted multiple times: 40962 in the case of the analyses to follow. By simple probability a confidence region of 5% (the default $\alpha < 0.05$ used by convention in most statistical tests) will result in about 2050 rejections of the null hypothesis even if there is just random noise in the data.

There are five commonly used ways of controlling for this type of false positive when considering multiple comparisons:

1. setting a more conservative rejection region.
2. Bonferroni's correction.
3. Random Field Theory.
4. False Discovery Rate.
5. Permutation tests.

Setting a more conservative rejection region involves a different alpha for the significance test. This does not eliminate the problem, it simply reduces the scope. For example, by setting $\alpha < 0.001$ one would encounter on average only 41 false positives in the image. Finding the right balance in selecting thresholds in this fashion is, however, somewhat arbitrary.

Bonferroni's correction is one of the most standard ways of accounting for multiple comparisons. It is very simple to compute: one divides the desired α by the number of comparisons to be performed. In order to obtain $\alpha < 0.05$ corrected using Bonferroni's method one would thus need $\alpha < 0.0000012$ uncorrected. This is clearly a very stringent thresholding technique, liable to remove true positives along with false positives.

Random Field Theory is an extension of Bonferroni's correction. It makes the assumption, however, that not all 40962 points are independent, but that the actual number of independent units or resels is much smaller due to the imperfect alignment between surfaces and the blurring used (Worsley et al., 1992; Worsley et al., 1996; Worsley et al., 1999; Worsley et al., 2004). In one example described in the population simulation paper (Lerch and Evans, 2005) there were 450 resels in a dataset blurred with a 30mm diffusion smoothing kernel.

The False Discovery Rate (FDR) takes a different approach to thresholding. Whereas the previous tests used a strong correction for multiple comparisons, FDR is more lenient. Given an $\alpha < 0.05$, for example, random field theory states that 19 times out of 20 there will be no false positive in the resulting statistics threshold. FDR, on the other hand, would state that of all the results shown in an image, 5% will on average be false positives. One thus gains increased sensitivity while losing the same level of certainty as provided with random field theory (Genovese et

al., 2002).

The last method available is permutation testing. Permutation testing is a non-parametric technique used to determine the exact distribution of a particular statistical test within a given sample, hence providing the p-value of the test result. The statistical test is repeated multiple times, exchanging the association between the subject's cortical thickness data and their group label (such as patient or control) each time. This is ideally repeated so that all possible combinations are reproduced. If the sample is too large for that to be reasonable, the distribution can be approximated by taking a large number of random permutations. The maximum accuracy across all 40,962 vertices is kept for each one of the permutation tests; these then form the distribution of the accuracy test-statistic. A p-value of 0.05, for example, is then computed by sorting the maximum accuracy values and taking the fifth percentile from the top. This provides strong control over experiment-wise Type I error. One of the great advantages of permutation tests is that they are applicable even in the absence of theoretical solutions using random field theory (Nichols and Holmes, 2002; Nichols and Hayasaka, 2003).

A final important consideration for statistical analysis of cortical thickness is the issue of whether to scale thickness values based on cerebral volume. Voxel based morphometry, for example, always scales its density values based on the transformation to a standard template. There are reasons to believe, however, that the thickness of the cortex does not scale linearly with brain size. This is true across evolution (see figure 2.6) as well as within human populations, as evidenced by the lack of correlation between cerebral volume and cortical thickness (Ad-Dabbagh et al., 2005). Scaling cortical thickness values to a common space is thus likely to introduce distortions in the data rather than correct for a possible confound, as smaller brains would feature exaggerated cortical thickness

values after scaling.

Chapter 3: Validation Study

3.1. Preface

Cortical thickness measurements from MRI are relatively new; one of the first goals was thus to validate our methodology. A previous study by Kabani et al. had addressed the accuracy of thickness measurements (Kabani et al., 2001). This work should ideally be expanded by looking at high resolution post-mortem data long with MRIs acquired in the same subject. Unfortunately such data was not available. Instead the focus was placed on addressing the precision and reproducibility of cortical thickness.

There are multiple ways of defining cortical thickness given the existence of two surfaces. There are also choices to be made in the width of the blurring kernel to use, or whether to use one at all. Multiple comparisons also must be accounted for in the statistical analysis; here too multiple choices are at ones disposal. A way to address all these questions needed to be found.

One of the most common uses of brain imaging techniques is the comparison of morphological measures with external information about the

subjects, be they behavioural (e.g. cognitive performance), clinical (e.g. neurological deficit score), or demographic (e.g. age). In order to test the efficacy of cortical thickness for such usage a simulated population study was created by artificially inducing anatomical change in one part of the study sample. The different cortical thickness metrics were compared using different blurring kernels and varying statistical thresholds in order to test for the best combination of parameters that can recover the induced change. Repeated scans of a single subject were also used to test the robustness of the different thickness metrics.

The results in this paper established that the metric which defines thickness as the distance between the inside and outside cortical surfaces, t_{ink} , had the highest precision. Smoothing the thickness maps along the surface was also clearly shown to be necessary, a kernel of between 30-35 mm found optimal in this simulation. The tradeoff between different corrections for multiple comparisons was illustrated, with the more stringent random field theory indeed removing all false positives, but the more lenient false discovery rate capturing a greater number of true positives. A power analysis found that, given two groups of 25 subjects each, a 0.6mm (15%) change in cortical thickness can be found at a significance level of 0.05 (corrected for multiple comparisons using Random Field Theory).

The optimal parameters for cortical thickness analysis were established for the remaining studies used in this thesis and for various collaborations with external scientists. The results in this simulation validation paper also gave confidence in the overall methodology, showing an overall sensitivity of 0.93 while maintaining 100% specificity.

Cortical Thickness Analysis examined through power analysis and a population simulation.

Jason P. Lerch and Alan C. Evans

McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill
University

Published in NeuroImage, 24(1), p. 163-73, 2005

3.2. Abstract

We have previously developed a procedure for measuring the thickness of cerebral cortex over the whole brain using 3D MRI data and a fully-automated surface-extraction algorithm (ASP). This paper examines the precision of this algorithm, its optimal performance parameters and the sensitivity of the method to subtle, focal changes in cortical thickness.

The precision of cortical thickness measurements was studied using a simulated population study and single subject reproducibility metrics. Cortical thickness was shown to be a reliable method, reaching a sensitivity (probability of a true positive) of 0.93. Six different cortical thickness metrics were compared. The simplest and most precise method measures the distance between corresponding vertices from the white matter to the gray matter surface. Given two groups of 25 subjects, a 0.6mm (15%) change in thickness can be recovered after blurring with a 3D Gaussian kernel (full-width half max = 30mm). Smoothing across the 2D surface manifold also improves precision; in this experiment the optimal kernel size was 30mm.

3.3. Introduction

The measurement of cortical thickness has long been of interest to the neurosciences, starting with the early reconstructions of Brodmann (Brodmann, 1909) and von Economo (Economo and Koskinas, 1925). Recent advances in image processing and image acquisition has allowed for the automatic extraction of cortical thickness from MRI (MacDonald, 1997; Fischl and Dale, 2000; MacDonald et al., 2000). This paper investigates and summarizes current methodology and evaluates the

power and sensitivity of the different techniques.

The study of the morphometry of the cerebral cortex at the macroscopic level visible in current MRI provides the neurosciences with an opportunity to investigate both normal and abnormal change. Most such investigations use a combination of semi-automatic techniques, usually focusing on the manual delineation of structures of interest, followed by statistical comparisons of volumes (c.f. (Pruessner et al., 2001)). This approach, while clearly quite capable of providing important information about the population under investigation, has several disadvantages. It is very labor intensive, it suffers from intra and inter-rater reliability issues, and, most importantly, it restricts the analysis to predetermined regions of interest.

Several fully automated approaches have also been developed; the most widely used of these is Voxel Based Morphometry (VBM) (Ashburner and Friston, 2000). At its most generic, VBM is the comparison of voxels in a series of linear models. Most methods (c.f. (Paus et al., 1999; Ashburner and Friston, 2000; Baron et al., 2001)) employ a standard set of image processing steps involving linear registration, tissue classification, and creation of “voxel density” maps representing tissue concentration in a local neighborhood. The usual end result is an image that contains regions that have significantly increasing or decreasing signal that correlates with some independent neurobiological parameter. This latter parameter may be just a categorical difference between two groups, e.g. separated by disease status or gender, or, more generally, will be a continuous variable, such as age or behavioral performance, in which case a regression of image signal against that variable is plotted at each voxel (Wright et al., 1995; Paus et al., 1999).

Cortical thickness analysis is similar to VBM, albeit the analysis is performed at the nodes of a three-dimensional polygonal mesh rather than

on a 3D voxel grid, but it has the advantage of providing a direct quantitative index of cortical morphology. The metric captures the distance between the white matter surface and the gray-CSF intersection according to some geometric definition; the output is a scalar value measured in millimeters. The regression slope at each vertex across the cortex in a statistical analysis is meaningful: not only can one determine that cortical thickness is significantly different between groups, but one can also measure that difference. This naturally leads to the ability to define clinical as well as statistical significance.

The use of cortical thickness analysis in MRI studies is relatively new, with only a small number of studies published on the methodology (Meyer et al., 1996; MacDonald, 1997; Zeng et al., 1999; Fischl and Dale, 2000; Jones et al., 2000; MacDonald et al., 2000; Miller et al., 2000; Duygu Tosun et al., 2001; Kabani et al., 2001; Yezzi and Prince, 2003), and even fewer on normal or abnormal populations (Fischl and Dale, 2000; MacDonald et al., 2000; Rosas et al., 2002). This is due to the difficult nature of extracting the inner and outer surfaces of the cerebral cortex at the limited resolution provided by today's MRI machines (usually one cubic millimeter), where the fine details of sulcal anatomy are often obscured by the partial volume effect. Moreover, manual delineation of cortical thickness is very difficult (whether from MRI or post-mortem samples) due to the necessity of creating a correct cut or slice plane perpendicular to the surfaces.

Defining cortical thickness, even when models of the inner and outer surfaces are present, is not trivial. Cortical thickness is a distance metric but there are multiple ways of defining corresponding points on the two surfaces between which that distance is to be measured. Moreover, the distance need not be measured in a straight line, but can be the result of a

more complicated equation, such as fluid flow lines.

This paper examines the power of cortical thickness as an analysis tool; it compares the various definitions of cortical thickness proposed in the literature; the effect of different size blurring kernels; and analyzes the effect of correcting for multiple comparisons. These studies were performed using a simulated population study where the true difference between the two groups is artificially induced and therefore known. Furthermore, repeat scans of a single subject will be used to examine the variability inherent in the different cortical thickness metrics and make a first attempt at defining the power of the method.

Rather than addressing accuracy we focus on the question of precision. The distinction between the two is subtle but crucial:

Accuracy: The ability of a metric to capture the correct distance between the pial and white matter surfaces, as defined by anatomical criteria and validated through manual measurements or accurate MR simulations.

Precision: The ability of a metric to provide reproducible results from repeated estimations and thereby differentiate between two measures known to be different.

A metric can therefore be declared most accurate through the comparison of automated and manual measurements, or through directly simulating a cortical sheet with a known thickness and validating each metric against such a construct. Each of these methods has to overcome significant challenges. Manual measurements of cortical thickness are tremendously difficult to undertake, being highly dependent on a perfectly perpendicular cutting angle. Moreover, even using the exact same post-mortem slice, individual raters can easily differ by over 0.5mm at any one location due to the blurred cortical boundary at the white matter surface (Economo and Koskinas, 1925). Furthermore, the traditional thickness measurements

derived from post-mortem slices are dependent on a straight-line measurement of cortical thickness, as these measurements are always carried out in two dimensions. Accurate validation of MR measurements of cortical thickness thus requires a three-dimensional reconstruction of high resolution post-mortem data. The alternative of validation through construction of a cortical sheet with known thickness is very attractive, but difficult to use in comparing thickness metrics. The reason is that the construction of such a cortex would be dependent on a pre-existing definition of cortical thickness, and would therefore be biased towards that metric from the beginning. Furthermore, accurately simulating MRI from polygonal models of the cortex has to first address the issue of correctly incorporating partial volume into the tissue model. We plan to address the question of accuracy both through the use of a simulator as well as high resolution post-mortem reconstructions; that, however, is the subject of future work. This paper examines the precision of cortical thickness analysis through the use of repeated acquisitions of the same subject as well as a population simulation.

3.4. Methodology

3.4.1. Measuring Cortical Thickness

Measuring cortical thickness is a complex process involving multiple image processing steps. The native data, usually consisting of a T1 MRI per subject but optionally includes any number of modalities. These one or more images of the brain parenchyma are used to provide an anatomic label for each voxel (typically this means classification into gray matter, white matter, CSF, and non-brain classes). Prior to this classification step, intensity and spatial normalization must be performed. Intensity correction

for non-uniformity is obtained using the N3 algorithm (Sled et al., 1998); spatial normalization is done to the ICBM 152 average using a nine parameter linear registration (Collins et al., 1994). If there is more than one image per subject, any additional MRIs are registered to the first MRI using mutual information registration (Collins et al., 1994).

Each subject's brain is classified into white matter, gray matter, CSF, and background using all available imaging modalities and a classifier trained by stereotaxic space probability maps (Kollokian, 1996; Zijdenbos et al., 2002). These probability maps were created from 305 classified samples; 1000 points per tissue class were randomly chosen from areas having a greater than 90% chance of being of the correct tissue type in that location in stereotaxic space. Prior to classification the training tag points are pruned for each individual subject to remove any outliers.

The inner and outer cortical surfaces are then extracted using the ASP algorithm (MacDonald et al., 2000). The essence of ASP is the creation of simple (non self-intersecting) surfaces with spherical topologies using deformable models. The classified volume is taken as input, and the process begins with the deformation towards the white matter surface. Along with the image information, $T_{boundary-dist}$, several model terms are used to constrain the fit, and self-intersection is explicitly prohibited. The model terms are: $T_{stretch}$, constraining distances between neighboring vertices; T_{bend} constraining deviation from model shape; and $T_{self-proximity}$, constraining the proximity of pairs of non-adjacent polygons. The grey matter surface is obtained using the same constrains as listed above along with $T_{surface-surface}$, preventing the two surface from coming within a certain distance of each other, and $T_{vertex-vertex}$, which penalizes corresponding vertices as they deviate from an ideal distance. This last constraint allows for sulcal penetration of the grey matter surface even when the sulcus in question

has been obscured due to partial volume blurring of the CSF space.

The creation of the two surfaces then allows for the measuring of cortical thickness using various distance metrics. They are summarized in Table 3.1 and described in more detail below.

The first method is t_{link} . It is conceptually very simple, measuring the distance between linked nodes on the inner and outer surface. The correspondence between such nodes is created by the expansion of the outer surface from the inner surface, each polyhedron having the same topology and number of vertices. This method is inherently very robust: the model constraints that govern the expansion will guarantee low variability, minimizing large errors and outliers. However, there is no guarantee that the linked method will produce a distance measure corresponding to what an anatomist would chose.

Name	Description	Citation
tlink	Distance between linked nodes	(MacDonald et al., 2000)
tnear	Distance to nearest node	(MacDonald et al., 2000)
tnormal	Distance along surface normal	(MacDonald et al., 2000)
tlayered-normal	Distance along iteratively computed normal	NA
taverage-near	Distance to nearest node computed twice, averaged	(Fischl and Dale, 2000)
tlaplace	Distance solved using Laplace's equation	(Jones et al., 2000)

Table 3.1: description of cortical thickness metrics used in this study.

The t_{near} method performs a simple search across the opposite surface and picks the vertex that is the shortest (Euclidian) distance away. While intuitive, this method has the potential for gross errors, such as jumping across gyri, as there is no guarantee that the nearest point is the

anatomically most sensible one. t_{normal} constrains the point that can be found to lie along the intersection of the surface normal. $t_{layered-normal}$ creates a series of nested surfaces first by a process of weighted averages of the inner and outer surfaces evaluated between corresponding nodes on each surface. The surface normal is computed at each of these nested surfaces and then averaged, thereby producing a constraint for finding the corresponding point on the opposite surface that is less prone to producing outliers than the simple surface normal. $t_{average-near}$, first published in (Fischl and Dale, 2000) computes t_{near} twice, once from the outside to the inside surface, and once from the inside to the outside surface. These two values are then averaged to produce a thickness value at that node.

The last method is $t_{laplace}$, first published in (Jones et al., 2000) and reimplemented locally. Two boundaries, the white matter volume and the extra-cortical volume as defined by the two surfaces extracted using the ASP algorithm, are defined and fixed. Laplace's equation, shown in Equation 3.1, is then iteratively solved across the entire volume using the Jacobi method. Iterations continue until the change across each iteration becomes smaller than a preset threshold. Gradients are then computed using two point differences, and the gradient vectors normalized to produce tangent vector fields. Streamlines are computed at every voxel in the cortical volume by integrating towards each of the boundaries using Euler's method, and the two path-lengths added together to produce a thickness value at that point.

$$\Delta^2\Psi = \frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}$$

Equation 3.1: Laplace's Equation

The final step before analysis is blurring the thickness data. A surface

based diffusion smoothing kernel is used, which generalizes Gaussian kernel smoothing and makes it applicable to any arbitrary curved surface (Chung et al., 2002). It has to be remembered that this blurring kernel used on the surface has a different meaning from the standard volumetric kernels since surface curvature is followed as illustrated in figure 3.1.

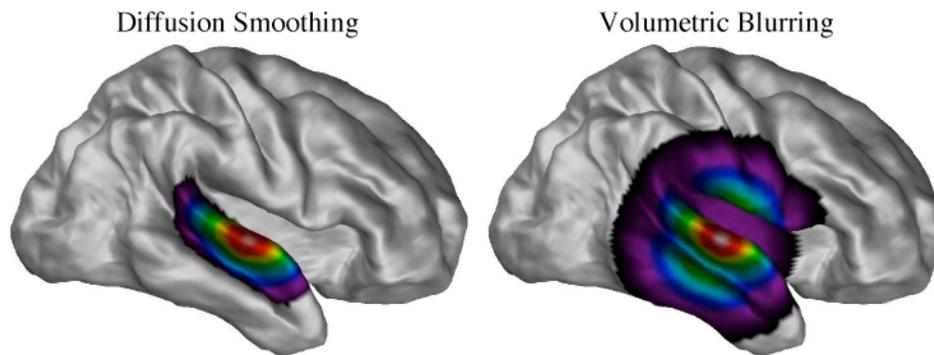


Figure 3.1: An illustration of the difference between the geometry preserving diffusion smoothing blurring over a 2D surface manifold and the more commonly-employed 3D volumetric blurring kernels. The FWHM was set at 30mm in both cases. One can see how anatomically disparate areas such as the inferior motor and sensorimotor areas are influenced by the volumetric kernel but not by diffusion smoothing.

The arguments for blurring are fourfold:

1. By the central limit theorem, smoothing has the effect of rendering the data more normally distributed, thereby increasing the validity of statistical tests.
2. It reduces the impact of imperfect alignment between cortices by replacing individual vertex values with neighborhood averages.
3. It reduces noise in the measurement of cortical thickness. The fact that the average cortex is only a few voxels thick leads to some variability in thickness measures due to the inadequate MRI sampling. By

averaging neighboring vertices in the diffusion smoothing operation this noise is reduced.

4. Since blurring increases the interdependence of the neighboring vertices it also reduces the number of comparisons to be controlled for using Random Field Theory (see section 3.4.2).

These improvements in signal-to-noise and statistical normality are of course obtained at the cost of a degradation in image resolution in the classical image analysis trade-off. The choice of optimal blurring kernel width is discussed below.

3.4.2. **Statistical Analysis**

Once the thickness maps have been generated and optionally smoothed for each subject, statistical tests can be performed. A linear model is applied separately at each vertex t : $Y(t) = X\beta(t) + \varepsilon(t)$, where $Y(t)$ is the measure of cortical thickness, X is the matrix of explanatory variables, β represents the slope to be estimated for each explanatory variables, and $\varepsilon(t)$ is the normally distributed error. A series of statistical tests, such as a t , F , or adjusted R^2 values, can be applied. The regression slope, β , can also be plotted at every vertex. The ability to derive meaning out of the regression slope is one of the key strengths of cortical thickness analysis, since that slope can be expressed as millimeters change. Accurate estimation and interpretation of the slope will be influenced by the kernel used, since blurring causes thickness to be estimated across areas of cortex rather than individual vertices. There is thus potential for underestimation of local change should the kernel size be too large.

The challenge, also faced by VBM and functional imaging techniques, is to correct for the multiple comparisons that are undertaken in the analysis. For the cortical thickness analysis experiments described here the number

of nodes in the cortical mesh resulting from ASP is 40962. The most common method to control for multiple testing is to adjust the required significance threshold such that Type I error is controlled for across the entire brain. The by now default method for such correction uses Random Field Theory thresholding (Worsley et al., 1992; Worsley et al., 1996), which takes the smoothness of the data into account in determining the number of Resels, thereby reducing the effective number of tests to be controlled for. The resulting threshold states that, in an area where the null hypothesis is true, the chance of rejecting one or more of the tests is less than or equal to α , the preset level of confidence. This type of control is quite stringent, providing the benefit that vertices where the null hypothesis is rejected are highly likely to be true positives, but also leaves a high likelihood of false negatives. Implementation of Random Field Thresholding on the surface is made more complex by the non-isotropic nature of the images. The solution is to estimate the effective FWHM (eFWHM) (determined through the normalized residuals of the fitted model) along the edge of each vertex, and to warp (in a statistical sense) the coordinates of each vertex so that the eFWHM is approximately constant (Worsley et al., 1999).

3.4.3. Analyzing the Variance

In order to quantify the normal variance expected for cortical thickness estimation, 19 different T1 MRIs with 1mm isotropic sampling were acquired from the same subject over a short period of time (Holmes et al., 1998). The cortical thickness pipeline as described above was run on each acquisition. All the different metrics were used, each blurred with different sized kernels. Means and standard deviations were computed both at every vertex along with each subject's mean thickness. A normalized standard deviation map was produced by dividing the standard deviation

by the mean at each vertex. The procedure was then repeated for 25 normal subjects taken from the ICBM database (Mazziotta et al., 2001) in order to capture the variance inherent in a normal population.

Power calculations were performed on both sets of data. The standard deviation of cortical thickness was modeled at every vertex to answer the questions:

1. What N is needed in order to recapture a change of x millimeters?
2. Given two equal groups of N subjects each, what change can be recaptured?

For both cases the significance level (type I error probability) was set to 0.05 and the power (1 – type II error probability) at 0.95 multiple comparisons corrected for with Random Field Theory. The interpretation of the results is dependent on the blurring kernel: in the unblurred data we are looking at the sample size needed to recapture a change of a certain magnitude, where that change is isolated from any neighbors. The addition of blurring, however, modifies the value at each vertex to reflect a weighted neighborhood average.

3.4.4. Population Simulation

The goal of many imaging studies in neurology is to assess morphometric differences between two populations. In order to evaluate the utility of cortical thickness analysis in such a scenario an artificial "patient" population was created through induced thinning of the cortex. 50 subjects from the ICBM (Mazziotta et al., 2001) database were taken. All the MRIs were corrected for non-uniformity artifacts, linearly registered into stereotaxic space, and classified into their component tissue types, all as described above. 25 of these subjects were randomly chosen and designated as patients. In this group the MRIs were segmented using

ANIMAL (Collins et al., 1995), and the right superior temporal gyrus (rSTG) arbitrarily chosen for thinning. Thinning was induced through a six neighbor dilation of the white matter into the rSTG as defined by ANIMAL (see Figure 3.2). The cortical surfaces were then fit on all subjects and cortical thickness measured with each of the available metrics. Statistical analysis was performed at every vertex to assess if the induced change can be recaptured. Prior to the induced thinning the two populations were compared, and no statistically significant differences were found ($p > 0.3$).



Figure 3.2: An illustration of the rSTG thinning procedure. Green is gray matter, white is white matter, and gray is the gray matter that was removed in the "patient" population.

In order to evaluate the performance of the different metrics and blurring kernels the standard epidemiological terms true positives, false positives, true negatives, and false negatives were defined (see Table 3.2). The definition of truth for the purposes of this simulation was based on the probability map of the rSTG. Since the rSTG is the site of the induced thinning, it should also be the area exhibiting significant results. The rSTG was defined individually for each subject, however, and is not perfectly aligned in stereotaxic space. Hence a probability map was created for the rSTG in which each vertex value represented the proportion of subjects for whom that vertex was labeled as rSTG. Truth, for the purposes of the simulation experiment, was thus defined as a statistically significant vertices that intersect the rSTG probability map.

		Test	
		+	-
True	+	a	b
State	-	c	d

Sensitivity: $\frac{a}{a+b}$ (Probability of a true positive)

Specificity: $\frac{d}{c+d}$ (Probability of a true negative)

True Positive	Positive Test where true state is positive
True Negative	Negative Test where true state is negative
Sensitivity	Probability of a true positive
Specificity	Probability of a true negative

Table 3.2: Epidemiological statistics used in the study.

3.5. Results

3.5.1. Variability

The standard deviation of cortical thickness was measured at each point on the cortex in repeated scans across one subject as well as across a normal population. The results across metrics are summarized in table 3.3.

	19 scans of Same Subject					25 Normal Subjects				
	Mean	Std	Nrm. std	δ	n	Mean	Std	Nrm. std	δ	n
t_{near}	2.46	0.21	0.09	21%	20	2.38	0.36	0.15	38%	52
t_{normal}	4.39	0.50	0.11	28%	31	4.30	0.70	0.16	41%	59
$t_{layered-normal}$	3.99	0.28	0.07	18%	15	3.95	0.47	0.12	30%	34
$t_{average-near}$	2.53	0.18	0.07	18%	15	2.48	0.31	0.13	31%	37
$t_{laplace}$	3.71	0.27	0.07	18%	15	3.63	0.44	0.12	30%	35
t_{link}	3.93	0.22	0.06	14%	11	3.88	0.35	0.09	22%	21

Table 3.3: Where δ is the minimum percentage change that can be recaptured when $n = 25$, and n is the change that can be recaptured when $\delta = 25\%$. Nrm. Std is the normalized standard deviation. All comparisons made using 30mm blurring kernel, $p=0.05$ after correction for multiple comparisons using random field theory ($t=4.67$).

Variability differs across different cortical thickness metrics.

t_{normal} has the highest standard deviation, $t_{average-near}$ the lowest. Due to the different definitions for the thickness metrics, the mean thickness is quite variable across the different methods, ranging from a high of 4.39 mm in t_{normal} to a low of 2.46 mm in t_{near} (which by definition must have the lowest value). After normalizing to account for these differences by dividing the standard deviation with mean thickness, t_{normal} once again has the worst performance, t_{link} the best. The same pattern emerges whether these metrics are investigated across repeat scans of one subject, or across 25 different young normals (see table 3.3).

Variability is not uniform across the cortex.

Variance in thickness differs dependent on location in the cortex, being highest along the superior aspects of the central sulcus, lowest in the

prefrontal cortex; see figure 3.3, which shows the standard deviation of cortical thickness at each vertex in both 25 normal subject as well as the 19 repeated scans of a single subject. Standard deviation is marginally related to cortical thickness at any vertex ($R^2=0.0005$). Normalizing the standard deviation increases the effect ($R^2=0.06$). Thinner cortical areas are thus more variable than their thicker counterparts, though thickness itself only explains a small part of the heterogeneity of variability. While the overall magnitude of the deviation varies across metrics, the spatial pattern is similar. As seen in the right hand column of figure 3.3, the pattern of variability remains stable across blurring kernels, even as the overall variability decreases with increased smoothing. The exception to this rule is that at high blurring kernels the representation of non-cortical areas (such as the brain-stem cut) where cortical thickness measurements are meaningless and therefore highly variable begins to influence the standard deviation of their cortical neighbors.

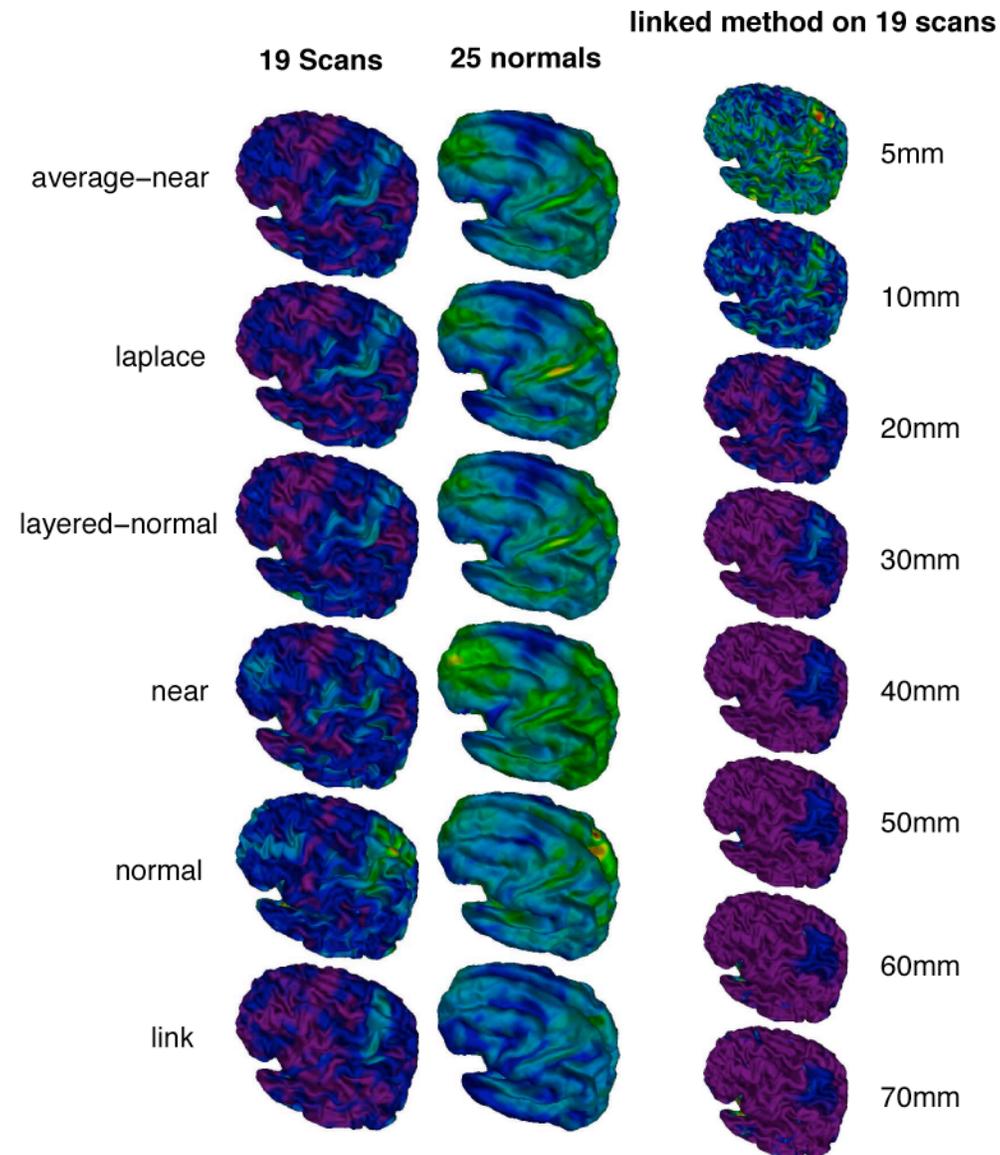


Figure 3.3: Normalized Standard Deviation across different thickness metrics and blurring kernels. The two leftmost columns show the normalized standard deviation across different metrics, first in 19 scans taken of the same subject, second across 25 normal subjects. The last column shows the change in normalized standard deviation across blurring kernels, using the tlink method and 19 scans of the same subject.

Variability declines with increased blurring – up to a point.

Normalized standard deviation ($\frac{std}{mean}$) declines with blurring up to a 40 - 50 mm kernel, after which it increases again, as shown in figure 3.4. This increase appears to be due to an increased spatial representation of the non-cortical areas. Normalized standard deviation in cortical areas spatially removed from non-cortical regions continues to decline with increased blurring.

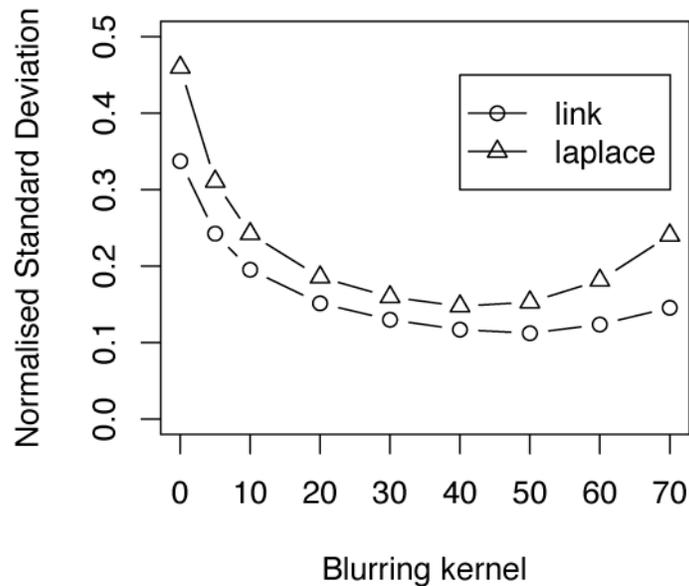


Figure 3.4: Graph of normalized standard deviation across blurring kernels, using the t_{link} and $t_{laplace}$ methods computed in 19 scans of the same subject. There is a similar pattern to the one seen in figure 3.6 – a decrease up to a kernel size of 40mm followed by increasing standard deviations, once again indicating that the optimal blurring kernel for minimizing variance is in the 35 – 40mm range.

Power calculations.

Power calculations for the different metrics are given in table 3.3 and

illustrated for the tlink metric in figure 3.5. Given two groups of more than 100 subjects each a change of 0.29mm can be recovered. Conversely, given two small groups of 20 subjects each a change of 1mm would reach statistical significance. These numbers assume Random Field Theory corrections for multiple comparisons. Since power calculations are dependent on variance the exact change required to reach significance is not uniform across the cortex, being highest in the superior aspects of the central sulcus, lowest in the prefrontal areas (see figure 3.3). Moreover, statistical significance is dependent on the eFWHM, which in turn is influenced by the amount of blurring. and the resulting resels and statistical thresholds are plotted in figure 3.6. Threshold and resels are minimized at a kernel size of 35mm.

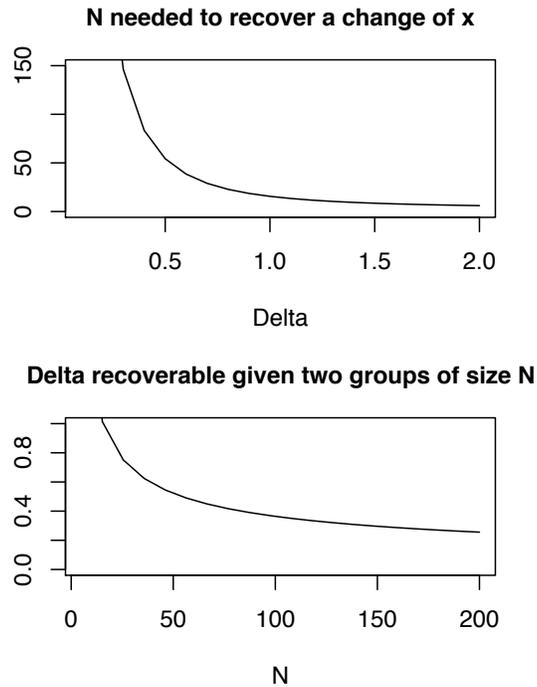


Figure 3.5: An illustration of the power of the t_{link} method, standard deviation computed after application of a 30mm blurring kernel and using a standard deviation of 0.27mm (computed from the single subject variance shown in table 3.3). The top graph shows the number of subjects needed to recover a thickness change of size Delta. The bottom graph shows the size delta that can be recovered given two equal groups of subjects of size N. One can see that, given an $N > 100$, a 0.35mm change can be recovered, two small groups ($N = 20$) a 1mm change would reach significance.

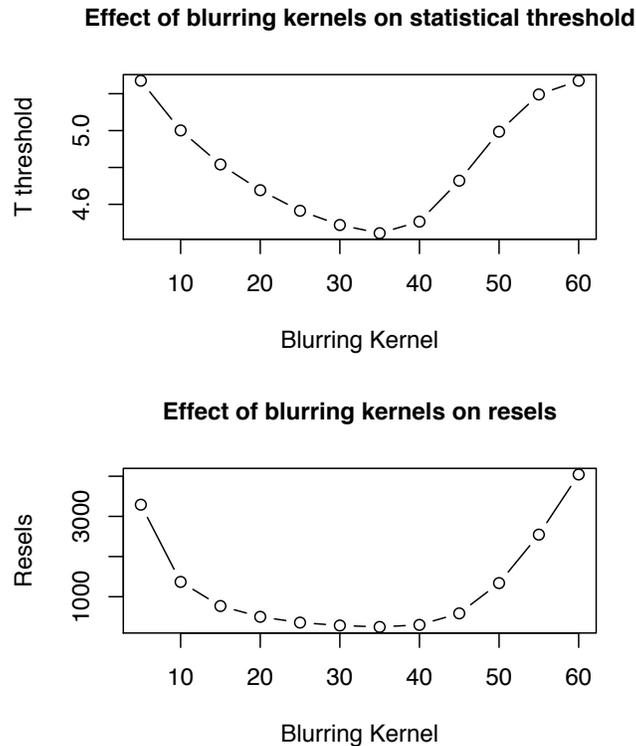


Figure 3.6: The effect of blurring kernels on resels and the t-statistics threshold using Random Field Theory. The smoothness of the thickness maps as shown by the number of resels decreases up until a 35mm kernel, increases again thereafter. This increase could be due to increasing influence of non-cortical areas (such as the corpus callosum and brain-stem cut) included in the mesh. The maximum kernel size that should be used is thus 35mm.

3.5.2. Population Simulation

The simulation was modeled after a comparison between two groups ("controls" and "patients"), where the patient group had their rSTG artificially thinned.

30 millimeters is the optimal kernel FWHM

An evaluation of different blurring kernels reveals the classic trade-off: increasing kernel size improves sensitivity but also decreases the ability to accurately estimate the regression slope.

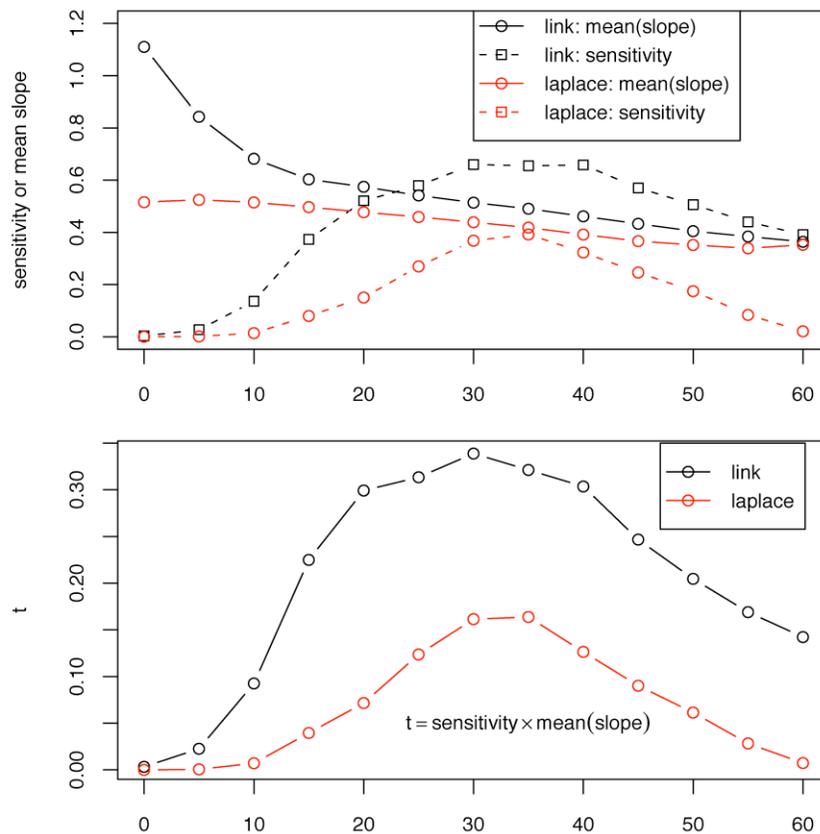


Figure 3.7: The effect of increasing blurring kernels as a function of sensitivity and mean regression slope. The maxima of the function $t = \text{sensitivity} \times \text{mean(slope)}$ is to be found at 30mm, indicating that this is the optimal blurring kernel size for this study. All calculations were made under the assumption that truth is the rSTG probability map thresholded at 1%.

As shown in figure 3.7, sensitivity increases up to a blurring kernel of 35mm, and then declines. The mean slope, which should approach one millimeter since one layer of voxels was removed in the rSTG, declines

steadily with increasing kernel sizes. Optimizing the tradeoff between estimation of slope and sensitivity is performed with the equation , whose maxima is found at 30mm. The same pattern hold for both t_{link} and $t_{laplace}$, though the decline in the regression slope is much more noticeable for t_{link} .

t_{link} is the most sensitive method

The six different cortical thickness metrics were all compared at 30mm blurring. The performances were compared by evaluating the sensitivity (ability to recover true change) of the different metrics at increasing truth thresholds. Figure 3.8 shows the thresholded t-statistics maps superimposed onto a sphere along with the probability map of the rSTG that represents the vertices to be recaptured. Visual inspection of these maps shows that all methods show significant results in the correct anatomical region, and that t_{link} provides the most convincing overlap. Figure 3.9 describes the same results by graphing sensitivity against probability map threshold and statistical threshold. t_{link} is clearly the most sensitive method. $t_{laplace}$ is second, except at high thresholds where t_{normal} surpasses it. This figure also illustrates that regions of greater overlap of rSTG are easier to recover. Four of the six metrics show a decline at high probability map thresholds. The origin of this change is unknown, but probably results from an interaction between the shape of the rSTG, the blurring kernel, and the geometric definitions of the thickness metrics. When the truth threshold is set at 50% and the statistical threshold is varied (as shown in the bottom part of figure 3.9), t_{link} is again the most sensitive.

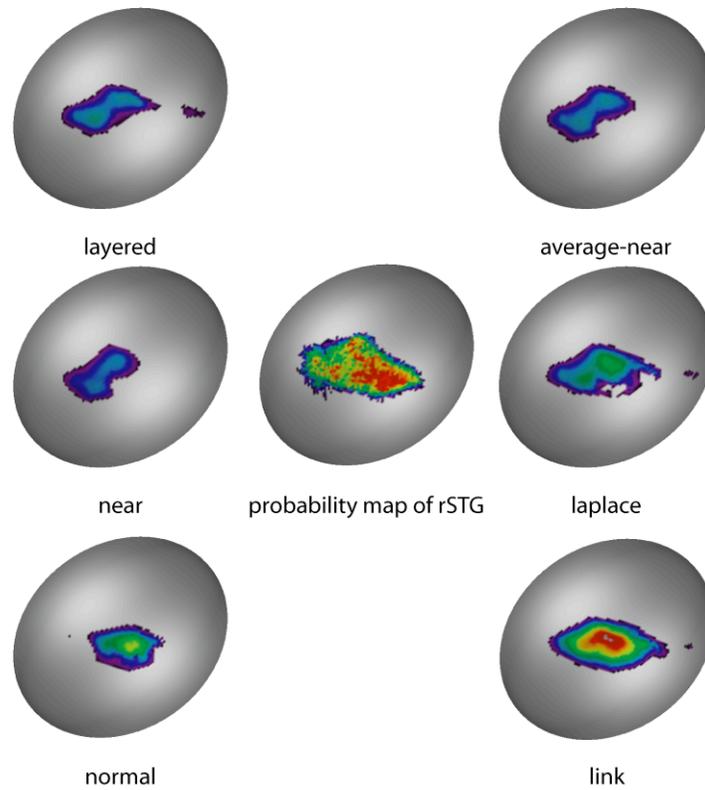


Figure 3.8: Results of each of the different metrics at 30mm blurring, thresholded at $t \geq 2.5$. The results are displayed on a sphere; the sphere in the centre shows the probability map of the rSTG, and thus the area of the cortex to be recaptured. Qualitative assessment of the shapes of the t-statistics maps indicate that the t_{link} and $t_{laplace}$ produce the closest match to the probability map.

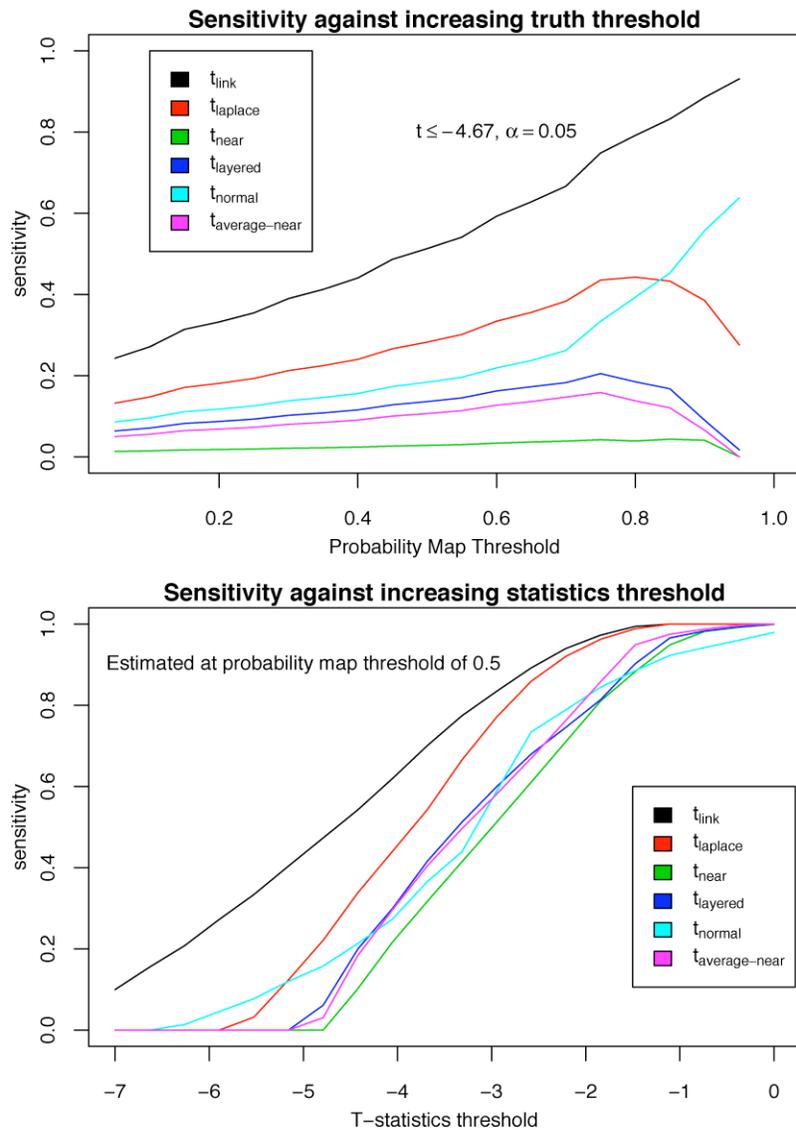


Figure 3.9: Sensitivity of the six different cortical thickness metrics at 30 mm blurring, graphed against ever more stringent statistical thresholds (bottom panel) and percentage overlap of the rSTG (top panel). The superiority of the t_{link} metric is clearly noticeable, attaining both a higher sensitivity across different thresholds of the rSTG probability map as well as higher t-statistics values compared to the other metrics.

Controlling for multiple comparisons

The results for any metric can be decomposed into its component statistical measurements such as true positives, false positives, and false negatives, and their respective values evaluated across changing statistical thresholds. This is shown for the tlink metric in figure 3.10. False positives (FP) decline rapidly with increasing thresholds, true positives (TP) and false negatives (FN) decrease and increase linearly. The relationship can be described in the thresholding index $index = \frac{TP}{FP+FN}$, which maximizes true positives while simultaneously minimizing false positives and false negatives. This function has a maxima at $t=-3.3$, considerably below the random field threshold of $t=-4.67$.

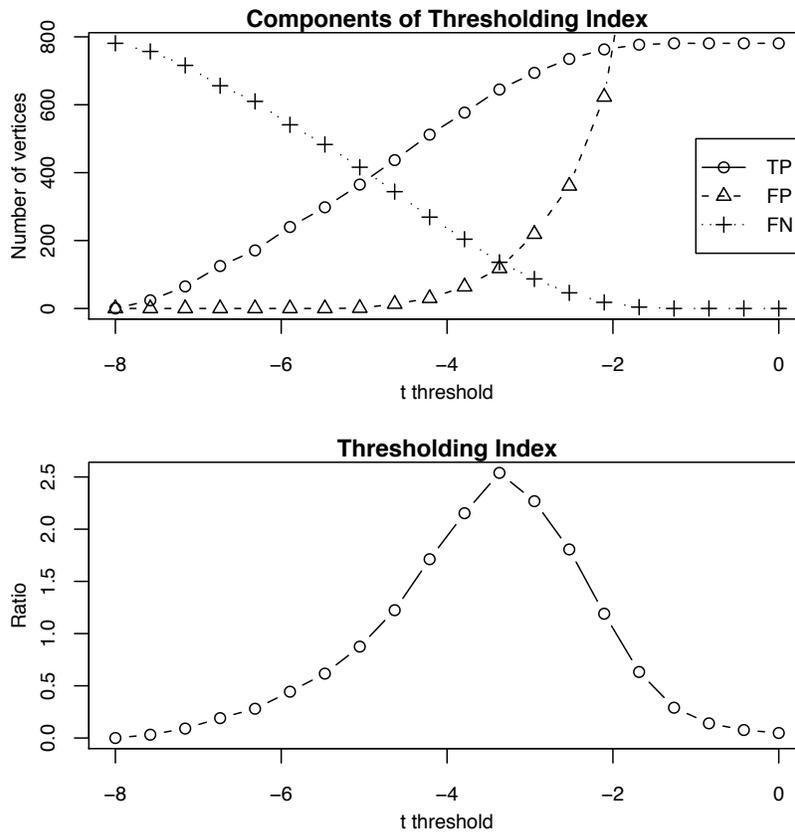


Figure 3.10: The top graph shows the individual components that make up the thresholding index shown in the bottom graph. This lower graph was generated at a rSTG threshold of 0.58, chosen since that is where the ratio reaches its maximum (2.54) when searching across all possible rSTG thresholds. Definitions: TP = True Positives, FP = False Positives, FN = False Negatives.

3.6. Discussion

The goal of this study was to examine a fully automated cortical thickness analysis system, to differentiate between multiple cortical thickness metrics, and to investigate the ability of cortical thickness to differentiate between different populations. Accuracy of the different metrics was never

under investigation. Instead, we addressed the question of precision.

Addressing precision is important in its own right, as it can compare the relative ability of different metrics in differentiating between groups of variable thicknesses. For example, it is conceivable that the t_{normal} metric is the most accurate, yet its high variability would reduce its value in population studies. The precision of a metric measures its usefulness in the multiple subject studies so often undertaken in brain-imaging; and while high precision with low accuracy is certainly undesirable, so is high accuracy with low precision.

3.6.1. Comparing the different metrics.

We compared six different metrics from three different labs. Each metric was evaluated in terms of variance across a population or a single subject as well as performance in the population simulation. The results from these two tests are related, with normalized standard deviation proving a significant predictor of sensitivity in the population simulation ($t=-3.48$, $p=0.026$, $df=5$, $R^2=0.75$). One conclusion is that future thickness methods development should keep the goal of minimizing variability clearly in sight, and if necessary increase the complexity of the algorithm to achieve that goal.

The different metrics can be ordered from best to worst by comparing their performance in the variability analysis (see Section 3.5.1) and the population simulation (see Section 3.5.2). The results of these two analyses can be summarized by dividing the sensitivity of a metric by the percentage change that can be recovered given two groups of 25. The ranking is the following:

1. t_{link}

2. t_{aplace}
3. t_{normal}
4. $t_{\text{layered-normal}}$
5. $t_{\text{average-near}}$
6. t_{near}

These metrics were all compared using the default parameters; it is conceivable that tuning the image analysis pipeline in different ways will change the exact results produced in this study. An especially important topic to pursue for further study is the impact that closer spacing of vertices in the cortical meshes has on the variability of thickness metrics (i.e. increasing our vertex count).

Six methods from three different labs were studied; there are, however, others described in the literature on measuring cortical thickness from MRI. The six methods were chosen for their ability to be easily incorporated into our image processing pipeline; any methods that inherently rely on different cortical tessellations, such as (Zeng et al., 1999; Miller et al., 2000) were therefore excluded. Other methods, such as the one introduced by (Yezzi and Prince, 2003) extend methods described and tested in this paper; their improvements might very well lead to improved results in our simulation. We do believe that the population simulation framework used in this study is an elegant way to compare different cortical thickness metrics in a controlled yet realistic manner. We have therefore made the volumes used in this study available in order to encourage comparisons of different methods to the six metrics described herein. The data can be found at: http://www.bic.mni.mcgill.ca/thickness_population_simulation/.

3.6.2. **Varying variability**

A noticeable trend in studying the variability across the cortex is its regional variation. The most variable areas of the cortex are the pre and post-central gyri, the primary visual areas, and the anterior medial temporal lobes. Two explanations are likely to play a role. First, areas with the thinnest cortex (the motor and visual areas) have high variability. This can be accounted for by the variance induced by the 1 mm sampling in the datasets. Since the sampling stays uniform, but thickness varies, areas with thin cortex are likely to have the highest normalized standard deviation. The second argument relates to the difficulty in segmenting certain areas. The medial temporal lobes, for example, is one of the most challenging for the cortical fitting (Kabani et al., 2001).

3.6.3. **Effects of blurring**

An open question often asked in Voxel Based Morphometry applies to cortical thickness analysis as well: what amount of smoothing is desirable? Our results suggest three conclusions in this regard.

- An increase in the FWHM decreases variability, leading to improved sensitivity up until a kernel size of 35mm (see figure 3.7).
- Increasing FWHM changes the interpretation of the regression slope, potentially underestimating the amount of localized cortical thickness change as described in Section 3.5.2 and shown in figure 3.7.
- Kernels larger than 35mm actually decrease sensitivity (see figure 3.7).

When there is prior information about the extent of the signal (area of thinning) to be detected, the size of the blurring kernel should match the

size of the putative area of change (i.e. matched filter theory, see (Pratt, 1991) for an overview). However, for exploratory searches over the whole brain where there is no prior expectation of signal extent, this concept is meaningless. Tuning of blurring kernel size should thus be driven by the desire to limit the FWHM in order to allow for accurate estimation of μ , while at the same time staying large enough to retain sensitivity. Ultimately, therefore, the size of the FWHM should be driven by the number of subjects in the study: a large n allows for a smaller kernel, which in turn allows for accurate estimation of the amount of local thickness change, whereas small n still needs larger FWHM in order to retain adequate sensitivity. This equation should be balanced by prior hypotheses about the expected area of change.

3.6.4. Thresholding statistical maps

In the implementation described above, 40962 linear models are analyzed - one for each vertex - with every statistical analysis. Multiple comparisons thus have to be corrected for. The prevailing philosophy in brain imaging to date has been to provide stringent control for Type I error, most commonly implemented through applications of Random Field Theory or Bonferroni correction. This stringency, as the population simulation shows, has its costs, as it allows a high percentage of false negatives. More liberal thresholding techniques, such as the False Discovery Rate (Genovese et al., 2002), might prove attractive for exploratory studies using cortical thickness analysis. Incidentally, the thresholding index maxima at $t \leq -3.3$ found in Section 3.5.2 corresponds exactly to a False Discovery Rate q value of 0.05, indicating that this new technique more closely approximates our ideal thresholding index than the Random Field Theory.

3.7. Conclusions

We have shown cortical thickness to be a reliable method, reaching a sensitivity of 0.93. The most precise method is t_{link} . This is due to its ability to minimize variance leading to higher statistical sensitivity. All the metrics had a specificity of 1. While this may seem like a useless index for comparing and contrasting the different metrics, it does indicate a high degree of confidence in any results that are obtained regardless of the metric employed.

Blurring along the surface was shown to be critical, as it significantly increases the sensitivity of cortical thickness analysis. The optimal blurring kernel in our simulation was 30mm (see figure 3.7). An optimum thresholding index which maximizes true positives against both false negatives and false positives was found to lie at $t = 3.3$ (see figure 3.10). Given these optimal parameters and two groups of 25 subjects, a 0.6mm (15%) change in thickness after 30mm blurring can be recovered. Increasing the number of subjects to 100 in each group allows for a 0.29mm (7%) change to be recovered.

In order to validate our methodology a framework was created to capture the precision of the different thickness metrics, and to test the effect of changing parameters for image blurring and statistical thresholding in the analysis pipeline. This general framework can be used to examine future advances in the entire pipeline, such as the impact of different tissue classification methodologies and non-linear alignment techniques. More work is to be done in validating the accuracy of different metrics, and possibly in creating new metrics based on higher resolution anatomical information, which should in turn be evaluated using the precision criteria

illustrated in this paper. The data used for this paper has also been made available online to encourage comparisons of other cortical thickness metrics against the ones tested herein.

3.8. Acknowledgements

The authors would like to thank Drs. Alex Zijdenbos, Louis Collins, Jens Pruessner, Yasser Ad-Dab'bagh and Jean-Francois Mangin for their suggestions and comments. Jason Lerch is funded by a K.M. Hunter/CIHR Doctoral Research Award.

Chapter 4: Alzheimer's Disease

4.1. Preface

Having established the validity of cortical thickness measurements for human brain imaging studies by simulation, the next task was to apply this algorithm to real data. We chose to study Alzheimer's Disease (AD), a patient group where cortical pathology was implicated and to some extent understood, so as to be able to reproduce the assumed spatial distribution of the pathology, quantify it, and better understand the mechanism of cortical pathology in the disorder.

AD is known to have a stereotypical developmental pattern, with the pathological hallmarks of the disease originating in the entorhinal cortex before spreading to the rest of the limbic system, followed by the associated cortices and then the rest of the cortex (Braak and Braak, 1991; Braak and Braak, 1995; Braak and Braak, 1996; Delacourte et al., 1999). Moreover, it had been shown that the spread of the neurofibrillary plaques was associated with neuronal loss and shrinkage of the cortex (Gomez-Isla et al., 1997; Buldyrev et al., 2000). Cortical thickness analysis should thus show the greatest loss in limbic and association areas

implicated earliest and least in the motor regions which are the last to feature pathology.

This is indeed what was found, with the greatest thickness loss - 1.25mm - occurring in the entorhinal cortex, the least in the motor and sensori-motor cortex. We were moreover able to correlate cortical thickness with Mini-Mental State Exam (MMSE) score as well as disease progression measured over longitudinal acquisitions, localizing thickness loss to the parahippocampal gyrus in both cases, with additional MMSE results in the lateral temporal lobes and progression associated loss in the frontal and anterior temporal lobes. These results both increased our knowledge about Alzheimer's Disease and gave further face validity to the cortical thickness algorithm.

Focal decline of cortical thickness in Alzheimer's Disease identified by computational neuroanatomy

Jason P Lerch¹, Jens C Pruessner^{1,2}, Alex Zijdenbos¹, Harald Hampel²,
Stefan J Teipel², Alan C Evans¹

¹ McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

² Alzheimer Memorial Center, Dementia Research Section and Memory Clinic, Department of Psychiatry, Ludwig-Maximilian University, Munich, Germany

Published in Cerebral Cortex 15 (7), 2005.

4.2. Abstract

Alzheimer's disease (AD) is characterized by a heterogeneous distribution of pathological changes throughout the brain. Magnetic Resonance Imaging (MRI) can be used to investigate the regional distribution of cortical atrophy in AD in vivo. One marker for the disease-specific atrophy is the thickness of the cortical mantle across the brain, obtained with automated 3D image processing.

Here, we present data from 36 subjects (17 controls and 19 patients diagnosed as probable AD) investigated for cortical thickness across the entire brain. We show significant cortical thickness decline in AD in temporal, orbitofrontal and parietal regions, with the most pronounced changes occurring in the allocortical region of the medial temporal lobes, outlining the parahippocampal gyrus, and representing a loss of over 1.25 millimeters of cortical thickness. Moreover, focal cortical areas decline with progression of the disease as measured by time from baseline scan as well as the Mini-Mental State Exam.

The results demonstrate the ability of this method to detect changes in cortical thickness in AD, across the entire brain, without need of prior anatomical definitions. The regional distribution of changes reported here is consistent with independent findings on the distribution of neuropathological alterations in AD. Using cortical thickness, moreover, we provide a direct quantitative index of atrophy in the disease.

Keywords: Alzheimer's Disease, Automated 3D image analysis, Cortical Thickness, MRI, Neuroinformatics, Neuropathology

Alzheimer's disease is characterized by the formation of neurofibrillary plaques and tangles, and neuronal loss across the central nervous system (Small et al., 2002). The histopathological changes show a characteristic sequence, with the entorhinal cortex and the hippocampus being among the first affected regions of the brain, followed by selected regions of the neocortex (Braak and Braak, 1991; Braak and Braak, 1996; Juottonen et al., 1998). Cognitive decline correlates with cortical atrophy (Mungas et al., 2002), which can be investigated with Magnetic Resonance Imaging (MRI) in vivo.

MRI provides insight into the temporal sequence of AD-related regional atrophy, using region specific as well as global search algorithms. Within region specific protocols, recent effort has focused on using MRI to detect early morphological changes in AD pathology; attention has specifically focused on developing precise segmentation protocols for the early affected temporal lobe structures such as the entorhinal cortex (EC), or the hippocampus (HC) (Van Hoesen, 1995; de Leon et al., 1997; Mori et al., 1997; Krasuski et al., 1998; De Toledo-Morrell et al., 2000).

In comparison, whole brain imaging analysis allows detection of changes throughout the entire cerebrum. So far, the methodology of whole brain analysis approaches has been restricted to voxel-based morphometry (VBM)(Ashburner and Friston, 2000), implemented volumetrically (Baron et al., 2001), across the cortex (Thompson et al., 2001), or as deformation analysis in longitudinal datasets (Good et al., 2002). In VBM, the concepts of gray-matter density and gray matter concentration are central for the interpretation of the results; however, voxel density at any one point for any one subject is, unlike cortical thickness, meaningless.

Using specific algorithms to analyze cortical thickness across the entire

cortex is a complementary method to the established research paradigms, offering a direct quantitative index of cortical atrophy that can be applied to single subjects and to group analysis. Cortical atrophy is reflected in loss of gray matter which will result in a reduction of cortical thickness. Measuring cortical thickness across the entire cerebrum establishes a marker for the AD-related cortical atrophy.

4.3. Materials and Methods

36 subjects were investigated. MRI scans were acquired from 19 patients with a combined 31 acquisitions (up to three scans per subject). The patients had the clinical diagnosis of probable AD according to the NINCDS-ADRDA (McKhann et al., 1984). For comparison of baseline MRI measures, 17 healthy volunteers with one acquisition each were recruited and subsequently scanned using identical acquisition parameters. Patients were recruited from the Department of Psychiatry, Alzheimer Memorial Center, Dementia and Imaging Research Group, University of Munich, Germany. Further sociodemographic information of the subjects is shown in Table 4.1. Cognitive impairment in the AD patients was assessed using the Mini Mental State Examination (MMSE) (Folstein et al., 1975). The average MMSE score in the Alzheimer group was 21.2 (10 to 29). In the control group, the MMSE average score was 29.3 (range 28 to 30).

	# Subjects	# Scans	Age	MMSE
Controls	17	17	61.0±9.1	29.3±0.6
Patients	19	31	68.8±6.9	21.2±4.6

Table 4.1: Demographics of the study population.

Significant medical co-morbidity in the AD patients and controls was excluded by interviews on medical history, physical and neurological

examination, psychiatric evaluation, chest X-ray, ECG, EEG, brain MRI, and laboratory tests (complete blood count, sedimentation rate, electrolytes, glucose, blood urea nitrogen, creatinine, liver-associated enzymes, cholesterol, HDL, triglycerides, antinuclear antibodies, rheumatoid factor, VDRL, HIV, serum B12, folate, thyroid function tests and urine analysis). None of the AD patients had hypertension or diabetes. All subjects or the holders of their Durable Power of Attorney provided written informed consent for the study. The protocol was approved by the Ethical Review Board of the Faculty of Medicine, Ludwig Maximilian University, Munich, Germany.

MRI examinations were performed on a 1.5 T Siemens Magnetom Vision MRI scanner (Siemens, Erlangen, Germany). All subjects were investigated with a volumetric T1 weighted sagittal oriented MRI sequence (TR = 11.6 ms, TE = 4.9 ms, resolution = 0.94 by 0.94 by 1.2 mm). The rectangular field of view (FOV) for the sagittal images was 256 mm (SI) x 204 mm (AP). Additionally, an axial oriented fast FLAIR sequence (TR = 9000 ms, TE = 110 ms, resolution = 0.94 by 0.94 by 6 mm) was obtained. For the purpose of this study, only the T1 weighted images entered further processing. The native MRI were registered into standardized stereotaxic space using a linear transformation (Collins et al., 1994). Simultaneously, the images were corrected for non-uniformity artifacts (Sled et al., 1998). The registered and corrected volumes were segmented into white matter, gray matter, cerebro-spinal fluid and background using an advanced neural net classifier (Zijdenbos et al., 2002). The white and gray matter surface were then fitted using deformable models (MacDonald et al., 2000), resulting in two surfaces with 81920 polygons each. The surface deformation algorithm works by first fitting the white matter surface, then expanding outward to find the gray matter CSF intersection. One characteristic of this procedure is that each vertex of the white matter

surface is closely related to its gray matter surface counterpart; cortical thickness can thus be defined as the distance between these linked vertices. The relevant parts of the processing pipeline are shown schematically in Figure 4.1.

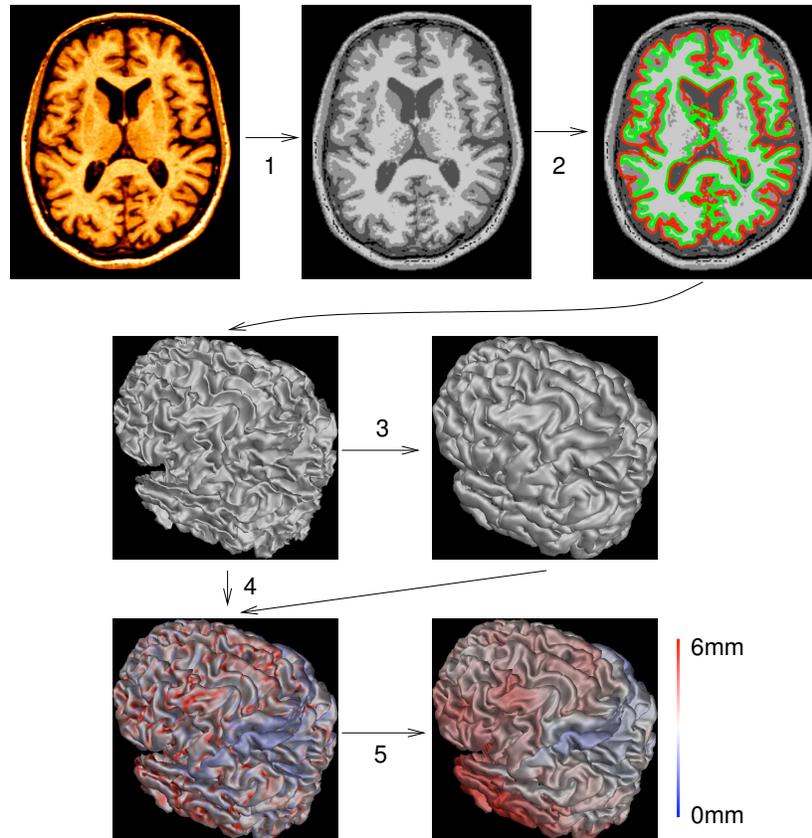


Figure 4.1: An overview of the steps involved in cortical thickness analysis. First, the images are non-uniformity corrected and registered into stereotaxic space. They are then classified (1) and fit with a white matter surface (2). The gray surface is found by expanding out from the white (3). Cortical thickness is measured at every vertex (4), and blurred using a 20mm surface based kernel (5).

In order to improve the ability to detect population changes, each subject's cortical thickness map was blurred using a 20 millimeter surface based blurring kernel (Chung et al., 2002). Diffusion smoothing, unlike the volumetric blurring used in VBM, follows the curvature of the surface and

thus respects anatomical boundaries. Twenty millimeters was chosen as the kernel size in order to maximize statistical power while still minimizing false positives.

Statistical analysis was performed at every vertex, regressing cortical thickness against clinical state, MMSE scores, or time from baseline. Multiple time-points were used where available in order to provide greater stability in the estimation of the fixed effects than a purely cross-sectional model would have allowed. Linear mixed models using the restricted maximum likelihood (REML) estimation method were employed to account for the within subject correlations present due to the repeated acquisitions in the AD cohort (Jose C. Pinheiro and Douglas M. Bates, 2000). Mixed models extend linear models by incorporating random effects, which can best be regarded as an additional error term. The model used is the following:

$$Y_i = X_i\beta + Z_ib_i + \varepsilon_i, i = 1, \dots, M$$

where β is the vector of fixed effects, $b_i, i = 1, \dots, M$ is the vector of random effects describing a shift in the intercept for each subject (M taking the value of either 19 or 36, depending on whether only the patients or all subjects were used), X_i (of size $n \times p$) and Z_i (of size $n_i \times q$) are known fixed-effects and random-effects regressor matrices, and ε_i is the n_i dimensional within group error vector, where n_i takes on values between 1 and 3, describing the number of acquisitions per subject. y_i is the estimated cortical thickness vector for subject i . Each of the statistical models used age as a covariate, thereby controlling for the main effects of age on the dependent variables; the value of p is therefore 3. The random-effects matrices (Z) included the intercept term; the value of q is therefore 1. For the regression with the MMSE scores, only the AD subjects were selected in order to avoid confounding of the clinical state with the MMSE analysis.

To illustrate with an example, the matrices for the MMSE analysis would be:

$$X_i = \dots = X_{19} = \begin{bmatrix} 1 & 66.9 & 24 \\ 1 & 67.8 & 20 \\ 1 & 69.5 & 20 \end{bmatrix}, Z_i = \dots = Z_{19} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

with the example subject 19 being one subject with three acquisitions taken at age 66.9, 67.8, and 69.5, and having MMSE scores of 24, 20, and 20 at each respective acquisition. The one dimensional random effects matrix Z is the intercept for each subject – i.e. a common slope across subjects was assumed, only the intercepts were allowed to vary. For the other analyses (clinical state and time from baseline) the third column of the X matrix would change to be a 0 or 1 for control or patient in the clinical state analysis (i.e. “treatment” contrasts were used) or the time, measured in years, of the follow up scan minus the baseline scan (i.e., for subject 19, these would be 0, 0.9, 2.6). The terms of would thus be: β_1 , the mean intercept; β_2 , for the common effect of age; and β_3 , the common effect of MMSE/group/time from baseline. The one dimensional vector b_i , $i = 1, \dots, 19$ (in the case of the MMSE or time since baseline analysis, since only the 19 patients are considered) or b_i , $i = 1, \dots, 36$ (in the case of the clinical state analysis, since all 36 subjects are considered) describes a shift in the intercept for each subject. In the parlance of the computational system used to solve these equations (Jose C. Pinheiro and Douglas M. Bates, 2000), the function call is $lme(y \sim 1 + age + MMSE, random = \sim 1|ID)$, where y is the cortical thickness at the vertex in question, age is the age at the time of each acquisition, $MMSE$ is the MMSE score at the time of each acquisition, the random effect is the intercept, and all subjects are grouped by the subject ID.

The resulting statistical maps were thresholded using the False Discovery Rate (FDR) theory (Genovese et al., 2002) at a q value of 0.05 after

pooling the p-values from all regressions run in this analysis. The interpretation of the maps is therefore that 5% of the results shown across all regressions are, on average, false positives. The figures also show the regression slope, which is the change in millimeters with each unit of measurement (group difference, point of MMSE, or in years).

Cortical thickness methods applied to human Magnetic Resonance Images can be divided into three broad categories. The first measures thickness on a voxel by voxel level, and is illustrated by two papers solving a partial differential equation across the cortex (Jones et al., 2000; Yezzi and Prince, 2003). The second set of methods use advanced versions of the marching cubes algorithm, creating two cortical surfaces with variable tessellation (c.f. (Zeng et al., 1999; Miller et al., 2000)). The last group uses deformable models to create white and grey matter cortices, and includes our work as outlined above as well as that of Fischl and Dale (Fischl and Dale, 2000; Rosas et al., 2002). One characteristic of using deformable models is that the number of polygons will always be identical across subjects, allowing for the easy creation of a surface coordinate system that can be used to run statistical analyses of cortical thickness at every vertex of the surface.

Our cortical thickness algorithm has been validated by comparison to manual measurements (Kabani et al., 2001). Furthermore, extensive analysis of the precision of our cortical thickness analysis framework was recently performed (Lerch and Evans, NeuroImage, in press). Two tests were run: a study of reliability where 19 scans of the same subject were processed separately to assess for methodological variability and a population simulation. In the latter, 50 normal scans were processed. 25 of them had the right superior temporal gyrus (rSTG) artificially thinned by one six neighbor dilation of the white matter in that region, and that change recovered statistically. In the 19 repeat scans the mean standard

deviation across all vertices was 0.27mm. The population simulation resulted in 93% sensitivity in areas of perfect overlap of the rSTG along with 100% specificity (no single vertex outside of the rSTG was found to be significant between the two groups). This simulation study provides confidence that results shown using the AD data described in this paper are valid.

4.4. Results

4.4.1. Group Differences (Normals vs. AD)

The results from the analysis clearly show significant differences in cortical thickness between the two groups. The average thickness across the entire cortex was significantly thinner in AD patients (3.1 ± 0.28 mm) compared to controls (3.74 ± 0.32 mm; $t=-3.8$, $p<.0007$), resulting in an average difference of 0.47 mm after removal of age effect by regression. Furthermore, the resulting maps of atrophy clearly show region specificity of thickness decline in AD. The most significant changes were found in the medial temporal lobes, the anterior and posterior cingulate region, the frontal lobes, the inferior parietal lobes, the orbitofrontal cortex, and the visual association cortex (Figures 4.2-4.4). Most of these effects were found bilaterally except for the insula region, where only the left hemisphere appeared significant.

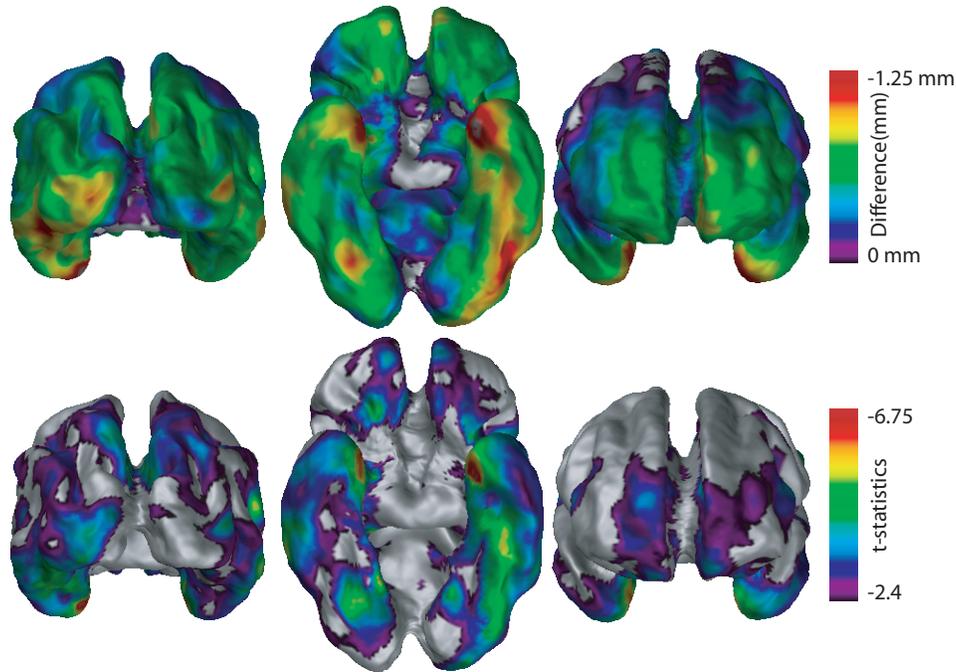


Figure 4.2: T-statistical (lower panel) and cortical thickness (upper panel) difference maps in 19 AD subjects versus 17 controls (age effect removed by regression). Results from the statistical analysis are displayed at each vertex of the surface of a standardized brain in terms of t-statistical color maps as well as color maps displaying the estimated cortical thickness difference in millimeters (the regression slope) between the two groups. Significant differences can be seen in the temporal lobes, especially the entorhinal and perirhinal cortices, as well as medial frontal and parietal lobes and left associative visual areas.

Location	Talairach Coordinates			t-statistics	mm difference	% atrophy
	X	Y	Z			
PHG	27	2	-39	-6.7	-1.25 ± 0.22	22 ± 3.8
ITG	-47	-55	-16	-5.3	-1.00 ± 0.18	22 ± 4.0
MOG	-13	31	-21	-4.5	-0.86 ± 0.23	18 ± 4.9
STG	59	-42	14	-4.4	-0.65 ± 0.14	17 ± 2.7
Pos. Cing	4	-30	37	-4.3	-0.86 ± 0.19	22 ± 5.0
LOG	-20	-99	0	-4.1	-1.03 ± 0.25	24 ± 5.8
MTG	63	-39	-14	-3.8	-0.85 ± 0.22	20 ± 5.2
Lt. IFG	-46	36	11	-3.7	-0.76 ± 0.22	22 ± 5.6
Ant. Cing.	-11	50	5	-3.4	-0.95 ± 0.28	19 ± 5.8

Table 4.2: Results sorted by t-statistics. PHG: parahippocampal gyrus. Ant. Cing.: anterior cingulate. MTG: middle temporal gyrus. ITG: inferior temporal gyrus. LOG: lateral occipital gyrus. MFG: middle frontal gyrus. MOG: medial orbital gyrus. STG: superior temporal gyrus. Pos. Cing: posterior cingulate. Lt IFG: left inferior frontal gyrus.

In the frontal lobes, the effects were most pronounced in the left anterior cingulate region (0.95 ± 0.28 mm loss in AD), the dorsolateral prefrontal cortex (0.76 ± 0.22 mm loss), and the orbitofrontal cortex (0.86 ± 0.23 mm). The effects in the dorsolateral prefrontal cortex appeared in the vicinity of Brodmann area (BA) 45. The effects were stronger in the left than in the right hemisphere. In the parietal lobe, the effects were most pronounced in the posterior cingulate region (0.86 ± 0.19 mm loss in cortical thickness in AD), and the visual association areas (1.0 ± 0.25 mm loss).

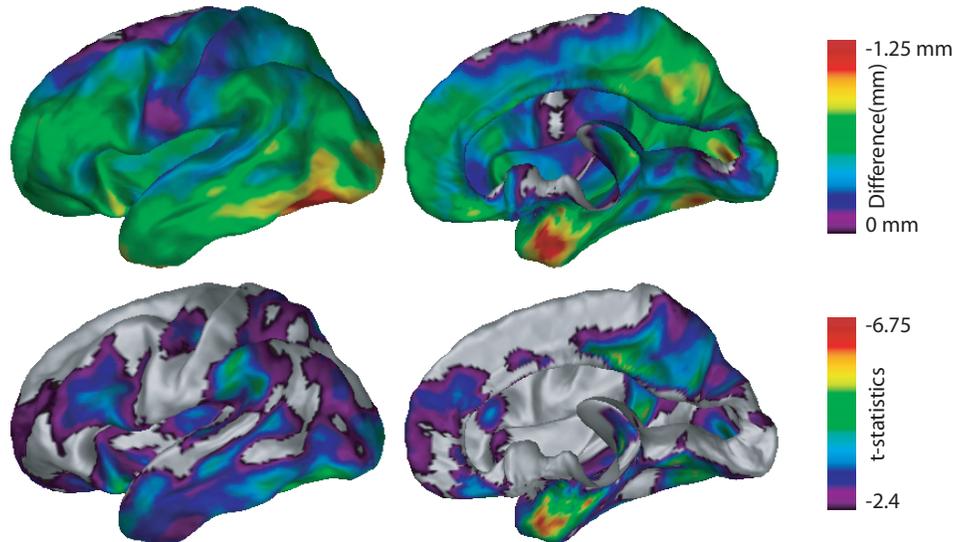


Figure 4.3: T-statistical (lower panel) and cortical thickness (upper panel) difference maps in 19 AD subjects versus 17 controls (age effect removed by regression). The display emphasizes the differences in the left lateral and right medial views. Significant differences can be seen in the posterior and anterior cingulate, the left dorsolateral prefrontal cortex, most of the temporal lobes, and the left supramarginal gyrus.

The entire medial temporal lobe appeared to be severely affected. The average difference between cortical thickness in AD and in the control group ranged from 0.5 mm to 1.3 mm, with the strongest differences emerging in the area of the parahippocampal gyrus (PHG). This finding is consistent with the results from recent studies suggesting that the PHG, especially the entorhinal cortex (EC), is affected early in the course of AD (Krasuski et al., 1998; Van Hoesen et al., 2000; Callen et al., 2001). To investigate cortical thickness decline in this area in our subjects more closely, manual segmentation of the structures of the PHG was used to trace the EC in the MRI of all subjects using a recently developed protocol (Pruessner et al., 2002). The labels of the EC from all subjects were then used to create a customized probabilistic map of the EC for the subjects in this study, which was overlaid onto the t-statistics map from the thickness

analysis. This allowed an accurate evaluation of the decline in cortical thickness within the EC in our study sample. The result of this procedure is shown in Figure 4.5, indicating that the most striking difference between the two groups occurred in the anterior portion of the EC in the left hemisphere, with a difference of 1.25 ± 0.22 mm in cortical thickness between the two groups. The lateral temporal lobes also showed significant interaction with clinical state, with a peak of 0.8 ± 0.22 mm in the left medial temporal gyrus (BA 21).

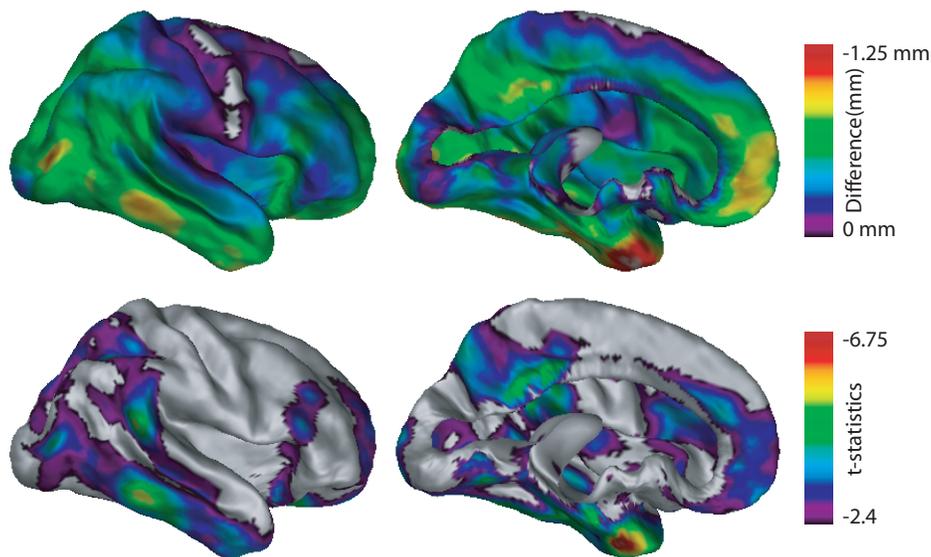


Figure 4.4: T-statistical (lower panel) and cortical thickness (upper panel) difference maps in 19 AD subjects versus 17 controls (age effect removed by regression). The display emphasizes the differences right lateral and left medial views. Significant differences can be found throughout the temporal lobes, the posterior and anterior cingulate. Compared to the left hemisphere the difference in the supramarginal gyrus and the dorsolateral prefrontal cortex is reduced.

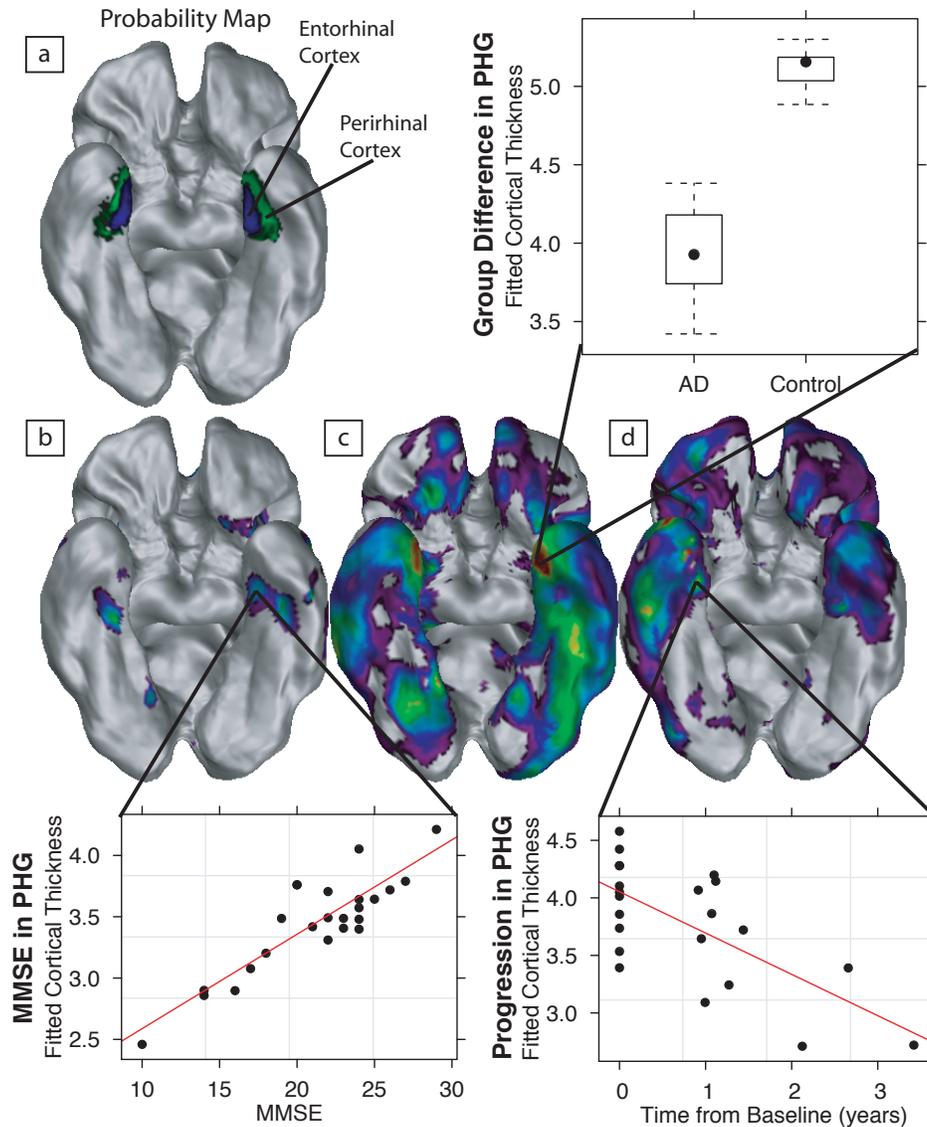


Figure 4.5: Regional analysis using the cortical thickness methodology in AD versus normal controls displaying differences in the entorhinal cortex. The four cortical views show: a) the probability maps of the entorhinal and perirhinal cortices in this population; b) the t-statistics of the MMSE regression; c) the group analysis; and d) the time difference from baseline. The color scales are the same as in figures 4.2-4.4, 4.6, and 4.7 respectively. The graphs illustrate the effects at the vertex indicated by the black lines. Atrophy of the PHG is clearly implicated in each of the three analyses shown.

In the parietal and occipital lobes, significant differences further appeared in the transition area between parietal and occipital lobes in the area of the visual association cortex (BA 18-19; between 0.4 – 1.0 mm loss of cortical thickness in the AD group). Interestingly, the primary visual cortex (BA 17) appeared to be spared, showing no interaction with group ($t=-1.6$).

4.4.2. Regression of Cortical Thickness versus MMSE

In order to investigate whether cortical thickness measures vary with disease progression in the AD patients, a regression of cortical thickness against the MMSE score was performed in the AD patients at every vertex of the surface model. Mean cortical thickness was marginally associated with MMSE scores ($t= 2.24$, $p=.06$). Regionally, however, it was found that lower MMSE scores were associated with significantly thinner cortical thickness in the bilateral PHG, the left superior temporal gyrus, left insula, and left anterior cingulate gyrus (Figure 4.6).

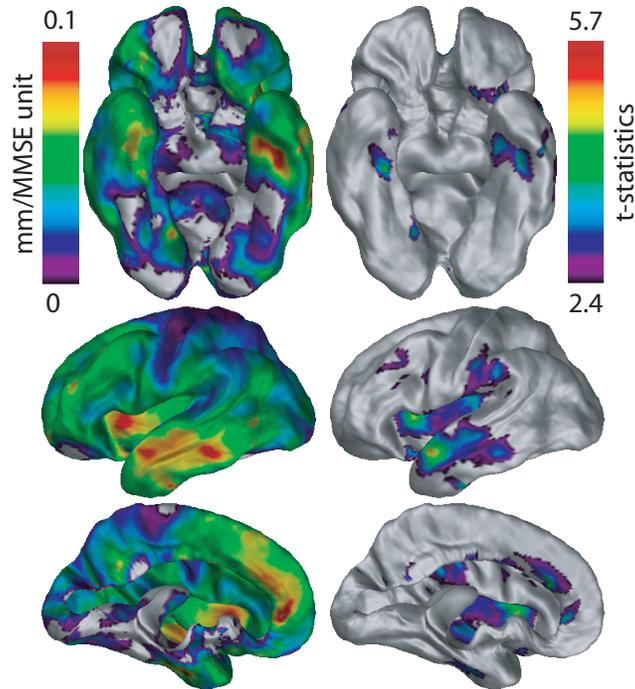


Figure 4.6: T-statistical (right column) and cortical thickness (left column) maps of the MMSE regression against cortical thickness in 19 AD subjects. Significant correlations can be seen in the lateral temporal lobes, the left insula, left anterior cingulate, and bilateral para-hippocampal gyri.

4.4.3. Regression of Cortical Thickness versus progression.

Further analysis of the effect of disease progression was performed by regressing cortical thickness against the time difference from baseline in the follow-up scans within the AD cohort. There was a significant correlation between time from baseline and thickness ($t=-3.47$, $p=0.006$) featuring a decline of 0.18 mm per year. Significant regional correlations were found in the anterior temporal lobes including the para-hippocampal gyrus, the anterior frontal lobes, and the anterior cingulate (Figure 4.7). A post-hoc test was also conducted to ascertain the rate of thinning in cortical areas that featured an initial group difference versus those where the Alzheimer's patients did not have significantly thinner cortex than the normal controls. Both these areas declined significantly with time from

baseline; the rate was faster in cortical regions which were also significantly thinner in the patients (0.21mm per year, $t=-4$, $p=0.004$) than the rest of the cortex (0.16mm per year, $t=-3.4$, $p=0.009$).

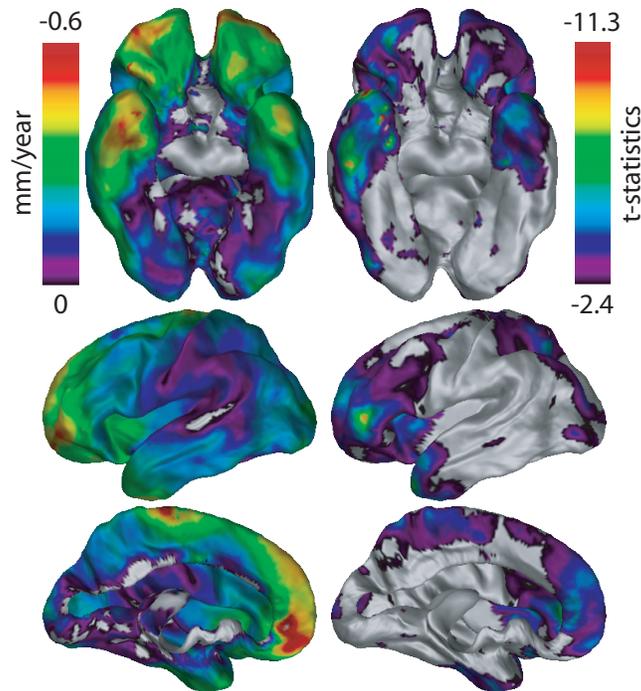


Figure 4.7: T-statistical (right column) and cortical thickness (left column) maps of the regression against time from baseline. Significant correlations can be seen in the anterior temporal lobes, anterior frontal lobes, and lateral posterior cortex.

4.5. Discussion

We used a fully automated method to measure cortical thickness across the entire brain to investigate differences in a group of AD patients versus age-matched controls. We further investigated the correlation between cortical thickness and MMSE scores as well as time since baseline in the patient population.

Our results clearly show AD related decline in cortical thickness in multiple areas of the brain, many of which have been reported in previous MR studies. Cortical thickness of the medial temporal lobes was most severely reduced in AD patients, and, within the medial temporal lobes, the parahippocampal gyrus was most affected. This is consistent with previous MR and histopathological studies showing that this region of the brain is affected early and profoundly in the course of the disease (Juottonen et al., 1999; Xu et al., 2000). The results from this study extend previous findings by showing that in the left hemisphere, the posterior portions of the PHG also seem to be strongly affected in AD. For the anterior portion of the PHG, in the area of the entorhinal cortex, the most significant differences occurred in the left hemisphere, though the effects were largely bilateral.

Cortical atrophy in AD is not limited to the medial temporal lobe; the remaining limbic system, the lateral temporal lobes, and certain associative visual areas correlate significantly with the disease as well. The finding of significant cortical thickness decline in the area of the visual association cortex is consistent with other studies showing this region to be strongly affected by AD (Cronin-Golomb et al., 1991; Thulborn et al., 2000), as are the results in the lateral temporal lobes (Busatto et al., 2003). Within the limbic system the orbito-frontal cortex (Van Hoesen et al., 2000) and the anterior and posterior cingulate cortices (Valla et al., 2001; Matsuda et al., 2002) too have been shown to correlate with the disease in ways that support the results of this study. Combined with spared areas of the cortex, such as the motor, sensory, and primary visual cortices (BA17) (Buckner et al., 2000; Rizzo et al., 2000; Prvulovic et al., 2002), a pattern of atrophy in AD emerges that is focal to multiple cortical areas while centered on the medial temporal lobes.

The results from the MMSE regression within the AD group indicate that

the decline in the PHG is correlated with the MMSE score (Fields et al., 1992). Significant associations between MMSE scores and cortical thickness were found in bilateral medial temporal areas as well as temporal, anterior cingulate, and insular regions of the left hemisphere. The results demonstrate that cortical thickness analysis can be used to link clinical information with decline in particular cortical areas when available. In the case of MMSE scores, the significant findings with cortical atrophy were left hemisphere dominant. Progression as measured by time from baseline in repeated scans showed a more significant effect than the MMSE analysis, indicating continuing atrophy that is not entirely captured by changes in cognitive state as captured by the MMSE. Progressive atrophy was particularly strong in the anterior frontal and temporal lobes as well as the posterior cingulate. Cortical areas which were significantly thinner in the patients thinned at a faster rate than those that were not, though this difference was small and would therefore suggest that the entire cortex is declining with disease progression. The longitudinal aspect of our analysis should be explored in greater detail and compared non-linear registration methods that have recently been used in the analysis of AD populations and normal aging (Fox et al., 2000; Scahill et al., 2002; Scahill et al., 2003).

AD is commonly subdivided into six stages representing advancing pathology as defined by the progression of neurofibrillary tangles (NFTs) and senile plaques (Braak and Braak, 1991; Braak and Braak, 1996). The patients used in this study were diagnosed with probable AD according to NINCDS-ARDR criteria and thus believed to be in stage V or VI (Nagy et al., 1999). These two end stages, known as the neocortical stages, feature plaque and NFT involvement in virtually all subdivisions of the cerebral cortex with a particular emphasis on association areas and the medial temporal lobes. NFT presence is strongly correlated with

neuronal loss and cortical atrophy (Gomez-Isla et al., 1997; Grignon et al., 1998). Moreover, increased duration of NFT presence is associated with increased atrophy (Grignon et al., 1998). This leads to the hypothesis that, in patients diagnosed with probable AD, cortical areas involved earlier in the disease progression will show greater atrophy than areas involved later (Gomez-Isla et al., 1997). Our results, which reveal greatest MRI detectable atrophy in areas implicated earlier (such as the entorhinal cortex), largely support this hypothesis. The ability to stage AD based on in-vivo imaging should be investigated further. Moreover, some of the findings in this study may be due to the small sample size, since occasional atypical representation of AD (such as the visuo-spatial variant, c.f. (Cronin-Golomb et al., 1991)) may have an undue influence on our results. This analysis should therefore be replicated in a larger sample.

The major difference between previous findings and the results obtained in this study lies in the use of the cortical thickness analysis technique. This strategy has a number of advantages when compared to other MR methods used to investigate AD with MRI. Cortical thickness analysis allows searching for associations between the depth of the cortex and sociodemographic, clinical or psychological variables across the entire surface of the brain. It shares this advantage with other global search algorithms like VBM and deformation field analysis. Unlike VBM, however, it is based on the description of the actual thickness of the cortex in millimeters and thus allows meaningful quantitative description of the results (difference in cortical thickness in millimeters). The major disadvantage is that the focus is entirely on the cortex; changes in subcortical structures, white matter, or cerebro-spinal fluid will not be picked up.

Finally, cortical thickness analysis is fully automated and thus reaches 100% operator independence. The results are operator-independent and

do not rely on the correct interpretation of regional anatomical variations, in contrast to manual segmentation protocols.

Cortical thickness analysis may serve as a surrogate marker for the neuronal loss that accompanies the histopathological changes in the cortex which occur in AD and their temporal ordering. For example, the highly significant findings in the PHG, in the area of the EC, support the view (Juottonen et al., 1999; Xu et al., 2000) about the importance of medial temporal lobe morphometry in the diagnosis of AD. Future studies can compare the relative efficacy of cortical thickness measures with volumetric analyses of the hippocampus (Pruessner et al., 2001).

Continuing work will investigate associations between cortical thickness and other sociodemographic or clinical variables. Owing to its fully automatic implementation, the method can be used to characterize the healthy and pathological ranges of cortical thickness in specific age groups. The development of normative data for specific age and disease groups will allow direct comparison of individual subjects with cortical thickness norms in health and disease. This process has the potential of aiding in the early diagnosis of dementia. Furthermore, unlike commonly used region of interest measures, cortical thickness analysis provides coverage of the complete cerebrum, and could thus be used for differential diagnosis of the various types of dementia as well. Finally, future studies will have to show the value of this new method in monitoring of the progression of dementia across the entire cortex.

Acknowledgements: ICBM grant PO1MHO52176-11, Principal Investigator Dr. John Mazziotta; CIHR grant MOP-34996. Jason Lerch is funded by a K.M. Hunter/CIHR Doctoral Research Award.

Chapter 5: AD prediction

5.1. Preface

One of the pivotal roles of brain imaging in Alzheimer's Disease is early diagnosis. It is currently predicted, though not known with certainty, that pathological changes in the cortex precede the clinical symptoms used for the diagnosis of probable or possible AD. Even though current treatments for AD are largely symptomatic, the fact remains that the earlier treatment commences, the better the chances of success or the higher the quality of life that can be maintained.

Having described and quantified cortical thickness loss in AD in the previous paper, the natural extension was to assess how capable thickness is at reproducing the clinical diagnosis. The AD cases available for this study allowed us to ask the important question of how good cortical thickness analysis is at diagnosis in established AD using statistically robust discrimination techniques. Established AD cases were used to prove that it can work; future work needs to address whether cortical thickness analysis can predict progression to AD from mild cognitive

impairment, and whether data from a prospective study can predict AD before the clinician can.

We showed that cortical thickness can perfectly separate patients from controls if one uses the parahippocampal area along with one of six other structures. Moreover, we were able to provide maps of relative diagnostic accuracy, sensitivity, and specificity of each vertex. Lastly, this study compared the relative efficacy of three different discriminant techniques - linear discriminants, quadratic discriminants, and logistic regression, and found them to provide comparable performance.

Future work should extend this to prospective data: can brain imaging predict AD before the clinical diagnosis can be made, and if so, how much earlier? The analysis framework developed herein is generic and can therefore also be applied to other clinical datasets.

Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls

Authors: Jason P. Lerch¹, Jens Pruessner^{1,2}, Alex P. Zijdenbos¹, D. Louis Collins¹, Stefan J. Teipel², Harald Hampel², and Alan C. Evans¹

¹ McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada.

² Alzheimer Memorial Center, Dementia Research Section and Memory Clinic, Department of Psychiatry, Ludwig-Maximilian University, Munich, Germany

Submitted to *Neurobiology of Ageing*.

5.2. Abstract

We investigated the potential of fully automated measurements of cortical thickness to reproduce the clinical diagnosis in Alzheimer's Disease (AD). Thickness maps were analyzed using three different discriminant techniques to separate patients from controls. All analyses were performed using leave-one-out cross-validation to avoid overtraining of the discriminants. The results show regionally variant patterns of discrimination ability, with over 90% accuracy obtained in the medial temporal lobes and other limbic structures. Multivariate discriminant analysis produced 100% accuracy with six different combinations, all involving the parahippocampal gyrus. We therefore propose automated measurements of cortical thickness as a tool to improve the clinical diagnosis of probable AD, as well as a research method to gain unique insight into the etiology of cortical pathology in the disease.

Keywords: Alzheimer's Disease, Cortical Thickness, Diagnosis, Magnetic Resonance Imaging.

5.3. Introduction

Brain imaging has played an increasingly important role in the study of dementias over the last decade. Within the realm of magnetic resonance imaging (MRI), work has predominantly focused on the segmentation of structures known to be involved in dementia, especially the parahippocampal gyrus and the hippocampus (Killiany et al., 1993; Bobinski et al., 1999; Killiany et al., 2000; Pruessner et al., 2002; Chetelat and Baron, 2003). These methods have recently been complemented with

techniques employing whole brain analyses (Baron et al., 2001; Janke et al., 2001; Thompson et al., 2001; Chetelat et al., 2002; Gee et al., 2003; Lerch et al., 2004).

One of the key roles that imaging can play in dementia research as well as clinical practice is to aid the early diagnosis process (see Scheltens et al., 2002; Chetelat and Baron, 2003; Zakzanis et al., 2003 for reviews of the field). The crucial task is to identify image-based metrics that can accurately differentiate patients from normal elderly controls. Ultimately, the goal will be to discover a metric that can detect dementia at an earlier stage than a standard neurological diagnosis.

Discriminating normal elderly controls from patients diagnosed with probable AD using MR imaging techniques achieves accuracies ranging from 58% to 100% (Chetelat and Baron, 2003). This wide range in accuracies reported indicates that the search for good discriminants continues. Moreover, almost all methods described to date involve manual measurements, which are time consuming, subject to inter-rater variability, and require expert anatomists in order to be performed correctly. Normal analysis is furthermore often restricted to a pre-defined subset of brain regions, which has a number of limitations. First is the theoretical challenge in performing differential diagnosis of the various dementias, since each have their unique anatomical presentations, featuring different rates and foci of cortical atrophy (Good et al., 2002; Gee et al., 2003). The second issue is that using prior regions of interest restricts the use of additive discriminant models. For example, it is conceivable that the best discriminant is not just the volume of the entorhinal cortex, but rather a model containing the entorhinal cortex and prefrontal cortex volumes.

Discriminant analyses need to separate the training and validation datasets to be methodologically robust. Lack of such separation limits the

generalizability to the population at large and risks classifier overtraining. Given an arbitrarily complex classification decision boundary any number of groups can be correctly separated; such boundaries are, however, unlikely to correctly classify new data not used in the creation of that boundary. Not separating training and validation datasets therefore results in inflated reports of accuracy. The training and validation datasets are ideally entirely separate; given a small study population, however, a leave-one-out or jack-knife approach can be employed. Here all but one of the members of the population are used to train the classifier, and the thus created decision boundary is used to classify the subject who was left out. The process is then repeated for all subjects, each being left out of the training once and subsequently classified. Such techniques avoid the danger of inflating accuracies while still retaining maximum power and should therefore always be used when performing discriminant analyses.

Existing approaches for dementia classification in the imaging literature typically rely on logistic regression. Other types of discriminant analysis techniques exist, such as linear and quadratic discriminants or neural networks, some of which may lead to better discrimination for the application at hand. More importantly, quadratic and linear discriminants, for example, have the potential of separating more than two groups simultaneously, a property lacking in logistic regression and necessary for differential diagnosis of different dementias. Different discriminant techniques should thus be assessed in the simple case of separating AD patients from controls using imaging data in order to evaluate their relative abilities as dementia classifiers.

The objective of this study was to investigate the ability of automated cortical thickness measurements to discriminate Alzheimer's patients from normal elderly controls. Cortical thickness has ideal properties for this

task: it covers the entire cortex, it is fully automated and thus not subject to inter-rater variability, and it reflects a fundamental neuro-anatomical property known to be affected in Alzheimer's disease (Gomez-Isla et al., 1997; Grignon et al., 1998; Van Hoesen et al., 2000). Recent work using the techniques described in this paper has shown it to be significantly different in normal elderly controls and Alzheimer's Disease patients (Lerch et al., 2004). Three different discriminant functions were tested: 1) linear discriminant analysis (LDA), 2) quadratic discriminant analysis (QDA), and 3) logistic regression implemented through neural nets. In order to not over-train the classifiers, leave-one-out (jack-knife) cross-validation was performed throughout.

5.4. Methods

5.4.1. Participants

36 subjects were studied. MRI scans were acquired from 19 patients (mean age 68.8 ± 6.9) and 17 controls (mean age 61.0 ± 9.1). The patients had the clinical diagnosis of probable AD according to the NINCDS-ADRDA (McKhann et al., 1984). Patients were recruited from the Department of Psychiatry, Alzheimer Memorial Center, Dementia and Imaging Research Group, Ludwig Maximilian University, Munich, Germany. Cognitive impairment in the AD patients was assessed using the Mini Mental State Examination (MMSE) (Folstein et al., 1975). The average MMSE score in the Alzheimer group was 21.2 (10 to 29). In the control group, the MMSE average score was 29.3 (range 28 to 30).

Significant medical co-morbidity in the AD patients and controls was excluded by interviews on medical history, physical and neurological examination, psychiatric evaluation, chest X-ray, ECG, EEG, brain MRI,

and laboratory tests (complete blood count, sedimentation rate, electrolytes, glucose, blood urea nitrogen, creatinine, liver-associated enzymes, cholesterol, HDL, triglycerides, antinuclear antibodies, rheumatoid factor, VDRL, HIV, serum B12, folate, thyroid function tests and urine analysis). None of the AD patients had hypertension or diabetes. All subjects or the holders of their Durable Power of Attorney provided written informed consent for the study. The protocol was approved by the Ethical Review Board of the Faculty of Medicine, Ludwig Maximilian University, Munich, Germany.

5.4.2. Image acquisition

MRI examinations were performed on a 1.5 T Siemens Magnetom Vision MRI scanner (Siemens, Erlangen, Germany). All subjects were investigated with a volumetric T1-weighted sagittal oriented MRI sequence (TR = 11.6 ms, TE = 4.9 ms, resolution = 0.94 by 0.94 by 1.2 mm). The rectangular field of view (FOV) for the sagittal images was 256 mm (SI) x 204 mm (AP). Additionally, an axial oriented fast FLAIR sequence (TR = 9000 ms, TE = 110 ms, resolution = 0.94 by 0.94 by 6 mm) was obtained. For the purpose of this study, only the T1-weighted images entered further processing due to the low resolution of the fast FLAIR acquisition.

5.4.3. Image processing

The native MRI were registered into standardized stereotaxic space using a linear transformation (Collins et al., 1994). Simultaneously, the images were corrected for non-uniformity artifacts (Sled et al., 1998). The registered and corrected volumes were segmented into white matter, gray matter, cerebro-spinal fluid and background using an advanced neural net classifier (Zijdenbos et al., 2002). The white and gray matter surface were then fitted using deformable models (MacDonald et al., 2000), resulting in

two surfaces with 81,920 polygons (40,962 vertices) each. The surface deformation algorithm works by first fitting the white matter surface, then expanding outward to find the gray matter CSF intersection. One characteristic of this procedure is that each vertex of the white matter surface is closely related to its gray matter surface counterpart; cortical thickness can thus be defined as the distance between these linked vertices (Lerch and Evans, 2005). Two example cortical thickness maps are shown in figure 5.1. The use of the cortical thickness algorithms described here has been validated using both manual measurements (Kabani et al., 2001) and a population simulation (Lerch and Evans, 2005).

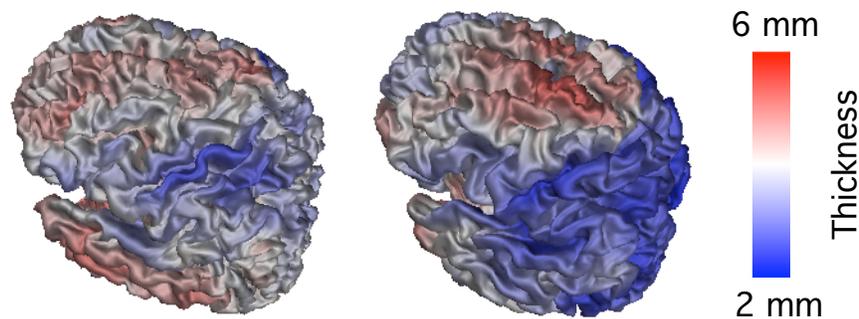


Figure 5.1: Two example cortical thickness maps, showing a normal control on the left, an AD patient on the right.

In order to improve the ability to detect population changes, each subject's cortical thickness map was blurred using a 20 millimeter surface-based diffusion blurring kernel (Chung et al., 2003). Diffusion smoothing, unlike the volumetric blurring used in voxel based morphometry, follows the curvature of the surface and thus respects anatomical boundaries. Twenty millimeters was chosen as the kernel size in order to maximize statistical power while still minimizing false positives and allowing for accurate estimation of the millimeter difference in cortical thickness (Lerch and Evans, 2005).

Using an automatic segmentation algorithm to identify 25 separate cortical

structures (ANIMAL - a 3D image-warping algorithm, Collins et al., 1995; Robbins et al., 2004) we were able to match each subject's MRI to a pre-segmented template using non-linear deformations. An average cortical segmentation was produced by finding the anatomic label with the highest occurrence at each vertex (figure 5.2). The cortical ROI map provided the regional parcellation scheme for the analyses in table 5.1. The reasons for performing analyses on ROIs are twofold: 1) to limit the number of potential combinations when performing multivariate discrimination; and 2) to provide anatomical definitions more in line with the existing literature. It should be noted that the area defined as the parahippocampal gyrus comprises parts of the entorhinal cortex, perirhinal cortex, parahippocampal cortex, and those parts of uncus and hippocampus proper that are on the cortical surface in the medial temporal lobe.

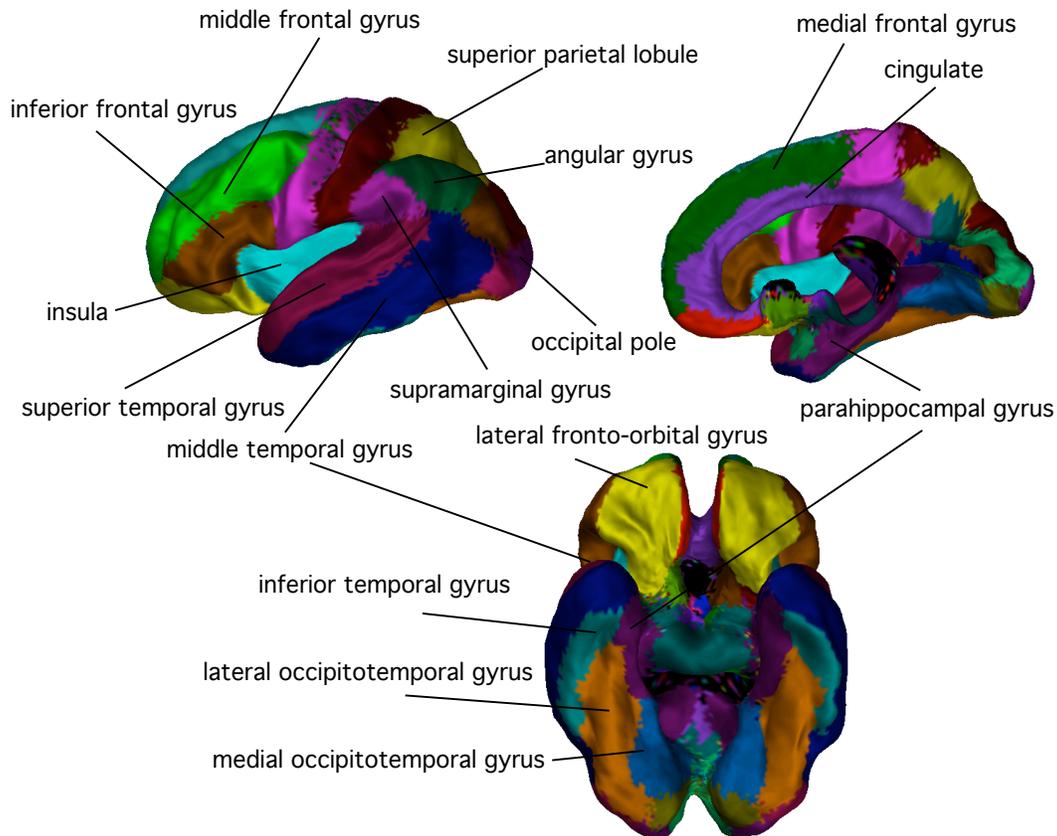


Figure 5.2: The surface segmentation used for the definitions of cortical areas displayed on the average cortex, each colour representing a different label. Thickness for each area was defined as the mean thickness of all vertices defined as belonging to a particular region.

5.4.4. Discriminant analysis

Discriminant analysis was performed using three different techniques: linear discriminant analysis, quadratic discriminant analysis, and logistic regression. The discriminant functions were tested at all 40,962 vertices of the surface. Results were computed as sensitivity, the proportion of AD patients classified as patients; specificity, the proportion of controls classified as controls; and accuracy, the proportion of all subjects correctly classified (see equation 5.1). All discriminant analysis was implemented

using the functions available in the MASS and nnet libraries available in the R Project for Statistical Computing (<http://www.r-project.org>) (Venables and Ripley, 2002).

		Discriminant Results		$Specificity = \frac{d}{c + d}$
		AD	Cntrl	
Diagnosis Result	AD	A	B	$Sensitivity = \frac{a}{a + b}$
	Cntrl	C	D	

Equation 5.1: Measurements of Accuracy, Sensitivity, and Specificity

Both linear and quadratic discriminants aim to find the best separation between the classes so that any new observation has the greatest likelihood of being added to the correct class. In the case of linear discriminants the class covariance matrices are assumed to be equal, whereas they are separately estimated for quadratic discriminants. The most popular method in the literature of prediction of clinical state through imaging in dementia is logistic regression, which models the posterior probabilities of the different classes through linear functions while at the same time ensuring that they sum to 1. This is described in more detail in (Hastie et al., 2001; Venables and Ripley, 2002).

The discriminant analysis used for this study performs 40,962 tests at points equally spaced across the cortex. Controlling for multiple comparisons using parametric tests such as Bonferroni's Correction or Random Field Theory (Worsley et al., 1999; Worsley et al., 2004) has to date not been developed for discriminant tests. In order to address the problem of chance findings due to multiple comparisons we employed permutation tests. Permutation testing is a non-parametric technique used

to determine the exact distribution of a particular statistical test within a given sample, hence providing the p-value of the test result. The statistical test is repeated multiple times, exchanging the association between the subject's cortical thickness data and their group label (patient or control) each time. This is ideally repeated so that all possible combinations are reproduced; in the case of the population presented here, however, that leads to a prohibitively large number of permutations (${}_{36}C_{17}$), hence we restrict ourselves to an approximate or Monte-Carlo permutation test, where only a subset - in our case 1000 - of all possible permutations is randomly chosen. The maximum accuracy across all 40,962 vertices is kept for each one of the 1000 permutation tests; these then form the distribution of the accuracy test-statistic. A p-value of 0.05, for example, is then computed by sorting the 1000 maximum accuracy values and taking the 50th highest. This provides strong control over experiment-wise Type I error. See (Nichols and Holmes, 2002) for a review of permutation testing in neuroimaging.

Mean thicknesses across different cortical structures were assessed separately. Each structure was tested both individually and in a multivariate test in combination with every other structure. This leads to ${}_nC_r$ possible combinations, where n is the number of structures (25) and r is the number of structures in each combination (2), expanded in equation 5.2 and totalling 300 combinations.

$${}_nC_r = \frac{n!}{r!(n-r)!}$$

Equation 5.2: Combinatorics This equation gives the number of possible combinations.

Leave-one-out cross-validation was used; i.e. the discriminant is trained using all the subjects save one; that subject is then classified as either patient or control based on the thus trained discriminant. This is repeated

for every subject, so as to not bias the results by using the same data for the results as for the training.

5.5. Results

5.5.1. Controlling for multiple comparisons.

Multiple comparisons were controlled for using 1000 permutation tests. The fifth percentile of the resulting distribution - reflecting a p-value of 0.05 with strong control of Type I error - was found at an accuracy of 0.84%. No random permutation exceeded an accuracy of 0.85%, any accuracies higher than 0.85 can thus strongly reject the null hypothesis of no accurate discrimination between patients and controls.

5.5.2. Separating AD patients from normal elderly controls.

When applied to the real cortical thickness data, mean cortical thickness alone resulted in an accuracy of 75%, sensitivity of 79%, and specificity of 71%. No single structure was able to perfectly separate the two groups. The parahippocampal gyrus had the highest accuracy at 94%. The lowest accuracy was 50% in the superior frontal gyrus. The results for all structures are listed in table 5.1.

area	accuracy	sensitivity	specificity
parahippocampal gyrus	0.94	0.94	0.95
medial occipitotemporal gyrus	0.89	0.88	0.89
cingulate region	0.83	0.88	0.79
inferior temporal gyrus	0.83	0.88	0.79
lateral occipitotemporal gyrus	0.83	0.82	0.84
medial fronto-orbital gyrus	0.81	0.88	0.74
middle temporal gyrus	0.81	0.82	0.79
occipital pole	0.81	0.82	0.79
insula	0.81	0.82	0.79
precuneus	0.78	0.71	0.84
inferior occipital gyrus	0.78	0.82	0.74
inferior frontal gyrus	0.75	0.71	0.79
lateral fronto-orbital gyrus	0.75	0.88	0.63
supramarginal gyrus	0.75	0.65	0.84
angular gyrus	0.75	0.65	0.84
superior temporal gyrus	0.75	0.71	0.79
superior occipital gyrus	0.75	0.65	0.84
middle occipital gyrus	0.75	0.65	0.84
cuneus	0.75	0.71	0.79
middle frontal gyrus	0.72	0.71	0.74
superior parietal lobule	0.72	0.65	0.79
precentral gyrus	0.69	0.71	0.68
postcentral gyrus	0.69	0.59	0.79
medial frontal gyrus	0.61	0.65	0.58
superior frontal gyrus	0.50	0.59	0.42

Table 5.1: Results of QDA by cortical structure, sorted by accuracy. No single structure was able to give perfect discrimination.

When the discriminant techniques were applied at every vertex, perfect discrimination using QDA was obtained at 1 vertex, located in the right

entorhinal cortex. 85% accuracy was found in multiple areas in the temporal lobes, right posterior cingulate, right parietal lobes, and bilateral orbitofrontal regions. The primary motor cortex and medial frontal lobes provided the worst discrimination. Results are shown in figure 5.3.

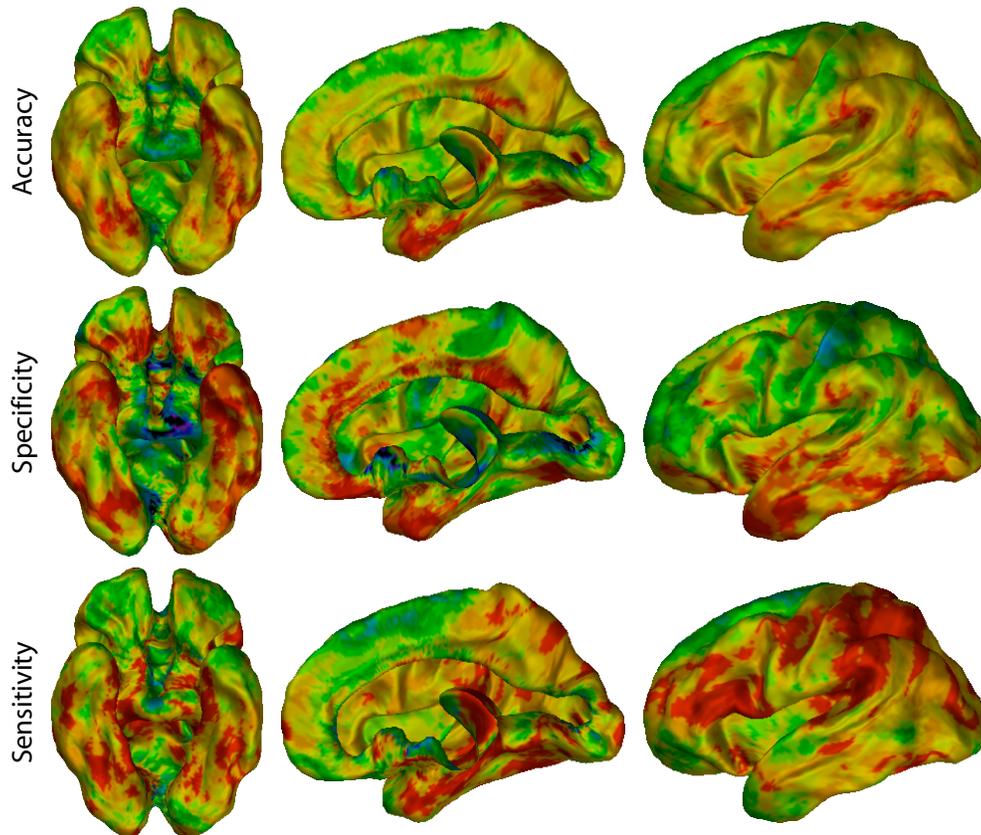


Figure 5.3: Results of the Quadratic Discriminant Analysis across the cortex, showing high accuracy in the temporal lobes, posterior cingulate, parietal lobules and dorsolateral prefrontal cortex. The maps of specificity resemble the expected distribution of AD pathology more closely than the sensitivity maps.

We also examined whether combining ROI data improved the group discrimination. All 300 possible two-way combinations were tested; six separate two-structure models obtained an accuracy of 100%. These two

structure models were always comprised of the parahippocampal gyrus together with either the superior frontal gyrus, inferior frontal gyrus, middle frontal gyrus, superior temporal gyrus, inferior temporal gyrus, or insula. Seven additional combinations featured accuracies of 97%. An example two-structure model is graphed in figure 5.4.

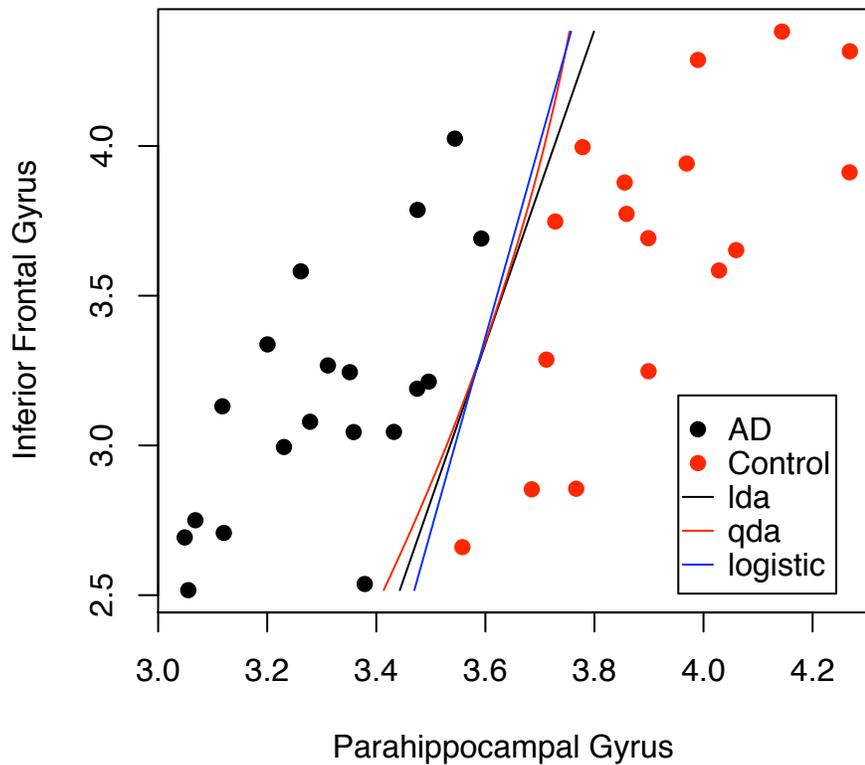


Figure 5.4: 100% accuracy can be obtained by using the mean thickness values of the parahippocampal gyrus and any one of six frontal or temporal lobe structures. The graph above shows the individual data-points from the parahippocampal gyrus and inferior frontal gyrus combination as well as the dividing lines chosen by the three discriminant techniques employed in this study.

5.5.3. Comparison of discriminant techniques

LDA and QDA had comparable performances, each obtaining a mean accuracy of 69% across the 40,962 vertices. Logistic regression's mean accuracy was 66%. Two sample t-tests showed no significant differences in accuracies between LDA and QDA, but both were significantly better than logistic regression ($p < 2.2 \times 10^{-16}$). QDA had the highest percentage of vertices with greater than 95% accuracy, and the only one to obtain 100% accuracy at a single vertex. In the two structure discrimination task logistic regression achieved the best results with 6 combinations featuring 100% accuracy, followed by LDA with 3 and QDA with 2.

5.6. Discussion

5.6.1. Discriminant ability of cortical thickness

The results show that cortical thickness was able to discriminate AD patients from normal elderly controls. The high accuracy obtained compares favorably with similar studies (Chetelat et al., 2002).

Previous studies have shown that accuracies range from 58% to 100% (Chetelat and Baron, 2003). Apparent 100% discrimination inevitably resulted from multivariate discriminant models (Bobinski et al., 1999; Killiany et al., 2000) or older studies (pre-1995) with a small study sample (less than 10 subjects in at least one of the two groups) (Killiany et al., 1993; Lehericy et al., 1994). The multivariate models used in the two studies that obtained perfect discrimination moreover did not employ any kind of separation of training and validation datasets. The higher the number of variables used in such a model, the more likely the chance of over-fitting the data if the validation and training sets are the same. To

illustrate this point with an example: in our datasets, if 10 most accurate vertices are used in a multivariate QDA, 100% accuracy is obtained. If, however, cross-validation is used on the same analysis, 8 false positives and 2 false negatives result – an accuracy of 72%. Even using just two vertices results in 2 false positives with cross-validation, none without. Hence, results obtained without separation of training and validation data should be confirmed with such separation in place.

Predictive ability is not uniform across the cortex. The best prediction is found for areas classically associated with AD, primarily located in the limbic circuit, specifically the parahippocampal gyrus. The poorest discrimination was found in the primary motor regions of the cortex, thought to be one of the last affected by the plaques and tangles characteristic of the disease (Braak and Braak, 1995; Braak and Braak, 1996; Delacourte et al., 1999). Furthermore, the cortical localization of these results closely resembles previous work using standard linear and mixed model techniques (Lerch et al., 2004).

5.6.2. On the use of multivariate discriminants

The best results obtained in this study came from multivariate discriminant analysis, performed by combining mean thickness across 25 different cortical structures. The ability to use multivariate methods is one of the key strengths in using cortical thickness measures for subject classification, as it makes the method easily expandable to discriminating between different types of dementia as well as differentiating patients from normal elderly controls. As discussed here, care has to be taken to separate the training from the testing dataset. Due to obtaining 100% accuracy with two structures we did not explore more complex models. Further work should investigate using more than two structures for this type of analysis; the key will be to reduce the dimensionality of the problem and find

computationally efficient means of selecting which structures or vertices to use for multivariate discriminants.

The anatomical definitions used for the structure labels were based on an automatic procedure that has not been validated for all structures, and is therefore likely to include an element of error. The main reason for using them was to reduce the dimensionality of the problem. The exact boundaries of the anatomical definitions used should not be taken as definite.

5.6.3. On the use of different types of discriminants

Three types of discriminants were used in this analysis: linear discriminants, quadratic discriminants, and logistic regression. The results obtained from each were comparable, with QDA best at discriminating based on single values, logistic regression best in the multivariate task. There is a further theoretical advantage in using either LDA or QDA, and that is their ability to discriminate between more than two groups at a time. LDA or QDA could be employed when analyzing different types of dementias or more than two groups within the same dementia. In AD, it would be especially interesting to investigate the classification of mild cognitive impairment using these analysis techniques.

5.7. Conclusions

In summary, we presented a set of fully automatic techniques to extract cortical thickness from an AD cohort with normal elderly controls, and employed statistically well-controlled discriminant techniques at every point on the cortex to separate patients from controls. This allows for maps showing the distribution of accuracy, sensitivity, and specificity. Initial

multivariate analyses combining automatically segmented cortical regions produced 100% accuracies in six different combinations, all involving the parahippocampal gyrus. These results are a clear indication that the use of advanced computational techniques applied to brain imaging data carries significant promise as a diagnostic tool. The potential of this technique can best be appreciated when employed with populations at risk using longitudinal designs, which is why future studies should focus on mild cognitive impairment and very early AD subjects.

5.8. Acknowledgements

ICBM grant PO1MHO52176-11, principal investigator Dr John Mazziotta; CIHR grant MOP-34996. Jason Lerch is funded by a K.M. Hunter/CIHR Doctoral Research Award.

Chapter 6: MACACC

6.1. Preface

The last study in this thesis departs considerably from the previous three. The main question is generic: given that we have shown cortical thickness to be precise, capable of quantifying their characteristic cortical pathology of specific disorders, and has a potentially role in AD diagnosis, can we begin using cortical thickness maps to generate new hypotheses about neuroanatomy in cortical development or degeneration?

The specific question addressed was how do the thickness of different areas of the cortex relate to each other. Different cognitive functions localize to different areas of the cortex, and high domains, such as language or vision, are organized into hierarchical domains spanning multiple cortical areas. Such cortical organization leads to a corollary hypothesis: as the thickness of the cortex increases in one part of a network, the other centres in the same network will feature a correlated change. These correlations will furthermore be influenced by developmental patterns and cognitive factors.

We developed a series of statistical methods to test these hypotheses,

called Mapping Anatomical Correlations Across Cerebral Cortex (MACACC), and applied them to a normal paediatric population of 292 subjects. The results showed a robust relationship between known fibre tracts and patterns of cross-cortical thickness correlations in the case of Brodmann Area 44, our chosen seed region. These relationship furthermore changed as the total population was subdivided into different groupings: the correlation between BA 44 and the superior temporal gyrus was tighter in older than younger subjects, and BA 44 featured stronger correlations to the anterior cingulate and lateral parietal areas in subjects with higher IQs than those in the low IQ grouping.

The purpose of this paper was to develop and illustrate the MACACC methods. The convincing nature of these preliminary results indicate that this field of study has potential in multiple normal and disease populations. For instance, a degenerative disorder such as AD might show a breakdown of the relationship between the cortical thickness of different limbic areas. Changing correlation patterns across normal ageing are also open to investigation, as are the cross-cortical correlations in disorders such as Schizophrenia or even ADHD hypothesized to have aberrant connectivity.

Mapping Anatomical Correlations Across Cerebral Cortex (MACACC) using Cortical Thickness from MRI.

Jason P. Lerch¹, Keith Worsley¹, W. Philip Shaw², Deanna Greenstein²,
Rhoshel K. Lenroot², Jay Giedd², Alan C. Evans^{1*}

¹ McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill
University, Montreal, Quebec.

² National Institute of Mental Health, National Institutes of Health,
Department of Health and Human Services, Bethesda, Maryland.

Submitted to NeuroImage.

6.2. Abstract

We introduce MACACC - Mapping Anatomical Correlations Across Cerebral Cortex - to study correlated changes within and across different cortical networks. The principal topic of investigation is whether the thickness of one area of the cortex changes in a statistically correlated fashion with changes in thickness of other cortical regions. We further extend these methods by introducing techniques to test whether different population groupings exhibit significantly varying MACACC patterns. The methods are described in detail and applied to a normal childhood development population ($n=292$), and show that association cortices have the highest correlation strengths. Taking Brodmann Area (BA) 44 as a seed region revealed MACACC patterns strikingly similar to tractography maps obtained from diffusion tensor imaging. Furthermore, the MACACC map of BA 44 changed with age, older subjects featuring tighter correlations with BA 44 in the anterior portions of the superior temporal gyri. Lastly, IQ-dependent MACACC differences were investigated, revealing steeper correlations between BA 44 and multiple frontal and parietal regions for the higher IQ group, most significantly ($t=4.0$) in the anterior cingulate.

6.3. Introduction

The cerebral cortex is organized into networks of functionally complementary areas. Classic examples include the dorsal and ventral visual streams, the limbic system, and the language networks. These networks are traditionally studied using functional paradigms designed to reveal the particular role of a certain area. Examples of such studies include the role of Broca's Area in word repetition, synonym generation (Klein et al., 1997), verbal fluency (Frith et al., 1991), speech production (Buckner et al., 1995) and silent word production (Friedman et al., 1998).

Functional specialization can also lead to related anatomical change. A recent study investigated the size of the hippocampus, a structure involved in spatial navigation, in London taxi drivers and found increases in size correlating with increased experience in navigating the streets of London (Maguire et al., 2000; Maguire et al., 2003). Similarly, faster phonetic learners were found to have greater white matter density in parietal regions than slower learners (Golestani et al., 2002). Trained musicians feature enlarged primary motor and sensorimotor areas, premotor areas, anterior superior parietal areas, and inferior temporal gyri (Schlaug, 2001; Gaser and Schlaug, 2003a; Gaser and Schlaug, 2003b). This last example involving musicians is particularly informative as it involves multiple cortical areas, including motor, sensorimotor, and multimodal sensory areas, collaborating. Increases in anterior corpus callosum size further suggests that the intra-hemispheric connectivity of the brain is enhanced in trained musicians (Schlaug, 2001).

We propose to address a related but less explored topic: as the anatomy of one cortical area changes, are there correlated morphological changes

in other cortical areas? An example hypothesis from the language-processing domain might be that a population with thicker cortices in Broca's Area will have a correspondingly larger Wernicke's Area.

The connectivity of the human cerebral cortex is not a new topic of investigation. It has traditionally been studied using fiber tracing, wherein a seed region is injected with a retrograde tracer in order to determine which areas have direct fibre connections to the seed region (Romanski et al., 1999; Petrides and Pandya, 2002). More recently the notion of functional connectivity has been promoted, i.e. areas that are functionally related will feature correlated change in a fMRI or PET functional activation study (Friston et al., 1993; Friston et al., 1996; Koski and Paus, 2000; Friston, 2002; Friston et al., 2003; Horwitz, 2003; Lee et al., 2003; Ramnani et al., 2004).

We propose to study correlated anatomical changes using methods related to functional connectivity but employing morphometric data; we have dubbed this approach Mapping Anatomical Correlation Across Cerebral Cortex (MACACC). Of particular interest will be not just testing which areas of the cortex correlate with which other areas, but also whether the MACACC patterns vary across different categorical groupings based on demographical variables (age, gender, socio-economic status) or clinical diagnoses.

The methods used in MACACC involve statistical analyses of data extracted from anatomical MRI using the metric of cortical thickness, which features several advantageous properties for MACACC:

1. It covers the entire cortex.
2. It provides a biologically meaningful measurement (cortical thickness).

3. It has a reduced number of points compared to volumetric data (40,962 vertices versus 1,000,000 voxels).

The methods developed herein should, however, also be applicable to other anatomical data such as, for example, voxel density measures from Voxel Based Morphometry (VBM) (Ashburner and Friston, 2000; Watkins et al., 2001). Some statistical issues have been explored in (Worsley et al., 2005).

We use this technique to address two core developmental issues. Firstly, several strands of evidence suggest that anatomical and functional interconnections between Broca's and Wernicke's areas may increase with age (Paus et al., 1999). Language becomes more lateralized (Holland et al., 2001), functional imaging studies demonstrate increased activity of Brodmann Area 44 (BA 44) during language processing (Amunts et al., 2004), and DTI studies of the white matter tract connecting Broca's and Wernicke's areas shows changes in indirect measures of integrity (fractional anisotropy) (Schmithorst et al., 2002). We thus predict that the correlation in cortical thickness of these two language processing regions would show a developmental gradient, increasing with age. Secondly, intelligence is linked to brain activity. Greater brain activity during intellectually demanding activity has been shown using fMRI in individuals of greater general intelligence in an extensive network comprising lateral frontal and parietal regions (Gray et al., 2003). Others report similar alterations in functional connectivity linked to intelligence (Haier et al., 2004) which may be underpinned by enhanced structural connectivity. We hypothesize that, as seen in the motor learning examples given above, the functional enhancements are mirrored by concomitant change in the underlying neural substrate which can be detected by MRI-based morphometry, e.g. MACACC. The population used herein comes from a

study of normal childhood brain development (Giedd, 2004; Gogtay et al., 2004). The purpose is not to provide a definitive account of these latter two questions but rather to illustrate the concept of MACACC, leaving more detailed study of brain development for future work.

6.4. Methods

The methods used to study MACACC can be subdivided into: 1) the extraction of the morphometric data and, 2) the statistical techniques used to ascertain the correlations. A summary of all the definitions used throughout is included in table 6.2.

6.4.1. Extraction of Morphometric Data

The input metric for MACACC studies is cortical thickness as measured from MRI. The choice of MR sequence is immaterial as long as sufficient resolution and grey/white matter contrast are provided. The native MR image is corrected for non-uniformity artefacts (Sled et al., 1998) and registered into stereotaxic space using a nine parameter linear transformation (Collins et al., 1994). Cerebral tissue is classified into white matter, gray matter, spinal fluid and background using a neural net classifier (Zijdenbos et al., 2002). The inner and outer cortical surfaces are then extracted using deformable surface-mesh models (MacDonald et al., 2000; Kim et al., 2005) and non-linearly aligned towards a standard template surface (Robbins et al., 2004). Cortical thickness is measured in native-space millimetres using the linked distance between the white and pial surfaces, t_{link} (MacDonald et al., 2000; Lerch and Evans, 2005). The thickness map is blurred using a 30mm surface based diffusion smoothing kernel (Chung et al., 2003). These methods have been validated using both manual measurements (Kabani et al., 2001) and a population

simulation (Lerch and Evans, 2005), and used in an Alzheimer's Disease population study (Lerch et al., 2004). Closely related methods have also been applied to Huntington's (Rosas et al., 2002) and normal ageing (Salat et al., 2004) among others. Example output can be seen in figure 6.1.

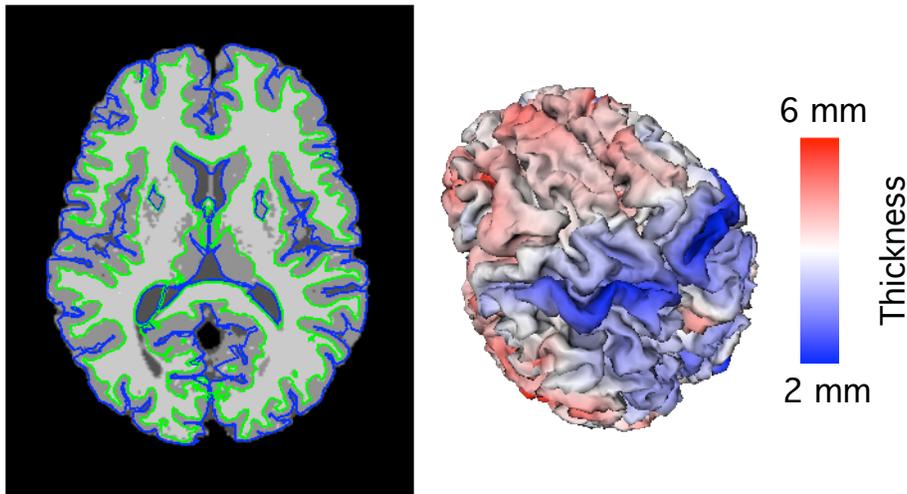


Figure 6.1: Example cortical thickness map, showing (on the left) the intersection of the pial surface (blue) and the white matter surface (green) on a transverse slice of the tissue classification map. On the right is a cortical surface colour coded with cortical thickness.

6.4.2. Statistical Techniques

Measuring MACACC can be subdivided into two main components: assessing cross-cortical correlations as well as quantifying differences in MACACC maps across groups.

MACACC methods – correlations

Cortical cross-correlations are obtained using simple linear correlations whose strength is measured using Pearson's r .

$$r = \frac{\sum T_i T_j - \frac{\sum T_i \sum T_j}{N_s}}{\sqrt{(\sum T_i^2 - \frac{(\sum T_i)^2}{N_s})(\sum T_j^2 - \frac{(\sum T_j)^2}{N_s})}}$$

Equation 6.1: Pearson's r. Where T_i and T_j represent cortical thickness at the two vertices to be correlated with each other and N_s is the total number of subjects.

Pearson's r takes on a value between 1 and -1, the sign referring to whether the correlation is positive or negative, and the closer to 1 or -1 the more significant the correlation. Equivalently, though slower computationally, is the use of a linear model wherein the seed region T_i (the area whose MACACC one is measuring) is modelled by target region T_j (every vertex on the rest of the cortex).

$$T_i = \beta_0 + \beta_1 T_j + \varepsilon$$

Equation 6.2: Linear Model. Where the thickness of seed region T_i is modelled by a linear model of thickness of T_j plus a normally distributed error ε .

The significance of the linear model is measured using the student's t statistics with $N-1$ degrees of freedom.

If a hypothesis exists, then one can target a seed region or vertex and explore its cortical cross-correlations. In the absence of a hypothesis a less directed search is needed. We therefore introduce the concept of MACACC-strength which measures the mean r value for the correlation against all other vertices at each vertex.

$$MACACC_{s_i} = \frac{\sum_{j=1}^{N_v} \left(\frac{\sum T_i T_j - \frac{\sum T_i \sum T_j}{N_s}}{\sqrt{(\sum T_i^2 - \frac{(\sum T_i)^2}{N_s})(\sum T_j^2 - \frac{(\sum T_j)^2}{N_s})}} \right)}{N_v} = \frac{\sum_{j=1}^{N_v} r}{N_v}$$

Equation 6.3: MACACC-strength: Where N_v is the number of vertices, N_s is the number of subjects, T_i is the thickness at seed point i , and T_j is the thickness at target point j .

MACACC-strength is an $N \times N$ process and therefore computationally expensive. The current surface extraction procedure results in 40,962 vertices, a total of 40962×40962 correlations are needed. The true measure of MACACC-strength can be approximated, however, by correlating the thickness at every vertex against each subject's mean cortical thickness averaged over all 40,962 vertices. The results will be identical if the standard deviation is exactly the same at every node, which is known to not be the case (Lerch and Evans, 2005). The approximation is, however, a lot quicker to compute and should give a good indication of the true MACACC-strength result.

Assessing group differences

The second component of MACACC is testing for group differences. The question to be addressed is whether the MACACC maps between two or more groups is significantly different. For example, does the MACACC map for BA 44, which shows a significant correlation with the superior temporal gyrus, change across different age ranges? There are two different hypotheses that can be tested:

1. Differences in MACACC slope.
2. Differences in MACACC variance.

MACACC-slope

In the case of a MACACC-slope the type of correlation between groups is different, i.e. it is steeper in one group than another (figure 6.2a), whereas in MACACC-variance the slope is potentially the same but the goodness of fit of the linear correlation is different (figure 6.2b).

Testing for the difference in slope is performed using a classic interaction linear model.

$$T_i = \beta_0 + \beta_1 T_j + \beta_2 \text{Group} + \beta_3 (T_j * \text{Group}) + \varepsilon$$

Equation 6.4: Interaction Model testing for differences in the relationship between T_i and T_j cortical thickness among the different factors present in Group. The Group component is modelled using treatment contrasts.

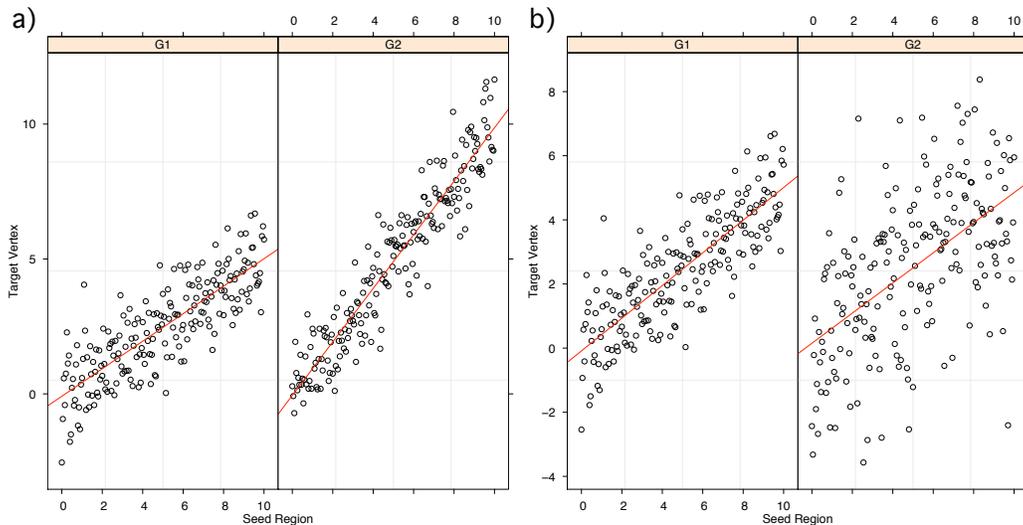


Figure 6.2: Group differences: The two graphs above show the difference between MACACC-slope and MACACC-variance using simulated data. Graph a shows a correlation between target and seed regions where the variance between the two groups is the same, the slope different, being steeper in group G2. Graph b shows the same slope in the target vertex to seed region correlation, but the variance is higher in group G2 than G1. Graph a would thus be captured using MACACC-slope tests (interactions are significant at $p < 2.2 \times 10^{16}$, variance not significant), graph b using MACACC-variance tests (interactions not significant, variance significant at $p < 3.9 \times 10^8$).

The factors in the Group variable in the linear model described above are created using treatment contrasts, significance tested using student's t statistic. One can also think of this problem with the following formulation: is the data better fit by a model that contains separate slopes for each group rather than a common slope?

MACACC–variance

One of the assumptions behind least-squares fits is that the error distribution along each dimension have equal variance, i.e. homoscedasticity. Given the example in figure 6.2b, this assumption will not necessarily hold in MACACC analyses. We therefore use a general least squares approach allowing unequal variances for each group in the classification factor (Jose C. Pinheiro and Douglas M. Bates, 2000). Testing for differences in MACACC-variance is more complex than MACACC-slope, as the statistical methods are less well established. The criteria for a valid test are that it:

1. Capture difference in variances across multiple groups.
2. Has a defined model rejecting the assumption of equal variance H_0 .
3. Be robust against outliers.

After the correlation slope has been determined, we must then test for whether the residuals by group are homoscedastic (have equal variance) or heteroscedastic (have unequal variance). Residuals are obtained by fitting a linear model relating T_i and T_j and subtracting the observed response from the predicted response.

$$\begin{aligned} \text{Linear Model: } T_i &= \beta_0 + \beta_1 T_j + \varepsilon \\ \text{Residual: } e_i &= T_i - \beta_0 - \beta_1 T_j \end{aligned}$$

Equation 6.5: Linear model residuals.

Three tests were assessed in this study:

1. F-test for variance.
2. Bartlett's test.

3. Levene's test.

The F-test for variance is the simplest of the three, containing solely the ratio of two group variances.

$$s_1^2 = \frac{1}{N_1 - 1} \sum_{i=1}^{N_1} (G1_{e_i} - \hat{G1}_e)$$

$$s_2^2 = \frac{1}{N_2 - 1} \sum_{i=1}^{N_2} (G2_{e_i} - \hat{G2}_e)$$

$$F = \frac{s_1^2}{s_2^2}$$

Equation 6.6: F-test for variance. G1 and G2 represent the two groups, N1 and N2 the respective group sizes, e_i the residuals as determined in equation 6.5.

The F-test has two fundamental limitations: it is limited to two groups and is highly non-robust against outliers.

Bartlett's test and Levene's test are designed to overcome limitations of the F-test. Bartlett's test measures the squared deviations from the group mean, Levene's the absolute deviation from the group mean, median, or trimmed mean (NIST/SEMATECH, 2005).

$$T = \frac{(N-k) \ln s_p^2 - \sum_{i=1}^k (N_i - 1) \ln s_i^2}{1 + (1/3(k-1)) \left(\left(\sum_{i=1}^k 1/(N_i - 1) \right) - 1/(N-k) \right)}$$

Where N is the total sample size
and k is the number of groups
and N_i is the sample size of the i th group
and s_i^2 is the pooled variance

Equation 6.7: Bartlett's Test, which results in a X^2 distribution of $k-1$ degrees of freedom (NIST/SEMATECH, 2005).

$$W = \frac{(N-k) \sum_{i=1}^k N_i (\hat{Z}_{i.} - \hat{Z}_{..})^2}{(k-1) \sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - \hat{Z}_{i.})^2}$$

Where N_i is the sample size of the i th group
and k is the number of groups
and $\hat{Z}_{i.}$ are the group means of Z_{ij}
and $\hat{Z}_{..}$ is the overall mean of the Z_{ij}
and $Z_{ij} = |Y_{ij} - \tilde{Y}_{i.}|$
and $\tilde{Y}_{i.}$ is the median of i th group

Equation 6.8: Levene's Test, which results in a F distribution of $(k-1, N-k)$ degrees of freedom (NIST/SEMATECH, 2005).

The main difference between Bartlett's and Levene's test is their respective robustness to non-normality and to outliers. Levene's test, emphasizing the absolute difference from either the median or trimmed mean should theoretically be less susceptible to outliers. The question of which of these tests to use for MACACC analysis will be addressed in the section on applications to normal development.

Multiple Comparisons

In MACACC, as for any statistical technique in brain-imaging, the error rate due to multiple comparisons has to be controlled for. Depending on the exact test used, MACACC produces p-values generated from t, F, or X^2 distribution. The resulting values can be corrected using standard techniques, such as Bonferroni correction, Random Field Theory (Worsley et al., 1992; Worsley et al., 1996; Worsley et al., 1999; Worsley et al., 2004), Permutation testing (Nichols and Holmes, 2002; Nichols and Hayasaka, 2003), or the False Discovery Rate (FDR) (Genovese et al., 2002). We used FDR since it provides the greatest ability to threshold data stemming from multiple different statistical tests simultaneously, as described in (Lerch et al., 2004). The interpretation by FDR theory is thus that, of all the vertices shown as significant at the 0.05 level in this paper, 5% will be false positives.

Using Residuals

MACACC analyses performed on raw cortical thickness data can be dominated by main effects occurring within the population being studied, most obviously age related changes (Thompson et al., 2000; Giedd, 2004; Gogtay et al., 2004). To remove this confound a linear regression of age was first performed at every vertex, and the residuals of that regression substituted for the raw cortical thickness values. Similar procedures removed gender differences.

6.5. Application to normal development

The methods described above will be illustrated in a normal development population. The sample consists of 292 adolescents, characterized in table 6.1.

	N	Age	IQ
Male	157	11.1±3.8	116.4±12.7
Female	135	12.0±3.9	112.5±13.4

Table 6.1: Population Characteristics

MR Images were acquired on a 1.5-T General Electric Signa scanner (Milwaukee, Wis). T1-weighted images with contiguous 1.5-mm slices in the axial plane and 2.0-mm slices in the coronal plane were obtained using 3-dimensional spoiled gradient recalled echo in the steady state. Imaging parameters were echo time of 5 ms, repetition time of 24 ms, flip angle of 45°, acquisition matrix of 256 x 192, number of excitations equals 1, and 24 cm field of view. All subjects were tested using age appropriate versions of the Weschler intelligence scales (the WPPSI-III for children 4 to 6; the WISC-III or WASI for 6 to 17 years and the WAIS-III for 18 and

above (Wechsler, 1991)).

6.5.1. Correlation Strength

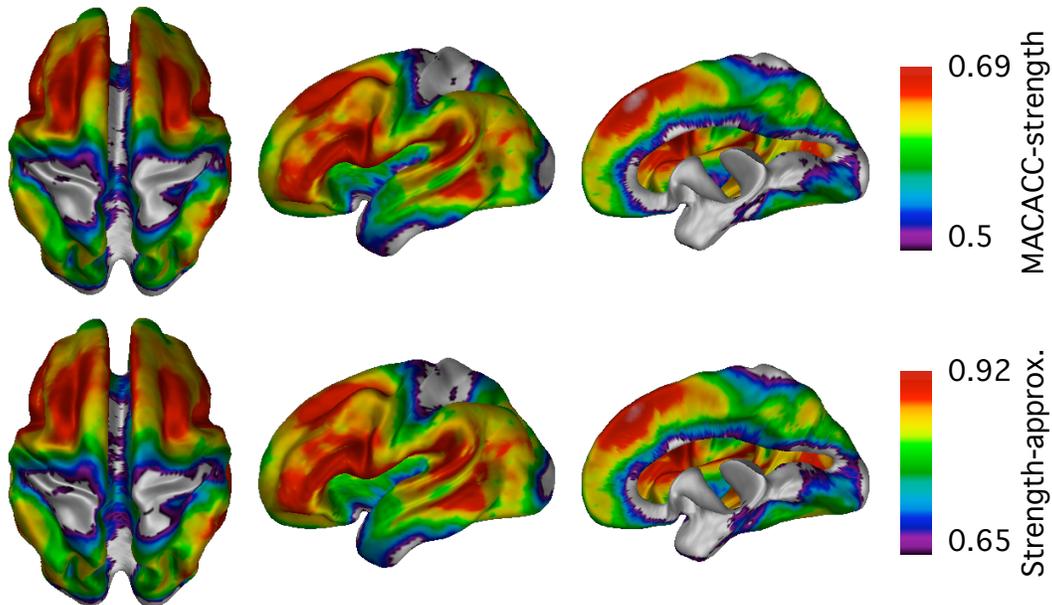


Figure 6.3: MACACC-strength: Three view of the correlation strength of the entire population sample ($n=292$) are shown in the top part of the figure. The association cortices clearly have the highest strength, primary motor, sensorimotor, and visual cortices the least. The bottom figure shows the MACACC-strength approximation, computed by correlating every vertex against mean cortical thickness. The results clearly look similar, albeit at different thresholds, and took 30 seconds to compute, as compared to the three days for the correct measure of MACACC-strength.

MACACC-strength measures the degree to which any part of the cortex correlates with the rest of the cortex. If one assumes that areas of the cortex with related function should correlate with each other to a greater degree, then one can expect that those areas of the cortex which relate to the greatest number of cortical regions will feature the highest MACACC-strength: the association cortices. These serve an integrative function,

receiving inputs from multiple cortical and non-cortical sources, and in turn distributing information to multiple areas. The results of MACACC-strength in the entire population sample are shown in figure 6.3. As expected, the highest MACACC-strength was found in the association cortices, the lowest in the primary motor, sensorimotor, and visual areas.

6.5.2. Seed region: BA 44

As an example of the use of MACACC to assess a specific cortical area's correlation with other cortical areas, we used left Brodmann Area 44 (BA 44). We used a probability map derived from cytoarchitectonic data and non-linearly transformed into stereotaxic space for our regional definition (Amunts et al., 1999), as shown in figure 6.4a. BA 44 is part of Broca's area and primarily implicated in high-level aspects of speech production (Klein et al., 1995; Klein et al., 1997; Amunts et al., 2004; Fiebach et al., 2005), as well as potentially involved in imitation (Heiser et al., 2003; Makuuchi, 2004) and music perception (Platel et al., 1997). It is connected through the arcuate and uncinate fasciculi to other language areas, including Wernicke's and BA 40 (Petrides and Pandya, 1988; Chertkow and Murtha, 1997; Duffau et al., 2002; Parker et al., 2005).

We first used a BA 44 seed region to map the MACACC of BA 44 in the entire population sample, followed by an investigation of age differences for this map. Given its involvement in the language circuit, one would expect BA 44 to correlate strongly with functionally related areas such as Wernicke's area, the superior temporal gyrus, and BA 40 and 39. The results of this analysis are shown in figure 6.4b. The cortical thickness of BA 44 correlates with neighbouring frontal regions, the large extent of these correlations is possibly due to the expansive definition of BA 44 used in this study. The MACACC map outside of the frontal lobes shows a striking similarity to tractographic maps of the arcuate fasciculus,

encompassing inferior parietal areas along with the lateral temporal lobes. and containing much of the language network. These correlations were highly significant, exceeding a correlation coefficient of 0.8.

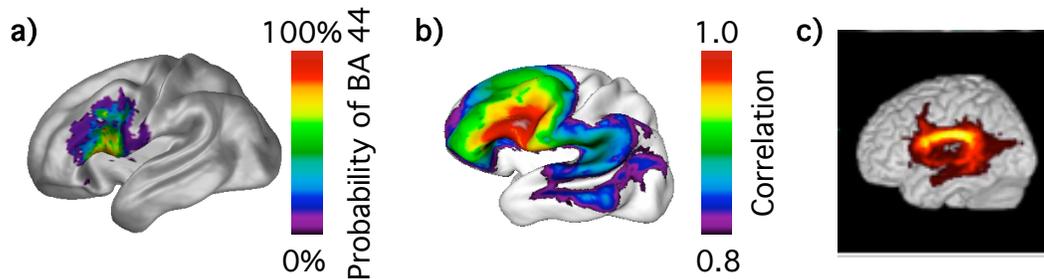


Figure 6.4: MACACC of BA 44. Part a shows the probability map of Brodmann Area 44 used as the seed region, obtained from (Amunts et al., 1999). Part b shows the MACACC map generated from the 292 subjects, and part c, reproduced from (Parker et al., 2005), shows the diffusion tensor imaging probability map of Broca's area connections through the arcuate and uncinate fasciculi. The correlation map features strong correlations throughout the frontal lobes, possibly due to the expansive BA 44 probability map used, but the extra-frontal lobe correlations show a pattern remarkably similar to the tractography of the arcuate fasciculus.

6.5.3. Developmental MACACC differences

The cerebral cortex undergoes profound change throughout normal childhood development, particularly through the reorganization of synaptic and axonal contacts. Synaptic pruning is especially relevant for our dataset, as it is a continuous process from ages two to sixteen (Rivkin, 2000). The thickness of the cortex declines throughout most of the cerebrum with the exception of the temporal poles where continuous growth can be observed.

Given the degree of structural change observed throughout development and the coincident maturation of language abilities one can hypothesize

age related MACACC changes within our sample. Recent diffusion tensor imaging work has shown that the arcuate fasciculus changes in fractional anisotropy during normal development within an age range similar to that of our sample (Schmithorst et al., 2002).

Furthermore, fMRI has shown that lateralization of language function increases, and that BA 44 BOLD activity correlates with age (Holland et al., 2001). To test for MACACC age-related differences we subdivided our population into age-specific grouping, the young group containing subjects of less than 8.8 years, the old group older than 14.2 years, the middle group in between. Age was controlled for within each of the groupings. Both MACACC-slope and MACACC-variance were and the results are show in figure 6.5. Significant differences were observed between the older and younger subjects. The slope relating BA 44 and the bilateral superior temporal gyri (STG) is steeper in older subjects. Moreover, the variance differs in the STG, the medial frontal lobes, right Broca's Area, right orbito-frontal cortex, and bilateral lateral parietal lobes. In all cases the variance is less in the older subjects. These maps thus indicate an increasing correlation with age between the left BA 44 and its counterparts in the language network and frontal lobe circuits.

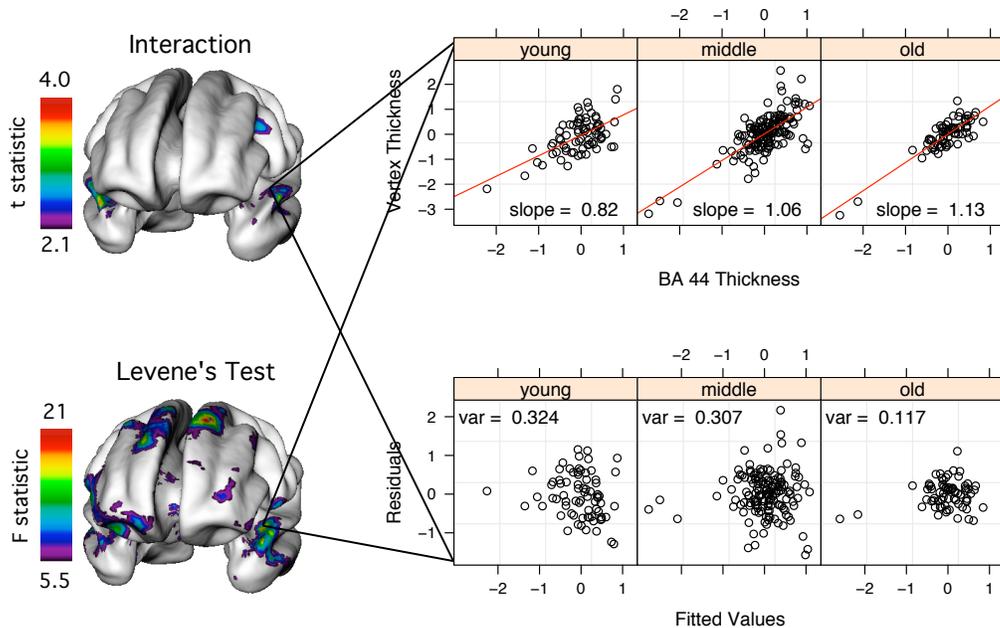


Figure 6.5: MACACC age difference. Significant age by BA 44 interactions can be found in the bilateral STG, though the main effect appears to be in the differences in variance between the groups, where the bilateral STG are significant as well as multiple frontal and lateral temporal lobe areas. The plot shows one vertex in the left STG, describing both the the slopes (top) as well as variance in a plot of residuals against fitted values (bottom) for each of the three age groups (young, middle, old; see text).

6.5.4. Variance tests

As discussed in the methods section, multiple tests for group differences in variance were assessed. The results of different variance tests are shown in figure 6.6; the overall patterns are clearly similar. Robustness to outliers, however, is where the tests differ ed most. Figure 6.6 shows that a single outlier can have tremendous influence on the estimation of variance, and that Levene's Test is the most robust to this outlier. For that reason Levene's Test was chosen for testing MACACC-variance.

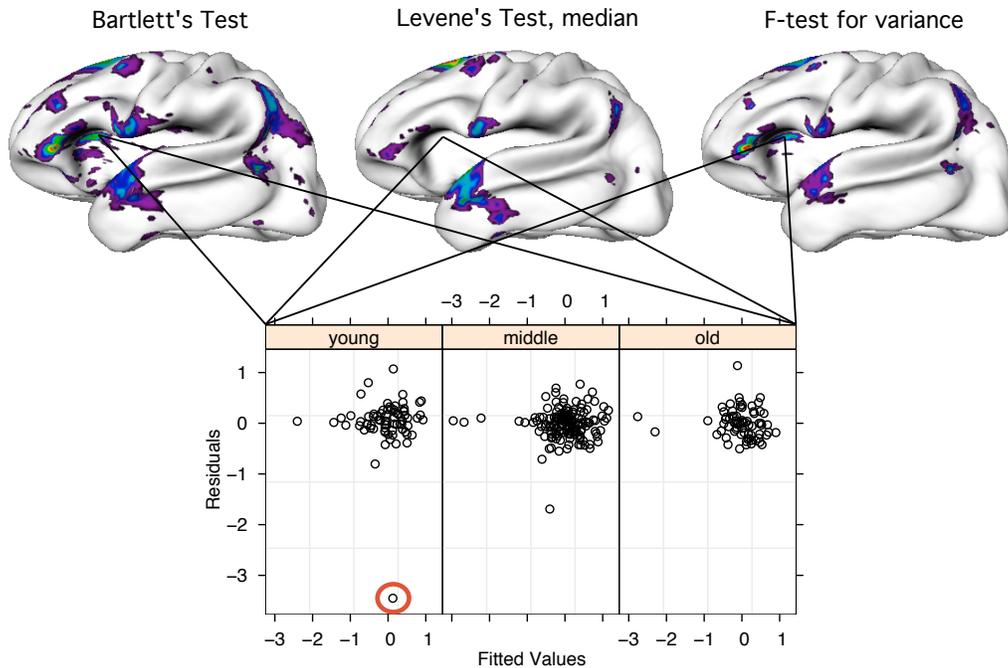


Figure 6.6: MACACC-variance tests: the differences in variance as estimated using three statistical tests are shown, revealing similar patterns of significance. One difference, however, can be found in the left insula, and a plot of the residuals for all 292 subjects at this vertex shows that a single outlier (circled in red, and likely a methodological error) is driving significance in the F-test for variance and Bartlett's test. Levene's test, on the other hand, is robust against the influence of this outlier.

6.5.5. IQ MACACC differences

General measures of intelligence correlate with global grey matter volumes (Andreasen et al., 1993; Reiss et al., 1996; Pennington et al., 2000; Posthuma et al., 2003; Wilke et al., 2003; Haier et al., 2004). Correlations with particular cortical areas are, however, limited. This leads to the hypothesis that IQ is a property of the cortex better understood in terms of interactions between different cerebral regions (Wilke et al.,

2003). We tested this hypothesis by dividing our population into low (< 100) and high (>120) IQ sub-samples and testing for MACACC-slope and MACACC-variance differences with left BA 44 as the seed region. The results are shown in figure 6.7. The principal differences were found in the MACACC-slope analysis in the ventro- and dorso-lateral prefrontal cortex, the lateral parietal lobes, and the anterior cingulate. The focus on the cingulate confirms previous IQ morphometry studies (Wilke et al., 2003), and extends them by showing a network of regions differentiating low and high IQ subjects involving prefrontal, lateral parietal, and anterior temporal lobes.

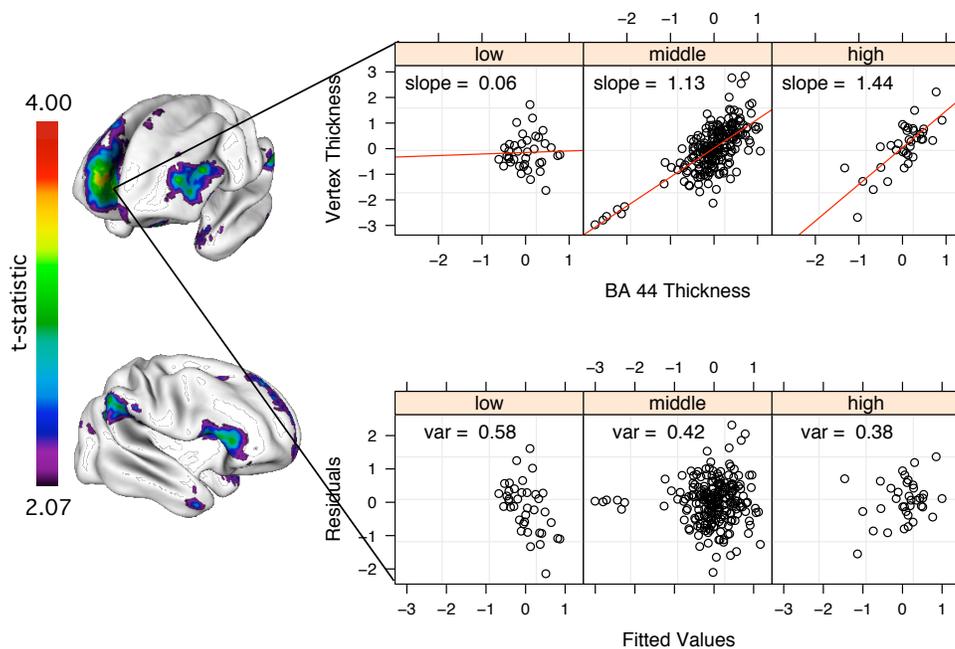


Figure 6.7: IQ differences: This figure shows the MACACC differences between low and high IQ grouping in the population. There are significant MACACC-slope differences in the ventro and dorsolateral prefrontal cortex, the lateral parietal lobes, and, most significantly, the right anterior cingulate. The graph shows the BA 44 to cingulate slope (top) and variance (bottom) differences. Note the difference in slopes between the IQ groups.

6.6. Discussion

We have introduced a new strategy for analysing correlated changes in cortical morphology within a population sample. The MACACC methods can be used to study the correlated connectivity of a seed region, to assess the total correlation strength at every point in the cortex, and to investigate differences in cross-cortical correlations between subject groupings. We chose the correlations approach as it was found to have greater regional specificity than related techniques such as singular value decomposition (Worsley et al., 2005). MACACC-strength tests, which reveal the relative inter-connectivity of every vertex, require no prior hypotheses, and can thus be used to assess global population differences. MACACC-slope and MACACC-variance tests, on the other hand, are designed to be used with a seed region, and thus require a prior hypothesis. Given a seed region, precise investigations of group differences within a cortical network involving the seed region can then be obtained and tested for significance.

The results from the normal brain development population are intriguing and encouraging. While MACACC maps are not the same as anatomical connectivity measured with fibre tracking, the correlation maps of BA 44 do bear a striking resemblance to the diffusion tensor maps of the arcuate fasciculus (figure 6.4), suggesting that they are measuring two aspects of the same underlying processes. In the population studied, different cortical areas within the language network correlated tightly; we thus predict that future MACACC studies will show similar structural relationships in different cortical networks.

The MACACC map of BA 44 showed an age-related tightening of the

correlation with the superior temporal gyrus (figure 6.5), which is connected to BA 44 through the arcuate and uncinate fasciculi (Parker et al., 2005). The STG, like BA 44, is involved in the anatomical network underlying language (Dronkers et al., 2004; Matsumoto et al., 2004; Parker et al., 2005) and auditory function (Platel et al., 1997; Romanski et al., 1999; Kiehl et al., 2001; Jeffries et al., 2003; Kim et al., 2004). The tightening of correlations is occurring concurrently with expanding language function within the age range studied (Holland et al., 2001; Schmithorst et al., 2002). This suggests that MACACC is capturing an important part of the normal brain development process, wherein the thickness of the cortex between areas subserving related functional specializations show greater similarity to each other with age, while simultaneously white matter tracts connecting these areas mature (Paus et al., 1999; Paus et al., 2001; Schmithorst et al., 2002).

We furthermore report stepwise increases in connectivity between BA 44, the ACC and parietal cortex linked to increasing intellectual ability, reflected by IQ (figure 6.7). Meta-analyses of fMRI studies have implicated the inferior frontal gyrus as a core region in supporting diverse cognitive tasks which supports our use of BA 44 as a 'seed' region (Duncan et al., 2000). The strengthened pattern of correlations between this hub and fronto-parietal regions bear a striking resemblance to the functional activation network shown by Gray to vary with intelligence (Gray et al., 2003). We identify the ACC and parietal cortex as structural correlates of the lateral frontal cortex regions which have been implicated in the allocation of attentional resources, planning and cognitive control, all core facets of intelligence (Andreasen et al., 1993; Reiss et al., 1996; Rubia et al., 2000; Paus, 2001; Booth et al., 2003; Wilke et al., 2003; Haier et al., 2004; Toga and Thompson, 2004).

What is particularly interesting is that the type of group differences in the IQ and age studies varies: the main results for age differences shows a change in MACACC-variance, as the correlations tighten with age. With IQ, on the other hand, the key findings are in MACACC-slope, as the type of correlation with the seed region is different in the high and low IQ groups. The tightening correlations with age capture an ongoing developmental process. Different slopes in the IQ case suggest that the cortical areas implicated interact in a differently coupled fashion. The relationship between BA 44 and the anterior cingulate graphed in figure 6.7 indicates that, in the low IQ group, the thickness of the cingulate barely increases with an increase in the cortical thickness of BA 44. In the high IQ group, on the other hand, an increase of 1mm in BA 44 predicts an increase of 1.44mm in the cingulate. The fact that such different relationships can exist within the same population divided by IQ suggests that the biological underpinnings of such changes should be further investigated.

Based on the above discussion we predict that degenerative disorders such as Alzheimer's Disease will show increasing dissociation between related areas, resulting in significant MACACC-variance differences. Disorders such as schizophrenia, hypothesized to be a connectivity disorder (Pearlson, 1999; Penn, 2001; Innocenti et al., 2003; de Haan and Bakker, 2004), would likely feature both MACACC-slope and variance differences.

Studying MACACC opens a new field of study to the imaging neurosciences. It is conceivable that certain anatomical patterns cannot be detected by simple models which relate morphometry to group or morphometry to cognitive/clinical variable, but instead lie in the relationship between regions of the cortex. To give an example,

Alzheimer's Disease involves characteristic thinning of the cortex (Gomez-Isla et al., 1997; Lerch et al., 2004). What is less certain is how early into the disease progression cortical thinning becomes detectable, especially given the noise inherent in an in-vivo imaging study. A potentially more sensitive indicator of onset of the AD pathological processes is the statistical relationship between the limbic system structures known to be involved in the early stages of the disease (Braak and Braak, 1991; Delacourte et al., 1999).

The MACACC methods described here still need to be extended and further explored. The most important addition will be to extend MACACC to longitudinal and mixed models, where one can look for changes in correlation not just across populations but within single subjects. Further work also needs to include better handling of continuous data for MACACC-variance tests. Methods such as Levene's and Bartlett's test need discrete groupings. One potential solution is to use general least squares with a weights structure reflecting change in variance with the continuous variable, and perform a likelihood ratio test to ascertain whether allowing for non-uniform error improves the model fit.

To conclude, we propose that studying MACACC opens a new field of study to the imaging neurosciences, providing a structural methodology for delineating cortical correlations which complements current techniques defining anatomical and functional connectivity.

6.7. Acknowledgments

ICBM grant PO1MHO52176-11, principal investigator Dr John Mazziotta; CIHR grant MOP-34996. Jason Lerch is funded by a K.M. Hunter/CIHR Doctoral Research Award. We thank Dr. Alex Zijdenbos for his many

insightful comments on this work.

6.8. Appendix

Name	Definition
Correlation	Measured using Pearson's r , relates to the strength of the correlation between a seed region and the rest of the cortex.
MACACC-strength	The average correlation coefficient when every vertex is correlated with every other vertex.
MACACC-slope	Shows a group difference in the estimated slope between the seed region and target vertex.
MACACC-variance	Shows a group difference in the variance around the estimated slope between the seed region and target vertex.

Table 6.2: Definition of terms: a list of the definitions used throughout the paper.

Chapter 7: Conclusions

7.1. Summary and Implications

This dissertation described a series of investigations designed to establish in-vivo analysis of cortical thickness using magnetic resonance images (MRI) as a viable technique in human brain imaging. The goals were to establish confidence in cortical thickness as a viable metric through both simulations and analyses of real data, to consider possible clinical applications of thickness measurements, and to expand our knowledge of cortical morphology by examining within cortex correlations.

Chapter 3 was the key in establishing the methodological validity of cortical thickness measurements. A previous study of the accuracy of cortical thickness (Kabani et al., 2001) provided a necessary base, but did not answer the question of how precise thickness measures could be at recovering a known change in a simulated population study. Moreover, our study design could examine which of the six available cortical thickness measures provided optimal performance. Choices of blurring kernel size and multiple comparisons correction methods were also amenable to investigation. The results showed that the best metric, t_{link} , could obtain a

high sensitivity of 0.93 with 100% specificity, that blurring on the surface is clearly necessary and should ideally be in the 20-35mm range, and that the different multiple comparisons correction techniques provided varying sensitivity to both false positives and false negatives. A power analysis found that, given two groups of 25 subjects each, a 0.6mm (15%) change in cortical thickness can be found. Along with giving confidence in the use of cortical thickness for populations studies, this paper also provided the parameters and choice of thickness metric that were used for the remaining studies and ongoing collaborations.

In Chapter 4 cortical thickness analysis was applied to an Alzheimer's Disease dataset. We hoped to both increase face validity of our image processing pipeline by generating believable results as well as increase our understanding of the role of the cortex in AD. The results showed a pattern of cortical atrophy in line with prediction based on the known advance of histopathology in the disease (Braak and Braak, 1995; Braak and Braak, 1996; Delacourte et al., 1999). The amount of thinning was quantified, reaching a maximum of 1.25mm in the parahippocampal gyrus. Correlations of cortical thickness with disease progression and MMSE scores showed increasing atrophy in the parahippocampal gyrus, lateral temporal and frontal lobes.

The same dataset was used to assess the potential clinical use of cortical thickness for disease diagnosis purposes, described in Chapter 5. Given mean thickness measures over the parahippocampal gyrus and one of six other structures we showed cortical thickness to be capable of perfectly reproducing the initial clinical diagnosis. We further provided maps showing the diagnostic accuracy of every single vertex. Here maps of specificity resembled the expected pattern of AD atrophy, whereas maps of sensitivity were more diffuse. Different discriminant techniques were

compared and found to be roughly comparable in performance.

The last analysis, presented in Chapter 6, created a framework for studying cross-cortical correlations of cortical thickness and applied these techniques to a normal brain development dataset of 292 subjects. This analysis, dubbed Mapping Anatomical Correlations Across Cerebral Cortex (MACACC), provided a correlation map of Brodmann Area 44 which bore striking resemblance to diffusion tensor fibre maps of the same area. These correlational maps moreover tightened with age, showing a developmental pattern of cross-cortical correlations. They also changed with IQ, the high IQ sample having strong correlations with the anterior cingulate and lateral parietal lobes, something absent in the low IQ group. Lastly, a map of correlation strength of each vertex was shown, which featured the association cortices as the most correlated areas.

The results described above were underpinned by methodological development, particularly in the areas of image processing pipelines (tying together the individual tools that process the raw MR data to produce maps of cortical thickness), statistics of data on the surface, and visualization. The analysis framework described herein is available for automated analysis of MR data from arbitrary populations. Cortical thickness analysis has thus become a tool applicable to wide-scale research in human neuroanatomy.

7.2. Future Work

The findings described in this dissertation call for a number of future investigations. They include further validation work, applications to new datasets, and continuing advances in the underlying image processing algorithms. The most important extension is to investigate the accuracy of

our automated cortical thickness measurements using high-resolution post-mortem datasets together with MRIs in the same subject in order to allow for correlations between MR thickness and histology. It will be important to perform these analyses entirely in three dimensions to avoid some of the confounds usually associated with thickness measurements from post-mortem slices. Such a dataset would also inform about the neuroanatomical factors that decrease accuracy, and shed light on some of the discrepancies between the maps of cortical thickness produced from post-mortem brains by von Economo and Koskinas (Economo and Koskinas, 1925) and modern MR. Lastly, post-mortem data could be used to construct new geometric definitions of cortical thickness based upon information not available with the current MR resolution such as cortical layers and potentially even columns.

The population-simulation based validation described in Chapter 3 can also be extended to test for additional variables in the image processing pipeline. How does increasing the number of polygons in the cortical mesh change precision? What effect does non-linear alignment of cortical meshes (Robbins et al., 2004) have on sensitivity and specificity? Does the number of polygons or the use of non-linear alignment change the optimal blurring kernel size? Do other types of surface based smoothing increase or decrease performance?

Our investigations of the cortex in Alzheimer's Disease leave many questions that can be answered with additional data. The progression of AD measured in Chapter 4, for example, should be repeated in a larger sample with multiple timepoints per subject in order to provide a better understanding of the potential non-linear nature of atrophy in the disease. The MMSE is a weak indicator of mental health; multiple other cognitive tests are available whose scores could be correlated with cortical thickness. In the realm of diagnosis the population should be expanded to

include different types of dementia in order to allow for differential discrimination. Earlier onset AD cases should be used, and ideally even prospective data ought to be examined to see if cortical morphometry can detect dementia onset before the clinical diagnosis can be made.

The number of additional application domains to apply cortical thickness analysis to is large. Multiple diseases and disorders implicate the cortex; all could be analysed using the methods described in this dissertation. Changes in the structure of cortex in normal populations are also fascinating avenues of investigation. Does the acquisition of new skills increase cortical thickness? Several studies would suggest that this is indeed the case (Schlaug, 2001; Golestani et al., 2002; Gaser and Schlaug, 2003a; Gaser and Schlaug, 2003b; Draganski et al., 2004). If so, what is the timecourse of these changes? Do different types of skill acquisition merely change the location of cortical thickness alterations, or is the extent and type of change variant as well?

MACACC maps are especially open to further investigations due to their novel nature. Multiple previous studies have investigated morphological changes with a disorder or correlated against cognitive variables, but very few have looked at changes in cross-cortical correlations within these same populations. Are MACACC strength maps, for example, more sensitive to early disease detection than standard point by point discriminants? One would expect that the standard correlations between limbic lobe structures would break down as AD pathology advances, but this needs to be tested. We have so far only tested MACACC maps of BA 44 in a normal development population; the number of extensions to different populations, different cortical areas, and different groupings based on clinical or cognitive variables is staggering.

The image processing that underlies cortical thickness analysis, while very

robust and powerful, can always be improved as well. As resolution and field strength of MR increases some of the standard assumptions made will have to be adapted. The non-uniformity correction approach used herein (Sled et al., 1998) will have to be modified to deal with the large distortions in 3T and higher MR machines. The assumption, made by the tissue classification algorithm, that grey matter, white matter, and CSF are uniform throughout the brain is clearly wrong. As the resolution and SNR of MRI improves, this will have to be modified. The various algorithms in the pipeline will also have to be tested in abnormal cortices, such as those featured in cases of polymicrogyria or double cortices. Lastly, smaller changes to each part of the pipeline can always be made to improve its overall performance.

The goal of this thesis, as mentioned above, was to establish confidence in cortical thickness as a viable metric in human brain imaging. The four studies that compose the core of the dissertation showed that cortical thickness analysis produces precise results in a simulation, can quantify atrophy in pathology, can act as a diagnostic aid, and shows patterns of cross-cortical correlations whose potential is just beginning to be explored. The growing number of studies employing cortical thickness using both the methods developed at the Montreal Neurological Institute as well as elsewhere show that the main goal of this thesis has been accomplished.

Bibliography

Ad-Dabbagh, Y., Singh, V., Robbins, S., Lerch, J. P., Lyttelton, O., Fombonne, E., et al. (2005). Native-Space Cortical Thickness Measurement and the Absence of Correlation to Cerebral Volume. (Eds.), 11 Toronto:

Amunts, K., Weiss, P. H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J. M., et al. (2004). Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space--the roles of Brodmann areas 44 and 45. *Neuroimage*, 22(1), 42-56.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B., & Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *J Comp Neurol*, 412(2), 319-41.

Anderton, B. H. (2002). Ageing of the brain. *Mech Ageing Dev*, 123(7), 811-17.

Andreasen, N. C., Flaum, M., Swayze, V. n., O'Leary, D. S., Alliger, R., Cohen, G., et al. (1993). Intelligence and brain structure in normal individuals. *Am J Psychiatry*, 150(1), 130-34.

Antel, S. B., Bernasconi, A., Bernasconi, N., Collins, D. L., Kearney, R. E., Shinghal, R., et al. (2002). Computational models of MRI characteristics of focal cortical dysplasia improve lesion detection. *Neuroimage*, 17(4), 1755-60.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry--the methods. *Neuroimage*, 11(6 Pt 1), 805-21.

Ashtari, M., Kumra, S., Bhaskar, S. L., Clarke, T., Thaden, E., Cervellione, K. L., et al. (2005). Attention-deficit/hyperactivity disorder: a preliminary diffusion tensor imaging study. *Biol Psychiatry*, 57(5), 448-55.

Baron, J. C., Chetelat, G., Desgranges, B., Perchey, G., Landeau, B., de la Sayette, V., et al. (2001). In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. *Neuroimage*, 14(2), 298-309.

Barta, P., Miller, M. I., & Qiu, A. (2005). A stochastic model for studying the laminar structure of cortex from MRI. *IEEE Trans Med Imaging*, 24(6), 728-42.

Bobb, A. J., Addington, A. M., Sidransky, E., Gornick, M. C., Lerch, J. P., Greenstein, D. K., et al. (2005). Support for association between ADHD and two candidate genes: NET1 and DRD1. *Am J Med Genet B Neuropsychiatr Genet*,

Bobinski, M., de Leon, M. J., Convit, A., De Santi, S., Wegiel, J., Tarshish, C. Y., et al. (1999). MRI of entorhinal cortex in mild Alzheimer's disease. *Lancet*, 353(9146), 38-40.

Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Trommer, B. L., Davenport, N. D., et al. (2003). Neural development of selective attention and response inhibition. *Neuroimage*, 20(2), 737-51.

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol (Berl)*, 82(4), 239-59.

Braak, H., & Braak, E. (1995). Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol Aging*, 16(3), 271-8; discussion 278-84.

Braak, H., & Braak, E. (1996). Evolution of the neuropathology of Alzheimer's disease. *Acta Neurol Scand Suppl*, 165, 3-12.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth.

Buckner, R. L., Snyder, A. Z., Sanders, A. L., Raichle, M. E., & Morris, J. C. (2000). Functional brain imaging of young, nondemented, and demented older adults. *J Cogn Neurosci*, 12 Suppl 2, 24-34.

Buckner, R. L., Raichle, M. E., & Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *J Neurophysiol*, 74(5), 2163-73.

Buldyrev, S. V., Cruz, L., Gomez-Isla, T., Gomez-Tortosa, E., Havlin, S., Le, R., et al. (2000). Description of microcolumnar ensembles in association cortex and their disruption in Alzheimer and Lewy body dementias. *Proc Natl Acad Sci U S A*, 97(10), 5039-43.

Busatto, G. F., Garrido, G. E., Almeida, O. P., Castro, C. C., Camargo, C. H., Cid, C. G., et al. (2003). A voxel-based morphometry study of temporal lobe gray matter reductions in Alzheimer's disease. *Neurobiol Aging*, 24(2), 221-31.

Buxhoeveden, D., Lefkowitz, W., Loats, P., & Armstrong, E. (1996).

The linear organization of cell columns in human and nonhuman anthropoid Tpt cortex. *Anat Embryol (Berl)*, 194(1), 23-36.

Callen, D. J., Black, S. E., Gao, F., Caldwell, C. B., & Szalai, J. P. (2001). Beyond the hippocampus: MRI volumetry confirms widespread limbic atrophy in AD. *Neurology*, 57(9), 1669-74.

Cannistraro, P. A., & Rauch, S. L. (2003). Neural circuitry of anxiety: evidence from structural and functional neuroimaging studies. *Psychopharmacol Bull*, 37(4), 8-25.

Casanova, M. F., Araque, J., Giedd, J., & Rumsey, J. M. (2004). Reduced brain size and gyrification in the brains of dyslexic patients. *J Child Neurol*, 19(4), 275-81.

Chen, J. T., Narayanan, S., Collins, D. L., Smith, S. M., Matthews, P. M., & Arnold, D. L. (2004). Relating neocortical pathology to disability progression in multiple sclerosis using MRI. *Neuroimage*, 23(3), 1168-75.

Chertkow, H., & Murtha, S. (1997). PET activation and language. *Clin Neurosci*, 4(2), 78-86.

Chetelat, G., & Baron, J. C. (2003). Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. *Neuroimage*, 18(2), 525-41.

Chetelat, G., Desgranges, B., De La Sayette, V., Viader, F., Eustache, F., & Baron, J. C. (2002). Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. *Neuroreport*, 13(15), 1939-43.

Chung, M. K., Robbins, S. M., Dalton, K. M., Davidson, R. J., Alexander, A. L., & Evans, A. C. (2005). Cortical thickness analysis in autism with heat kernel smoothing. *Neuroimage*, 25(4), 1256-65.

Chung, M. K., Worsley, K. J., Paus, T., Robbins, S., Evans, A. C., Taylor, J., et al. (2002). Tensor-Based Surface Morphometry.

Chung, M. K., Worsley, K. J., Robbins, S., Paus, T., Taylor, J., Giedd, J. N., et al. (2003). Deformation-based surface morphometry applied to gray matter deformation. *Neuroimage*, 18(2), 198-213.

Cohen, L. G., Bandinelli, S., Findley, T. W., & Hallett, M. (1991). Motor reorganization after upper limb amputation in man. A study with focal magnetic stimulation. *Brain*, 114 (Pt 1B), 615-27.

Collins, D. L., Holmes, C. J., Peters, T. M., & Evans, A. C. (1995). Automatic 3D model-based neuroanatomical segmentation. *Human Brain Mapping*, 3, 190-208.

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994).

Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr*, 18(2), 192-205.

Cronin-Golomb, A., Corkin, S., Rizzo, J. F., Cohen, J., Growdon, J. H., & Banks, K. S. (1991). Visual dysfunction in Alzheimer's disease: relation to normal aging. *Ann Neurol*, 29(1), 41-52.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, 9(2), 179-94.

Davatzikos, C., & Bryan, R. N. (2002). Morphometric analysis of cortical sulci using parametric ribbons: a study of the central sulcus. *J Comput Assist Tomogr*, 26(2), 298-307.

de Haan, L., & Bakker, J. M. (2004). Overview of neuropathological theories of schizophrenia: from degeneration to progressive developmental disorder. *Psychopathology*, 37(1), 1-7.

de Leon, M. J., Convit, A., DeSanti, S., Bobinski, M., George, A. E., Wisniewski, H. M., et al. (1997). Contribution of structural neuroimaging to the early diagnosis of Alzheimer's disease. *Int Psychogeriatr*, 9 Suppl 1, 183-90; discussion 247-52.

De Toledo-Morrell, L., Goncharova, I., Dickerson, B., Wilson, R. S., & Bennett, D. A. (2000). From healthy aging to early Alzheimer's disease: in vivo detection of entorhinal cortex atrophy. *Ann N Y Acad Sci*, 911, 240-53.

Delacourte, A., David, J. P., Sergeant, N., Buee, L., Wattez, A., Vermersch, P., et al. (1999). The biochemical pathway of neurofibrillary degeneration in aging and Alzheimer's disease. *Neurology*, 52(6), 1158-65.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature*, 427(6972), 311-12.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D. J., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1-2), 145-77.

Drury, H. A., Van Essen, D. C., Anderson, C. H., Lee, C. W., Coogan, T. A., & Lewis, J. W. (1996). Computerized mappings of the cerebral cortex: a multiresolution flattening method and a surface-based coordinate system. *J Cogn Neurosci*, 8(1), 1-28.

Du, A. T., Schuff, N., Kramer, J. H., Ganzer, S., Zhu, X. P., Jagust, W. J., et al. (2004). Higher atrophy rate of entorhinal cortex than hippocampus

in *AD. Neurology*, 62(3), 422-27.

Duffau, H., Capelle, L., Sichez, N., Denvil, D., Lopes, M., Sichez, J. P., et al. (2002). Intraoperative mapping of the subcortical language pathways using direct stimulations. An anatomo-functional study. *Brain*, 125(Pt 1), 199-214.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science*, 289(5478), 457-60.

Tosun, D., Rettman, M. E., Tao, X., Han, X., Xu, C., & Prince, J. L. (2001). Calculation of human cerebral cortical thickness on opposing sulcal banks. *Proceedings of 7th International Conference on Functional Mapping of the Human Brain*,

Economo, C., & Koskinas, G. N. (1925). *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*. Wien und Berlin: J. Springer.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, 270(5234), 305-07.

Elston, G. N. (2003). Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb Cortex*, 13(11), 1124-38.

Ezekiel, F., Chao, L., Kornak, J., Du, A. T., Cardenas, V., Truran, D., et al. (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 Dis Assoc Disord*, 18(4), 196-201.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Hum Brain Mapp*, 24(2), 79-91.

Fields, S. D., Fulop, G., Sachs, C. J., Strain, J., & Fillit, H. (1992). Usefulness of the Neurobehavioral Cognitive Status Examination in the hospitalized elderly. *Int Psychogeriatr*, 4(1), 93-102.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3), 341-55.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2), 195-207.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*, 97(20), 11050-55.

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. *J Psychiatr Res*, 12, 189-98.

Fox, N. C., Cousens, S., Scahill, R., Harvey, R. J., & Rossor, M. N. (2000). Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects. *Arch Neurol*, 57(3), 339-44.

Friedman, L., Kenny, J. T., Wise, A. L., Wu, D., Stuve, T. A., Miller, D. A., et al. (1998). Brain activation during silent word generation evaluated with functional MRI. *Brain Lang*, 64(2), 231-56.

Friston, K. (2002). Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annu Rev Neurosci*, 25, 221-50.

Friston, K. J., Frith, C. D., Fletcher, P., Liddle, P. F., & Frackowiak, R. S. (1996). Functional topography: multidimensional scaling and functional connectivity in the brain. *Cereb Cortex*, 6(2), 156-64.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. (1993). Functional connectivity: the principal-component analysis of large (PET) data sets. *J Cereb Blood Flow Metab*, 13(1), 5-14.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, 19(4), 1273-302.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. (1991). A PET study of word finding. *Neuropsychologia*, 29(12), 1137-48.

Gaser, C., & Schlaug, G. (2003a). Gray matter differences between musicians and nonmusicians. *Ann N Y Acad Sci*, 999, 514-17.

Gaser, C., & Schlaug, G. (2003b). Brain structures differ between musicians and non-musicians. *J Neurosci*, 23(27), 9240-45.

Gee, J., Ding, L., Xie, Z., Lin, M., DeVita, C., & Grossman, M. (2003). Alzheimer's disease and frontotemporal dementia exhibit distinct atrophy-behavior correlates: a computer-assisted imaging study. *Acad Radiol*, 10(12), 1392-401.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4), 870-78.

Geschwind, N., & Levitsky, W. (1968). Human brain: left-right asymmetries in temporal speech region. *Science*, 161(837), 186-87.

Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Ann N Y Acad Sci*, 1021, 77-85.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A*, 101(21), 8174-79.

Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron*, 35(5), 997-1010.

Gomez-Isla, T., Hollister, R., West, H., Mui, S., Growdon, J. H., Petersen, R. C., et al. (1997). Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease. *Ann Neurol*, 41(1), 17-24.

Good, C. D., Scahill, R. I., Fox, N. C., Ashburner, J., Friston, K. J., Chan, D., et al. (2002). Automatic differentiation of anatomical patterns in the human brain: validation with studies of degenerative dementias. *Neuroimage*, 17(1), 29-46.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nat Neurosci*, 6(3), 316-22.

Grignon, Y., Duyckaerts, C., Bennechib, M., & Hauw, J. J. (1998). Cytoarchitectonic alterations in the supramarginal gyrus of late onset Alzheimer's disease. *Acta Neuropathol (Berl)*, 95(4), 395-406.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2004). Structural brain variation and general intelligence. *Neuroimage*, 23(1), 425-33.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning*. New York: Springer.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., & Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *Eur J Neurosci*, 17(5), 1123-28.

Holland, S. K., Plante, E., Weber Byars, A., Strawsburg, R. H., Schmithorst, V. J., & Ball, W. S. J. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage*, 14(4), 837-43.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *J Comput Assist Tomogr*, 22(2), 324-33.

Horwitz, B. (2003). The elusive concept of brain connectivity. *Neuroimage*, 19(2 Pt 1), 466-70.

Hund-Georgiadis, M., Yves Von Cramon, D., Kruggel, F., & Preul, C. (2002). Do quiescent arachnoid cysts alter CNS functional organization?: A fMRI and morphometric study. *Neurology*, 59(12), 1935-39.

Hurdal, M. K., & Stephenson, K. (2004). Cortical cartography using the discrete conformal approach of circle packings. *Neuroimage*, 23 Suppl 1, S119-28.

Innocenti, G. M., Ansermet, F., & Parnas, J. (2003). Schizophrenia, neurodevelopment and corpus callosum. *Mol Psychiatry*, 8(3), 261-74.

Janke, A. L., de Zubicaray, G., Rose, S. E., Griffin, M., Chalk, J. B., & Galloway, G. J. (2001). 4D deformation modeling of cortical disease progression in Alzheimer's dementia. *Magn Reson Med*, 46(4), 661-66.

Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003). Words in melody: an H(2)15O PET study of brain activation during singing and speaking. *Neuroreport*, 14(5), 749-54.

Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Gamst, A. C., Stout, J. C., Bonner, J., et al. (2001). Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol Aging*, 22(4), 581-94.

Jones, S. E., Buchbinder, B. R., & Aharon, I. (2000). Three-dimensional mapping of cortical thickness using Laplace's equation. *Hum Brain Mapp*, 11(1), 12-32.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.

Juottonen, K., Laakso, M. P., Insausti, R., Lehtovirta, M., Pitkanen, A., Partanen, K., et al. (1998). Volumes of the entorhinal and perirhinal cortices in Alzheimer's disease. *Neurobiol Aging*, 19(1), 15-22.

Juottonen, K., Laakso, M. P., Partanen, K., & Soininen, H. (1999). Comparative MR analysis of the entorhinal cortex and hippocampus in diagnosing Alzheimer disease. *AJNR Am J Neuroradiol*, 20(1), 139-44.

Jurgens, U., Kirzinger, A., & von Cramon, D. (1982). The effects of deep-reaching lesions in the cortical face area on phonation. A combined case report and experimental monkey study. *Cortex*, 18(1), 125-39.

Kabani, N., Le Goualher, G., MacDonald, D., & Evans, A. C. (2001). Measurement of cortical thickness using an automated 3-D algorithm: a validation study. *Neuroimage*, 13(2), 375-80.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of*

neural science. New York: McGraw-Hill, Health Professions Division.

Kiehl, K. A., Laurens, K. R., Duty, T. L., Forster, B. B., & Liddle, P. F. (2001). Neural sources involved in auditory target detection and novelty processing: an event-related fMRI study. *Psychophysiology*, 38(1), 133-42.

Killiany, R. J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., et al. (2000). Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. *Ann Neurol*, 47(4), 430-39.

Killiany, R. J., Moss, M. B., Albert, M. S., Sandor, T., Tieman, J., & Jolesz, F. (1993). Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. *Arch Neurol*, 50(9), 949-54.

Kim, D. E., Shin, M. J., Lee, K. M., Chu, K., Woo, S. H., Kim, Y. R., et al. (2004). Musical training-induced functional reorganization of the adult brain: functional magnetic resonance imaging and transcranial magnetic stimulation study on amateur string players. *Hum Brain Mapp*, 23(4), 188-99.

Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., Macdonald, D., et al. (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage*,

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. *Proc Natl Acad Sci U S A*, 92(7), 2899-903.

Klein, D., Olivier, A., Milner, B., Zatorre, R. J., Johnsrude, I., Meyer, E., et al. (1997). Obligatory role of the LIFG in synonym generation: evidence from PET and cortical stimulation. *Neuroreport*, 8(15), 3275-79.

Kollokian, V. (1996). Performance analysis of automatic techniques for tissue classification in magnetic resonance images of the human brain (Doctoral dissertation, McGill University, 1996).

Koski, L., & Paus, T. (2000). Functional connectivity of the anterior cingulate cortex within the human frontal lobe: a brain-mapping meta-analysis. *Exp Brain Res*, 133(1), 55-65.

Krasuski, J. S., Alexander, G. E., Horwitz, B., Daly, E. M., Murphy, D. G., Rapoport, S. I., et al. (1998). Volumes of medial temporal lobe structures in patients with Alzheimer's disease and mild cognitive impairment (and in healthy controls). *Biol Psychiatry*, 43(1), 60-68.

Kuperberg, G. R., Broome, M. R., McGuire, P. K., David, A. S., Eddy,

M., Ozawa, F., et al. (2003). Regionally localized thinning of the cerebral cortex in schizophrenia. *Arch Gen Psychiatry*, 60(9), 878-88.

Le Goualher, G., Argenti, A. M., Duyme, M., Baare, W. F., Hulshoff Pol, H. E., Boomsma, D. I., et al. (2000). Statistical sulcal shape comparisons: application to the detection of genetic encoding of the central sulcus shape. *Neuroimage*, 11(5 Pt 1), 564-74.

Le Goualher, G., Procyk, E., Collins, D. L., Venugopal, R., Barillot, C., & Evans, A. C. (1999). Automated extraction and variability analysis of sulcal neuroanatomy. *IEEE Trans Med Imaging*, 18(3), 206-17.

Lee, L., Harrison, L. M., & Mechelli, A. (2003). A report of the functional connectivity workshop, Dusseldorf 2002. *Neuroimage*, 19(2 Pt 1), 457-65.

Lehericy, S., Baulac, M., Chiras, J., Pierot, L., Martin, N., Pillon, B., et al. (1994). Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. *AJNR Am J Neuroradiol*, 15(5), 929-37.

Lerch, J. P., & Evans, A. C. (2005). Cortical thickness analysis examined through power analysis and a population simulation. *Neuroimage*, 24(1), 163-73.

Lerch, J. P., Pruessner, J. C., Zijdenbos, A., Hampel, H., Teipel, S. J., & Evans, A. C. (2004). Focal Decline of Cortical Thickness in Alzheimer's Disease Identified by Computational Neuroanatomy. *Cereb Cortex*,

Liu, T., Shen, D., & Davatzikos, C. (2004). Deformable registration of cortical structures via hybrid volumetric and surface warping. *Neuroimage*, 22(4), 1790-801.

Lohmann, G., Preul, C., & Hund-Georgiadis, M. (2003). Morphology-based cortical thickness estimation. *Inf Process Med Imaging*, 18, 89-100.

MacDonald, D. (1997). A method for identifying geometrically simple surfaces from three dimensional images (Doctoral dissertation, McGill University, 1997).

MacDonald, D., Kabani, N., Avis, D., & Evans, A. C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage*, 12(3), 340-56.

Magnotta, V. A., Andreasen, N. C., Schultz, S. K., Harris, G., Cizadlo, T., Heckel, D., et al. (1999). Quantitative in vivo measurement of gyrification in the human brain: changes associated with aging. *Cereb Cortex*, 9(2), 151-60.

Magoun, H. W., Ph.D., Marshall, L. H., Ph.D., Marshall, L. H., &

Magoun, H. W. (1998). *Discoveries in the Human Brain: Neuroscience Prehistory, Brain Structure, and Function*. Humana Press.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proc Natl Acad Sci U S A*, 97(8), 4398-403.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S., & Burgess, N. (2003). Navigation expertise and the human hippocampus: a structural brain imaging analysis. *Hippocampus*, 13(2), 250-59.

Makuuchi, M. (2004). Is Broca's Area Crucial for Imitation? *Cereb Cortex*,

Mangin, J. F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., et al. (2004). A framework to study the cortical folding patterns. *Neuroimage*, 23 Suppl 1, S129-38.

Matsuda, H., Kitayama, N., Ohnishi, T., Asada, T., Nakano, S., Sakamoto, S., et al. (2002). Longitudinal evaluation of both morphologic and functional changes in the same individuals with Alzheimer's disease. *J Nucl Med*, 43(3), 304-11.

Matsumoto, R., Nair, D. R., LaPresto, E., Najm, I., Bingaman, W., Shibasaki, H., et al. (2004). Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain*, 127(Pt 10), 2316-30.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., et al. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos Trans R Soc Lond B Biol Sci*, 356(1412), 1293-322.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34(7), 939-44.

Meyer, J. R., Roychowdhury, S., Russell, E. J., Callahan, C., Gitelman, D., & Mesulam, M. M. (1996). Location of the central sulcus via cortical thickness of the precentral and postcentral gyri on MR. *AJNR Am J Neuroradiol*, 17(9), 1699-706.

Miller, M. I., Massie, A. B., Ratnanather, J. T., Botteron, K. N., & Csernansky, J. G. (2000). Bayesian construction of geometrically based

cortical thickness metrics. *Neuroimage*, 12(6), 676-87.

Mori, E., Yoneda, Y., Yamashita, H., Hirono, N., Ikeda, M., & Yamadori, A. (1997). Medial temporal structures relate to memory impairment in Alzheimer's disease: an MRI volumetric study. *J Neurol Neurosurg Psychiatry*, 63(2), 214-21.

Morris, R., Paxinos, G., & Petrides, M. (2000). Architectonic analysis of the human retrosplenial cortex. *J Comp Neurol*, 421(1), 14-28.

Mountcastle, V. B. (1998). *Perceptual neuroscience : the cerebral cortex*. Cambridge, Mass.: Harvard University Press.

Mungas, D., Reed, B. R., Jagust, W. J., DeCarli, C., Mack, W. J., Kramer, J. H., et al. (2002). Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. *Neurology*, 59(6), 867-73.

Nagano-Saito, A., Washimi, Y., Arahata, Y., Kachi, T., Lerch, J. P., Evans, A. C., et al. (2005). Cerebral atrophy and its relation to cognitive impairment in Parkinson disease. *Neurology*, 64(2), 224-29.

Nagy, Z., Hindley, N. J., Braak, H., Braak, E., Yilmazer-Hanke, D. M., Schultz, C., et al. (1999). Relationship between clinical and radiological diagnostic criteria for Alzheimer's disease and the extent of neuropathology as reflected by 'stages': a prospective study. *Dement Geriatr Cogn Disord*, 10(2), 109-14.

Narr, K. L., Bilder, R. M., Toga, A. W., Woods, R. P., Rex, D. E., Szeszko, P. R., et al. (2005). Mapping cortical thickness and gray matter concentration in first episode schizophrenia. *Cereb Cortex*, 15(6), 708-19.

Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat Methods Med Res*, 12(5), 419-46.

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp*, 15(1), 1-25.

NIST/SEMATECH. (2005). *e-Handbook of Statistical Methods*. <http://www.itl.nist.gov/div898/handbook/>.

O'Donnell, S., Noseworthy, M. D., Levine, B., & Dennis, M. (2005). Cortical thickness of the frontopolar area in typically developing children and adolescents. *Neuroimage*, 24(4), 948-54.

Obermayer, K., & Blasdel, G. G. (1993). Geometry of orientation and ocular dominance columns in monkey striate cortex. *J Neurosci*, 13(10),

4114-29.

Palmen, S. J., Hulshoff Pol, H. E., Kemner, C., Schnack, H. G., Durston, S., Lohuis, B. E., et al. (2005). Increased gray-matter volume in medication-naive high-functioning children with autism spectrum disorder. *Psychol Med*, 35(4), 561-70.

Parker, G. J., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A., Ciccarelli, O., & Lambon Ralph, M. A. (2005). Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage*, 24(3), 656-66.

Pascual-Leone, A., Nguyet, D., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., & Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J Neurophysiol*, 74(3), 1037-45.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci*, 2(6), 417-24.

Paus, T., Collins, D. L., Evans, A. C., Leonard, G., Pike, B., & Zijdenbos, A. (2001). Maturation of white matter in the human brain: a review of magnetic resonance studies. *Brain Res Bull*, 54(3), 255-66.

Paus, T., Zijdenbos, A., Worsley, K., Collins, D. L., Blumenthal, J., Giedd, J. N., et al. (1999). Structural maturation of neural pathways in children and adolescents: in vivo study. *Science*, 283(5409), 1908-11.

Pearlson, G. D. (1999). New insights on the neuroanatomy of schizophrenia. *Curr Psychiatry Rep*, 1(1), 41-45.

Penn, A. A. (2001). Early brain wiring: activity-dependent processes. *Schizophr Bull*, 27(3), 337-47.

Pennington, B. F., Filipek, P. A., Lefly, D., Chhabildas, N., Kennedy, D. N., Simon, J. H., et al. (2000). A twin MRI study of size variations in human brain. *J Cogn Neurosci*, 12(1), 223-32.

Petrides, M., & Pandya, D. N. (1988). Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J Comp Neurol*, 273(1), 52-66.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur J Neurosci*, 16(2), 291-310.

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., et al. (1997). The structural components of music perception. *A*

functional anatomical study. *Brain*, 120 (Pt 2), 229-43.

Posthuma, D., Baare, W. F., Hulshoff Pol, H. E., Kahn, R. S., Boomsma, D. I., & De Geus, E. J. (2003). Genetic correlations between brain volumes and the WAIS-III dimensions of verbal comprehension, working memory, perceptual organization, and processing speed. *Twin Res*, 6(2), 131-39.

Pratt, W. K. (1991). *Digital Image Processing*. New York: John Wiley & Sons, Inc.

Preul, C., Lohmann, G., Hund-Georgiadis, M., Guthke, T., & von Cramon, D. Y. (2005). Morphometry demonstrates loss of cortical thickness in cerebral microangiopathy. *J Neurol*, 252(4), 441-47.

Pruessner, J. C., Collins, D. L., Pruessner, M., & Evans, A. C. (2001). Age and gender predict volume decline in the anterior and posterior hippocampus in early adulthood. *J Neurosci*, 21(1), 194-200.

Pruessner, J. C., Kohler, S., Crane, J., Pruessner, M., Lord, C., Byrne, A., et al. (2002). Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. *Cereb Cortex*, 12(12), 1342-53.

Prvulovic, D., Hubl, D., Sack, A. T., Melillo, L., Maurer, K., Frolich, L., et al. (2002). Functional imaging of visuospatial processing in Alzheimer's disease. *Neuroimage*, 17(3), 1403-14.

Rakic, P. (1974). Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. *Science*, 183(123), 425-27.

Rakic, P. (1976). Prenatal genesis of connections subserving ocular dominance in the rhesus monkey. *Nature*, 261(5560), 467-71.

Rakic, P. (1995). A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci*, 18(9), 383-88.

Ramnani, N., Behrens, T. E., Penny, W., & Matthews, P. M. (2004). New approaches for exploring anatomical and functional connectivity in the human brain. *Biol Psychiatry*, 56(9), 613-19.

Ratnanather, J. T., Botteron, K. N., Nishino, T., Massie, A. B., Lal, R. M., Patel, S. G., et al. (2001). Validating cortical surface analysis of medial prefrontal cortex. *Neuroimage*, 14(5), 1058-69.

Rauch, S. L., Wright, C. I., Martis, B., Busa, E., McMullin, K. G., Shin, L. M., et al. (2004). A magnetic resonance imaging study of cortical

thickness in animal phobia. *Biol Psychiatry*, 55(9), 946-52.

Redcay, E., & Courchesne, E. (2005). When is the Brain Enlarged in Autism? A Meta-Analysis of All Brain Size Reports. *Biol Psychiatry*,

Reiss, A. L., Abrams, M. T., Singer, H. S., Ross, J. L., & Denckla, M. B. (1996). Brain development, gender and IQ in children. A volumetric imaging study. *Brain*, 119 (Pt 5), 1763-74.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J Neurosci*, 23(8), 3295-301.

Rettmann, M. E., Han, X., Xu, C., & Prince, J. L. (2002). Automated sulcal segmentation using watersheds on the cortical surface. *Neuroimage*, 15(2), 329-44.

Rivkin, M. J. (2000). Developmental neuroimaging of children using magnetic resonance techniques. *Ment Retard Dev Disabil Res Rev*, 6(1), 68-80.

Rizzo, M., Anderson, S. W., Dawson, J., & Nawrot, M. (2000). Vision and cognition in Alzheimer's disease. *Neuropsychologia*, 38(8), 1157-69.

Robbins, S., Evans, A. C., Collins, D. L., & Whitesides, S. (2004). Tuning and comparing spatial normalization methods. *Med Image Anal*, 8(3), 311-23.

Rockel, A. J., Hiorns, R. W., & Powell, T. P. (1980). The basic uniformity in structure of the neocortex. *Brain*, 103(2), 221-44.

Romanski, L. M., Bates, J. F., & Goldman-Rakic, P. S. (1999). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J Comp Neurol*, 403(2), 141-57.

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., et al. (2002). Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology*, 58(5), 695-701.

Ross, M. E., Swanson, K., & Dobyns, W. B. (2001). Lissencephaly with cerebellar hypoplasia (LCH): a heterogeneous group of cortical malformations. *Neuropediatrics*, 32(5), 256-63.

Rottkamp, C. A., Atwood, C. S., Joseph, J. A., Nunomura, A., Perry, G., & Smith, M. A. (2002). The state versus amyloid-beta: the trial of the most wanted criminal in Alzheimer disease. *Peptides*, 23(7), 1333-41.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S. C., Simmons, A., et al. (2000). Functional frontalisation with age: mapping neurodevelopmental trajectories with fMRI. *Neurosci Biobehav Rev*, 24(1),

13-19.

Sailer, M., Fischl, B., Salat, D., Tempelmann, C., Schonfeld, M. A., Busa, E., et al. (2003). Focal thinning of the cerebral cortex in multiple sclerosis. *Brain*, 126(Pt 8), 1734-44.

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., et al. (2004). Thinning of the cerebral cortex in aging. *Cereb Cortex*, 14(7), 721-30.

Santangelo, S. L., & Tsatsanis, K. (2005). What is known about autism: genes, brain, and behavior. *Am J Pharmacogenomics*, 5(2), 71-92.

Sbarbati, A., Pizzini, F., Fabene, P. F., Nicolato, E., Marzola, P., Calderan, L., et al. (2004). Cerebral cortex three-dimensional profiling in human fetuses by magnetic resonance imaging. *J Anat*, 204(6), 465-74.

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. *Arch Neurol*, 60(7), 989-94.

Scahill, R. I., Schott, J. M., Stevens, J. M., Rossor, M. N., & Fox, N. C. (2002). Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. *Proc Natl Acad Sci U S A*, 99(7), 4703-07.

Scheltens, P., Fox, N., Barkhof, F., & Carli, C. D. (2002). Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion. *The Lancet Neurology*, 1, 13-21.

Schlaug, G. (2001). The brain of musicians. A model for functional and structural adaptation. *Ann N Y Acad Sci*, 930, 281-99.

Schmithorst, V. J., Wilke, M., Dardzinski, B. J., & Holland, S. K. (2002). Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: a cross-sectional diffusion-tensor MR imaging study. *Radiology*, 222(1), 212-18.

Selemon, L. D., & Goldman-Rakic, P. S. (1999). The reduced neuropil hypothesis: a circuit based model of schizophrenia. *Biol Psychiatry*, 45(1), 17-25.

Sexton, P. M., Paxinos, G., Kenney, M. A., Wookey, P. J., & Beaumont, K. (1994). In vitro autoradiographic localization of amylin binding sites in rat brain. *Neuroscience*, 62(2), 553-67.

Shen, D., & Davatzikos, C. (2004). Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping.

Neuroimage, 21(4), 1508-17.

Skoglund, T. S., Pascher, R., & Berthold, C. H. (1996). Heterogeneity in the columnar number of neurons in different neocortical areas in the rat. *Neurosci Lett*, 208(2), 97-100.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*, 17(1), 87-97.

Small, G. W., Agdeppa, E. D., Kepe, V., Satyamurthy, N., Huang, S. C., & Barrio, J. R. (2002). In vivo brain imaging of tangle burden in humans. *J Mol Neurosci*, 19(3), 323-27.

Sowell, E. R., Delis, D., Stiles, J., & Jernigan, T. L. (2001). Improved memory functioning and frontal lobe maturation between childhood and adolescence: a structural MRI study. *J Int Neuropsychol Soc*, 7(3), 312-22.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *J Neurosci*, 24(38), 8223-31.

Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System : An Approach to Cerebral Imaging*. Thieme Medical Publishers.

Thompson, P. M., Mega, M. S., Woods, R. P., Zoumalan, C. I., Lindshield, C. J., Blanton, R. E., et al. (2001). Cortical change in Alzheimer's disease detected with a disease-specific population-based brain atlas. *Cereb Cortex*, 11(1), 1-16.

Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., & Toga, A. W. (2000). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*, 404(6774), 190-93.

Thompson, P. M., Lee, A. D., Dutton, R. A., Geaga, J. A., Hayashi, K. M., Eckert, M. A., et al. (2005). Abnormal cortical complexity and thickness profiles mapped in Williams syndrome. *J Neurosci*, 25(16), 4146-58.

Thulborn, K. R., Martin, C., & Voyvodic, J. T. (2000). Functional MR imaging using a visually guided saccade paradigm for comparing activation patterns in patients with probable Alzheimer's disease and in cognitively able elderly volunteers. *AJNR Am J Neuroradiol*, 21(3), 524-31.

Toga, A. W., & Thompson, P. M. (2004). *Genetics of Brain Structure and Intelligence*. *Annu Rev Neurosci*,

Tohka, J., Zijdenbos, A., & Evans, A. (2004). Fast and robust

parameter estimation for statistical partial volume models in brain MRI. *Neuroimage*, 23(1), 84-97.

Tosun, D., Rettmann, M. E., Han, X., Tao, X., Xu, C., Resnick, S. M., et al. (2004). Cortical surface segmentation and mapping. *Neuroimage*, 23 Suppl 1, S108-18.

Valla, J., Berndt, J. D., & Gonzalez-Lima, F. (2001). Energy hypometabolism in posterior cingulate cortex of Alzheimer's patients: superficial laminar cytochrome oxidase associated with disease duration. *J Neurosci*, 21(13), 4923-30.

van der Valk, P. H., Snoeck, I., Meiners, L. C., des Portes, V., Chelly, J., Pinard, J. M., et al. (1999). Subcortical laminar heterotopia in two sisters and their mother: MRI, clinical findings and pathogenesis. *Neuropediatrics*, 30(3), 155-60.

Van Essen, D. C. (2004). Surface-based approaches to spatial localization and registration in primate cerebral cortex. *Neuroimage*, 23 Suppl 1, S97-107.

Van Essen, D. C., Drury, H. A., Joshi, S., & Miller, M. I. (1998). Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc Natl Acad Sci U S A*, 95(3), 788-95.

Van Hoesen, G. W. (1995). Anatomy of the medial temporal lobe. *Magn Reson Imaging*, 13(8), 1047-55.

Van Hoesen, G. W., Parvizi, J., & Chu, C. C. (2000). Orbitofrontal cortex pathology in Alzheimer's disease. *Cereb Cortex*, 10(3), 243-51.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. New York: Springer.

Watkins, K. E., Paus, T., Lerch, J. P., Zijdenbos, A., Collins, D. L., Neelin, P., et al. (2001). Structural asymmetries in the human brain: a voxel-based statistical analysis of 142 MRI scans. *Cereb Cortex*, 11(9), 868-77.

Wechsler, D. (1991). *WISC-III : Wechsler intelligence scale for children : manual*. San Antonio: Psychological Corp., Harcourt Brace Jovanovich.

White, T., Andreasen, N. C., Nopoulos, P., & Magnotta, V. (2003). Gyrification abnormalities in childhood- and adolescent-onset schizophrenia. *Biol Psychiatry*, 54(4), 418-26.

Wiegand, L. C., Warfield, S. K., Levitt, J. J., Hirayasu, Y., Salisbury, D. F., Heckers, S., et al. (2004). Prefrontal cortical thickness in first-episode psychosis: a magnetic resonance imaging study. *Biol Psychiatry*, 55(2),

131-40.

Wilke, M., Sohn, J. H., Byars, A. W., & Holland, S. K. (2003). Bright spots: correlations of gray matter volume with IQ in a normal pediatric population. *Neuroimage*, 20(1), 202-15.

Willis, W. G., & Weiler, M. D. (2005). Neural substrates of childhood attention-deficit/hyperactivity disorder: electroencephalographic and magnetic resonance imaging evidence. *Dev Neuropsychol*, 27(1), 135-82.

Woo, C. L., Chuang, S. H., Becker, L. E., Jay, V., Otsubo, H., Rutka, J. T., et al. (2001). Radiologic-pathologic correlation in focal cortical dysplasia and hemimegalencephaly in 18 children. *Pediatr Neurol*, 25(4), 295-303.

Worsley, K., Chen, J.-I., Lerch, J., & Evans, A. C. (2005). Comparing connectivity via thresholding correlations and SVD. *Philosophical Transactions of the Royal Society*, in Press,

Worsley, K. J., Andermann, M., Koulis, T., MacDonald, D., & Evans, A. C. (1999). Detecting changes in nonisotropic images. *Hum Brain Mapp*, 8(2-3), 98-101.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab*, 12(6), 900-18.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signal in images of cerebral activation. *Human Brain Mapping*, 4, 58-73.

Worsley, K. J., Taylor, J. E., Tomaiuolo, F., & Lerch, J. (2004). Unified univariate and multivariate random field theory. *Neuroimage*, 23 Suppl 1, S189-95.

Wright, I. C., McGuire, P. K., Poline, J. B., Traverso, J. M., Murray, R. M., Frith, C. D., et al. (1995). A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. *Neuroimage*, 2(4), 244-52.

Xu, Y., Jack, C. R., Jr., O'Brien, P. C., Kokmen, E., Smith, G. E., Ivnik, R. J., et al. (2000). Usefulness of MRI measures of entorhinal cortex versus hippocampus in AD. *Neurology*, 54(9), 1760-67.

Yezzi, A. J., Jr., & Prince, J. L. (2003). An Eulerian PDE approach for computing tissue thickness. *IEEE Trans Med Imaging*, 22(10), 1332-39.

Zakzanis, K. K., Graham, S. J., & Campbell, Z. (2003). A meta-

analysis of structural and functional brain imaging in dementia of the Alzheimer's type: a neuroimaging profile. *Neuropsychol Rev*, 13(1), 1-18.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058), 846-49.

Zeng, X., Staib, L. H., Schultz, R. T., & Duncan, J. S. (1999). Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation. *IEEE Trans Med Imaging*, 18(10), 927-37.

Zhang, W., Simos, P. G., Ishibashi, H., Wheless, J. W., Castillo, E. M., Kim, H. L., et al. (2003). Multimodality neuroimaging evaluation improves the detection of subtle cortical dysplasia in seizure patients. *Neurol Res*, 25(1), 53-57.

Zijdenbos, A. P., Forghani, R., & Evans, A. C. (2002). Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Trans Med Imaging*, 21(10), 1280-91.

Zilles, K. (1992). Neuronal plasticity as an adaptive property of the central nervous system. *Anat Anz*, 174(5), 383-91.

Zilles, K., Armstrong, E., Schleicher, A., & Kretschmann, H. J. (1988). The human pattern of gyrification in the cerebral cortex. *Anat Embryol (Berl)*, 179(2), 173-79.

Zilles, K., Palomero-Gallagher, N., & Schleicher, A. (2004). Transmitter receptors and functional anatomy of the cerebral cortex. *J Anat*, 205(6), 417-32.