Computer aided diagnosis in temporal lobe epilepsy and Alzheimer's dementia

Simon Duchesne, ing., M.Sc.

McGill University

Montreal, Quebec, Canada

October 23, 2005

A thesis submitted to McGill University

in partial fulfilment of the requirements of the degree of

Doctor in Philosophy (Biomedical Engineering)

Your file Votre référence
ISBN: 978-0-494-25135-5
Our file Notre référence
ISBN: 978-0-494-25135-5

Dedication

This thesis is dedicated first to those that made it possible:

Lynn, Chahaya, Jalani, and Kian;

my family;

and Louis.

Then, as it reaches out, my dedication will go to those in need of the help it contains.

Men build too many walls and not enough bridges

-Sir Isaac Newton
(1642 – 1727)

Acknowledgments

I would like to aknowledge:

All of my co-authors: Neda Bernasconi, Andrea Bernasconi, Christian Bocti, Howard Chertkow, Kathy De Sousa, Harald Hampel, Jens Pruessner, Andre Olivier, and Stefan Teipel;

The incredibly diverse and resourceful people at the Image Processing Laboratory and the Brain Imaging Center for all those years before, during and after my time, in particular Rosanne Aleong, Jennifer Chew, Andrew Janke, Jason Lerch, and Tomas Paus;

Support from the Canadian Institute for Health Research, the Fonds pour la Recherche en Santé du Québec (FRSQ), Manulife Financial / Canadian Council of Professional Engineers and the McGill Alma Mater Fund; and

Drs. Mazziotta and Evans from the International Consortium for Brain Mapping project for permission to use the data.

Table of contents

List of Tables

List of Figures

Abstract

Computer aided diagnosis within neuroimaging must rely on advanced image processing techniques to detect and quantify subtle signal changes that may be surrogate indicators of disease state. This thesis proposes two such novel methodologies that are both based on large volumes of interest, are data driven, and use cross-sectional scans: appearance-based classification (ABC) and voxel-based classification (VBC).

The concept of appearance in ABC represents the union of intensity and shape information extracted from magnetic resonance images (MRI). The classification method relies on a linear modeling of appearance features via principal components analysis, and comparison of the distribution of projection coordinates for the populations under study within a reference multidimensional appearance eigenspace. Classification is achieved using forward, stepwise linear discriminant analyses, in multiple cross-validated trials. In this work, the ABC methodology is shown to accurately lateralize the seizure focus in temporal lobe epilepsy (TLE), differentiate normal aging individuals from patients with either Alzheimer's dementia (AD) or Mild Cognitive Impairment (MCI), and finally predict the progression of MCI patients to AD. These applications demonstrated that the ABC technique is robust to different signal changes due to two distinct pathologies, to low

resolution data and motion artifacts, and to possible differences inherent to multi-site acquisition.

The VBC technique relies on voxel-based morphometry to identify regions of grey and white matter concentration differences between co-registered cohorts of individuals, and then on linear modeling of variables extracted from these regions. Classification is achieved using linear discriminant analyses within a multivariate space composed of voxel-based morphometry measures related to grey and white matter concentration, along with clinical variables of interest. VBC is shown to increase the accuracy of prediction of one-year clinical status from three to four out of five TLE patients having undergone selective amygdalo-hippocampectomy. These two techniques are shown to have the necessary potential to solve current problems in neurological research, assist clinical physicians with their decision-making process and influence positively patient management.

Abrégé

Les systèmes experts d'aide au diagnostic en neuroimagerie doivent composer avec des techniques avancées de traitement de l'image pour détecter et quantifier des differences subtiles dans le signal, qui peuvent être des indicateurs indirects de l'état de la maladie. Cette thèse propose deux nouvelles méthodes à cet effet, toutes deux basées sur de larges volumes d'intérêt contenant les données intrinsèques, et qui font usage d'un seul volume d'acquisition: la classification basée sur l'apparence (CBA), et la classification basée sur les voxels (CBV).

Le concept d'apparence en CBA représente l'union de l'information d'intensité et de forme, informations continues dans les images obtenues par résonance magnétique (IRM). La méthode de classification se base sur un modèle linéaire de vecteurs d'apparence via une analyse par composantes principales, et la comparaison des distributions de coordonnées de projection – pour les groupes sous étude – dans un espace référentiel multidimensionnel d'apparence. La classification elle-même est obtenue grâce à des analyses discriminantes linéaires graduées, dans de multiples essais de validation. Dans ce travail, il est démontré que la méthode CBA parvient à classifier de façon précise le foyer épileptogène en épilepsie du lobe temporal, de différencier le vieillissement normal de la maladie d'Alzheimer et du déficit cognitive léger, ainsi que de prédire la progression vers la démence d'Alzheimer

pour les patients atteints d'un déficit cognitif léger. Ces applications démontrent que la technique CBA est robuste face à des changements de signaux dus aux deux pathologies distinctes, à des données à basse résolution et porteuses de défauts reliés au mouvement, et potentiellement aux différences inhérentes dans des acquisitions multi-sites.

La technique CBV se fonde sur la morphométrie par voxel pour identifier les régions de différence en concentration de matière grise et matière blanche entre les images recalées de deux groupes d'individus. Suit la modélisation linéaire de variables mesurées dans ces régions, incluant des variables cliniques d'intérêt. La classification est effectuée à l'aide d'analyses discriminantes linéaires à l'intérieur d'un espace multivarié, formé des mesures reliées à la concentration de matière grise ou blanche et des variables cliniques. Il est démontré que CBV permet d'augmenter la précision de la prédiction du résultat post-opératoire pour des patients souffrant d'épilepsie du lobe temporal et ayant subit une résection du complexe amygdalo-hippocampique, et ce dans une proportion de trois à quatre patients sur cinq.

Ces deux techniques ont donc le potentiel nécessaire pour résoudre des problèmes actuels en recherche neurologique, d'assister les spécialistes dans leur processus décisionnel et d'influencer de façon positive la gestion des patients.

Original contributions

The central purpose of this work was to develop automated aid to diagnosis methodologies for neurological diseases, and to test and validate these techniques with real-life data sets. More specifically, the following original contributions are claimed.

Main contributions:

1) Created an appearance-based classification (ABC) methodology using MR images for the purpose of aid to diagnosis in neurological diseases;
2) Determined that ABC can accurately lateralize the seizure focus in intractable temporal lobe epilepsy (TLE);
3) Determined that ABC can accurately discriminate normal aging from Alzheimer's dementia (AD) and mild cognitive impairment (MCI);
4) Determined that ABC can accurately predict progression to AD on average 2.6 years before clinical diagnostic, in a given cohort of amnestic MCI patients;
5) Created a voxel-based morphometry classification (VBC) methodology using MR images for the purpose of aid to diagnosis in neurological diseases;

6) Determined that VBC can improve the accuracy of predicting one-year post-operative clinical status in patients having undergone selective amygdalo-hippocampectomy (SAH) for intractable TLE;
7) Proposed and verified experimentally that the cumulative microscopic effect of AD and TLE influence the T1w MR signal and can be estimated by appearance-based models;
8) Proposed and verified experimentally that large, nonspecific volumes of interest centered on the medial temporal lobes (MTL) contain additional discriminatory information when compared to hippocampal and/or amydgala volumetry, in the tested TLE and AD applications.

Secondary contributions:

9) Defined a class of appearance parameters composed of a combination of T1-weighted MR intensity and shape measures;
10) Extended 2D appearance-based segmentation approaches into a 3D classification methodology;
11) Evaluated that assessing shape changes in the right medial temporal lobe is essential for accurate lateralization of the seizure focus in TLE;
12) Determined that there are extra-hippocampal areas of grey and white matter concentration changes that are related to one-year SAH outcome;

13) Determined that shape characteristics within the MTL are critical in the differentiation of AD and MCI from normal aging;

14) Determined that ABC is sufficiently robust to reject site acquisition differences in a classification task;

15) Determining that T1-weighted intensity characteristics within the right MTL are critical in predicting MCI progression to AD;

16) Used a novel visualization tool to demonstrate the regions of intensity changes in the right MTL that are linked to progression of MCI to AD;

17) Created a pipeline for the rapid, reproducible and automated processing of MR images across multiple datasets.

Disclaimer

A U.S. patent application has been submitted on parts of this work for protection of intellectual property ("Systems and Methods of Classification Utilizing Intensity and Spatial Data", Co-inventors D.L. Collins, Ph.D., S. Duchesne, ing., M.Sc., U.S. Patent Pending, No 10/990396).

Author contributions

I am the first author of all four manuscripts included in this thesis and have performed all of the methodological developments, data processing, statistical analysis and interpretation of the data. Software development was undertaken where needed, particularly for the image processing pipeline and high-dimensional modeling. The contributions of co-authors included supervision, data collection, and advice on style and content of the manuscripts.

The following list, by manuscript, summarizes the contributions of co-authors:

Chapter 4 – MR-based neurological disease classification: application to temporal lobe epilepsy

- Authors: **S. Duchesne**, N. Bernasconi, A. Bernasconi, D.L. Collins
- Guarantors of integrity of entire study, all authors; study concepts and design, all authors; literature research, S.D.; clinical studies, N.B., A.B.; data acquisition, N.B., A.B.; methods, analysis and interpretation, S.D.; statistical analysis, S.D.; manuscript preparation, S.D.; manuscript revision/review, all authors; manuscript definition of intellectual content, editing, and final version approval, S.D., D.L.C.

Chapter 5 – Automated MR analysis helps predicting surgical outcome in selective amygdalo-hippocampectomy patients

- Authors: **S. Duchesne**, N. Bernasconi, A. Olivier, A. Bernasconi, D.L. Collins

- Guarantors of integrity of entire study, all authors; study concepts and design, all authors; literature research, S.D.; clinical studies, A.O., N.B., A.B.; data acquisition, N.B., A.B., A.O.; methods, analysis and interpretation, S.D.; statistical analysis, S.D.; manuscript preparation, S.D.; manuscript revision/review, all authors; manuscript definition of intellectual content, editing, and final version approval, S.D., D.L.C.

Chapter 6 – Differentiation of normal aging from Alzheimer's dementia and Mild Cognitive Impairment using a novel MR-based classification methodology

- Authors: **S. Duchesne**, J.C. Pruessner, H. Hampel, S. Teipel, D.L. Collins;
- Guarantors of integrity of entire study, all authors; study concepts and design, all authors; literature research, S.D.; clinical studies, H.H., S.T., J.C.P.; data acquisition, H.H., S.T., J.C.P.; methods, analysis and interpretation, S.D.; statistical analysis, S.D.; manuscript preparation, S.D.; manuscript revision/review, all authors; manuscript definition of intellectual content, editing, and final version approval, S.D., D.L.C.

Chapter 7 – Predicting MCI progression to AD at baseline MRI

- Authors: **S. Duchesne**, K. De Sousa, C. Bocti, H. Chertkow, D. L. Collins;
- Guarantors of integrity of entire study, all authors; study concepts and design, all authors; literature research, S.D.; clinical studies, K.D.S., C.B., H.C.; data acquisition, K. D.S., C.B., H.C.; methods, analysis and interpretation, S.D.;

statistical analysis, S.D.; manuscript preparation, S.D.; manuscript revision/review, all authors; manuscript definition of intellectual content, editing, and final version approval, S.D., D.L.C.

Chapter 1

Introduction

Neuroimaging techniques, in particular structural or anatomical magnetic resonance imaging (MRI), allow the visualization and quantification of pathologically induced brain changes in the living subject; it then becomes possible to employ metrics capturing these changes as surrogate measures of the state of the disease. These techniques can help to achieve earlier diagnosis, characterize the time-course of the illness and generally increase therapy efficacy.

Many neurological diseases exhibit pathologically specific discriminatory information in the form of local intensity variations and shape changes when observed on MRI. Such diseases include for example schizophrenia, Alzheimer's dementia (AD) or epilepsy with a seizure focus that can be lateralized to one of the medial temporal lobes (MTL).

The purpose of this thesis is to develop and validate, using real data, novel automated image processing techniques that would exploit these intensity and shape changes, in order to ultimately answer important clinical problems in neurology. For this thesis, the image input data will be limited to structural MRI, whose availability is constantly increasing in major Canadian centers. As such, the techniques presented will allow for automated, reproducible, and robust analysis of images as an aid to diagnosis and eventually assist clinicians in their decision process regarding patient management and care.

Thesis overview

This thesis is organized into eight major sections. Chapter 2, containing the background about basic and advanced MRI techniques necessary to the understanding of the novel methodologies, follows this introduction. Chapter 3 includes a detailed overview of the two methodologies that have been developed and validated in this thesis: (a) appearance-based classification (ABC), based on multi-dimensional analysis of intensity and shape measures from large, non-specific Volumes of Interest; and (b) voxel-based classification (VBC), based on individual-to-group comparison of regional variations in grey and white matter concentrations, targeted using between-group voxel-based morphometry.

Following Chapter 3, four separate manuscripts, each preceded by a preface, form additional chapters. Each manuscript consists in the validation of one of the

aforementioned methodologies with real data, the idea being that, for each of two diseases, one article attempts to aid the immediate diagnostic, while the second endeavors the prediction of future clinical status.

In the first article, published in *NeuroImage*, a study is performed to lateralize the seizure focus in temporal lobe epilepsy (TLE) patients. The second article, to be submitted to *NeuroSurgery*, recounts our attempt at predicting surgical outcome (clinical status at one-year follow-up) for a sub-group of the same TLE cohort. The third article, to be submitted to *Alzheimer's and Dementia*, is concerned with improving the diagnosis of normal aging individuals *vs.* patients with either Alzheimer's dementia (AD) or mild cognitive impairment (MCI). Finally, the last article, to be submitted to *Lancet Neurology*, details our work in predicting the future clinical status and progression of MCI patients to AD.

The thesis concludes with a last chapter containing a general discussion and suggestions for future research. References for all chapters, as well as recurring abbreviations, have been collected in a single list at the end of the thesis for brevity.

Chapter 2

Background

In Canada, for the year 2001, approximately 647 000 individuals over 15 years of age received a magnetic resonance scan, 12% of those for neurological or brain-related issues (CIHI 2003). As of 2003, there were 147 magnetic resonance (MR) scanners in the country, most of them acquired and installed within the past 5 years (CIHI 2003). With no side effects, high-resolution and high contrast between soft tissues, MRI is a modality that has dramatically improved the ability to perform *in vivo* analysis and diagnosis, particularly in neurology and neurosurgery.

The burden of interpretation of MRI rests on the many radiologists working in various hospitals and clinical centers. Most medical problems find their solution through the experienced judgment and careful observation that these professionals bring to the task; in some instances however, advanced image processing techniques are necessary to help them reach a confident decision. Qualitative analysis often

leads to many more quantitative questions (*Which of these two hippocampi is smallest? What is the size of this multiple sclerosis plaque?*), while the biggest problem remains to answer accurately the most fundamental clinical question: does the patient have, or not, the disease? In these instances, computerized systems offer an automated, unbiased and objective assessment of the images that hopefully can provide factual evidence supporting the diagnosis.

In order to understand the advanced image processing techniques presented in this work, one needs to summarize the context for their evolution, the raw data that will be used as input, and some of the previous techniques and procedures on which they are based. The purpose of this chapter is to fulfill this need. It is composed of four sections, beginning with an overview of nuclear magnetic resonance (NMR), continuing with a more in-depth description of basic MR signals, and moves on to basic and more advanced processing techniques that are used to maximize the information content of the MR images. This summary of the path from initial data to processed output is done with the goal of highlighting the usefulness of each step in the context of clinical aid to diagnosis. The chapter closes with two case studies on TLE and AD, which are the two main applicative areas of this thesis.

Nuclear Magnetic Resonance

The process of Nuclear Magnetic Resonances (NMR) centers on the interaction of an atomic nucleus with a magnetic field. Spin, a fundamental property of the nuclei

that describes its angular momentum, must be half-integer in order for the nucleus to experience resonance. This implies therefore that nuclei such as carbon-13, nitrogen-15, fluorine-19, sodium-23, phosphorus-31 and, of course, the hydrogen nucleus, can be used for NMR. The latter is the most relevant nuclei for in-vivo medical imaging, because it is present in large concentrations in water and fat, which are major constituents of human tissue.

In the presence of an external magnetic field B_0 in the z-direction, the magnetic moment of the spinning nucleus combines with B_0, and the nucleus experiences a torque. This will flip the nucleus in one of two states: a low energy state, aligned with B_0, and a high energy state, away from B_0. Bulk magnetization refers to the total alignment of multiple nuclei in a sample; the low energy state will slightly dominate the high energy one, and the sample becomes magnetized; it is expressed as a vector M, aligned with B_0.

If one were to apply a transverse magnetic field B_1 to the magnetized sample, with energy corresponding to the difference between low and high states, then some nuclei in the sample would flip and bulk magnetization would be reduced. The vector Mwould be seen to tilt away from the direction of B_0, and spiral towards the "transverse" or perpendicular plane. As soon as M has a component in the transverse plane (M_{xy}), it will generate a detectable signal (Tofts 2003).

When B_1 is removed, the absorbed energy in the system will dissipate in order for the system to return, or "relax", to the previous state of equilibrium. This dissipation

occurs by a number of processes, known collectively as relaxation mechanisms (Tofts 2003), whose exponential decay are governed by various time constants *T*. Longitudinal, or T1-relaxation, refers to energy transfer from the spins to the surrounding lattice, and therefore describes the return to equilibrium of the longitudinal component M_z of the bulk magnetization vector *M*. Transverse, or T2-relaxation, refers to spin-spin interactions, and explains the changes in the perpendicular component to the background field.

First level: MR data analysis

MR signals, acquired as a function of the relaxation mechanisms described above, are indicative of the biological composition of the tissue and of its physiological status; changes can be correlated with the pathology of interest. The first level of MR data analysis usually consists in the radiological interpretation of the images or the analysis of the measured signal with minimal processing (usually only performed to correct for acquisition-specific distortions).

T1 imaging

In normal tissues T1 values are related to macromolecular concentration, water binding and water content (Tofts 2003). This is the reason for the basic T1 contrast in the brain: the fat in myelin causes white matter (WM) to have a shorter T1 value

than grey matter (GM), and in turn GM has a shorter value than cerebro-spinal fluid (CSF) (see Figure 1)

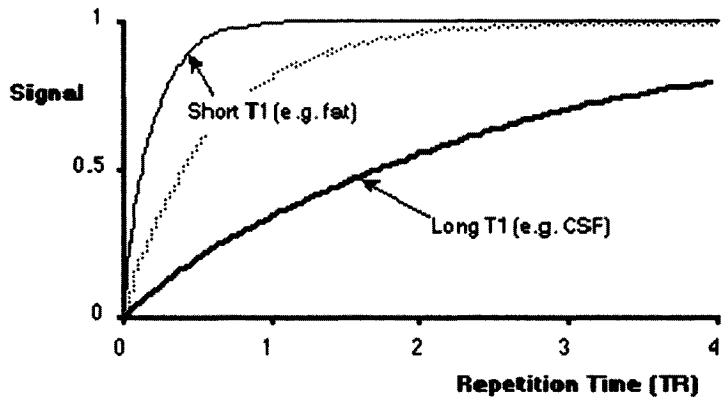

Figure 1 - Approach to magnetic equilibrium of an initially unmagnetized sample. In this graph, a value of 1.0 for Mo represents equilibrium magnetization, and the time is shown in seconds (Source: http://airto.bmap.ucla.edu/BMCweb/CourseWork/M285/MRI/MRBasics.html)

In practice T1-weighted images have become the standard for anatomical MRI, offering the best possible means of assessing GM and WM boundaries and therefore allowing the parcellation or segmentation of the brain into different morphological regions.

Pathologies that affect the water distribution will change the T1 signal. For instance, edema around tumors or in inflammatory acute multiple sclerosis (MS) lesions lead to an increase in T1 (Tofts 2003). T1 changes in TLE and AD will be addressed specifically later in this chapter.

T2 imaging

Transverse relaxation relates to the spin-spin interactions of the decaying transverse magnetization and is found to be a function of the tissue and/or the local magnetic environment of the nuclei (Tofts 2003). As such this opens a window to the microscopic environment of the hydrogen nuclei under investigation, and allows for a differentiation in tissue types, in health or in disease. T2 relaxation, like T1, is strongly correlated to the bulk water content in the cells however, contrary to T1, fatty tissue will have a longer T2 signal. The contrast in T2-weighted images of the brain is therefore reversed from that of a T1-weighted image: CSF appears bright, and WM dark.

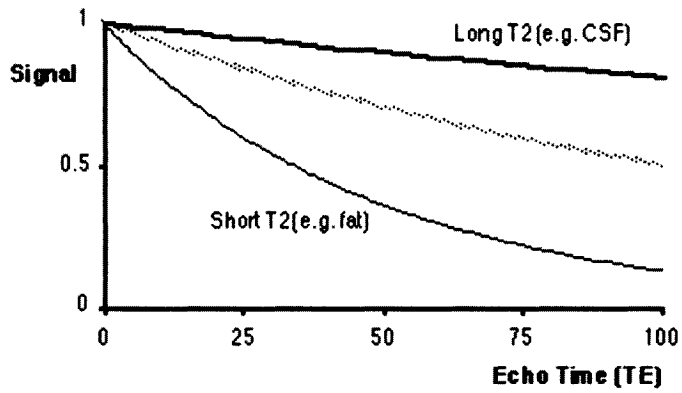

Figure 2 - Decay of transverse component of MR signal
This graph shows the magnitude of the MR signal as it decays over time from its initial maximum (Source: airto.bmap.ucla.edu/BMCweb/CourseWork/M285/MRI/MRBasics.html)

In the normal brain, heterogeneity in T2 values is attributed to variations in water content, cytoarchitecture, iron concentration, degree of myelination and increased vascularity (Tofts 2003). Abnormalities identified on T2-weighted images reflect

altered tissue water content, which can be associated with axonal loss, gliosis, demyelination and edema, which are seen in or around MS lesions (Larsson, Barker et al. 1998)(Larsson et al 1989). In TLE, increases in T2-weighted signal intensity have been reported in hippocampal sclerosis, related to the gliosis in the dentate gyrus (Briellmann, Kalnins et al. 2002). T2 is not a good indicator of AD (Laakso, Partanen et al. 1996).

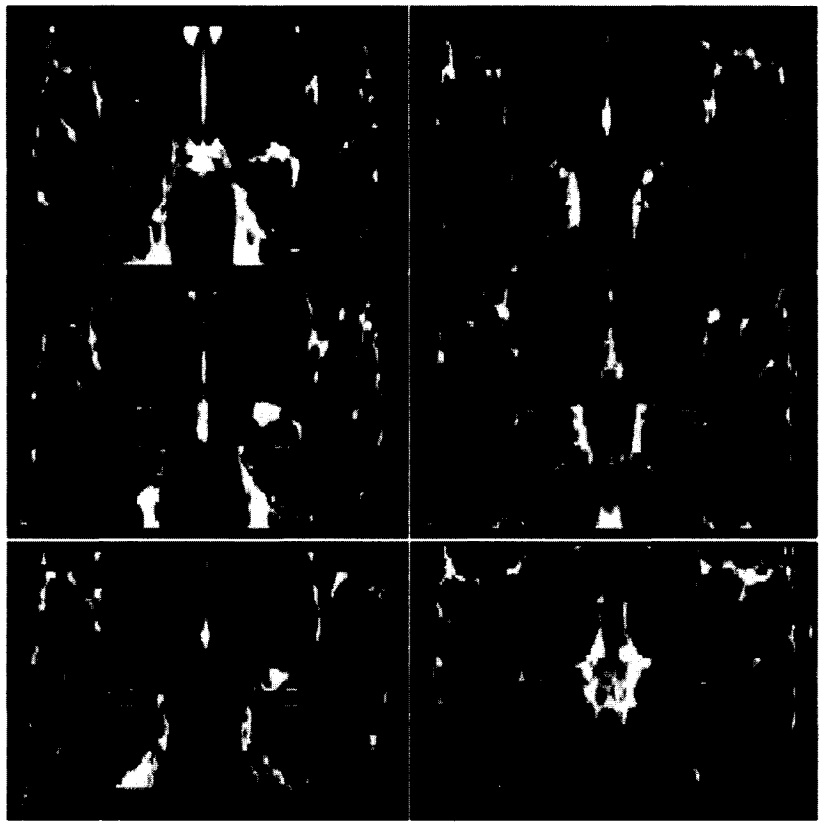

Figure 3 – T2 imaging in TLE

Coronal T2 maps calculated from 23 images of a control subject. Pixel intensity represents T2 relaxation time calculated by fitting a single exponential to corresponding pixels from each constituent image. The white squares within the hippocampal head (panels 1 and 2), body (panels 3–5), and tail (panel 6) represent the region of interest for the T2 relaxation time measurements. Significantly higher T2 relaxation time was found in TLE patients when compared to NC. Source: (Bernasconi, Bernasconi et al. 2000)

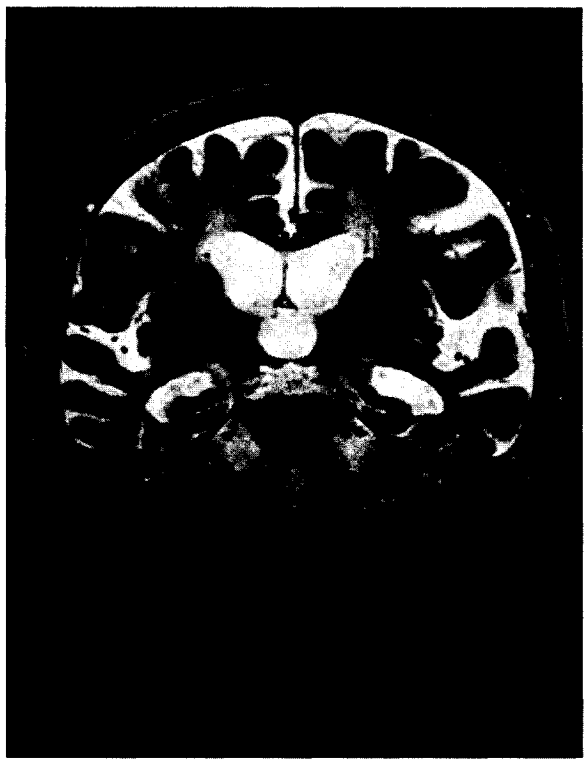

Figure 4 – T2-weighted imaging in AD

This coronal T2-weighted image shows extensive white matter hyperintensity and severe medial temporal lobe atrophy, which is suggestive of combination between degenerative and vascular brain pathology.
Source: (Bastos Leite, Scheltens et al. 2004)

Proton density imaging

Proton density (PD) refers to the overall concentration of MRI-visible protons (hydrogen nuclei) in tissue. As discussed previously, most of the hydrogen nuclei are located in water, and virtually all tissue water is visible. There is also a large pool of non-aqueous protons (30% of protons in white matter) that is MRI-invisible (Tofts 2003).

Increases in PD correspond largely to edema and are seen in longitudinal examination of MS plaques (Narayanan, Fu et al. 1997). Changes in PD often correlate closely with T1 changes.

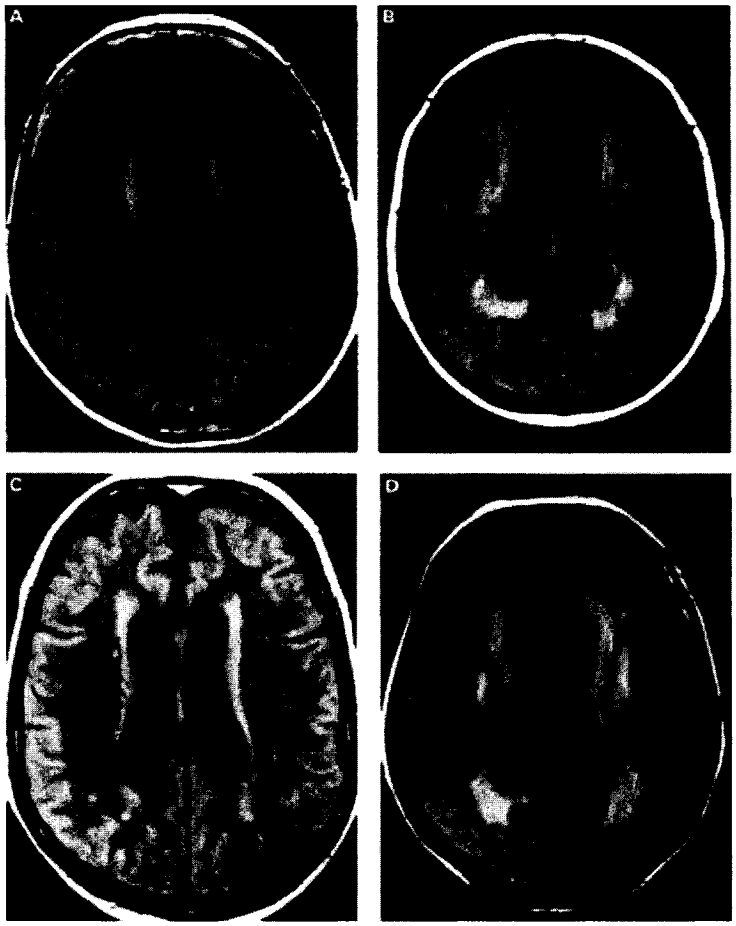

Figure 5 – PD imaging

Axial proton density MRI showing white matter lesions: (A) normal control; (B) patient with vascular dementia (extensive frontal, occipital and parietal hyperintensities); (C) patient with pathologically confirmed dementia with Lewy bodies (primarily frontal and parietal abnormalities); (D) patient with pathologically confirmed AD (frontal, parietal and occipital changes) Source: (Barber, Scheltens et al. 1999)

Diffusion Imaging

Free water will diffuse isotropically, that is equally in all directions, due to Brownian motion. Barriers, such as cell membranes in biological tissue, will hinder this diffusion and make the process anisotropic. The higher the anisotropy, the more restricted the water pool; anisotropy can serve therefore indirectly to measure cellular membrane integrity. Diffusion imaging measures the microscopic translational motion of hydrogen nuclei in water molecules and from this we can infer the degree of anisotropy via the apparent diffusion coefficient (ADC). In healthy tissue, myelin will have high anisotropy, particularly in large fiber bundles or tracts (such as the corpus callosum), whereas in CSF, diffusion will be nearly isotropic.

Decreases in anisotropy in disease are indicative of the destruction of tissue, such as in demyelination in MS (Cassol, Ranjeva et al. 2004). Further, by measuring the diffusion behavior in a range of directions (the diffusion tensor) we obtain a measure that is independent of the orientation of the subject in the scanner (Tofts 2003). By following the direction of maximum diffusion, we can identify the connectivity of major nerve fiber bundles, a process known as tractography. Decreased anisotropy in the hippocampi in both AD and MCI when compared to normal aging have been reported (Kantarci, Jack et al. 2001), with other differences elsewhere in the brain between AD and normal aging (see Figure 6). The reliability

for individual prognosis is poor however, as there is considerable overlap between the anisotropy values of the three groups.

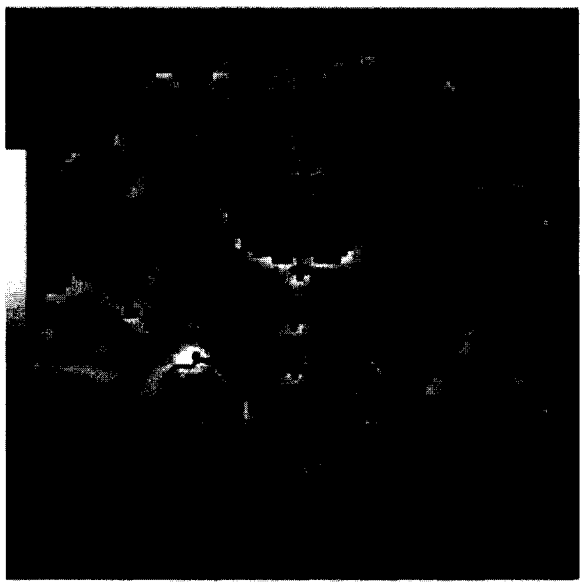

Figure 6 – Diffusion imaging in AD

Apparent diffusion coefficient (average) map, with regions of interest drawn on the anterior cingulated WM, thalamic, temporal stem and hippocampi. AD patients exhibit decreased anisotropy in these areas when compared to normal aging. Source (Kantarci, Jack et al. 2001)

Magnetization Transfer Imaging

Magnetization transfer imaging (MTI) enables the properties of protons bound to macromolecules to be probed (Tofts 2003). These bound protons (the 30% MRI-invisible protons referred to in the Proton Density Imaging section above) compose a different pool than the free water protons observed so far via T1, T2, PD or DTI; they are largely located in fat and proteins, and therefore in myelinated white matter. The magnetization transfer ratio (MTR) between free and bound protons gives an

indication of the quantity of bound protons present in tissue. CSF being essentially free water, its MTR will be close to zero, whereas MTR in WM will be high. MTR is reduced in many pathologies, in part due to demyelination, such as the corpus callosum of AD patients (Hanyu, Asano et al. 1999). There are also noticeable changes in MTR in MCI patients (Kabani, Sled et al. 2002). MTR measurements were initially thought to be useful for lateralization of seizure focus in TLE (Tofts, Sisodiya et al. 1995), but more extensive research refuted that claim (Li, Narayanan et al. 2000).

MR Spectroscopy

Hydrogen is, by far, the most commonly observed nuclei in MRI however, as mentioned previously, any half-integer spin nuclei can theoretically be observed. Bulk magnetization actually results in a spectrum with peaks corresponding to various other chemical compounds or metabolites that resonate at different frequencies. Magnetic Resonance Spectroscopy (MRS) is concerned with the understanding and tracking of resonant peaks from other molecules; the major difficulty of course consisting in the identification of spectral peaks corresponding to a given molecule. Once identified, the resulting specificity is high however, and absolute concentration measurements become possible.

Various chemicals involved in brain morphological and physiological processes have been identified. These include N-acetyl-aspartate (NAA), thought to give an

indication of neuronal density, choline-containing compounds, creatine, myo-inositol, glutamate and glutamine, lactate and lipids (Tofts 2003).

Of all the compounds studied by Kantarci *et al.*, (Kantarci, Smith et al. 2002), the NAA/Myo-inositol ratio may be the most efficient predictor of memory and cognitive function in patients with MCI and AD. MRS has also been shown to be sensitive in detecting hippocampal sclerosis in TLE (Li, Caramanos et al. 2000)(see Figure 7).

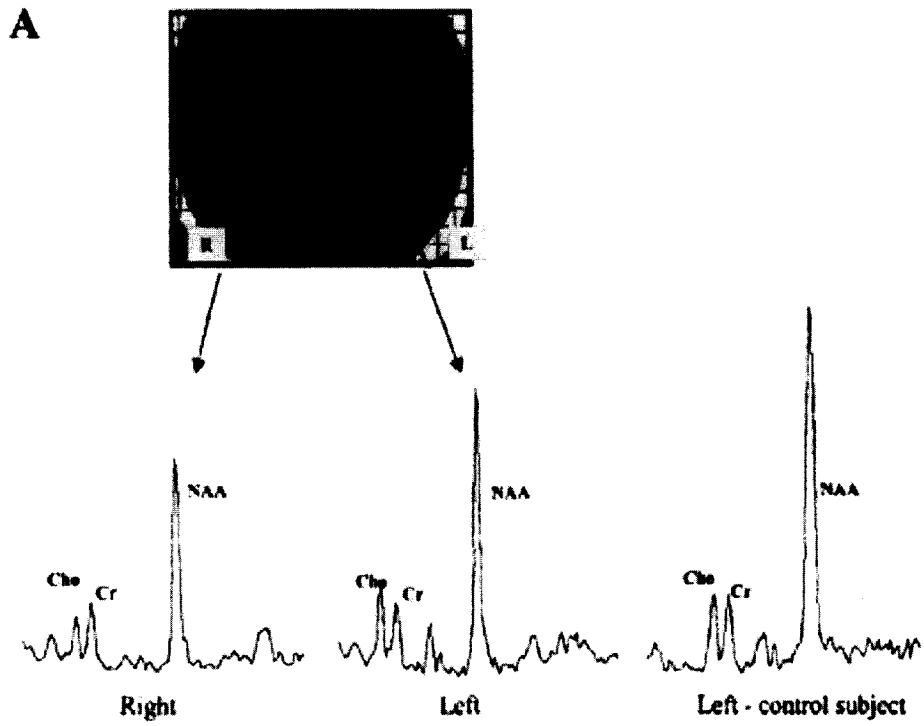

Figure 7 – MRS imaging in TLE

Follow-up temporal lobe 'H-MRS imaging of a 26-year-old patient with newly diagnosed TLE, 7 months after complete seizure control by carbamazepine, still shows low NAA/Cr values in both temporal regions
Source: (Bernasconi, Antel et al. 2001)

Functional MR imaging

Functional MRI (fMRI) shows the location and sometimes magnitude of increased neuronal activation arising from specific tasks (e.g. visual, motor or cognitive). Whereas all previous MR signals are directly related to the concentration of nuclei in the tissue, albeit in a macroscopic fashion, fMRI is an indirect measure of physiological response to an activation paradigm. As its name implies, it is a functional modality and should therefore be compared to other such imaging techniques, such as nuclear radiation (e.g. gamma cameras, positron emission tomography (PET) and single photon emission computed tomography (SPECT)) or electro-magnetic imaging (e.g. electro-encephalography (EEG) and magneto-encephalography (MEG)). All functional modalities, when compared to structural imaging such as x-ray based computed tomography (CT) or high resolution MRI, offer limited spatial resolution. Yet, their informational content is of paramount importance, as it enables a window into the time-dependent physiological processes in health and in disease.

There are several ways of detecting the increased metabolic demand associated with activity using MRI, the most common being the blood oxygen dependent (BOLD) signal (Logothetis and Pfeuffer 2004). When nerve cells are active they consume oxygen, supplied by local capillaries and carried by hemoglobin. A hemodynamic response occurs a few seconds after a burst of neuronal activity. The difference in the level of blood oxygenation, related to the fluctuations in diamagnetic oxygenated

hemoglobin and paramagnetic deoxygenated hemoglobin, results in a contrast (the BOLD signal) that can be used for imaging purposes.

Aside from acquisition parameters, the relevance of the neuropsychological activation paradigm must also be considered when evaluating results from fMRI. If the test has poor specificity/sensitivity, then the acquisition will not compensate for this poor design. It must also be closely coordinated with the acquisition sequence.

FMRI has been used extensively in many areas of the neurosciences; for example, localization of language regions for surgical planning (Poliakov, Hinshaw et al. 1999), and assessment of differences in cognitive abilities between normal aging and AD (Johnson, Saykin et al. 2000) and MCI (see Figure 8).

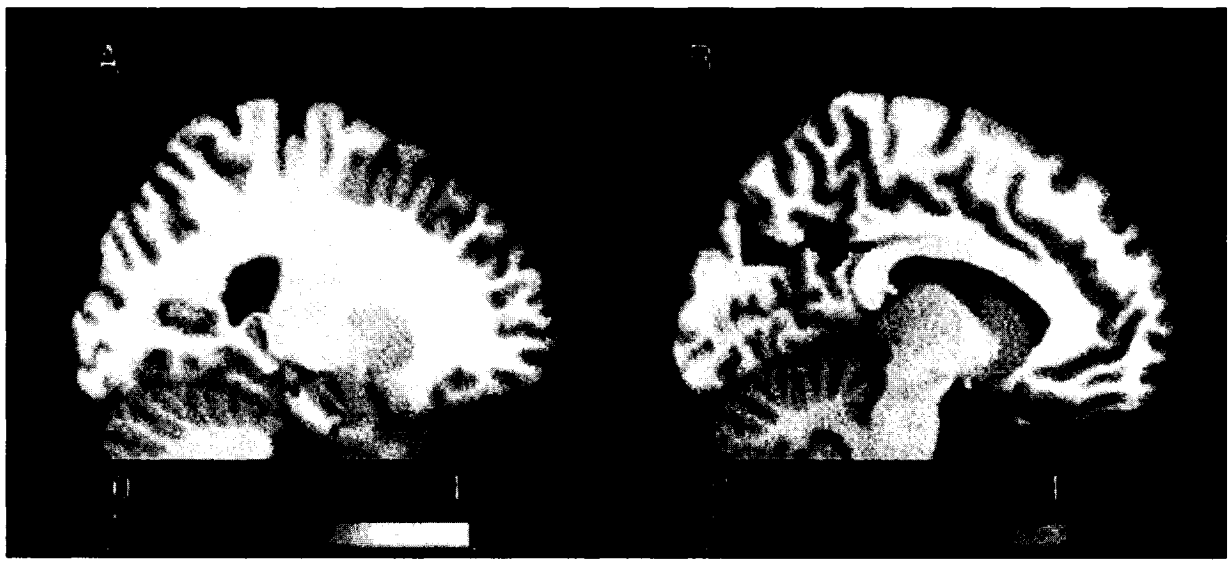

Figure 8 – fMRI in AD

Areas of significance where age-matched controls are significantly more active than patients with MCI on the two comparisons: (A) novel items: the hippocampus is significantly more active in controls than MCI and (B) previously learned items: the posterior cingulate and precuneus are more active in controls during recognition. Source: (Johnson, Schmitz et al. 2005)

Second level: MR data-derived features analysis

While the radiological analysis of MR data does give a number of answers to clinical questions, it is generally not sufficient when dealing with complex problems, such as assessing the rate of atrophy for structures in AD, or measuring the reduction in plaque size in MS through time during treatment. Further processing must be done to the original data, to put images in correspondence, and pinpoint subtle elements of interest on the images.

Texture

The concept of texture in medical imaging refers to the relationship between neighboring pixels (in 2D) or voxels (3D), under the hypothesis that some of the tissue characteristic giving rise to the signal within that image element will be related to its surrounding. Extracting local neighborhood intensity co-occurrences via *n*-th order statistics therefore serves to highlight changes in this interdependence, ideally related to the cumulative effect of microscopic pathological disturbances. Texture in MR images may depend, to varying degrees, on local water content, on the nature, concentration and distribution of protein and lipid molecules, on the chemical constituents and structure of molecules other than water, and substances which affect the local magnetic field, on perfusion, diffusion and flow and on the presence

of contrast agents (Tofts 2003). Different image resolutions will also restrain the comparative ability of texture metrics across subjects or groups. The most widely used texture metrics are derived from the spatial gray-level dependence matrix, often referred to as the co-occurrence matrix, which can be computed in 3D (Kovalev, Kruggel et al. 2001).

Neurodegenerative diseases of the brain, such as AD, are characterized by complex microscopic processes which result in cell death and physiological changes to brain tissue. It has been suggested that this may result in detectable changes in the 2D/3D macroscopic texture visible in MR (Freeborough and Fox 1998). First-order texture measures also show promise for assisting in locating focal cortical dysplasia in extra-temporal epilepsy (Bernasconi, Antel et al. 2001; Srivastava, Maes et al. 2005).

GM/WM/CSF classification

At current imaging resolutions ($\sim$1mm^3), within-brain voxels in MRI may broadly be thought as composed of either one of three classes: white matter (WM), grey matter (GM), or cerebro-spinal fluid (CSF). The GM is primarily composed of neuronal bodies, while the WM consists essentially of closely packed axonal tracts, heavily myelinated, forming cortico-cortical or cortico- sub-cortical connections. It is readily acknowledged that the actual *in vivo* cellular arrangement is not as clearly defined, as within any given voxel, a large number of axons and cell bodies co-exist; however, the overall signal being an average of each tissue contributions within that volume

element, the majority of voxels will have a signal heavily weighted towards either GM, WM, or CSF. For some boundary voxels (especially the GM/WM interfaces), the difficulty remains to separate or assign them a particular label, due to partial volume effects (i.e. proportions of different tissue classes within a single volume element).

At present however, this trichotomous view of brain tissue serves many purposes, and therefore much work has been dedicated to the separation of images into those major classes, a process referred to as tissue classification or segmentation.

Tissue classification techniques have typically used either multi-modality images, in an effort to parcellate the multi-spectral intensity space into three distinctive classes, and/or used *a priori* information about the known spatial distribution of GM/WM/CSF in control subjects. This task is made even more complicated by the fact that the spatial distribution of GM/WM/CSF is highly variable, no two brains being topologically alike; and further, GM and WM intensities do vary across a single brain, in part due to the acquisition or improper correction of its bias, and otherwise because of the heterogeneousness of tissue within volume elements throughout the brain. The reader is referred to (Pham, Xu et al. 2000) for a comprehensive review of segmentation algorithms.

In and of itself, the assignment of classes to tissue does not necessarily lead to a useful measurement that can aid the diagnosis; further analysis of the tissue maps

however does so, such as brain matter volumetry, ratio maps, and voxel-based morphometry (VBM), discussed in the next section.

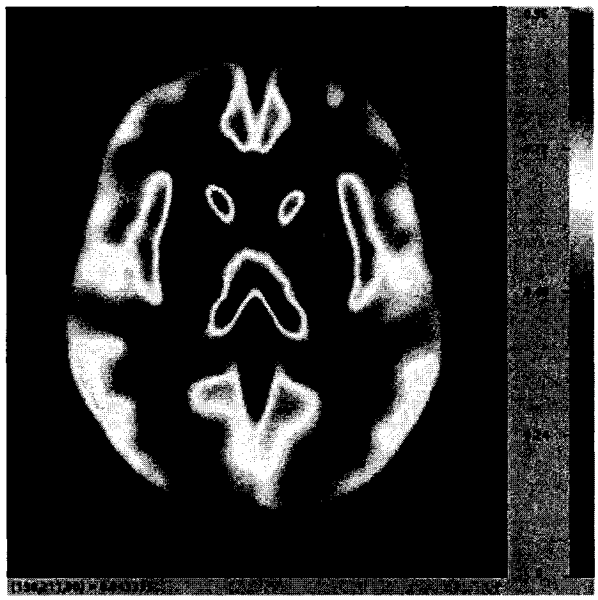

Figure 9 – GM average map for TLE subjects with left hippocampal atrophy.

This transverse image shows the average, smoothed distribution of GM in a group of 47 TLE patients with left hippocampal atrophy. In such an image voxels take a value between 0 (no GM present on average) and 1 (GM present in all cases).

Registration

Registration is the process by which two or more images are put in correspondence, i.e. aligned with respect to each other; when successful, it allows the comparison of spatial elements – and the signal within – on a point-by-point basis between the two or more images.

Scans from a single subject can be registered, either cross-sectionally across different modalities (intra-subject, inter-modality)(e.g. T1 and T2 images acquired in a single session) or longitudinally (intra-subject, intra-modality)(multiple images acquired at

different time points in the course of health or disease). Inter-subject registration allows image data from a group of subjects to be pooled. This is done for the purpose of group comparisons, or in order to compare scans from an individual to those of a group. Again, this can be performed cross-sectionally or longitudinally, within or across modalities. For a thorough review of image registration algorithms, the reader is referred to articles by Maintz *et al.* and Hill *et al.* (Maintz and Viergever 1998; Hill, Batchelor et al. 2001).

Whether or not registration is successful is not in itself indicative of a particular pathological state. Rather, the strength of co-registered images resides in the ability to perform intra-subject or inter-subject comparisons, in a mono- or multi-modality fashion. Further, there is much information embedded in the registration process itself, which can be exploited using deformation-based morphometry (see next section).

Registration is necessary step and a key enabling methodology in the study of health and disease. The co-registration possibilities are too numerous to discuss here: given N independently acquired modalities (T1w, T2w, PD, DTI, MTR, MRS, fMRI, SPECT, PET (one tracer each!), EEG, MEG, CT), one quickly reaches an astounding number of bi-modality combinations (in this case, 66). Of course, not all possibilities are necessarily relevant, as the informational value of certain modalities overlap.

Third level: MR model-based analysis

Even data-driven techniques don't necessary get to the essence of the most often asked clinical question: does the subject have, or not, the disease? This often requires some form of meta or higher-order analysis before it can be distilled into a binary output suitable for clinical consumption. The following model-based techniques stand at the top of the image processing food chain, and typically result in a dichotomous output that can aid the diagnosis.

Volumetry

Volumetry, essentially the measurement of the volume of a particular bounded structure of interest (Figure 10), is the primary indicator of structure integrity. Due to its prevalence in the medical literature and its acceptance as a gold standard in many reports, it will be studied here in greater detail.

Volumetry relies on accurate segmentation of the structures of interest, which is either performed manually or automatically; the two approaches will be discussed. While undeniably useful for understanding disease-driven changes in the chosen structure(s), volumetry suffer from significant drawbacks related to its reliance on segmentation; the most important ones being lack of knowledge about interrelations between neighboring tissues, and the incapacity at spatially locating the detected volume change.

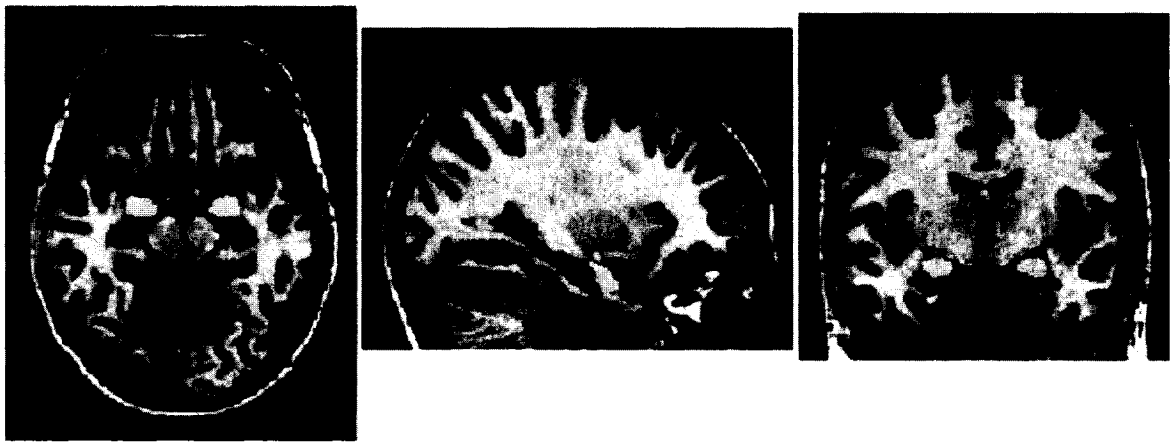

Figure 10 – Hippocampal and amygdala volumetry

This image shows manually determined labels for the amygdala (yellow), hippocampal head (red), body (green) and tail (blue). This separation allows the analysis of sub-components of the HC however, precise inter-individual boundaries are hard to define and assess with reliability.
Source: N. Bernasconi

Manual segmentation

Manual segmentation techniques involve contour delineation of structures by one or more trained neuroanatomists. Stereology, i.e. the technique of proper sampling based on the Cavalieri principle (Cavalieri 1635), had been the preferred approach to calculating volumes before the advent of high-resolution images, where volumes can be reliably estimated directly from the (often isotropic) voxel count and slice thickness.

While expert human intervention remains the most accurate segmentation technique and is considered the gold standard in many studies, serious drawbacks undermine its usefulness in a number of situations. The main difficulty resides in the subjective interpretation of anatomical variations. In the case of HC and AG for example,

differences in border definitions amongst research groups - and for that matter, amongst investigators of a same group – have hindered the comparison of results. The reported intra-class inter-rater overlap coefficients in a segmentation study of the HC are in the range K= 0.83-0.94 (Pruessner, Li et al.). Commonality in the definition of segmentation protocols is needed if results are to be compared. Secondly, research groups use different software packages to trace the targeted structure. Most employ 2D visualization tools for brain images, without the possibility of adjusting resolution or image contrast. A common 2D error is inter-slice misregistration, which leads to non-smooth 3D surfaces. On the other hand, scalable 3D imaging is available, allowing for precise display and enlargement of regions of interest in coronal, sagittal and transverse orientation (Pruessner, Li et al.). Finally, manual processing is time-consuming, as the specialist must delineate the structures on a slice-by-slice basis. Taking the HC as an example, its longest axis of 4-5 cm generates in the order of 40-50 slices with isotropic $1mm^3$ voxels. Hogan (Hogan, Mark et al. 2000) reports a total segmentation time of 2 hours per HC. This result has been corroborated by Pruessner et al. (Pruessner, Li et al.) in their manual segmentation study.

Automated segmentation

Computer-based segmentation can be roughly divided in two broad approaches, based on the reliance or not on *a priori* information.

- *Forward segmentation approaches* use a pre-defined template to match the new volume to achieve segmentation. In most cases experts are required to initialize the segmentation process by choosing landmarks. The matching process can be done in the way of an initial contour which will be propagated through some form of elastic matching onto the image until a proximity criteria is maximized. An example of this technique is the High-Dimensional Brain Mapping technique using fluid transformations, proposed by Christensen (Christensen, Joshi et al.). Other techniques, such as the Active Shape Model of Cootes and Taylor (Cootes, Taylor et al.), will rely on the placement of landmarks on the image to derive a model that is then globally deformed to match some intensity features of the new image.
- *Backwards segmentation*, on the other hand, can be thought as a reversion of the original paradigm. In those cases the new image is registered through various means unto a reference volume; atlases that have been previously defined on the reference volume can be propagated back unto the new subject, using the inverse transformation matrix. This paradigm can be observed in Thirion's *demons* (Thirion) and Collins' ANIMAL (Collins, Holmes et al.) algorithms.

While objective and repeatable, automated segmentation of structures is not yet perfectly accurate, as evidenced by the reduced overlap with expert-based manual segmentation (Duchesne, Pruessner et al. 2002).

Morphometry

Voxel-based morphometry

Voxel-based morphometry (VBM) consists in the statistical analysis of generalized linear model (GLM) results performed on a voxel-by-voxel basis on combined cohorts of co-registered subject imaging data. The procedure has been used extensively for the study of GM or WM concentration volumes across subjects, and originates from the study of longitudinal activation maps in fMRI. In fact, one of the most popular standard software package for VBM is that of Ashburner *et al.*, (Ashburner and Friston 2000), which consists in an adaptation of their statistical parametric mapping (SPM) technique, initially designed for functional activation (fMRI, PET) analysis. The method used in this thesis, called fMRIStat, was developed by Worsley et al. (Worsley, Marrett et al. 1996) for the same purpose. The standard approach involves co-registering images in a linear and/or non-linear fashion to improve the spatial co-location of brain structures; then tissue (GM, WM) are segmented into maps, spatially smoothed to form so-called "concentration" maps, with the understanding that they do not represent an absolute measurement of cell body counts or axonal densities, but rather an approximation of the likelihood of finding GM or WM in that area. Parameters of a general linear model are estimated at each voxel, across subjects, possibly with additional co-variates if applicable (e.g. age, sex). Voxels with statistically significant differences are identified based on a threshold corrected for multiple comparisons. Rather than

using Bonferroni correction, fMRIStat relies on Gaussian Random Field theory for this correction (Worsley, Marrett et al. 1996). The final result is a volume containing only those voxels (or peaks) in which there are statistically significant differences between the tissue concentrations for the groups under study (see Figure 11).

VBM's detection accuracy is limited by sources of spatially dependent and independent noise that compromise the statistical results. Misregistration is a key factor which has been discussed (Ashburner and Friston; Bookstein; Davatzikos 2004) and addressed using linear and nonlinear registration in recent VBM implementations (Ashburner and Friston 2000); fMRIStat allows users to decide which degree of normalization should take place.

Smoothing is generally used as a panacea to account for local misregistrations in the analysis, reduce noise in the input data and make the GM/WM concentrations normally distributed. This in turn improves the detection of regions of differences but comes at the cost of decreasing localization accuracy. The size of the smoothing kernel should be matched to the expected difference between groups (Ashburner and Friston 2000).

In order to report ensembles, or clusters, of differing voxels, as opposed to peaks, the issue of data nonisotropy must be successfully resolved (Worsley, Andermann et al.). Concentration volumes of GM / WM are highly nonisotropic since the noise component has nonconstant smoothness. One needs to correct the image and make the data isotropic in a statistical sense (Worsley, Andermann et al.) before assessing

results. Whereas many studies do not explicitly mention such correction, one must take care to incorporate a methodology of the kind proposed by Salmond (Salmond, Ashburner et al. 2002) or Worsley (Worsley, Andermann et al.) before reporting cluster statistics. These authors overcome nonisotropy by measuring cluster size in resels - resolution elements - rather than voxels. This is shown below, where the total cluster size **S** equals the summation of resels at these voxels **i** , part of cluster *cl*, above the cluster t-statistics threshold. Unitless resels are calculated by taking the ratio of the voxel volume *dx dy dz* in mm^3 over the cubed effective smoothing kernel full-width at half-maximum (FWHM) at that point, also in mm^3 (Worsley, Andermann et al.):

$$S = \sum_{i \in cl} \frac{dxdydz}{FWHM_i^3}$$

Deformation-based morphometry

The term "deformation-based morphometry" will be used to refer to various techniques that have emerged in the literature whose common trait is the exploitation of the information contained within the high-dimensional, dense deformation field obtained by nonlinear registration of brain images. Assessment of shape differences across individuals used to be done via the analysis and tracking of landmarks (Bookstein 1984) , but those are difficult to identify reliably and consistently across the human brain. In essence, by calculating the

registration vectors required to align every point in an image with those of a common reference (Figure 12), the deformation field contains continuous landmarks, and thus captures individual brain shape characteristics that can be compared within a group in the reference space.

Three morphometry techniques will be discussed: deformation, tensor, and surface-based morphometry.

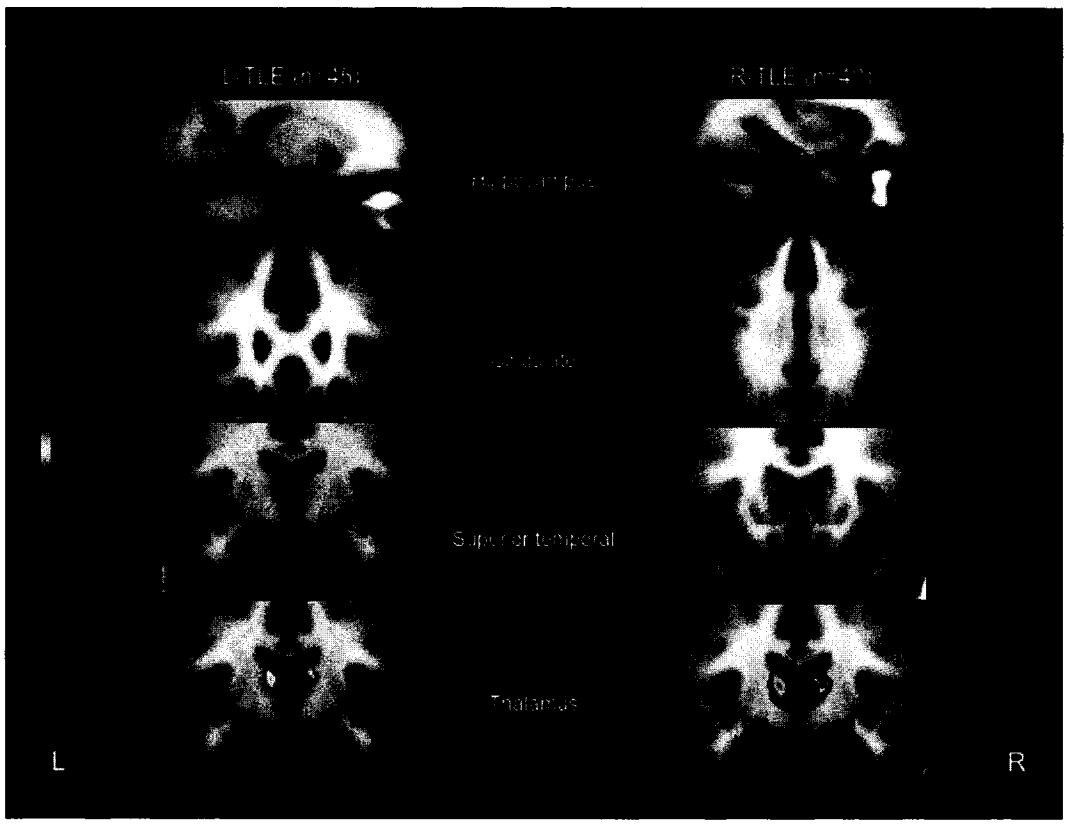

Figure 11 – VBM in TLE

Statistically significant peaks of gray matter decrease in patients with left and right temporal lobe epilepsy (L-TLE and R-TLE) in temporolimbic and frontal areas superimposed on the ICBM 152 average template of healthy controls for anatomical reference.
Source (Bernasconi, Duchesne et al. 2004)

Figure 12 – Deformation field

Computation of the nonlinear transformation required to align two images results in a dense deformation field, assessed at every voxel, capturing individual brain shape differences.
Source: D.L. Collins

- *Deformation-based morphometry* – When comparing groups, deformation-based morphometry uses registration-derived deformation fields to identify differences in the relative positions of structures within the subjects' brains. The 3D vector transformations can be analyzed as a multivariate statistical inference problem and solved using Hotelling's T^2 statistic (Thompson and Toga 1997); the result can be turned into a probabilistic atlas of deformations for a group of individuals. Ashburner and Friston (Ashburner and Friston 2000) point out that such statistics based on the displacement field do not directly localize regions within different

structures, but rather identify brain structures that have translated to different positions.

- *Tensor-based morphometry* – In those methods differences in the local shape of brain structures are derived from a tensor field (Chung, Worsley et al. 2001; Shen and Davatzikos 2003). By definition, the Jacobian of the deformation is the volume of the unit-cube after the deformation (Chung, Worsley et al. 2001). Assuming that one can find the deformation field at any voxel, volume change can thus be detected at a voxel level. In tensor-based morphometry (Ashburner and Friston 2000), the nine components of the Jacobian form scalar fields used to measure the second-order morphological variabilities. The advantage of this technique over the classical MRI-based volumetry is that it does not require *a priori* knowledge of the region of interest to perform the morphological analysis. Moreover, the deformation-based volumetry improves the power of detecting the regions of volume change within the limits of the accuracy of the registration algorithm (Chung, Worsley et al. 2001). To detect statistically significant local volume change, the T random field with its P value of the maximum field can be used (Chung, Worsley et al. 2001).

- *Surface-based morphometry* – Finally, the anatomical variations associated with the deformation of a surface can be statistically quantified; studies have looked at individual structure surfaces (Csernansky, Wang et al. 2000; Wang, Joshi et al.; Csernansky, Schindler et al. 2004), and others have applied this idea to the whole

cortex, modeled as a continuous sheet (MacDonald, Kabani et al. 2000; Chung, Worsley et al. 2003; Lerch, Pruessner et al. 2005).

High-dimensional classification

Only recently have classification methods emerged that use high-dimensional information (e.g. multiple 3D image features), typically reduced to a lower dimensional domain.

Techniques proposed by Golland et al. (Golland, Grimson et al. 2005) and Joshi et al. (Joshi, Pizer et al. 2002) rely on shape characteristics of single objects. So far these methods have shown value with discrete structures, easily identifiable, and with simple topologies: the hippocampus (Joshi, Pizer et al. 2002), hippocampus and the corpus callosum (Golland, Grimson et al. 2005), and the thalamus (Csernansky, Schindler et al. 2004). Those method become limited by much the same considerations as volumetry: 1) they do not represent interactions between neighboring, and biomechanically dependent, tissue; and 2) they are prone to being misled by segmentation errors and inter/intra-rater variability, due to inconsistent formulation of the anatomical boundary, reduced contrast (which happens frequently in multiple pathologies), and possibly poor segmentation technique.

The approach taken by Lao et al. (Lao, Shen et al. 2004)as well as Liu et al. (Liu, Teverovskiy et al. 2004) most closely resembles the methodologies proposed in this thesis. They both rely on information from a wide image area, as opposed to a

specific structure; they extract information directly from image parameters, and attempt the classification on the reduction parameters.

Lao et al. use wavelet decomposition of mass-preserved GM, WM and CSF 3D concentration maps as input to a Support Vector Machine classifier. Liu et al. uses 2D texture features on selected slices fed to a linear discriminant classifier.

Other techniques

New methods continuously emerge as more interdisciplinary collaborations exploit various elements of mathematical and computer vision theories, and older ones are continuously improved and brought in to aid the diagnosis of various neurological diseases. Those include medial sheets models (Joshi, Pizer et al.; Styner and Gerig; Bouix, Pruessner et al. 2005), 2D-cortical flattening (Drury, Van Essen et al. 1996; Fischl, Sereno et al. 1999; Van Essen, Drury et al. 2001), and sulci/gyri mapping. The latter attempts to capture the normal topological variability in sulcal and gyral patterns in the human cortex, considered by many to be one of the biggest challenges facing neurological imaging in the current era. Manual extraction of common topologies (Ono, Kubik et al. 1990) , along with automated techniques (Thompson, Schwartz et al. 1996; Le Goualher, Barillot et al. 1997; Lohmann and von Cramon 2000; Cachia, Mangin et al. 2003; Mangin, Riviere et al. 2004) are trying to capture the full extent of normal variability (Dameron, Gibaud et al. 2004).

Case studies

We will now turn our attention to outstanding problems within two neurological diseases whose characteristics make them ideal test cases to illustrate the capabilities of the methodologies developed in this thesis: temporal lobe epilepsy (TLE) and Alzheimer's dementia (AD).

Case study: temporal lobe epilepsy

Etiology and treatment

Microscopically, the most commonly described findings in MTLE are neuronal loss and gliosis of the hippocampus (HC) and the parahippocampal (PHC) region (Falconer, Serafetinides et al. 1964). The central goal of surgery in MTLE has been the removal of this affected tissue, in an attempt to remove the site of seizure generation and thus achieve seizure control. As the understanding of the disease and image-guidance improved (Olivier, Alonso-Vanegas et al. 1996), MTLE surgery has moved from complete to partial lobectomy (anterior temporal lobectomy or ATL), to selective amygdalo-hippocampectomy (SAH) (Olivier 2000). The latter approach consists in the resection of the amygdala (AG) and HC complex, and often parts of the neighboring structures, such as the entorhinal cortex (EC). Vickrey et al. (Vickrey, Hays et al. 1995) have assessed MTLE surgery success and demonstrated that surgery patients had greater decline in average monthly seizure frequency and

took fewer antiepileptic medications, when compared to patients that had not received surgery.

Imaging characteristics

HC neuronal loss found in surgical specimens obtained from patients with pharmacologically intractable TLE has been shown to correlate with HC atrophy on volumetric Magnetic Resonance Images (MRI) (Cascino, Jack et al. 1991). MRI studies in MTLE have also shown volume reduction ipsilateral to the side of the seizure focus in the PHC region (Bernasconi, Bernasconi et al. 2003); there is also evidence for extra-limbic and extra-temporal gray and white matter reductions in epilepsy patients when compared to controls (Woermann, Free et al. 1999; Keller, Mackay et al. 2002; Bernasconi, Duchesne et al. 2004). Nearly 80% of MTLE cases exhibit lateralized HC atrophy, strongly correlated with an ipsilateral seizure focus (see Figure 13), while approximately 5% of intractable TLE patients demonstrate bilateral atrophy without significant differences between hemispheres, and the remaining 15% of cases have non-atrophic HC (Bernasconi, Bernasconi et al. 2003).

Outstanding challenges

Lateralization of seizure focus in TLE

Presurgical evaluation of patients with refractory epilepsy is a necessary step, whose main outcome is the localization of the resection target. This thesis addresses the problem of patients with seizure focus, or foci, within the medial temporal lobes; one

of these will be selected as the resection target, a process referred to as *lateralization of seizure focus.*

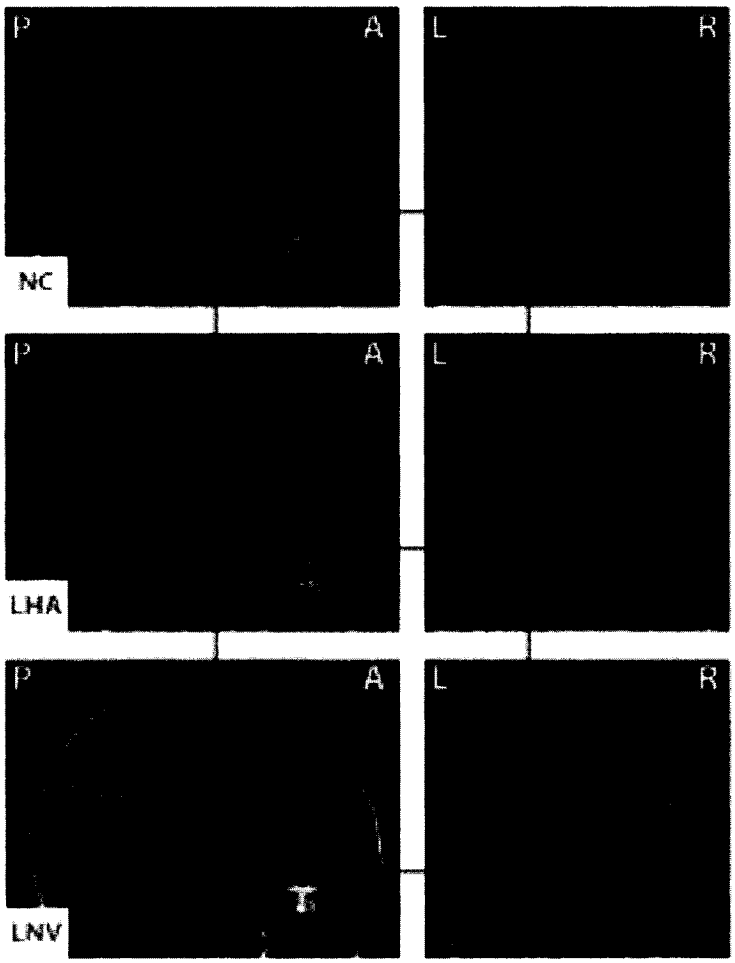

Figure 13 - T1 images of TLE

(Top) Sagittal and coronal T1 MRI of normal control subject (Middle) Left hippocampal atrophy is readily noticed on these images (same orientation as above) of a TLE patient as a enlargement of the lateral ventricles, along with increased atrophy of the hippocampal head (red lines intersection). (Bottom) On the other hand, around 15% of intractable TLE patients exhibit normal hippocampal volumes, therefore making lateralization impossible in those cases on the basis of HC volumetry alone. This is exemplified in these T1w MR images (same orientation as above) of a TLE patient with hippocampal normal volume.

As mentioned, HC atrophy is strongly correlated with an ipsilateral seizure focus. Volumetry of the hippocampus has therefore been used in many centers for pre-surgical evaluation however, it is a time-consuming, manual procedure requiring expert intervention, and subject to the drawbacks outlined in a preceding section. Further, approximately 20% of intractable TLE patients demonstrate either bilateral atrophy or normal volumes, making volumetry-based lateralization impossible in these cases. The alternative to HC volumetry is to proceed with surgically implanted, intra-cranial electro-encephalography (sEEG). While very sensitive, sEEG is a very invasive approach that is time and resource consuming, notwithstanding its impact on the patient's quality of life.

Our goal is to lateralize the seizure focus in TLE patients with and without hippocampal atrophy on cross-sectional T1-weighted MRI. The ability to perform accurate lateralization in a non-invasive manner could significantly alter current clinical practice. This forms the subject of Chapter 4 of this thesis; the methodology that has been developed and used to achieve lateralization is appearance-based classification, and is described in Chapter 3 in greater details than found in the published manuscript.

Prediction of outcome for TLE surgery

While the goal of surgery is to eradicate seizures, approximately 20 to 30% of patients are not seizure-free following standard surgery (Antel, Li et al. 2002). Predicting surgical outcome in the treatment of medial temporal lobe epilepsy

(MTLE) remains an outstanding challenge. Since the advent of high-resolution neuroimaging, there have been numerous attempts at finding structural or functional markers that can be used as a predictive surrogate of surgical outcome, but no completely reliable indicator has been found to date.

While the majority of patients undergoing surgery have positive outcome, it is difficult at present to determine *a priori* if the procedure will be successful. MRI has been used to predict surgical outcome in ATL (Jack, Sharbrough et al. 1992); (Radhakrishnan, So et al. 1998) (Kobayashi, Lopes-Cendes et al. 2001) with an accuracy ranging from 62% to 96%(Gilliam, Faught et al. 2000). Fewer authors have attempted to predict surgical outcome for SAH based on MRI: Abosch et al. (Abosch, Bernasconi et al. 2002) have used MRI, and Antel et al. (Antel, Li et al. 2002) have used MRI in combination with MR spectroscopy. The seizure-free classifier developed by Antel correctly predicted the surgical outcomes of 39 of 52 (75%) of patients who became seizure free and 21 of 29 (72%) of patients who did not.

There remains a need to increase the accuracy of non-invasive, pre-operative surgical outcome prediction in SAH. This represents the goal of the work presented in Chapter 5; more information on the necessary background is included in that chapter, and the voxel-based classification methodology used to achieve this goal is described in detail in Chapter 3.

Case study: Alzheimer's dementia and Mild Cognitive Impairment

Etiology and treatment

Alzheimer's dementia (AD) is a progressive neurodegenerative disorder associated with disruption of neuronal function and gradual deterioration in cognition, function, and behavior (Khachaturian 1985) . The progression of AD is gradual, and the average patient lives 8 to 10 years after onset of symptoms (Petrella, Coleman et al. 2003). It is the most common cause of dementia in the elderly (65 years and older), responsible for 75% of all dementia cases (Group 2000; J. R. Petrella 2003). As the global life expectancy rises and populations age, the prevalence of AD is expected to triple over the next 50 years (Carr, Goate et al. 1997); improving care while reducing the socio-economic impact of AD is therefore an important and necessary topic of research.

The etiology of AD can be summarized as neuronal dysfunction and eventual loss due to abnormal accumulation of Aβ and Tau proteins (Thal, Rub et al. 2000; Thal, Rub et al. 2002; Giannakopoulos, Herrmann et al. 2003). Neuropathological studies in AD have shown in fact that brain degeneration occurs very early in the course of the disease, even before the first clinical signs, in certain regions such as the medial temporal lobe (MTL)(Cummings, Pike et al. 1996; Nagy, Hindley et al. 1999; Thal,

Rub et al. 2000; Thal, Rub et al. 2002; J. R. Petrella 2003). Microscopically the strongest predictor of premortem cognitive dysfunction appears to be the relative area of entorhinal cortex occupied by beta-amyloid deposition(Cummings, Pike et al. 1996). To date however the diagnosis of clinically probable AD can be made with high accuracy in living subjects only once the stage of dementia has been reached, and requires clinical, neuropsychological and imaging assessments (Petrella, Coleman et al. 2003). It can only be confirmed by postmortem histopathology (Risse, Raskind et al. 1990).

Imaging characteristics

AD is associated with progressive cerebral atrophy, which can be seen on MRI with high resolution (Fox and Schott 2004)(see Figure 14). This atrophy is preferentially located in early AD in the temporal and limbic areas (hippocampus, temporal pole, insula, with co-occurring enlargment of the lateral ventricles), while in later disease stages, the rest of the cortex (frontal, occipital lobes) is affected, with relative sparing of the sensori-motor cortex (Karas, Burton et al. 2003; Thompson, Hayashi et al. 2003; Thompson, Hayashi et al. 2004).

Longitudinal analysis of MRI has been proposed to differentiate between aging, MCI, and AD (Fox and Freeborough 1997; Fox and Schott 2004) with high accuracy however, by its very nature such a method implies a delay between scans before any assessment can be made. Cross-sectional measurements of the hippocampus have achieved classification between controls and AD patients (90.7%), and between

individuals with MCI and AD patients (82.3%) (Pennanen, Kivipelto et al. 2004). Longitudinally-assessed MR-based atrophy rates of the hippocampus is also a reliable indicator of the state of the disease (Jack, Slomkowski et al. 2003).

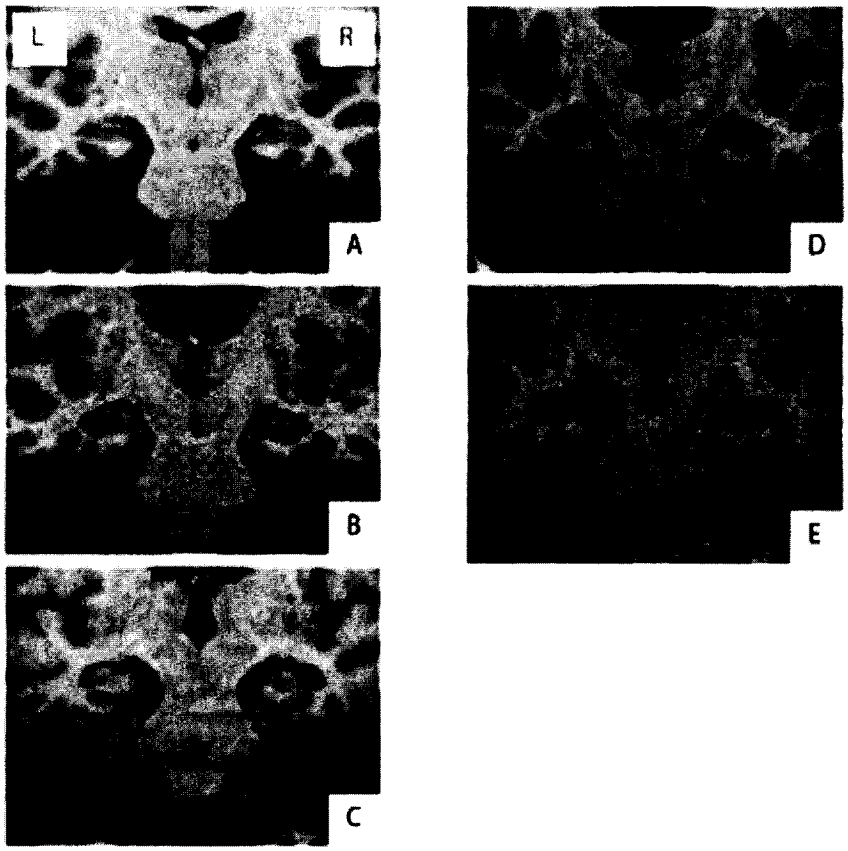

Figure 14 - T1 weighted images of AD patients

Those images perpendicular to the long axis of the temporal lobe show the different degrees of medial temporal lobe atrophy (MTA), according to the visual rating scale proposed by Scheltens (A) absence of atrophy (MTA=0); (B) minimal atrophy (MTA=1); (C) mild atrophy in the right side (MTA = 2), severe atrophy on the left (MTA=4); (D) moderate atrophy (MTA=3); and (E) severe atrophy (MTA=4) (Source: (Bastos Leite, Scheltens et al. 2004)

Outstanding challenges

Differentiation of AD and MCI from normal aging

The difficulties in early clinical detection lie for the most part in the similarities between cognitive impairment due to normal aging (NA) processes and the initial manifestations of AD (Chetelat and Baron 2003). The ability to differentiate AD from NA, in an objective, repeatable and accurate fashion, would act as a potent factor in the therapy management of these individuals. Additional background on this topic is included in Chapter 6, where the differentiation of AD and MCI from NA is attempted using the same appearance-based technique (described in Chapter 3) that was used for lateralization of seizure focus in TLE in Chapter 4.

Early prediction of MCI progression to AD

Mild cognitive impairment (MCI) is a condition referring to patients with significant but isolated memory impairment relative to subjects of identical age (Flicker, Ferris et al. 1991; Petersen, Doody et al. 2001). MCI individuals are considered an at-risk group for progression to Alzheimer's dementia (AD)(DeCarli 2003; Dubois and Albert 2004). Early prediction of progression to AD in MCI patients is therefore an important research goal that is met in Chapter 7, and which may have significant impact on the therapy course for these individuals. Again, further relevant material is included in the introduction of this manuscript, and the appearance-based classification methodology is described in the next chapter.

Chapter 3

Methods

In this chapter, two novel image analysis techniques are introduced as the methodological contributions of this thesis. Their application to existing diagnostic problems in classification and prediction of neurological diseases will be discussed in subsequent chapters.

There are three major design considerations underlining both techniques. The first decision is to move from structure-centric approaches to one based on a large volume of interest (VOI), hoping to increase the informational value (see Figure 15). For example, in volumetric studies of the HC, between 3000 to 4000 voxels are compressed in a single scalar measurement, expressed in mm^3. Even though essential information about disease processes can be gained through volumetric studies (e.g. atrophy rates in AD), much information is lost in this process; the location of atrophy being a prime example in this case. In the methodologies being proposed

here, more voxels are processed and linearly modeled to yield additional important information towards our classification/prediction goal.

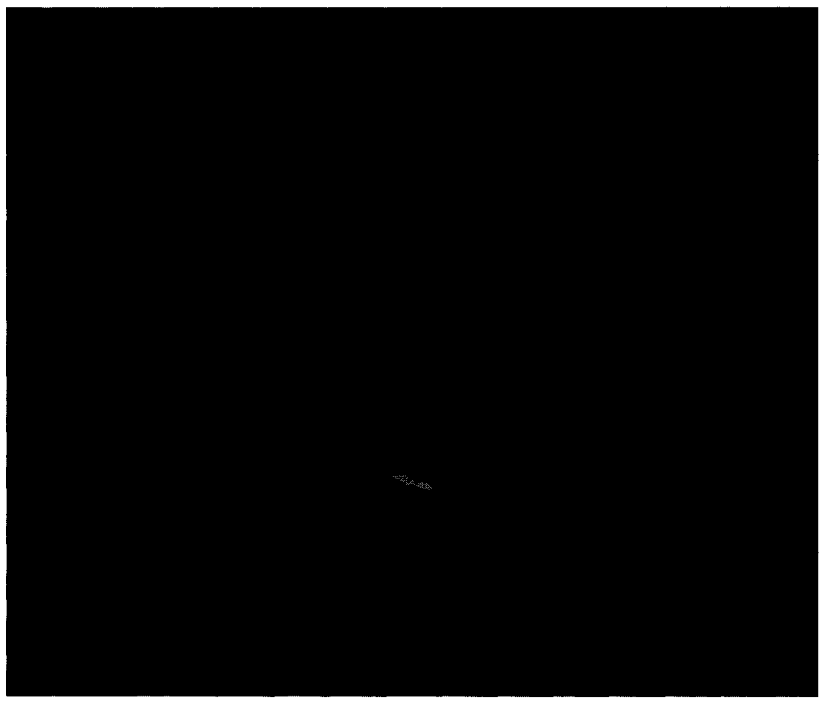

Figure 15 – Structure-centric *vs.* volume of interest approaches
Volumetry of individual structures (e.g. hippocampus, in green) has been used extensively in the analysis of medical images but offers limited informational value. A large, non-specific 3D volume of interest (shown here in blue) would contain additional information potentially useful for classification and prediction.

The second design choice is a move from *a priori* knowledge-based approaches to a completely data-driven technique. Our goal is not to preselect a structure for analysis but rather to investigate the relationship between voxels within the volume of interest, as they relate to the pathology. This is the common biological hypothesis underlying the two methods to be presented in this section: that differences in MRI signals are likely due to pathology-related microscopic and macroscopic changes or

alterations in the various tissues within these voxels. It is therefore a conscious decision to not presuppose the pathological effect on the signal intensity, or its affect on any structure. Again, we can contrast this choice with volumetry, which works on the basic assumption that some pathologies result in cumulative tissue loss, measurable as a dichotomous change in boundary voxels and therefore structure volume. It is likely however that these pathologies alter the signal intensity in an incremental fashion, and thus, by compressing continuous change into a single binary step, one reduces the dynamic range and gets low sensitivity. A data-driven approach, where image features are extracted without *a priori* hypothesis, allows the modeling of subtle signal changes that, arguably, may be harder to interpret biologically, but serve our ultimate purpose of classification and prediction of pathological effects by virtue of being potentially more sensitive.

The third major consideration is to restrict the analysis to cross-sectional data, as opposed to serial or longitudinal studies. This stems from a clinical need in rapid assessment of pathological effects in patients, but implies that both methodologies will rely on individual-to-group comparisons for classification and prediction. This of course assumes that the disease effect within each group will be sufficiently homogeneous. One way to improve upon this homogeneity in the four applications presented in Chapters 4 to 7 was to carefully assign patients in cohorts with similar age and/or clinical condition.

It should be stated at this point that the first methodology (appearance-based classification) was sufficiently developed and its application satisfactorily proven in TLE and AD as to warrant protection for intellectual property, resulting in a patent application to the United States Patents and Trademarks Office on 19 Nov. 2004 (no 10/990396).

MRI acquisition and initial image processing

Pre-operative 3D T1-weighted (T1w) MR images in all of the following studies were acquired on multiple 1.5T scanners with different sequences however, initial image processing steps were common to all datasets and will now be described.

Global MRI scans were first corrected for intensity non-uniformity due to scanner variations (Sled, Zijdenbos et al. 1998), and then globally aligned using a 9-degrees of freedom (DF) linear registration (Collins, Neelin et al. 1994) into a standard reference space. Following registration the data was resampled onto a $1mm$ isotropic grid (Collins, Neelin et al. 1994). The reference image used for the linear registration and resampling was the ICBM 152 T1-weighted target, a voxel-by-voxel average of the 152 normal subjects previously registered in the Talairach-like stereotaxic space in the context of the ICBM project (Mazziotta, Toga et al. 1995). This standard pipeline has been in use at our Institute and is described elsewhere (Zijdenbos, Forghani et al. 2002)(see Figure 16).

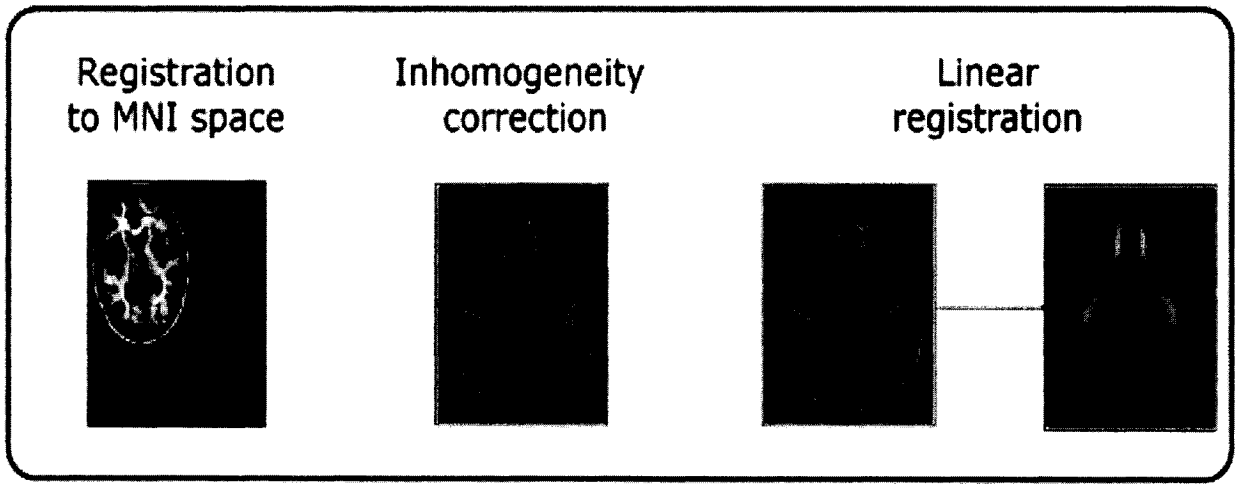

Figure 16 – Initial image processing pipeline

Appearance-based classification (ABC)

The first technique presented is based on the concept of appearance, which encompasses not only the shape of the structure or volume of interest, but also its grey-level intensities.

The field of appearance-based analysis in medical imaging owes much to the contemporary work of Cootes and colleagues, whose first seminal publication on the topic of point distribution models (PDM) dates back to 1991 (Cootes, Cooper et al. 1991). This segmentation concept was extended to Active Shape Models (Cootes, Taylor et al. 1995), which creates a model of the shape of the structure of interest that describes both typical shape and typical variability, using statistical analyses of previous examples from a training set. Finally, Active Appearance Models (AAM) were introduced in 1998 (Cootes, Edwards et al. 1998) in part to further their

original method and make more use of the intensity information in the images. Knowledge about the shape is incorporated into the AAM via a PDM. The intensity under the landmarks in the PDM is sampled and used to generate a gray-level model. The two models—gray-level and PDM—are concatenated, and a supermodel is created from principal components analysis (PCA). The resulting principal components that explain the most variation in the supermodel are then selected. Those principal components can be considered as eigenmodes of appearance variation, embedding shape and intensity variability.

Their goal in developing the PDM, ASM and AAM was to match a full, photorealistic model directly on a new image instance by minimizing the difference between the image under interpretation and one synthesized by the model. We have retained the concept of appearance (intensity and shape) but have widen its applicability from segmentation (Duchesne, Pruessner et al. 2002) to classification. The method can be summarized as follows (see Figure 17) and is described later in subsequent paragraphs. First, from processed data **(A)** we generate a non-pathological appearance eigenspace **(B)** from a large training group of young, neurologically healthy training subjects ($N = 152$). This multidimensional eigenspace is created by uniting results from four distinct PCA of (i) linearly registered intensity images of the left and right VOIs from T1w MRI of these training subjects **(C)**; and (ii) an approximation of the Jacobian of the deformation fields for the same VOIs **(D)**, fields obtained via non-linear registration to a

common reference image. Secondly, VOIs from patients are projected in the multidimensional appearance eigenspace created **(E)**. The normality of the distribution of the projected eigencoordinates is assessed using quantile plots and Shapiro-Wilke statistics. The last step **(F)** consists in linear discriminant analyses (LDA) in a leave-one-out, forward stepwise approach, using eigencoordinates in the reference appearance eigenspace to classify the patients. While the eigenspace may not be optimal to represent the patient population, it forms the basis for a comparative evaluation of the different groups that leads to their classification. The methodological details are elaborated in the following sections. The reader will note that some of this work was first presented by the author as preliminary works in progress in conference proceedings (Duchesne, Bernasconi et al. 2002; Duchesne, Bernasconi et al. 2003), and finally as a peer-reviewed article (Duchesne, Bernasconi et al.), which forms the next chapter of this thesis.

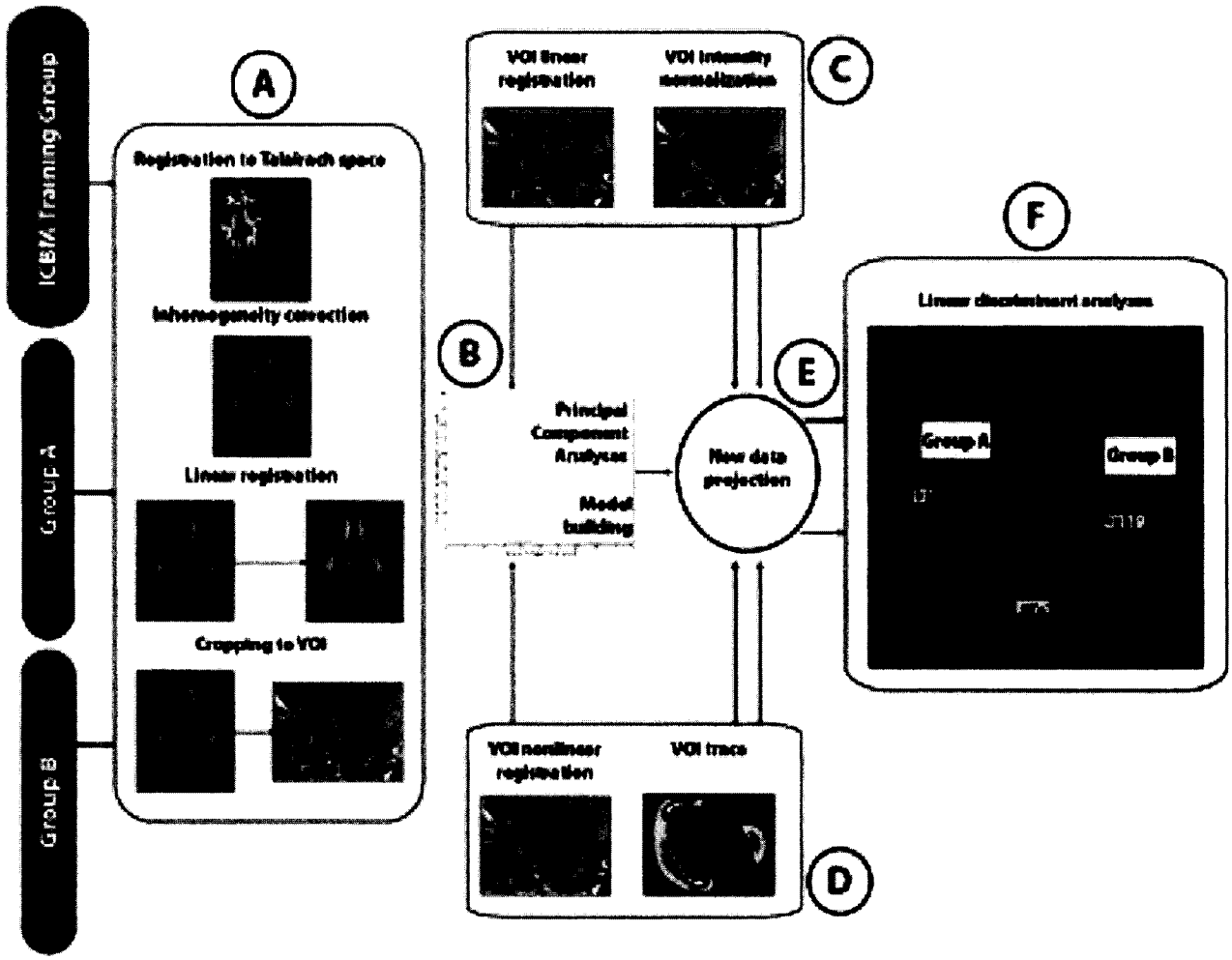

Figure 17 – ABC methodological pipeline

Volume of interest

As mentioned, the approach is not structure-centric, but rather based on a regional analysis within a 3D VOI. Two VOIs were selected in fact, centered on the left (Figure 18) and right medial temporal lobe respectively, using Talairach-like MNI space coordinates (start coordinates x = [−53,+2] for the left and right side, y = −53 and z = −52). Each VOI measured n = 55 X 82 X 80 = 360800 voxels. The VOI was selected so that its extent captured the hippocampus and neighboring MTL

structures, such as the parahippocampal gyrus, irrespective of normal inter- and intra-individual variability.

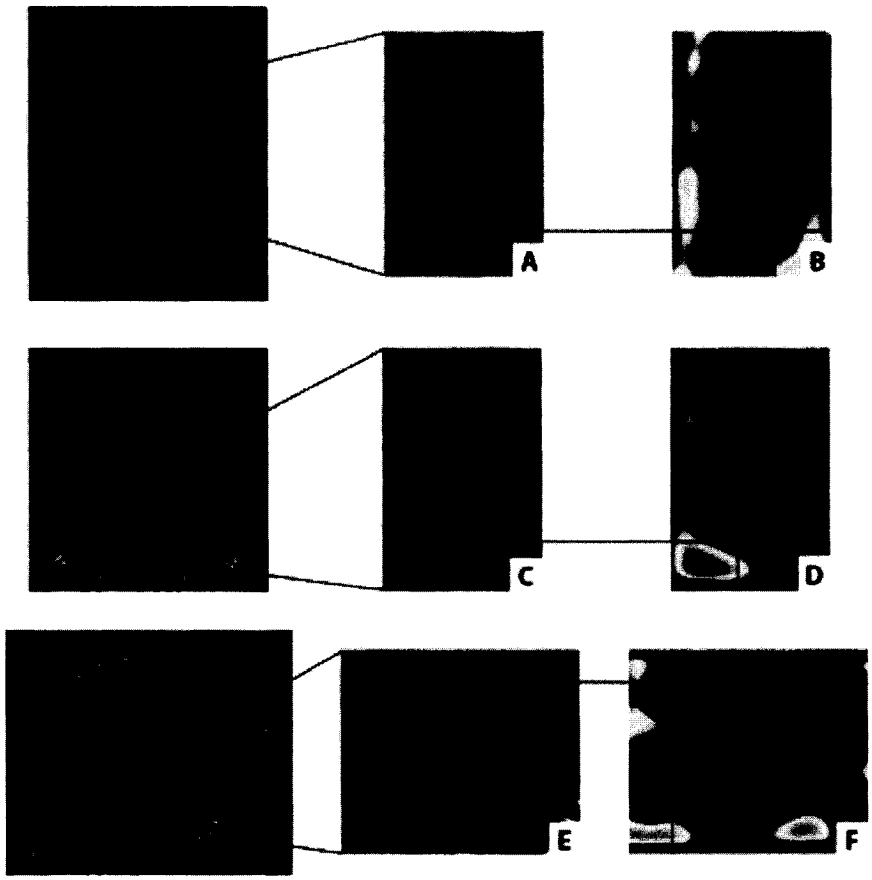

Figure 18 - Intensity and shape features within the left VOI in ABC methodology

(Left column) Transverse, coronal and sagittal whole-brain views for TLE subject #28 with LHA. (Middle column) Identical views through the VOI, as defined in MNI space on the left MTL (red boxes on whole-brain views). (Right column) Identical views trough the trace VOI for the same subject. Green voxels do not move. From green to white (maximum) via yellows and reds indicate increase or expansion. From green to black (minimum), via blues and purples indicate decrease or contraction. The direction of movement is defined as the deformation which the subject's VOI seen in (A), (C) and (E) must accomplish in order to align with the corresponding VOI extracted from the reference volume. Contraction represents atrophy in the case of tissue. Notice in the transverse view in (A) and (B) the contraction of the lateral ventricular space, necessary to correct its enlargement, also noticeable in (E) and (F), while in (C) and (D), the coronal view displays the superiocentral shift due to hippocampal atrophy.

Intensity data

Two image features at each voxel location were retained for modeling. The first feature used in our classifier is *grey level intensity*, denoted by script ***g***. Scanner output volumes are preprocessed in the manner described in the preceding section on initial image processing, and VOIs were extracted based on stereotaxic coordinates. To further reduce positional variations, which would propagate as unwanted noise in the morphometric modelling, a second local affine transformation (DF = 12) is applied to co-register the individual subject's VOI with the corresponding VOI on the ICBM 152 target. Finally, VOI intensity mean is also linearly scaled to the reference target. The input to the intensity model therefore consists in the rasterized intensity data from the VOIs processed in this fashion (Figure 18). Using intensity features from MRI raises the question of calibration and normalization. Absolute intensities are rarely used in MRI, since they vary with machine calibration, shimming, and patient-induced variations. We have tried to limit those variations by (1) using the same scanner for every individual in the study; (2) ensuring that the same quality assurance procedures were followed for each acquisition; and (3) scanning the patients in random group order.

Shape data

The second feature is the *trace*, denoted by script ***t***. The trace is the first-order approximation of the Jacobian of a non-linear registration-derived deformation field,

which maps a subject to our reference ICBM target, and thus forms an estimate of local volume change.

Nonlinear registration attempts to match image features from a source volume to those of the reference image at a local level, typically in a hierarchical fashion, with the aim of reducing a specific cost function. Whereas many nonlinear registration processes exist, the one chosen for this study was ANIMAL (Collins and Evans 1997). This algorithm attempts to match image grey-level intensity features at a local level (voxel) in successive blurring steps, by maximizing the cross-correlation function of voxel intensities between the source and reference images. The result is a dense deformation field capturing the displacements required to align all voxels within the subject VOI with those of the reference image. In our application, we estimate the nonlinear transformation between the VOI of an individual subject to the ICBM 152 target VOI. Even though the method does not require that the same VOI be used for both trace and intensity input, we have used the same in the four studies proposed in this thesis.

Our implementation of trace calculation was discussed in (Janke, de Zubicaray et al. 2001) and follows the notation developed by (Chung, Worsley et al. 2001). If U is the displacement field which matches homologous points between two images, then the local volume change of the deformation in the neighborhood of any given voxel is determined by the Jacobian J (Chung, Worsley et al. 2001), which is defined as

$$J(x,t) = \det(\mathrm{I} + \frac{\partial U}{\partial x})$$

where **I** denotes the identity matrix and $\frac{\partial U}{\partial x}$ is the 3 X 3 displacement gradient matrix of U. For relatively small displacements (i.e. for small rotations)

$$J \approx 1 + tr(\nabla U)$$

as suggested by (Chung, Worsley et al. 2001). Hence the trace $tr(\nabla U)$ represents a crude yet indicative measure of local volume change. When the change is near zero in the neighborhood of *x*, the deformation is *incompressible* so there is no volume change. However, if the trace is positive, the volume increases while when negative, the volume decreases after the deformation.

It could be argued that the determinant should be used, rather than the trace; first, the determinant measures a biologically meaningful quantity, that is brain tissue volume change; second, the latter is the sum of the eigenvalues whereas the former is their product. In the case of brain deformations, the numerical difference between the two measurements would be small. Our computing the determinants for each reference MRI and comparing them with the trace volumes proved this fact. The average difference between trace and determinant shows a mean of 0.006 (indicative of no bias) and a standard error of 0.011, pointing to a limited variability between the two measures. The trace of the Jacobian was retained for our modeling since the calculation time was one to two orders of magnitude faster than the determinant at the time; it is proposed to move to determinant calculations as a marginally better approximation of local volume change in the future, and as a way of ensuring that one extracts biologically meaningful information from the data.

The input to the PCA therefore consists in rasterized vectors of calculated trace volumes, for which examples are shown in Figure 18.

Multidimensional spaces

Principal Components Analysis (PCA) is used to reduce the dimensionality of the input data (intensity *g*, trace *t*) and generate linear variation models based on the N=152 datasets from our ICBM normal control subjects.

Each intensity or shape VOI can be represented by a single point in an *n* dimensional space (*n*-D), where *n* is the number of voxels in the VOI. Thus the ensemble of reference intensity or shape VOIs gives a cloud of points in this multidimensional space. It is assumed that these points lie within some region of the space, which are called "Allowable Domains" (similar to "Allowable Shape Domains" in Cootes' work (Cootes, Edwards et al. 2001)) and that the points give an indication of the form and size of this region. Every point within this *n*-D domain relates back to an intensity or shape VOI whose characteristics are broadly similar to that of those in the original reference set upon which the Domain has been defined.

The following mathematical reasoning is aimed at constructing independent linear models for each intensity or shape Allowable Domains. As such, it is given in its general form, with a notation identical to that developed by Cootes *et al.* (Cootes, Edwards et al. 2001).

Given a set of N reference examples x_i, the mean is calculated using:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

where x_i is the column vector of intensity or trace values for a given VOI instance, of dimension n. The principal axes of an ellipsoid fitted to the ensemble data can be calculated by applying PCA, where each Principal Component (PC) axis yields a direction of a mode of co-variation, a way in which the intensity or shape instances tend to move together.

For each example i in the reference set, the deviation dx_i from the mean x_i is calculated, where

$$dx_i = x_i - \bar{x}$$

In order to find the basis for this space, the $n \times n$ covariance matrix, S, is calculated using:

$$S = \frac{1}{N}\sum_{i=1}^{N} dx_i dx_i^T$$

The principal axes of the hyper-ellipsoid, giving the modes of variation of the intensity or trace examples, are described by p_k (k=1,2,...n), the unit eigenvectors of S such that:

$$Sp_k = \lambda_k p_k$$

where λ_k is the k-th eigenvalue of S, $\lambda_k \geq \lambda_{k+1}$.

The ensemble of PCs from the left and right VOI grey-level intensity models defines an Allowable Grey Domain **G** as the space of all possible elements expressed by the intensity eigenvectors λ_G. Likewise, an Allowable Trace Domain **T** is created as the space of all possible elements expressed by trace eigenvectors λ_T for the left or right VOIs.

Within these Domains, linear models allow any new vector instance x, be it trace or intensity vectors from the left or right VOI, to be approximated by $\bar{x}$, the mean normalised input vector; P_x, the set of orthogonal modes of variation; and B_x, the set of parameters:

$$x = \bar{x} + P_x B_x$$

Each mode of variation is orthogonal to others within the same feature model. Covariances between features are not excluded, but most of them should be accounted for in the forward stepwise discriminant process.

The assumption is made that the form of these Domains in a high dimensional space is approximately ellipsoidal (Cootes, Edwards et al. 2001). It can be shown that the eigenvectors of the covariance matrix corresponding to the largest eigenvalues describe the longest axes of the ellipsoid, and thus the most significant modes of variation in the variables used to derive the covariance matrix (Cootes, Edwards et al. 2001). Most of these variations can usually be explained by a smaller number of modes, l, where $l << n$ and $l < p$. This means that if the n dimensional ellipsoid is approximated by an l dimensional ellipsoid, then the original ellipsoid has a relatively small width along axes with indexes $l + 1$ and greater (Cootes, Edwards et al. 2001).

One method for calculating l is to chose the smallest number of modes such that the sum of their variances explains a sufficiently large proportion of λ, the total variance of all the variables, where

$$\lambda = \sum_{k=1}^{n} \lambda_k$$

In order to know the variance of each eigenvector, or how each PC contributes to the description of the total variance of the system, the ratio of relative importance of the eigenvalue λ_k associated with the eigenvector k is used:

$$r_k = \frac{\lambda_k}{\sum_{j=1}^{p} \lambda_j}$$

where r_k is the fraction relative importance for eigenvalue λ_k, over the sum of all λ, and p is the total number of eigenvectors. Incidentally, $p=N-1$ in this implementation, adapted from that described in Appendix A of Cootes *et al.*'s technical report (Cootes and Taylor 2000). This algorithm reduces the dimensionality of the matrices to be estimated and hence ensures that substantial changes in VOI size can be easily accommodated.

Aiming for a given percentage F of explained variation, one wants the smallest f eigenvectors such that the following condition is satiseizure focusied:

$$r_1 + r_2 + ... + r_f > \frac{F}{100}$$

The choice of threshold F will be dependent on the total amount of variation that the model is asked to represent. We define a restricted version of the domains **G** and **T** as **G*** and **T***, such that

$$G^* \subset G$$
$$T^* \subset T$$

where the upper bound on the dimensionality of **G** or **T** is p, as defined above, while

the upper bound for **G*** or **T*** is *f*, set by choosing the threshold F.

In all applications of this methodology herein, we proceeded in selecting a total of 538 eigenvectors out of a possible $(152-1)\times 4 = 604$ for the four models (left/right intensity/trace VOIs) built from our reference group, corresponding to F=99.7% of the per-model variance.

Rasterized vectors of the processed VOI intensity and trace data for each study subject are projected into Domains **G*** and **T***, and thus form the eigencoordinate vectors. While a number of possible features can be calculated based on the distribution of the projected data, our classifier is based on the position along the PC axes. The distribution of eigencoordinates along any principal component for a given population should be normally distributed and this normality is assessed via quantile plots and Shapiro-Wilke statistics. Classification, based on linear discriminant analyses of the eigencoordinate distributions, will be discussed in a subsequent section in this chapter.

Methodological considerations

Representation - Our approach has been to create a model eigenspace based on subjects from the ICBM database, in which we projected our cohorts of patients. While such a space built from young, neurologically healthy individuals may not be optimal to represent the test populations, it should be noted that our primary goal was to find an independent basis for a comparative evaluation of our study cohorts

and its subsequent classification, but not the best mathematical representation. This is exemplified in Figure 19, where a cartoon-like representation of our training set is drawn, aligned and centered with the first two principal components since they were extracted from that data. Our goal is to use this space for the separation of pathology #1 subjects into two groups, 1a and 1b. While the space may not be optimal for pathology 1, it forms a reference basis built from young, neurologically healthy individuals, that can also be used for other pathologies, such as pathology #2 that may or may not overlap with pathology #1.

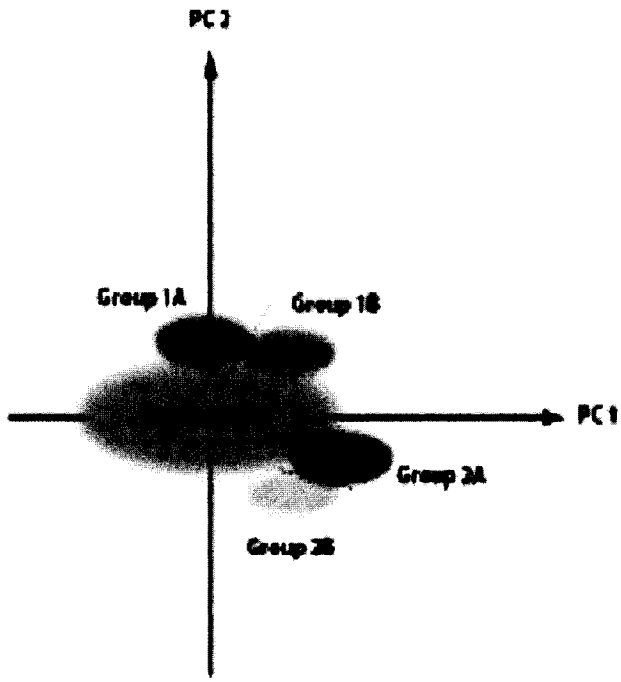

Figure 19 – Cartoon-like representation of classification of patient groups.

The reference set is represented by the green ellipse, aligned and centered with the first two principal components since the eigenspace originates from these data. Our goal however is to use the same space for the separation of many pathologies. For example, pathology #1, while not optimally represented by the eigenvectors of the reference space, can still be separated in two groups, 1A and 1B. The same can be said of a second pathology, which may or may not overlap with the previous one. Using a normative space built from young, neurologically healthy individuals should give the ability to perform classification studies in multiple pathologies.

Independence - As the reference set is composed of ICBM subjects that are separate from the trial subjects, there is no issue of over determination in the creation of the reference space. Further, by using cross-validation trials, separating trial subjects into training and test sets, we ensured that there was no over learning in the classification stage while maximizing the amount of information available in the trials subjects.

Automation - The methodology is formulated to be simple, flexible, and easily scalable, allowing a number of possible features to be used without extensive reengineering. We achieved substantial data reduction, starting from a model of 152 reference images (152 X 2 VOIs X 360800 voxels = $1.1 \text{ X } 10^7$ points in MRI space) down to 538 eigencoordinate points in a predefined PC space.

RAVENS-based classification - There are similarities and differences between the ABC methodology and the classification method proposed by Lao et al. (Lao, Shen et al. 2004). Both reduce a high-dimensional input vector and attempt the classification on the reduction parameters, albeit using a different technique (PCA for ABC; wavelet decomposition for Lao et al.). As measured in Liu et al. (Liu, Teverovskiy et al. 2004), the strength of the classification does not reside in the optimality of the classification function but rather in the appropriate choice of features. It is there that the main difference between the two techniques resides. Lao et al. uses mass-preserved GM, WM and CSF maps, which are essentially trinarized

versions of the original intensity image, whereas the ABC technique makes use of the full dynamic range of the data, and therefore has access to additional information. Further, the ABC technique explicitly models local volume changes, as opposed to its implicit effect in the mass-preserving framework of the RAVENS maps (Davatzikos, Genc et al. 2001).

VBM-based classification (VBM)

The second methodology that forms part of this thesis is based on the well-known statistical analysis approach called voxel-based morphometry (VBM). Previously used solely for the purpose of finding and locating group differences in anatomy, it is extended here to the classification of individual MRIs based on group patterns and therefore constitutes to our knowledge one of the only application of VBM for assessment of images from a single individual.

Detection of concentration differences related to grouping variable

The primary objective is to determine regions of gray and white matter concentration differences that are related to the grouping variable. To this end multiple between-group VBM studies are performed (Ashburner and Friston 2000),

on combined cohorts of co-registered subject image data. The software package used in this thesis is fMRIStat from Worsley et al. (Worsley, Marrett et al. 1996).

The approach involves co-registering images in a linear fashion (affine registration, 9 DF) to improve the spatial co-location of brain structures, with a common reference target of 152 young healthy volunteers, the so-called MNI brain or ICBM 152 average (Mazziotta, Toga et al. 1995). Brain tissue is then segmented (GM, WM) in volumes (Zijdenbos and Dawant 1994) where each voxel is assigned to one class. GM/WM maps are then blurred using an isotropic Gaussian kernel of $10mm$ full-width at half-maximum, under the assumption that this spatial kernel size matches the expected changes (Ashburner and Friston 2000). In those smoothed "concentration maps", each voxel takes on a value between (0,1) indicative of the presence or not of GM or WM in that voxel. An example average concentration map for a group of TLE patients with left hippocampal atrophy is shown in Figure 9.

Parameters of a general linear model are estimated at each voxel, across subjects, possibly with additional co-variates if applicable (e.g. age, sex). Voxels with statistically significant differences are identified based on a threshold corrected for multiple comparisons. Rather than using Bonferroni correction, fMRIStat relies on Gaussian Random Field theory for this correction (Worsley, Marrett et al. 1996). The final result is a volume containing only those voxels (or peaks) in which there are statistically significant differences between the tissue concentrations for the groups under study (see Figure 11).

Significant clusters, formed by a sufficiently large number ("extent") of contiguous voxels above a given threshold t-value (not necessarily peak voxels under the previous threshold for statistical significance), are obtained following the procedure developed by Worsley *et al.* (Worsley, Andermann et al. 1999) to correct for data anisotropy.

Similarity measures

Our hypothesis for VBM-based classification is that areas of differences in GM or WM concentration related to the grouping variable, which have been extracted in the previous step, can be exploited for classification purposes. We thus proceed in defining a global region of interest as the ensemble of all voxels *j* above the cluster statistical significance threshold *t* in each of the four concentration difference maps (one each for the combination of increase/decrease in GM/WM concentrations). The first measure selected was the straightforward calculation of the total number of voxels *j* within the region, *V*, and the second feature was the mean GM or WM concentration of all region-of-interest voxels *j*. Thus, for subject *i*, with concentration map $[GM]^{+}_{i}$ (for increases in grey matter concentration), the measure M_i was calculated as follows:

$$\overline{GM}_i = \frac{1}{V}\sum_{j=1}^{V}[GM]^{+}_{ij}$$

with similar measures derived for WM and the same for decreases in

concentration ($[GM]^+$, $[GM]^-$, $[GM]^+$, $[WM]^-$).

Other metrics are possible; we have proposed a "weighted" concentration measure (Duchesne, Bernasconi et al. 2004), consisting in multiplying the GM or WM value in the region of interest voxels by its t-statistic and taking the mean weighted concentration value. This measure was rejected experimentally, as it did not increase the accuracy of classification.

Ideally there would be two groups of subjects, one to extract the global area of interest, and the second to test the classification. However, in cases where there is a limited number N of patients, one can proceed in a leave-one-out fashion where multiple VBM group comparison analyses are carried out, with each subject removed in turn from the group; the concentration difference maps are therefore independently assessed from that subject. The global region of interest can be calculated as defined above for each trial, and the number of voxels in the region, along with the mean GM/WM concentration within that region, measured on the subject's own MR image. This approach is repeated N times, in order to estimate the classifier's accuracy.

Classification

The classification approach used for the applications presented in this thesis was similar for the two methodologies described above. Every experiment was dichotomous, and while many possible choices of classifier existed, linear

discriminant analyses (LDA) were chosen due to their simplicity and ease of implementation.

Closely following the notation of Duda *et al.* (Duda, Hart et al. 2001), we defined two states of nature ω_1, ω_2 in our experiments (e.g. left MTL *vs.* right MTL seizure focus subjects, or AD and MCI patients *vs.* normal aging). In each case, the subjects being classified are part of the trial group, and not the same as the reference group used to compute the MR-based features defining the reference space.

The prior probabilities $p(\omega_1)$, $p(\omega_2)$ are usually fixed at 0.5, and do not necessarily always represent the normal population incidence. Exceptions to this rule will be noted in the text, and consisted in adjusting the prior probabilities to reflect the composition of the trial groups.

The design of our classifier is simple. For each variable, the LDA function can be expressed as follows:

$$g(v) = w^t v + w_0$$

where v is the feature vector (e.g. eigencoordinates, VBM-based measures), w is the weight vector and w_0 the bias or threshold weight. For a two-category classifier we implement the following decision rule: decide ω_1 if $g(v) > 0$ and ω_2 if $g(v) < 0$.

There were two possible implementations for LDA: complete estimation and forward, stepwise analysis.

- *Complete estimation* - all variables were included in the model;

- *Forward, stepwise analysis* - variables were entered stepwise into the analysis by Wilk's λ method; that is, at each step the variable that resulted in the smallest Wilk's λ for the discriminant function was selected for entry. Wilk's λ shows the proportion of the total variance in the discriminant scores not explained by differences among groups. A small λ value (near 0) indicates that the group's mean discriminant scores differ. A λ of 1 occurs when the mean of the discriminant scores is the same in all groups and there is no between-groups variability; therefore, the larger λ is, the less discriminating power is present.

Classification experiments were computed using SYSTAT 10.2 (SSI, Richmond, CA).

Cross-validation trials

The final reported result in each manuscript consisted in the mean accuracy of multiple cross-validated trials (Figure 20). In each of those trials, the study data is *randomly* split in two groups: the first group is the *training set* and is used to estimate the classification functions; the second group is the *test set*, classified using the functions derived from the training set. This procedure is repeated multiple times and the average proportion of correct classifications of the test set is an empirical measure for the success of the discrimination. The cross-validation experiments were completed using SYSTAT 10.2 (SSI, Richmond, CA).

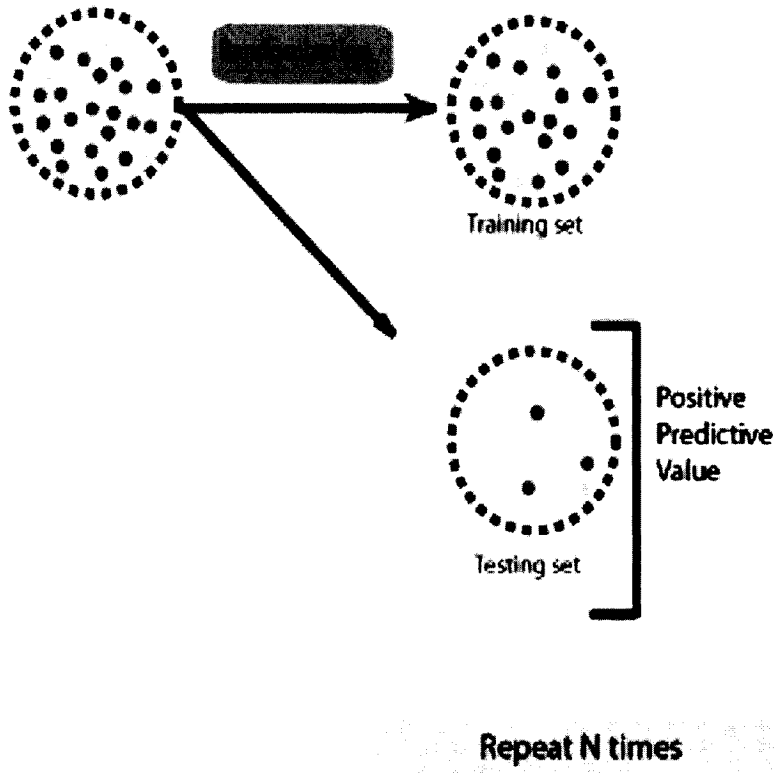

Figure 20 – Cross-validation trials

Cross validation trials proceed as follows. The study population is randomly split into a training and a test set, with the linear discriminant functions estimated on the training set. The test set is used to evaluate how good the classification is, and this process is repeated multiple times to get a good estimate of the overall accuracy for classification.

Results will be given in terms of their sensitivity (the proportion of individuals with a true positive result), their specificity (the proportion of individuals with a true negative result) and their accuracy (the proportion of all subjects correctly classified), as described in Table 1 below.

		Test result	
		Group 1	Group 2
Diagnosis	Group 1	TP	FN
	Group 2	FP	TN

Sensitivity

$$SN = \frac{TP}{TP + FN}$$

Specificity

$$SP = \frac{TN}{TN + FP}$$

Accuracy

$$A = \frac{TP + TN}{TP + FP + TN + FN}$$

Table 1 - Sensitivity, specificity, and accuracy measurements
(TP = true positive, TN = true negative, FP = false positive, FN = false negative)

In the following chapters, these methods will be used for classification and prediction purposes in two different pathologies. In Chapter 4, the ABC technique will be used to lateralize the seizure focus on T1w images from TLE patients with and without hippocampal atrophy; in Chapter 6, the same methodology will be used to differentiate AD and MCI patients from normal aging individuals, and in Chapter 7, to predict cognitive decline to AD in a cohort of MCI patients. Finally, the VBC methodology will be used in Chapter 5 for the prediction of surgical outcome in a cohort of TLE patients having undergone selective amygdalo-hippocampectomy.

Chapter 4

Lateralization of seizure focus in TLE

Foreword

We will recall that the primary goal of pre-operative evaluation in intractable TLE is to confirm which of the two amygdalo-hippocampal complexes is most affected, and therefore becomes the resection target, a process commonly referred to as lateralization of seizure focus. Further, we will recollect that nearly 80% of TLE cases exhibit HC atrophy (HA) predominantly in the left (L) or right (R) MTL, strongly correlated with an ipsilateral seizure focus. However, around 5% of intractable TLE patients demonstrate bilateral atrophy without significant differences between hemispheres, and the remaining 15% of cases have non-atrophic HC (hippocampal normal volume, or HNV). In these latter two cases volume-based lateralization of seizure focus is at best equal to chance.

Our primary goal in the manuscript presented in this chapter was to verify that there existed appearance differences (intensity and shape) related to seizure focus localization. We purposefully moved away from a structure-centric approach to selecting a large, non-specific VOI, centered on the MTL, and inclusive of the HC as well as the AG, the parahippocamapl (PHC) gyrus, other nearby gyri and ventricular areas. This choice was motivated by volumetry results obtained in the same cohort of patients by one of our co-author (Bernasconi, Bernasconi et al. 2003), as well as results from our voxel-based morphometry (VBM) study, also within the same cohort (Bernasconi, Duchesne et al. 2004) (see Figure 11), all pointing to the existence of extra-hippocampal differences in patients with left or right seizure focus when compared to normal controls.

At the same time, we wished to verify that these appearance differences could be used for the purpose of lateralization. We hypothesized that they could be modeled linearly with principal components, or at least sufficiently well estimated to allow for the correct classification of patients with or without HA into the proper lateralized group. The standard of comparison for this task was clinical evaluation, including extensive video-EEG recordings, and in some cases backed-up by surgically implanted EEG (sEEG).

Our results indicate that our assumptions were correct. We found differences in appearance parameters that were sufficient to achieve 100% accurate lateralization of seizure focus in patients with or without HA (HA vs HNV, LHA vs RHA, LNV vs

RNV); further, we found that the system was 96% accurate at lateralizing seizure focus even when given no *a priori* information about structure atrophy (L vs R). This is to be compared with HC volumetry, which achieves 81% accuracy at the same task in our cohort.

Early forms of this work were printed as conference proceedings (Duchesne, Bernasconi et al. 2002; Duchesne, Bernasconi et al. 2003), before being published in the journal *NeuroImage*.

The reader will notice that the Methods section of this manuscript, as well as the Discussion – Methodological Considerations, have already been addressed in the previous Methods chapter of this thesis; figures and figure captions have also been seen earlier. The reader may therefore choose to skip the Methods section of this chapter without loss of information.

For the sake of brevity, keywords have not been included, and a list of common abbreviations can be found at the end of this thesis.

MR-based neurological disease classification methodology: application to lateralization of seizure focus in temporal lobe epilepsy

Simon Duchesne, Neda Bernasconi, Andrea Bernasconi and D. Louis Collins

McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University

NeuroImage, 29(2):557-566, 2006

Abstract

Classification of neurological diseases tends to concentrate on specific structures such as the hippocampus (HC). The hypothesis for the novel classification methodology presented in this work is that pathologies will impact large tissue areas with detectable variations of T1-weighted MR signal intensity and registration metrics. The technique is applied to the lateralization of seizure focus in 127 patients with intractable temporal lobe epilepsy (TLE), in which the site of seizure onset was determined by comprehensive evaluation (69 with left MTL seizure focus (seizure focus) (group "L") and 58 with right seizure focus (group "R")). The method analyses large, non-specific Volumes of Interest (VOI) centered on the left and right medial temporal lobes (MTL) (55 x 82 x 80 voxels) of pre-processed scans in stereotaxic space. The VOIs are then linearly and nonlinearly registered to a reference target image. Principal Components Analyses of (i) the normalized intensity and (ii) the trace, a measure of local volume change, are used to generate a multidimensional reference space from a set of 152 neurologically healthy subjects. Data from TLE patients are projected in this space, and leave-one-out, forward stepwise linear discriminant analysis of the eigencoordinate distributions is used to classify TLE patients. Following MRI volumetric analysis, 80 patients had HC atrophy (group "HA") ipsilateral to the seizure focus (42 with left seizure focus or "LHA", and 38 with right or "RHA"), and the remaining 47 had normal HC

volumes (group "HNV") (27 with left seizure focus or "LNV", and 20 with right seizure focus or "RNV"). The method was 100% accurate at separating "HA" *vs* "HNV", "LHA" *vs* "RHA", and "LNV" *vs* "RNV". It was also 96% accurate at separating "L" *vs* "R".

Our results indicate that MR data projected in multidimensional feature domains can lateralize seizure focus in epilepsy patients with a high accuracy, irrespective of HC volumes. This single scan, practical and objective method holds promise for the pre-surgical evaluation of TLE patients.

Introduction

Many neurological diseases exhibit pathologically-specific discriminatory information in the form of local intensity variations and shape changes when observed on magnetic resonance images (MRI). This is the case for schizophrenia, Alzheimer's dementia (AD) or epilepsy with a seizure focus that can be lateralized to one of the medial temporal lobes (MTL). MRI-based computer-aided diagnosis approaches have usually focused on those attributes. An example is the T1-weighted (T1w) signal intensity, that may serve as an indicator of disease progression since subtle changes may indicate an underlying pathological process before structure integrity is lost. This would be the case for tissue atrophy in AD or temporal lobe epilepsy (TLE) for example, where authors have used the signal intensity directly (Webb, Guimond et al. 1999; Duchesne, Bernasconi et al. 2002; Thomaz, Boardman

et al. 2004), while others performed texture analysis to show evidence of alterations that corroborate the hypothesis of tissue damage in conjunction with intensity information (Liu, Teverovskiy et al. 2004). Other techniques employ registration, a process where individual subject images are aligned into a reference space, to make spatial comparisons between cohorts at the voxel level, such as in voxel-based morphometry (Ashburner and Friston 2000), deformation-based morphometry (Chung, Worsley et al. 2001; Chung, Worsley et al. 2003; Shen and Davatzikos 2003; Lao, Shen et al. 2004), or surface analysis of individual structures (Csernansky, Wang et al. 2000; Wang, Joshi et al. 2001; Csernansky, Schindler et al. 2004). Appearance-based approaches combine intensity and local shape information, but their application has been limited to segmentation (Cootes, Edwards et al. 2001; Duchesne, Pruessner et al. 2002).

The most common neuroimaging technique for computer-aided diagnosis in pathologies such as AD or TLE has consisted in the study of individual structures, such as the hippocampus (HC), mostly via volumetry (Jack, Theodore et al. 1995) and now with newer methods based on shape characteristics, such as medial surfaces (Joshi, Pizer et al. 2002; Styner, Gerig et al. 2003; Bouix, Pruessner et al. 2005). While undeniably useful for understanding disease-driven change in the chosen structure(s), these analyses suffer from significant drawbacks: they rely heavily on segmentation, a process with its own limitations, do not capture interrelations between neighboring tissues, critical in many pathologies, and usually require expert

knowledge, making them unsuitable for large scale studies. Finally, apart from appearance-based approaches, most of the techniques reviewed use either intensity or registration information, but rarely both.

Lateralization of seizure focus in TLE

We present intractable TLE as a pathology of interest whose characteristics make it an ideal test case for our proposed methodology. Microscopically, the most commonly described pathological findings in TLE are neuronal loss and gliosis of the HC and the parahippocampal region (Falconer, Serafetinides et al. 1964). As discussed, MR-based volumetry is a primary indicator of structure integrity (Bernasconi, Bernasconi et al. 2003), and hippocampal neuronal loss found in surgical specimens obtained from patients with pharmacologically intractable TLE has been shown to correlate with hippocampal atrophy on volumetric MRI (Cascino, Jack et al. 1991). Studies in TLE have also shown volume reduction ipsilateral to the side of the seizure focus in the parahippocampal region (Bernasconi, Bernasconi et al. 2003). The advent of high resolution MRI has, therefore, had a major impact on the presurgical evaluation of patients with refractory epilepsy. Nearly 80% of TLE cases exhibit lateralized HC atrophy (see Figure 21), strongly correlated with an ipsilateral seizure focus (Bernasconi, Bernasconi et al. 2003). HC volumetry is therefore used in many centers for the pre-surgical evaluation however, it is a time-consuming, manual procedure requiring expert intervention, and subject to the drawbacks outlined above. Further, around 5% of intractable TLE patients

demonstrate bilateral atrophy without significant differences between hemispheres, and the remaining 15% of cases have nonatrophic HC, making volume-based lateralization impossible. Finally, neuropathological studies have shown that more than one brain structure are affected in TLE, while voxel-based morphometry studies of MR images (Woermann, Free et al. 1999; Keller, Mackay et al. 2002; Bernasconi, Duchesne et al. 2004) have reported significant areas of grey matter and white matter concentration decrease throughout the medial temporal lobe as well as in extra-temporal lobe brain areas.

Hypothesis and objective of research

Our general hypothesis is that in the case of many neurological diseases, microscopic changes will impact the T1w MRI signal intensity sufficiently to be detected, while macroscopic alterations in structure shape will be noticed via registration-based metrics. Consequently, we are proposing in this article a classification methodology that moves away from the single structure-based paradigm by using (i) a large, non-specific Volume of Interest (VOI) containing sufficient discriminatory information, and (ii) by combining intensity and registration-based shape features. A preliminary implementation has been described in Duchesne *et al.* (Duchesne, Bernasconi et al. 2003). From the preceding literature on TLE, we concluded that a single-structure approach for the lateralization of seizure focus was not adequate and that, in order to be succesful, classification should be based on the information contained in a VOI centered on the MTL, incorporating the HC along with other limbic structures. In

this paper, lateralization of seizure focus in TLE is used as a test case for our proposed classification methodology.

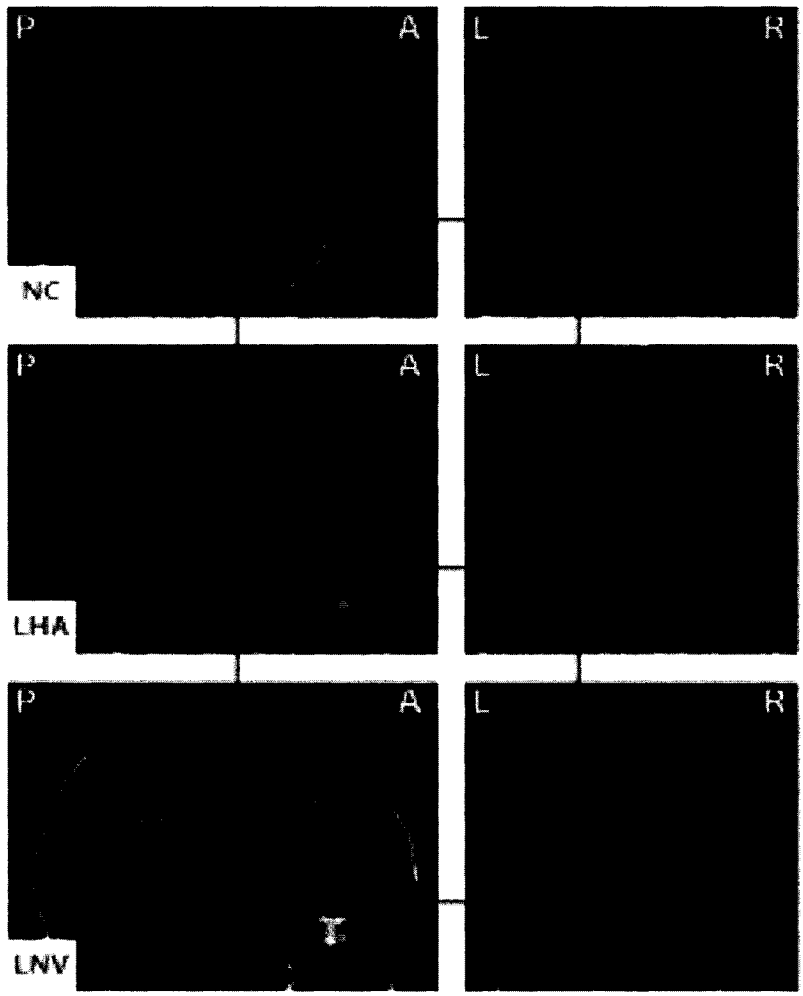

Figure 21 – T1 images in TLE

(Top) Sagittal (x = 62) and coronal (y = 117) T1-weighted (T1w) magnetic resonance images (MRI) of normal control subject #18. (Middle) Hippocampal neuronal loss found in surgical specimens obtained from patients with pharmacologically intractable temporal lobe epilepsy (TLE) has been shown to correlate with hippocampal atrophy (HA) on volumetric MRI (Cascino et al., 1991). For example, left HA is readily observed on these T1w MR images (same orientation as above) of patient #18 as an enlargement of the lateral ventricles, along with increased atrophy of the hippocampal head (red lines intersection). (Bottom) On the other hand, around 15% of intractable TLE patients exhibit normal hippocampal volumes, therefore making lateralization impossible in those cases on the basis of HC volumetry alone. This is exemplified in these T1w MR images (same orientation as above) of HNV patient #1.

The objectives of this work are therefore to describe the methodology and demonstrate its usefulness by proceeding with the lateralization of seizure focus in TLE patients with and without hippocampal atrophy, based on their pre-operative T1w MRIs.

Methods

Our method can be summarized as follows (see Figure 22). First, from processed data **(A)** we generate a non-pathological eigenspace **(B)** from a large training group of young, neurologically healthy training subjects (N = 152). This multidimensional eigenspace is created by uniting results from four distinct Principal Component analyses of (i) linearly registered intensity images of the left and right VOIs from T1wMRI of these training subjects **(C)**; and (ii) an approximation of the determinant of the Jacobian matrix of the deformation field for the same VOIs **(D)**, where the deformation fields are obtained by non-linear registration to a common reference image. Secondly, VOIs from our patients are projected in the multidimensional eigenspace created **(E)**. The normality of the distribution of the projected eigencoordinates is first assessed using quantile plots and Shapiro-Wilke statistics. The last step consists in linear discriminant analyses (LDA) in a leave-one-out, forward stepwise approach, using eigencoordinates in the non-pathological eigenspaces to classify the patients. While the non-pathological eigenspace may not be optimal to represent the patient population, it forms the basis for a comparative

evaluation of the different groups that leads to their classification. The methodological details are elaborated in the following sections.

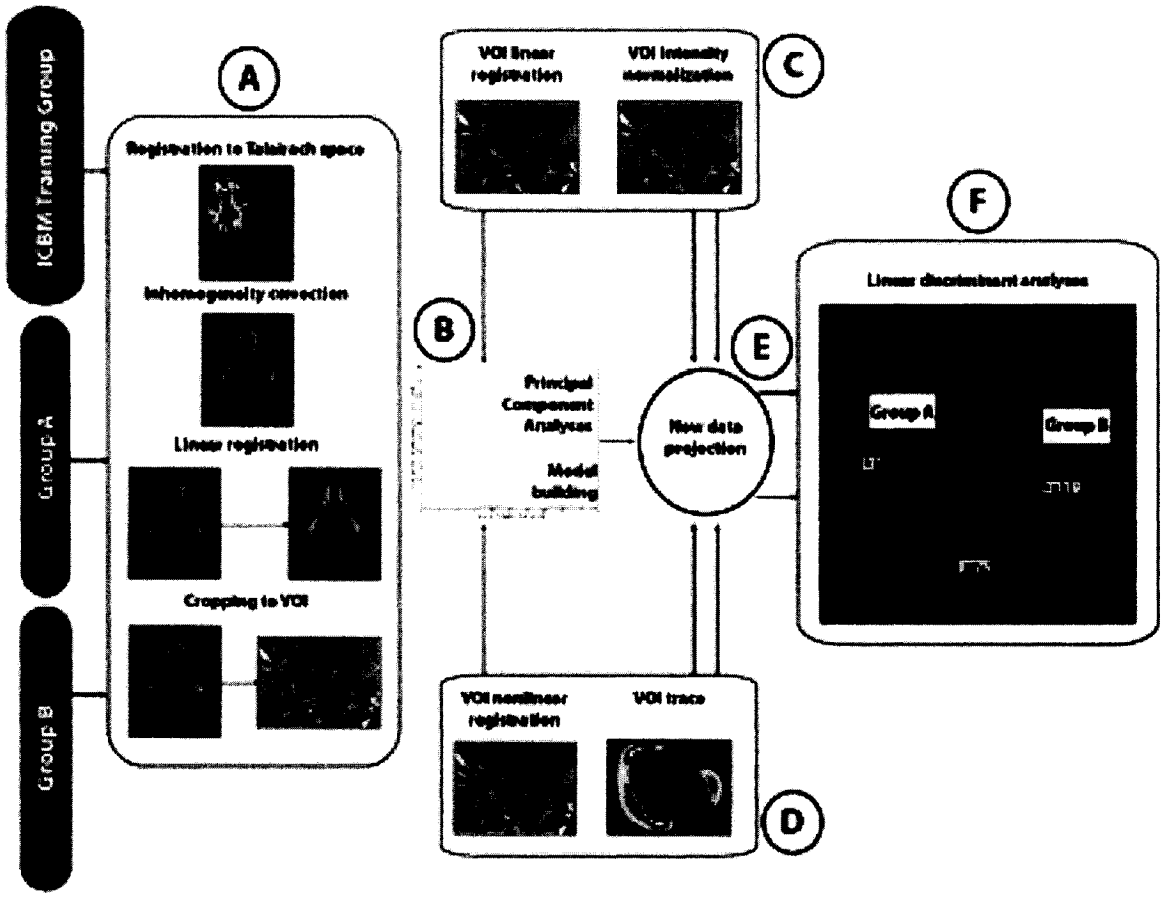

Figure 22 – Classification methodology

Our hypothesis for the proposed classification system is that pathologically-induced microscopic changes will impact the T1w MR signal intensity, while macroscopic alterations in structure shape can be detected via image registration. In this methodology, two volumes of interest (VOI), centered on the left and right medial temporal lobes, are extracted from pre-processed data (A). At first, a large group of young, healthy control subjects' images (N = 152) is used to generate a normal, non-pathological eigenspace (B) that will serve as a basis for comparison of the TLE study subjects. This multidimensional reference eigenspace is created by uniting results from four distinct Principal Component analyses of (i) linearly registered intensity images of the VOIs from T1w MRI (C); and (ii) an approximation of the determinant of the Jacobian matrix of the deformation field within those VOIs (D). The data from the study subjects (Groups A and B) are projected in the reference space (E). Classification is based on leave-one-out, forward stepwise linear discriminant analyses of the eigencoordinate distributions (F).

Subjects

The Ethics Committee of the Montreal Neurological Institute and Hospital approved the study and informed consent was obtained from all participants. A total of 279 subjects were included in this study. The training group consisted in 152 young, neurologically healthy volunteers from the International Consortium for Brain Mapping project (Mazziotta, Toga et al. 1995), whose scans were used to create the non-pathological model. The test population consisted in 127 patients with pharmacologically intractable TLE. Seizure type and the site of seizure onset were determined by a comprehensive evaluation including detailed clinical history, neurological examination, review of medical records and neuropsychological evaluation. The seizure focus was determined by predominantly ipsilateral interictal epileptic abnormalities (70% cutoff) and unequivocal unilateral seizure onset recorded during prolonged video-EEG monitoring in all patients. Based on these criteria, TLE patients were divided into those with a left-sided (group "L", n= 69 [54%]) or a right-sided (group "R", n=58 [46%]) seizure focus.

MRI acquisition and initial processing

MRI data for our 279 subjects was collected with a T1-weighted MRI protocol on a 1.5 T scanner (Philips Gyroscan, Best, Netherlands) using a fast gradient echo sequence (TR = 18*ms*, TE = 10*ms*, 1 NEX pulse sequence, flip angle=30°, matrix size=256 X 256, FOV = 256*mm*, slice thickness=1*mm*). Following acquisition, intensity inhomogeneities due to scanner variations were corrected (Sled, Zijdenbos

et al. 1998). A 9-degrees of freedom (DF) linear registration was then used for global alignment (Collins, Neelin et al. 1994) into a standard reference space. The data was finally resampled onto a $1mm$ isotropic grid (Collins, Neelin et al. 1994). The reference image used for the linear registration and resampling was the ICBM 152 T1w target, a voxel-by-voxel average of the 152 normal subjects previously registered in the Talairach-like stereotaxic space in the context of the ICBM project (Mazziotta, Toga et al. 1995).

Input data for multi-dimensional space creation

Two VOIs were selected for this study, centered on the left and right medial temporal lobe, using Talairach coordinates (start coordinates x = $[-53,+2]$ for the left and right side respectively, y = -53 and z = -52). Each VOI measured n = 55 X 82 X 80 = 360800 voxels. The VOI was selected so that its extent captured the hippocampus and neighboring MTL structures, such as the parahippocampal gyrus, irrespective of normal inter- and intra-individual variability (see Figure 23).

Two image features at each voxel location were retained for modeling. The first feature used in our classifier is *grey level intensity*, denoted by script $\boldsymbol{g}$. Scanner output volumes are preprocessed in the manner described in the preceding section, which includes intensity inhomogeneity correction and global linear registration into Talairach space (DF = 9). The VOIs are then extracted based on their Talairach coordinates. To further reduce positional variations, which would propagate as unwanted noise in the morphometric PCA modelling, a second local affine

transformation (DF = 12) is applied to co-register the individual subject's VOI with the corresponding VOI on the ICBM 152 target. Finally, VOI intensity mean is also linearly scaled to the reference target.

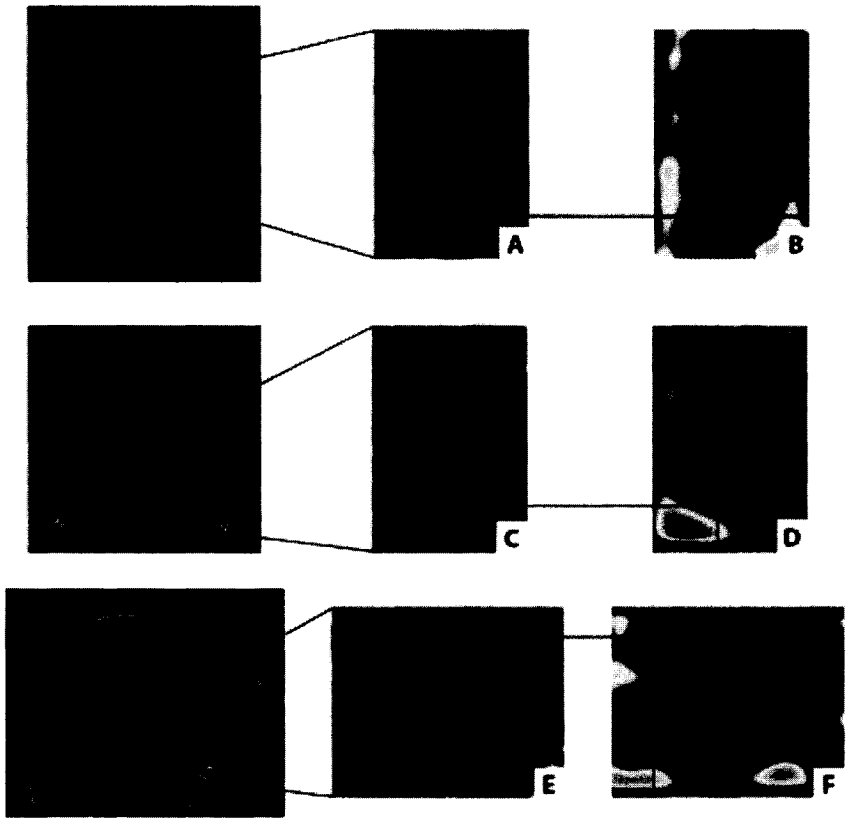

Figure 23 – Intensity and shape features within the volume of interest

(Left column) Transverse, coronal and sagittal whole-brain views for TLE subject #28 with LHA. (Middle column) Identical views through the Volume of Interest (VOI), as defined in stereotaxic space on the left medial temporal lobe (red boxes on whole-brain views). (Right column) Identical views trough the trace VOI for the same subject. Green voxels do not move. From green to white (maximum) via yellows and reds indicate increase or expansion. From green to black (minimum), via blues and purples indicate decrease or contraction. The direction of movement is defined as the deformation which the subject's VOI seen in (A), (C) and (E) must accomplish in order to align with the corresponding VOI extracted from the reference volume. Contraction represents atrophy in the case of tissue. Notice in the transverse view in (A) and (B) the contraction of the lateral ventricular space, necessary to correct its enlargement, also noticeable in (E) and (F), while in (C) and (D), the coronal view displays the superiocentral shift due to hippocampal atrophy.

The input to the intensity model therefore consists in the rasterized intensity data from the VOIs processed in this fashion. The second feature is the *trace*, denoted by script ***t***. The trace is the first-order approximation of the determinant of the Jacobian matrix of a non-linear registration-derived deformation field, which maps a subject to our reference ICBM target, and thus forms an estimate of local volume change.

Nonlinear registration attempts to match image features from a source volume to those of the reference image at a local level, typically in a hierarchical fashion, with the aim of reducing a specific cost function. Whereas many nonlinear registration processes exist, the one chosen for this study was ANIMAL, developed by (Collins and Evans 1997). This algorithm attempts to match image grey-level intensity features at a local level (voxel) in successive blurring steps, by maximizing the cross-correlation function of voxel intensities between the source and reference images. The result is a dense deformation field capturing the displacements required to align the subject VOI with that of the reference image. In our application, we estimate the nonlinear transformation between the VOI of an individual subject to the ICBM 152 target VOI. Note that the trace input VOI for our subjects is the same as the one used for the intensity modeling.

Our implementation of trace calculation was discussed in (Janke, de Zubicaray et al. 2001) and follows the notation developed by (Chung, Worsley et al. 2001). If U is the

displacement field which matches homologous points between two images, then the local volume change of the deformation in the neighborhood of any given voxel is determined by the Jacobian J (Chung, Worsley et al. 2001), which is defined as

$$J(x,t) = \det(\mathrm{I} + \frac{\partial U}{\partial x})$$

where $\mathbf{I}$ denotes the identity matrix and $\frac{\partial U}{\partial x}$ is the 3 X 3 displacement gradient matrix of U. For relatively small displacements,

$$J \approx 1 + tr(\nabla \mathrm{U})$$

as suggested by (Chung, Worsley et al. 2001). Hence the trace $tr(\nabla \mathrm{U})$ represents a crude yet indicative measure of local volume change. This is well suited to brain applications as the level of atrophy present between controls and patients is small yet significant and should therefore be apparent in the results of a local non-linear registration process. Examples are shown in Figure 23. The input to the trace PCA therefore consists in rasterized vectors of calculated trace volumes.

Multidimensional spaces and classification

Principal Components Analysis (PCA) is used to reduce the dimensionality of the input data (intensity **g**, trace **t**) and generate linear variation models based on the N=152 datasets from our ICBM normal control subjects. We use here the notation identical to (Cootes, Edwards et al. 2001):

$$x = \bar{x} + P_x B_x$$

which allows any vector instance x, be it trace or intensity vectors from the left or

right VOI, to be approximated by $\bar{x}$, the mean normalised input vector; P_x, the set of orthogonal modes of variation; and B_x, the set of parameters. The resulting four PC models were each $p=N-1$ (or 151-dimensional).

The ensemble of Principal Components from the left and right VOI grey-level intensity models defines an Allowable Grey Domain **G** as the space of all possible elements expressed by the intensity eigenvectors λ_G. Likewise, an Allowable Trace Domain **T** is created as the space of all possible elements expressed by trace eigenvectors λ_T for the left or right VOIs. It can be shown that the eigenvectors of the covariance matrix corresponding to the largest eigenvalues describe the longest axes of the ellipsoid, and thus the most significant modes of variation in the variables used to derive the covariance matrix (Cootes, Edwards et al. 2001). Most of the variation can usually be explained by a smaller number of modes, l, where $l << n$ and $l < p$. This means that if the $2n$ dimensional ellipsoid is approximated by a l dimensional ellipsoid, then the original ellipsoid has a relatively small width along axes with indexes $l + 1$ and greater (Cootes, Edwards et al. 2001).

The total variance of all the variables is equal to:

$$\lambda = \sum_{k=1}^{2n} \lambda_k$$

whereas for l eigenvectors, explaining a sufficiently large proportion of λ, the sum of their variances, or how much these principal directions contribute in the description of the total variance of the system, is calculated with the ratio of relative importance of the eigenvalue λ_k associated with the eigenvector k:

$$r_k = \frac{\lambda_k}{\sum_{j=1}^{p} \lambda_j}$$

We proceeded in selecting eigenvectors for each of our four models (left/right intensity/trace VOIs), corresponding to a per-model variance ratio of $r = 0.997$ for our normative training group. Again, our goal was not to find the optimal space to represent the study subjects, but rather to have a common basis in which to compare study populations. This is exemplified in Figure 24, where a cartoon-like representation of our training set is drawn, aligned and centered with the first two principal components since they were extracted from that data. Our goal is to use this space for the separation of pathology #1 subjects into two groups, 1a and 1b. While the space may not optimal for pathology 1, it forms a normative basis built from young, neurologically healthy individuals that can also be used for other pathologies, such as pathology #2 that may or may not overlap with pathology #1. The reader should also note that in this study we will not discuss the case of classifying normal control subjects from TLE patients on the basis of their MRI characteristics, as it offers no clinical value.

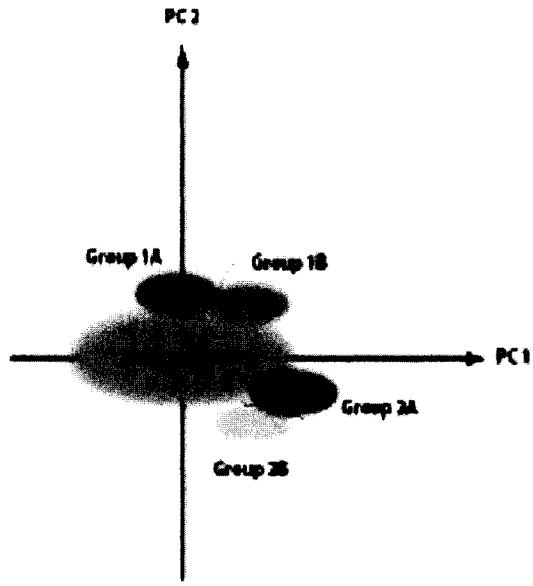

Figure 24 – Cartoon-like representation of classification of patient groups

The reference set is represented by the green ellipse, aligned and centered with the first two principal components since the eigenspace originates from these data. Our goal however is to use the same space for the separation of many pathologies. For example, pathology #1, while not optimally represented by the eigenvectors of the reference space, can still be separated in two groups, 1A and 1B. The same can be said of a second pathology, which may or may not overlap with the previous one. Using a normative space built from young, neurologically healthy individuals should give the ability to perform classification studies in multiple pathologies.

Rasterized vectors of the processed VOI intensity and trace data for each study subject are projected into Domains **G** and **T**, and thus form the eigencoordinate vectors. While a number of possible features can be calculated on the distribution of the projected data, our classifier is based on the position along the PC axes. The distribution of eigencoordinates along any principal component for a given population should be normally distributed and this normality is assessed via quantile plots and Shapiro-Wilke statistics.

The classification is based on leave-one-out, forward stepwise linear discriminant analyses of eigencoordinates along the significant eigenvectors (SYSTAT 10.2,

Georgia, PA; P−to−enter < 0.0001). Results for classification are then compared to the pre-operative clinical evaluations in terms of accuracy.

Results

As mentioned earlier, the patient population consisted in 127 individuals with pharmacologically intractable epilepsy, 69 (54%) with a left MTL seizure focus (group "L") and 58 (46%) in the right (group "R"). MRI volumetric analyses were performed on T1w MRI for all patients using a previously published protocol (rater: N.B.) (Bernasconi, Bernasconi et al. 2003). The HC was determined atrophic if its volume fell below two standard deviations from the mean HC volume of a matched group of control subjects. Eighty (80) patients (63%) had hippocampal atrophy (group "HA") ipsilateral to the seizure focus (42 with left seizure focus and HA or group "LHA", and 38 with right seizure focus and HA or group "RHA"). The remaining 47 patients (37%) had normal hippocampal volumes (group "HNV") as seen on MRI (27 with left seizure focus and HNV or group "LNV", and 20 with right seizure focus and HNV or group "RNV"). Table 2 contains demographic information on the different groups of TLE patients as well as the training set.
By definition, HC volumetry based on MR images was therefore 100% accurate in this cohort to lateralize seizure focus in those patients with HA, as the seizure focus was always ipsilateral, and 50% accurate (as good as chance) for those cases with HNV. The overall accuracy of HC volumetry for the group of patients in this study

was (103/127) = 81%.

	N	Mean Age	Males	Females
LHA	42	35	22	20
RHA	38	32	17	21
LNV	27	39	11	16
RNV	20	32	7	13
Training set	152	25	87	65

Table 2 - Description of patient and training population

Our overall model, formed by the union of the intensity and trace models for the left and right VOIs, was built using our training set of 152 healthy control subjects from the ICBM database. For each of these models all eigenvectors accounting for 99.7% of the variance of the neurologically healthy data set were chosen, for a total of 538 eigenvectors. Normality of the eigencoordinates from the projected study data was confirmed for each PC using quantile plots and Shapiro-Wilke test ($P > 0.05$).

In order to assess the classification accuracy of our automated methodology, we performed four experiments as follows:

- **Experiment 1** - Classification (HA *vs* HNV): determine the classifier's ability at separating the HA *vs* HNV groups based solely on MR imaging data;
- **Experiment 2** - HA Lateralization (LHA *vs* RHA): determine the classifier's ability at lateralizing the seizure focus within the hippocampal atrophy group;
- **Experiment 3** - HNV Lateralization (LNV *vs* RNV): determine the classifier's ability at lateralizing the seizure focus within the normal hippocampal volume group; and

• **Experiment 4** - Lateralization (L *vs* R): determine the classifier's ability at lateralizing the seizure focus irrespective of hippocampal volume.

For each experiment a response curve was built of model accuracy *vs* total number of eigenvectors included in the classification process (see Figure 25). The best classification results following leave-one-out, forward stepwise linear discriminant analysis (P−to−enter < 0.0001) on the total number of eigenvectors included in the classification are reported in Table 3, with the resulting number of eigenvectors retained by the classifier to reach the best classification results and the corresponding Wilk's λ statistics. Highlights of Figure 25 and Table 3 are 100% classification accuracy for **Experiments 1, 2** and **3**, and 96% accuracy for **Experiment 4**.

Expt	Group A	Group B	Total # eigen.	# Classif. Eigen.	Wilk's	DF	Accuracy
1	HA (80)	HNV (47)	10	4	0.029	4	100%
2	LHA (42)	RHA (38)	538	13	0.083	13	100%
3	LNV (27)	RNV (20)	8	5	0.078	5	100%
4	L (69)	R (58)	240	12	0.242	12	96%

Table 3 - Classification results by experiments

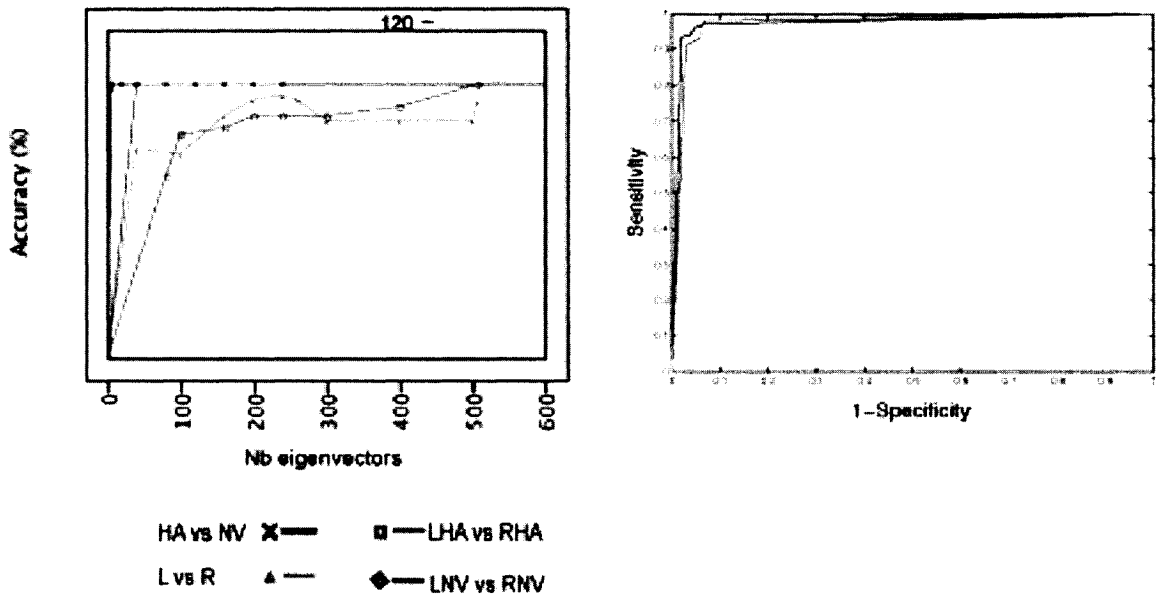

Figure 25 – Experimental results

(Left) Model response curves for the 4 classification experiments. The classifier was presented between 4 and 538 eigenvectors, for each classification experiment, with the accuracy generally increasing as variables were added. The optimal number of eigenvectors presented to the classifier, along with the number of eigenvectors retained by the classifier ($P < 0.0001$), are reported in Table 3. It is important to note that the order with which the eigenvectors were presented to the classifier is not optimal for the task of separating TLE patients, but rather is optimal for the representation of the reference set data (decreasing eigenvalue). This explains why some of the curves are not monotonically increasing (e.g. L vs. R). (Right) Receiver Operating Characteristic (ROC) curve for the Left vs. Right seizure focus experiment (Experiment 4). Forward, stepwise analysis identified 12 significantly discriminant eigenvectors as shown in Table 3. We plotted sensitivity vs. 1—specificity for Left seizure focus discrimination (blue curve) and Right seizure focus discrimination (green curve) for the 12 variable model.

Discussion

For many neurological diseases with specific intensity and shape changes, as seen on MRI, the classical approach for computer-aided diagnosis remains volumetry of single structures, such as the HC, or parahippocampal gyrus (Bernasconi, Bernasconi et al. 2003). MR-based HC volumetry, while greatly beneficial, is not

sufficiently specific nor sensitive, especially in cases where the HC volume falls within normal ranges.

We are proposing a classification methodology that uses intensity and registration-based shape features from a large, non-specific VOI. Our results indicate that these features are sufficient to adequately discriminate between different groups of subjects. In our study population, the accuracy of classification of patients into HA or HNV, the lateralization of the HA group into LHA and RHA, and of HNV into LNV and RNV, was 100%. We were also able to lateralize the seizure focus irrespective of HC volume with 96% accuracy.

Methodological considerations

The methodology is formulated to be simple and flexible, allowing a number of possible features to be used without extensive reengineering. We achieved substantial data reduction, starting from a model of 152 VOIs (152 X 2VOIs X 360800 voxels = 1.1 X 10^7 points in MRI space) down to 538 eigencoordinate points in a predefined PC space. As the training set is composed of ICBM subjects that are separate from the test set subjects, there is no issue of overdetermination. Further, by using leave-one-out or jacknife linear discriminant analyses, we ensured that there was no overlearning in the classification stage.

From the response curves of Figure 25 for each classification experiment an initial conclusion can be made, in that the optimal number of eigenvectors varies depending on the classification problem. This result is not surprising and means that

distinct covarying intensity and/or shape features characterize the different two groups under study. Assessment of the variations in intensity and shape embedded in the significant eigenvectors is beyond the scope of this paper, and will be performed, along with their clinical correlate, in a separate communication.

Using intensity features from MRI raises the question of calibration and normalization. Absolute intensities are rarely used in MRI, since they vary with machine calibration, shimming, and patient-induced variations. We have tried to limit those variations by (1) using the same scanner for every individual in the study; (2) ensuring that the same quality assurance procedures were followed for each acquisition; and (3) scanning the patients in random group order. Our approach has been to create a model eigenspace based on subjects from the ICBM database, in which we projected our cohorts of TLE patients. While such a space built from young, neurologically healthy individuals may not be optimal to represent the test population, it should be noted that our primary goal was to find an independent basis for a comparative evaluation of our study cohort and its subsequent classification, but not the best mathematical representation. This is demonstrated in Figure 25, especially for **Experiment 4** (L *vs* R). If the space was optimal for representing the TLE patients one would expect a monotonous increase in accuracy as the number of eigenvectors presented to the classifier increased. However, such is not the case since we were presenting to the classifier eigenvectors ordered along their weight for a young, healthy population. This explains why, for example, in the

case **Experiment 4**, the accuracy of the classification does not monotonically increase, since quite obviously some of the eigenvectors that improve the classification are found in later reference space eigenvectors, that do not carry much eigenweight for normal, healthy subjects, but they do carry discriminatory information for TLE patients.

Clinical considerations

The prevalence of HA as reported in the literature for cohorts of TLE patients ranges from 50%(Ng, Tang et al. 2000) through 90%(Cascino 1995), with a "consensus range" emerging between 75%(Spencer, McCarthy et al. 1993) or 76%(Jack, Theodore et al. 1995) to 85%(Cheon, Chang et al. 1998; Bernasconi, Bernasconi et al. 2003). HA prevalence in our cohort fell lower than this accepted range at 63%, but this complicates, rather than simplifies, the task for the classification system, as there are less HA examples to learn from in the leave-one-out analyses. Our first experiment was aimed at determining if HC atrophy was detectable with our methodology. By definition, HC volumetry determines atrophy with 100% accuracy. Our results show that we are able to differentiate between HA and HNV with the same accuracy. From a clinical point of view however, the advantage of the automated technique relies in its independence from manual intervention, outside of the one-time definition of the VOI in Talairach-like coordinates. HA is typically associated with ipsilateral seizure and therefore volumetry seems like a good indicator of lateralization. The accuracy of

lateralization in TLE patients with HC atrophy is at a minimum 80% (Bernasconi, Bernasconi et al. 2003) or 90% when including structures other than the HC, such as the amygdala (Cendes, Andermann et al. 1993), entorhinal and parahippocampal formation (Bernasconi, Bernasconi et al. 2003), or combining with T2-weighted MRI hyperintensities in the medial temporal lobes (Kuzniecky, Suggs et al. 1991) or spectroscopic data (Li, Caramanos et al. 2000). Other authors report a maximum of 98% (Spencer, McCarthy et al. 1993) or 100% in (Jack, Theodore et al. 1995) for lateralization via volumetry. In this cohort, since we selected individuals with atrophy ipsilateral to seizure focus, the accuracy of MRI-based volumetry was 100%. Our classification results achieved the same accuracy for lateralization of seizure focus in patients with HA, using 13 eigenvectors.

Also by definition, the lateralization of seizure focus in patients with normal hippocampal volume cannot be done on the basis of HC volumetry and therefore is at best equal to chance (50%). In contrast, our methodology surpasses this result by achieving 100% lateralization accuracy with 5 eigenvectors.

Finally, we attempted to lateralize seizure focus from the onset, that is without *a priori* separation of the groups based on hippocampal volume. The accuracy of HC volumetry in this instance can be derived by combining the consensus ranges of prevalence for HA, along with the lateralization accuracy of HC volumetry in the presence of HA or HNV, that is [(75 − 85%) X (85% − 98%)] + [(50%) X (25% − 15%)]. Therefore, lateralization of seizure focus with volumetry alone should range

between 76% and 91%, and as mentioned in this cohort is 81%. Our results show an ability to lateralize seizure focus with 96% accuracy with 12 eigenvectors, but, contrary to volumetry, without the need for additional measurements or manual segmentation. The non-invasiveness of MR-based techniques makes them preferrable alternatives for the presurgical evaluation of patients with TLE if proven to be as accurate. We believe that the proposed methodology for automated analysis of T1w MRIs has reached the level at which it needs to be evaluated further in prospective studies to determine its possible inclusion in routine clinical use.

We chose intractable temporal lobe epilepsy (TLE) as a pathology of interest since its characteristics made it an ideal test case for our proposed methodology. Clearly, other neurological diseases that exhibit related morphological changes detectable via T1w MRI intensity and shape may benefit from this approach. As an example, preliminary results for this methodology in the study of Alzheimer's dementia and Mild Cognitive Impairment are discussed in a separate communication (Duchesne, Pruessner et al. 2005).

Conclusion

The underlying assumption for this work is that there exists sufficient information in VOIs extracted from T1w MRI global volumes to be used for the correct classification of our study subjects. Our results demonstrate that multidimensional spaces can be created from a PCA of extracted features of interest from training set

images, and that such a space forms an adequate basis for the discrimination between subject projections. Our results further indicate that LDA can classify groups using the selected features with a high degree of success. The proposed methodology does not rely on segmentation, requires no user input and is data-driven.

The ability to lateralize seizure focus in intractable TLE based solely on pre-operative T1wMR imaging has great potential to improve current diagnostic procedures. It is anticipated that such systems will gain widespread use in the next few years in a wide range of neurological diseases.

Chapter 5

Prediction of outcome for TLE surgery

Foreword

Having demonstrated in the previous chapter that there were appearance differences in intensity and shape related to seizure focus localization, and that these differences could be used for pre-surgical lateralization, we turned our attention to the problem of predicting the long-term, post-surgical clinical status.

The main objective of selective amygdala-hippocampectomy (SAH) performed in intractable TLE patients is one of seizure control; in the majority of cases, but not in all, this goal is achieved. Being able to predict for which patients the current procedure would be most beneficial would significantly alter patient management.

Our first attempt at predicting post-surgical clinical status at one year follow-up was to use the same technique as used in the previous chapter, that is a linear model of appearance principal components extracted from the MTL. The results we obtained

were inconclusive, and we could do no better than chance. This was not entirely surprising, given the literature on the subject (see Chapter 2), which indicates that the extent of disease damage in extra-temporal areas is probably the most determinant predictor of outcome. We therefore decided to turn to the second, VBM-based methodology described in Chapter 3.

Our primary goal in this article was to determine if there existed regions of grey and white matter concentration differences related to surgical outcome outside of the temporal lobe. At the same time, we wished to verify if these differences could be used for the purpose of predicting surgical outcome. Again, we hypothesized that they could be modeled linearly, or at least sufficiently well estimated to allow for the correct classification of patients into positive or negative outcome groups. The standard of comparison for this task was clinical evaluation of seizure control at one-year follow-up.

Our results indicate that our assumptions were correct. We found differences in GM and WM concentrations mostly outside of the hippocampal area, and outside of the MTL, that were indicative of post-surgical outcome. A model based on the age of seizure onset, duration of seizures, ipsi and contra-lateral hippocampal volumes reached a 72.6% accuracy; this was increased to 78.9% when incorporating GM and WM concentration measures within these areas.

An early form of this work was printed as a conference proceeding(Duchesne, Bernasconi et al. 2004), before being submitted for publication in the journal *NeuroSurgery.*

As the methodology has been extensively described in Chapter 3, the reader may choose to skip the Methods section of this chapter without loss of information. For the sake of brevity, keywords have not bee included, and a list of abbreviation can be found at the end of this thesis.

Automated MRI analysis increases prediction accuracy of surgical outcome in selective amygdalo-hippocampectomy patients

Simon Duchesne[1], Neda Bernasconi[1], Andre Olivier[2], Andrea Bernasconi[1], D. Louis Collins[1].

[1] Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, QC, Canada

[2] Department of Neurosurgery, Montreal Neurological Hospital, McGill University, Montreal, QC, Canada

To be submitted to NeuroSurgery, 2006

Abstract

Background – For patients with intractable temporal lobe epilepsy (TLE), resection of the epileptogenic seizure tissue has been shown to be a beneficial approach, with the majority of patients experiencing seizure control (positive outcome) after surgery. Predicting intervention success however remains a difficult task.

Objective – We wished to study pre-operative T1w MRI of TLE patients who had undergone selective amygdala-hippocampectomy (SAH) as part of their surgical treatment, in order to predict clinical status at one year follow-up.

Methods – We performed voxel-based morphometry (VBM) studies of gray and white matter (GM/WM) concentration changes by comparing patients with negative outcome to those with positive outcome to determine regions of concentration differences related to the surgical result. A region of interest was defined as the union of all voxels above a threshold for significance. Its size, the average GM and WM concentration within it, were measured for each patient. A model was built by adding VBM-based measures to non-VBM data (age of onset, duration of epilepsy, ipsi- and contra-lateral hippocampal volumes as measured on MRI). Multiple cross-validation trials were used to assess the predictive ability of the model.

Results – GM concentration changes were primarily located in the left fronto- and inferio-temporal region, while WM concentration changes were found in the left mid-temporal region. Predicting surgical outcome using non-VBM based measures

resulted in an accuracy of 72.6%, while adding VBM-based features improved the this predictive accuracy of future clinical status to 78.9%.

Conclusion – This article demonstrates the idea that pre-operative MRI analysis can increase the accuracy of surgical outcome prediction for an individual patient, and therefore could be useful for patient selection in order to maximize surgical benefit. Of course, conventional or depth EEG remain critical tools for the understanding of the temporo-spatial pattern of the epileptic discharges; it is only proposed here that MRI can play a potent role in pre-surgical planning and evaluation.

Introduction

Predicting surgical outcome in the treatment of medial temporal lobe epilepsy (MTLE) remains an outstanding challenge. Since the advent of high-resolution neuroimaging, there have been numerous attempts at finding structural or functional markers that can be used for that purpose, but no completely reliable indicator has been found to date.

Microscopically, the most commonly described findings in MTLE are neuronal loss and gliosis of the hippocampus (HC) and the parahippocampal (PHC) region (Falconer, Serafetinides et al. 1964). The central goal of surgery in MTLE has consisted in the removal of this affected tissue, in an attempt to remove the site of

seizure generation and thus achieve seizure control. As the understanding of the disease and image-guidance improved (Olivier, Alonso-Vanegas et al. 1996), MTLE surgery has moved from complete to partial lobectomy (anterior temporal lobectomy or ATL), to selective amygdalo-hippocampectomy (SAH) (Olivier 2000). The latter approach consists in the resection of the amygdala (AG) and HC complex, and often parts of the neighboring structures, such as the entorhinal cortex (EC). For most intractable MTLE patients, surgery is the best possible treatment choice (McLachlan 2001). Vickrey *et al.* (Vickrey, Hays et al. 1995) have assessed MTLE surgery success and demonstrated that surgery patients had greater decline in average monthly seizure frequency and took fewer antiepileptic medications, when compared to patients that had not received surgery.

For the purposes of this study, we will define surgical outcome, or postoperative seizure control (Jack, Sharbrough et al. 1992), as either positive (complete remission and disappearance of all seizures, corresponding to Engel Group I) or negative (all levels of seizure recurrences, or Engel Groups II/III/IV)(Engel, Van Ness et al. 1993).

Surgical Outcome Prediction in MTLE

While the majority of patients undergoing surgery have positive outcome, there remains a significant proportion (20-30%) where the procedure will not completely eliminate seizures (Antel, Li et al. 2002). It is difficult at present to determine *a priori* if the procedure will be successful. MTLE surgical outcome prediction has

been examined mainly in subjects undergoing ATL, and remains marginal when using neuropsychological testing (Sawrie, Martin et al. 1998), shows correlation to interictal EEG (Bautista, Cobbs et al. 1999), and appears successful (>90% accuracy) when using PET (Dupont, Semah et al. 2000) (Choi, Kim et al. 2003) or pathological assessment of the resected tissue (Berg, Walczak et al. 1998).

Magnetic Resonance Imaging (MRI) has seen its use increase in the pre-surgical evaluation of MTLE patients. HC neuronal loss found in surgical specimens obtained from patients with pharmacologically intractable TLE has been shown to correlate with HC atrophy on volumetric MRI (Cascino, Jack et al. 1991). MRI studies in MTLE have also shown volume reduction ipsilateral to the side of the seizure focus in the PHC region (Bernasconi, Bernasconi et al. 2003); there is also evidence for extra-temporal gray and white matter (GM, WM) reductions in epilepsy patients when compared to controls (Woermann, Free et al. 1999) (Keller, Wilke et al. 2004) (Bernasconi, Duchesne et al. 2004). Nearly 80% of MTLE cases exhibit lateralized HC atrophy (Bernasconi, Bernasconi et al. 2003), strongly correlated with an ipsilateral seizure focus.

Surgical outcome prediction via MRI has been attempted in ATL (Jack, Sharbrough et al. 1992) (Radhakrishnan, So et al. 1998) (Kobayashi, Lopes-Cendes et al. 2001) with an accuracy ranging from 62% to 96%(Gilliam, Faught et al. 2000). Fewer authors have attempted to predict surgical outcome for SAH based on MRI. Abosch *et al.* (Abosch, Bernasconi et al. 2002) reviewed EEG, histopathological studies and

pre-operative MRI for 54 patients, half of whom had good post-operative seizure control (Engel Groups I/II). Their main conclusion is that SAH outcome can be best predicted when patients exhibit clear unilateral TLE; there remains a subpopulation (nearly 50% of cases), however, that meets the criteria for TLE, but does not become free from seizure following SAH (Abosch, Bernasconi et al. 2002). Antel *et al.* (Antel, Li et al. 2002) have used a classifier based on MRI and MR spectroscopy, correctly predicting the surgical outcomes of 39 of 52 (75%) of patients who became seizure free and 21 of 29 (72%) of patients who did not.

There remains a need therefore to increase the accuracy of non-invasive, pre-operative surgical outcome prediction in SAH.

Hypothesis and objectives of research

Our research hypothesis is that there exist anatomical regions of the brain that are linked to post-surgical outcome, either inside or outside of the MTL, and whose image informative content has the potential to predict surgical outcome.

Our goal is to demonstrate the usefulness of a new methodology, based on pre-operative MRI, in the prediction of surgical outcome in a retrospective study of MTLE patients having undergone SAH. In the first part we will describe our attempt at identifying those regions by comparing pre-operative MRIs of positive outcome patients to those of negative outcome patients. In the second part we describe our methodology for predicting surgical outcome based on simple MR image features computed within the identified regions.

Methods

Subjects

The study population consisted of 39 patients with intractable, non-foreign tissue MTLE seen at the Montreal Neurological Institute for SAH. All patients gave written consent for our study, which was approved by our Institute Review Board. Lateralization of seizure focus in MTLE patients was determined by a comprehensive evaluation including prolonged video-electroencephalogram (EEG) telemetry. The EEG focus was defined as right or left if more than 70% of seizures were recorded from one side.

Pre-operative T1-weighted (T1w) MRI 3D images were acquired on a 1.5T scanner (Philips Gyroscan, Best, Netherlands) using a fast gradient echo sequence (TR=18 *ms*, TE=10 *ms*, 1 NEX pulse sequence, flip angle= 30°, matrix size=256 X 256, FOV=256*mm*, slice thickness=1*mm*). All global MRI data were processed to correct for intensity non-uniformity due to scanner variations (Sled, Zijdenbos et al. 1998), linearly registered into stereotaxic space and resampled onto a 1*mm* isotropic grid (Collins and Evans 1997). Manual MRI volumetry of the HC (rater: N.B.), based on a previously published protocol (Bernasconi, Bernasconi et al. 2003), showed atrophy ipsilateral to the seizure focus in all patients.

All SAH procedures were performed by the same surgeon (A.O.) with a standardized approach (Olivier 2000). Post-operative follow-up for a minimum of 12

months was used to consolidate patients in two outcome groups: seizure free (positive outcome, **n=25**) or not seizure free (negative outcome, **n=14**).

The clinical data collected consisted in lateralization of seizure focus, age of onset, duration of epilepsy, and outcome grouping, along with MR-based manually extracted ipsi- and contralateral HC volumes.

Detection of GM/WM concentration differences related to surgical outcome

Our primary objective was to determine regions of gray and white matter (GM / WM) atrophy related to MTLE surgical outcome. To this end we performed between-group voxel-based morphometry (VBM) (Ashburner and Friston 2000) studies. These are statistical analyses of generalized linear model results performed on a voxel-by-voxel basis on combined cohorts of co-registered subject imaging data. In this case, we compared the GM and WM concentration maps of negative outcome to positive outcome patients.

The GM/WM concentration maps were obtained after classification of T1w MRI volumes (Zijdenbos and Dawant 1994), linearly registered (Collins and Evans 1997) to a common reference target of 152 young healthy volunteers, the so-called MNI Brain or ICBM 152 average (Mazziotta, Toga et al. 1995). GM or WM concentration maps were blurred using an isotropic Gaussian kernel of 10*mm* full-width at half-maximum, under the assumption that this spatial kernel size matched the expected changes (Ashburner and Friston 2000). In those smoothed

concentration maps, each voxel takes on a concentration value between (0,1) indicative of the presence or not of GM or WM in that voxel. T-statistics maps were obtained by estimating a generalized linear model at each voxel, with age of onset and duration as covariates. Significant clusters, above a threshold t-value and composed of a sufficiently large number of contiguous voxels ("cluster extent"), were obtained following the procedure developed by Worsley *et al.* (Worsley, Andermann et al. 1999) to correct for data nonuniformity.

Prediction of surgical outcome

Our hypothesis for predicting surgical outcome is that regions of differences in GM or WM concentration related to surgical outcome, which have been extracted in the previous step, can be exploited for classification purposes. We thus proceeded to define a global region of interest as the ensemble of all voxels above the cluster statistical significance threshold in each of the four concentration difference maps (one each for the combination of increase/decrease in GM/WM concentrations). We measured the number of voxels within these regions of interest, along with the region's mean GM and WM concentration, on the patients' classified preoperative image, for a total of 8 VBM-based measurements.

Ideally there would be two groups of subjects, one to extract the global region of interest, and the second to test the classification. However, due to the limited number of patients ($\mathbf{N = 39}$), we decided to proceed in a leave-one-out fashion where multiple VBM group comparison analyses were carried out, with each subject

removed in turn from the group; the concentration difference maps were therefore independently assessed from that subject. The global region of interest was calculated as defined above, and the number of voxels in the region, along with the mean GM/WM concentration within that region, was measured on the subject's own MR image. This approach was therefore repeated 39 times, in order to benefit from making comparisons statistically independent.

To summarize, the 8 VBM measurements (average value and volumes for GM/WM increases/decreases in concentration) for each subject "left out" were measured on the regions defined by the VBM analyses of the remaining 38 subjects. When completed, we had age of onset, duration of seizures, ipsi and contra-lateral hippocampal volumes and those 8 VBM-based values for each subject that can be used as input to the classifier.

Classification was performed using linear discriminant analysis of the aforementioned VBM and non-VBM based measures to assess group membership, with prior probabilities set to 0.7 (positive) and 0.3 (negative) in accordance with data set frequencies. The final reported result consisted in the mean accuracy of multiple cross-validated trials (Figure 26). In each of those trials, the study data is randomly split in two groups: the first group is the *training set* and is used to estimate the classification functions; the second group is the *test set*, classified using the functions derived from the training set. This procedure is repeated multiple times and the average proportion of correct classifications of the test set is an

empirical measure for the success of the discrimination. The cross-validation experiments were completed using SYSTAT 10.2 (SSI, Richmond, CA).

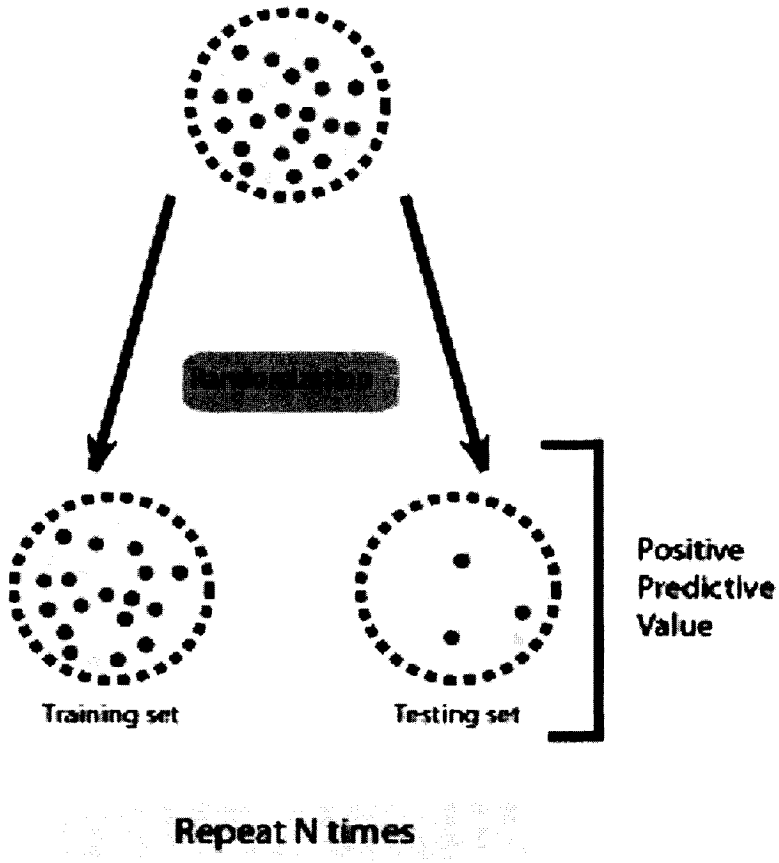

Figure 26 – Cross-validation trials

Cross validation trials proceed as follows. The study population is randomly split into a training and a test set, with the linear discriminant functions estimated on the training set. The test set is used to evaluate how good the classification is, and this process is repeated multiple times to get a good estimate of the overall accuracy for classification.

Experiments

Four experiments were run:

1) Prediction of surgical outcome using non-VBM based measures (age of onset, duration, ipsi and contralateral HC volumes);

2) Prediction of surgical outcome using VBM-based measures only;

3) Prediction of surgical outcome using age of onset, duration and VBM-based measures; and

4) Prediction of surgical outcome using all non-VBM and VBM-based measures.

Results

There were no statistical differences in age of onset, duration of epilepsy, ipsilateral or contralateral hippocampal volumes between the positive and negative outcome groups ($P > 0.05$, $DF = 37$) (see **Table 4**). Box plots of age of onset, duration, ipsi and contra-lateral volumes are shown in Figure 27.

Outcome	N	Age of onset (*yrs*)[SD]	Duration (*yrs*)[SD]	Ipsilateral HC volume (mm^3)[SD]	Contralateral HC volume (mm^3)[SD]
Positive	25	10.3 [6.0]	21.0 [12.1]	2588.8 [417.9]	3665.2 [669.1]
Negative	14	15.1[13.2]	24.1 [11.3]	2327.8 [449.3]	3333.4 [679.3]
P (Neg. to Pos.)		> 0.05	> 0.05	> 0.05	> 0.05

Table 4 - Clinical data for positive and negative outcome groups

There was an equal ($N = 7$) number of patients with left-sided seizure focus and ipsilateral HC atrophy than right-sided patients in the negative outcome groups,

while there were 14 left-sided and 11 right-sided patients in the positive outcome groups. Due to the limited number of patients in this study, the relationship between outcome and lateralization of seizure focus was not investigated further.

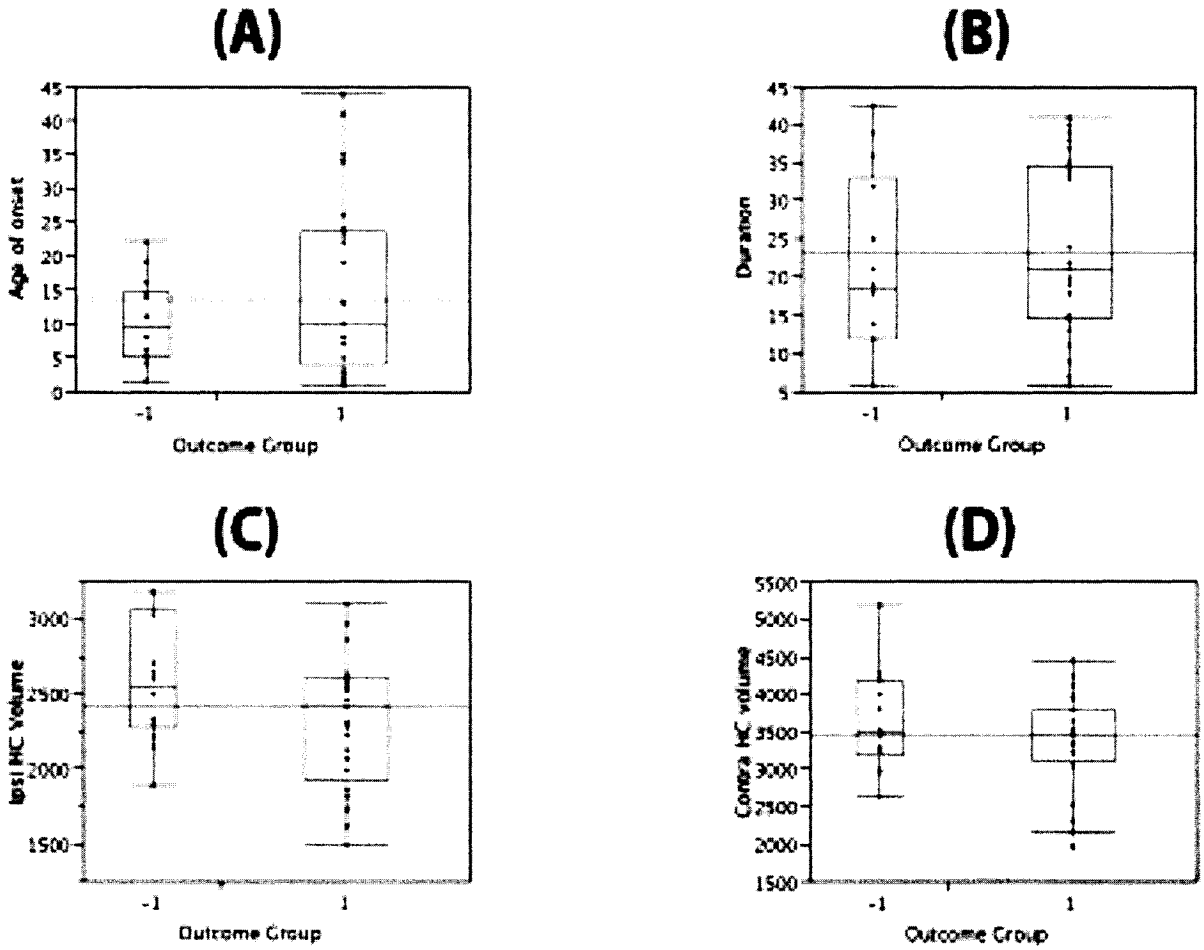

Figure 27 – Patient data

(A) and (B) Box-plots of age of onset and duration of epilepsy for the negative (group -1) and positive (group 1) outcome patients. (C) and (D) Box-plots of ipsi- and contra-lateral hippocampal volumes for the same patients. Experiment 1 assessed the accuracy of individual prediction using these data in multiple cross-validation trials. Specificity to positive outcome was 87%, sensitivity to negative outcome was 13.8%, resulting in a combined accuracy of 72.6%.

Figure 28 presents the results of between-group VBM analyses for GM and WM concentration differences, respectively, with all 14 negative outcome patients compared to all 25 positive outcome patients. All voxels displayed are above the cluster threshold for statistical significance (t=[-3.326, 3.326] for decrease or increase in concentration respectively, $P < 0.05$, corrected for data nonisotropy and multiple comparisons). Even though many voxels reached peak significance, and many more were above the cluster threshold, there were no clusters that reached significance, as none reached the constraint on their extent (number of contiguous voxels)(Worsley, Andermann et al. 1999); we will therefore report statistical trends. GM concentration decreases (negative relative to positive outcome patients) are detected primarily in the left fronto-temporal and left inferio-temporal regions, while concentration increases are found primarily in the basal forebrain region (see Discussion). Trends in WM concentration decreases (negative relative to positive outcome patients) are noted in the left mid-temporal area, caudal to the hippocampus; an area of increased WM concentration in negative outcome patients relative to positive can be found in the right cingulate at the parietal level.

It should be restated at this point that our definition of regions of interest is the ensemble of voxels above the t-statistic cluster threshold for significance; we did not impose as a condition that a voxel belong to a significant cluster. This relaxed condition entails that voxels included are not significant using standard VBM criteria however, as seen below, they can be used for classification.

The mean number of voxels per region was 3640 [SD 1120] for GM decrease and 973 [446] for GM increase, while the mean WM decrease region was 419 [215] voxels and the mean WM increase region was 2963 [959] voxels. Box plots of mean GM and WM concentration measures within the regions of interest are shown in **Figure 29**.

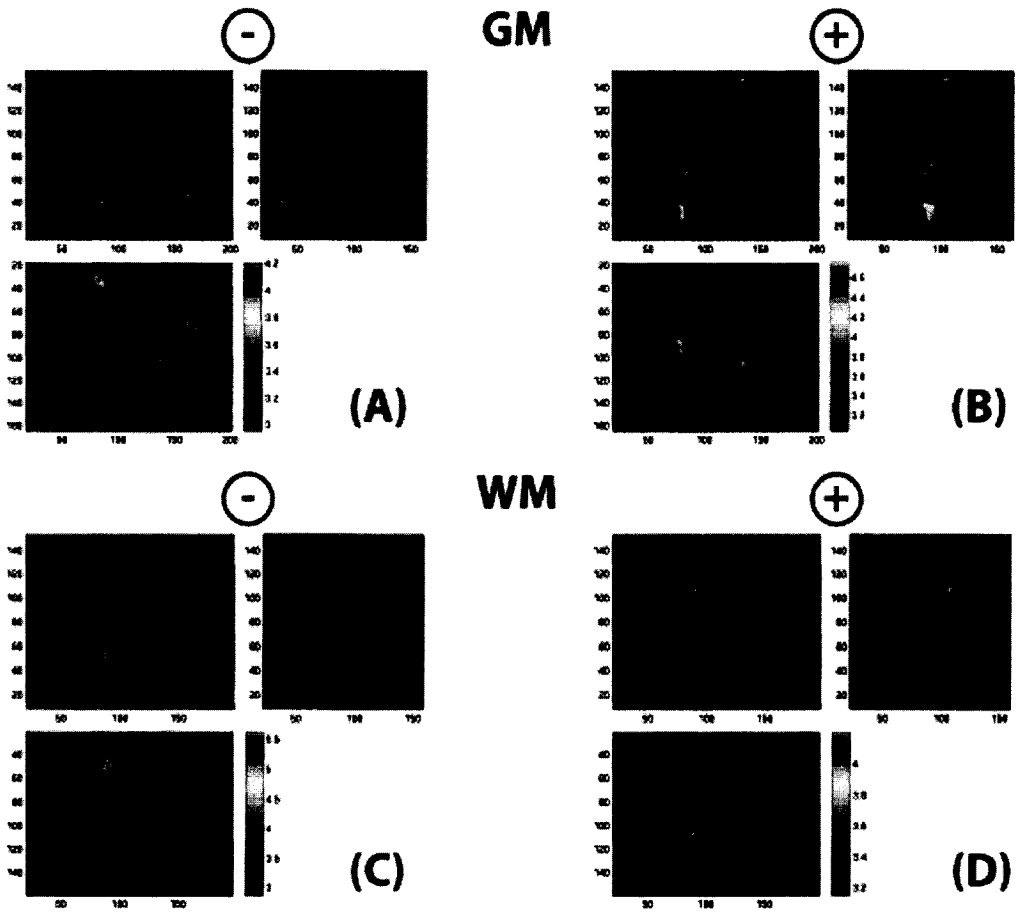

Figure 28 – VBM results related to surgical outcome

GM concentration decreases (A) and increases (B) in the negative outcome group when compared to positive outcome group as determined by VBM, shown in glass brain view; WM concentration decreases and increases are shown in (C) and (D), respectively. All voxels are above the cluster threshold for statistical significance (t > | 3.36|), but clusters are not significant as they do not meet the statistical extent criterium. GM concentration decreases are seen in the left fronto-temporal and left inferio-temporal regions, while increases are noticed in the basal forebrain region, likely due to misclassification due to poor contrast. Trends in WM concentration decreases are noted in the left mid-temporal area, caudal to the hippocampus; an area of increased WM concentration in negative outcome patients relative to positive

can be found in the right cingulate at the parietal level. Note that this VBM was completed using all 39 subjects. When extracting the VBM-based measures for each subject "left out", the VBM would only use the remaining 38 subjects.

Given those data, Experiments 1 through 4 assessed the accuracy of individual prediction using a model built either with non-VBM (Experiment 1: 4 variables, age of onset, duration, ipsi and contralateral HC volumes), or VBM measures (Experiment 2: 8 variables, number of voxels, GM/WM mean concentration in the increase/decrease regions of interest), or a combination of the two (Experiment 3: 10 variables, age of onset, duration, 8 VBM-based measures; Experiment 4: 12 variables, age of onset, duration, ipsi and contralateral HC volumes, and 8 VBM-based measures), with fifteen cross-validation trials run in each experiment. The median number of subjects in the training sample and the test sample, the specificity for positive outcome and sensitivity to negative outcome, as well as the combined prediction accuracy, were recorded and are shown in

Table 5.

Expt	Variables	Median number of subjects		SP	SN	Accuracy
		Training set	*Test set*			
1	4	37	2	87%	13.8%	72.6%
2	8	37	2	91.7%	25.0%	66.7%
3	10	37	2	90.5%	50.0%	76.5%
4	12	37	2	90.9%	58.3%	78.9%

Table 5 – Summary of results [Specificity (SP) to positive outcome, sensitivity (SN) to negative outcome, and mean accuracy]

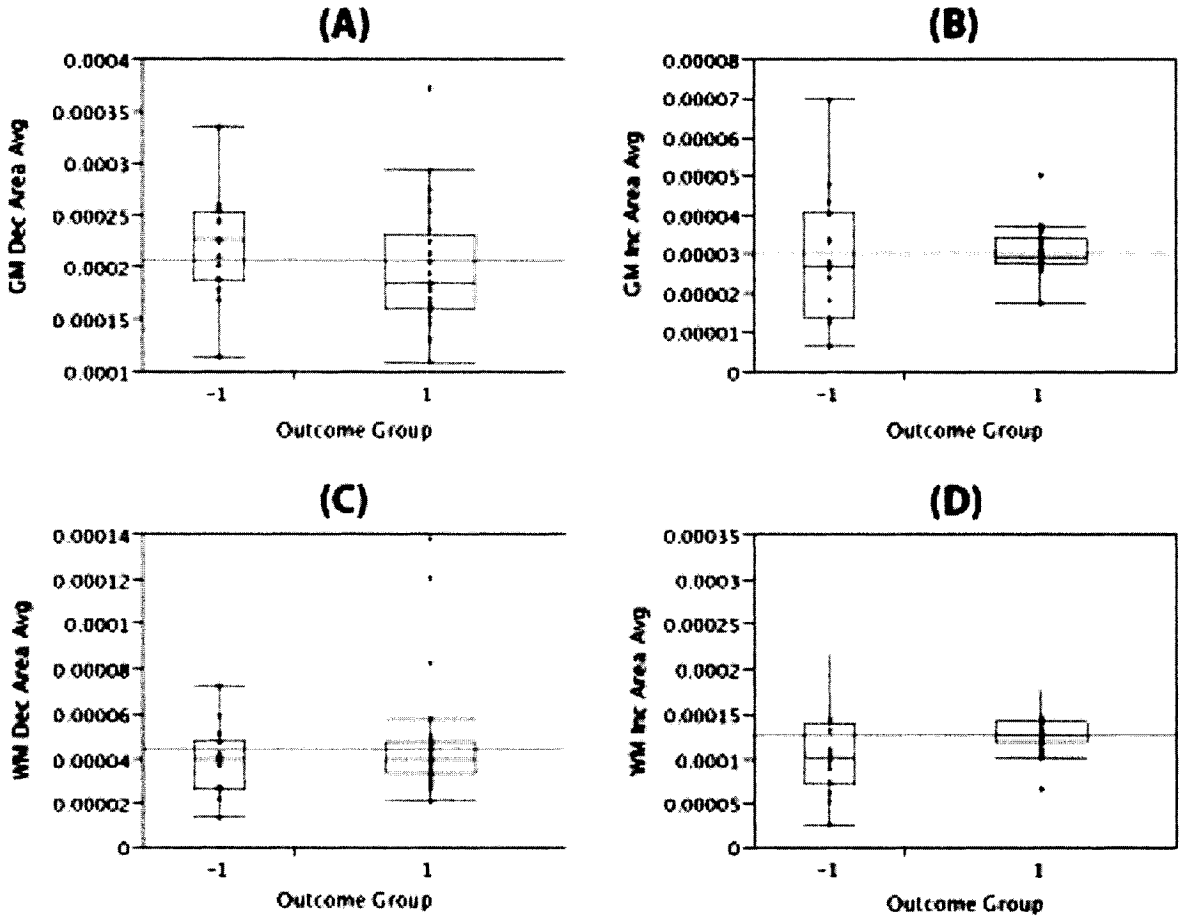

Figure 29 – VBM-based measurements

(A) and (B) Box-plots of GM area average decreases and increases in concentration values, respectively, for all voxels belonging to the global region of interest, in individual subjects volumes; WM area average increases and decreases are shown in (C) and (D) respectively. In this figure group (-1) represents negative outcome while group (1) represents positive outcome. While the spread in the measures is large, there are sufficient differences between the distributions to increase the specificity to positive outcome to 90.9%, the sensitivity to negative outcome to 58.3%, and the combined accuracy to 78.9%, when adding these features to those of Fig. 2.

Discussion

Hippocampal and extra-hippocampal involvment in MTLE

A consensus exists on the importance of HC resection in MTLE surgery, especially in cases where there is concordance between HC atrophy assessed on MRI and focal epileptic discharges measured via EEG. In such unequivocal cases, outcome can be reliably predicted (Jack, Sharbrough et al. 1992; Bautista, Cobbs et al. 1999); discordant assessments are not as reliable.

The situation outside of the HC is not as well defined. Bonilha *et al.* (Bonilha, Kobayashi et al. 2004) report that the extent of the AG and HC resection is clearly associated with surgical outcome, whereas pre-operative variables and interictal EEG abnormalities are not. They hypothesize that an incomplete resection of atrophic HC may explain most surgical failures in patients with MTLE due to unilateral HC sclerosis (Bonilha, Kobayashi et al. 2004). However, in the retrospective study by Abosch *et al.* (Abosch, Bernasconi et al. 2002) mentioned earlier, of 27 SAH patients, 13 of these subsequently having undergone either extension of the SAH (six cases) or a cortico-SAH (seven patients), showed that near-total resection of mesial structures did not necessarily improve outcome. The authors conclude by stating that the majority of patients receiving suboptimal seizure control following SAH did not meet the criteria for unilateral MTLE, and are therefore unlikely to benefit from additional resection of mesial structures. There remains a subpopulation, however, that meets the criteria for MTLE, but does not

become free from seizure following SAH. This is also supported by an analysis of the discharge patterns by Bautista *et al.* (Bautista, Cobbs et al. 1999), concluding that patients with focal interictal epileptiform discharges included in surgical resection have good surgical outcomes, while those with discharges extending beyond the area of resection experience poor surgical outcome.

The importance of extra-hippocampal anatomy for the prediction of postoperative outcome has been shown in a (18)fluorodeoxyglucose (FDG) positron emission tomography (PET) study by Dupont *et al.* (Dupont, Semah et al. 2000), where they correctly predicted 2 -year prognosis in 100% of the patients using 4 regions: the temporal pole, the medial temporal region, the anterior part of the lateral temporal neocortex, and the basofrontal region.

We have identified areas of GM and WM concentration differences that were related to surgical outcome in TLE patients having undergone SAH using a standard VBM approach. The fact that there exist regions of differences nearly reaching significance, regardless of the mixed nature of the groups (left and right seizure focus), is further indication of extra-hippocampal involvement in TLE. The basal forebrain differences are more than likely due to errors in the automated tissue classifier (Zijdenbos and Dawant 1994) due to decreased tissue contrast in this area. WM and GM concentration decreases however in the fronto-, inferio and mid-temporal areas are indicative of possible increased pathological affect in these areas.

By including this extra-hippocampal information to the other non-VBM information (age of onset, duration of seizure, ipsi/contralateral HC volume) we increase the accuracy of MR-based classifiers to 4 out of 5 patients (**Experiment 4**), primarily by increasing the sensitivity to negative surgical outcome. This accuracy is maintained even if we automate completely the image processing, that is without resorting to manual segmentation of the hippocampus (**Experiment 3**). We hypothesize that our results reflect the case that patients with poor surgical outcome have more disease-driven damage in extra-hippocampal areas, including extra-temporal areas.

Methodological Considerations

Automated techniques for surgical outcome prediction are by nature objective; the proposed methodology is also data-driven, requires no user intervention (Experiment 3 paradigm), and is based on a standard T1w MR acquisition.

We have been successful in identifying significant peaks and trend areas of GM and WM concentration differences using a standard VBM approach. Figure 28 displays the VBM results for the between-group differences for all subjects. Our classification technique relies on a leave-one-out approach for purposes of independence however, we do not feel that the exclusion of one individual from the group comparison significantly changes the regions of GM and WM concentration differences when compared to the whole-group results. These areas could be refined if we were to use other approaches which improve the accuracy of VBM, such as

modulation by the Jacobian determinant (Good, Johnsrude et al. 2001), flipping left/right depending on side of seizure focus, or restricting the search area for VBM by eliminating regions known not to be affected in epilepsy (e.g. cerebellum, brainstem).

A limitation of this study is the relatively small patient sample size. Great care was taken to avoid dependence and overlearning of the training set, by performing multiple leave-one-out VBM analyses and subsequent cross-validation trials for classification. A larger database than currently available would allow the separation into subgroups based on seizure focus laterality and/or to consider the four main Engel (Engel, Van Ness et al. 1993) outcome classes separately.

Conclusion

In this work we have set the basis for surgical outcome prediction for temporal lobe epilepsy patients undergoing selective amygdala-hippocampectomy, based on voxel-based morphometry analysis of pre-operative grey and white matter concentration maps. Regions of difference exist between positive and negative outcome patients, and can be used to predict surgical outcome with high accuracy using a completely automated methodology; the accuracy is increased marginally when including manual, expert-based volumetric information.

MR-based preoperative planning, including seizure focus lateralization, is an established procedure. This article presents the idea that pre-operative MRI can be

used to increase the accuracy of surgical outcome prediction for an individual patient. Using MRI to predict outcome, without depending on additional minimally-invasive (SPECT, PET) or invasive (surgically implanted EEG or sEEG) modalities, could reduce the significant resources and patient burden devoted to pre-surgical evaluation, while improving success. Of course, conventional and/or sEEG remain critical tools for the understanding of the temporo-spatial pattern of the epileptic discharges; it is only proposed here that MRI can play a potent role in pre-surgical planning and evaluation.

Chapter 6

Differentiation of AD and MCI from normal aging

Foreword

We have already mentioned that the difficulties in early clinical detection of AD lie for the most part in the similarities between its initial manifestations and cognitive impairment due to normal aging (NA) processes (Chetelat and Baron 2003). In order to resolve this situation an objective classification technique is required. Having previously demonstrated that existing appearance differences in TLE could be used for classification, we wished to test whether or not the same process could be used in the identification of AD or MCI patients when compared to normal aging individuals. The aspect of the method being investigated here is robustness, not to noise but rather to a change in the signal: given a completely dissimilar disease, with a radically different pathological process, would the classification be successful? We thought that our methodology was just as relevant as in TLE, as there were many

reports in the literature of intensity (Freeborough and Fox 1998) and shape changes (Fox and Schott 2004), to cite but a few, occurring with time between AD, MCI and NA. Further, there was ample evidence that the earliest pathological deposition of Aβ in the brain occurred in various structures within the MTL (Thal, Rub et al. 2002), and therefore our large volume approach, rather than a structure-centric one, seemed equally applicable.

Our goal remained to perform this task in a cross-sectional fashion, departing from the usual research paradigm of longitudinal examination. While we expected a loss of sensitivity and/or specificity due to this strategy, we believed that the clinical reality could still be enhanced by a technique that would be readily applicable for the treating physician. The standard of comparison for this task was a clinical diagnosis of probable AD or MCI according to criteria from NINCDS-ADRDA for the former and Petersen et al. (Petersen, Doody et al. 2001) for the latter.

Our results indicate that our methodology was indeed robust to a different pathological process. We found that the appearance-based approach was 92.6% accurate in separating a group of 20 normal aging individuals *vs* 17 AD and MCI patients. This was superior to the 82.1% accuracy of a model composed of age, MMSE scores, hippocampal and amygdala volumetric measurements.

An early form of this work was published as a conference proceeding (Duchesne, Pruessner et al. 2005), before being submitted for publication in the journal *Alzheimer's and Dementia.*

After submission of this manuscript, we became aware of a fundamental flaw in the image acquisition. Whereas we believed a priori that the scans had been acquired at a single site due to blinding during initial analysis, it appeared that such was not the case as patients and controls were actually scanned in 3 different centers within the Munich area. This could have had a significant impact on the results.

We have since rerun our analyses, including this element as a new variable. We are pleased to report in an addendum to this manuscript that the system was sufficiently robust to maintain an accuracy of 90.9%, while rejecting the institution as a discriminating variable. This would therefore indicate that the methodology could handle multi-site data, an important consideration in clinical research.

As the methodology has been extensively described in Chapter 3, the reader may choose to skip the Methods section of this chapter without loss of information. For the sake of brevity, keywords have not bee included, and a list of abbreviation can be found at the end of this thesis.

Differentiation of Alzheimer's dementia and Mild Cognitive Impairment from Normal Aging with a novel cross-sectional MR-based classification technique

Simon Duchesne [1], Jens C. Pruessner [2], Stefan Teipel [3], Harald Hampel [3], D. Louis Collins [1].

[1] Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, QC, Canada

[2] Brain Imaging Group, Douglas Hospital, McGill University, Montreal, QC, Canada

[3] Alzheimer Memorial Center, Dementia Research and NeuroImaging section, Ludwig-Maximilian University, Munich, Germany

To be submitted to Alzheimer's and Dementia, 2006

Abstract

Background - Neuropathological studies in Alzheimer's dementia (AD) have shown brain degeneration early in the disease course in the medial temporal lobes (MTL).

Objective - We aimed at differentiating AD and mild cognitive impairment (MCI) from normal aging (NA) with > 90% accuracy using a recently developed automated classification technique based on analysis of MTL information from cross-sectional MR data.

Methods

Subjects - 13 clinical AD and 4 MCI patients [age 68.0 (8.2), MMSE 20.2(5.0)] were compared together against 20 NA subjects [age 60.8 (8.6), MMSE 29.8(0.6)]. T1w MRI data were acquired after informed consent on a 1.5T scanner.

Analysis – All scans were corrected for intensity signal inhomogeneity, linearly registered in stereotaxic space and intensity normalized. Hippocampal (HC) and amygdala (AG) volumes were obtained by manual segmentation. Rectangular volumes of interest (VOI) were defined on the left and right MTL (80 X 52 X 60 voxels), linearly and nonlinearly registered to a reference target image. The classification features were the normalized intensity and the trace of the Jacobian of the nonlinear registration deformation fields. Normative spaces for group comparison were created from 152 normal young subjects using principal components analysis of their intensity and trace VOIs, and the scans from the NA

and MCI+AD cohort subjects were projected in these spaces. Forward stepwise linear discriminant analyses of the eigendata was used to classify the subjects in multiple cross-validation trials.

Results – There were significant group differences between NA and AD+MCI for age ($P = 0.0138$) as well as MMSE, HC and AG volumes ($P < 0.0001$). For classification using non-MR features (age, MMSE, left and right HC and AG volumes) the mean accuracy was 82.1%, sensitivity to AD+MCI 81.2%, and specificity to NA of 85.9%. For classification using only MR features, the mean accuracy was 92.6%, mean sensitivity 78.2% and mean specificity 97.3%.

Conclusion - Our results indicate that cross-sectional MR data can adequately discriminate NA from AD+MCI, at an improved accuracy when compared to HC and AG volumetry. This single-scan, practical and objective method holds promise for AD or MCI patient management.

Introduction

Neuroimaging techniques, in particular structural or anatomical magnetic resonance imaging (MRI), allow the visualization and quantification of pathologically induced brain changes in the living subject; it then becomes possible to employ those metrics as surrogate measures of the state of the disease. These techniques can help to achieve earlier diagnosis, characterize the time-course of the illness and generally increase therapy efficacy.

Alzheimer's disease (AD) is a progressive neurodegenerative disorder associated with disruption of neuronal function and gradual deterioration in cognition, function, and behavior (Khachaturian 1985) . The progression of AD is gradual, and the average patient lives 8 to 10 years after onset of symptoms (Petrella, Coleman et al. 2003). It is the most common cause of dementia in the elderly (65 years and older), responsible for 75% of all dementia cases (Group 2000; J. R. Petrella 2003). As the global life expectancy rises and populations age, the prevalence of AD is expected to triple over the next 50 years (Carr, Goate et al. 1997); improving care while reducing the socio-economic impact of AD is therefore an important and necessary topic of research.

The diagnosis of clinically probable AD can be made with high accuracy in living subjects only once the stage of dementia has been reached, and requires clinical, neuropsychological and imaging assessments (Petrella, Coleman et al. 2003). It can only be confirmed by postmortem histopathology. Neuropathological studies have shown however that brain degeneration occurs very early in the course of the disease, even before the first clinical signs, in certain regions such as the medial temporal lobe (MTL) (Petrella, Coleman et al. 2003) (Figure 30). The etiology can be summarized as neuronal dysfunction and eventual loss due to abnormal accumulation of beta-amyloid proteins.

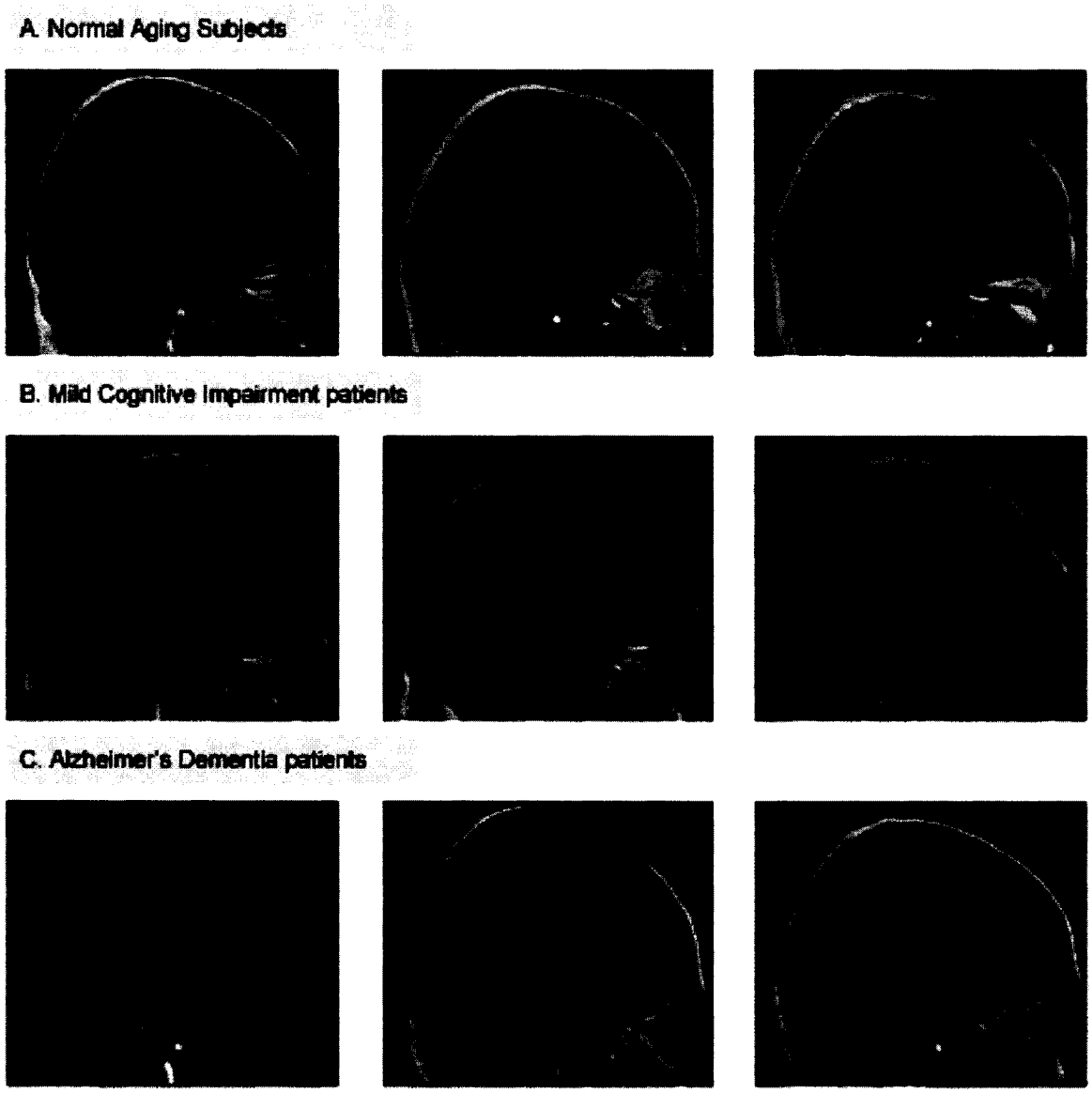

Figure 30 – T1w imaging in normal aging, MCI and AD

Anatomical magnetic resonance imaging (MRI) enables us to visualize pathologically induced brain changes in the living subject and use those as surrogate measures of the state of the disease. In general, radiological assessment of cross-sectional MR images demonstrates typical enlargement of ventricular and sulcal spaces in Alzheimer's dementia patients (AD, bottom row) when compared to normal aging individuals (NA, top row), but it is easy to find individuals with similar characteristics across both groups. The radiological situation is even more complex in the assessment of Mild Cognitive Impairment (MCI, middle row), an at-risk group of progression to AD. This significant overlap between the cross-sectional MRI appearance of NA, MCI and AD makes differentiation challenging for even the best trained observers.

The difficulties in early clinical detection lie for the most part in the similarities between cognitive impairment due to normal aging (NA) processes and the initial manifestations of AD (Chetelat and Baron 2003). Beyond changes due to normal aging, but before AD, mild cognitive impairment (MCI) is a condition referring to patients with significant but isolated memory impairment relative to subjects of identical age (Petersen, Doody et al. 2001). MCI individuals are considered an at-risk group for progression to AD, and therefore early prediction of progression to AD in those patients is an important research goal.

Our general hypothesis is that cross-sectional, single-scan classification may be possible based on detectable T1w MR signal intensity differences that would possibly be due to pathology-related microscopic changes, combined with macroscopic alterations in structure shape, measurable via registration-based metrics. These features would be extracted a large, non-specific Volume of Interest (VOI) centered on the MTL (see Figure 31).

Our goal is to demonstrate that it is possible to differentiate NA subjects from AD and MCI patients with high accuracy based solely on cross-sectional MRI data using an automated classification technique.

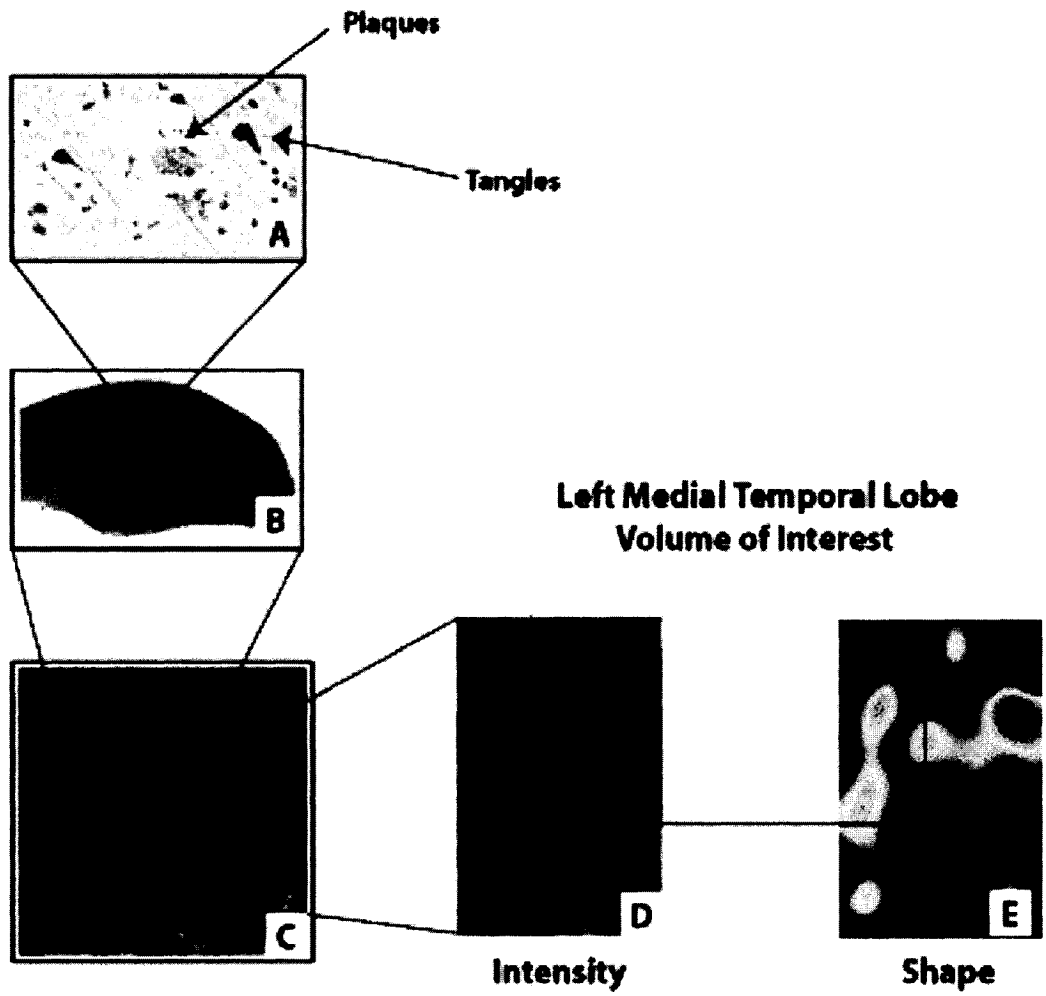

Figure 31 – Hypothesis for classification

Histological data for AD shows that tangles and plaques start to accumulate primarily in the medial temporal lobe, and are ultimately responsible for neuronal dysfunction and eventual loss. The hypothesis for our methodology is that pathology related microscopic cellular changes such as those would eventually affect the T1-weighted MR signal on a macroscopic level. The resulting intensity and shape changes can be exploited for classification purposes, even though the tangles or plaques cannot be seen on T1w MRI at that resolution.

Method

Study group

A total of 194 subjects from three different groups were included in this study. The reference group consisted in 152 young normal subjects from the International Consortium for Brain Mapping database (ICBM)(Mazziotta, Toga et al. 1995), whose scans were used to create the reference model. The study population consisted in 20 normal aging subjects (NA), 13 AD and 4 MCI patients. Our goal was to demonstrate the ability of our classification methodology to differentiate NA from MCI or AD using MRI and therefore, we merged the latter two groups into a single patient group with 17 subjects.

The patients had clinical diagnosis of probable AD according to the NINCDS-ADRDA (McKhann, Drachman et al. 1984), recruited from the Alzheimer Memorial Center, Ludwig Maximilian University (Munich, Germany). Further sociodemographic information on study subjects is shown in

Table 6. Cognitive impairment in the AD patients was assessed using the Mini-Mental State Examination (MMSE)(Folstein, Folstein et al. 1975). MCI was diagnosed according to the criteria by Petersen et al. (Petersen, Doody et al. 2001). MCI patients performed below the age-adjusted average in tests on memory impairment in the absence of global cognitive impairment and impaired activities of daily living. Normal aging subjects were free of memory complaints and scored within 1 standard deviation on all axes of the CERAD (Consortium to Establish a

Registry for Alzheimer's Disease) cognitive battery. Significant medical co-morbidity in the AD/MCI patients and controls was excluded by 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, creatine, 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/MCI 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).

	N	Age (*yrs*)	MMSE	HC volume (mm^3)		AG volume (mm^3)	
				Left	Right	Left	Right
NA	20	60.8 (8.6)	29.8 (0.6)	3332 (570)	3144 (594)	970 (255)	986 (299)
AD+MCI	17	68.0 (8.2)	20.2 (5.0)	2060 (638)	1996 (628)	602 (274)	549 (236)
P (NA vs AD+MCI)		**0.0138 ***	**<0.0001***	**<0.0001***	**<0.0001***	**<0.0001***	**<0.0001***

Table 6 – Information on NA, AD, MCI cohort

Data acquisition and initial processing

The reference ICBM subjects were scanned on a Philips Gyroscan 1.5T scanner (Best, Netherlands) using a T1w fast gradient echo sequence (TR=18 *ms*, TE=10

ms, flip angle=30°, matrix size=256 x 256, FOV=256 *mm*, slice thickness=1 *mm*, sagittal acquisition). MRI examinations for the study subjects were performed on a 1.5T Siemens Magnetom Vision MRI Scanner (Siemens, Erlangen, Germany), with a volumetry volumetric T1w sagittally oriented MRI sequence (TR=11.6*ms*, TE=4.9*ms*, resolution 0.94 x 0.94 x 1.2 mm^3, rectangular FOV for sagittal slices 256 *mm* (SI) x 204 (AP) *mm*]. Initial MRI processing was identical for all subjects in the study. Following acquisition, intensity inhomogeneities due to scanner variations were corrected (Sled, Zijdenbos et al. 1998) and the mean grey-level intensity was scaled to that of the standard reference for all subjects. Linear registration was used for global alignment into the same standard reference space (Collins, Neelin et al. 1994), that corresponds to a voxel-by-voxel average of the 152 reference subjects previously registered in the Talairach-like stereotaxic space in the context of the ICBM project (Mazziotta, Toga et al. 1995). The data was finally resampled onto a $1mm^3$ isotropic grid (Mazziotta, Toga et al. 1995).

MRI-based hippocampi and amygdalae volumes

MRI volumetry of left and right hippocampi and amygdalae was performed by one rater (J.C.P.) for the NA and AD+MCI patients using a previously published protocol (Pruessner, Li et al. 2000).

MRI-based multi-dimensional classification

The classification method can be summarized as follows (see Figure 32). First, from pre-processed data **(A)** volumes of interest centered on the left and right MTL were

extracted using stereotaxic-like (Talairach and Tournoux 1988) coordinates (start coordinates x=[+2,57] for the left and right sides respectively, y=[-53] and z=[-52]). Each VOI thus measured 55 x 82 x 80 = 360800 voxels. The extent of the VOI captured the hippocampus and neighboring MTL structures, such as the entorhinal, perirhinal and parahippocampal gyri, irrespective of normal inter- and intra-individual variability. To further reduce positional variations, which would propagate as unwanted noise in the morphometric modeling, a local affine transformation was applied to co-register the individual subject's VOI with the corresponding VOI of the reference target.

The second step was to generate a normal, non-pathological eigenspace **(B)** from the ICBM reference group of young, normal control subjects (N = 152). This multidimensional eigenspace was created by uniting results from four distinct Principal Component analyses of (i) intensity images of the left and right VOIs **(C)**; and (ii) a measure of local volume change for the left and (Worsley, Marrett et al. 1996) right VOIs **(D)**. The latter estimate was derived from the determinant of the Jacobian of the deformation field, obtained following non-linear registration of the subject volume to the reference image.

Nonlinear registration attempts to match image features from a source volume to those of the reference image at a local level, typically in a hierarchical fashion, with the aim of reducing a specific cost function. Whereas many nonlinear registration processes exist, the one chosen for this study was ANIMAL, developed by Collins et

al. (Collins and Evans). This algorithm attempts to match image grey-level intensity features at a local level (voxel) in successive blurring steps, by maximizing the cross-correlation function of voxel intensities between the source and reference images. The result is a dense deformation field, capturing the displacements required to align the subject VOI with that of the reference image.

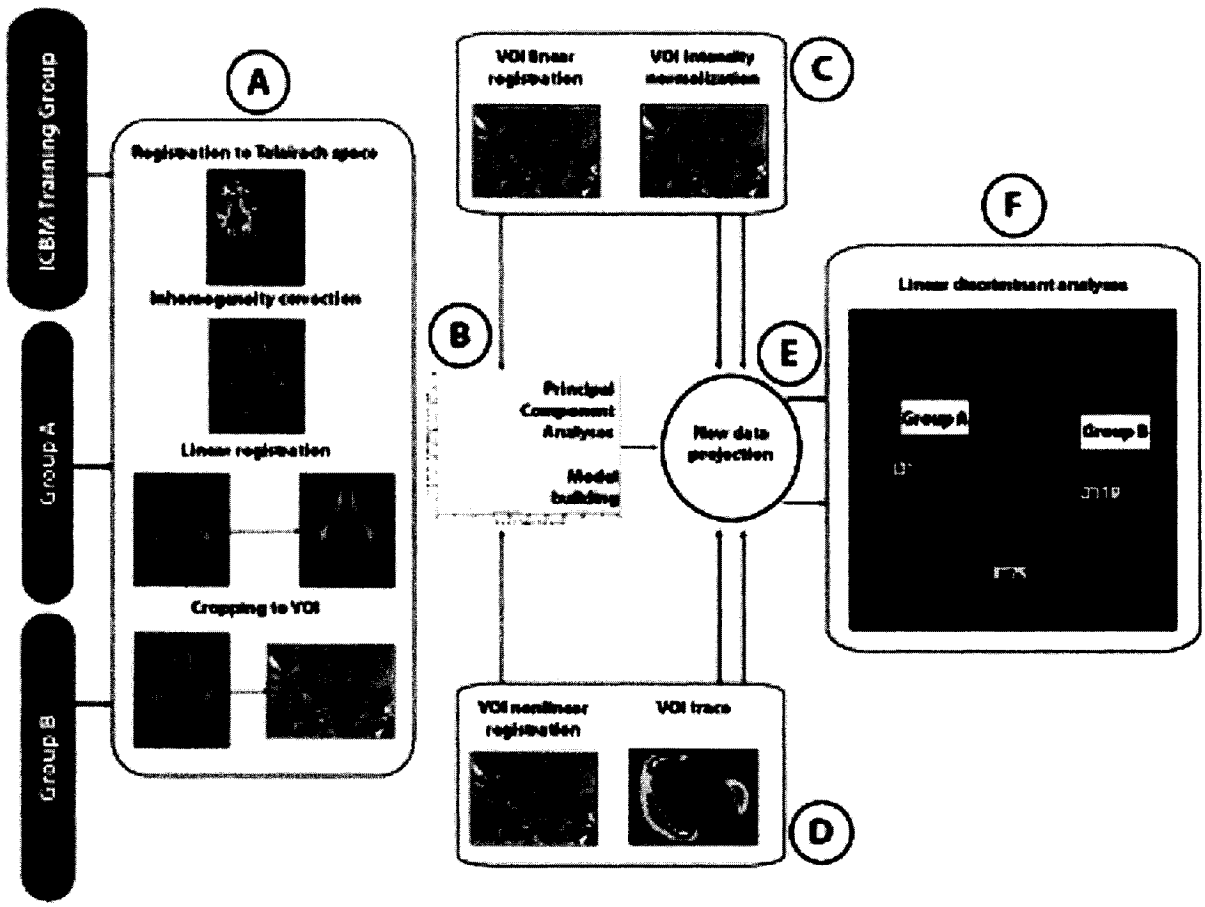

Figure 32 - Methodology

The classification methodology used in this article can be broken down in the following steps. First, all images are pre-processed in an identical fashion, which includes signal inhomogeneity correction, global alignment to a standard stereotaxic space, and extraction of two volumes of interest, centered on the left and right medial temporal lobes, shown in (A). A local affine transformation is applied to the local VOI to improve the alignment with the reference volume, and the mean intensity of the VOI is scaled with respect to the mean intensity of the reference volume in (C). We also extract, for each of those volumes of interest, a measure of

shape variation in (D). Intensity and shape models for each VOI are built using Principal Components Analysis in (B), in order to generate a reference space from a group of images of 152 young healthy volunteers. Our ultimate goal being to discriminate between groups A and B (NA *vs* AD+MCI), we process their images in a similar fashion and project the intensity and shape volumes in the multidimensional space of the reference group in (E). For classification, we exploit the differences in the projection eigencoordinates using crossvalidated, linear discriminant analyses in (F).

The classification of NA and AD+MCI subjects was accomplished by projecting their VOIs in the multidimensional reference eigenspace **(E)**, and then by estimating the classification functions with forward stepwise linear discriminant analyses. The goal is not to find a feature space that optimally represents the two groups but rather to find a suitable basis for comparison, allowing for effective classification (Figure 33).

The final reported result consisted in the mean accuracy of multiple cross-validated trials (Figure 34). In such designs, the study data is randomly split in two samples, the first (training set) to estimate the classification functions, used to classify the second sample (test set). The proportion of correct classification for the test set is an empirical measure for the success of the discrimination.

Experiments

There were two classification experiments performed:

- Experiment 1 was aimed at separating the NA and AD+MCI groups on the basis of non-MR features (baseline age, MMSE, left and right AG and HC volumes); and

- Experiment 2 was aimed at separating the NA and AD+MCI groups on the basis of MR features, extracted in the fashion described above.

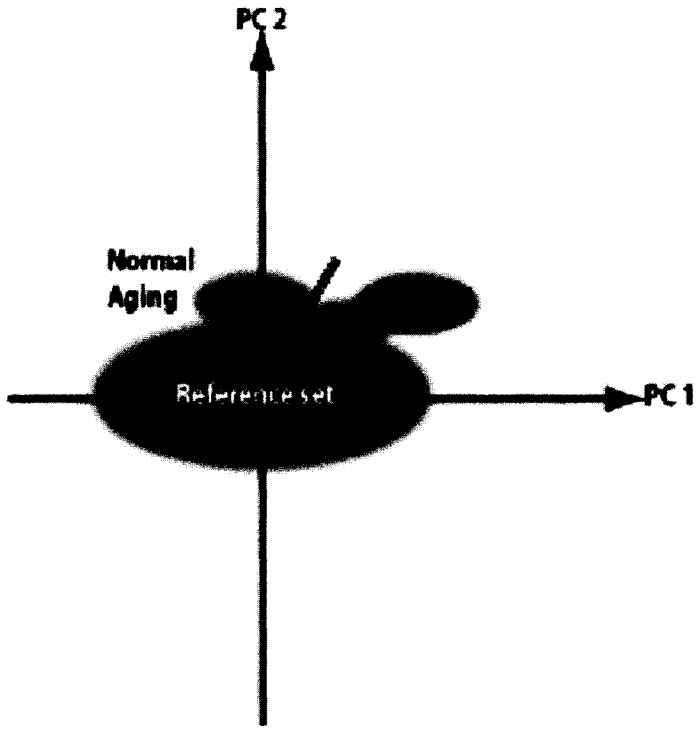

Figure 33 – Cartoon-like representation of classification of patient groups

The reference set is represented by the blue ellipse, aligned and centered with the first two principal components since the eigenspace originates from this data. Our goal however is to use the same space for the separation of NA individuals (green) from AD+MCI patients (orange and red). The goal is not to find the optimal space in which to represent the NA, AD or MCI subjects, but rather to have a common basis to compare populations.

Results

Subjects age, MMSE and volumetry

There was a significant group difference for baseline age (DF=36, P=0.0138) when comparing the NA and AD+MCI groups (Table 6). The average MMSE score in

the AD group was 18.7, which would qualify them as being "moderate" AD (Petrella, Coleman et al. 2003); the MCI group MMSE average was 25.5. When combined the AD+MCI MMSE group average was 20.2, significantly different (DF=36, P<0.0001) from the NA group MMSE average of 29.4.

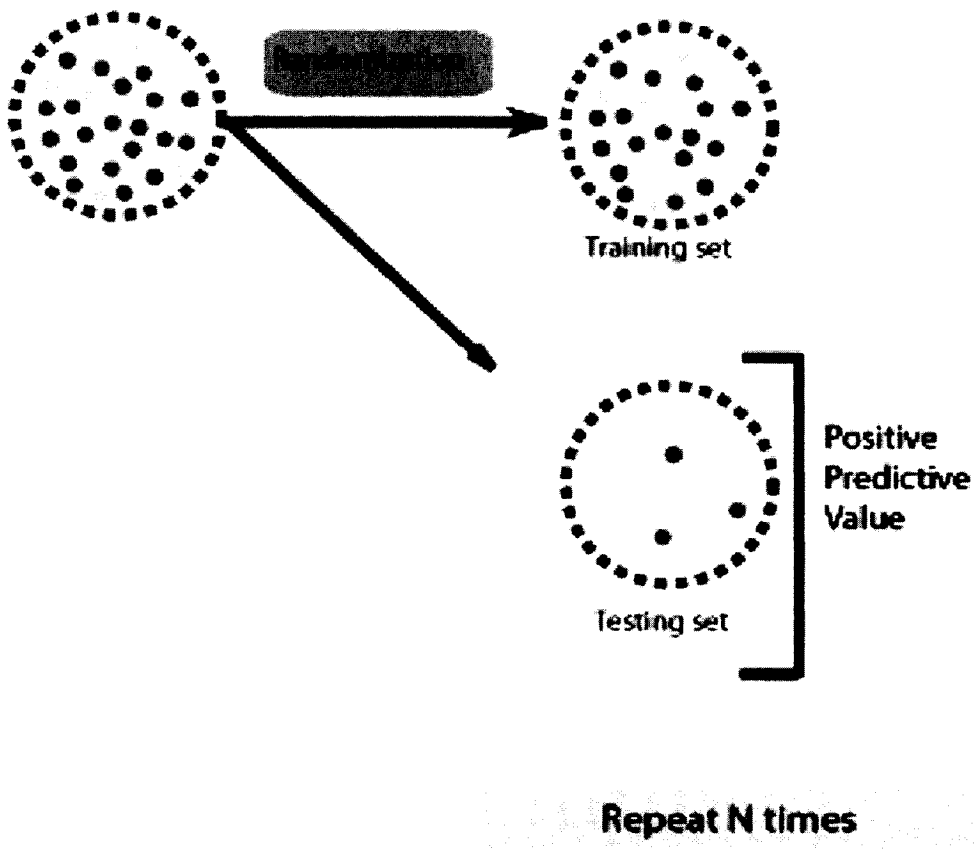

Figure 34 – Cross-validation trials

The study population is randomly split into a training and a test set, with the linear discriminant functions estimated on the training set. The test set is used to evaluate how good the classification is, and this process is repeated multiple times to get a good estimate of the overall accuracy for classification.

There were significant volume differences between the two groups under study for all structures (Table 6): left HC (DF=36, P<0.0001), right HC (DF=36, P<0.0001), left AG (DF=36, P=0.0002), and right AG (DF=36, P<0.0001).

Given those group differences, **Experiment 1** was designed to assess the accuracy of individual classification using these 6 features (age, MMSE, left and right HC, left and right AG volumes). Forward, stepwise linear discriminant analysis (P-to-enter = 0.15) was used to select the discriminatory variables from the training set and compute the classification function for each trial of the test set. Fifteen cross-validation trials were run, with a median of 30 subjects in the training sample and 7 in the test sample. The mean accuracy for individual classification was 82.1%, the mean sensitivity to AD+MCI was 81.2%, and the mean specificity to NA 85.9%. The median number of variables retained by the forward, stepwise linear discriminant analysis was 1, with never more than 2 variables retained by the model. More specifically, the left HC volume was always chosen; in 9 trials out of 15, it was the only variable retained. In 5 trials, age was also significant, and in one trial, left AG volume.

Experiment 2 was designed to assess the accuracy of individual classification using MR-based features. Our overall MRI-based model, formed by the union of the intensity and trace principal components for the left and right VOIs, was built using our training set of 152 young control subjects. A total of 538 principal components were retained, accounting for 99.7% of the variance of the *reference* data set.

Fifteen cross-validation trials were run, with a median of 32 subjects in the *training* sample and 5 in the *test* sample. Forward, stepwise linear discriminant analysis (P-to-enter = 0.15) was used to select the discriminatory variables from the training set and compute the classification function for each trial of the test set. Normality of the eigencoordinates along those eigenvectors was confirmed using quantile plots and Shapiro-Wilke test ($P > 0.05$). The mean accuracy for individual classification was 92.6%, the mean sensitivity to AD+MCI was 78.2%, and the mean specificity for NA 97.3%.

The median number of variables retained by the forward, stepwise process was 17. A plot of the NA *vs* AD+MCI eigencoordinate data along the three most occurring discriminative eigenvectors is shown in Figure 6. All but two of the 7 most occurring eigenvectors throughout the trials (occurrence > 2) were shape eigenvectors.

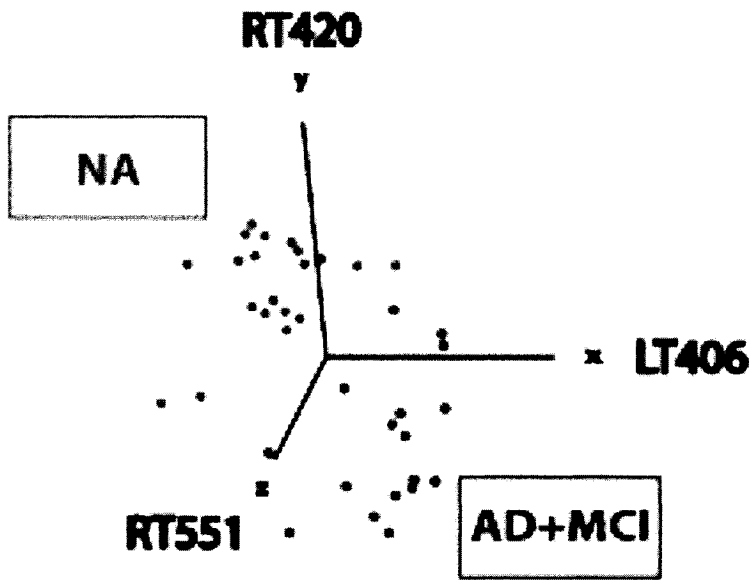

Figure 35 – Patients data projected along most discriminant eigenvectors
The NA (green) and AD+MCI (red) data is shown here plotted along the three most

discriminating eigenvectors in the multidimensional reference space of cross-sectional MR intensity and shape information. One can estimate a plane that would optimally separate the two groups in one of the cross-validation trials. The resulting average discriminatory accuracy for 15 trials was 92.6%, with a mean sensitivity of 78.2% and mean specificity of 97.3%.

Discussion

Our results indicate that the position information (eigencoordinates) of new data projected in a multidimensional normative domain is sufficient to adequately discriminate between normal aging individuals and AD or MCI patients. The mean accuracy of the technique, given the study population, was 92.6%, its mean sensitivity 78.2% and mean specificity 97.3%.

Methodological implications

In order for automated, quantitative neuroimaging methods to gain widespread acceptance in the scientific and clinical community, they must meet and exceed the lower bound on overall diagnosis accuracy set by current clinical performance, estimated at 78% in the case of AD (Weiner, Albert et al. 2005). Techniques must also be reproducible, practical, objective and sufficiently accurate to detect changes due to the pathology of interest.

Accuracy - The identification of group membership based on MR features is on average 10% more accurate than if it were done with baseline age, MMSE and

manually segmented HC and AG volumes. It is also 15% more specific. However, it has to be noted that the technique is less (7.7%) sensitive.

The increased accuracy can readily be explained when one considers the informational value of the two approaches. In HC volumetry, between 3000 to 4000 voxels are compressed in a single scalar measurement, expressed in mm^3, while in the methodology being proposed, more than 4 X 300000 voxels are processed and expressed in terms of 538 principal components coordinates. Clearly, additional important information is embedded in MTL voxels, and is included in our analysis in the form of intensity and shape covariates.

Reproducibility - Being a completely automated methodology, the technique presented in this article is therefore completely reproducible. Volumetry, on the other hand, shows intra (0.91-0.95) and inter-rater (0.83-0.91) variability, as demonstrated in a study on hippocampal segmentation by Pruessner *et al* (Pruessner, Li et al. 2000). This added variability in segmentation measurements – volumetric noise - reduces the classification accuracy of any methodology.

Practicality - It is important to note that the PCA technique as employed here requires no manual intervention other than the definition of the VOI, which is done once and based on Talairach coordinates on the reference image. This is a significant reduction of required resources when comparing such a technique to volumetry or other expert-based techniques that may require 2 hours per hippocampus for manual segmentation (Hogan, Mark et al. 2000).

Objectivity - The underlying assumption for this work is that there exists sufficient information in a chosen VOI from MRI to be used for the accurate determination of group membership in all subjects. As there is no further user interaction other than the delineation of the VOI, the process is completely data-driven and therefore objective.

Clinical implications

Use of cross-sectional MRI vs other imaging modalities for AD diagnosis

For patients with AD or at an increased risk of developing AD (e.g. MCI) a number of approaches can be considered in order to aid the diagnosis, all with varying degrees of success. Neuropsychological screening tests are necessary to recognize and monitor these subjects, with an accuracy of 78 % (Weiner, Albert et al. 2005). Neuroimaging approaches must be considered, however disadvantages of techniques other than routine MRI preclude them from being widespread AD diagnosis-enabling technologies. PET (Klunk, Engler et al. 2004) and SPECT (Kogure, Matsuda et al. 1999; Herholz 2003) offer sufficient specificity and sensitivity in the differential diagnosis of AD from other cortical and subcortical dementias and also offer prognostic value however, both modalities are minimally invasive procedures with radiation dose limitations and therefore cannot be repeatedly performed on a single patient for longitudinal studies nor used as a screening mechanism for large

populations. Other techniques such as fMRI (Johnson, Saykin et al. 2000), MR spectroscopy (Kantarci, Smith et al. 2002), MR diffusion tensor (Stahl, Dietrich et al. 2003) and MR magnetization transfer (Kabani, Sled et al. 2002) show promise for the future, but they are difficult to implement in a clinical setting without a dedicated research group for technical support. Finally, as compared to MRI, CT images lack the detailed soft-tissue information necessary for detecting subtle structure changes associated with the disease, especially at an early stage, even though late-stage measures show promise(Frisoni, Rossi et al. 2002).

Longitudinal analysis of MRI has been proposed to differentiate between aging, MCI, and AD (Fox and Freeborough 1997; Fox and Schott 2004) with high accuracy. However, by its very nature such a method implies a delay between scans before any assessment can be made. That delay, required to detect a change induced by the pathology in serial scans, reduces the window of opportunity for early treatment or optimal patient management. Cross-sectional analysis, on the other hand, by generating an immediate assessment, would allow for better patient management if shown to be sufficiently accurate.

Automated classification vs expert-based volumetry

Volumetry, that is volume measurement of a particular bounded structure of interest, is the primary indicator of structure integrity. Stereology, i.e. the technique of proper sampling based on the Cavalieri principle (Cavalieri 1635), has been the preferred approach to calculating volumes before the advent of high-resolution

images, where volumes can be reliably estimated directly from the (often isotropic) voxel count, slice thickness and pixel size. Cross-sectional measurements of the hippocampus have achieved classification accuracies between controls and AD patients of nearly 90%, and between individuals with MCI and AD patients near 80%, with hippocampal volumes and/or deformation analysis (Csernansky, Wang et al. 2000; Pennanen, Kivipelto et al. 2004) Longitudinally-assessed MR-based atrophy rates of the hippocampus is also a reliable indicator of the state of the disease (Jack, Slomkowski et al. 2003) .

While undeniably useful for understanding disease-driven changes in the chosen structure(s), volumetry relies on accurate segmentation, a process subject to inter/intra-rater variability (Pruessner, Li et al. 2000) as mentioned earlier. This variability is likely to increase in MCI and AD as boundaries shift due to atrophy and become less clearly delineated as intensity changes. This added variability in segmentation measurements, or volumetric noise, will reduce the classification accuracy of any methodology based on such information. Finally, it requires expert knowledge, making it unsuitable for large-scale studies and may limit its availability in a clinical setting.

Morphological findings

From the results of Experiment 1, we can reasonably conclude that left HC volume is the main variable differentiating NA and MCI+AD groups, in line with other observations from different groups (Visser, Scheltens et al. 1999; Mungas, Reed et al.

2002; Thompson, Hayashi et al. 2004). Likewise, from Experiment 2, local volume changes are the strongest discriminators between our two groups since shape eigenvectors are mostly used in the classifier. This is also supported by the literature that shows different atrophy patterns between AD and NA (Fox and Schott 2004; Thompson, Hayashi et al. 2004).

A more detailed assessment of the variations in intensity and shape embedded in the significant eigenvectors should yield important information regarding the pathophysiology of the disease, but is beyond the scope of this paper. This assessment should provide important clues to the on-going disease processes common across individuals and their implications on brain morphology. This task will be performed, and interpreted with their clinical correlates, in a separate communication.

Conclusion

We report results from a quantitative MRI technique that achieved the required sensitivity and specificity in the cohort studied by classifying AD+MCI *vs* NA with 92.1% accuracy, 97.3% specificity and 78.2% sensitivity. The methodology does not rely on segmentation, requires no user input and is data-driven. Such a system stands to benefit patients and health care systems immediately.

Addendum

After submission of this manuscript, we became aware of a fundamental flaw in data analysis due to blinding to subject data. Whereas we believed *a priori* that the scans had been acquired at a single site, it appeared that such was not the case as patients and controls were actually scanned in 3 different centers within the Munich area (see Table 7 for breakdown). We therefore needed to assess the impact on the results.

Group	Site 1	Site 2	Site 3
NA	15	5	0
MCI	0	0	4
AD	0	4	9

Table 7 - Acquisition sites

Leave-one-out classification based solely on acquisition sites separates individuals in our two groups (NA vs AD+MCI) with 76% accuracy.

Experiment 1 was rerun including acquisition site as a variable. Forward, stepwise linear discriminant analyses (P-to-enter = 0.15) were used in fifteen cross-validation trials, with a median of 33 subjects in the training sample and 4 in the test sample. The mean accuracy for individual classification was 83% (compared to 82.1% previously), the mean sensitivity to AD+MCI was 83.4% (compared to 81.2% previously), and the mean specificity to NA 79.6% (compared to 85.9% previously).

The median number of variables retained by the forward, stepwise linear discriminant analysis was 2, and in all trials, acquisition site was retained as a discriminant variable.

Experiment 2 was repeated as well, including acquisition site along with MR features. Fifteen cross-validation trials were run in a similar fashion as **Experiment 1**, with a median of 33 subjects in the training sample and 4 in the test sample. The mean accuracy for individual classification was 90.9% (compared to 92.6% previously), the mean sensitivity to AD+MCI was 76.2% (compared to 78.2%), and the mean specificity for NA 97.3%, identical to the previous result. The median number of variables retained by the forward, stepwise process was 17, and in 14 out of 15 trials, acquisition site was not retained by the stepwise process; in the one trial where it was, it was of low rank.

This would therefore indicate that our methodology is sufficiently robust to multi-site differences in scanner hardware and acquisition variations. While more balanced and planned testing is required, this fortuitous discovery bodes well for the ability to handle multi-site data, an important consideration in clinical research.

Chapter 7

Prediction of MCI progression to AD

Foreword

Outside of therapeutic research, one of the most important questions in the field of AD is to determine who, within at-risk groups, and eventually the general population, will develop the disease. The study of the morphological differences between MCI patients that will progress to AD (decline in cognitive status towards dementia) *vs.* those that do not progress is still in its infancy, and no consensus has emerged yet regarding possible different appearance patterns between these patients. Longitudinal structural analysis does indicate a higher rate of atrophy in structures such as the hippocampus for progressors, however cross-sectional analysis is not equally sensitive (Chong and Sahadevan 2005).

We confirmed in the last chapter that differences related to AD pathology could be estimated with a linear modeling approach of appearance principal components.

Given these, the dimensions that we sought to investigate in this last article were two-fold: on the one hand, extend this work to address the important issue of prediction of progression to AD in a cohort of MCI patients, and on the other, assess the system's robustness to noise and low resolution data.

The cohort we investigated was part of a larger study on memory loss in the elderly, arguably one of the few studies of this kind in the world with such long-term clinical follow-up. We selected 24 MCI patients from this study, 2/3 of whom progressed to AD within a mean time of 2.6 years, and the remaining patients remaining stable after a mean follow-up of 5.7 years. Incidentally, baseline neuropsychological testing proved insufficient for predicting progression to AD (Chertkow, Bocti et al. 2005). The MR data we used had $1mm$ in-plane resolution but $5mm$ slice thickness. Even though we resampled that data to a $1mm$ isotropic grid, the classification was essentially presented with 20% less information when compared to the three studies in previous chapters. We also realized that there were many movement artifacts, noise contributors to the images. Those artifacts were responsible for rejecting most (about 2/3) of the patients from the parent memory study.

Using noisy, low resolution data served two objectives: by itself, it was a test of how much data was actually needed for the classifier, as well as giving an idea of its robustness. Secondly, it was also a real-world test, where the clinical reality dictates that acquisitions must be performed quickly, especially in aging subjects and even

more so in patients with dementia; if proven successful in this data set, it would bode well for its use outside of the research and into the clinic.

The standards of inclusion and progression for patients in this study were based on extensive clinical diagnosis of MCI and probable AD according to criteria including Petersen et al. (Petersen, Doody et al. 2001), the Montreal Cognitive Assessment scale (Nasreddine, Phillips et al. 2005) and the NINCDS-ADRDA.

We are pleased to report that the methodology was robust and highly successful to predict future clinical state in our cohort of MCI patients, with an accuracy of 82.6%. This essentially means that, given only baseline scans for these individuals, an automated analysis technique could predict progression to dementia on average 2.6 years before the clinical diagnosis in four patients out of five. We also established that the main areas of signal difference were related to T1w intensity in the right MTL, calling for future work in biological interpretation.

An early form of this work was published as a conference abstract (Duchesne, De Sousa et al. 2005), before being submitted for publication in the journal *Lancet Neurology*.

As the methodology has been extensively described in Chapter 3, the reader may choose to skip the Methods section of this chapter without loss of information. For the sake of brevity, keywords have not been included, and a list of abbreviation can be found at the end of this thesis.

Predicting MCI progression to AD with appearance-based classification technique at baseline MRI

Simon Duchesne [1], Howard Chertkow [2], Christian Bocti [2], Kathy De Sousa [2], ,

D. Louis Collins[1]

[1] Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, Canada

[2] Center for Aging, Lady Davis Institute, McGill University, Montreal, Canada

To be submitted to Lancet Neurology, 2006

Abstract

Background - Amnestic Mild Cognitive Impairment (MCI) individuals are known to be at risk for progression to Alzheimer's Dementia (AD). There is evidence that in those who will progress, measurable hippocampal and entorhinal cortex atrophy from MRI serves as a moderate, though labor-intensive, predictor.

Objective - We aimed at predicting progression to AD in MCI patients using a recently developed MRI-based automated classification technique, which uses all MRI data within a 3D rectangular volume of interest, automatically positioned on the medial temporal lobes (MTL).

Methods

Subjects –24 amnestic MCI patients gave informed consent to participate in this study, which was approved by our IRB. Mean follow-up after baseline neuropsychological testing and MRI was 5.7 (4.0) yrs. 16 patients progressed to AD after 2.6 (1.6) yrs [mean age 77.7, sd (5.1) yrs, MMSE 26.9 (2.4) at baseline], and 8 remained non-demented at date of latest follow-up [age 80.3 (6.3) yrs, MMSE 28.5 (1.2) at baseline].

Analysis – Baseline T1-weighted MRIs were acquired on a 1.5T scanner (5 mm slices), corrected for intensity inhomogeneity, linearly registered in stereotaxic space, resampled to a $1mm^3$ grid and intensity normalized. Rectangular volumes of interest (VOI) were defined on the left and right MTL (80 x 52 x 60 voxels). Each VOI was further linearly and non-linearly registered to a reference target image.

Multidimensional reference spaces were built using principal component analysis of (i) the normalized intensity and (ii) a shape metric, corresponding to local volume change, calculated via the trace of the Jacobian of the dense deformation field from nonlinear registration. This was done first for a reference set of 152 normal young subjects. Afterwards, scans from the study subjects were projected in these spaces to extract features for each data set. Forward stepwise linear discriminant analyses were used on the projection coordinates in multiple cross-validation trials to assess the classifier performance.

Conclusion - Results of 15 cross-validation trials show an average sensitivity to progression to AD of 94% (15/16), average specificity to non-progression of 75% (6/8) and overall average accuracy of 88% (21/24). These results indicate that single-scan MR data projected in multidimensional feature domains has the potential to adequately predict future clinical status of amnestic MCI patients. This practical, automatic, and objective method therefore holds promise for early AD detection as an aid to diagnosis and patient management.

Introduction

Mild cognitive impairment is a term that describes elderly individuals with memory complaints and significant objective cognitive impairment relative to subjects of identical age (Flicker, Ferris et al. 1991; Petersen, Doody et al. 2001). Individuals with mild cognitive impairment are considered an at-risk group for progression to dementia (DeCarli 2003; Dubois and Albert 2004). For those individuals where memory loss is the predominant complaint, the term amnestic MCI (hereafter referred to simply as MCI) is now used. MCI strongly predicts progression to dementia due to Alzheimer's disease (AD)(Petersen 2004). AD is a progressive neurodegenerative disorder associated with disruption of neuronal function and gradual deterioration in cognition, function, and behavior (Khachaturian 1985). Early prediction of progression to AD in MCI patients is therefore an important research goal.

The etiology of AD is not clearly known, but most researchers stress the abnormal accumulation of Aβ and Tau proteins (Thal, Rub et al. 2000; Thal, Rub et al. 2002; Giannakopoulos, Herrmann et al. 2003) leading to neuronal dysfunction and eventual cell loss and abnormal cell death. Neuropathological studies in AD have shown in fact that abnormal accumulation of proteins and brain degeneration occurs very early in the course of the disease, even before the first clinical signs. Neurofibrillary tangles accumulate early on in certain regions such as the medial temporal lobe (MTL)(Cummings, Pike et al. 1996; Nagy, Hindley et al. 1999; Thal,

Rub et al. 2000; Thal, Rub et al. 2002; J. R. Petrella 2003; Petrella, Coleman et al. 2003). Microscopically a strong predictor of premortem cognitive dysfunction appears to be the relative area of entorhinal cortex occupied by beta-amyloid deposition (Cummings, Pike et al. 1996). Macroscopically, there is a growing body of evidence indicating that entorhinal cortex atrophy is a strong predictor of progression to AD in MCI individuals (Du, Schuff et al. 2003; deToledo-Morrell, Stoub et al. 2004; Pennanen, Kivipelto et al. 2004)

To date the diagnosis of clinically probable AD can be made with high accuracy in living subjects only once the stage of dementia has been reached, and requires clinical, neuropsychological and imaging assessments (J. R. Petrella 2003); it can only be confirmed by postmortem histopathology (Risse, Raskind et al. 1990).

For the detection of initial manifestations of AD in MCI patients a number of approaches are being proposed with varying degrees of success (Chertkow 2002; Chong and Sahadevan 2005). Clinical prediction and use of neuropsychological tests have been used to assess risk of progression in group studies, but have not yet proven their utility in individual cases (Daly, Zaitchik et al. 2000). The accuracy of neuropsychological cognitive markers for AD is approximately 78 % (Weiner, Albert et al. 2005). CSF biomarkers reach high sensitivity and specificity, but have associated complications in clinical practice and might not be acceptable on a repeated basis (Blennow and Hampel 2003).

Neuroimaging must be considered (Chetelat and Baron 2003) yet, disadvantages of techniques other than routine anatomical MRI preclude them from being widespread AD prognosis-enabling technologies. Positron Emission Tomography (PET) (Klunk, Engler et al. 2004; Nordberg 2004) and Single Photon Emission Computer Tomography (SPECT) (Kogure, Matsuda et al. 1999; Herholz 2003) have been used with varying degrees of success in diagnosing and predicting AD but are expensive and not standardized. Furthermore, both modalities are minimally invasive procedures with radiation dose administration and associated cumulative limitations, and therefore cannot be repeatedly performed on a single patient for longitudinal studies nor used as a screening mechanism for large populations. Other techniques such as fMRI (Johnson, Saykin et al. 2000), MR spectroscopy (Kantarci, Jack et al. 2000), MR diffusion tensor (Stahl, Dietrich et al. 2003) and MR magnetization transfer (Kabani, Sled et al. 2002) imaging show promise for the future, but are all difficult to implement in a clinical setting without a dedicated and sophisticated research group for technical support. Finally, as compared to anatomical MRI, CT images lack the detailed soft-tissue information necessary for detecting subtle structure changes associated with the disease at a very early stage, even though late-stage AD measurements show promise (Frisoni, Rossi et al. 2002). Cross-sectional and longitudinal analysis of MRI for measurement of volumes and atrophy rates of the hippocampi (HC) and other structures have been proposed to differentiate between aging, MCI, and AD (Fox and Freeborough 1997; Jack,

Slomkowski et al. 2003; Fox and Schott 2004) . There is evidence that in those who will progress, measurable hippocampal and entorhinal cortex atrophy as demonstrable on T1-weighted (T1w) MRI, serve as a moderate, though labor-intensive, predictor (Du, Schuff et al. 2001; Grundman, Sencakova et al. 2002; deToledo-Morrell, Stoub et al. 2004; Korf, Wahlund et al. 2004; Pennanen, Kivipelto et al. 2004). While undeniably useful for understanding disease-driven changes in the chosen structure(s), volumetry relies on accurate segmentation, a time-consuming process subject to inter/intra-rater variability (Pruessner, Li et al. 2000). More importantly perhaps, it requires expert knowledge, making it unsuitable for large-scale studies and this will limit its widespread availability in a clinical setting. Further, by its very nature longitudinal analysis implies a delay between scans before any assessment can be made. That delay, required to detect a change induced by the pathology in serial scans, reduces the window of opportunity for early treatment or optimal patient management. Cross-sectional analysis, on the other hand, by generating an immediate assessment using a single scan, may allow for better patient management if shown to be sufficiently accurate.

Our general hypothesis is that prediction of future clinical state may be possible based on T1w MRI signal differences, detectable on cross-sectional or single-scan images. These would possibly be due to pathology-related microscopic and macroscopic changes or alterations in structure shape, measured via intensity and shape-based features and forming a measure of the appearance of an image. Moving away from structure-centric approaches, appearance features extracted from a large, non-specific Volumes of Interest (VOI) centered on the MTL of the subject (see Figure 36) can be compared to multiple MR images of reference, and training and test groups are used to build a model for prediction.

The overall technique, called appearance-based classification (ABC), was applied successfully in the lateralization of the seizure focus in temporal lobe epilepsy (Duchesne, Bernasconi et al. 2006).

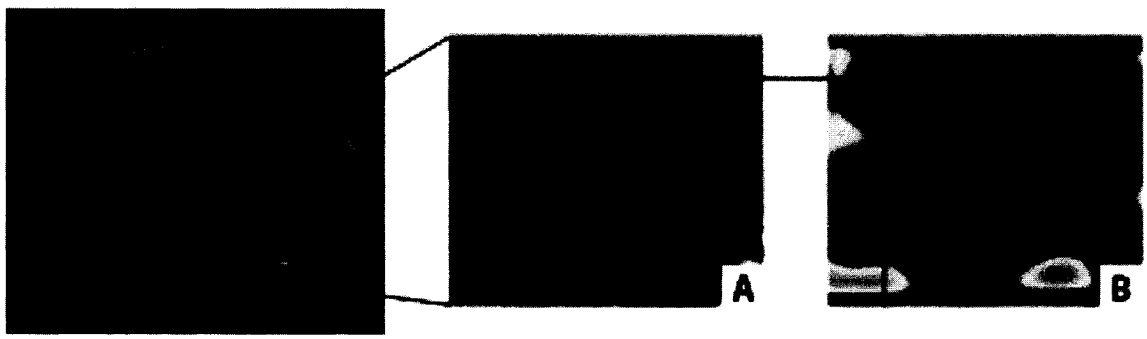

Figure 36 - Intensity and shape features within the volume of interest

(Left) Sagittal whole-brain view for subject with hippocampal atrophy. (Middle) Closer view of the Volume of Interest (VOI), as defined in stereotaxic space on the left medial temporal lobe. (Right) Identical view trough the trace VOI for the same subject. Green voxels do not move. From green to white (maximum) via yellows and reds indicate increase or expansion. From green to black (minimum), via blues and purples indicate decrease or contraction. The direction of movement is defined as the deformation that the subject's VOI seen in (A) must accomplish in order to align with the corresponding VOI extracted from the reference volume.

Our goal is to demonstrate the ability of this novel MRI-based automated classification technique in predicting progression to AD in a cohort of MCI patients.

Methods

Subjects and neuropsychological evaluation

The McGill University Study of Mild Memory Loss in the Elderly has followed a set of 89 amnestic MCI (with or without other cognitive domain problems) over 12 years. All subjects had presented to their family physician with complaints (from the individuals themselves or their families) of memory loss, and were referred for further evaluation. All were assessed by a neurologist or geriatrician skilled in assessment of memory-impaired elderly individuals, and were judged as also able to meet the criteria for Mild Cognitive Impairment (MCI) defined in the working group of Winblad and colleagues (Winblad, Palmer et al. 2004) as well as Petersen in his original operational definition of MCI (Petersen, Smith et al. 1995; Petersen, Smith et al. 1999). All had a history of memory decline in the last 1-4 years reported by the patient, caregiver (usually the spouse), or both, of a sufficient degree to bring them to medical attention. All subjects were documented to have objective memory impairment on a standardized mental status exam. Deficits in other cognitive areas (outside of short term memory) were judged as minimal and not sufficient to imply a

clinical diagnosis of dementia. In all cases, there was supporting objective evidence on clinical mental status evaluation as well as formal neuropsychological testing. The subjects did not meet the NINCDS-ADRDA criteria for the diagnosis of probable AD or the DSM-3 criteria for dementia, due to the lack of significant associated other cognitive deficits or to the lack of impairment of daily functioning (McKhann, Drachman et al. 1984) ((APA) 1994) . All were classified according to the Washington University clinical dementia rating (CDR) scale and met the criteria for "0.5" on that scale as described above (Hughes, Berg et al. 1982). There was no evidence on clinical evaluation of systemic or other neurological disease sufficient to interfere with cognitive function. CT and/or MRI excluded structural brain disease, and blood work was done including CBC, routine chemistry, thyroid function, serum B12, folate, and VDRL. All subjects scored less than 4 on the Hachinski ischemic scale (Hachinski, Iliff et al. 1975). The Mini-mental status exam (MMSE) (Folstein, Folstein et al. 1975) was carried out by the clinician as a global assessment tool.

Subjects received standardized clinical, mental status, neuropsychological, and cognitive evaluation. All subjects underwent neuropsychological evaluation which included the Logical Memory I and II components of the Weschler Memory Scale (Wechsler 1987), the RAVLT (Rey Auditory Verbal Learning Test), the FOME (Fuld Object Memory Examination), the Knopman and Ryberg test of verbal memory (Knopman and Ryberg 1989). A shortened version of the Boston

Naming test (Kaplan, Goodglass et al. 1978), letter and category fluency , tests of block design (Wechsler 1981) and clock drawing (Freedman and Leach 1994) as well as the digit symbol and digit span sub-tests of the WAIS- R verbal intelligence scale (Wechsler 1981) were also administered. Depression was excluded via the Yesavage Geriatric Depression Scale (Geriatric Depression scale > 15) (Yesavage, Brink et al. 1983).

The diagnosis of amnestic MCI in our cohort was supported by a neuropsychological evaluation establishing that a) there was memory performance at least 1 s.d. below the mean on one of the RAVLT (Rey Auditory Verbal Learning Test), the FOME (Fuld Object Memory Examination), or the Logical Memory (paragraph recall) subtest of the Wechsler Memory Scale – Revised.

The subjects were followed at 12-month intervals with both clinical assessment and repeat neuropsychological examination. At each visit, the clinical diagnosis was re-evaluated by the multi-disciplinary team headed by a neurologist or a geriatrician. The presence or not of clinical dementia was assessed. All subjects were followed for a minimum of 24 months unless death intervened.

Differences between subjects who declined and those who did not were assessed by t-test. Outcome data were analyzed by analysis of variance, using both outcomes of decline documented on MMSE, and clinical diagnosis of dementia, as outcome variables. Correlations of the presence of clinical features with outcome were assessed using Chi-square analyses.

All subjects provided written informed consent for the study and the McGill University Review Board approved the protocol. Readers should note that detailed results from this neuropsychological data have been presented in a separate communication (Chertkow, Bocti et al. 2005).

MRI acquisition and initial processing

MRI data for all subjects were acquired between 1993 and 1998 on a 1.5T GE Signa 5 scanner (GE Healthcare, Milwaukee, WI) using a 3D sequence (TR=300*ms*, TE=4.2*ms*, FA=90, sagittal acquisition with 256 (SI) x 256 (AP) 0.86 *mm* pixels, 5 *mm* slice thickness) at the Montreal General Hospital.

Following acquisition, intensity inhomogeneities due to scanner variations were corrected (Sled, Zijdenbos et al. 1998). Grey-level intensities were then commonly scaled across subjects. Affine (9 degrees of freedom or DF) linear registration was used for global alignment into a standard reference space (Collins, Neelin et al. 1994). The data was finally resampled onto a $1mm^3$ isotropic grid (Mazziotta, Toga et al. 1995). The reference image used for the linear registration and resampling was the International Consortium for Brain Mapping (ICBM) T1w target, known as icbm152, a voxel-by-voxel average of the 152 normal subjects previously registered in the Talairach-like stereotaxic space in the context of the ICBM project (Mazziotta, Toga et al. 1995).

MRI-based classification of subjects

The classification method used to predict future clinical state of the subject can be summarized as follows (Duchesne, Bernasconi et al. 2005). First, from pre-processed data, volumes of interest (VOI) centered on the left and right MTL were extracted using stereotaxic coordinates (start coordinates in Talairach space: x=[+2,57] for the left and right sides respectively, y=[-53] and z=[-52]). Each VOI measured 55 x 82 x 80 = 360800 voxels. The extent of the VOI captured the hippocampus and neighboring MTL structures, such as the entorhinal, perirhinal and parahippocampal gyri, irrespective of normal inter- and intra-individual variability. To further reduce positional variations, which would propagate as unwanted noise in the morphometric modeling, a local affine transformation (12 DF) was applied to co-register the individual subject's VOI with the corresponding VOI of the reference target.

The second step was to generate a normal, non-pathological eigenspace from the ICBM reference group of young, normal control subjects (N = 152). This multidimensional eigenspace was created by uniting results from four distinct Principal Component (PC) analyses of (i) intensity images of the left and right VOIs; and (ii) a shape metric, corresponding to a measure of local volume change for the left and right VOIs. The latter estimate was derived from the trace of the Jacobian of the deformation field, obtained following non-linear registration of the subject volume to the reference image. The classification of **P** *vs* **NP** groups of MCI patients

was accomplished by projecting the MRI VOIs from these subjects, processed in a similar fashion, in the multidimensional reference eigenspace; and then by estimating classification functions on the projection coordinates with forward stepwise linear discriminant analyses. The goal was not to find a feature space that optimally represented the two groups but rather to find a suitable basis for comparison, allowing for effective classification by deriving prognostic information correlated to a particular distribution of coordinates along a PC (see Figure 37).

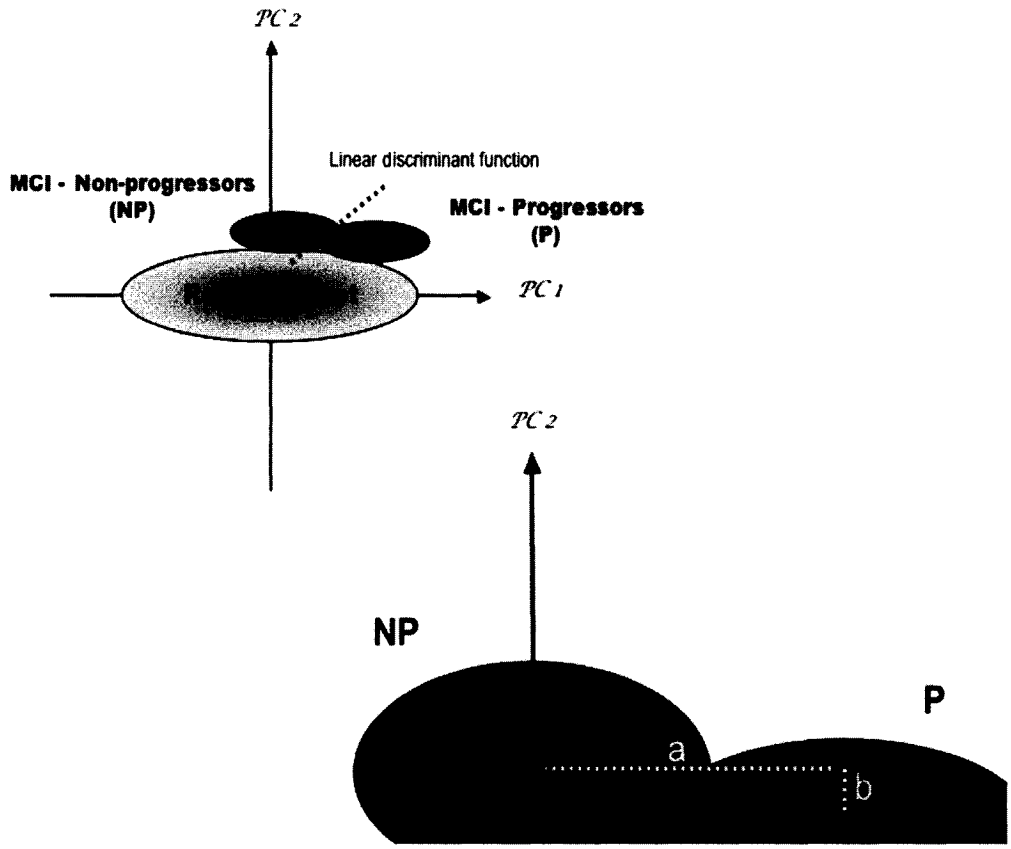

Figure 37 - Cartoon-like representation of classification of patient groups (Top) The reference set is represented by the green ellipse, aligned and centered with the first two principal components since the eigenspace originates from this

data. Our goal however is to use the same space for the separation of MCI progressors (red) from non-progressors (pink). (Bottom) The eigendistance between centers of the two populations can be expressed back in terms of image intensity and shape variations; these in turn can help with the biological interpretation of the results (see Figure 39).

The final reported results consist in the mean accuracy of multiple cross-validated trials. In each of these trials, the study data was *randomly* split in two groups: the first group was the *training set* and was used to estimate the classification functions; the second group was the *test set*, classified using the functions derived from the training set. This procedure was repeated multiple times and the average proportion of correct classifications of the test set was used as an empirical measure for the success of the discrimination. The cross-validation experiments were completed using SYSTAT 10.2 (SSI, Richmond, CA).

Experiments

There were two classification experiments performed:

- Experiment 1 was aimed at separating **P** and **NP** groups on the basis of two non-MR features, baseline age and MMSE; and
- Experiment 2 was aimed at separating **P** and **NP** groups on the basis of MR features, extracted in the fashion described above.

Results

Demographics

With the design of this longitudinal study, subjects were followed until they progressed to dementia, and then dropped from the cohort.

All subjects showed neuropsychological deficits in the memory domain greater than 1.5 standard deviation on age-matched tests. While clinically mild, these were at times equivalent or greater than the degree of memory loss.

In both **P** and **NP** subgroups, between 1/2 and 2/3 of subjects had additional domain deficits. There were all combinations of isolated memory as well as non-memory deficits accompanying the memory loss. The multiple patterns were encountered both in progressors as well as non-progressors. Chi square analysis of all combinations failed to find significant differences in the distribution of neuropsychological profiles, in any set of groupings. As mentioned previously, detailed results of the full subject cohort have been presented in a separate communication (Chertkow, Bocti et al. 2005).

Only 24 subjects out of the larger group of 89 had baseline MRI that could be recovered, and therefore form the sub-group investigated in this study. Other images were either lost in digital format due to backward hardware and software incompatibilities, or discarded due to the presence of extensive motion artifacts. Sixteen of these 24 subjects progressed to AD on average in 2.6 (std dev 1.6) yrs after an initial baseline visit (Group **P**), while the remaining eight subjects have

clearly not progressed during the follow-up period (Group **NP**, mean follow-up 5.7 (std dev 4.0) yrs).

Age and MMSE of the progressors and non-progressors is shown in Table 8. There was a significant group difference for baseline age (DF=22, P=0.0358) when comparing the **P** [mean age 78.5, sd (5.3) yrs] *vs* **NP** groups [72.0 (6.8) yrs]. The average MMSE score in the progressor group was 26.5 (2.3) at baseline scan, significantly different (DF=22, P=0.0111) from the non-progressor group MMSE average of 28.5 (1.2) at baseline scan.

	N	**Age** *(yrs)*	**MMSE**	**Conversion** *(yrs)*
Progressors (P)	16	78.5 (5.3)	26.5 (2.3)	2.6 (1.6)
Non-progressors (NP)	8	72.0 (6.8)	28.5 (1.2)	-
P *(P vs NP)*		**0.0138 ***	**<0.0001***	

Table 8 – Demographic and MMSE information on study subjects

Given the age and MMSE group differences, **Experiment 1** was designed to assess the accuracy of individual classification using these 2 features. For all cross-validation trials a model using both features was used to compute the classification function from the training set and apply it to the test set. Fifteen trials were run, with a median of 23 subjects in the *training* sample and 1 in the *test* sample. The

mean accuracy for individual classification was 63%, the mean sensitivity for progression 59%, and the mean specificity for non-progression 57%.

Experiment 2 was designed to assess the accuracy of individual classification using MR-based features. Our overall MRI-based model, formed by the union of the intensity and shape principal components for the left and right VOIs, was built using our reference set of 152 young control subjects. A total of 538 principal components were retained, accounting for 99.7% of the variance of the *reference* data set. Fifteen cross-validation trials were run, with a median of 23 subjects in the *training* sample and 1 in the *test* sample. Forward, stepwise linear discriminant analysis (P-to-enter = 0.15) was used to select the discriminatory variables from the training set and compute the classification function for each trial of the test set. Normality of the eigencoordinates along those eigenvectors was confirmed using quantile plots and Shapiro-Wilke test ($P > 0.05$).

The mean accuracy for individual classification was 88% (21/24), the mean sensitivity for progression 94% (15/16), and the mean specificity for non-progression 75% (6/8). The median number of variables retained by the forward, stepwise process was 10, with the three most occurring discriminating eigenvectors describing covariances in right MTL intensity (see Figure 38).

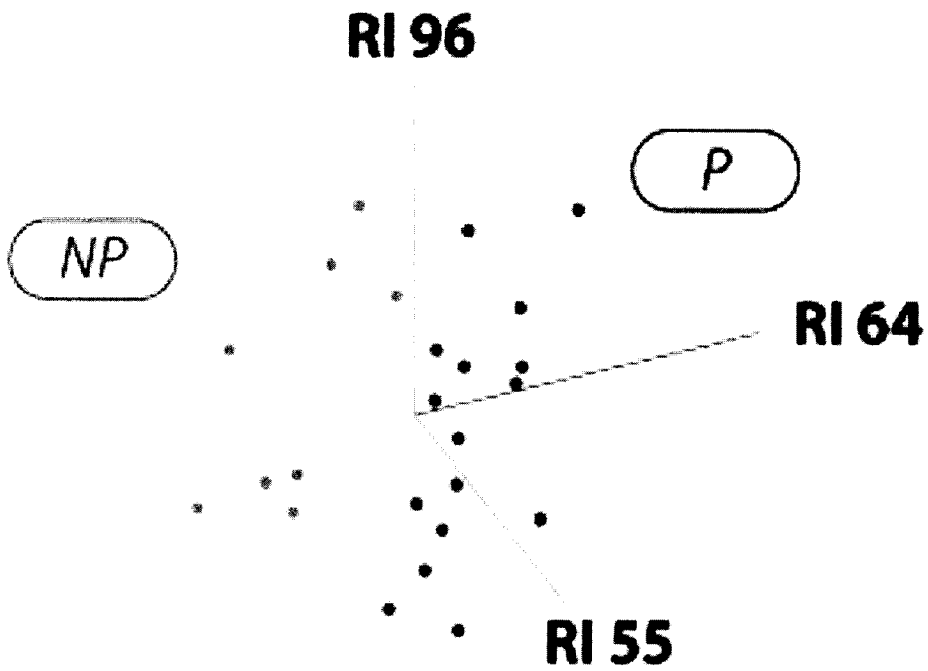

Figure 38 – Patients data plotted along most discriminat eigenvectors

The NP (green) and P (red) patients data are shown here plotted along the three most discriminating eigenvectors in the multidimensional reference space of cross-sectional MR intensity and shape information. One can estimate a plane that would optimally separate the two groups in one of the cross-validation trials. The resulting average discriminatory accuracy for 15 trials was 86%, on average 2.6 years before clinical diagnosis. Mean sensitivity for progression was 84.6% and mean specificity 83%.

Discussion

Our results indicate that the position information (eigencoordinates) of new data projected in a multidimensional reference domain has the potential to adequately predict progression of MCI to AD, on average 2.6 yrs before a clinical diagnosis is made. The mean accuracy of the technique, given the study population, was 86%, its mean sensitivity to progression 84.6% and mean specificity 83%.

The ability to predict progression in MCI has been explored in terms of clinical, neuropsychological and biological parameters (Bozoki, Giordani et al. 2001; Chertkow 2002). Although promising, none of these to date have produced a simple, reliable, and inexpensive predictive measure.

This reinforces the notion that additional information must be collected at baseline and follow-up, using other independent or orthogonal biomarker(s) such as neuroimaging. We have already discussed the limitations of functional and longitudinal structural imaging; we will focus our attention in this section to available structural techniques that are applicable to cross-sectional MR data.

Volumetry of the HC and surrounding structures is the most-often used approach in research and clinical practice. It has been shown to suffer from significant intra (0.91-0.95) and inter-rater (0.83-0.91) variability in a hippocampal segmentation study in normal controls by Pruessner *et al.* (Pruessner, Li et al. 2000). This variability is likely to increase in MCI and AD as boundaries shift due to atrophy and become less clearly delineated as intensity changes. This added variability in segmentation measurements, or volumetric noise, will reduce the classification accuracy of any methodology based on such information.

By contrast, the technique presented in this article is completely automated and therefore completely reproducible. Further, it is important to note that the technique requires no manual intervention other than the definition of the VOI, which is done once and based on stereotaxic coordinates on the reference image.

This is a significant reduction of required resources when comparing to volumetry or other expert-based techniques that may require 90-120 minutes per hippocampus for manual segmentation. This independence from user interaction also ensures that the process remains objective.

One might wonder if this classification procedure is overdetermined when using many eigencoordinates for classification, and may risk over-learning of the data, the independence of the reference group from the training set and the independence of the training set from the test set in the multiple trials both ensure that such will not be the case. In such a setting, the cross-validation procedure makes the most of the limited data (24 subjects).

The increased accuracy of this method when compared to single-scan volumetry can readily be explained when one considers the difference in informational value of the approaches. Wherease the current technique uses 4 X 300000 voxels, processed and expressed in terms of 538 principal components coordinates, HC volumetry compresses 3000 to 4000 voxels in a single scalar measurement, expressed in mm^3. Clearly, additional important information is embedded in voxels outside of HC but within the MTL, included in our analysis in the form of intensity and shape covariances.

Other cross-sectional structural imaging techniques have been suggested for the differentiation of AD from normal aging and/or MCI. Cortical thickness, known to be affected in AD (Shefer 1973; Duyckaerts, Hauw et al. 1985), can be estimated via

automated analysis of in vivo MRI. Its successful application to AD research has so far been limited to the differentiation of normal aging from AD (Lerch, Pruessner et al. 2005); it has not yet been applied to the successful prediction of MCI progression to AD. In order to detect cortical thickness changes, high-resolution scans are required (at least with $1mm$ isotropic acquisition); conversely, the technique proposed in this article is shown to be accurate even when using low resolution MRI (slice thickness 5mm).

Assessment of the covariations in intensity and shape embedded in the most discriminating principal components should yield important information regarding the pathophysiology of the disease. To this effect we have imaged the three most discriminating eigenvectors by expressing the difference between centers of the progressing and non-progressing groups in the image domain, along each of these principal components. While there were on average 10 discriminating functions for our validation trials, we chose to image the first three most discriminating principal components for the sake of clarity. The resulting image (see Figure 39) shows covarying areas of maximal difference between the groups that drive the classification.

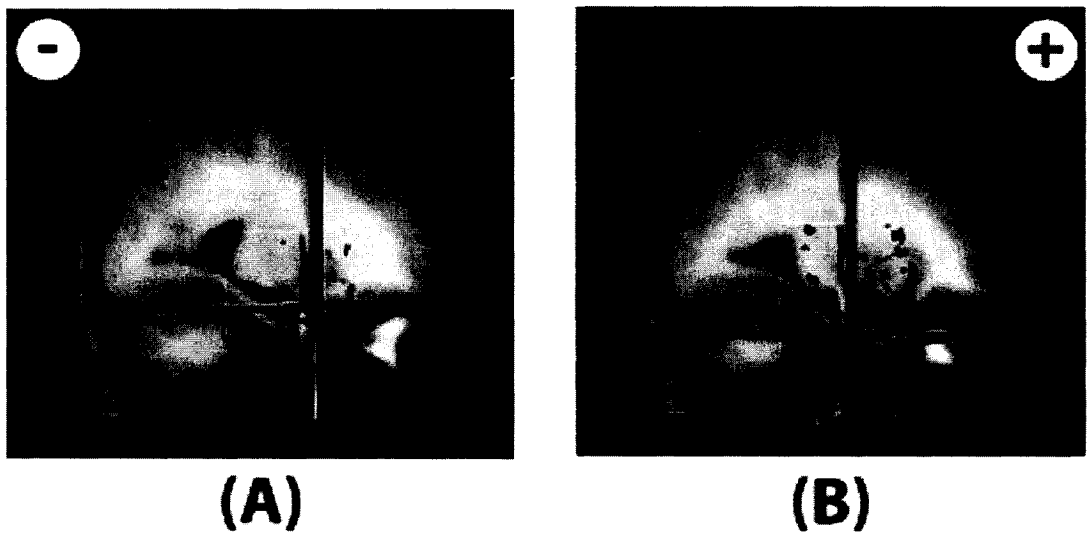

Figure 39 - Areas of discriminating intensity covariances in right medial temporal lobe

Eigendistance (see Figure 37) between centers of the two populations along the three most discriminating PCs of Figure 38 have been expressed back in the image domain (PC55 in blue, PC96 in red and PC64 in green). All three eigenvectors were modeling intensity variations in the right MTL. Our results therefore suggest that the degree of pathological impact in the right MTL is indicative of future progression to AD.

Of primary interest is the fact that all three PCs were modeling intensity variations in the right MTL. This does not imply that intensity changes are not apparent in the left MTL but rather, that they do not have strong discriminating power in separating progressors vs non-progressors. Accepting our hypothesis that pathology-related microscopic changes affect the T1w signal, this finding would imply that the main difference between the two groups is therefore related to the degree of disease spread in the right MTL. A second possible conclusion is that discriminative intensity variations will occur before discriminative shape changes. The latter is supported by the fact that the degree of atrophy of MTL structures is smaller in MCI

than in AD patients, when compared to normal controls (Pennanen, Kivipelto et al. 2004; Thompson, Hayashi et al. 2004).

Differences in gray and white matter in the right MTL have also been reported in the literature since such intensity changes are reflected in decreases in gray matter concentrations. VBM studies have shown patterns of differences primarily in the right MTL when looking at MCI vs AD (Karas, Scheltens et al. 2004; Pennanen, Testa et al. 2005) as well as, and more importantly, in progressor vs non-progressor MCI patients (Chetelat, Landeau et al. 2005).

As mentioned, discriminative changes in shape should follow those in intensity, and indeed there have been reports of the discriminatory ability of shape measures in AD, MCI and normal aging differentiation studies (Csernansky, Wang et al. 2000; Duchesne, Pruessner et al. 2005) but not within an MCI population as it relates to cognitive decline.

Chetelat and colleagues report regions of significant greater GM loss in converters relative to non-converters in the hippocampal area, inferior and middle temporal gyrus, posterior cingulate, and precuneus (Chetelat, Landeau et al. 2005). This pattern is echoed by areas of intensity covariances shown in Figure 39. Areas of difference between the two groups are primarily located in the HC, PHC and medial occipito-temporal gyrus. It is not clear if these intensity differences are indicative of GM or WM atrophy, as this cannot be determined by this methodology; more analysis is therefore required.

Overall, we speculate that these intensity differences result from advanced extra-cellular plaque formation, neurofibrillary tangles accumulation and other pathological processes in the progressors.

Conclusion

We report results from a MRI analysis technique that achieved high sensitivity and specificity in predicting MCI progression or non-progression to AD with 86% accuracy in the cohort studied on average 2.6 years before clinical diagnosis. We found that right MTL intensity changes were most discriminative between the two groups. The importance of the right MTL in the prognosis of MCI and early detection of AD is therefore supported. Clinically, this methodology, which does not rely on segmentation and requires no user input, holds promise for the development of an inexpensive, reliable and safe biomarker to be used in predicting progression in MCI individuals.

Chapter 8

General Discussion and Conclusion

General discussion

The purpose of this thesis was to develop and validate novel, automated image processing techniques that would answer important clinical problems in neurological diseases. Two methods have been proposed: one based on the appearance (intensity and shape features), and a second, on grey and white matter voxel-wise characteristics. Both methods were designed to be volume-centric, data-driven, and cross-sectional.

Chapter 4 was instrumental in presenting the ABC methodology and its application to the difficult task of lateralization of seizure focus in temporal lobe epilepsy. While MRI is not a sensitive tool for measuring the temporo-spatial pattern of seizure onset, it is fit for the task of assessing its impact on the morphology within the

medial temporal lobe. This, in turn, seems sufficient for resolving the clinical problem of lateralization.

The ABC technique was as accurate as expert-based volumetry of the hippocampus in separating patients with or without hippocampal atrophy (127/127), as well as lateralizing seizure focus in subjects with hippocampal atrophy (80/80). It also achieved perfect accuracy (47/47) at determining the seizure focus in patients without hippocampal atrophy, well above volumetry (equal to chance) in this respect. Finally, it achieved 96% accuracy in lateralizing seizure focus when given no *a priori* information about hippocampal volume; this is a major improvement over the 81% accuracy for expert-based volumetry in the same cohort. These results suggest that MRI analysis could replicate the gold standard measurement (clinical evaluation, extensive video-EEG monitoring, and in some cases surgically implanted EEG), while offering considerable reductions in time, resources, and patient discomfort.

In Chapter 5 the second, VBM-based methodology was used to increase our ability at predicting the future clinical status of patients undergoing SAH, from three to four out of five patients when added to the clinical information available (age of onset, duration of epilepsy, contra and ipsi lateral hippocampal volumes). It also fared better in specificity (90.5% *vs.* 75%), slightly better in accuracy (78.9% *vs.* 74.1%), but poorer for sensitivity (50% *vs.* 72%) than a similar study conducted by Antel et al. (Antel, Li et al. 2002) using MRI and MRS imaging. Contrary to the

latter, the VBC method relies on standard T1w imaging, whereas MRS demands significant expertise and long acquisition times.

This study also underlined the fact that there appears to be extra-hippocampal GM and WM differences between patients with negative *vs* positive outcome; this finding should be further explored. It would imply that some of the candidates undergoing SAH have more extensive pathology-related damage; in these cases, more extensive pre-surgical evaluation would be required and possibly the resection target modified to suit their particular condition.

Chapter 6 was key in assessing the robustness of the method to a change in the signal brought about by the application to a completely different pathology. The ABC methodology was proven to be accurate in differentiating AD and MCI from normal aging individuals, on the basis of a single MR scan, with a mean accuracy for individual classification of 90.9%, mean sensitivity to AD+MCI of 76.2% and mean specificity for NA of 97.3%. This is to be compared to the ability of hippocampal and amygdala volumetry at the same task, with a mean accuracy of 82.1%, mean sensitivity to AD+MCI 81.2%, and mean specificity to NA of 85.9% within that cohort.

This has immediate clinical implications for memory clinics; for example, the high specificity of this MRI-based technique could offer a reliable and objective mean of reducing the number of false positives, given inconclusive neuropsychological assessments. It also enlarges the scope of the technique potentially for other

neurological diseases, as these results add further credence to support our hypothesis that intensity and shape changes can be used for discriminative tasks. The surprising find that ABC seems also robust to signal differences due to acquisition hardware bodes well for the future application of this technology to multi-site data, which is routinely done in large scale clinical studies and pharmaceutical trials.

Chapter 7 demonstrated that the ABC technique with MR data could be a potent imaging biomarker, by accurately predicting future clinical state in a cohort of MCI patients. The search for such a biomarker is at the center of MCI research in a wide variety of areas: neuropsychology, proteomics and genomics, and neuroimaging; results such as those confirm the ability of MRI measurements to act as a surrogate indicators for the disease.

Further, the resulting evidence from the experiments in that manuscript demonstrates that the right MTL holds the key information to differentiate or predict future clinical status, and will also be useful for the design of new techniques or analysis of other structural and functional data.

Finally, this study highlighted the fact that the ABC methodology is robust to noise (from movement artifacts) and reduced information found in clinical thick-slice data. Some would imply that this essentially indicates that a structure-centric approach could be used rather than the generic volume of interest. This claim is not necessarily true. First, the accuracy of HC and AG volumetry was lower than the

VOI-centered ABC approach, and by all accounts those are privileged structures in AD research. Second, *a priori* knowledge of the pathophysiological processes of the disease are required in order to select the structure(s) of interest within the convoluted brain morphology. The ABC methodology requires no such preliminary exploratory work. On the other hand, a sensitivity analysis to a change in VOI parameters should be completed; while its extent may remain the same, different sampling schemes could be investigated.

Future work

In fact, there are probably more open questions to be answered now that the techniques have proven successful than at the start of this thesis.

The location of the VOI should be reexamined. One practical consideration when selecting the left or the right MTL was the available memory to perform the necessary computations. The constant upgrading of hardware capabilities make even larger, less specific VOIs possible, which could potentially encompass most sub-cortical structures.

The extension or relocation of the VOI is directly linked to the application. There are a number of other neurological disorders in which intensity and/or shape differences exists, and therefore that could benefit from the ABC technique: schizophrenia and depression are but two examples. One attempt is currently under way for the study of movement disorders, with a VOI centered on the thamalus and

basal ganglia. One could argue that a single model should be done for the whole brain; while possible, the extensive, normal cortical variability may mask out the potential pathological signal.

In any case, the technique should also be tested using a VOI located in a brain region where the a priori hypothesis shows that the pathological effect is null or extremely limited (e.g. occipital lobe in AD); this would be an interesting confirmatory experience to validate the sensitivity of the method.

Further sensitivity analysis of the ABC methodology would include the different parameters of registration (see Robbins et al. (Robbins, Evans et al. 2004)) and intensity normalization. Likewise, the effect of Gaussian blurring kernel size, as well as parameters involved in the tissue classification, should be exhaustively verified in further studies using the VBC methodology.

The same technique is not only scalable in terms of the size of the feature vector, but also to its content. So far the applications have been mono-modality. Clearly, other imaging (T2, PD, PET, SPECT) and/or other descriptors of the images (texture vectors) can be incorporated in the modeling and used for discrimination.

Modeling itself could be revisited. The underlying assumption in both methodologies was that linear models (of intensity and shape covariances, or large scale GM or WM differences) would be sufficient for the discrimination task. It is to be expected that many nonlinear processes result from the pathology; accounting for those nonlinearities using other modeling techniques (nonlinear PCA, for example) could

potentially reduce even further the discriminatory space.

Most of the experiments in these manuscripts confronted MR analysis against other clinically oriented variables. Reality dictates that all available information should be combined in order to maximize the benefit to the patient. Of course, this was done purposefully within these articles to single out the increase in classification or prediction accuracy that could be attributed solely to the MR-based methodologies. In the future, all available information should be combined in a proper statistical framework to maximize the accuracy of the classifier. It is likely that other approaches, e.g. volumetry, redundantly explain some aspects of the underlying pathological process, such as local volume change. In such cases, a stepwise process would reject the redundant features and select only the most discriminating one. Careful retrospective analysis of the discriminant variables would yield the appropriate cocktail of measurements commensurate to the task at hand.

The road ahead for research is therefore centered on converting these methodologies to nonlinear modeling; to combine clinical variables with MRI information; and to extend the analysis to multi-modality data. As a first step, this could entail using co-registered T2 and PD images, and if possible PET imaging (e.g. PIB compound in AD) or inter-ictal SPECT (e.g. Technetium-99m HMPAO in TLE). It remains that many centers will not have access to such resources, lacking the facilities or the expertise and may have only access to simple MR sequences. One development goal should be therefore to tie together the methodologies set

forth in this thesis into an easy-to-use package, in order for users to incorporate them easily into current clinical practice.

Finally, much work remains in the biological interpretation of the discriminative intensity and shape information embedded in the eigenvectors. This step is necessary, it is realized, for acceptance of the methodologies in the wider clinical community. As an example, the corroboration of findings on T1w MRI – or other MR sequences – with functional data, such as PET images with PIB, would serve to bridge the gap between intensity differences seen on MRI in MCI progressors *vs* non-progressors and the microscropic pathological advances in Aβ load.

Conclusion

The goal of this thesis was to present the development and validation of two novel methodologies for automated analysis of MRI, to be used as aid to diagnosis in clinically relevant neurological problems. The four studies presented have shown that these techniques can lateralize the seizure focus and predict surgical outcome in temporal lobe epilepsy, differentiate AD and MCI from normal aging, and predict future clinical status for MCI patients. The potential for future studies and development seems clear, and should aid clinicians in making important diagnostic decisions regarding patient management and care.

References

(APA), A. P. A. (1994). Diagnostic and Statistical Manual of Mental Disorders, 4th ed.: DSM-IV. Washington, D.C.

Abosch, A., N. Bernasconi, et al. (2002). "Factors predictive of suboptimal seizure control following selective amygdalohippocampectomy." J Neurosurg **97**(5): 1142-51.

Antel, S. B., L. M. Li, et al. (2002). "Predicting surgical outcome in temporal lobe epilepsy patients using MRI and MRSI." Neurology **58**(10): 1505-12.

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

Ashburner, J. and K. J. Friston (2001). "Why Voxel-based Morphometry Should be Used." NeuroImage **14**: 1238-1243.

Barber, R., P. Scheltens, et al. (1999). "White matter lesions on magnetic resonance imaging in dementia with Lewy bodies, Alzheimer's disease, vascular dementia, and normal aging." J Neurol Neurosurg Psychiatry **67**(1): 66-72.

Bastos Leite, A. J., P. Scheltens, et al. (2004). "Pathological aging of the brain: an overview." Top Magn Reson Imaging **15**(6): 369-89.

Bautista, R. E., M. A. Cobbs, et al. (1999). "Prediction of surgical outcome by interictal epileptiform abnormalities during intracranial EEG monitoring in patients with extrahippocampal seizures." Epilepsia **40**(7): 880-90.

Berg, A. T., T. Walczak, et al. (1998). "Multivariable prediction of seizure outcome one year after resective epilepsy surgery: development of a model with independent validation." Epilepsy Res **29**(3): 185-94.

Bernasconi, A., S. B. Antel, et al. (2001). "Texture analysis and morphological processing of magnetic resonance imaging assist detection of focal cortical dysplasia in extra-temporal partial epilepsy." Ann Neurol **49**(6): 770-5.

Bernasconi, A., N. Bernasconi, et al. (2000). "T2 relaxometry can lateralize mesial temporal lobe epilepsy in patients with normal MRI." Neuroimage **12**(6): 739-46.

Bernasconi, N., A. Bernasconi, et al. (2003). "Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region." Brain **126**(Pt 2): 462-9.

Bernasconi, N., S. Duchesne, et al. (2004). "Whole-brain voxel-based statistical analysis of gray matter and white matter in temporal lobe epilepsy." Neuroimage **23**(2): 717-23.

Blennow, K. and H. Hampel (2003). "CSF markers for incipient Alzheimer's disease." Lancet Neurol **2**(10): 605-13.

Bonilha, L., E. Kobayashi, et al. (2004). "Value of extent of hippocampal resection in the surgical treatment of temporal lobe epilepsy." Arq Neuropsiquiatr **62**(1): 15-20.

Bookstein, F. L. (1984). "A statistical method for biological shape comparisons." J Theor Biol **107**(3): 475-520.

Bookstein, F. L. (2001). ""Voxel-Based Morphometry" Should Not Be Used with Imperfectly Registered Images." NeuroImage **14**: 1454-1462.

Bouix, S., J. C. Pruessner, et al. (2005). "Hippocampal shape analysis using medial surfaces." Neuroimage **25**(4): 1077-89.

Bozoki, A., B. Giordani, et al. (2001). "Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss." Arch Neurol **58**(3): 411-6.

Briellmann, R. S., R. M. Kalnins, et al. (2002). "Hippocampal pathology in refractory temporal lobe epilepsy: T2-weighted signal change reflects dentate gliosis." Neurology **58**(2): 265-71.

Cachia, A., J. F. Mangin, et al. (2003). "A generic framework for the parcellation of the cortical surface into gyri using geodesic Voronoi diagrams." Med Image Anal **7**(4): 403-16.

Carr, D. B., A. Goate, et al. (1997). "Current concepts in the pathogenesis of Alzheimer's disease." Am J Med **103**(3A): 3S-10S.

Cascino, G. D. (1995). "Clinical correlations with hippocampal atrophy." Magn Reson Imaging **13**(8): 1133-6.

Cascino, G. D., C. R. Jack, Jr., et al. (1991). "Magnetic resonance imaging-based volume studies in temporal lobe epilepsy: pathological correlations." Ann Neurol **30**(1): 31-6.

Cassol, E., J. P. Ranjeva, et al. (2004). "Diffusion tensor imaging in multiple sclerosis: a tool for monitoring changes in normal-appearing white matter." Mult Scler **10**(2): 188-96.

Cavalieri, B. F. (1635). Geometria Indivisibilibus Continuorum Nova Quadam Ratione Promota ("A Method for the Determination of a New Geometry of Continuous Indivisibles"). Bologna, University of Bologna.

Cendes, F., F. Andermann, et al. (1993). "MRI volumetric measurement of amygdala and hippocampus in temporal lobe epilepsy." Neurology **43**(4): 719-25.

Cheon, J. E., K. H. Chang, et al. (1998). "MR of hippocampal sclerosis: comparison of qualitative and quantitative assessments." AJNR Am J Neuroradiol **19**(3): 465-8.

Chertkow, H. (2002). "Mild cognitive impairment." Curr Opin Neurol **15**(4): 401-7.

Chertkow, H., C. Bocti, et al. (2005). "MCI progressors and non-progressors: Neuropsychological profiles fail to guarantee the prognosis." Alzheimer's & Dementia: The Journal of the Alzheimer's Association **1**(1): 21.

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

Chetelat, G., B. Landeau, et al. (2005). "Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: A longitudinal MRI study." Neuroimage **27**(4): 934-46.

Choi, J. Y., S. J. Kim, et al. (2003). "Extratemporal hypometabolism on FDG PET in temporal lobe epilepsy as a predictor of seizure outcome after temporal lobectomy." Eur J Nucl Med Mol Imaging **30**(4): 581-7.

Chong, M. S. and S. Sahadevan (2005). "Preclinical Alzheimer's disease: diagnosis and prediction of progression." Lancet Neurol **4**(9): 576-9.

Christensen, G. E., S. C. Joshi, et al. (1997). "Volumetric Transformation of Brain Anatomy." IEEE Transactions on Medical Imaging **16**: 864--877.

Chung, M. K., K. J. Worsley, et al. (2001). "A unified statistical approach to deformation-based morphometry." NeuroImage **14**(3): 595-606.

Chung, M. K., K. J. Worsley, et al. (2003). "Deformation-based surface morphometry applied to gray matter deformation." Neuroimage **18**(2): 198-213.

CIHI (2003). Medical Imaging In Canada. Ottawa, ON, Canadian Institute for Health Information: 114.

Collins, D. L. and A. C. Evans (1997). "ANIMAL: Validation and Applications of Non-linear Registration Based Segmentation." International Journal of Pattern Recognition and Artificial Intelligence **11**(8): 1271--1294.

Collins, D. L., C. J. Holmes, et al. (1995). "Automatic 3D Model-Based Neuroanatomical Segmentation." Human Brain Mapping **3**: 190--208.

Collins, D. L., P. Neelin, et al. (1994). "Automatic 3D Intersubject Registration of MR Volumetric Data in Standardized Talairach Space." Journal of Computer Assisted Tomography **18**: 192--205.

Cootes, T. F., D. H. Cooper, et al. (1991). A Trainable Method of Parametric Shape Description. British Machine Vision Conference, Springer-Verlag.

Cootes, T. F., G. J. Edwards, et al. (1998). Active Appearance Models. Proceedings of the European Conference on Computer Vision, Verlag

ECCV: 484-498.

Cootes, T. F., G. J. Edwards, et al. (2001). " Active Appearance Models." IEEE Transactions on Pattern Analysis and Machine Intelligence **23:6**: 681-685.

Cootes, T. F. and C. J. Taylor (2000). Statistical Models of Appearance for Computer Vision, University of Manchester.

Cootes, T. F., C. J. Taylor, et al. (1995). " Active Shape Models, Their Training and Application." Computer Vision and Image Understanding **61**: 38--59.

Csernansky, J. G., M. K. Schindler, et al. (2004). "Abnormalities of thalamic volume and shape in schizophrenia." Am J Psychiatry **161**(5): 896-902.

Csernansky, J. G., L. Wang, et al. (2000). "Early DAT is distinguished from aging by high-dimensional mapping of the hippocampus. Dementia of the Alzheimer type." Neurology **55**(11): 1636-43.

Cummings, B. J., C. J. Pike, et al. (1996). "Beta-amyloid deposition and other measures of neuropathology predict cognitive status in Alzheimer's disease." Neurobiol Aging **17**(6): 921-33.

Daly, E., D. Zaitchik, et al. (2000). "Predicting conversion to Alzheimer disease using standardized clinical information." Arch Neurol **57**(5): 675-80.

Dameron, O., B. Gibaud, et al. (2004). "Numeric and symbolic knowledge representation of cerebral cortex anatomy: methods and preliminary results." Surg Radiol Anat **26**(3): 191-7.

Davatzikos, C. (2004). "Why voxel-based morphometric analysis should be used with great caution when characterizing group differences." Neuroimage **23**(1): 17-20.

Davatzikos, C., A. Genc, et al. (2001). "Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy." Neuroimage **14**(6): 1361-9.

DeCarli, C. (2003). "Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment." Lancet Neurol **2**(1): 15-21.

deToledo-Morrell, L., T. R. Stoub, et al. (2004). "MRI-derived entorhinal volume is a good predictor of conversion from MCI to AD." Neurobiol Aging **25**(9): 1197-203.

Drury, H. A., D. C. Van Essen, et al. (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., N. Schuff, et al. (2001). "Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease." J Neurol Neurosurg Psychiatry **71**(4): 441-7.

Du, A. T., N. Schuff, et al. (2003). "Atrophy rates of entorhinal cortex in AD and normal aging." Neurology **60**(3): 481-6.

Dubois, B. and M. L. Albert (2004). "Amnestic MCI or prodromal Alzheimer's disease?" Lancet Neurol **3**(4): 246-8.

Duchesne, S., N. Bernasconi, et al. (2002). On the classification of temporal lobe epilepsy using MR image appearance. 16th International Conference on Pattern Recognition, Quebec City, QC.

Duchesne, S., N. Bernasconi, et al. (2003). TLE lateralization using MR image intensity and registration features. Medical Image Computing and Computer-Assisted Intervention, Montreal, QC, Springer-Verlag.

Duchesne, S., N. Bernasconi, et al. (2004). TLE surgery outcome prediction. Medical Image Computing and Computer-Assisted Intervention, St-Malo, FR, Springer-Verlag.

Duchesne, S., N. Bernasconi, et al. (2006). "MR-based neurological disease classification methodology: Application to lateralization of seizure focus in temporal lobe epilepsy." Neuroimage **29**(2): 557-66.

Duchesne, S., K. De Sousa, et al. (2005). Predicting MCI progression to AD via automated analysis of T1 weighted MR image intensity. Alzheimer's Association International Conference on Prevention of Dementia, Washington, D.C.

Duchesne, S., J. Pruessner, et al. (2002). "Appearance-based segmentation of medial temporal lobe structures." Neuroimage **17**(2): 515-31.

Duchesne, S., J. Pruessner, et al. (2005). Successful AD and MCI differentiation from normal aging via automated analysis of MR image features. Alzheimer's Association International Conference on Prevention of Dementia, Washington, D.C.

Duda, R. O., P. E. Hart, et al. (2001). Pattern Classification, Wiley-Interscience.

Dupont, S., F. Semah, et al. (2000). "Accurate prediction of postoperative outcome in mesial temporal lobe epilepsy: a study using positron emission tomography with 18fluorodeoxyglucose." Arch Neurol **57**(9): 1331-6.

Duyckaerts, C., J. J. Hauw, et al. (1985). "Cortical atrophy in senile dementia of the Alzheimer type is mainly due to a decrease in cortical length." Acta Neuropathol (Berl) **66**(1): 72-4.

Engel, J., Jr., P. C. Van Ness, et al. (1993). Outcome with respect to epileptic seizures. Surgical treatment of the epilepsies. J. Engel, Jr. New York, Raven Press: 609-621.

Falconer, M. A., E. A. Serafetinides, et al. (1964). "Etiology and Pathogenesis of Temporal Lobe Epilepsy." Arch Neurol **10**: 233-48.

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

Flicker, C., S. H. Ferris, et al. (1991). "Mild cognitive impairment in the elderly: predictors of dementia." Neurology **41**(7): 1006-9.

Folstein, M. F., S. E. Folstein, et al. (1975). ""Mini-mental state". A practical method for grading the cognitive state of patients for the clinician." J Psychiatr Res **12**(3): 189-98.

Folstein, M. F., S. E. Folstein, et al. (1975). ""Mini-mental state": A practical method for grading the cognitive state of patients for the clinician." Journal of Psychiatric Research **12**(3): 189-198.

Fox, N. C. and P. A. Freeborough (1997). "Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease." J Magn Reson Imaging **7**(6): 1069-75.

Fox, N. C. and J. M. Schott (2004). "Imaging cerebral atrophy: normal ageing to Alzheimer's disease." Lancet **363**(9406): 392-4.

Freeborough, P. A. and N. C. Fox (1998). "MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease." IEEE Trans Med Imaging **17**(3): 475-9.

Freedman, M. and L. Leach (1994). Clock drawing: A neuropsychological analysis. New York, NY, USA, Oxford University Press.

Frisoni, G. B., R. Rossi, et al. (2002). "The radial width of the temporal horn in mild cognitive impairment." J Neuroimaging **12**(4): 351-4.

Giannakopoulos, P., F. R. Herrmann, et al. (2003). "Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease." Neurology **60**(9): 1495-500.

Gilliam, F., E. Faught, et al. (2000). "Predictive value of MRI-identified mesial temporal sclerosis for surgical outcome in temporal lobe epilepsy: an intent-to-treat analysis." Epilepsia **41**(8): 963-6.

Golland, P., W. E. Grimson, et al. (2005). "Detection and analysis of statistical differences in anatomical shape." Med Image Anal **9**(1): 69-86.

Good, C. D., I. S. Johnsrude, et al. (2001). "A voxel-based morphometric study of ageing in 465 normal adult human brains." Neuroimage **14**(1 Pt 1): 21-36.

Group, T. C. S. o. H. a. A. W. (2000). "The incidence of dementia in Canada." Neurology **55**: 66-73.

Grundman, M., D. Sencakova, et al. (2002). "Brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial." J Mol Neurosci **19**(1-2): 23-7.

Hachinski, V. C., L. D. Iliff, et al. (1975). "Cerebral blood flow in dementia." Arch Neurol **32**(9): 632-7.

Hanyu, H., T. Asano, et al. (1999). "Diffusion-weighted and magnetization transfer imaging of the corpus callosum in Alzheimer's disease." J Neurol Sci **167**(1): 37-44.

Herholz, K. (2003). "PET studies in dementia." Ann Nucl Med **17**(2): 79-89.

Hill, D. L., P. G. Batchelor, et al. (2001). "Medical image registration." Phys Med Biol **46**(3): R1-45.

Hogan, R. E., K. E. Mark, et al. (2000). "Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients." Radiology **216**(1): 291-7.

Hughes, C., L. Berg, et al. (1982). "A new clinical scale for the staging of dementia." British Journal of Psychiatry **140**: 566-572.

J. R. Petrella, R. E. C., P.M. Doraiswamy (2003). "Neuroimaging and Early Diagnosis of Alzheimer Disease: A Look to the Future." Radiology **226**(2): 315-336.

Jack, C. R., Jr., F. W. Sharbrough, et al. (1992). "Magnetic resonance image-based hippocampal volumetry: correlation with outcome after temporal lobectomy." Ann Neurol **31**(2): 138-46.

Jack, C. R., Jr., M. Slomkowski, et al. (2003). "MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD." Neurology **60**(2): 253-60.

Jack, C. R., Jr., W. H. Theodore, et al. (1995). "MRI-based hippocampal volumetrics: data acquisition, normal ranges, and optimal protocol." Magn Reson Imaging **13**(8): 1057-64.

Janke, A. L., G. de Zubicaray, et al. (2001). "4D deformation modeling of cortical disease progression in Alzheimer's dementia." Magn Reson Med **46**(4): 661-6.

Johnson, S. C., A. J. Saykin, et al. (2000). "The relationship between fMRI activation and cerebral atrophy: comparison of normal aging and alzheimer disease." Neuroimage **11**(3): 179-87.

Johnson, S. C., T. W. Schmitz, et al. (2005). "Activation of brain regions vulnerable to Alzheimer's disease: The effect of mild cognitive impairment." Neurobiol Aging.

Joshi, S., S. Pizer, et al. (2001). Multi-scale 3D Deformable Model Segmentation Based on Medial Description. Proceedings of the Information Processing in Medical Imaging Conference, Springer

IPMI: 64-77.

Joshi, S., S. Pizer, et al. (2002). "Multiscale deformable model segmentation and statistical shape analysis using medial descriptions." IEEE Trans Med Imaging **21**(5): 538-50.

Kabani, N. J., J. G. Sled, et al. (2002). "Magnetization transfer ratio in mild cognitive impairment and dementia of Alzheimer's type." Neuroimage **15**(3): 604-10.

Kantarci, K., C. R. Jack, Jr., et al. (2000). "Regional metabolic patterns in mild cognitive impairment and Alzheimer's disease: A 1H MRS study." Neurology **55**(2): 210-7.

Kantarci, K., C. R. Jack, Jr., et al. (2001). "Mild cognitive impairment and Alzheimer disease: regional diffusivity of water." Radiology **219**(1): 101-7.

Kantarci, K., G. E. Smith, et al. (2002). "1H magnetic resonance spectroscopy, cognitive function, and apolipoprotein E genotype in normal aging, mild cognitive impairment and Alzheimer's disease." J Int Neuropsychol Soc **8**(7): 934-42.

Kaplan, E. F., H. Goodglass, et al. (1978). The Boston Naming Test. Boston, E.Kaplan & H.Goodglass.

Karas, G. B., E. J. Burton, et al. (2003). "A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry." Neuroimage **18**(4): 895-907.

Karas, G. B., P. Scheltens, et al. (2004). "Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease." Neuroimage **23**(2): 708-716.

Keller, S. S., C. E. Mackay, et al. (2002). "Voxel-based morphometric comparison of hippocampal and extrahippocampal abnormalities in patients with left and right hippocampal atrophy." Neuroimage **16**(1): 23-31.

Keller, S. S., M. Wilke, et al. (2004). "Comparison of standard and optimized voxel-based morphometry for analysis of brain changes associated with temporal lobe epilepsy." Neuroimage **23**(3): 860-8.

Khachaturian, Z. S. (1985). "Diagnosis of Alzheimer's Disease." Archives of Neurology **42**: 1097-1105.

Klunk, W. E., H. Engler, et al. (2004). "Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B." Ann Neurol **55**(3): 306-19.

Knopman, D. S. and S. Ryberg (1989). "A verbal memory test with high predictive accuracy for dementia of the Alzheimer type." Archives of Neurology **46**(2): 141-5.

Kobayashi, E., I. Lopes-Cendes, et al. (2001). "Seizure outcome and hippocampal atrophy in familial mesial temporal lobe epilepsy." Neurology **56**(2): 166-72.

Kogure, D., H. Matsuda, et al. (1999). "[Longitudinal evaluation of early dementia of Alzheimer type using brain perfusion SPECT]." Kaku Igaku **36**(2): 91-101.

Korf, E. S., L. O. Wahlund, et al. (2004). "Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment." Neurology **63**(1): 94-100.

Kovalev, V. A., F. Kruggel, et al. (2001). "Three-dimensional texture analysis of MRI brain datasets." IEEE Trans Med Imaging **20**(5): 424-33.

Kuzniecky, R., S. Suggs, et al. (1991). "Lateralization of epileptic foci by magnetic resonance imaging in temporal lobe epilepsy." J Neuroimaging **1**(4): 163-7.

Laakso, M. P., K. Partanen, et al. (1996). "MR T2 relaxometry in Alzheimer's disease and age-associated memory impairment." Neurobiol Aging **17**(4): 535-40.

Lao, Z., D. Shen, et al. (2004). "Morphological classification of brains via high-dimensional shape transformations and machine learning methods." Neuroimage **21**(1): 46-57.

Larsson, H. B., G. J. Barker, et al. (1998). "Nuclear magnetic resonance relaxation in multiple sclerosis." J Neurol Neurosurg Psychiatry **64 Suppl 1**: S70-6.

Le Goualher, G., C. Barillot, et al. (1997). "Modeling cortical sulci with active ribbons." IJPRAI **11**(8): 1295-1315.

Lerch, J. P., J. C. Pruessner, et al. (2005). "Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy." Cereb Cortex **15**(7): 995-1001.

Li, L. M., Z. Caramanos, et al. (2000). "Lateralization of temporal lobe epilepsy (TLE) and discrimination of TLE from extra-TLE using pattern analysis of magnetic resonance spectroscopic and volumetric data." Epilepsia **41**(7): 832-42.

Li, L. M., S. Narayanan, et al. (2000). "Magnetization transfer ratio is unable to lateralize epileptic foci in patients with temporal lobe epilepsy." AJNR Am J Neuroradiol **21**(10): 1853-6.

Liu, Y., L. Teverovskiy, et al. (2004). Discriminative MR Image Feature Analysis for Automatic Schizophrenia and Alzheimer,Äôs Disease Classification.

Logothetis, N. K. and J. Pfeuffer (2004). "On the nature of the BOLD fMRI contrast mechanism." Magn Reson Imaging **22**(10): 1517-31.

Lohmann, G. and D. Y. von Cramon (2000). "Automatic labelling of the human cortical surface using sulcal basins." Med Image Anal **4**(3): 179-88.

MacDonald, D., N. Kabani, et al. (2000). "Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI." Neuroimage **12**(3): 340-56.

Maintz, J. B. and M. A. Viergever (1998). "A survey of medical image registration." Med Image Anal **2**(1): 1-36.

Mangin, J. F., D. Riviere, et al. (2004). "Coordinate-based versus structural approaches to brain image analysis." Artif Intell Med **30**(2): 177-97.

Mazziotta, J. C., A. W. Toga, et al. (1995). "A probabilistic atlas of the human brain: theory and rationale for its development. The International Consortium for Brain Mapping (ICBM)." Neuroimage **2**(2): 89-101.

McKhann, G., D. Drachman, et al. (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.

McKhann, G., D. Drachman, et al. (1984). "Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Health and Human Services Task Force on Alzheimer's Disease." Neurology **34**: 939-944.

McLachlan, R. S. (2001). "Commentary on Epilepsy Surgery in Canada." Can. J. Neurol. Sci. **28**: 4-5.

Mungas, D., B. R. Reed, et al. (2002). "Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease." Neurology **59**(6): 867-73.

Nagy, Z., N. J. Hindley, et al. (1999). "The progression of Alzheimer's disease from limbic regions to the neocortex: clinical, radiological and pathological relationships." Dement Geriatr Cogn Disord **10**(2): 115-20.

Narayanan, S., L. Fu, et al. (1997). "Imaging of axonal damage in multiple sclerosis: spatial distribution of magnetic resonance imaging lesions." Ann Neurol **41**(3): 385-91.

Nasreddine, Z. S., N. A. Phillips, et al. (2005). "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment." J Am Geriatr Soc **53**(4): 695-9.

Ng, K. K., K. W. Tang, et al. (2000). "Brain MRI of hippocampal volumetry in patients with refractory temporal lobe epilepsy." Chin Med J (Engl) **113**(3): 254-6.

Nordberg, A. (2004). "PET imaging of amyloid in Alzheimer's disease." The Lancet Neurology **3**(9): 519.

Olivier, A. (2000). "Transcortical selective amygdalohippocampectomy in temporal lobe epilepsy." Can J Neurol Sci **27 Suppl 1**: S68-76; discussion S92-6.

Olivier, A., M. Alonso-Vanegas, et al. (1996). "Image-guided surgery of epilepsy." Neurosurg Clin N Am **7**(2): 229-43.

Ono, M., S. Kubik, et al. (1990). Atlas of the Cerebral Sulci, Thieme Verlag.

Pennanen, C., M. Kivipelto, et al. (2004). "Hippocampus and entorhinal cortex in mild cognitive impairment and early AD." Neurobiol Aging **25**(3): 303-10.

Pennanen, C., C. Testa, et al. (2005). "A voxel based morphometry study on mild cognitive impairment." J Neurol Neurosurg Psychiatry **76**(1): 11-4.

Petersen, R. C. (2004). "Mild cognitive impairment as a diagnostic entity." J Intern Med **256**(3): 183-94.

Petersen, R. C., R. Doody, et al. (2001). "Current concepts in mild cognitive impairment." Arch Neurol **58**(12): 1985-92.

Petersen, R. C., G. E. Smith, et al. (1995). "Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals [published erratum appears in JAMA 1995 Aug 16;274(7):538]." Journal of the American Medical Association **273**(16): 1274-8.

Petersen, R. C., G. E. Smith, et al. (1999). "Mild cognitive impairment: clinical characterization and outcome [published erratum appears in Arch Neurol 1999 Jun;56(6):760]." Archives of Neurology **56**(3): 303-8.

Petrella, J. R., R. E. Coleman, et al. (2003). "Neuroimaging and early diagnosis of Alzheimer disease: a look to the future." Radiology **226**(2): 315-36.

Pham, D. L., C. Xu, et al. (2000). "Current methods in medical image segmentation." Annual Review of Biomedical Engineering **2**: 315-337.

Poliakov, A. V., K. P. Hinshaw, et al. (1999). "Integration and visualization of multimodality brain data for language mapping." Proc AMIA Symp: 349-53.

Pruessner, J. C., L. M. Li, et al. (2000). "Volumetry of Hippocampus and Amygdala with High-Resolution MRI and Three-Dimensional Analysis Software: Minimizing the Discrepancies between Laboratories." Cerebral Cortex **10**: 433-442.

Radhakrishnan, K., E. L. So, et al. (1998). "Predictors of outcome of anterior temporal lobectomy for intractable epilepsy: a multivariate study." Neurology **51**(2): 465-71.

Risse, S. C., M. A. Raskind, et al. (1990). "Neuropathological findings in patients with clinical diagnoses of probable Alzheimer's disease." Am J Psychiatry **147**(2): 168-72.

Robbins, S., A. C. Evans, et al. (2004). "Tuning and comparing spatial normalization methods." Med Image Anal **8**(3): 311-23.

Salmond, C. H., J. Ashburner, et al. (2002). "Distributional assumptions in voxel-based morphometry." Neuroimage **17**(2): 1027-30.

Sawrie, S. M., R. C. Martin, et al. (1998). "Contribution of neuropsychological data to the prediction of temporal lobe epilepsy surgery outcome." Epilepsia **39**(3): 319-25.

Shefer, V. F. (1973). "Absolute number of neurons and thickness of the cerebral cortex during aging, senile and vascular dementia, and Pick's and Alzheimer's diseases." Neurosci Behav Physiol **6**(4): 319-24.

Shen, D. and C. Davatzikos (2003). "Very High-Resolution Morphometry Using Mass-Preserving Deformations and HAMMER Elastic Registration." NeuroImage **18**: 28-41.

Sled, J. G., A. P. Zijdenbos, et al. (1998). "A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data." IEEE Transactions on Medical Imaging **17**: 87-97.

Spencer, S. S., G. McCarthy, et al. (1993). "Diagnosis of medial temporal lobe seizure onset: relative specificity and sensitivity of quantitative MRI." Neurology **43**(10): 2117-24.

Srivastava, S., F. Maes, et al. (2005). "Feature-based statistical analysis of structural MR data for automatic detection of focal cortical dysplastic lesions." NeuroImage **27**(2): 253.

Stahl, R., O. Dietrich, et al. (2003). "[Assessment of axonal degeneration on Alzheimer's disease with diffusion tensor MRI]." Radiologe **43**(7): 566-75.

Styner, M. and G. Gerig (2001). Medial Models incorporating object variability for 3D shape analysis. Proceedings of the Information Processing in Medical Imaging Conference, Springer

IPMI: 502-516.

Styner, M., G. Gerig, et al. (2003). "Statistical shape analysis of neuroanatomical structures based on medial models." Med Image Anal **7**(3): 207-20.

Talairach, J. and P. Tournoux (1988). Co-planar stereotaxic atlas of the human brain: an approach to medical cerebral imaging. Stuttgart, Thieme.

Thal, D. R., U. Rub, et al. (2002). "Phases of A beta-deposition in the human brain and its relevance for the development of AD." Neurology **58**(12): 1791-800.

Thal, D. R., U. Rub, et al. (2000). "Sequence of Abeta-protein deposition in the human medial temporal lobe." J Neuropathol Exp Neurol **59**(8): 733-48.

Thirion, J. P. (1996). Non-rigid Matching Using Demons. Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition, CVPR: 245-251.

Thomaz, C. E., J. P. Boardman, et al. (2004). Using a Maximum Uncertainty LDA-Based Approach to Classify and Analyse MR Brain Images.

Thompson, P. M., K. M. Hayashi, et al. (2003). "Dynamics of gray matter loss in Alzheimer's disease." J Neurosci **23**(3): 994-1005.

Thompson, P. M., K. M. Hayashi, et al. (2004). "Mapping hippocampal and ventricular change in Alzheimer disease." Neuroimage **22**(4): 1754-66.

Thompson, P. M., C. Schwartz, et al. (1996). "Three-dimensional statistical analysis of sulcal variability in the human brain." J Neurosci **16**(13): 4261-74.

Thompson, P. M. and A. W. Toga (1997). "Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations." Med Image Anal **1**(4): 271-94.

Tofts, P. (2003). Quantitative MRI of the brain: measuring changes caused by disease. Chichester, West Sussex, John Wiley and Sons Ltd.

Tofts, P. S., S. Sisodiya, et al. (1995). "MR magnetization transfer measurements in temporal lobe epilepsy: a preliminary study." AJNR Am J Neuroradiol **16**(9): 1862-3.

Van Essen, D. C., H. A. Drury, et al. (2001). "An integrated software suite for surface-based analyses of cerebral cortex." J Am Med Inform Assoc **8**(5): 443-59.

Vickrey, B. G., R. D. Hays, et al. (1995). "Outcomes in 248 patients who had diagnostic evaluations for epilepsy surgery." Lancet **346**(8988): 1445-9.

Visser, P. J., P. Scheltens, et al. (1999). "Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment." J Neurol **246**(6): 477-85.

Wang, L., S. C. Joshi, et al. (2001). "Statistical Analysis of Hippocampal Asymmetry in Schizophrenia." NeuroImage **14**: 531-545.

Webb, J., A. Guimond, et al. (1999). "Automatic detection of hippocampal atrophy on magnetic resonance images." Magn Reson Imaging **17**(8): 1149-61.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale - Revised. New York, The Psychological Corporation.

Wechsler, D. (1987). Wechsler Memory Scale - Revised. San Antonio, The Psychological Corporation.

Weiner, M., M. Albert, et al. (2005). The Use of MRI and PET for Clinical Diagnosis of Dementia and Investigation of Cognitive Impairment: A Consensus Report, Alzheimer's Association.

Winblad, B., K. Palmer, et al. (2004). "Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment." J Intern Med **256**(3): 240-6.

Woermann, F. G., S. L. Free, et al. (1999). "Voxel-by-voxel comparison of automatically segmented cerebral gray matter--A rater-independent comparison of structural MRI in patients with epilepsy." Neuroimage **10**(4): 373-84.

Worsley, K. J., M. Andermann, et al. (1999). "Detecting Changes in Nonisotropic Images." Human Brain Mapping **8**: 98-101.

Worsley, K. J., S. Marrett, et al. (1996). "A unified statistical approach for determining significant signals in images of cerebral activation." Human Brain Mapping **4**(1): 58-73.

Yesavage, J. A., T. L. Brink, et al. (1983). "Development and validation of a geriatric depression screening scale: A preliminary report." Journal of Psychiatric Research **17**: 37-49.

Zijdenbos, A. P. and B. M. Dawant (1994). "Brain segmentation and white matter lesion detection in MR images." Crit Rev Biomed Eng **22**(5-6): 401-65.

Zijdenbos, A. P., R. Forghani, et al. (2002). "Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis." IEEE Trans Med Imaging **21**(10): 1280-91.

Other Publications

Other publications can be divided into three components: peer-reviewed journal articles, peer-reviewed conference proceedings and short conference abstracts

Journal articles

N. Bernasconi, **S. Duchesne**, et al. "Whole-brain voxel-based statistical analysis of gray matter and white matter in temporal lobe epilepsy", *NeuroImage*, 23 :717-723, 2004

S. Duchesne, J.C. Pruessner, D.L. Collins, "Appearance-based segmentation of medial temporal lobe structures", *NeuroImage*, 17(2):515-531, 2002

Conference proceedings

S. Duchesne et al., « Predicting clinical variable from MRI features : application to MMSE in MCI ». Proc. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, J. Duncan, G. Gerig, Eds., Springer-Verlag, LNCS 3749(1): 392-399, 2005

S. Duchesne et al., "TLE surgery outcome prediction ", Proc. *MICCAI*, C. Barillot, D.R. Haynor, and P. Hellier, Eds., Springer-Verlag, LNCS 3217 :696-702, 2004

S. Duchesne et al, "TLE lateralization using MR image intensity and registration features", *Proc. MICCAI*, T. Peters, R.E. Ellis, Eds., Springer-Verlag, 2879(1):367-374 2003

S. Duchesne *et al*, "On the classification of Temporal Lobe Epilepsy based on MR image appearance", *Proc. International Conference on Pattern Recognition*, D. Laurendeau, Eds., IEEE Computer Society, 1: 520-523,2002

S. Duchesne *et al*, "Apperance-based modeling and segmentation of the hippocampus from MR images", *Proc. 23rd Annual IEEE Engineering in Medicine and Biology Society Conference*, IEEE EMB Society, Istanbul, Turkey, 2001

S. Duchesne *et al*, "Analysis of deformation fields for apperance-based segmentation", *Proc. MICCAI*, W.J. Niessen, M.A. Viergever, Eds., Springer-Verlag, 2208:1189-1190, 2001

Conference abstracts

R. Al Sufayan, **S. Duchesne**, J. Jaffey, N. Duchesne, "Pure Ductal Carcinoma In Situ: Trend towards a Reduced upgrading when using vacuum-assisted devices in ultrasound-guided breast biopsies". Submitted to *American Roentgen Ray Society*, 2005

R. Al Sufayan, **S. Duchesne**, N. Duchesne, "Improving Detection of Pure DCIS at Ultrasound: Review of Imaging Criteria Frequency title ". Submitted to *American Roetgen Ray Society*, 2005

S. Duchesne et al., "Successful AD and MCI differentiation from normal aging via automated analysis of MR image features". *AA Int'l Conf. on Prevention of Dementia*, Washington, D.C., July 2005.

S. Duchesne et al., "Predicting MCI progression to AD via automated analysis of T1 weighted MR image intensity". *AA Int'l Conf. on Prevention of Dementia*, Washington D.C., July 2005.

S. Duchesne et al., "Predicting MCI progression to AD via automated analysis of T1 weighted MR image intensity". *4th Annual Research Day*, RQRV, Montreal, April 2005

N. Bernasconi, **S. Duchesne**, et al. « VBM of grey and white matter in patients with unilateral temporal lobe epilepsy". Proc. Of Am. Epilepsy Society, Boston, MA, 2003

R. Aleong, **S. Duchesne**, *et al*, "Assessment of Adolescent Body Perception: Development of a Novel Adolsescent Body Morphing Tool", *CIHR 2nd Annual Research Award Recipient Symposia*, CIHR, 2003

S. Duchesne et al, "Within-group non-linear registration improves VBM results", *Proc. ISMRM*, ISMRM, Toronto, Canada , 2003

N. Bernasconi, **S. Duchesne** *et al*, "Voxel-based statistical analysis of grey matter and white matter in patients with unilateral temporal lobe epilepsy". *NeuroImage* 19 (Suppl 1):575, 2003

Recurring Abbreviations

Alzheimer's disease	**AD**
Amygdala	**AG**
Degrees of Freedom	**DF**
Electroencephalogram	**EEG**
Gray Matter	**GM**
Hippocampal normal volume	**HNV**
Hippocampus	**HC**
Hippocampal Atrophy	**HA**
Hippocampal Normal Volume	**HNV**
Left seizure focus with Hippocampal Atrophy	**LHA**
Left seizure focus with Normal HC Volume	**LNV**
International Consortium for Brain Mapping	**ICBM**
Magnetic Resonance Imaging	**MRI**
Medial Temporal Lobe	**MTL**
Mild Cognitive Impairment	**MCI**
Mini-Mental State Examination	**MMSE**
Non-Progressors	**NP**
Normal Aging	**NA**
Normal Controls	**NC**
Parahippocampus	**PHC**
Principal Components	**PC**
Principal Components Analysis	**PCA**
Progressors	**P**
Right SF with hippocampal atrophy	**RHA**
Right SF with normal HC volume	**RNV**
Selective Amygdalo-Hippocampectomy	**SAH**
Seizure focus	**SF**
Temporal Lobe Epilepsy	**TLE**
Volume of Interest	**VOI**
Voxel-Based Morphometry	**VBM**
White Matter	**WM**