Improved methods for analyzing MR spectroscopic imaging in order to better understand neurological injury

by

Li Xu

Department of Biomedical Engineering
McGill University, Montreal

December 23, 2005

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

Master of Engineering

Your file Votre référence
ISBN: 978-0-494-25018-1
Our file Notre référence
ISBN: 978-0-494-25018-1

Abstract

This thesis describes a method to analyze proton magnetic resonance spectroscopy imaging (MRSI) of multiple sclerosis (MS) patients. Multiple sclerosis is a chronic disease of the central nervous system (CNS). MRSI can non-invasively measure the metabolites in the brains and is helpful in research of progression of MS. Conventional approaches to analyze MRSI data are either using region-of-interest (ROI) methods or averaging the resonance intensities over the whole study area. This thesis documents an approach to use all reliable resonance intensities in MRSI based on multivariate mixed effect statistical models for repeated measurement. We applied the method in a series of studies and demonstrated that the distribution of brain metabolites was different among MS patients in different disease phases. These studies also showed the correlations between the brain metabolites and clinical data such as disease duration and clinical disability.

Résumé

Cette thèse décrit une méthode pour analyser l'imagerie spectroscopique par résonance magnétique de patients atteints de sclérose en plaques. La sclérose en plaques est une maladie chronique du système nerveux central. L'imagerie spectroscopique par résonance magnétique permet de mesurer de façon non invasive le métabolisme du cerveau et est utile dans la recherche sur la progression de la sclérose en plaques. Les approches conventionnelles pour analyser les données d'imagerie spectroscopique sont soit les méthodes utilisant des régions d'intérêt, soit des méthodes moyennant les intensités de résonance sur toutes la région étudiée. Cette thèse documente une approche qui utilise toutes les intensités de résonance, basée sur des modèles statistiques d'effet mixte multivariable pour mesures répétées. Nous appliquons cette méthode à une série d'études et démontrons que le métabolisme du cerveau sont différents parmi les patients à différentes phases de la maladie. Ces études ont également montrées des corrélations entre le métabolisme du cerveau et des données cliniques, comme la durée de la maladie et les incapacités cliniques.

Table of Contents

Acknowledgements

This thesis would not be possible without the support and help from many people.

First I thank my supervisor, Dr. Douglas Arnold for being supportive in me and my research and his confidence in me. I thank Dr. Louis Collins, my co-supervisor for all the help in my MR imaging processing. He also introduced me to the MR spectroscopy lab.

I would like to thank Dr. Alan Vandal for his insight guidance in the research of statistical modeling. I would also like to thank Dr. Sridar Narayanan for his endless help throughout my research. Dr. Andrea Bernasconi was the first person who introduced me all the Neuroimaging techniques in neurological research. Dr. Bruce Pike, a master of MR physics, offered me chance to understand MR mechanism. Dr. Robert Kearney taught me not only the signal processing, but also rigorous attitude in science research.

Thank Berengere Aubert Broche for translating the abstract into French. Thank my friends and colleagues in the MR spectroscopy lab. They are Simon Francis, Zografos Caramanos, Jacqueline Chen, Samson Antel, Neda Bernasconi, Amy Wang, Carmela Tarteglia, Ellie Tobman, Alexandre Carmel-Veilleux and Robin Wang. Without them, my research experience would not be so enjoyable.

Thank my parents for their constant helps and support. Thank my wife, Jing Zou for her encouragement and advice, and my son, Sean Xu, who made my life so enjoyable.

1. Introduction

Proton magnetic resonance spectroscopy (MRS) is a valuable technique for imaging neuronal injury in the brains of multiple sclerosis (MS) patients. Recent studies have shown that there are statistically significant decreases of N-acetylaspartate (NAA) in MS lesions, and also in normal-appearing white matter (NAWM) and gray matter. Current approaches to the analysis of MRS data suffer from several limitations. Usually voxels from lesions or NAWM are selected arbitrarily or average values are used. Inconsistent voxel selection has several limitations:

a) These methods are prone to be subjectively biased.

b) They use only part of the available information, which may decreases statistical power.

c) They normally do not take into consideration the spatial correlation between voxels.

The goal of this thesis is to improve the analysis method that uses the information from all the voxels in the MRS data. We propose to analyze MRS data using extensions of the multivariate mixed effect models for repeated measurement. This method takes into consideration the spatial correlations between the voxels. Inferences can be made based on general linear models, which allow one to assess the effect of several independent variables (e.g. disease duration or group) on MRS signal intensities. We hypothesize that this method will produce better results by improved

statistical power than the conventional MRS analysis methods to detect difference between different groups and difference between lesions and NAWM.

The thesis is organized as follows. Chapter 2 introduces multiple sclerosis disease and relevant Neuroimaging techniques, both magnetic resonance imaging (MRI) and MRS. Chapter 2 also overviews brain MRI registration and classification techniques. The last part of chapter 2 introduces the concepts of multivariate mixed effect models for repeated measurement.

Chapter 3 presents data (MRI and MRS) acquisition processes and post-processing, especially resampling data to standard space to facilitate comparison across different persons.

Chapter 4 describes our statistical methods to analyze both cross-sectional and longitudinal data, which includes structured covariance matrix modeling, parameter estimation and statistical inference.

Chapter 5 presents the results of our three studies using our analysis method. There are one cross-sectional study, one longitudinal study and one clinical study with patients treated with drug.

Chapter 6 presents discussion and conclusion.

2. Review of the literature

2.1 Multiple Sclerosis

2.1.1 Epidemiology of MS

Multiple Sclerosis (MS) is an inflammatory demylinating disease of the central nervous system. There are about 300,000 MS patients in the North America and over 2 millions worldwide. Females are affected more frequently (2-3 times) than males. Almost 70% of patients show symptoms between ages 21 and 40. MS is the leading cause of non-traumatic neurological disability in young and middle-aged adults [1].

2.1.2 Phases of MS

About 85% of MS patients experience acute symptoms followed by partial or complete remission, entering the relapsing-remitting (RR) stage. The relapsing-remitting cycles continue and cause chronic accumulation of clinical disability from incomplete remissions. After 10 years, roughly 50% of these RR patients will enter the secondary-progressive (SP) stage of the disease [2], which is characterized by a gradual worsening of the disease without apparent exacerbation or remission.

2.1.3 Pathology of MS

One characteristic pattern of MS is multi-focal demylinated lesions. The typical pathological patterns of white matter MS lesions are: [3,4]

- Inflammation with T cells, B cells and macrophages/microglia.
- Demyelination and a variable degree of remyelination.

- Oligodendrocyte loss.
- Axonal loss and degeneration.
- Gliosis with astrocyte proliferation and intensive glial fibre production.

Lesions in gray matter have also been found in MS [5,6,7], which are associated with extensive demyelination, neuro-axonal loss and microglial activation. Cortical lesions are generally not associated with increased lymphocyte infiltration [8]. Also there are findings of lesions in spinal cord [9,10].

Even the normal appearing white matter (NAWM) is not 'normal' as compared to the normal control. The finding in NAWM includes [11,12,13]

- Perivascular inflammation.
- Marked astrocytic proliferation.
- Reduced axonal density.
- Sclerosis in blood vessels.

These findings suggest that MS is pathologically diffuse and heterogeneous.

2.2 Neuroimaging and MS

2.2.1 Magnetic Resonance Imaging (MRI)

Proton (^{1}H) nuclear magnetic resonance (MR) originates from the interaction between proton nuclei (mainly from water) and an external magnetic field. Relaxation is described as the process by which spins return to equilibrium after an initial perturbation (excitation). Conventional MRI makes use of different T1 and T2 relaxation times between different brain tissue types. In T2-weighted images, tissues

that have long T2 relaxation times (such as fluids) appear bright. In T1-weighted images, tissues that have short T1 relaxation times (such as fat) present as bright signal.

T2-weighted images are highly sensitive for the detection of hyper-intense MS lesions as in Figure 1(a). This makes them useful for diagnosing MS [14]. Hypo-intense lesions on T1-weighted MRI, as in Figure 1(b) normally represent areas where there are severe demyelination and axonal loss [15].

(a) (b)

Figure 1: (a) Axial T2-weighted MRI and (b) T1-weighted MRI. They presented the same brain location of a sigle MS patient. In (a) multiple T2 hyperintense MS lesions are visible. In (b) corresponding multiple T1 hypo-intese lesions.

One limitation of T2-weighted MRI is that it lacks specificity with regard to the heterogeneous pathological substrates of individual lesions. Specifically, edema, inflammation, demyelination, remyelination, gliosis, and axonal loss, all lead to a similar appearance of hyper-intensity on T2-weighted images.

2.2.2 Magnetic Resonance spectroscopy (MRS) Imaging

2.2.2.1 Chemical Shift

Chemical compounds containing a particular nucleus can have slightly different resonance frequencies than those predicted for the nucleus alone due to the interactions of the negatively charged electrons that surround the nucleus. Electrons have spin properties similar to the protons and the neutrons in the nucleus of the atom. When placed in an externally applied magnetic field, electrons precess and generate a small magnetic field around the nucleus. These local magnetic fields created by the electrons can variably decrease the strength from the external field. As a result, the nucleus experiences a slightly altered magnetic field, and resonates at a slightly shifted frequency [16]. This phenomenon is called chemical shift and is one of the important properties of MR spectroscopy. Chemical shift determines the resonance frequency position of each peak on a MR spectrum, which, when expressed as parts per million (ppm) from a reference substance, is independent of the strength of the magnetic field [17] applied.

2.2.2.2 Long Echo MRS

There are several important parameters for MRS sequence. Repetition time (TR) is the time between two consecutive radiofrequency (RF) pulses measured in milliseconds. Clinical MRS normally uses TR of 2000, which is long enough to get reasonable signal intensities. Echo time (TE) is the time from the application of an RF pulse to the measurement of the spectroscopy signal and is also in units of milliseconds.

Magnetic Resonance spectroscopy (MRS) at long echo times (TE >135) normally displays three major resonances as in Figure 2:

- Choline-containing phospholipids (Cho) at 3.24 ppm
- Creatine and phosphocreatine (Cr) at 3.02 ppm
- N-acetyl groups, mainly N-acetyl-aspartate (NAA) at 2.02 ppm.

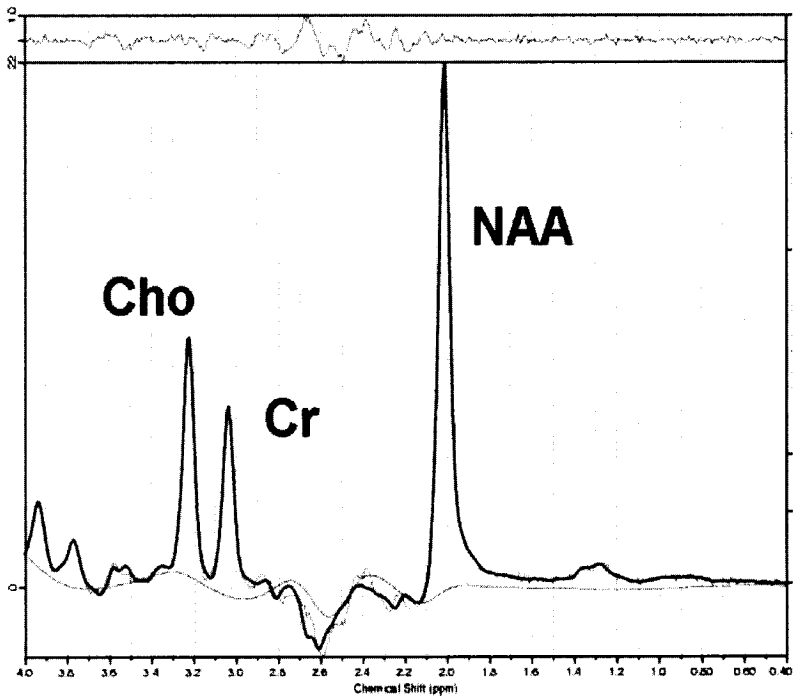

Figure 2: Long echo brain MRS Imaging

NAA is a metabolite found primarily in neurons of mature brains [18,19,20]. Changes of NAA inside the brain can provide information about axonal injury or neuronal dysfunction. The precise role of NAA is not clear. It has been implicated in several neural processes such as: regulation of neuronal protein synthesis [21], fatty acid synthesis [22,23,24] and metabolism of neurotransmitters such as N-acetyl-aspartyl-glutamate (NAAG) [25].

In brains of MS patients, NAA reductions have been found in lesions [26], in normal- appearing white matter (NAWM) [27] and even in cortical gray matter (GM) [28]. NAA in NAWM can be abnormally low in the early stages of MS, even before significant clinical disability is evident [29] and before clinically definite MS has been established [30]. NAA levels are significantly correlated with patients' clinical disability [31], selective motor impairment [32] and cognitive dysfunction [33].

Choline-containing phospholipid (Cho) is mainly in the cell membranes and can be considered an index of phospholipids metabolism. Increased Cho probably suggests increased membrane synthesis or breakdown, and may also indicate increased number of cells. Cho has been shown to increase following inflammation and demyelination in lesions and NAWM [34,35].

Creatine/phosphocreatine (Cr) is known to play an important role in energy metabolism and represent high-energy phosphates reserves that provide for homeostasis and energy needs. Cr levels have been shown to be the highest in astrocytes and oligo-dendrocytes [36]. In normal and some pathological conditions, Cr is homogeneously distributed inside brain and has been considered stable enough to be used as an intrinsic reference for reporting relative concentrations of other metabolites in the brain.

2.2.2.3 Short echo MRS

Using a short echo time (TE< 30 ms) allows observation of additional metabolites with short T2 relaxation times, such as myoinositol, glutamate, glutamine and glycine. Because these metabolites have very low concentrations in the human brains, quantification of them could be very complex. In our clinical studies, we mainly use long echo MRS.

2.2.2.4 Single-voxel and multi-voxel MRS

In the clinical application of MR spectroscopy, localization techniques allow the definition of small tissue volumes of interest (VOI), using anatomical MR images as reference. Localization techniques may be characterized as single-volume (or voxel) or multi-voxel. There are three main single voxel techniques, but only two of them are commonly employed clinically. One is stimulated echo acquisition mode (STEAM) and the other is point resolved spectroscopy (PRESS) [67] technique. Both techniques use frequency selective radiofrequency pulses to excite three orthogonal planes. A spectrum can then be obtained from a small volume of tissue defined by the intersection of three orthogonal planes.

Both STEAM and PRESS are highly effective volume localization schemes, but there are some major differences between them. The STEAM technique is better for measurements involving short echo times (TE < 30 ms), but the STEAM sequence loses signal intensity by a factor of two and is highly susceptible to motion and diffusion processes. PRESS is the choice of volume localization method for long echo times (TE > 135 ms). This sequence also has better signal intensity and is less sensitive to patient motion effects.

Multi-voxel spectroscopy can be one-dimensional (1D), two-dimensional (2D), or three dimensional (3D), and is usually performed using chemical shift imaging (CSI). CSI is a technique for collecting MR spectra from multiple contiguous voxels covering a large region of interest. Data is obtained as spectral maps or metabolite images, and they can be superimposed on conventional MR images to compare changes in spectra

from adjacent voxels, or to obtain the distributional pattern of a particular metabolite within the interested tissue segment.

The dimensionality of CSI techniques refers to the number of spatial dimensions that are phase encoded in the measurement sequence. 1D CSI, 2D CSI and 3D CSI generate spectra localized to slices, columns and voxels respectively. Normally 1D CSI and 2D CSI are combined with other localization methods to define the dimensions that are not phased encoded. For example, in 2D CSI, a slice of tissue is defined using a slice selective excitation in one dimension and using phase encoded gradients in the other two dimensions.

Single-voxel spectroscopy with automated processing is widely available for clinical use because it is simple and easy to implement, and produces a single spectrum that is immediately accessible for interpretation.

In comparison, multi-voxel spectroscopy allows acquisition of data from many locations at the same time and improves detection of regional variability. Normally the voxel size of CSI is much smaller compared to the volume of single-voxel spectra. A smaller voxel volume of CSI leads to lower signal intensity. Furthermore the clinical scan time limits the number of averages of signals available for large-matrix CSI. So multi-voxel MRS signals generally have lower signal-to-noise ratio (SNR). In order to reach reasonable SNR, the resolutions of CSI images are normally lower (10 x 10 x 10 mm^3) as compared to that of conventional MRI images, which can reach to a resolution of 1 x 1 x 1 mm^3.

2.2.2.5 Quantification of MRS

In most studies of MS using MRS, voxels are arbitrarily selected from specific region of interest (ROI). This may be subject to selection bias and partial volume effect. Partial volume effect occurs because each MRS voxel may contain different tissue types.

Absolute quantification of metabolites in MRS can be difficult. A more practical and easier way to assess variations in metabolite levels is by calculating peak area ratios. Metabolite ratios, like those calculated in this study using Cr as a reference, are favored for clinical MRS because they are simple to obtain, are not dependent on changes in coil loading among different individuals, and are relatively unaffected by relaxation-times changes. Also, ratios are not susceptible to cerebral spinal fluid (CSF) partial volume effect. The disadvantage of using metabolic ratios is that they are sensitive to changes in the concentration of both metabolites in the ratios. Using Cr as a reference in our study assumes that its concentration is stable. This is more or less the case [85]. However interpretation of NAA and other metabolites relative to Cr should be made with care.

2.3 Overview of Brain Imaging Processing

2.3.1 Image Registration

Registration is the determination of a transformation between the coordinates in one space and those in another, such that points in the two spaces that correspond to the same anatomic point are mapped to each other.

There are four basic types: a) Intra-subject, intra-modality: alignment of studies in the same subject acquired in the same modality. b) Intra-subject, inter-modality: alignment of studies in the same subject across different modalities. This is relevant to the registration of a functional (functional MRI) study with an anatomical (MRI) study. c) Inter-subject, intra-modality: alignment of studies of different subjects in the same modality. This is commonly needed if group comparison is of interest such as using statistical parametric mapping (SPM). d) Inter-subject, inter-modality: alignment of studies of different subjects across different modalities. It can be useful, for example, for registration of a positron emission tomography (PET) study to an MRI template, although more usually it is achieved via intra-modality registration as an intermediate step.

Most registration algorithms require one or several quantitative similarity measures of the alignment between the two images to be matched (usually defined in the overlapping volume). The similarity measures can be categorized according to the type of information used to construct the measures, as follows:

a) Landmark measures: The similarity measure would typically represent the average distance between the corresponding landmarks. The subclasses of landmarks are: extrinsic [37,38,39], anatomical [40,41] and geometrical landmarks [42].

b) Surface or edge measures: They require a preliminary processing step to extract surfaces or edges in both images. The similarity measure quantifies an average "distance" between the corresponding surfaces. There is a wide range of techniques such as "crest lines" [43,44,45] and points on the surface [46].

c) Voxel intensity measures: where subclasses are c.1) Principal axes: By determining the center of mass (counts) and orientation (i.e. principal axes) of the images to be registered, a direct transformation can be calculated. Registration is then achieved by simply aligning both the centers of mass and orientation [47,48]. c.2) Cross-correlation can be applied for both intra- and inter-modality registration problems [49,50]. c.3) Use of the voxel intensity histogram: These include joint entropy [51], mutual information [52,53,54] and the correlation ratio [55,56].

There are two types of transformations for registrations. The most common one is the rigid body transformation, which assumes that the distances between any two points in the body are preserved. It requires six parameters: three rotations and three translations. The other is non-rigid registration, which can be used to register brain images from different subjects or to monitor the change of the brain over time.

During registration, 3-dimentional interpolations are required to estimate the values of the transformed image. The two simplest methods are nearest-neighbor and linear interpolation. Nearest-neighbor interpolation takes the value from single closest neighboring voxels. The linear interpolation takes the weighted values from four nearest voxels. The weighed values are reverse to the Euclidean distances from the resampled position to these discrete neighboring samples. More complicated interpolation methods

are the higher order interpolations where their weights are based on higher order polynomial functions of distances. These include quadratic interpolation [57], cubic B-spline interpolation [58], Lagrange Interpolation and Gaussian Interpolation [59].

2.3.2 Brain MRI Segmentation and Classification

Segmentation is the process of separating the images into groups of voxels based on homogenous features. There are many features that can be used, such as signal intensity, spatial location, edge, context. Classification is the process of labeling each MRI voxel to different tissue types such as gray matter, white matter, cerebral-spinal fluid (CSF) or MS lesions.

For automatic and semi-automatic MS lesion identification processes, most of the segmentation algorithms are mainly based on multi-spectral signal intensities and in combination of other features. These are:

- Statistically modeling [60]: parametrically models the multi-modality intensities of the different tissues types, while some treat MS lesion as outlier for the models.
- Fuzzy connectedness [61,62]: assigns fuzzy affinities to the target object during classification. The affinity between the two given voxels is defined as a combined weighted function of the coordinate space adjacency, the intensity space adjacency, and the intensity gradient space adjacency to the corresponding target object features.
- Artificial neural networks (ANN) [63]: non-parametric analysis based on a system of parallel and connected nodes that process information to make decisions indicated at output nodes, which mimics the real human neural systems.

- K-Nearest Neighbor (KNN) classification [64]: a technique that is used for the identification of clusters that occur naturally in some form of feature space.
- Bayesian classifier [65]: uses Bayes' formula to calculate probability that a given voxel belongs to a certain tissue class given its intensity.

2.4 Multivariate Statistical Models

2.4.1 Normal Distribution

The normal distribution is the most used statistical distribution because normality arises naturally in many physical, biological, and social measurement situations and normality is important in statistical inference.

Let x be a scalar random variable. The general formula for the probability density function of the normal distribution is:

$$f(x) = \frac{\exp[-(x-\mu)^2/(2\sigma^2)]}{\sqrt{2\pi}\sigma}$$

where u is the mean of the normal distribution and σ^2 is the variance, which describes the degree of variation of the distribution around mean.

2.4.2 Multivariate Normal (MVN) Distribution

Let $\boldsymbol{X} = (X_1, X_2 \ldots X_n)'$ be a random vector composed of random variables X_i, where $i = 1, 2, \ldots, n$. We denote the n-dimensional joint-normal distribution with mean vector μ and covariance matrix Σ as $N_n(\mu, \Sigma)$. If Σ is positive definite, the probability density function of the MVN is:

$$f(x) = \frac{\exp[-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)]}{\sqrt{2\pi^n |\Sigma|}}$$

Given any pair of components X_i and X_j, we denote their covariance as *cov* (X_i ,X_j). The covariance is defined by the expectation:

$$\text{cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)]$$

where μ_i and μ_j are the means of X_i and X_j. By definition, covariance is symmetric, with *cov* (X_i ,X_j) = *cov* (X_j ,X_i). Also, the covariance of any component with itself is the variance of the component:

$$\text{cov}(X_i, X_i) = E[(X_i - \mu_i)(X_i - \mu_i)] = \text{var}(X_i)$$

So the detailed structure of the covariance of matrix of random vector $\boldsymbol{X}$ is:

$$\Sigma = \begin{pmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \dots & \text{cov}(X_2, X_n) \\ \dots & \dots & \dots & \dots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \dots & \text{cov}(X_n, X_n) \end{pmatrix}$$

Analogous to variance, σ^2 of scalar random variable x, which is a measure of variation relative to the mean of x in 1-D real space, the covariance matrix, $\boldsymbol{\Sigma}$ measures variation of the random vector $\boldsymbol{X}$ relative to the mean vector, $\boldsymbol{\mu}$ in a real space of dimension n.

2.4.3 Multivariate Mixed Effect Model

Mixed effect linear models incorporate both fixed effects and random effects. Fixed effects are associated with the groups as a whole or with levels of experimental factors, which are normally what we are interested in. Random effects describe the variation of individual within his own group. Mixed effect models are suitable to

characterize the common structure of repeated measures, growth curves or serial measurement data.

Let $\boldsymbol{y}_i$ be a $t_i \times 1$ vector containing the responses for subject i, where $i = 1,\ldots,n$, and t_i is the number of repeated or serial measurements for subject i. $\boldsymbol{y}_i$ are assumed to follow the model

$$y_i = X_i\beta + \varepsilon_i ,$$

where $\mathbf{X}_i$ is a $t_i \times p$ known design matrix, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown regression parameters (fixed effect) and the $\boldsymbol{\varepsilon}_i$ is independently distributed as $\boldsymbol{N_n}$ **(0,** $\boldsymbol{\Sigma_i}$**).** And the elements of each covariance matrix $\boldsymbol{\Sigma_i}$, for $i = 1, ..,n,$ are known functions of q unknown covariance parameters contained in the vector $\boldsymbol{\theta}$.

The ability to model those $\boldsymbol{\Sigma_i}$ allows one to examine several alternative structures for $\boldsymbol{\Sigma_i}$, each structure having important subject matter interpretations. Even if the interest is mainly in the regression parameter $\boldsymbol{\beta}$, efficiency of their estimates may be improved considerably by modeling them parsimoniously. This is especially likely when sample sizes are small and the data are unbalanced.

The simplest form for $\boldsymbol{\Sigma_i}$ is one that arise from independent, constant variance observations:

$$\Sigma_i = \sigma^2 I_i$$

where $\boldsymbol{I}_i$ is the a $t_i \times t_i$ identity matrix. This is just the ordinary linear regression model. One may generalize this by allowing σ^2 to vary from group to group, or by allowing the variance to vary from observation to observation within subject. The latter still assumes that each $\boldsymbol{\Sigma_i}$ is diagonal.

Other important special class is the random-coefficients model [66]:

$$\Sigma_i = Z_i \phi Z_i^{'} + \sigma^2 I_i .$$

This model arises by assuming:

$$\varepsilon_i = Z_i b_i + \mu_i ,$$

where $\mathbf{Z_i}$ is a known $t_i \times k$ matrix, $\boldsymbol{b_i}$ and $\boldsymbol{\mu_i}$ are independent random vectors with $b_i \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{\phi})$ and $\mu_i \sim \boldsymbol{N}(\mathbf{0}, \sigma^2 \mathbf{I_i})$. $\boldsymbol{b_i}$ is the parameters for the random effect, which describes the variation of each individual within group.

Another important class is called incomplete data model, which is obtained by assuming each $\boldsymbol{\Sigma_i}$ is actually a sub-matrix of a $T \times T$ matrix $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$. This model arises in the situations where a fixed number T of measurements, corresponding to different times of experimental conditions, are to be collected on each of n subjects, but not all of the subjects' responses are observed. This model allows us to directly deal with missing data.

3. Data Acquisition and Processing

3.1 MR and MRS examination

Conventional proton MRI and MRS examinations of the brain were obtained in a single session for each examination using a Philips Gyroscan S15 operating at 1.5T (Philips Medical Systems, Best, The Netherlands). A sagittal survey image was used to identify the anterior commissure (AC) and posterior commissure (PC). Multi-slice images were obtained in coronal and transverse planes, perpendicular and parallel to the AC-PC line, respectively (TR = 2075, TE1 = 30.6, TE2 = 90, slice thickness 3 mm). These images were used to select an intracranial volume of interest (VOI) for spectroscopy. We used a VOI angled parallel to the AC and PC line measuring approximately 90 mm anteroposterior × 20 mm craniocaudal × 90 mm left-right and offset craniocaudally so that it was centered on the corpus callosum. The VOI was kept constant in size and position after the first examination for each subject.

Proton spectroscopy images were acquired using a 90° - 180° - 180° PRESS sequence for volume selection [67] (TR = 2000, TE = 272). They are long echo CSI. Magnetic field homogeneity was optimized to a line-width of about 5 Hz over the VOI using the proton signal from water. Water suppression was achieved by a chemical shift selective saturation (CHESS) pulse followed by a dephasing gradient pulse [68].

MRSs were generated by two-dimensional phase encoding (250 mm × 250 mm field of view, 32 × 32 phase encoding steps and one signal average per step). After a water-suppressed acquisition was completed, another MRS was acquired without water

suppression using TR 2000, TE 272, field of view of 250 mm ×250 mm and 16 ×16 phase encoding steps.

3.2 MRS Post-processing

Post-processing of the raw spectrum data was done on a SUN/SPARC system, using Xunspec1 software (Philips Medical System, Best, The Netherlands). The non-water-suppressed MRS were interpolated to 32 × 32. A mild Gaussian k-space filter and an inverse two-dimensional Fourier transformation was then applied to both the water suppressed and unsuppressed MRS. Artifacts present in the time-domain water-suppressed signal due to magnetic field inhomogeneities were corrected by dividing the water-suppressed MRS signal by the non-water-suppressed signal [69]. The residual water signal was then fitted and removed from the water-suppressed data using the Hankel singular-value decomposition (HSVD) procedure [70]. To enhance the resolution of the spectral peaks, a Lorentzian-to-Gaussian transformation was applied prior to Fourier transformation to the spectral domain. The result was 1024 voxels (32 × 32) each containing data ready for quantification and subsequent generation of the MRS. The nominal voxel size in plane was about 8 × 8 × 22 mm^3, giving a resolution of about 12 × 12 × 22 mm^3 after k-space filtering.

Peaks for NAA and Cr were detected using locally developed software. The resonance intensity was determined from the area of each peak that was bounded below by a spline-corrected baseline. Chemical shifts were calculated relative to the NAA resonance at 2.02 ppm.

After quantification, one MRS expert (Dr. Sridar Narayanan) would review the results. The MRS intensities with poor qualities (poor baseline shape, bad phase

correction, poor signal-to-noise ratio) were rejected and all the others were treated as valid intensities.

3.3 MRI Post-processing

Multiple sclerosis lesion classification was performed mainly by Simon Francis and it is composed of two processes. First, the data was run through an automatic Bayesian classifier, which offered the ability to toggle between the proton density, T1 and T2-weighted images (to facilitate discrimination between grey matter and CSF) [71]. Next, readers reviewed the resulting lesion voxels and corrected errors.

The automatic Bayesian classifier is composed of three stages:

1) Pre-classification processing: includes intensity bias-field corrections, aligning (registering) the image modalities, isolating the brain parenchyma and ensuring that the images have a standardized intensity range.
2) Classification: First using a k-means classifier in conjunction with domain-specific heuristic and connectivity rules to estimate the mean and covariance of multiple sub-classes per tissue type. Then the probabilities for each tissue class are calculated using Bayes theorem. The class with the highest probability, the maximum *a posteriori* (MAP), is selected for each voxel.
3) Post-classification processing: analyzes voxel connectivity to identify misclassified tissue utilizing a set of heuristic rules which mimics the 'common sense' judgment of a trained human expert MRI reader.

After tissue type classification, the MS lesion volumes were calculated by multiplying the total number of identified lesion voxels and the unit voxel volume.

We also extracted the distance to MS lesion by calculating the chamfer distance to the lesion mask using software called 'mincchamfer', which was one of the applications from MNI Automated Linear Registration Package Version 0.98k developed by developed by Dr. Louis Collins [72].

3.4 Standard Space

Both MRI and MRS were acquired from a three-dimensional volume whose size and orientation relative to the scanner isocenter were known. Normally MRS data are usually presented in two-dimensional format. We oriented the spectroscopy image into three-dimension according to its position parameters, and resampled it to the same spatial resolution as corresponding MRI using nearest-neighbor interpolation. This placed the resulting 3-D spectroscopy image in 'native space' (a coordinate frame relative to the scanner isocenter). Thus MRI and MRS were in register with each other as in Figure 3.

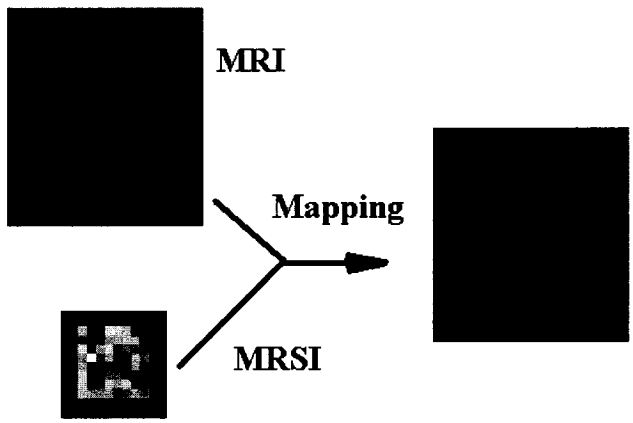

Figure 3: MRI and MRS mapping

Because there are spatial variations in the MRS between subjects due to anatomical variations, signals from different brain locations are not directly comparable.

We needed to register the individual MRI volume to an average three-dimensional MRI brain volume in a standard anatomical space (MNI/ICBM 152) using a transformation consisting of linear translation, rotation and scaling [73,74]. With the same transformation, we also transformed the MRS box and MRI lesion volumes to standard space. After registration of MRI to the average brain map in standard space, we significantly reduced the differences between individual brains in terms of orientation, size and shape. Thus all MRS boxes, MS lesions volumes and MRI anatomical data from different individuals at different time points were anatomically matched with each other so that the voxel to voxel comparison could be possible.

Then we redefined a new VOI in standard space that covered most of the individual MRS boxes. The voxel size of this new VOI was close to that in the original MRS ($8 \times 8 \times 22$ mm^3). MRS data were resampled within the new VOI by averaging the valid resonance intensities within each new voxel. So signals from voxels at the same resampled location in different individual images could be matched with each other. Now the resonance intensity from each voxel in the new VOI was associated with the location of that voxel in standard space and the percentage of MS lesion in that voxel. Appropriate statistical methods could then be applied.

4. Statistical Modeling

Our goal is to quantify the relationships between metabolite concentrations (e.g. NAA/Cr) in brain and other independent variables including clinical diagnostic subgroup, duration of disease and hyper-intensity lesion on T2-weighted MRI. In addition, we want to monitor changes of metabolites during a longitudinal study of 2 years.

4.1 Previous Methods for MRS

Multi-voxel MRS has an advantage over localized single voxel MRS in that it presents spatial information on biochemical pathological changes in brain. This is especially true for diseases like MS which have multiple pathological foci. A common approach to analyze MRS is to manually choose several voxels at different locations in each spectroscopic image and compare the difference in signal intensities either between clinical subgroups or between normal and abnormal brain tissues. Working in this way, valuable information from unselected voxels is ignored, diminishing the advantage offered by multi-voxel MRS. Another common method is to average all the valid MRS intensities to get the general indication of the metabolite level within the brain.

4.2 Statistical Models for the Cross-sectional Study

We propose to use all the information from MRS and treat it with the multivariate mixed effect model. The model can incorporate the spatial correlations in

the MRS and the variations between individuals. And the model enables us to assess the relationships between resonance intensities and other independent clinical or MRI variables. The model is

$$Y_{iv} = X_{iv}\beta + \gamma_i + \varepsilon_i(v),$$

where $i=1, 2, ..., N$, indicates subjects and $v = 1, 2, ..., V$ denotes the voxels in the spectroscopic image in standard space. X_{iv} are $1 \times p$ vectors containing independent variables for each individual and β is $p \times 1$ vector defines the fixed effect parameters. γ_i indicates the random effect for each individual , which is independent and identically distributed normal variable ($\gamma_i \sim N(0, \tau^2)$). γ_i describe the deviation of individual level of MRS intensities from the corresponding group level. $\varepsilon_i(v)$ indicates spatially correlated residual which incorporates the intra-subject spatial correlation in MRS. Specifically, $E[\varepsilon_i(v)] = 0$ and the correlation between $\varepsilon_i(v)$ and $\varepsilon_i(u)$, $Cor(\varepsilon_i(v), \varepsilon_i(u))$, is a monotonic decreasing function of the distance as shown below.

Let $\boldsymbol{Y_i} = [Y_{i1}, Y_{i2}, ..., Y_{iV}]'$, for $v = 1, 2, ..., V$ to indicates the resonance intensities (or ratio) for all the voxels within subject i. $\boldsymbol{Y_i}$ is then a multivariate normal (MNV) variable with structured covariance, $\boldsymbol{\Sigma}$.

$$\sum = \tau^2 J + R$$

where $\boldsymbol{J}$ is a V by V matrix with all elements equal to one, τ^2 measures the variation between different subjects, i.e. the random effect and $\boldsymbol{R}$ is the matrix describes the spatial correlations between voxels, where each element Rij is the covariance between voxel i and j, $Cov(v_i, v_j)$.

4.2.1 Spatial Correlation

The causes for the spatial correlations between voxels in MRS are two-fold. First, the distributions of metabolites are continuous and smooth within the brains. Second, the interpolation function of the CSI spatial sampling extends outside the nominal voxel as in Figure 4. The resulting resonance intensity in one voxel partially contains information from surrounding areas.

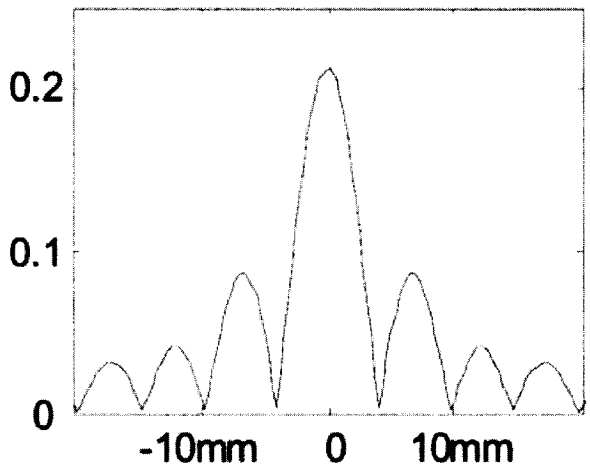

Figure 4: 2-D view of interpolation function of CSI sampling

One way to measure the spatial correlations is semi-variogram. The idea was borrowed from geostatistics. $\gamma(h)$, the semi-variogram at lag h is defined as:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{|s_i - s_j| = h} (y_i - y_j)^2 ,$$

where h is the lag, the distance between locations s_i and s_j; y_i and y_j are the values of variable y at location s_i and s_j respectively; $N(h)$ is the number of pairs of observed data points separated by a lag of h.

Here we demonstrate one 1-dimensional semi-variogram by a simulation of a sequence of 1,000 samples from one uniform distributed random variable. Because each

sample was drawn independently, the correlation between any two samples, y_i and y_j would be the same. So the semi-variogram, $\gamma(h)$ had the same value around 820 and was independent of the lag, h as in Figure 5(a) and Figure 5(b). The only trend along increasing lag, h was the increasing variation of sampled $\gamma(h)$. This was because with larger lag, the number of pairs of observed data, $N(h)$ was smaller.

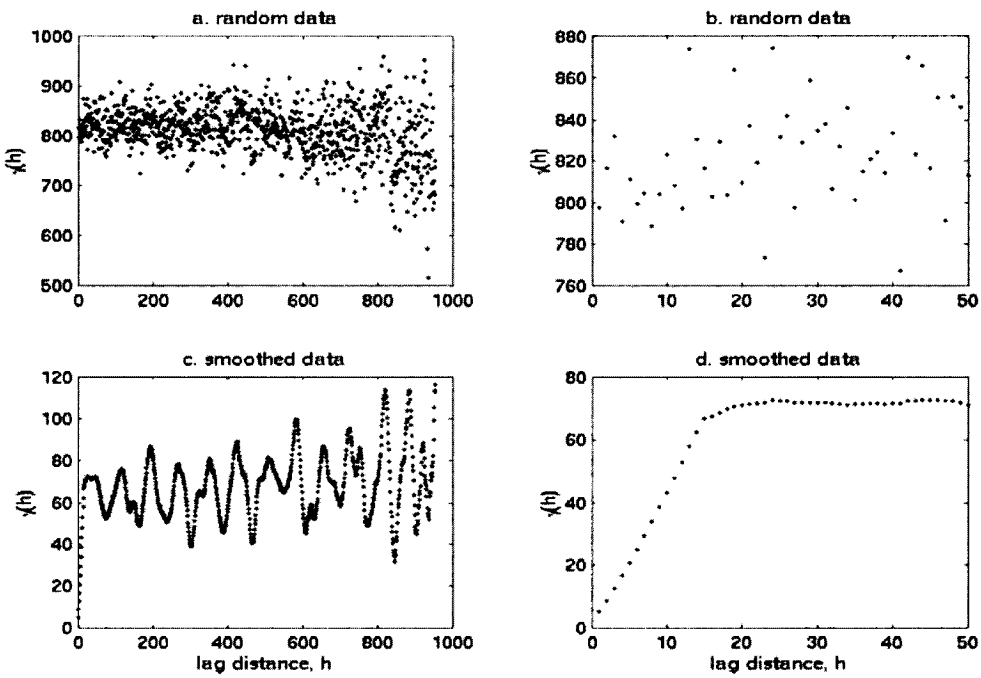

Figure 5: Simulation of semi-variogram (a) γ(h) of raw sample; (b) γ(h) of raw sample - first 50 lags; (c) γ(h) of smoothed sample (d) γ(h) of smoothed sample - first 50 lags

Now we introduced artificial correlations between the raw samples by applying one 15-point moving average sequentially. The average $\gamma(h)$ decreased to around 70 as in Figure 5(c). This was because the moving-average sequence decreased the variance of the raw samples. We also notice that for the first 15 lag values ($h=1, 2, ..., 15$), $\gamma(h)$ increased almost linearly as shown in Figure 5(d) from 0 to 70. This indicated larger spatial correlations existed between samples with smaller lags.

The semi-variogram of MRS (Figure 6) clearly shown that there were significant spatial correlations between adjacent voxels in our MRS data, especially for voxels within distance of two-voxel size. Here the lags were measured as 3-dimensional distances between the center of each voxels and were in unit of single MRS voxel size (10 mm).

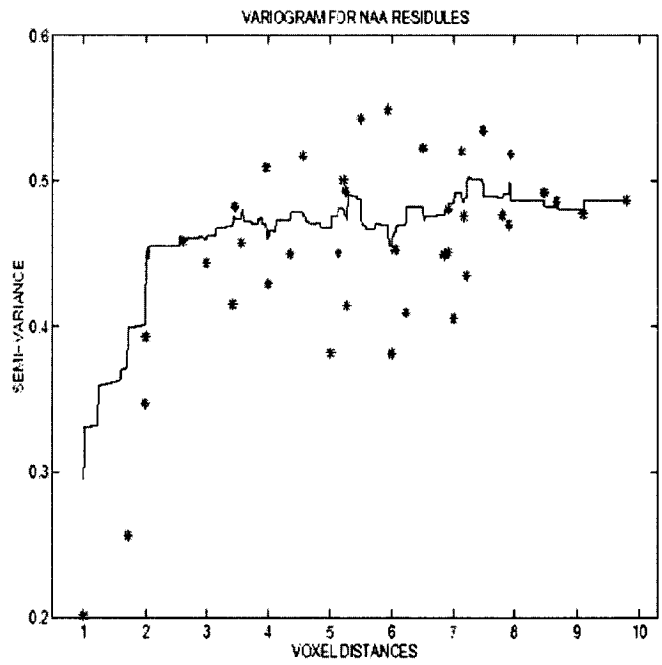

Figure 6: Sampled semi-variogram of MRS

Let $\boldsymbol{Y_i} = [Y_{i1}, Y_{i2}, ..., Y_{iV}]'$, for $v = 1, 2, ..., V$ to indicate the resonance intensities of voxel v within subject i. We modeled the structure covariance matrix, $\boldsymbol{\Sigma}$ of MNV $\boldsymbol{Y_i}$ as

$$\Sigma = \tau^2 J + R,$$

where $\boldsymbol{R}$ is the matrix defines the spatial correlations between voxels. Each element Rij of $\boldsymbol{R}$ is the covariance between voxel i and voxel j, Cov(v_i,v_j). One popular way to describe the spatial correlations is by a first-order auto-aggressive (AR1) model [75]:

$$Cov(v_i, v_j) = \begin{cases} \sigma^2, i = j \\ \sigma^2 \rho^{\|i-j\|}, i \neq j \end{cases},$$

where σ^2 is the variance of MRS voxels, ρ is the correlation parameter and $\|i\text{-}j\|$ is the distance between voxel i and j. Here we assume that the variances, σ^2, of voxels at different locations are the same.

Furthermore, if the variances of voxels differ considerably, we may not assume σ^2 are the same for voxels at different locations. In that case, we proposed a first-order ante-dependence (AD1) model [76] to describe the spatial correlation matrix $\boldsymbol{R}$:

$$Cov(v_i, v_j) = \begin{cases} \sigma_i^2, i = j \\ \sigma_i \sigma_j \rho^{\|i-j\|}, i \neq j \end{cases},$$

where each voxel, v_i has its own variance σ_i^2, and ρ is the correlation parameter.

4.2.2 Missing Data

Signals from MRS normally have low signal-to-noise ratio (SNR), which makes the quantification of intensities difficult. The resonance intensities from some of the MRS voxels could be invalid.

Utilizing the idea of incomplete data model, our mixed effect model could deal with missing data naturally [77]. We could exclude the missing voxels from the regression models by omitting the corresponding row and column from both the design matrix and the structured covariance matrix for that subject. The remaining sub-matrices can still contribute to the parameter estimation and inference.

4.3 Model fitting

One of the most popular model fitting methods is Maximum Likelihood (ML). Let $\boldsymbol{Y_i} = [Y_{i1}, Y_{i2}, ..., Y_{iV}]'$, for $v = 1, 2, ..., V$ to indicate the resonance intensities of voxel v within subject i. We treated $\boldsymbol{Y_i}$ as a multivariate normal (MNV) variable with structured covariance, $\boldsymbol{\Sigma} = \tau^2 \boldsymbol{J} + \boldsymbol{R}$. The mixed effect model for resonance intensities, $\boldsymbol{Y_i}$ for subject i is:

$$\boldsymbol{Y_i} = \boldsymbol{X_i}\boldsymbol{\beta_i} + \gamma_i + \boldsymbol{\varepsilon}_i,$$

where $\boldsymbol{X_i}$ is an $V_i \times p$ design matrix for ith subject, $\boldsymbol{\beta_i}$ is the fixed effect parameter $p \times 1$ vector, γ_i is the $V_i \times 1$ random effect vector and $\boldsymbol{\varepsilon}_i$ is $V_i \times 1$ residual vector. We could then express the marginal probability density of $\boldsymbol{Y_i}$ as follows:

$$f(Y_i) = \frac{1}{(2\pi)^{n_i/2} |\Sigma_i|^{1/2}} \exp[-\frac{1}{2}(Y_i - X_i\beta)' \Sigma_i^{-1}(Y_i - X_i\beta)]$$

where $i = 1,2,...n$, indicates the subject, n_i is the number of valid MRS voxel for ith subject, Σ_i is the sub-matrix of Σ, excluding the rows and columns corresponding to missing data.

Because the observations between different subjects are independent, we could obtain the joint likelihood function of all MRS observations by multiplying all the individual marginal probability density, $f(Y_i)$ and the log-likelihood is as follows:

$$l = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log|\Sigma_i| - \frac{1}{2}\sum_{i=1}^{n}(Y_i - X_i\beta)'\Sigma_i^{-1}(Y_i - X_i\beta).$$

Here l is the joint log-likelihood function; N is the total number of observations (valid voxels) of all subjects under study. We obtained the maximum-likelihood estimate (*mle*) of vector parameter $\boldsymbol{\theta}$ using the fisher-scoring algorithm. The parameter vector, $\boldsymbol{\theta}$

contains both the fixed effect estimate vector, β, and the covariance matrix parameters estimate vector, s.

4.3.1 Fisher Scoring

The first derivative of the log-likelihood function is called Fisher's score function, and is denoted by:

$$\mu(\theta) = \frac{\partial l(\theta; Y)}{\partial \theta},$$

where l is the joint log-likelihood function, Y is the observed dependent variable and θ is the estimate vector. The Fisher's score is a vector of first partial derivatives, one for each element of θ. Assuming the log-likelihood function is concave, we can find the *mle* by setting the score to zero by solving the equations:

$$\mu(\theta) = 0.$$

The Fisher's score is a random vector. When we have the true parameter θ, the score has mean zeros

$$E[\mu(\theta)] = 0,$$

and covariance matrix, which is called the information matrix:

$$\text{var}[\mu(\theta)] = E[\mu(\theta)\mu^t(\theta)] = \mathbf{I}(\theta).$$

Normally the information matrix can also be obtained as negative the expected value of the second derivatives of the log-likelihood:

$$I(\theta) = -E[\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta'}].$$

During the iterative procedures, we can expand the score function evaluated at the mle $\hat{\theta}$ around a trial value θ_0 using a first-order Taylor series:

$$\mu(\hat{\theta}) \approx \mu(\theta_0) + \frac{\partial \mu(\theta)}{\partial \theta}(\hat{\theta} - \theta_0).$$

Let **H** denote the Hessian matrix, which contains the second derivatives of the log-likelihood function:

$$H(\theta) = \frac{\partial^2 l(\theta)}{\partial \theta \partial \theta'} = \frac{\partial \mu(\theta)}{\partial \theta}.$$

Setting $\mu(\hat{\theta})$ to zero and solving $\hat{\theta}$ for the first-order approximation, we get

$$\hat{\theta} = \theta_0 - \mathbf{H}^{-1}(\theta_0)\, \mu(\theta_0).$$

The algorithm for computing the *mle* is called Newton-Raphson algorithm. In Fisher scoring algorithm, we replace the negative Hessian by its expected value, the information matrix. So that the improved estimate by each score step is given as:

$$\hat{\theta} = \theta_0 + \mathbf{I}^{-1}(\theta_0)\, \mu(\theta_0).$$

The Fisher scoring algorithm requires two conditions to be met at each iteration. First, each sub-covariance matrix, Σ_i should be positive definite. Second, the log-likelihood must increase at each step. When the size of covariance matrix is too large, or the number of estimated parameters is too large, sometimes the first condition may fail and cause the algorithm fail to converge.

4.4 Inference

After obtaining the estimated parameters using the Fisher Scoring algorithm, we could test if the parameters are statistically significant using the wald test. We may also want to know if one simple regression model is efficient as compared to more complex regression models. We could test this kind of hypotheses using the likelihood-ratio test.

4.4.1 Wald Test

Normally in large samples, the mle $\hat{\theta}$ has approximately a multivariate normal distribution:

$$\hat{\theta} \sim N_p(\theta, I^{-1}(\theta)).$$

Under the hypothesis:

$$H_0 : \theta = \theta_0$$

and for a fixed value θ_0, the Wald statistic is defined as:

$$W = (\hat{\theta} - \theta_0)' \operatorname{var}^{-1}(\hat{\theta})(\hat{\theta} - \theta_0),$$

which has approximately in large samples a chi-squared distribution with p degrees of freedom. p equals to the number of elements in parameter vector $\boldsymbol{\theta}$. In practice, we often replace the covariance matrix with the inverse of the expected information matrix evaluated at the mle, $\hat{\theta}$, i.e. $\operatorname{var}(\hat{\theta}) = I^{-1}(\hat{\theta})$, which can be obtained during the process of Fisher scoring algorithm. This facilitates estimation of the confidence intervals or P-values of coefficient estimates.

In particular if β is a single parameter and is the ith element of the parameter vector $\boldsymbol{\theta}$, under the hypothesis

$$H_0 : \beta = \beta_0,$$

we will have

$$(\hat{\beta} - \beta_0)\{I(\hat{\theta})_{ii}\}^{1/2} \sim N(0,1),$$

where $\hat{\beta}$ is mle of β, $I(\hat{\theta})_{ii}$ is the i^{th} element of the diagonal of the expected information matrix $I(\hat{\theta})$ and $N(0,1)$ is the standard normal distribution.

4.4.2 Likelihood Ratio Test

Suppose we have two models, M_1 and M_2, and M_1 is a subset of M_2, i.e. $M_1 \subset M_2$. We may obtain the simpler model M_1 by setting some of the parameters in M_2 to specific constants.

The maximum likelihood under the smaller model M_1 is

$$\max_{\theta \in M_1} L(\theta, y) = L(\hat{\theta}_1, y),$$

where $\hat{\theta}_1$ is the mle of θ under model M_1.

The maximum likelihood under the larger model M_2 is

$$\max_{\theta \in M_2} L(\theta, y) = L(\hat{\theta}_2, y),$$

where $\hat{\theta}_2$ is the mle of θ under model M_2. The ratio of these two likelihoods,

$$r = \frac{L(\hat{\theta}_1, y)}{L(\hat{\theta}_2, y)},$$

is between 0 and 1.

The twice the negative log of the likelihood ratio has approximately in large samples a chi-squared distribution. That is

$$-2\log r = 2\log L(\hat{\theta}_2, y) - 2\log L(\hat{\theta}_1, y) \sim \chi^2_{df},$$

where the degree of freedom, *df* is the difference of the number of parameters in these two models.

4.5 Statistical Models Extended for Longitudinal Studies

4.5.1 Longitudinal Statistical Models

Let $\boldsymbol{Y_i} = [Y_{i11}, Y_{i21}, ..., Y_{iV1}, Y_{i12}, Y_{i22}, ..., Y_{iVT}]'$, for $v = 1, 2, ..., V$ and $t = 1,2, ..., T$ to indicate the resonance intensities for voxel v at exam time t within subject i. We can treat $\boldsymbol{Y_i}$ as a multivariate normal (MNV) variable with structured covariance, $\boldsymbol{\Sigma} = \tau^2 \boldsymbol{J} + \boldsymbol{R}$. The mixed effect model for resonance intensities, $\boldsymbol{Y_i}$ for subject i will be:

$$\boldsymbol{Y_i} = \boldsymbol{X_i}\,\boldsymbol{\beta_i} + \gamma_i + \boldsymbol{\varepsilon}_i,$$

where $\boldsymbol{X_i}$ is a $V_i \times p$ design matrix for ith subject for all the exam times, $\boldsymbol{\beta_i}$ is $p \times 1$ fixed effect parameter vector, γ_i is $V_i \times 1$ random effect vector and $\boldsymbol{\varepsilon}_i$ is $V_i \times 1$ residual vector. We can then express the marginal probability density of $\boldsymbol{Y_i}$ as follows:

$$f(Y_i) = \frac{1}{(2\pi)^{V_i/2} |\Sigma_i|^{1/2}} \exp[-\frac{1}{2}(Y_i - X_i\,\beta)^t \Sigma_i^{-1} (Y_i - X_i\,\beta)]$$

where $i = 1,2,...n$, indicates the subject, V_i is the number of valid MRS voxel for ith subject, Σ_i is the sub-matrix of Σ, excluding the rows and columns corresponding to missing data.

4.5.2 Time Correlation

For the longitudinal study, we need to take care not only of the intra-subject spatial correlations, but also of the intra-subject time correlations. We could still use the semi-variogram, $\gamma(h)$ to explore the correlation if there were significant intra-subject time relationships between voxels at the same location. Here h is the time lag between examinations.

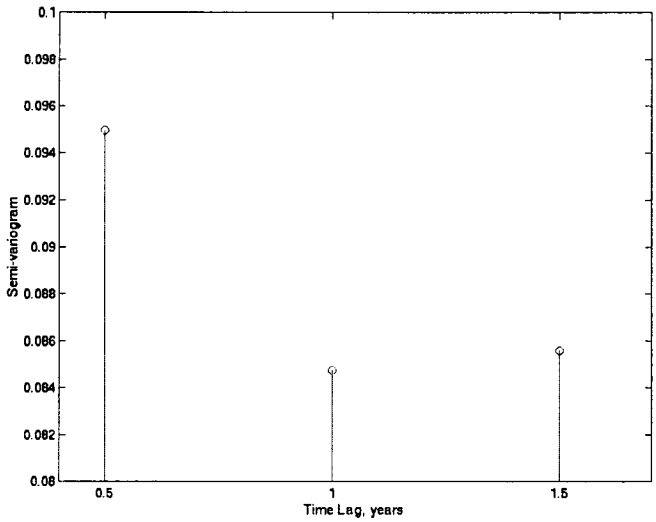

Figure 7: Time correlation over examination time

Based on the semi-variogram over examination time in Figure 7, there were no significant time correlations over time. We could now define the structured covariance matrix for the longitudinal study as:

$$\Sigma_{long} = \tau^2 \mathbf{J} + \mathbf{G},$$

where $\boldsymbol{J}$ is a $V \times T$ by $V \times T$ matrix with all elements equal to one, τ^2 measures the variation between different subjects, i.e. the random effect. $\boldsymbol{G}$ is the matrix describes both the possible time and spatial correlations between the voxels. Each element *Gij* of G denotes the covariance between voxel i and j. Cov(v_i,v_j) is defined as:

$$Cov(v_i, v_j) = \begin{cases} \sigma^2 \rho^{\|v_v - v_j\|}, & v_i \ and \ v_j \ within \ the \ same \ scan \\ 0, & otherwise \end{cases},$$

where σ^2 denotes the variances of MRS voxels, ρ is the correlation parameter and $\|i\text{-}j\|$ is the distance between voxel i and voxel j. Because there were no apparent time correlations, all the voxels from different exams were treated as being independent.

Here we developed a technique to analyze longitudinal MRS data. This method could deal with the spatial and time correlations within data and could naturally handle missing data. We applied this technique in our longitudinal study as shown in the next chapter. In our longitudinal study, both the model fitting algorithm and the inference tests were similar to those used in the cross-sectional study.

5. Experimental Results

5.1 Cross-sectional Study

5.1.1 Study Subjects

We chose 9 normal control subjects and 54 clinically definite MS patients. Thirty-seven of the patients were in the RR phase and seventeen were in the SP phase. The subjects are listed in Table 1.

Groups	Number	EDSS[a]	DURATION (YEARS)[b]	AGE (YEARS)	Averaged NAA/Cr Over Scan[c]
CONTROL	9	0	Not Applicable	34.5 ±9.14	3.1 ± 0.25
RR MS	37	2.4±1.28	9.9 ±7.1	30.8 ± 8.5	2.8 ± 0.23
SP MS	17	6.4 ±1.59	16.2 ± 8.2	46.5 ± 9.3	2.63 ± 0.22

Table 1: Cross-sectional Study Subjects

[a] EDSS is Kurtzke expanded disability status scale [79], a clinical rating scale ranging from 0 (normal neurological examination) to 10 (death due to MS).
[b] DURATION is the number of years after clinically diagnosed as MS.
[c] Averaged NAA/Cr over scan: the average value of NAA/Cr of all the valid voxels for each subject.

5.1.2 Structured Covariance Matrices

We chose a regression model with no independent variables

$$NAA/Cr = \beta_{01} + \beta_{02} + \ldots + \beta_{0V}$$

to estimate the covariance matrix. We used the metabolite ratio, *NAA/Cr* to represent the possible NAA concentration in the brains. β_{0v} for $v=1,2,\ldots,V$ denoted the interceps at each MRS location. This would be applied to all the regression models in the experiments.

The sampled unstructured covariance matrix, Σ_{SAMP} as in Figure 8(a) showed that the variances of voxels at different locations (diagonal elements of Σ_{SAMP}) were not homogenous and there were significant spatial correlations between adjacent voxels (off diagonal peaks).

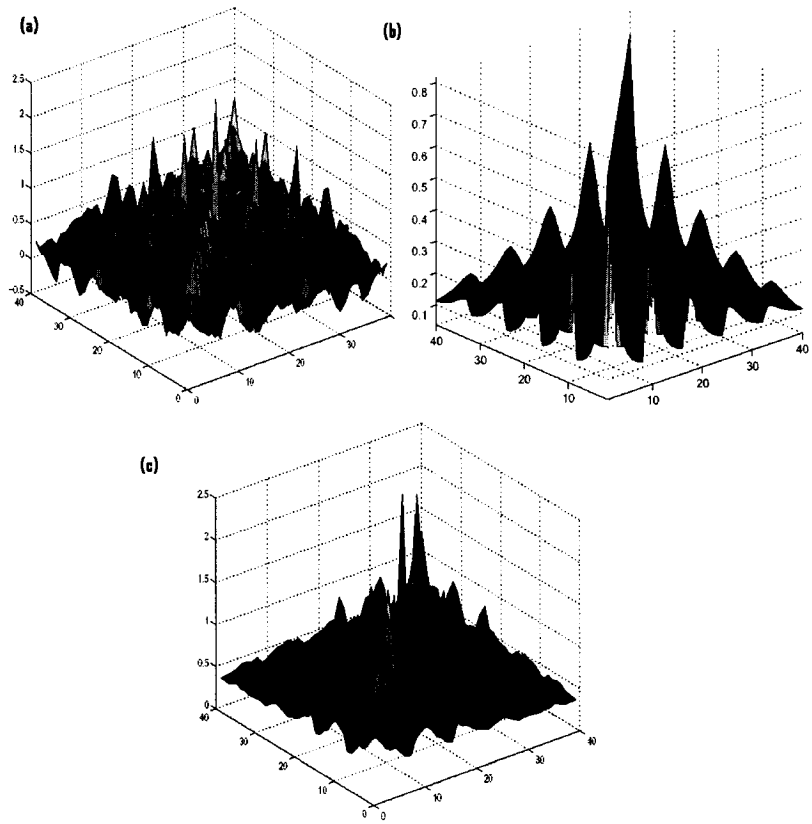

Figure 8: Covariance Matrices. (a) Sampled unstructured Σ_{SAMP}; (b) First-order Auto-aggressive modeled Σ_{AR1}; (C) First-order Ante-dependent modeled Σ_{AD1}.

The first-order auto-aggressive modeled covariance matrix, Σ_{AR1} in Figure 8(b) assumed that the variances of voxels at different locations were similar. Σ_{AR1} required three parameters, τ^2 for random effect, σ^2 for variance and ρ for spatial correlation. Compared to Σ_{AR1}, the first-order anti-dependent modeled covariance matrix, Σ_{AD1} in Figure 8(c) was closer to the sampled Σ_{SAMP}, in the expense of 35 extra parameters to be

estimated. Σ_{AD1} needed 36 parameters to measure the variances at 36 different voxel locations in 6 × 6 MRS matrix.

The comparison between Σ_{AR1} and Σ_{AD1} using likelihood ratio test was

$$-2\log r = 2\log L(\hat{\theta}_{AD1}, y) - 2\log L(\hat{\theta}_{AR1}, y)) = (594.6) - (-3020.5) = 3615.1,$$

which had a chi-square distribution with a degree of freedom of 35. The p-value was smaller than 0.001. So Σ_{AD1} was preferred.

We also tested the two types of structured covariance matrices using a series of regression models:

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 LESION \quad \text{and}$$

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 LESION^2,$$

where the independent variable '*LESION*' denoted the percentage of lesion volume within each resampled MRS voxel. This variable had a range from 0 to 1.

In both Σ_{AR1} and Σ_{AD1} modeled covariance matrices, the estimates of spatial correlation parameter, ρ were both around 0.80, which were very significant. The results for the regression models were shown in Table 2 and Table 3.

By the likelihood ratio tests, the regression models based on Σ_{AD1} were still preferred. In addition, the standard errors of the estimated parameters in the regression models using Σ_{AD1} were smaller than those in the models using Σ_{AR1}. This indicated that proper modeling of covariance matrices could improve inference of regression models.

Regression models with Σ_{AR1}	$-2*L$	$\beta_{1(LESION)}$ (SE)[a]	$\beta_{1(LESION^2)}$ (SE)[a]
$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 Lesion$	3005.1	-0.167 (0.041)	NA[b]
$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 Lesion^2$	3012.9	NA[b]	-0.155 (0.050)

Table 2: Regression result with AR1 modeled covariance matrix

[a] SE is standard error. [b] NA is not applicable.

Regression models with Σ_{AR1}	$-2*L$	$\beta_{1(LESION)}$ (SE)[a]	$\beta_{1(LESION^2)}$ (SE)[a]
$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 Lesion$	-594.6	-0.159 (0.034)	NA[b]
$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 Lesion^2$	-587.3	NA[b]	-0.162 (0.044)

Table 3: Regression result with AD1 modeled covariance matrix

[a] SE is standard error. [b] NA is not applicable.

The tradeoff in using the Σ_{AD1} based regression models were the CPU times. Generally it took more than 10 times longer for the algorithm to converge in the Σ_{AD1} based models as compared to the time used in the Σ_{AR1} based regression models. The other drawback in using Σ_{AD1} was that occasionally the Fisher Score algorithm failed to converge because too many parameters were required to be estimated, especially for the regression models in the longitudinal study. So the results shown in the following sections were all based on the regression models using Σ_{AR1}.

5.1.3 NAA/Cr vs. Lesion

We checked the relationship between NAA/Cr ratios and percentage of lesion volume within each resampled MRS voxel using a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 DIAG(RR) + \beta_2 DIAG(SP) + \beta_3 LESION \times DIAG(RR) + \beta_4 LESION \times DIAG(SP)$$

In this cross-sectional study, '*DIAG*' was a dummy variable to specify three types of study subjects: normal controls, RR MS patients and SP patients. Here we used the "treatment" coding to create dichotomous variables where each level of the categorical variable was contrasted to a specified reference level. In the case of variable '*DIAG*', which had three levels, we specified normal controls as the reference level. We created

two dichotomous variables, one would contrast RR MS patients with the reference level (normal controls) and the other would contrasts SP MS patients with the reference level.

The results are shown in Table 4. NAA/Cr ratios in normal appearing white matter (NAWM) of RR patients decreased by 12.1% and NAA/Cr ratios in NAWM of SP patients decreased by 14.9% as compared to those in white matter (WM) of normal controls. NAA/Cr ratios in the lesions of the MS patients were further decreased, 18.1% (12.1%+6%) for RR patients and 20.6% (14.9%+5.7%) for SP MS patients as compared to those in WM of normal brains.

Variables	Estimated β	Standard Error	Strength [a]	P-value
DIAG(RR)	-0.38	0.095	-12.1%	<0.001
DIAG(SP)	-0.47	0.097	-14.9%	<0.001
LESION×DIAG(RR)	-0.19	0.049	-6.0%	<0.001
LESION×DIAG(SP)	-0.18	0.045	-5.7%	<0.001

Table 4: Cross-sectional study, NAA/Cr vs. LESION

[a] Strength is compared to NAA/Cr in WM of normal control, 3.148(0.4184)

5.1.4 NAA/Cr vs. Disease Duration

We studied the relationship between NAA/Cr ratios and disease duration using a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 DURATION \times DIAG(RR) + \beta_2 DURATION \times DIAG(SP).$$

The variable '*DURATION*' denoted the disease duration in units of years.

Variables	Estimated β	Standard Error	Strength [a]	P-value
DURATION×DIAG (RR)	-0.016	0.0052	-0.51%	0.002
DURATION×DIAG (SP)	-0.011	0.0036	-0.35%	0.003

Table 5: Cross-sectional study, NAA/Cr vs. DURATION

[a] Strength is compared to NAA/Cr in WM of normal control, 3.148(0.4184)

The results in Table 5 demonstrated that NAA/Cr ratios in NAWM of RR patients decreased about 0.51% per year; while NAA/Cr ratios in NAWM of SP patients decreased about 0.35% per year. They were statistically significant.

We would like to know if MS lesions were the major contributors in the decrease of NAA/Cr ratios over the disease durations. We tested the hypothesis using a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 LESION(RR) + \beta_2 LESION(RR) + \beta_3 DURATION \times DIAG(RR) + \beta_4 DURATION \times DIAG(SP).$$

The results in Table 6 showed that after controlling for MS lesions, NAA/Cr ratios in NAWM of RR MS patients decreased about 0.48% per year; while NAA/Cr ratios in NAWM of SP MS patients decreases about 0.32% per year. The results indicated that MS lesions do not change much the effect of disease duration over decreased NAA/Cr in NAWM.

Variables	Estimated β	Standard Error	Strength [a]	P-value
LESION(RR)	-0.179	0.0494	-5.69%	<0.001
LESION(SP)	-0.186	0.0447	-5.91%	<0.001
DURATION×DIAG (RR)	-0.015	0.0050	-0.48%	0.005
DURATION×DIAG (SP)	-0.010	0.0035	-0.32%	0.002

Table 6: Cross-sectional study, NAA/Cr vs. DURATION controlled for lesion

[a] Strength is compared to NAA/Cr of normal control, 3.148(0.4184)

5.1.5 NAA/Cr vs. EDSS

EDSS stands for Kurtzke expanded disability status scale [79], a clinical rating scale to evaluate the disability of MS patients. EDSS can have values from 0 to 10, the higher the EDSS value, the more severe the disability of MS patient is. The regression model to study the correlation between NAA/Cr ratios and disability of MS patients is

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 EDSS \times DIAG(RR) + \beta_2 EDSS \times DIAG(SP).$$

The results in Table 7 showed that for RR and SP patients, NAA/Cr ratios decreased 3.24% and 1.68% respectively for each unit scale increase in EDSS. This indicated that MS patients, who suffered from more sever disabilities as measured by EDSS, normally had lower NAA/Cr ratios in their brains.

Variables	Estimated β	Standard Error	Strength [a]	P-value
EDSS×DIAG (RR)	-0.102	0.024	-3.24%	<0.001
EDSS×DIAG (SP)	-0.053	0.011	-1.68%	<0.001

Table 7: Cross-sectional study, NAA/Cr vs. EDSS

[a] Strength is compared to NAA/Cr of normal control, 3.148(0.4184)

5.1.6 NAA vs. Distance to lesion

We tested the idea if MS lesions would have impact on their neighboring NAWM using a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 DIAG(RR) + \beta_2 DIAG(SP) + \beta_3 DISTANCE \times DIAG(RR) + \beta_4 DISTANCE \times DIAG(SP),$$

where the variable '*DISTANCE*' denoted the distance from the center of the resampled MRS voxel to the edge of the closest T2-weighted MS lesion.

Variables	Estimated β	Standard Error	Strength [a]	P-value
DIAG(RR)	-0.46	0.095	-14.6%	<0.001
DIAG(SP)	-0.57	0.097	-18.1%	<0.001
DISTANCE×DIAG (RR)	0.009	0.001	0.29%	<0.001
DISTANCE×DIAG (SP)	0.015	0.000	0.48%	<0.001

Table 8: Cross-sectional study, NAA/Cr vs. Distance to lesion

[a] Strength is compared to NAA/Cr of normal control, 3.148(0.4184)

The results in Table 8 indicated that NAA/Cr ratios in NAWM of RR MS patients would increase by 0.29% for each millimeter away from MS lesions; while

NAA/Cr ratios in NAWM of SP MS patients would increase by 0.48% for every millimeter away from MS lesions.

5.2 Longitudinal Study

5.2.1 Study Subjects

We selected thirty-seven clinical definite MS patients in the RR phase and seventeen MS patients in the SP phase. Each subject had four exams at 0, 0.5, 1 and 2 years of study. The subjects are listed in Table 9.

Groups	Number	EDSS	DURATION (YEARS)	AGE (YEARS)	Averaged NAA/Cr Over Scan at time 0
RR MS	37	2.4±1.28	9.9 ±7.1	30.8 ± 8.5	2.8 ± 0.23
SP MS	17	6.4 ±1.59	16.2 ± 8.2	46.5 ± 9.3	2.63 ± 0.22

Table 9: Longitudinal Study Subjects

5.2.2 NAA/Cr vs. Lesion

We first checked the relationship between NAA/Cr ratios and percentage of lesion partial volume within each resampled MRS voxel. The regression model used was

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 DIAG(SP) + \beta_2 LESION\ .$$

In this longitudinal study '*DIAG*' is a dummy variable to specify two types of study subjects, RR MS patients SP MS patients. This was different from the situation in the cross-sectional study where there were three types of subjects, including normal controls. We also utilized the "treatment" coding to create dichotomous variables. In

this case, we specified RR patients as the reference level. So we created only one dichotomous variable, which would contrast SP patients with RR patients.

The results are shown in Table 10. NAA/Cr ratios in NAWM of SP MS patients were 6.4% less than those in NAWM of RR patients. In addition, NAA/Cr ratios in lesions of MS patients were less than those in NAWM of RR patients. These results were similar to those from the cross-sectional study.

Variables	Estimated β	Standard Error	Strength [b]	P-value
DIAG(SP)	-0.18	0.064	-6.4%	0.028
LESION	-0.16	0.016	-5.7%	0.000

Table 10: Longitudinal Study, NAA vs. Lesion

[b] Strength is compared to NAA/Cr in NAWM of RR MS patients, 2.7969(0.4389)

5.2.3 NAA/Cr vs. EDSS

We studied the relationship between NAA/Cr ratios and clinical disability of MS patient using a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 EDSS + \beta_2 EDSS \times DIAG(SP).$$

The results are shown in Table 11. NAA/Cr ratios in NAWM of RR MS decreased 0.076 for each unit increase in EDSS. The parameter for the interaction term (EDSS×DIAG) was not significant, which indicated that the correlations between NAA/Cr and EDSS were similar for both RR and SP MS patients.

Variables	Estimated β	Standard Error	Strength [b]	P-value
EDSS	-0.076	0.016	-2.7%	0.027[†]
EDSS×DIAG(SP)	0.003	0.013	NA	0.766

Table 11: Longitudinal Study, NAA vs. EDSS

[b] Strength is compared to NAA/Cr of NAWM of RR-phased MS patients, 2.7969(0.4389)
[†] Statistically significant.

5.2.4 NAA/Cr changes over study time

We evaluated the change of NAA/Cr ratios over study time by a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 STUDYTIME + \beta_2 STUDYTIME \times DIAG(SP).$$

The variable '*STUDYTIME*' denoted the study period in units of years. The results are shown in Table 12. We did not find significant voxel-wise changes of NAA/Cr ratios over a study period of 2 years for both RR and SP patients.

Variables	Estimated β	Standard Error	Strength [b]	P-value
STUDYTIME	0.0377	0.0259	NA	0.1455
STUDYTIME×DIAG(SP)	0.0019	0.0456	NA	0.9669

Table 12: Longitudinal Study, NAA changes over study time

[b] Strength is compared to NAA/Cr of NAWM of RR-phased MS patients, 2.7969(0.4389)

5.3 Copaxone Study

Glatiramer acetate (Copaxone) is a random polypeptide that mimics the antigenic portion of myelin proteins and showed efficiency in the treatment of MS in several clinical trials [80,81]. The first purpose of this study was to see if the treatment of copaxone could have an effect on metabolites in brains of MS, measured as NAA/Cr ratios. The second purpose was to compare our voxel-by-voxel regression method with the classical method which makes use of averaged MRS values.

5.3.1 Study Subjects

We chose 9 untreated MS patients as controls and 15 MS patients treated with copaxone. All are RR MS patients. Each patient had two exams over a study period of

one year. At baseline, there was no significant difference between the two groups in EDSS, DURATION and averaged NAA/Cr. The subjects are listed in Table 13.

Groups	Number	EDSS	DURATION (years)	Averaged NAA/Cr over scan
Treated at baseline	9	2.70±1.74	8.11±8.49	2.81 ± 0.33
Treated at endpoint	9	2.33±1.88	9.11±8.49	2.83±0.36
Untreated at baseline	15	2.61±1.78	12.00±5.58	2.88 ± 0.21
Untreated at endpoint	15	2.94±2.04	13.00±5.58	2.79±0.18

Table 13: Copaxone study subjects

5.3.2 Change of NAA/Cr over study time

We examined the change of NAA/Cr ratios over a study period of 1 year by a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 STUDYTIME + \beta_2 STUDYTIME \times TREAT .$$

In this copaxone study, '*TREAT*' was a dummy variable to specify two types of subjects, treated and untreated MS patients. We used the "treatment" coding to create one dichotomous variable. We specified untreated MS patients as the reference level. So we created one dichotomous variable, which would contrast treated MS patients with untreated MS patients. And the variable '*STUDYTIME*' denoted the study period in unit of years.

We also used the averaged NAA/Cr ratios to check if we could obtain similar results and compare the efficiency between these two methods. The averaged NAA/Cr ratios (AVG_NAA/Cr) were the averaged values of all the valid MRS intensities in each exam. We also used a mixed effect model:

$$AVG\_NAA/Cr_{it} = \beta_0 + \beta_1 STUDYTIME + \beta_2 STUDYTIME \times TREAT + b_i + \varepsilon_{it},$$

where i=1,2, …., 24 for all the subjects in this study and t was 0 for the baseline of the study and 1 for the endpoint. Variable '*TREAT*' was the same dummy variable as in the above voxel-by-voxel case. $b_i \sim N(0,\sigma_b^2)$ denoted the random effect for each subject i, and $\varepsilon_{it} \sim N(0,\sigma^2)$ denoted the residuals. The results were obtained using the function called 'lme' in the 'nlme' library [82] of the statistical application R [83].

The results from both methods are shown in Table 14. Even though the estimated parameters from the two methods had similar signs, both estimated parameters were insignificant using our method. While both of the corresponding estimated parameters were statistically significant using averaged NAA/Cr as dependent variable and both had smaller standard errors.

	Our Voxel-by-Voxel method		Averaged NAA/Cr	
Variable	*STUDYTIME*	*STUDYTIME*× *TREAT*	*STUDYTIME*	*STUDYTIME*× *TREAT*
Estimates	-0.032	0.203	-0.094	0.108
Standard Error	0.1208	0.1462	0.0353	0.0444
P-value	0.789	0.164	**0.0142**†	**0.0238**†

Table 14: Copaxone study, NAA/Cr over study time

† Statistically significant.

5.3.3 NAA/Cr vs. DURATION

We also applied those two types of methods to study the correlation between NAA/Cr ratios and disease duration. Our method used a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 DURATION .$$

While the mixed effect model using averaged NAA/Cr ratios was

$$AVG\_NAA/Cr_{it} = \beta_0 + \beta_1 DURATION + b_i + \varepsilon_{it} .$$

The comparison results of both methods are shown in Table 15.

	Our Voxel-by-Voxel method	Averaged NAA/Cr
Variable	DURATION	DURATION
Estimates	-0.01445	-0.01669
Standard Error	0.00737	0.00405
P-value	**0.049**†	**0.0004**†

Table 15: Copaxone study, NAA/Cr vs. Duration

† Statistically significant.

In this case the parameters from both methods were statistically significant and were similar to the results from the cross-sectional study. While the standard errors of estimates using averaged NAA/Cr were still smaller than those obtained using our method.

5.3.4 NAA/Cr vs. EDSS

Again we compared the two types of methods in the study of the correlation between NAA/Cr ratios and disability of MS. Our method used a regression model

$$NAA/Cr = \beta_{01} + \beta_{02} + ... + \beta_{0V} + \beta_1 EDSS .$$

The mixed effect model using averaged NAA/Cr is

$$AVG\_NAA/Cr_{it} = \beta_0 + \beta_1 EDSS + b_i + \varepsilon_{it} .$$

The results are shown in Table 16. Our method generated a statistically significant estimate -0.058 for variable 'EDSS' while the estimate using averaged NAA/Cr was not significant. The standard errors of estimates from both methods were comparable.

	Our Voxel-by-Voxel method	Use Averaged NAA/Cr
Variable	EDSS	EDSS
Estimates	-0.058	-0.039
Standard Error	0.0282	0.0278
P-value	**0.0392**†	0.1733

Table 16: Copaxone study, NAA/Cr vs. EDSS

† Statistically significant.

6. Discussions and Conclusion

6.1Statistical modeling

The challenge in analysis of MR spectroscopy data is the low signal-to-noise (SNR) ratio of MRS signals. This is especially true for multi-voxel MRS data, which requires appropriate statistical analysis techniques.

Conventional approaches to analyze MRS data are either using region-of-interest (ROI) method or averaging the metabolite values over the whole study region. These methods may be prone to be subjectively biased or may suffer from decreased statistical power. We propose to analyze MRS data using all available information within MRS data. And we studied the relationships between brain metabolites measured as NAA/Cr ratios and the clinical information of MS patients based on the multivariate mixed effect model for repeated measurement.

While modeling repeated measurement data, we need to take into consideration the intra-subject spatial and time correlations. Borrowing an idea from geo-statistics, we used the semi-variogram to show that there were significant spatial correlations between the adjacent intensities of MRS. And the estimated coefficient of the spatial correlations was a significant value of 0.8.

We constructed the covariance matrices using either AR1 type spatial correlations, assuming homogenous variances over different locations, or AD1 type spatial correlations, in which case the homogenous variance assumption was relaxed. The sampled unstructured covariance matrix in Figure 8(a) demonstrated that the

variances at different locations were not homogenous. By the likelihood ratio tests, the regression models based on Σ_{AD1} were preferred over the simpler models based on Σ_{AR1}. Also the standard errors of estimates using Σ_{AD1} were smaller than those using Σ_{AR1}, which would facilitate better statistical inference. But there were several drawbacks in using Σ_{AD1} based regression models. We need to estimate 35 extra parameters, taking far more CPU time. In addition Σ_{AD1} based regression models were prone to fail to converge in some cases during parameter estimation. A better choice would be modeling the variances themselves based on voxel location or other possible information. However, we did not find any apparent trends of variances vs. voxel location.

We did not find the expected significant intra-subject correlations of NAA/Cr with time in our longitudinal study. The first cause may be that the time correlations were actually weak. The second would be that the resampling to standard space was not perfect. But in view of the relative large size of MRS voxel, the effect of registration errors should be small. The main reason is most likely the measurement error and MRS quantification error due to the low SNR of spectroscopic signals.

6.2 Study of MS patients with NAA/Cr

From the cross-sectional study, we found NAA/Cr ratios in NAWM of both RR and SP MS patients were lower than those in WM of normal controls. The decreases of NAA/Cr ratios in MS lesions were even larger, which reflected additional focal neuronal injury inside MS lesions. The decreases of NAA/Cr ratios in NAWM of MS could result from diffuse axonal injury and decreased axonal density, or neuronal metabolic dysfunction [84].

From the longitudinal study, we found NAA/Cr in NAWM of SP MS patients was lower than that in NAWM of RR patients. This would suggest that greater accumulated neuronal injury accompanies the longer disease duration of SP patients. The result was consistent with the finding of the negative correlation between NAA/Cr ratios and disease duration in the cross-sectional study.

In both the cross-sectional and the longitudinal studies, we found weak negative correlation between NAA/Cr ratios and EDSS. This implied that decrease of NAA/Cr ratios and increase in disability of patients is linked. Also the weak correlation suggests both variability in the distribution of NAA metabolite intensities among different patients and also heterogeneity in the progression of MS disease.

We found a small positive correlation between NAA/Cr ratios and the distances to focal MS lesions. This suggested that MS lesions have an impact on neighboring NAWM and play a role in the non-lesional, so called degenerative pathology in MS patients.

From our longitudinal study, NAA/Cr ratios of both RR and SP MS patients were stable over the 2-year study period. It could be that the change patterns of NAA/Cr ratios in the brains of patients were not homogenous and were not monotonic. So the 2-year study period could be too short to obtain any significant alterations among MS patients. Another possible explanation is that the changes of NAA/Cr ratios over this 2-year period were fairly small and our technique was not sensitive enough to detect them because of the low SNR of spectroscopic signals.

6.3 Comparison with simpler statistical methods

In the 1-year copaxone study, we compared our method with the classical method which uses averaged MRS intensities (or metabolite ratios). In the study of treatment effect on NAA/Cr ratios, our method did not find the decrease of NAA/Cr ratios for untreated MS patients and did not show the treatment effect on the changes of NAA/Cr ratios. While the method using averaged NAA/Cr ratios indicated that averaged NAA/Cr ratios for untreated MS patients decreased and also showed that the treatment effect was statistically significant. We also noticed that the standard errors of estimated parameters using our method were substantially larger than those using averaged NAA/Cr ratios in this case.

In the study of the relationship between NAA/Cr ratios and disease duration, the two methods produced similar results and the resulting estimates of the coefficients were similar to the results in the cross-sectional study.

In the study of the relationship between NAA/Cr ratios and disabilities of MS patients measured in EDSS, our method found a statistically significant negative correlation, which was consistent with both the results from the cross-sectional and the longitudinal study. The method using averaged NAA/Cr ratios did not find statistically significant correlation, while the standard errors of the estimated parameters from both methods were comparable.

The results from the copaxone study implied that with respect to statistical power, our method was not better than conventional methods using averaged metabolite values. Our voxel-by-voxel method did use all the information of MRS data and had a larger sample size. But the intra-subject variability of the intensities from the same

location was greater than the intra-subject variability of averaged NAA/Cr ratios because of the low SNR spectroscopic signals. This sizeable voxel-wise intra-subject variability eventually restrained the statistical power gained through utilizing all the MRS intensities.

Our method still has several advantages over conventional methods:

- Being able to incorporate the voxel position information and does not assume the distribution of metabolites are thoroughly identical over different locations in the brains.
- Being able to measure the partial volume effect of MS lesions directly.
- Being able to handle missing data naturally.
- Being able to take care of the possible intra-subject correlations (spatial or time) between MRS intensities.

Our method is not without limitations. One major restriction is that our method is not capable of dealing with high resolution imaging data, which is normally 3-diemensional $256 \times 256 \times 200$ volume.

6.4 Further Work and Possible Improvements

As we know, the variances of MRS intensities at different locations were not homogenous. The AD1 modeled structured covariance matrix (Σ_{AD1}) was too costly because each voxel had its own variance parameter. One possible improvement of our method is to model the variability of the variances of intensities at different locations. This would be more efficient and may converge more easily during the parameter estimation process.

Another possible extension of our method is to incorporate non-linear regression models. Because the relationships between the variables in the clinical studies could be higher order, instead of being simply linear.

6.5 Conclusion

The studies in this thesis demonstrated that MRS combined with conventional MRI techniques could be helpful in measuring diffuse neuronal injury in the brains of MS patients and could make them useful in monitoring the progression of multiple sclerosis. These non-invasive imaging techniques could facilitate more efficient ways of evaluating the treatment and prevention of those diseases.

Our method of analysis has several advantages over conventional methods. It can incorporate position information. It can directly measure the partial volume effect of different tissues (such as lesions). It does not suffer from subjective bias in selecting voxels; instead it makes use of all the information available in MRS. However our method does not show improvement in statistical power due to high voxel-wise variability of MRS intensities, which results from the low SNR MRS signals. The SNR issue should become less important in future with advances in MRS technology, such as improved MRS acquisition and quantification processes, and the use of 3T or higher magnetic field strength scanners.

References

1. Hauser SL. Multiple sclerosis and other demyelinating diseases. In: Isselbacher KJ, Wilson JD, Martin JB, Fauci AS, Kasper DL, eds. Harrison's Principles of Internal Medicine. New York, NY: McGraw-Hill, 1994; 2287-2295.

2. Weinshenker BG. Natural history of multiple sclerosis. Ann Neurol 1994; 36(suppl):S6-S11.

3. Lassmann H. Pathology of multiple sclerosis. In: Compston A, Ebers G, Lassmann H, *et al.*, editors. McAlpine's multiple sclerosis. 3rd ed. London: Churchill Livingstone; 1998. pp. 323–358.

4. Lassmann H. Axonal injury in multiple sclerosis. J Neurol Neurosurg Psychiatry 2003; 74:695–697.

5. Kidd D, Barkhof F, McConnell R, *et al.* Cortical lesions in multiple sclerosis. Brain 1999; 122:17–26.

6. Bö L, Vedeler CA, Nyland HI, *et al.* Subpial demyelination in the cerebral cortex of multiple sclerosis patients. J Neuropathol Exp Neurol 2003; 62:723–732.

7. Peterson JW, Bö L, Mörk S, *et al.* Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. Ann Neurol 2001; 50:389–400.

8. Bö L, Vedeler CA, Nyland H, *et al.* Intracortical multiple sclerosis lesions are not associated with increased lymphocyte infiltration. Mult Scler 2003; 9:323–331.

9. Ikuta F, Zimmerman HM (1976) Distributionof plaques in seventy autopsy cases of multiple sclerosis in the United States.Neurology 26:26–28 Neurol 14:279–287.

10. Lycklama à Nijeholt GJ, Bergers E, Kamphorst W, Bot J,Nicolay K, Castelijns JA, van Waesberghe JHTM, Ravid R, Polman CH, Barkhof F (2001)Post-mortem high-resolution MRI of the spinal cord in multiple sclerosis. Brain 124:154–166

11. Allen IV,McKeown SR (1979) A histological, histochemical and biochemical study of the macroscopically normal white matter in multiple sclerosis. J Neurol Sci 41:81–89.

12. Allen IV,Glover G, Anderson R (1981) Abnormalities in the macroscopically normal white matter in cases of mild or spinal multiple sclerosis.Acta Neuropathol (Berl) Suppl VII:176–178 158–64.

13. Evangelou N,Konz D, Esiri MM, Smith S, Palace J,Mathews PM (2000a) Regional axonal loss in the corpus callosum correlates with cerebral white matter lesion volume and distribution in multiple sclerosis. Brain 123:1845–1849.

14. McDonald WI, Compston A, Edan G *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. Ann Neurol 2001; 50: 121–7.

15. Van Walderveen MAA, Kamphorst W, Scheltens P *et al.* Histopathologic correlate of hypointense lesions on T1-weighted spin-echo MRI in multiple sclerosis. Neurology 1998; 50: 1282–8.

16. Kwock, L. Localized MR spectroscopy: basic principles. Neuroimaging Clinics of North America 8(1998), pp. 713-731.

17. Castillo, M., Kwock, L., and Mukherji, S.K. Clinical applications of proton MR spectroscopy. Ajnr: American Journal of Neuroradiology 17(1996); pp.1-15.

18. Moffett JR, Namboodiri MA, Cangro CB, Neale JH. Immunohistochemical localization of N-acetylaspartate in rat brain. Neuroreport 1991; 2:131-134.

19. Simmons MS, Frondoza CG, Coyle JT. Immunocytochemical localization of N-acetyl aspartate with monoclonal antibodies. Neuroscience 1991; 45:37-45.

20. Clark JB. N-acetylaspartate: a marker for neuronal loss or mitochondrial dysfunction. Dev Neurosci 1998; 20:271-276.

21. D.D. Clarke, S. Greenfield, E. Dicker and L.J. Tirri, A relationship of N-acetylaspartate biosynthesis to neuronal protein systhesis. J. Neurochem. 24 (1975), pp. 479–485.

22. A.F. D'Adamo and F.M. Yatsu, Acetate metabolism in the nervous system. N-acetyl--aspartic acid and the biosynthesis of brain lipid. Expl Brain Res. 13 (1966), pp. 961–965.

23. A.F. D'Adamo, L.I. Gidez and F.M. Yatsu, Acetyl transport mechanisms. Involvement of N-acetyl aspartic acid in de novo fatty acid biosynthesis in the developing rat brain. Expl Brain Res. 5 (1968), pp. 267–273.

24. Burri, C. Stefffen and N. Herschkowitz, N-acetyl--aspar-tate is a major source of acetyl groups for lipid synthesis during rat brain development. Devl Neurosci. 13 (1991), pp. 403–411.

25. C.B. Cangro, M.A.A. Namboodiri, L.A. Sklar, A. Corigliano-Murphy and J.H. Neale, Biosynthesis and immunohistochemistry of N-acetylaspartylglutamate in spinal sensory ganglia. J. Neurochem. 49 (1987), pp. 1579–1588.

26. Larsson HB, Christiansen P, Jensen M, Frederiksen J, Heltberg A, Olesen J, Henriksen O. Localized in vivo proton spectroscopy in the brain of patients with multiple sclerosis. *Magn Reson Med* 1991; 22: 23-31

27. Fu L, Matthews PM, De Stefano N, Worsley KJ, Narayanan S, Francis GS, Antel JP, Wolfson C, Arnold DL. Imaging axonal damage of normal appearing white matter in multiple sclerosis. *Brain* 1998; 121: 103-113.

28. Kapeller P, McLean MA, Griffin CM, Chard D, Parker GJ, Barker GJ, Thompson AJ, Miller DH. Preliminary evidence for neuronal damage in cortical gray matter and normal appearing white matter in short duration relapsing-remitting multiple sclerosis: a quantitative MR spectroscopy imaging study. *J Neurol* 2001; 248: 131-138

29. De Stefano N, Narayanan S, Francis GS, Arnaoutelis R, Tartaglia MC, Antel JP, Matthews PM, Arnold DL. Evidence of axonal damage in the early stages of multiple sclerosis and its relevance to disability. Arch Neurol 2001; 58: 65-70.

30. Tourbah A, Stievenart JL, Abanou A, Iba-Zizen MT, Hamard H, Lyon-Caen O, Cabanis EA. Normal-appearing white matter in optic neuritis and multiple sclerosis: a comparative proton spectroscopy study. Neuroradiology 1999; 41: 738-743

31. De Stefano N, Matthews PM, Antel JP, Preul M, Francis G, Arnold DL. Chemical pathology of acute demyelinating lesions and its correlation with disability. Ann Neurol 1995; 38: 901-909.

32. Lee MA, Blamire AM, Pendlebury S *et al.* Axonal injury or loss in the internal capsule and motor impairment in multiple sclerosis. Arch Neurol 2000; 57: 65–70

33. Pan JW, Krupp LB, Elkins LE, Coyle PK. Cognitive dysfunction lateralizes with NAA in multiple sclerosis. Appl Neuropsychol 2001; 8: 155–60.

34. Arnold DL, Matthews PM, Francis G, Antel J. Proton magnetic resonance spectroscopy of human brain in vivo in the evaluations of multiple sclerosis: assessment of the load of disease. Magn Reson Med 1990;14:154–159.

35. Husted CA, Goodin DS, Hugg JW, Maudsley AA, Tsuruda JS, de Bie SH, Fein G, Matson GB, Weiner MW. Biochemical alterations in multiple sclerosis lesions and normal-appearing white matter detected by in vivo 31P and 1H spectroscopic imaging. Ann Neurol 1994;36:157–165.

36. J Urenjak, SR Williams, DG Gadian and M Noble. Proton nuclear magnetic resonance spectroscopy unambiguously identifies different neural cell types. J. Neuroscience, Vol 13, 981-989.

37. Peters T, Davey B, Munger P, Comeau R, Evans A, Olivier A. Three-dimensional multimodal image-guidance for neurosurgery. IEEE Trans Med Imag 1996; 15:121–128.

38. Arun KS, Huang TS, Blostein SD. Least squares fitting of two 3-D point sets. IEEE Trans PAMI 1987; 5:698–700.

39. Besl PJ, McKay ND. A method for registration of 3D shapes. IEEE Trans PAMI 1992; 14:239–256.

40. Ge Y, Fitzpatrick JM, Kessler RM, Keske-Janicka M, Margolin RA. Intersubject brain image registration using both cortical and subcortical landmarks. In: Loew M, ed. Medical imaging. Bellingham: SPIE Press; 1995:2434.

41. Hill DLG, Hawkes DJ, Crossman JE, et al. Registration of MR and CT images for skull base surgery using point-like anatomical features. Br J Radiol 1991; 64:1030–1035.

42. Thirion J. New feature points based on geometric invariants for 3D image registration. Int J Comp Vision 1996; 18:121–137.

43. Monga O, Benayoun S. Using partial derivatives of 3D images to extract typical surface features. Comput Vision Image Understanding 1995; 61:171–189.

44. Gueziec A, Ayache N. Smoothing and matching of 3D space curves. Int J Comput Vision 1994; 12:79–104.

45. Maintz JBA, van den Elsen PA, Viergever MA. Comparison of edge-based and ridge-based registration of CT and MR brain images. Med Image Analysis 1996; 1:151–161.

46. Pelizzari CA, Chen GTY, Spelbring DR, Weichselbaum RR, Chen CT. Accurate three-dimensional registration of CT, PET and/or MR images of the brain. J Comput Assist Tomogr 1989; 13:20–26.

47. Alpert NM, Bradshaw JF, Kennedy D, Correia JA. The principal axis transformation – a method for image registration. J Nucl Med 1990; 31:1717–1722.

48. Slomka PJ, Hurwitz GA, Stephenson J, Cradduck T. Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm. J Nucl Med 1995; 36:1115–1122.

49. Junck L, Moen JG, Hutchins GD, Brown MB, Kuhl DE. Correlation methods for the centering, rotation and alignment of functional brain images. J Nucl Med 1990; 31:1220–1276.

50. Andersson JLR, Sundin A, Valind S. A method for coregistration of PET and MRI brain images. J Nucl Med 1995; 36:1307–1315.

51. Hill DLG, Studholme C, Hawkes DJ. Voxel similarity measures for automated image registration. Proc SPIE 1994; 2359:205–216.

52. Viola P, Wells III WM. Alignment by maximization of mutual information. Int Conf on Computer Vision. Los Alimitos: IEEE Computer Society Press; 1995:16–23.

53. Wells WM III, Viola P, Atsumi H, Nakajima S, Kikinis R. Multi-modal volume registration by maximization of mutual information. Med Image Anal 1996; 1:35–51.

54. Maes F, Collignon A. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997; 16: 187–198.

55. Roche A, Malandain G, Pennec X, Ayache N. The correlation ratio as a new similarity measure for multimodal image registration. In: Wells WM, Colchester A, Delp S, eds. Lecture notes in computer science. Proc MICCAI'98. Berlin Heidelberg New York: Springer; 1998; 1496:1115–1124.

56. Lau YH, Braun M, Hutton BF. Non-rigid registration using a median-filtered coarse-to-fine displacement filed and a symmetric correlation ratio. Phys Med Biol 2001; 46:1297–1319. coordinate transformation for 3-D image matching. IEEE Trans Med Imag 1997: 16:317–328.

57. N. A. Dodgson, "Quadratic interpolation for image resampling," IEEE Trans. Image Processing, vol. 6, pp. 1322–1326, 1997.

58. H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-26, no. 6, pp. 508–517, 1978.

59. C. R. Appledorn, "A new approach to the interpolation of sampled data," IEEE Trans. Med. Imag., vol. 15, pp. 369–376, 1996.

60. Van Leemput, K., et al., "Automated segmentation of multiple sclerosis lesions by model outlier detection", IEEE Trans. on Medical Imaging, vol. 20, pp. 677-688, Aug. 2001.

61. Y. Miki, R.I. Grossman, J.K. Udupa, S. Samarasekera, M.A. van Buchem, S. Cooney, S.N. Pollack, D.L. Kolson, C. Constantinescu, M. Polansky and L.J. Mannon, Computer-assisted quantitation of enhancing lesions in multiple sclerosis: Correlation with clinical classification. Am. J. Neuroradiol. 18 4 (1997), pp. 705–710.

62. Udupa, J.K., et al., Multiprotocol MR image segmentation in multiple sclerosis: experience with over 1,000 studies. Acad Radiol, 2001. 8(11): p. 1116-26.

63. Zijdenbos, A.P., R. Forghani, and A.C. Evans, Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. IEEE Trans Med Imaging, 2002. 21(10): p. 1280-91.

64. Anbeek P, Vincken KL, van Osch MJ, Bisschops RH, van der Grond J."Probabilistic segmentation of white matter lesions in MR imaging.",Neuroimage. 2004 Mar;21(3):1037-44.

65. Guttmann, C.R., et al., Quantitative follow-up of patients with multiple sclerosis using MRI: reproducibility. J Magn Reson Imaging, 1999. 9(4): p. 509-18.

66. Laid, N.M., and Ware, J.H., "Random Effects models for Longitudinal Data", Biometrics, 1982, 38, 963-974.

67. Ordigde, R.J., Mansfield, P., Lohman, J.A. and Prime, S.B. volume selection using gradients and selective pulses. Ann. NY Acad. Sci. 1987, 508, 376-385.

68. Haase A, Frahm J, Hanicke W, Matthaei D. 1H NMR chemical shift selective imaging. Phys Med Biol 1985; 30: 341-4.

69. Den Hollander JA, Oosterwaal B, Van Vroonhoven H, Luyten PR. Elimination of magnetic field distortions in 1H NMR spectroscopic imaging. Proc Soc Magn Reson Med 1991; 1: 472.

70. de Beer R, van Ormondt D. Analysis of NMR data using time domain fitting procedures. NMR Basic Principles and Progress. Vol 26. Berlin: Springer-Verlag, 1992: 204-248.

71. Simon J. Francis. Automatic Lesion Identification in MRI of Multiple Sclerosis Patients, Master's Thesis, Division of Neuroscience, Department of Neurology and Neurosurgery, McGill University, Montreal, 2004.

72. Collins, D.L., et al., Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput Assist Tomogr, 1994. 18(2): p. 192-205.

73. Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Stuttgart: Thieme, 1988.

74. Evans AC, Marrett S, Neelin P, Collins L, Worsley K, Dai W, et al. Anatomical mapping of functional activation in stereotactic coordinate space. NeuroImage 1992; 1: 43–53.

75. Jones R.H. Unequally spaced longitudinal data with AR(1) serial correlation. Biometrics 1991:47, 161-175.

76. Zimmerman, D. L., N'u nez-Ant'on, V. and ElBarmi, H., 1998. Computational aspects of likelihood based estimation of first-order antedependence models. J. Stat. Comp. Simul./ 60:67-84.

77. Little, R.J.A. and Rubin, D.B. Statistical Analysis with Missing Data. Wiley, New York (196).

78. Robert l. Jennrich, Mark D. Schlucher. Unbalanced Repeated-Measures Models with Structured Covariance matrices. Biometrics, 1986:42, 805-820.

79. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983; 33: 1444-1452.

80. Comi G, Filippi M, Wolinsky JS. European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging—measured disease activity and burden in patients with relapsing multiple sclerosis. Ann Neurol. 2001; 49: 290–297.

81. Flechter S, Vardi J, Pollak L, Rabey JM. Comparison of glatiramer acetate (Copaxone) and interferon beta-1b (Betaferon) in multiple sclerosis patients: an open-label 2-year follow-up. J Neurol Sci. 2002; 197: 51–55.

82. Jose C. Pinheiro, Douglas M. Bates, Mixed-Effects Models in S and S-PLUS, Springer Verlag, New York, 20000.

83. http://www.r-project.org/.

84. Matthews PM, Cianfaglia L, McLaurin J, Cashman N, Sherwin A, Arnold D, et al. Demonstration of reversible decreases in Nacetylaspartate (NAA) in a neuronal cell line: NAA decreases as a marker of sublethal neuronal dysfunction. Proc Soc Magn Reson Med 1995; 1: 147.

85. Caramanos, Z., S. Narayanan, D.L. Arnold. 1H-MRS quantitation of tNA and tCr in patients with multiple sclerosis: a meta-analytic review. Accepted on August 17, 2005 for publication in Brain.