Title
Analysis of longitudinal diffusion weighted imaging data

Author
Masaki, Fumitaro

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Analysis of longitudinal diffusion weighted imaging data

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Biomedical Engineering

by

Fumitaro Masaki

Thesis Committee:
Professor Frithjof Kruggel, Chair
Associate Professor Ana Solodkin
Assistant Professor Gultekin Gulsen

2014

Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Frithjof Kruggel for the useful comments, remarks and engagement through the learning process of this master thesis. His guidance helped me in all the time of research and writing of this thesis. Furthermore I would like to thank my thesis comittee: Dr. Solodkin and Dr. Gulsen for their encouragement. I express my warm thanks to Mr Miyake and Canon for their support. I also appreciate my fellow lab mates: Yang Zhang and Jerod Rasmussen, for the stimulating discussions.

Last but not least, I would like to thank my family for encouraging me to pursue my academic career to study here in the US.

Abstract of The Thesis

Analysis of longitudinal diffusion weighted imaging data

By

Fumitaro Masaki

Master of Engineering in Biomedical Engineering

University of California, Irvine, 2014

Professor Frithjof Kruggel, Chair

The linear model (LM) is typically used to analyze the relationship between imaging data and demographic/cognitive parameters. Each imaging measurement is considered as independent in this model. Recent neuroimaging studies collect time-series data, for which the assumption of independence is invalid. Instead, we use the linear mixed model (LMM) that gives us population effects and subject effects as regression coefficients. A downside of LMM is the computation burden. The purposes of this research are: (1) to develop the tools for analyzing large data, (2) to interpret results, and (3) to demonstrate how to use subject wise information.

We focused on a large dataset of diffusion weighted MR images. 730 images and demographic/cognitive tests were acquired in 176 subjects by the Alzheimer's Disease Neuroimaging Initiative (ADNI). Our code can successfully analyze the 30GB dataset within one day. We found that the population effects coefficients of LMM are roughly similar to the coefficients estimated by LM, and the statistical significance of regressors in the LMM is typically lower than in the LM. We demonstrated that subject wise information can be used to determine an onset of deterioration for each healthy control. Results suggest that this parameter is helpful to model the general aging process.

1 Introduction

1.1 Motivation

Alzheimer's disease (AD) is a type of dementia that leads to irreversible damage to the brain. At earlier stages of AD, patients forget their daily routine activities. Over the course of time, they lose understanding of where they are, are unable to understand time such as seasons and dates, are unable to have conversation with others, and are confused as to what to do next. In addition, they also experience a decrease in their emotional stability. For instance, they may show less facial expressions, along with exhibiting involuntary actions or loss of interests in surroundings and in themselves, etc. They may also show the symptom of depression, sleeping disorders, and eating disorders. Finally, they are fully dependent of help by their surroundings, and die.

According to the report of the Alzheimer's Association [1], the number of patients with AD in the United States is more than five million people, and AD is the sixth leading cause of death. Liesi, the author of this study, calculated the number of patients in 2040 using a longitudinal, population-based study, and he reported the expected number is 13.8 million unless preventive measures are developed [2]. Even now, the cost of AD is a huge burden for the health care system, and it is estimated to be $183 billion in the US [3] at this point. Of course, this burden will increase as the number of patients increase.

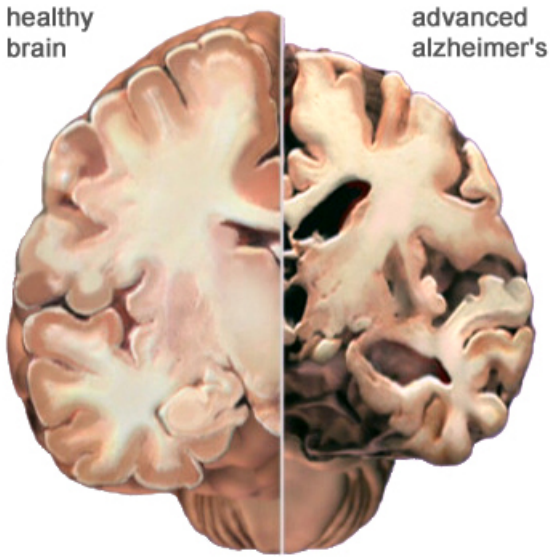

Figure 1: Comparison between healthy brain and AD. The brain of patients with AD is smaller than the brain of healthy people: [8]

Unfortunately, AD is considered as an incurable disease at this point, but several drugs approved by the Food and Drug Administration (FDA) can delay the progression of AD [4].

Slowing down the onset of AD can make it possible for families to make financial plans, organize their living environment, and to search for support from their communities.

What causes AD currently is unknown, but there are several aspects confirmed in the brain of AD patients, such as the presence of amyloid plaques (β-amyloid: Aβ), neurofibrillary tangles, neuronal loss, and cerebral amyloid angiopathy. Representatives from the Alzheimer's Association cite several methods for diagnosing AD, including a thorough medical history, mental status testing, a physical and neurological exam, and tests (such as blood tests and brain imaging) to rule out other causes of dementia-like symptoms. For this purpose, the NINCDS-ADRDA Alzheimer's Criteria is widely used [5].

Imaging methods widely used for diagnosing AD are Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) [7]. Several studies suggest that patients with AD show significant atrophy in the cerebral cortex of the brain, especially, the parahippocampal gyrus and posterior cingulate (Fig. 1). Due to the brain loss with aging, it is difficult to determine who will develop AD on the basis of images. PET, on the other hand, can detect the earliest onset of AD before the brain shows any significant atrophy [9]. β-amyloid (Aβ), one of the causative agents of AD, starts deposition before severe memory dysfunction occurs (Fig. 2). PET depicts how much Aβ is in the brain, the amount of which is associated with the progression of the AD. This is called amyloid PET imaging, and for example, Pittsburgh compound B (PiB) used as a tracer in PET imaging has shown excellent correspondence with Aβ deposition in the brain. While PET can be a powerful tool to find the earliest onset of AD, access to PET is limited. Now, there is great demand to develop the method that can find who will develop AD at an early stage, and which must also be accessible for everyone.

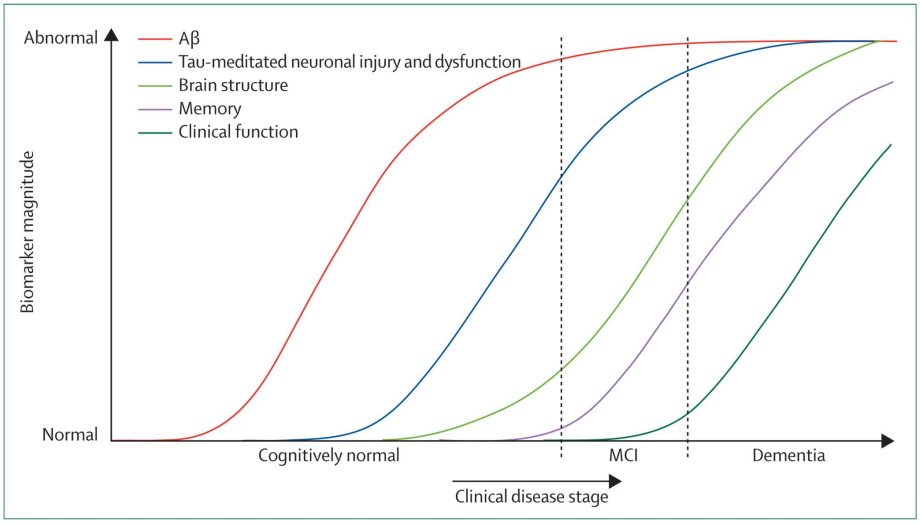

Figure 2: Dynamic biomarkers of the Alzheimer's pathological cascade: [9]

1.2 MRI

General Description
Magnetic Resonance Imaging (MRI) is an imaging modality, which utilizes signals from Nuclear Magnetic Resonance (NMR) with a strong magnetic field [6]. Two-thirds of the human body consist of water, and MRI can detect the signal from these water molecules. Since 1973, strong and stable magnets, reliable detectors, and other hardware have been developed for obtaining higher spatial resolution. Today, MRI is one of the most useful imaging modalities. Compared to other modalities, the advantages of MRI are as follows:

1. In addition to anatomical information, MRI can obtain functional information including brain activity and metabolism.
2. The procedure is non-invasive, and subjects are not exposed to ionizing radiation.
3. MRI has a high contrast between different tissues.
4. Compared to X-ray tomography, bones do not laed to imaging artifacts.

Imaging Protocol
Depending on the purpose of measurement, researchers can select an imaging acquisitions. There are many MRI protocols, and the development of new protocols are an active area of research. Typical imaging methods and their features are as follows.

1. T1 weighted imaging
   This imaging method shows anatomical structures of the brain. Water has low signal and appears dark.
2. T2 weighted imaging
   This image weighting is useful for detecting edema, revealing white matter lesions.
3. Fluid-attenuated inversion recovery (FLAIR)
   The FLAIR procedure suppresses signals emitted from cerebrospinal fluids (CSF), so this procedure is used for imaging periventricular hyperintense lesions.
4. Diffusion Weighted Imaging (DWI)
   DWI can visualize the diffusion of water molecules. Details are shown in the next section.

Diffusion Weighted Imaging (DWI)
DWI measures the diffusion of water molecules inside the body. Because the diffusion of the water molecules in the brain is restricted by obstructions in the environment, such as cell membranes, vascular structures and axons, DWI gives us information about microscopic details of the structure of the brain. The left panel of Fig. 3 shows schematic representation of elongate tissues such as the white matter (axon). The direction of greatest diffusivity is assumed to be parallel to the local direction of the white matter.

The signal ratio diffusion-weighted (S) to non diffusion-weighted signal (S_0) is:

$$\frac{S}{S_0} = exp[-\gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3}) D] = exp[-bD], \tag{1}$$

where γ is the gyromagnetic ratio, G is the strength of the gradient pulse, δ is the duration of the pulse, Δ is the time between the two pulses, and D is the diffusion coefficient. From the above equation, D is derived as:

$$D = -\frac{log(S/S_0)}{b} \tag{2}$$

The spatial anisotropy of the diffusivity is characterized by at least six different measurements with different magnetic gradients. Thus the diffusion is modeled as a 3x3 tensors. In order to analyze the results of DWI quantitatively, we use several index values, such as fractional anisotropy (FA), mean diffusivity (MD), and radial diffusivity (RD). The diffusivity of each voxel is described as an ellipsoid of three eigenvectors $(\epsilon_1, \epsilon_2, \epsilon_3)$ and eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ (shown in the right image of Fig. 3) [11]. The three eigenvectors correspond to the direction of diffusion, and the three eigenvalues describe the magnitude of diffusivity along each eigenvector. FA, MD, and RD are defined as follows:

$$FA = \sqrt{\frac{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{3}$$

$$MD = \frac{(\lambda_1 + \lambda_2 + \lambda_3)}{3} \tag{4}$$

$$RD = \frac{(\lambda_2 + \lambda_3)}{2}, \lambda_1 > \lambda_2, \lambda_3 \tag{5}$$

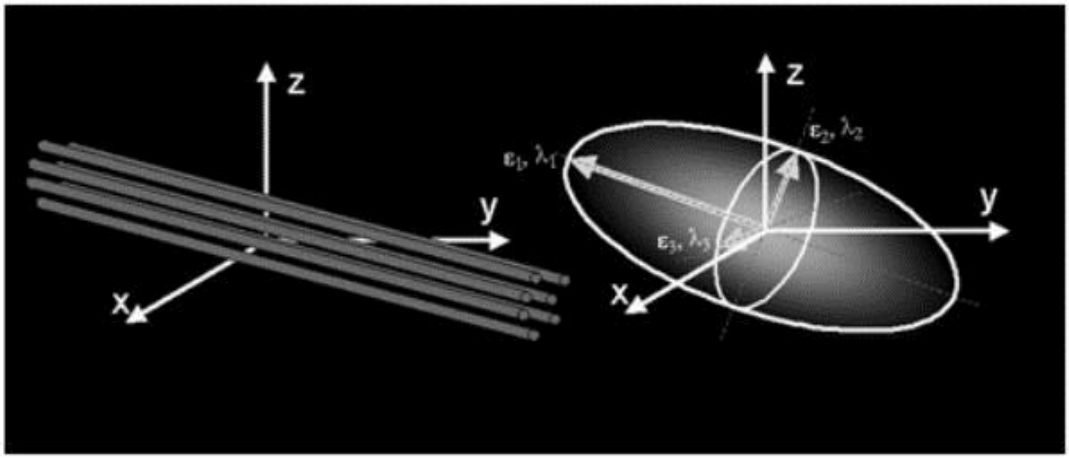

Figure 3: The three-dimensional diffusivity as an ellipsoid: [11]

By definition, FA shows a value from 0 to 1, the former meaning that the diffusion in the voxel is completely isotropic and the latter meaning that the direction of the water molecules in

the voxel is completely anisotropic. As explained above, motion of water molecules in the white matter area is restricted by the myelin. Therefore, if the myelin is damaged, FA will typically be smaller. MD will typecally be larger because MD shows how large the diffusion in the voxel is, and the damaged myelin will make more room for the extracellular water. Solodkin et al [10] reports that, by comparing the AD patients to health controls, some indexes of DTI (e.g. FA and MD) show significant group differences.

1.3 Linear Mixed Model (LMM)

1.3.1 Linear Model (LM) and its Limitations

Several studies show that the size of the brain shrinks as people age. For example, Hommer et al [12] reported that the brain size of male and female linearly decrease as they get older. We can assume the relationship between two or more parameters in order to obtain general trends and see how parameters affect each other. Because of its simplicity, the linear model (LM) is widely used. LM is described as:

$$y = X\beta + \epsilon, \tag{6}$$

where y is a vector of n observations, X is a $n * p$ regressor matrix, β is a p-dimensional vector of regression coefficients, and ϵ is a n-dimensional vector of residuals. The mean of y is calculated as $E(y) = X\beta$. The regression coefficients β are estimated by

$$\beta = (X^t X)^{-1} X^t y = X^+ y \tag{7}$$

Only q of p regressors may have a significance on the prediction of observations y. The F-test for a reduced model and the complete model is described as follows.

$$f = \frac{n-p}{q} \frac{SSH}{SSE}, \tag{8}$$

where SSH and SSE are the error variance for the reduced and complete model. A p-value is calculated from this f test, and then the p-value is converted to z-score using the inverse of the error function.

$$z(p) = \sqrt{2} erfc^{-1}(2p-1), p \in [0,1] \tag{9}$$

Typically, regression coefficients with $p < 0.05 (z > 1.96)$ are considered as significant.

subject ID	Exam no.	Age	GEN	WT	GRP
4119	40764	79.42	M	79.38	HC
4119	40960	79.95	M	75.3	HC
4119	41149	80.47	M	77.11	HC
4119	41486	81.39	M	75.75	HC
4136	40765	66.76	M	94.35	AD
4136	40913	67.17	M	88.45	AD
4136	41170	67.87	M	90.5	AD
4136	40989	67.38	M	91	AD
4288	40836	72.54	F	72.58	HC
4288	40953	72.86	F	86.18	HC
4288	41053	73.13	F	72.12	HC
4288	41214	73.58	F	73.48	HC
4288	41613	74.67	F	72.57	HC

Figure 4: A part of the data of this research: GEN=gender (Male or Female), WT=weight, GRP=clinical grouping (Healthy Controls (HC), Mild Cognitive Impairment (MCI), or Alzheimer's Disease (AD))

Fig. 4 shows a part of the data of this research. For example, age, gender, weight, and the clinical grouping can be used as regressors in X. Gender and the clinical grouping are categrical variables. As for gender, the column in X corresponding to gender has 0 for male, and 1 for female. As for the clinical grouping, two columns are assigned in X for describing the clinical grouping, one of which corresponds to MCI, and the other corresponds to AD. 1 in the column for MCI means that the subject is MCI at that measurement, 1 in the column for AD means that the subject is AD at that test.

The LM assumes that each observation is independent. In reality, however, patients in clinical studies often take several tests over the course of years. In this case, observations within each patient are not independent. This means that the LM cannot handle this model correctly. In our case, as shown in Fig. 4, each patient has taken several MRI acquisitions. In this case, we want to know the trends of DTI scalars that are common in all patients, as well as the differences between each patient. The former refers to population effects, and the latter refers to subject effects.

1.3.2 Linear Mixed Model (LMM)

The linear mixed model (LMM) is used to explicitly analyze the data where repeated measurements are made on the same statistical units. The details and theory of LMM will be described in the later section. Basically, the LMM extends the LM by considering the differences between each patient. Observations y are described as:

$$y = X\beta + U\gamma + \epsilon, \tag{10}$$

where y is a vector of observations, with mean $E(y) = X\beta$, β is a vector of regression coefficients for population effects, γ is a vector of regression coefficients for subject effects with mean $E(\gamma) = 0$, and variance-covariance matrix $\mathrm{var}(\gamma) = G$. ϵ is a vector of residuals with mean $E(\epsilon) = 0$ and variance $\mathrm{var}(\epsilon) = R$. X and U are matrices of regressors relating the observations y to β and γ, respectively.

The first term on the right-hand side corresponds to the population effects, and the second term represents the subject effects, which is unique in each subject. In the example of the brain measurement above, the first term corresponds to the dependence of brain size and age with respect to the whole sample, and the second term describes the difference in each subject. By analyzing the subject effects, we can detect individual inferences on brain degeneration. For our AD research, we will study both effects.

1.4 Aims

The final goal of developing the imaging technology for AD is to detect the onset of AD as early as possible, which will make it possible to delay or stop the progression of AD at the early stage, and increase patients' quality of life. LMM can greatly help to analyze images because the model gives us subject wise information as explained above. However, a downside of LMM is the much increased computation burden. So far, no useful LMM code was available that could be applied to big data. The aims of this research are:

1. To develop the tools that can analyze a large amount of images of the brain.

There are several software programs, like R, that allows us to analyze statistical data. These software programs are widely used in many academic fields, and greatly benefits us with this data. However, in regards to the data size, the computational burden is one of the most difficult problems in analyzing statistical data, and current software cannot complete the image analysis of our data within a reasonable time frame. Therefore, we employ compiled code and parallel processing to achieve an acceptable analysis time.

2. To interpret results.

We use an exploratory strategy to determine which variables such as age, weight and gender significantly influence brain structures.

3. To demonstrate how to use subject-wise information.

The LMM yields population information as β, and subject-wise information as γ. Because applying the LMM to big data about brain aging is new, it is unclear whether the subject related information formed in γ contains neurobiologically relevant information. Through this study, we try to demonstrate an application that provides meaningful interpretation of subject wise information.

2 Method

In the first part of this section, the theory of LMM will be explained, and then the brief introduction of the post-process module for spatial correlation will be described.

2.1 Theory of LMM

2.1.1 Definitions

Consider $i = \{1, \dots, m\}$ units measured at $j = \{1, \dots, n_i\}$ timepoints with observations y_{ij}. Let us denote x_{ij} as the covariate vector of the i-th unit at the j-th measurement for population effects $\boldsymbol{\beta} \in \mathbb{R}^p$, and u_{ij} as the covariate vector of the i-th unit at the j-th measurement for subject effects $\boldsymbol{\gamma}_i \in \mathbb{R}^q$ (i.e., the unit-specific effects). The general form of the linear mixed effects model writes [13]:

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{U}_i\boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i \quad \text{where} \quad \boldsymbol{\gamma}_i \sim \mathcal{N}_q(0, \mathbf{D}_i), \boldsymbol{\epsilon}_i \sim \mathcal{N}_{n_i}(0, \boldsymbol{\Sigma}_i), \operatorname{cov}(\boldsymbol{\gamma}_i, \boldsymbol{\epsilon}_i) = 0. \tag{11}$$

Now, denote the column vector of obervations:

$$\mathbf{y} := \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \in \mathbb{R}^n \quad \text{where} \quad n := \sum_{i=1}^{m} n_i; \tag{12}$$

the population effects matrix and coefficients:

$$\mathbf{X} := \begin{pmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_m \end{pmatrix} \in \mathbb{R}^{n \times p} \quad \text{and} \quad \boldsymbol{\beta} \in \mathbb{R}^p; \tag{13}$$

the subject effects matrix and coefficients:

$$\mathbf{U} := \begin{pmatrix} \mathbf{U}_1 & \mathbf{0}_{n_1 \times q} & \dots & \mathbf{0}_{n_1 \times q} \\ \mathbf{0}_{n_2 \times q} & \mathbf{U}_2 & & \\ \vdots & & \ddots & \\ \mathbf{0}_{n_m \times q} & & & \mathbf{U}_m \end{pmatrix} \in \mathbb{R}^{n \times (m\,q)} \quad \text{and} \quad \boldsymbol{\gamma} := \begin{pmatrix} \boldsymbol{\gamma}_1 \\ \vdots \\ \boldsymbol{\gamma}_m \end{pmatrix} \in \mathbb{R}^{m\,q}; \tag{14}$$

and errors:

$$\boldsymbol{\epsilon} := \begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \vdots \\ \boldsymbol{\epsilon}_m \end{pmatrix} \in \mathbb{R}^n; \tag{15}$$

and define $\mathbf{G}$ as the diagonal block matrix of m matrices $\mathbf{D}_i \in \mathbb{R}^{q \times q}$, and $\mathbf{R}$ as the diagonal block matrix of m matrices $\boldsymbol{\Sigma}_i \in \mathbb{R}^{n_i \times n_i}$.

2.1.2 Estimation

Consider the linear mixed effects model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \quad \begin{pmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\epsilon}\end{pmatrix} \sim \mathcal{N}_{mq+n}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\mathbf{G}(\boldsymbol{\theta}) & \mathbf{0}_{mq\times n}\\ \mathbf{0}_{n\times mq} & \mathbf{R}(\boldsymbol{\theta})\end{pmatrix}\right) \tag{16}$$

with covariance vector $\boldsymbol{\theta}$. Using $\mathbf{V}(\boldsymbol{\theta}) = \mathbf{U}\,\mathbf{G}(\boldsymbol{\theta})\,\mathbf{U}^T + \mathbf{R}(\boldsymbol{\theta})$, the population effects $\boldsymbol{\beta}$ and the subject effects $\boldsymbol{\gamma}$ are estimated by:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\,\mathbf{V}(\hat{\boldsymbol{\theta}})^{-1}\,\mathbf{X}\right)^{-1}\mathbf{X}^T\,\mathbf{V}(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y} \quad \text{and} \quad \hat{\boldsymbol{\gamma}} = \mathbf{G}\hat{\mathbf{U}}^T\mathbf{V}(\hat{\boldsymbol{\theta}})^{-1}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right). \tag{17}$$

The covariance vector $\boldsymbol{\theta}$ is typically estimated using Restricted Maximum Likelihood (REML) by:

$$\hat{\boldsymbol{\theta}} = \arg\max\left(l_p(\boldsymbol{\theta}) - \frac{1}{2}\ln\left|\mathbf{X}^T\,\mathbf{V}(\boldsymbol{\theta})^{-1}\,\mathbf{X}\right|\right). \tag{18}$$

Now, we spell out details of the iterative estimation process. For the special case $\boldsymbol{\Sigma}_i = \sigma^2\mathbf{I}_{n_i}$, the covariance vector $\boldsymbol{\theta}$ is comprised of $\boldsymbol{\theta} = \{\sigma^2, \mathbf{D}\}$ [14]. Note that this vector has $s = q(q+1)/2 + 1$ unique elements.

We start from OLS estimates for $\hat{\boldsymbol{\beta}}$ and compute residuals $\hat{\mathbf{r}}$, leading to OLS estimates for $\hat{\boldsymbol{\gamma}}$:

$$\hat{\boldsymbol{\beta}}_0 = \left(\sum_{i=1}^m \mathbf{X}_i^T\mathbf{X}_i\right)^{-1}\sum_{i=1}^m \mathbf{X}_i^T\mathbf{y}_i \quad \text{and} \quad \hat{\mathbf{r}}_{0,i} = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_0, \tag{19a}$$

$$\hat{\boldsymbol{\gamma}}_{0,i} = \left(\mathbf{U}_i^T\mathbf{U}_i\right)^{-1}\mathbf{U}_i^T(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_0) = \mathbf{U}_i^+\hat{\mathbf{r}}_{0,i}. \tag{19b}$$

Next, we compute starting values for $\hat{\sigma}^2$ and $\hat{\mathbf{D}}$:

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^m \mathbf{y}_i^T\mathbf{y}_i - \hat{\boldsymbol{\beta}}_0^T\sum_{i=1}^m \mathbf{X}_i^T\mathbf{y}_i - \sum_{i=1}^m \hat{\boldsymbol{\gamma}}_{0,i}^T\mathbf{U}_i^T\hat{\mathbf{r}}_{0,i}}{n - (m-1)q - p} \tag{20a}$$

$$\hat{\mathbf{D}}_0 = \frac{1}{m}\sum_{i=1}^m \hat{\boldsymbol{\gamma}}_{0,i}\hat{\boldsymbol{\gamma}}_{0,i}^T - \frac{\hat{\sigma_0}^2}{m}\sum_{i=1}^m \left(\mathbf{U}_i^T\mathbf{U}_i\right)^{-1}. \tag{20b}$$

Now, we can iterate the EM algorithm. First, the e-step consists of estimating the covariance matrix $\hat{\mathbf{V}}_t$, then $\hat{\boldsymbol{\beta}}_t$ and $\hat{\boldsymbol{\gamma}}_{t,i}$, where the subscript t denotes the iteration:

$$\hat{\mathbf{V}}_{t,i} = \mathbf{U}_i\hat{\mathbf{D}}_t\mathbf{U}_i^T + \hat{\sigma}_t^2\mathbf{I}_{n_i}, \tag{21a}$$

$$\hat{\boldsymbol{\beta}}_t = \hat{\mathbf{A}}_t^{-1}\sum_{i=1}^m \mathbf{X}_i^T\hat{\mathbf{V}}_{t,i}^{-1}\mathbf{y}_i \quad \text{using} \quad \hat{\mathbf{A}}_t = \sum_{i=1}^m \mathbf{X}_i^T\hat{\mathbf{V}}_{t,i}^{-1}\mathbf{X}_i \tag{21b}$$

$$\hat{\boldsymbol{\gamma}}_{t,i} = \hat{\mathbf{D}}_t\mathbf{U}_i^T\hat{\mathbf{V}}_{t,i}^{-1}\hat{\mathbf{r}}_{t,i}, \tag{21c}$$

Next, the m-step consists of estimating $\hat{\sigma}^2$ and $\hat{\mathbf{D}}$:

$$\hat{\mathbf{P}}_{t,i} = \hat{\mathbf{V}}_{t,i}^{-1}\left[\mathbf{I} - \mathbf{X}_i\hat{\mathbf{A}}_t^{-1}\mathbf{X}_i^T\hat{\mathbf{V}}_{t,i}^{-1}\right], \tag{22a}$$

$$\hat{\sigma}_{t+1}^2 = \frac{1}{n}\sum_{i=1}^{m}\left[(\hat{\mathbf{r}}_{t,i} - \mathbf{U}_i\hat{\boldsymbol{\gamma}}_{t,i})^T(\hat{\mathbf{r}}_{t,i} - \mathbf{U}_i\hat{\boldsymbol{\gamma}}_{t,i}) + \hat{\sigma}_t^2\,\mathrm{tr}(\mathbf{I} - \hat{\sigma}_t^2\hat{\mathbf{P}}_{t,i})\right], \tag{22b}$$

$$\hat{\mathbf{D}}_{t+1} = \frac{1}{m}\sum_{i=1}^{m}\left[\hat{\boldsymbol{\gamma}}_{t,i}\hat{\boldsymbol{\gamma}}_{t,i}^T + \hat{\mathbf{D}}_t(\mathbf{I} - \mathbf{U}_i^T\hat{\mathbf{P}}_{t,i}\mathbf{U}_i\hat{\mathbf{D}}_t)\right], \tag{22c}$$

Convergence is checked by observing the log-likelihood:

$$L_{REML}(\boldsymbol{\theta}) = \frac{1}{2}\left(\sum_{i=1}^{m}\left[\ln|\hat{\mathbf{V}}_i^{-1}| - \mathbf{r}_i^T\hat{\mathbf{V}}_i^{-1}\mathbf{r}_i\right] - \ln|\hat{\mathbf{A}}|\right). \tag{23}$$

2.1.3 Inference

Recall that:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\,\mathbf{V}(\hat{\boldsymbol{\theta}})^{-1}\,\mathbf{X}\right)^{-1}\mathbf{X}^T\,\mathbf{V}(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y} \quad \text{so} \quad \mathrm{var}(\hat{\boldsymbol{\beta}}) = \hat{\mathbf{A}}^{-1}. \tag{24}$$

Thus:

$$\hat{\beta}_j \pm z_{1-\alpha/2}\sqrt{\hat{\mathbf{A}}^{-1}} \tag{25}$$

gives an approximate $100(1-\alpha)\%$ confidence interval for β_j. Likewise, for $\boldsymbol{\gamma}$ we have:

$$\mathrm{var}(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) = \hat{\mathbf{D}}_i\left[\mathbf{I} - \mathbf{U}_i^T\hat{\mathbf{P}}_i\mathbf{U}_i\hat{\mathbf{D}}_i\right]. \tag{26}$$

Suppose that inferences are to be made simultaneously about l linear combinations of the elements of $\boldsymbol{\beta}$: $\mathbf{L}\boldsymbol{\beta}$, for a $(l \times p)$ population effects matrix $\mathbf{L}$. We examine the hypothesis $H_0 : \mathbf{L}\hat{\boldsymbol{\beta}} = \mathbf{d}$ vs. $H_1 : \mathbf{L}\hat{\boldsymbol{\beta}} \neq \mathbf{d}$ using a Wald-type test:

$$\frac{1}{l}\left(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{d}\right)^T\left[\mathbf{L}\hat{\mathbf{A}}^{-1}\mathbf{L}^T\right]^{-1}\left(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{d}\right) \sim F_{(\alpha,l,\nu)}, \tag{27}$$

where $l = \mathrm{rank}(L)$ denote the nominator and ν the denominator degrees-of-freedom, respectively.

For a single regressor of interest r and $d = 0$, this F-test simplifies to a t-test:

$$\frac{\beta_r}{\sqrt{\hat{\mathbf{A}}_{r,r}^{-1}}} \sim t_{(\alpha,\nu)}. \tag{28}$$

The degrees-of-freedom ν have to be estimated from the data. For m large, the t-distribution does not depend much on ν, so we can set $\nu = m - 1$ (inference option MLarge). A second

option is to use a Satterthwaite approximation for $\hat{\nu}$ (inference option Sat):

$$\hat{\nu_r} = \frac{\left(\hat{\mathbf{A}}^{-1}_{r,r}\right)^2}{\hat{\boldsymbol{g}}^T \hat{\mathbf{H}}^{-1} \hat{\boldsymbol{g}}}, \tag{29}$$

where $\hat{\boldsymbol{g}}$ corresponds to the vector of first partial derivates of $\hat{\mathbf{A}}^{-1}$ with respect to the s components of the covariance vector $\boldsymbol{\theta}$, i.e., $\hat{\boldsymbol{g}} = \{\frac{\partial \hat{\mathbf{A}}^{-1}}{\partial \boldsymbol{\theta}_1}, \ldots, \frac{\partial \hat{\mathbf{A}}^{-1}}{\partial \boldsymbol{\theta}_s}\}^T$, and $\hat{\mathbf{H}}$ corresponds to the matrix of second partial derivatives of the likelihood function L_{REML} with respect to the s components of the covariance vector $\boldsymbol{\theta}$, i.e.,

$$\hat{\mathbf{H}} = \begin{pmatrix} \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1} & \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_1} & \cdots & \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_s} \\ \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_1} & \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2} & & \\ \vdots & & \ddots & \\ \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_s \partial \boldsymbol{\theta}_1} & & & \frac{\partial^2 L_{REML}}{\partial \boldsymbol{\theta}_s \partial \boldsymbol{\theta}_s} \end{pmatrix} \tag{30}$$

Note that $\hat{\boldsymbol{g}}$ and $\hat{\mathbf{H}}$ can be computed just once and re-used for each regressor. Currently, derivatives are computed numerically using finite differences, although derivatives are available in explicit form [15].

For $l \geq 1$ and n small, Kenward and Roger [16] provided a modification (inference option KR):

$$\left(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{d}\right)^T \left[\mathbf{L}\hat{\mathbf{A}}^{-1}_{KR}\mathbf{L}^T\right]^{-1} \left(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{d}\right) = \lambda F_{(\alpha,l,m)}, \tag{31}$$

where $\hat{\mathbf{A}}^{-1}_{KR}$ denotes an adjusted covariance matrix:

$$\hat{\mathbf{A}}^{-1}_{KR} = \hat{\mathbf{A}}^{-1} + 2\,\hat{\mathbf{A}}^{-1}\hat{\mathbf{Z}}\hat{\mathbf{A}}^{-1}, \tag{32}$$

For the computation of matrix $\hat{\mathbf{Z}}$, refer to ([17], pp. 22ff). Form s symmetric incidence matrices $\mathbf{E}_i, i = \{1, \ldots, s\}$ of dimensions $q \times q$ with ones indicating the positions of unique elements in $\hat{\mathbf{D}}$. The last matrix (corresponding to $\hat{\sigma}^2$) equals I_q. Form the auxillary matrices:

$$\mathbf{G}_{i,k} = \mathbf{U}_k \mathbf{E}_i \mathbf{U}_k^T, \quad n_i \times n_i, \quad i = \{1, \ldots, s\}, k = \{1, \ldots, m\}, \tag{33a}$$

$$\hat{\mathbf{H}}_{i,k} = \mathbf{G}_{i,k}\hat{\mathbf{V}}_k^{-1}\mathbf{X}_k, \quad n_i \times p, \quad i = \{1, \ldots, s\}, k = \{1, \ldots, m\}, \tag{33b}$$

$$\hat{\mathbf{T}}_{i,k} = -(\hat{\mathbf{V}}_k^{-1}\mathbf{X}_k)^T \hat{\mathbf{H}}_{i,k}, \quad p \times p, \quad i = \{1, \ldots, s\}, k = \{1, \ldots, m\}, \tag{33c}$$

$$\hat{\mathbf{Q}}_{i,j,k} = \hat{\mathbf{H}}_{i,k}^T \hat{\mathbf{V}}_k^{-1} \hat{\mathbf{H}}_{j,k}, \quad p \times p, \quad i,j = \{1, \ldots, s\}, k = \{1, \ldots, m\}, \tag{33d}$$

$$\hat{K}_{i,j} = \sum_{k=1}^{m} \left[\mathrm{tr}(\hat{\mathbf{V}}_k^{-1}\mathbf{G}_{i,k}\hat{\mathbf{V}}_k^{-1}\mathbf{G}_{j,k}) - 2\,\mathrm{tr}(\hat{\mathbf{A}}^{-1}\hat{\mathbf{Q}}_{i,j,k}) + \mathrm{tr}(\hat{\mathbf{A}}^{-1}\hat{\mathbf{T}}_{i,k}\hat{\mathbf{A}}^{-1}\hat{\mathbf{T}}_{j,k})\right], \tag{33e}$$

$$s \times s, \quad i,j = \{1, \ldots, s\}, \quad \text{to finally yield:} \tag{33f}$$

$$\hat{\mathbf{Z}} = \sum_{i=1}^{s}\sum_{j=1}^{s} \hat{K}_{i,j}^{-1} \sum_{k=1}^{m} \left(\hat{\mathbf{Q}}_{i,j,k} - \hat{\mathbf{T}}_{i,k}\hat{\mathbf{A}}^{-1}\hat{\mathbf{T}}_{j,k}\right) \quad p \times p, \quad i,j = \{1, \ldots, s\}, \tag{33g}$$

The constant λ is computed as:

$$\hat{\boldsymbol{\Theta}} = \mathbf{L}^T \left(\mathbf{L} \hat{\mathbf{A}}^{-1} \mathbf{L}^T \right)^{-1} \mathbf{L}, \quad \hat{\boldsymbol{\Phi}}_{i,k} = \hat{\boldsymbol{\Theta}} \hat{\mathbf{A}}^{-1} \hat{\mathbf{T}}_{i,k} \hat{\mathbf{A}}^{-1} \tag{34a}$$

$$a_1 = \sum_{i=1}^{s} \sum_{j=1}^{s} \hat{W}_{i,j} \sum_{k=1}^{m} \mathrm{tr}(\hat{\boldsymbol{\Phi}}_{i,k}) \mathrm{tr}(\hat{\boldsymbol{\Phi}}_{j,k}), \tag{34b}$$

$$a_2 = \sum_{i=1}^{s} \sum_{j=1}^{s} \hat{W}_{i,j} \sum_{k=1}^{m} \mathrm{tr}(\hat{\boldsymbol{\Phi}}_{i,k} \hat{\boldsymbol{\Phi}}_{j,k}), \tag{34c}$$

$$b = \frac{1}{2l}(a_1 + 6a_2), \quad g = \frac{(l+1)\, a_1 - (l+4)\, a_2}{(l+2)\, a_2}, \tag{34d}$$

$$c_1 = \frac{g}{3l + 2(1-g)}, c_2 = \frac{l-g}{3l+2(1-g)}, c_3 = \frac{l+2-g}{3l+2(1-g)}, \tag{34e}$$

$$\hat{E} = \frac{1}{1 - a_2/l}, \quad \hat{V} = \frac{2}{l} \left[\frac{1 + c_1 b}{(1 - c_2 b)^2 \, (1 - c_3 b)} \right], \tag{34f}$$

$$\hat{\rho} = \frac{\hat{V}}{2\hat{E}^2}, \quad \hat{m} = 4 + \frac{l+2}{l\,\hat{\rho} - 1}, \quad \text{and} \quad \lambda = \frac{\hat{m}}{\hat{E}(\hat{m} - 2)}. \tag{34g}$$

2.1.4 Implementation

Interface definition:

- Input: unit-wise population and subject effect matrices $\mathbf{X}_i, \mathbf{U}_i$ (vector of matrices), unit-wise observations $\mathbf{y}_i$ (vector of vectors), convergence limit l, and inference option o.
- Output: fixed effects coefficients and associated significance $\hat{\boldsymbol{\beta}}, p_{\hat{\boldsymbol{\beta}}}$ (vectors).

Algorithm:

1. Compute n, determine m.
2. Initialize $\hat{\boldsymbol{\beta}}_0$ and $\hat{\boldsymbol{\gamma}}_i$ (Eq. 19), then $\hat{\sigma}_0^2$ and $\hat{\mathbf{D}}_0$ (Eq. 20).
3. Iterate:
4. Perform the E-step by computing $\mathbf{V}(\hat{\boldsymbol{\theta}})_{t,i}, \hat{\boldsymbol{\beta}}_t$, and $\hat{\boldsymbol{\gamma}}_{t,i}$ (Eq. 21).
5. Perform the M-step by computing $\hat{\mathbf{P}}_{t,i}, \hat{\sigma}_t^2$, and $\hat{\mathbf{D}}_t$ (Eq. 22).
6. Compute log-likelihood L_{REML} (Eq. 23).
7. If $dL < l$ exit, else goto step 3.
8. Compute inference on $\hat{\boldsymbol{\beta}}$ using inference option o (Eq. 27).

2.2 Spatial correlation

In this study, we adopt $p < 0.05$ as our criteria to decide whether the regressors have significance or not. However, as the number of voxels increases, type-1 error also increases. For example, when we discuss one million voxels, fifty thousand voxels would have type-1 error. One of the methods we used to avoid type-1 error is the Bonferroni correction. Because this method is relatively simple, it is widely used to correct type-1 error. However, the one disadvantage of the Bonferroni correction is that this correction is too conservative. In this research, we have developed two sets of codes for multiple comparisons. One is the FW method (Friston-Worsley method) [18], and the other is the BH procedure (Benjaminie-Hochberg procedure) [19].

Biologically, tissues are continuous objects, so voxels touching each other should have similar statistical results. If one voxel shows completely different statistical results from neighboring voxels, this statistical result may be wrong. The FW method has introduced this idea for the correction. On the other hand, the BH method is a powerful method for controlling the family-wise error rate. However, this correction can be used for any other purpose besides a brain image analysis. This spatial correlation module is optional in order to correct inferences as post-process procedure.

2.3 Implementation

2.3.1 Software and Validation

In statistics, R, a free software, is widely used. R has many functions for almost all statistical purposes including finance, agriculture, and medicine. Several implementations of LMM are available in R, and were validated by the researchers all over the world. lme4, one of the functions of R calculating LMM, is suitable for our purpose, but the only problem that R has for our research is its computational speed. R is an interpreted language, and because we deal with very large data sets, it is almost impossible for R to complete the iteration with a realistic time period. Thus, we decided to develop our own implementation in C++. In this research, we used results from R to validate our results.

As written in the previous section, the EM algorithm is repeated until the the difference in the log-likelihood dL reaches the convergence limit l. The smaller the convergence limit, the more accurate the result at the expense of a longer time to converge. In addition to the convergence limit, the floating point representation also affects the computation speed. Using double precision format gives us more accurate results, but the computation takes longer compared to a float precision.

We randomly chose 10 x 10 voxels in the brain from all image data, and compared the computational speed, and converge for several conditions. For this test, the dependent variable is the FA value, and total number of the image datasets is 680 (n in the above section), and the number of subjects are 172 (m in the above section). Each subject was scanned between two to

six times (n_i in the above section) over a period of up to three years.

In the first experiment, we set two regressors for the population effects (X) and subject effects (U), which corresponds to typical settings for the time-series analysis (intercept and aging rate).

Tab. 1 shows the result. In this experiment, all conditions could converge within a maximum of 1000 iterations. As for the computational time, if we choose a double precision and $dl = 1E - 10$, the computational time took almost three times longer than the rest of three conditions.

X	U	dl	precision	computational time [s]	average iteration #	notes
2	2	1E-6	float	24.43	108.3	
2	2	1E-10	float	33.74	168.7	
2	2	1E-6	double	27.48	119.7	
2	2	1E-10	double	102.07	651.2	

Table 1: Computational time and the average iteration number per voxel

In the second experiment, we assigned seven regressors for population effects (intercept, rate, age at the first test, two clinical grouping (Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD)), and two regressors for subject effects (intercept and age). The FA value will be described as follows.

$$\begin{aligned} FA = \beta_{Intercept} + \beta_{days} * days + \beta_{Age} * Age \\ + \beta_{MCI} + \beta_{AD} + \beta_{MCI} * days + \beta_{AD} * days \\ + \gamma_{Intercept} + \gamma_{days} * days \end{aligned}$$

As shown in Tab. 2, when we used a float precision, the optimization could not converge even if we chose $1E - 6$ as a relatively low convergence criteria. When we chose a double precision, the iteration could converge, but the computation time increased almost three times between $dl = 1E - 6$ and $dl = 1E - 10$.

X	U	dl	precision	computational time [s]	average iteration #	notes
7	2	1E-6	float	311.69	> 1000	no convergence
7	2	1E-10	float	319.14	> 1000	no convergence
7	2	1E-6	double	98.05	123.3	
7	2	1E-10	double	234.21	669.8	

Table 2: Computational time and the average iteration number per voxel

Finally, we wanted to learn how much the choice of the convergence limit influences the precision of the results. In order to validate our algorithm, we estimated the regression coefficients by R and our code. Tab. 3 shows the regression coefficients β calculated by R and our code with two different convergence criteria, $dl = 1E-6$ and $dl = 1E-10$. Results calculated by our code with the two different conditions are in good agreement with the regression coefficients calculated by R.

	R	1E-6, double	1E-10, double
Intercept	5.117E-01	5.117E-01	5.117E-01
Days	-7.744E-05	-7.736E-05	-7.744E-05
Age	-1.382E-03	-1.382E-03	-1.382E-03
GRP1_MCI	-5.215E-02	-5.214E-02	-5.215E-02
GRP2_AD	-5.528E-02	-5.507E-02	-5.528E-02
Days:GRP1_MCI	8.762E-05	8.780E-05	8.762E-05
Days:GRP2_AD	1.192E-04	1.177E-04	1.192E-04

Table 3: Regression coefficients calculated by R and our code with two different convergence criteria. The regression coefficients at the left column were calculated by R to validate our code.

Tab. 4 shows p-values calculated by R and our code with different convergence criteria. Because there is no difference in the comparison between $dl = 1E-6$ and $dl = 1E-10$, we selected a double precision and $dl = 1E-6$ as the criteria for our later research.

The p-values calculated by our code show larger values than the results from R, meaning our results are more conservative. The main reason of this difference is that R and our code adopt different optimization methods. R uses the Newton-Raphson method, and ours adopt an iterative REML method as described in the previous section.

	R	1E-6, double	1E-10, double
(Intercept)	2.000E-16	1.644E-12	1.692E-12
Days	4.009E-03	1.003E-02	8.989E-03
Age	6.419E-02	6.822E-02	6.846E-02
GRP1_MCI	7.760E-04	1.047E-03	1.002E-03
GRP2_AD	4.717E-03	5.907E-03	5.571E-03
Days:GRP1_MCI	5.566E-03	9.608E-03	8.754E-03
Days:GRP2_AD	1.270E-02	2.307E-02	2.060E-02

Table 4: p-values calculated by R and our code with two different convergence criteria. The regressors at the right column were calculated by R to validate our code.

2.3.2 Computational Framework

As discussed above, the big challenge for this research is to shorten the computational time for the statical analysis. Although we have decided to use C++, which is much faster than interpreted software such as R and MATLAB, we still need to care about the total amount of

time. A 3D image of the brain of a subject consists of $200 * 256 * 132$ voxels, meaning there are 132 slices of 2D image data with $200 * 256$ pixels for one brain slice. We have 730 MRI volumes acquired in 177 subjects. If we analyze one slice, it takes almost one hour to complete the iteration. Based on this result, it will take about 130 hours (almost a week) to complete to analyze the whole data set.

In order to speed up computation, we parallelize computation on a cluster. Our lab has a cluster of thirty processors. We use the Slurm (Simple Linux Utility for Resource Management) cluster scheduler [20] to distribute processes on our cluster. Each process computes a single image slice, so we use 132 processes per analysis. Each command line has the following information.

- input file name: define a slice for the statistical analysis
- output file name: define a file for the estimated 2D map of regressors and z-scores
- inference method: select from the following options, MLarge, Satterthwaite, and KR
- regressor matrix for population effects
- regressor matrix for subject effects
- mask image: skip voxels that do not have DTI information

The scheduler assigns a job to an idle processor. Roughly speaking, we can expect to finish the analysis for all slices around 4 hours (132 hours / 30 processors = 4 hours). Of course, the computation time depends on the number of regressors, the precision format, and the amount of brain voxels per slice. The larger number of regressors we use, the longer time it will take to complete. Details of computation time will be discussed in Section 4.

3 Experiments

3.1 Sample

3.1.1 General Description

The MRI acquisitions were taken at multiple sites. The total number of the subjects included in this study is 176 (108 male, 68 female), and 730 MRI images had been taken so far along with clinical diagnoses to determine the clinical grouping of three conditions (healthy controls, MCI, or AD) The first image was taken on June-16th, 2010, and the last image was taken on February-6th, 2014.

3.1.2 Demographic Information

As discussed in Section 2, the reason we use the LMM for this analysis is that the MRI acquisitions are time-series data. Fig. 5 shows the histogram of the number of MRI acquisitions each subject. All subjects had between two and six MRI acquisitions, with an average of four.

There were no fixed dates for subjects to take MRI acquisitions. Fig. 6 shows the days from the first tests of each subject. The MRI acquisitions were done throughout almost three years.The shortest interval was twenty two days from the former test, and the longest was 693 days.

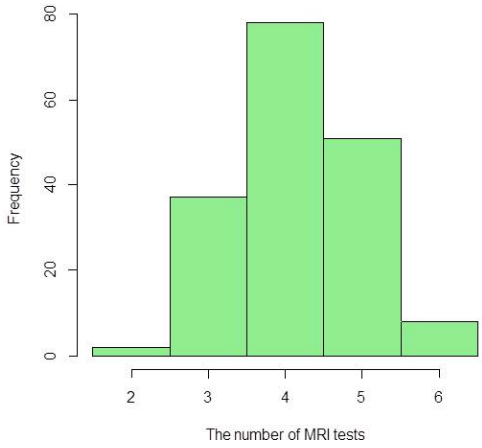

Figure 5: Histogram of the number of MRI acquisitions.

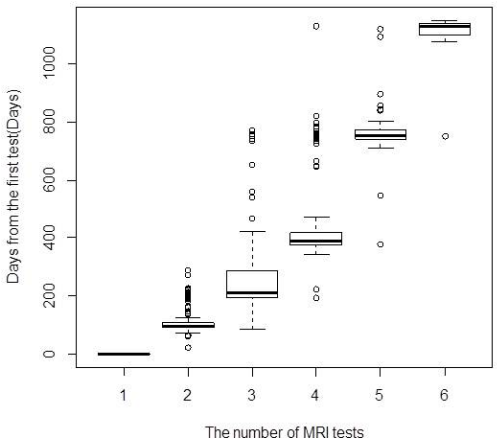

Figure 6: Days of each MRI acquisition from the first MRI acquisitions.

Subjects included in this study had an age between 48 and 90 years at the time of their first scan (average 73 years). Fig. 7 shows the histogram of the age when the subjects took the first MRI acquisition.

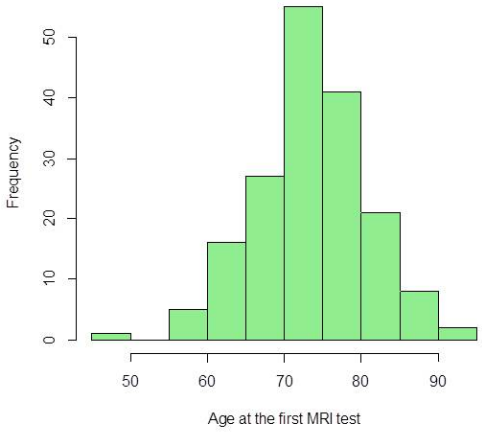

Figure 7: Histogram of Age.

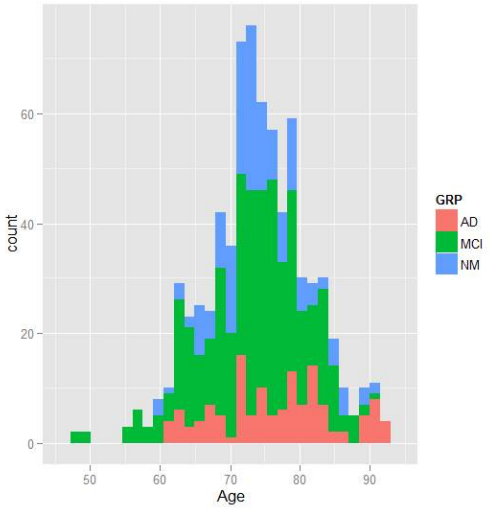

Figure 8: Histogram of Age of each MRI acquisition, color-coded according to clinical grouping.

The clinical grouping consists of three categories, including healthy controls (HC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). Grouping was assigned by physicians after testing subject's memory function. This sample includes 24% HC subjects, 56% MCI subjects, and 20 % AD subjects. Fig. 8 shows a subject's age at each MRI acquisition, and is color-coded according to the clinical grouping.

Alzheimer's Disease Assessment Scale (ADAS) is a clinical measure for the severity of AD. According to the ADNI General Procedure Manual, ADAS is described as "a brief cognitive test battery that assesses learning and memory, language production, language comprehension, constructional praxis, ideational praxis, and orientation" ([21]). Fig. 9 is a boxplot of the clinical grouping and ADAS.

Subjects have different brain volumes, so we use the peeled brain volume (PLV) for correction. Fig. 10 shows PLV for each subject.

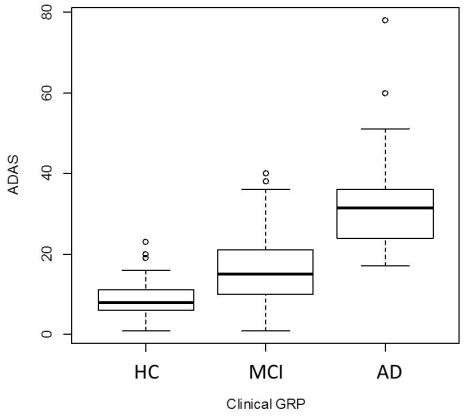

Figure 9: Boxplot of the clinical grouping and ADAS.

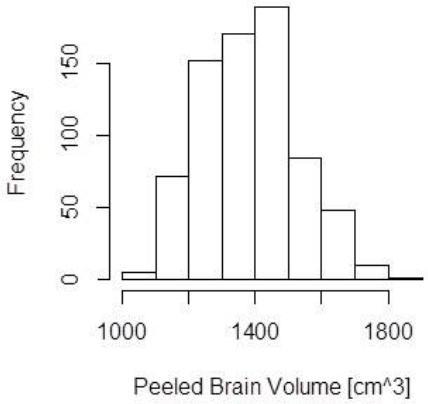

Figure 10: PLV for each subject.

3.1.3 MR Image Pre-Processing

People have different brain sizes. Even in one person, his/her brain size changes over time. One of our goals is to statistically identify which areas of the brain deteriorate when people develop AD. In order to analyze the brains of different sizes, we mapped all brain imaging data into a common space. The method is as follows:

T_1 weighted imaging data were analyzed using the following processing steps:

1. Scan data in DICOM format were converted into BRIAN format, aligned with the stereotactical coordinate system [25] and interpolated to an isotropical voxel size of 1mm using a fourth-order b-spline method.

2. All datasets were registered with the ICBM 2009c template [23] using a recent approach for nonlinear registration [27]. All registered head images were averaged, correcting for the mean intensity. The brain was extracted from the average to yield the brain template 1.

3. A mask of the intracranial volume was generated from each head dataset by a registration approach and used to extract the brain [24].

4. Data were corrected for intensity inhomogeneities using a newly developed technique that estimates the gain field by comparing the global intensity distribution with local ones. An intensity-corrected T_1-weighted image of the brain is obtained.

5. Data were segmented by a fuzzy approach using 3 classes [26], yielding a set of 3 probability images. Each voxel receives a probability for belonging to the intensity class 0: cerebrospinal fluid, CSF; 1: grey matter, GM, or 2: white matter, WM. As result, gross compartment volumes (in ml) for GM, WM, CSF, and peeled brain volume (PLV) were obtained for all subjects.

6. All intensity-corrected brain images were registered with template 1, and averaged to yield the brain template 2.

Diffusion-weighted imaging data were analyzed using the following processing steps:

1. Scan data in DICOM format were converted into BRIAN format, and image volumes corresponding to all gradient directions were registered with the gradient-free (T_2)-weighted image volume using affine registration and mutual information as an image similarity metric.

2. Diffusion tensors were computed from the registered diffusion-weighted images using a nonlinear procedure including anisotropic noise filtering [22]. Tensors were converted into fractional anistropy (FA), mean diffusivity (MD), and radial diffusivity (RD) values.

3. The T_2-weighted images were linearly registered with the (high-resolution) T_1-weighted obtained in step 4 above, and the resulting transformation was used to map FA, MD, and RD data into stereotaxic space. Next, the deformation field obtained in step 6 above was used to warp data into MNI space.

4. Data were smoothed using a Gaussian filter ($\sigma = 2$, FWHM of 4.7 mm), and volumes were re-grouped as stacks of slices of images.

3.2 Prototypical Example

3.2.1 Model

In order to demonstrate the usefulness of LMM, we now provide a typical use case, compare with results of the LM, and discuss results.

We are interested in analyzing how the white mater deteriorates as people age, thus a subject's age is one of the most important regressors. In addition, we are searching for the disease related effects, so the clinical grouping is included. As explained above, PLV should also be included because we use this for adjusting the brain volume of each subject. This matrix is shown in Fig. 11. The columns of MCI and AD show the clinical grouping, and 1 means the subject is AD (or MCI) at that MRI acquisition. If both of AD and MCI are 0, the subject was a HC.

Subject ID	Intercept	Age	MCI	AD	PLV
1	1	79.42	0	0	1453.68
1	1	79.95	0	0	1444.973
1	1	80.47	0	0	1477.432
1	1	81.39	0	0	1441.968
2	1	66.76	0	1	1546.645
2	1	67.17	0	1	1503.306
2	1	67.87	0	1	1426.136
2	1	67.38	0	1	1670.751
3	1	72.54	0	0	1277.517
3	1	72.86	0	0	1125.226

Figure 11: Part of Population effects matrix for a prototypical example. 1 in the column AD (or MCI) shows the subject was AD (or MCI) at that point. If both columns of AD and MCI are 0, the subject was a HC.

As we assume that people age differently, by including age in the subject effects matrix, we can model individual differences of their aging process.

Eq. 35 is the equation we used to evaluate the images. As described above, the values of MCI and AD are 0 or 1.

$$FA = \beta_{Intercept} + \beta_{Age} * Age + \beta_{MCI} * MCI + +\beta_{AD} * AD + \beta_{PLV} * PLV \\ +\gamma_{Intercept} + \gamma_{Age} * Age \tag{35}$$

3.2.2 Result

We obtained volumes of all regression coefficients and z-scores for FA, MD, and RD. We are interested in how aging affects the brain and which differences are found in MCI and AD patients with regard to healthy controls.

In this section, a hot iron colorbar is used to display the positive scale of the z-score, and a deep sea colorbar is used for the negative scale. Fig. 12 shows the scale of these color bars. Signs of z-scores match signs of β.

Figure 12: A hot iron colorbar for the positive scale (top) and a deep sea colorbar for the negative scale (bottom).

Here, we explain how to interpret the regression coefficients in Eq. 35. As mentioned in Section 1, when the direction that water molecules can move is limited by the axon, FA has a larger value. On the other hand, when the axon deteriorates, FA shows a smaller value because water molecules can freely move. At the same time, MD and RD have larger value because there is much space for the cellular water there.

β_{Age} corresponds to the change of FA (MD, RD) with age. Fig. 13 shows how to interpret the relationship between DTI scalars. Case 1 can be seen in the areas with the axon fibers which have the same direction (white matter). Case 2 can be seen in the areas with the axon fibers which have different directions and different deterioration speeds (cerebral cortex).

	FA	MD	RD	Interpretation
Case 1 Single axon	↓	↑	↑	When a single axon fiber deteriorates, the myelin sheath becomes damaged, so FA decreases. Less myelin means more space for extra-cellular water, so that MD and RD increase. (e.g. white matter)
Case 2 Two axons	↑	↑	↑	When one of two cable deteriorates, FA increases because the ratio of the size of eigen vectors increases. At the same time, the extra-cellular water increases, so MD and RD increases. (e.g. cerebral cortex)

Figure 13: Interpretation of the relationship between DTI scalars.

β_{MCI} and β_{AD} correspond to the difference of FA (MD, RD) between the MCI (AD) group and HC group. Positive means that the MCI (AD) group has larger FA (MD, RD) than HC group at that voxel.

Images and Interpretations: FA

Fig. 14 shows the significance of the age-related regressor expressed as a z-score. From the right panel, we can see that the corpus callosum (white arrow) significantly deteriorates as the result of aging process. In outer layer of the brain (gray matter), FA increases as people age. This effect can be explained as case 2 in Fig. 13.

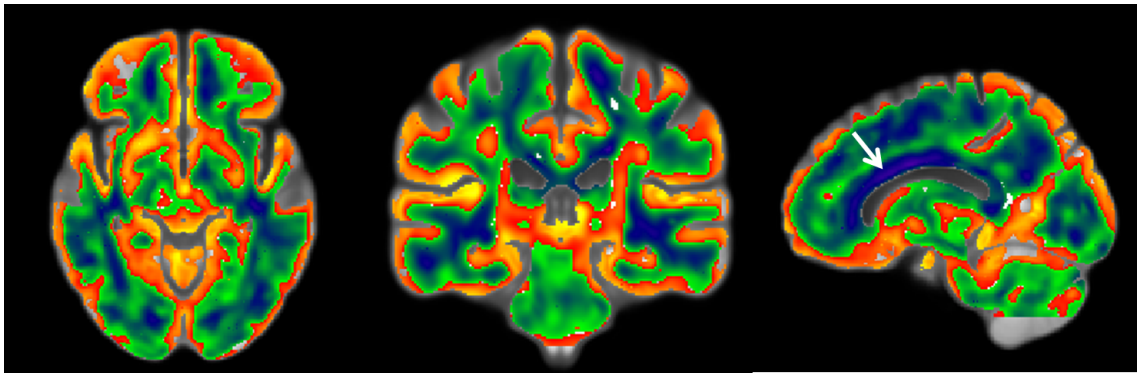

Figure 14: The significance of the age-related regressor expressed as a z-score (FA).

Fig. 15 shows the significance of the group-related regressor (MCI) expressed as a z-score. The left panel shows that both frontal white matter compartments (white arrows) have significantly smaller FA compared to healthy controls.

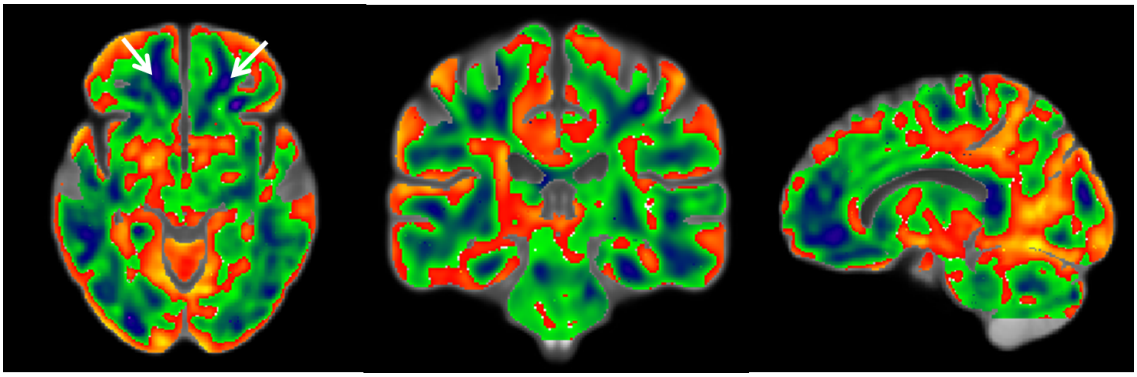

Figure 15: The significance of the group-related regressor (MCI) expressed as a z-score (FA).

Fig. 16 shows the the significance of the group-related regressor (AD) expressed as a z-score. The left panel shows that both frontal white matter compartments have significantly smaller FA compared to healthy controls (similar to Fig. 15). On the other hand, the striate (primary visual) cortex shows larger FA (white solid arrows) . From the center panel, we can see the internal capsule has significantly larger FA (white dash arrows). The right panels shows that the the corpus callosum has smaller FA.

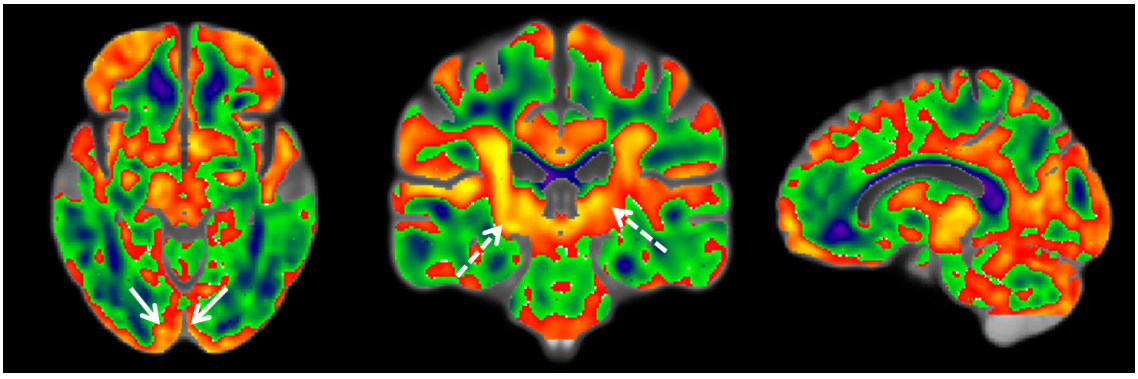

Figure 16: The significance of the group-related regressor (AD) expressed as a z-score (FA).

Images and Interpretations: MD

Fig. 17 shows the significance of the age-related regressor expressed as a z-score. Almost all areas show that MD increases as people age. Small orbitofrontal (white solid arrow in the

right panel) and occipital basal (white dash arrow) regions show a slight decrease of MD with age.

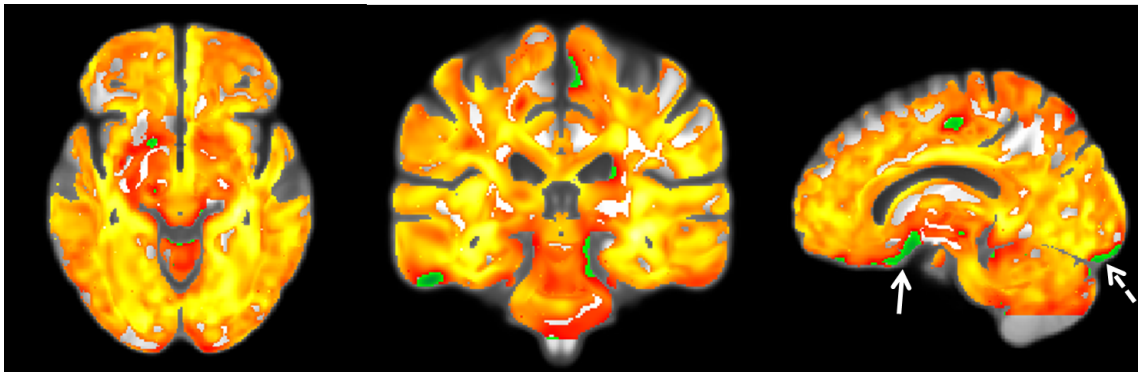

Figure 17: The significance of the age-related regressor expressed as a z-score (MD).

Fig. 18 shows the significance of the group-related regressor (MCI) expressed as a z-score. The center panel shows that the centrum semiovale (white solid arrows) has significantly larger MD. The right panel shows positive z-scores in the vicinity of the splenium (white dash arrow).

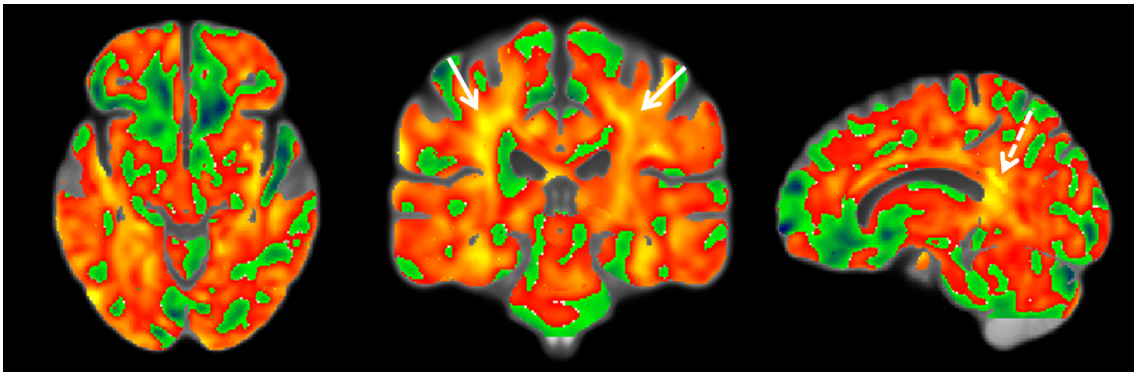

Figure 18: The significance of the group-related regressor (MCI) expressed as a z-score (MD).

Fig. 19 shows the significance of the group-related regressor (AD) expressed as a z-score. Almost all white matter areas have larger MD, and gray matter areas have smaller MD compared to healthy controls. Especially, from the center panel, we can see the high z-scores around the temporal lobe (white solid arrows), and also the right panel shows that the front lobe (white dash arrow) has high z-scores. Some cortical regions show lower MD although they barely reach statistical significance.

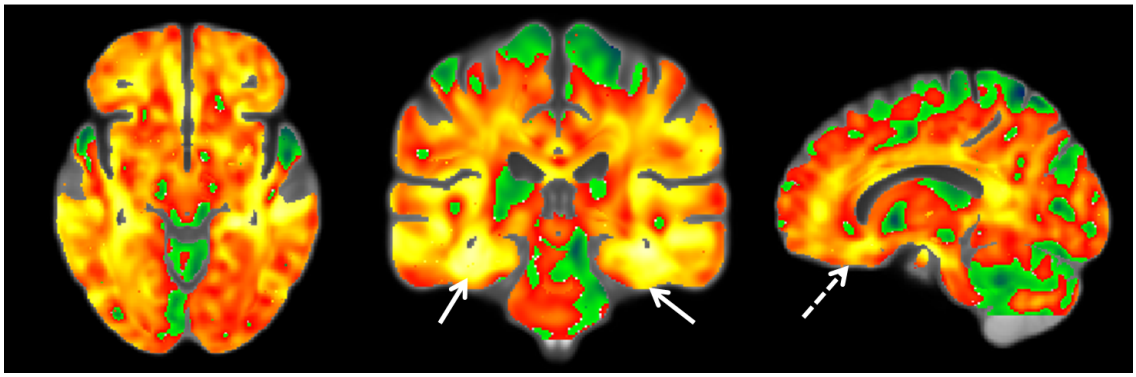

Figure 19: The significance of the group-related regressor (AD) expressed as a z-score (MD).

Images and Interpretations: RD

Fig. 20 shows the significance of the age-related regressor expressed as a z-score. Almost all areas show z-scores similar in magnitude and sign compared to the results for MD (Fig. 17).

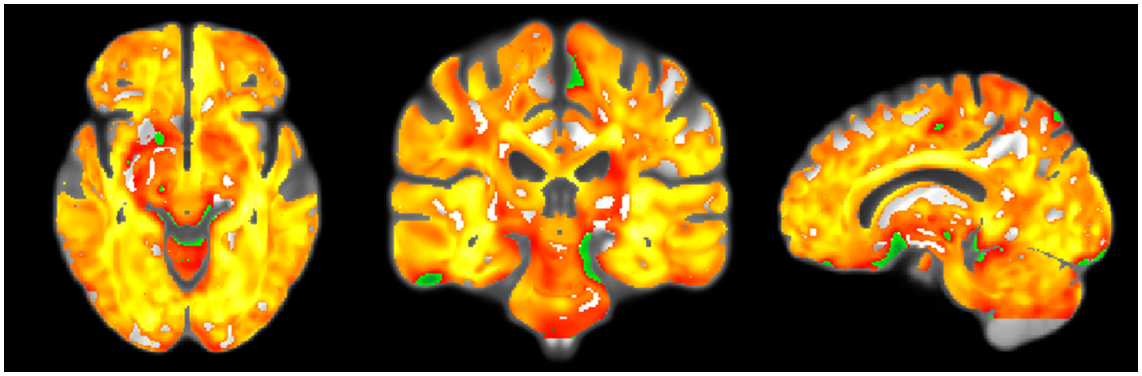

Figure 20: The significance of the age-related regressor expressed as a z-score (RD).

Fig. 21 shows the significance of the group-related regressor (MCI) expressed as a z-score. Almost all areas show z-scores similar in magnitude and sign compared to the results for MD (Fig. 18).

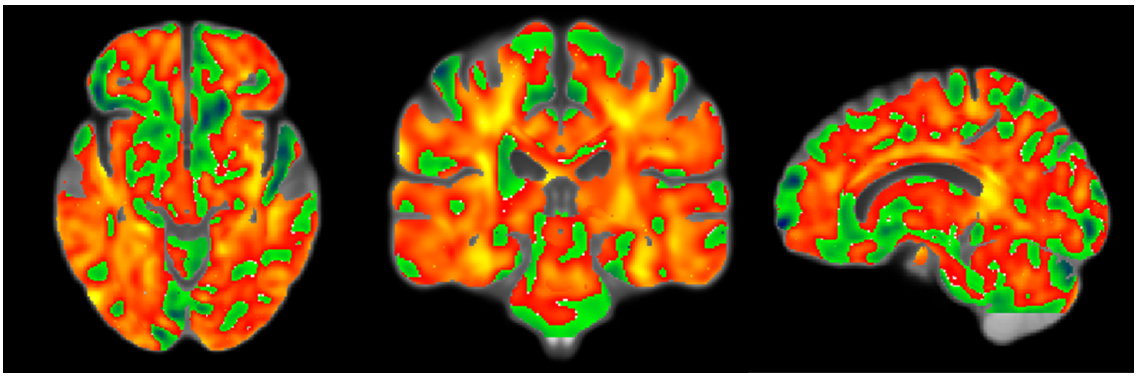

Figure 21: The significance of the group-related regressor (MCI) expressed as a z-score (RD).

Fig. 22 shows the significance of the group-related regressor (AD) expressed as a z-score. White matter areas show positive z-scores, and the areas around the temporal lobe (white arrows in the center panel) show high significance ($z > +1.96$).

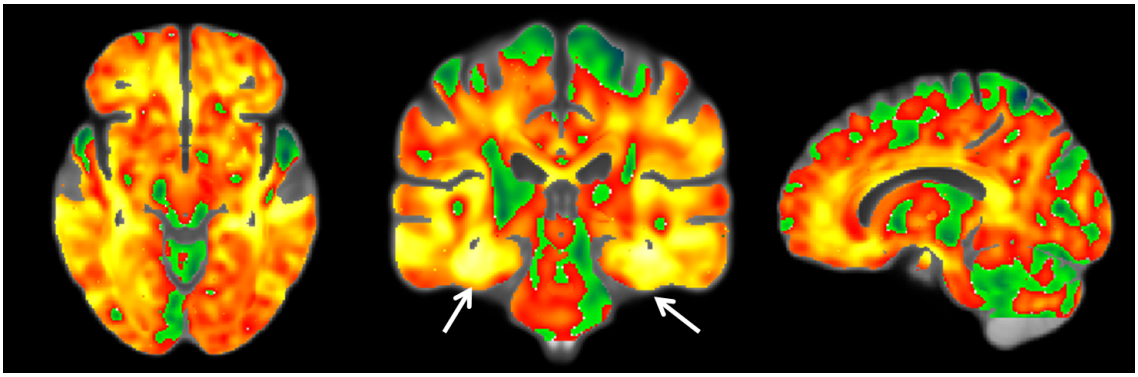

Figure 22: The significance of the group-related regressor (AD) expressed as a z-score (RD).

3.3 Subject-wise information

3.3.1 FA and Aging: A Closer Look

A main advantage of using LMM in this context is that subject-level results are obtained in the form of regression coefficients γ. We now try to explore how this information can be used. We consider FA as the quantity of interest, and focus on arbitrary selected single voxel at x=80, y=87, z=100. A plot of FA vs. age shows a considerable variability (Fig. 23).

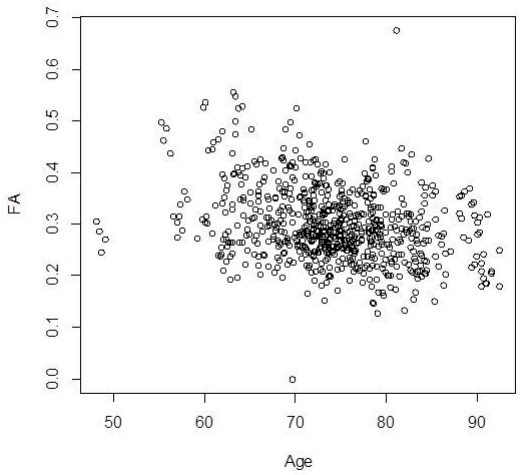

Figure 23: FA vs. age at (x,y,z)=(80,87,100) for all examinations

We computed the following LMM:

$$FA = \beta_{Intercept} + \beta_{Age} * Age + \gamma_{Intercept} + \gamma_{Age} * Age, \tag{36}$$

that has population and subject wise regressors for age-related effects.

Fig. 24 is the relationship between γ_{Age} and FA for each subject. Each point is color-coded according to the clinical grouping. In order to investigate whether these γ_{Age} have differences between the clinical groupings, we tested this result with the following linear model: $lm(\gamma_{Age}{}^{\sim}GRP)$.

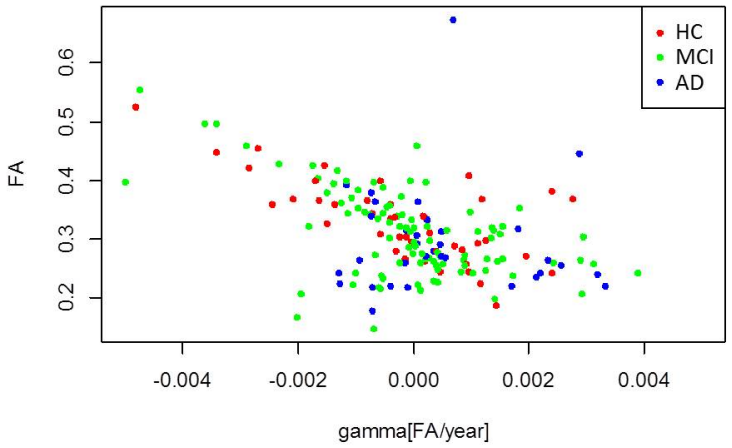

Figure 24: Relationship between FA and γ_{Age} : red, green and blue show HC, MCI, and AD subjects respectively. Note that this arbitrary selected location (80,87,100) is not expressed to show disease-related differences.

Tab. 5 shows that AD subjects have significantly larger γ_{Age} than healthy controls, and the average value of γ_{Age} for AD subjects is positive. This means that the average deterioration speed of AD subjects is slower than those of healthy controls.

Testing for group-related differences of FA with the following linear model: $(lm(FA \tilde{} GRP))$, we found that AD subjects have lower FA values, compared to both other groups (Tab. 6).

	Averaged γ_{Age}	p value
HC	-2.120E-04	0.351
MCI	-7.470E-05	0.614
AD	4.695E-04	0.046

Table 5: Averaged γ_{Age} for each clinical group.

	Averaged FA	p value
HC	0.331	$< 2e-16$
MCI	0.309	0.118
AD	0.289	0.017

Table 6: Averaged FA for each clinical group.

We found the following two facts on this voxel: (1) the brain of AD subjects deteriorates slower, and (2) FA value of AD subjects is smaller than healthy controls. Based on these two facts, we hypothesized a prototypical aging curve as shown in Fig. 25: initially, brains deteriorate at a certain rate. Later, this rate slows down.

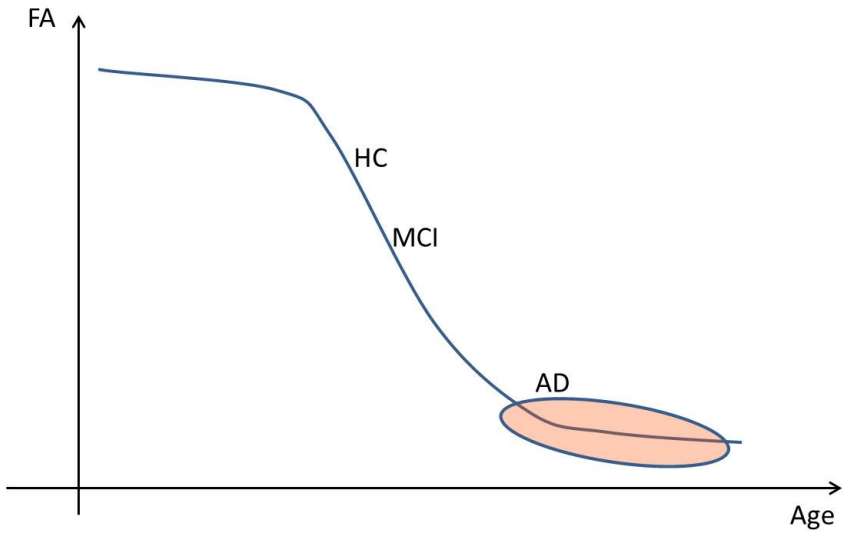

Figure 25: Prototypical aging curve.

3.3.2 Individual Aging Effects

In order to obtain clues for a prototypical aging curve, we now separate age related effects from disease related effects, and focus on healthy controls. We assume that some of the variance in Fig. 23 is due to an individually different onset in the aging process, called ΔAge. As shown in Fig. 26, ΔFA is defined as the difference between FA for each subject and the blue line at the average age. And then ΔFA for each subject is converted to ΔAge. The blue line shows the population effects of this group.

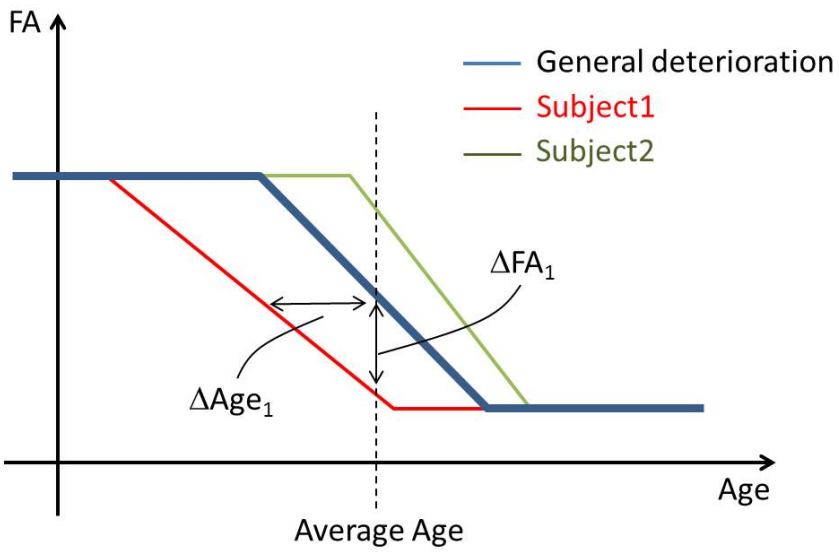

Figure 26: Each subject has an individual onset and speed of deterioration. ΔFA for each subject is converted to ΔAge.

$$\begin{aligned}\Delta FA_{subject_i} &= FA_{subject_i} - FA_{General\ deterioration} \\ &= X\beta + U\gamma - X\beta \\ &= U\gamma \quad \text{note: Age=average age of subjects} \\ \Delta Age_{subject_i} &= \frac{\Delta FA_{subject_i}}{\beta_{Age} + \gamma_{Age_i}}\end{aligned}$$

We tested the following two sets of population and subject effects matrices to calculate ΔAge, and computed it for the whole brain.

Model A: X=Intercept, Age, PLV and U=Intercept
Model B: X=Intercept, Age, PLV and U=Intercept, Age

Because we are interested in the voxels which show significance about their deterioration, ΔAge for each subject is calculated using only for voxels with $z < -1.96$. ΔAge is derived for all voxels for each subject.

First, we are interested to test whether ΔAge is random or may carry neurobiological information. If our hypothesis is true, ΔAge calculated for the whole brain of one subject should ideally have an unimodal distribution. If there are two peaks, this means that two areas deteriorate at different rates, or have different onsets. If there is no peak, no subject related onset can be derived from this model. Fig. 27 is the result of both model A and B for all healthy

controls, which shows single peak for each subject. This result suggests that we can use ΔAge as a subject related variable.

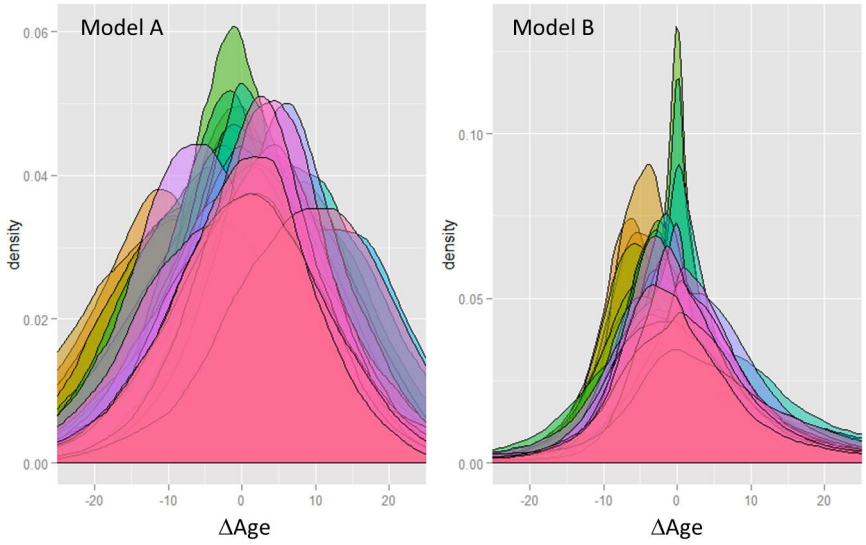

Figure 27: Density curve of ΔAge for each subject. Different colors show different subjects. The total number of healthy controls is 43.

In order to decide which model better describes the data, we computed the standard deviation of ΔAge for each subject for both models. Fig. 28 shows ΔAge and standard deviation for each subject for both models. The paired T-test showed that model B has significantly smaller standard deviation. Therefore, we adopted the model B for the next study.

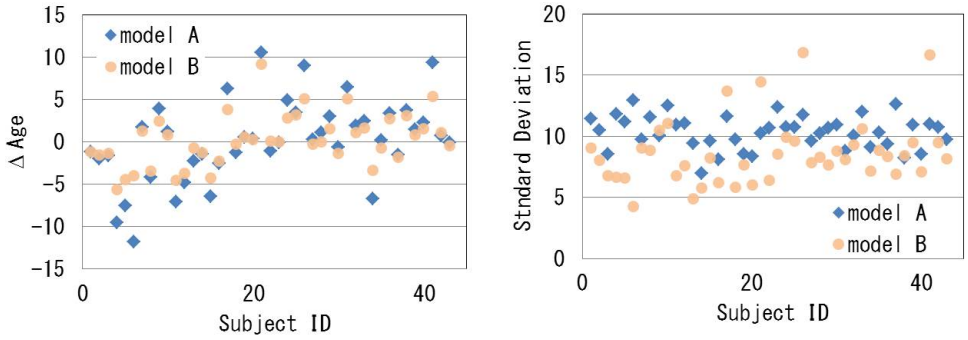

Figure 28: ΔAge (left) and standard deviation (right) for each subject.

Disease Onset: Prototypical deterioration

As we defined ΔAge as the difference from the prototypical deterioration curve shown in Fig.

25, we can use ΔAge to derive the prototypical deterioration curve by subtracting ΔAge from the age of each measurement. Fig. 29 shows the result of the relationship between FA (MD, RD) and the corrected age using ΔAge. The corrected relationship shows less variance.

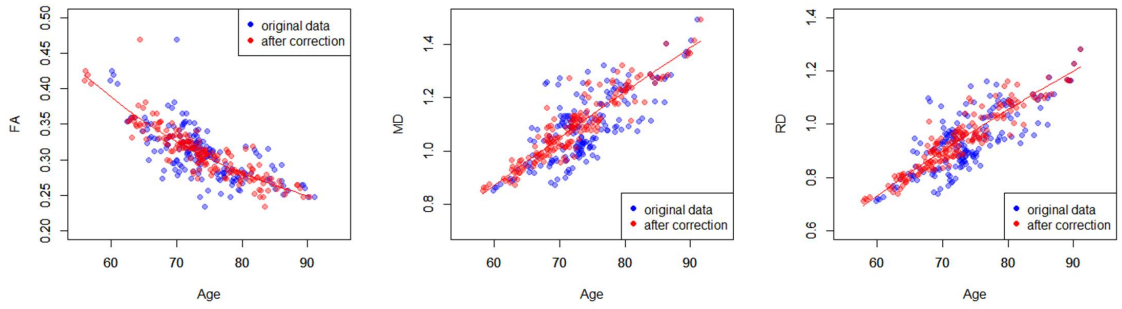

Figure 29: The relationship between averaged FA (left), MD (center), and RD (right) vs. age. Blue dots show original measurement, and red dots show corrected data with ΔAge. The red line is the fitting result.

In order to numerically describe this result, we fitted FA (MD, RD) vs. age by the model in Eq. 37. The result shows that the deterioration speed of FA becomes slower as people age, which matches our model described in Fig. 25.

$$DTIscalar = a + b * exp(c * Age) \tag{37}$$

	FA	MD	RD
a	0.167	5.896	3.088
b	1.593	-6.249	-3.670
c	-0.033	-0.0036	-0.0074

Table 7: Fitting coefficients for the corrected DTI scalars.

ΔAge is independently estimated for FA, MD, and RD. If ΔAge corresponds to an individual aging onset, ΔAge estimated from FA, MD, and RD should show similar results. Fig. 30 shows the relationship of ΔAge estimated by FA and MD (left panel), and FA and RD (right panel). Tab. 8 shows the fitting results of $lm(\Delta Age_{MD\ (RD)} \tilde{}\ \Delta Age_{FA})$. The correlation matrix of ΔAge estimated by each DTI scalar is shown in Tab. 9. In summary, the high correlation confirms the hypothesis that ΔAge has individual biological significance.

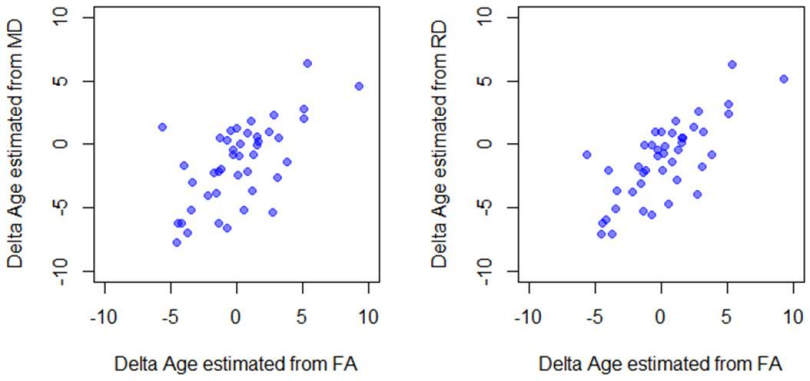

Figure 30: Relationship of ΔAge between FA and MD (left), and FA and RD (right).

	MD and FA		RD and FA	
	β	p-value	β	p-value
intercept	-1.5464	2.55E-4	-1.3745	9.15e-05
slope	0.6711	4.53e-06	0.7477	9.45e-09

Table 8: Linear fitting coefficients for MD and FA (left), RD and FA (right)

	FA	MD	RD
FA	1	0.636	0.746
MD	0.636	1	0.985
RD	0.746	0.985	1

Table 9: Correlation matrix for ΔAge.

Here, we examined how subject wise information can be used. We modeled a prototypical aging process, and found ΔAge carries individual biological information. In addition, we derived a prototypical age-dependency of FA, MD and RD in healthy subjects using ΔAge.

4 Discussion: Comparison of LM and LMM

Computation Speed

As described in Section 2, the estimation of the regression coefficients in the LMM is an iterative process, while in the LM, a matrix-vector product is sufficient for defining the regression coefficients. Thus, the computation time for LMM is always longer than LM. In order to compare the computation speed, we calculated LMM and LM for different conditions. The calculation was done for one slice using the sample group described in Section 3. Tab. 10 shows the result. Because data load and other operations take longer than the iteration process, computation times for LM are almost independent of the size of the regression matrix. However, the iteration speed of LMM strongly depends on the number of regressors in U (subject effects matrix). When only one regressor is used in U, it takes two minutes regardless of the number of regressors in X (population effects matrix), but if two regressors are used in U, the iteration time is 20 times longer than with one regressor in U.

LM			LMM		
X Population Effects	U Subject Effects	Time(min)	X Population Effects	U Subject Effects	Time(min)
Intercept Age	-	<1	Intercept Age	Intercept	2
Intercept Age ADAS	-	<1	Intercept Age ADAS	Intercept	2
Intercept Age Days ADAS	-	<1	Intercept Age Days ADAS	Intercept	2
Intercept Age	-	<1	Intercept Age	Intercept Age	50
Intercept Age ADAS	-	<1	Intercept Age ADAS	Intercept ADAS	56
Intercept Age ADAS	-	<1	Intercept Age ADAS	Intercept Age	86

Table 10: Computation speed for different conditions: This test was calculated for slice number=100, and DTI scalar=FA.

Regression Coefficient

Both LM and LMM calculate regression coefficients and z-scores for population effects. Here we will compare the results from both LM and LMM.

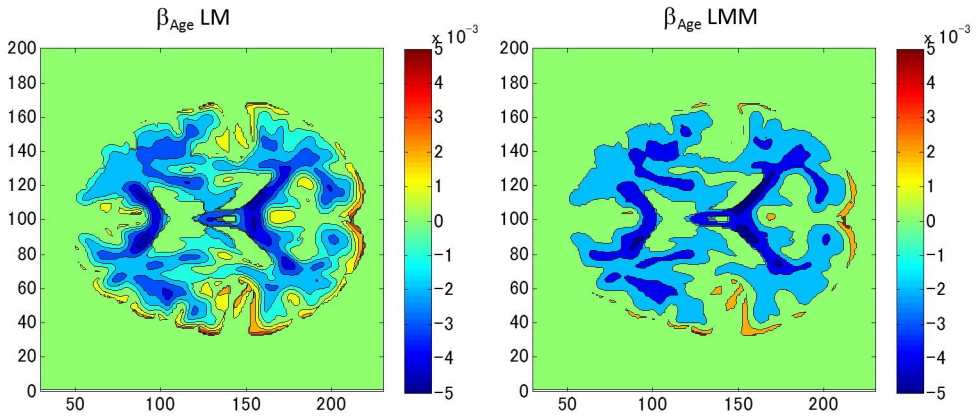

Figure 31: 2D maps of β_{Age} for LM (left) and LMM (right).

Fig. 31 shows the 2D maps of β_{Age} for one slice calculated by both models (LM is at the left side, and LMM is at the right side). Intercept and age are used in both X and U. β_{Age} estimated by LMM has less spatial details. Although voxel-wise computations are independent, the LMM map appears smoother than the LM map. Due to the additional flexibility of the LMM, some of the variance in the data may be found in the subject-level coefficients now.

In order to compare both results numerically, we calculated the ratio as obtained from both models. Fig. 32 shows the ratio of $\beta_{Age_{LMM}}$ and$\beta_{Age_{LM}}$. More than 84% of the voxels match each other within $\pm$ 20% accuracy.

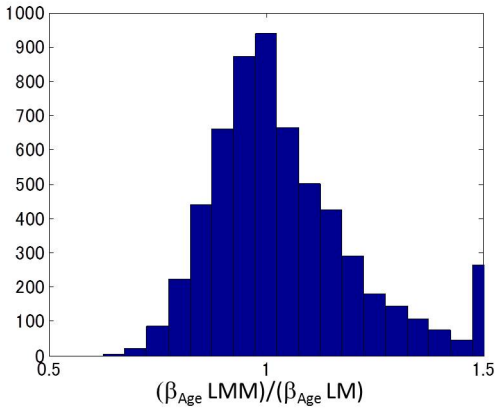

Figure 32: Histogram of $\frac{\beta_{Age_{LMM}}}{\beta_{Age_{LM}}}$. The voxels with $z < -1.96$ are used for this histogram.

This result is very useful when we focus on population effects. As we showed in the previous section, using LM is always faster than LMM, so we can use LM for exploration phases. For example we can test several models with LM, and then we can develop refined models using LMM.

Inference

As discussed the previous section, statistically, LMM can handle with time-series measurement correctly. Here, we compare inferences between both models.

Fig. 33 shows the results for maps of the z-score estimated by LM and LMM. As described in Section 2, we implemented three options to estimate the z-score in our code of LMM, which are based on different approaches to compute the degree of freedom. However, the Satterthwaite method was unstable during the optimization, because the computation of a numerical derivative is needed.

In LM, each MRI measurement is considered as an independent test, so the degree of freedom is the number of tests minus the number of regressors, while the maximum degree of freedom for LMM is the number of subjects minus the number of regressors (option name is MLarge). If we adopt Kenward-Roger approximation, we have to accept reduction of the degrees of freedom. Result shows that LM overestimates z-scores, but rough trends are similar to the results estimated by LMM. Therefore, we can use LM for our exploratory tests as we discussed in the previous section.

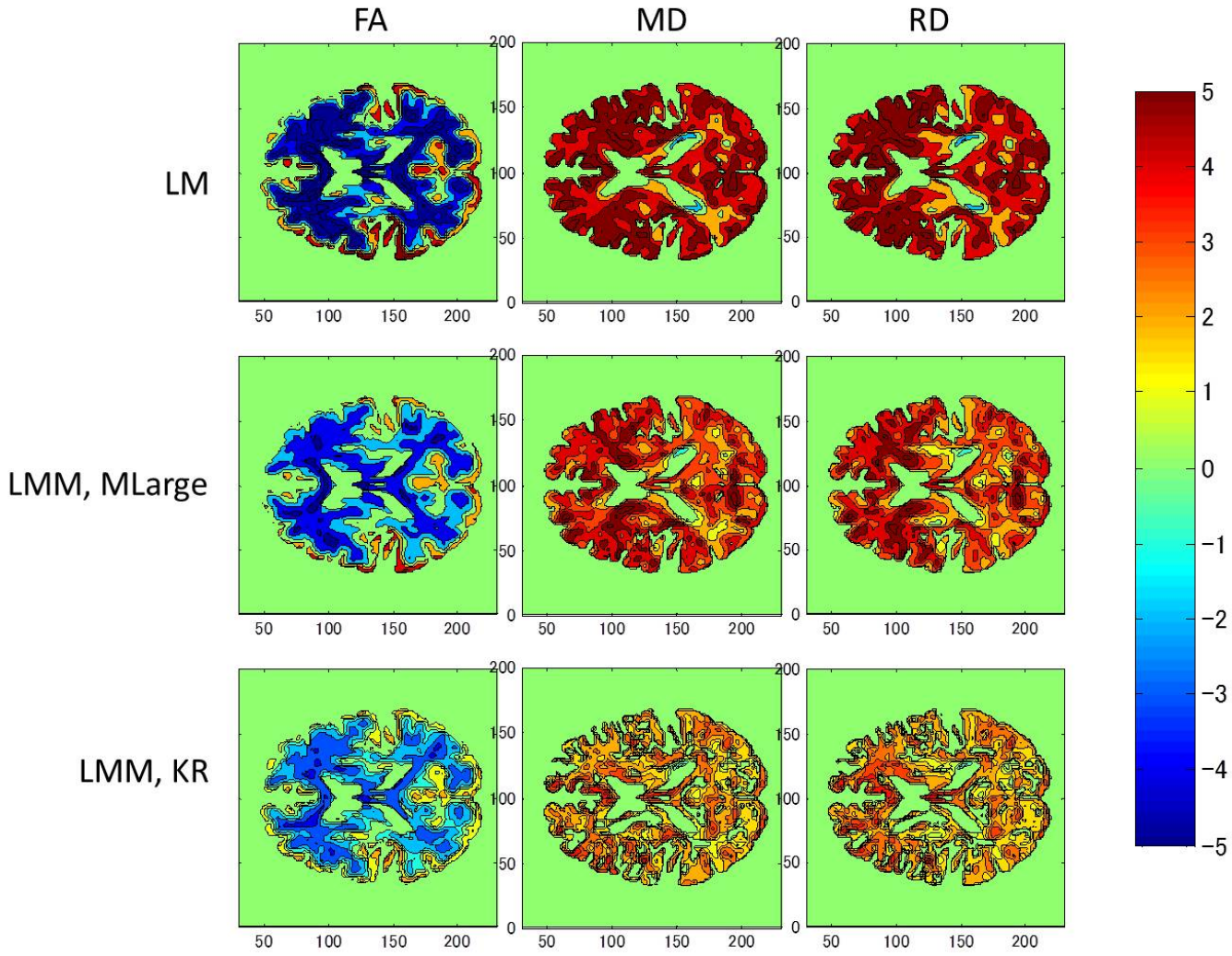

Figure 33: z-score maps. From top to bottom, LM, LMM with MLarge, and LMM with KR method are shown respectively. From right to left, FA, MD, and RD are shown.

5 Summary and Future Work

We have developed a C++ code for the LMM in order to analyze a large amount of MRI data, which cannot be handled well by using interpreted languages because of the slow computation speed. Our code was successful in reducing a considerable amount of computation time, and can analyze the whole dataset within one day. Through our experiment, we found that the population effects coefficients of the LMM are similar to the coefficients estimated by the LM, and that the LMM has lower inferences due to the statistically better modeling of time-series data. We have also implemented a spatial correlation module as an optional post-process procedure.

We analyzed 730 images taken from 176 subjects as prototypical examples. The statistical analysis with a large number of MRI images was first shown in this study. We interpreted our results and showed the interpretation of the relationship between DTI scalars.

The LMM made it possible for us to access subject wise information. We established that ΔAge can be used as the onset of deterioration for each subject. Based on this finding, we could model the general deterioration process.

In Section 2, we assumed that each measurement in one subject is independent, meaning $\mathbf{\Sigma}_i = \sigma^2 \mathbf{I}_{n_i}$. However, because measurements taken at short intervals are similar, we should model the temporal correlation in $\mathbf{\Sigma}_i$ as an autoregressive model. Here, off-diagonal elements will have non-zero values which corresponds to the correlation between serial examinations.

Considering usage of this tool for medical findings, we will try to find who will convert from MCI to AD. In Section 4, we focused on healthy controls to find the general deterioration model. Our goal is to find subjects who will develop AD at an earlier stage. In our current dataset, diagnostic labels are not revised. We can use the ADAS score instead. Of course, at the current stage, the clinical grouping is very useful because it tells us the clear category that each subject belongs to. We may be able to improve our understanding led by the clinical grouping using ADAS.

This time, we demonstrate that ΔAge is one of the subject related variable for each subject, but other subject-level effect might reveal other neurobiological characteristics.

References

[1] Alzheimer's Association
Alzheimer's Facts and Figures.
https://www.alz.org/alzheimers_disease_facts_and_figures.asp
(accessed: November 1, 2014).

[2] Hebert L., Weuve J., Scherr P., Evans D. (2013)
Alzheimer disease in the United States (2010−2050) estimated using the 2010 census
Neurology. 80, 1778-1783.

[3] Lopez O. (2011)
The Growing Burden of Alzheimer's Disease
Published Online: The American Journal of Managed Care
http://www.ajmc.com/publications/supplement/2011/a300_11nov/A300_11nov_Alzheimers_Lopez_S339
(accessed: November 1, 2014).

[4] The Alzheimer's Association
http://www.alz.org/research/science/alzheimers_disease_treatments.asp#approved
(accessed: November 1, 2014).

[5] McKhann G., Drachman D., Folstein M., Katzman R., Price D., Stadlan E.M. (1984)
Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease
Neurology. 34, 939-944.

[6] Wikipedia
http://en.wikipedia.org/wiki/Magnetic_resonance_imaging
(accessed: November 1, 2014).

[7] Ramachandran S. (2014)
Alzheimer Disease Imaging
Published Online: MedScape
http://emedicine.medscape.com/article/336281-overview
(accessed: November 1, 2014).

[8] Alzheimer's association
Brain Tour.
http://www.alz.org/braintour/healthy_vs_alzheimers.asp
(accessed: November 1, 2014).

[9] Clifford J., Knopman D., Jagust W., Shaw L., Aisen P., Weiner M., Peterson R., Trojanowski J. (2010)
Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade
The Lancet Neurology. 9, 119-128.

[10] Solodkin A., Chen E., Hoesen G., Heimer L., Shereen A., Kruggel F., Mastrianni J. (2013)
In Vivo Parahippocampal White Matter Pathology as a Biomarker of Disease Progression to Alzheimer's Disease
The Journal of Comparative Neurology — Research in Systems Neuroscience. 521, 4300-4317.

[11] Alexander A., Lee J., Lazar M., Field A. (2007)
Diffusion Tensor Imaging of the Brain
Neurotherapeutics. 4, 316–329.

[12] Hommer D. (2004)
Male and Female Sensitivity to AlcoholInduced Brain Damage
published online: http://pubs.niaaa.nih.gov/publications/arh27-2/181-185.htm
(accessed: November 1, 2014).

[13] Czado C. (2013)
Lecture 10: Linear Mixed Models.
https://www-m4.ma.tum.de/personen/professorinnen/claudia-czado/forschung/lecture-slides/ (accessed: November 1, 2014).

[14] Laird N., Lange N., Stram D. (1987)
Maximum Likelihood Computations with Repeated Measures: Application of the EM Algorithm.
Journal of the American Statistical Association. 82, 97–105.

[15] Lindstrom M.J., Bates D.M. (1988)
Newton-Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data.
Journal of the American Statistical Association. 83, 1014-1022

[16] Kenward M.G., Roger J.H. (1997)
Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood.
Biometrics. 53, 983–997.

[17] Halekoh U., Hojsgaard S. (2012)
A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Modelsthe R Package pbkrtest.
Submitted to: Journal of Statistical Software.

[18] Frinston K.J., Worsley R.S.J., Frackowiak J.C., Mazziotta J.C., Evans A.C. (1994)
Assessing the Significance of Focal Activations Using Their Spatial Extent
Human Brain Mapping. 1, 210-220

[19] Benjamini Y., Hochberg Y. (1995)
Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing
Journal of the Royal Statistical Society. Series B. 57, 289-300

[20] Lawrence Livermore National Laboratory
https://computing.llnl.gov/linux/slurm/
(accessed: November 1, 2014).

[21] ADNI General Procedure Manual
http://www.adni-info.org/Scientists/Pdfs/ADNI_GeneralProceduresManual.pdf.
(accessed: November 1, 2014).

[22] Fillard P., Pennec X., Arsigny V., Ayache N. (2007)
Clinical DT-MRI Estimation, Smoothing, and Fiber Tracking With Log-Euclidean Metrics.
IEEE TRANSACTIONS ON MEDICAL IMAGING. 26, 14761–1482.

[23] Fonov V.S., Evans A.C., McKinstry R.C., Almli C.R., Collins D.L. (2009)
Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. NeuroImage. 47, S102.

[24] Hentschel S., Kruggel F. (2004)
Determination of the intracranial volume: a registration approach.
International Workshop on Medical Imaging and Augmented Reality (MIAR 2004), Lecture Notes in Computer Science. 3150, 253-260.

[25] Kruggel F., von Cramon D.Y. (1999)
Alignment of magnetic-resonance brain datasets with the stereotactical coordinate system.
Medical Image Analysis. 3, 1-11.

[26] Pham D.L., Prince J.L. (1999)
An adaptive fuzzy segmentation algorithm of magnetic resonance images.
IEEE TRANSACTIONS ON MEDICAL IMAGING. 18, 737–752.

[27] Vercauteren T., Pennec X., Perchant A., Ayache N. (2009)
Diffeomorphic demons: Efficient non-parametric image registration.
NeuroImage. 49, 61-72.