MONITORING THE PROGRESSION OF ALZHEIMER'S DISEASE WITH LATENT TRANSITION MODELS

by

JIENA GU

B.Eng., South China Normal University, 2013

A REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:

Major Professor
Wei-Wen Hsu

Copyright

Abstract

BACKGROUND AND PURPOSE: Alzheimer's disease is currently a neurodegenerative diseases without any effective treatments to slow or reverse the progression. To develop any potential treatments, the need of a good statistical model to assess the progression of Alzheimer's disease is becoming increasingly urgent. This study proposed a latent transition model to monitor the progression of Alzheimer's disease which can help the development of a given proposed treatment.

METHOD: A latent transition model was used to assess the progression of Alzheimer's disease. The volume of Hippocampus and fluorodeoxyglucose-PET (FDG) were employed as biomarkers in this model. These two biomarkers are very sensitive to the pathological signs of the Alzheimer's disease. The proposed latent transition model was performed with real data from Alzheimer's Disease Neuroimaging Initiative (ADNI), which contain 2,126 participants from 2005 to 2014.

RESULTS/FINDINGS: The latent transition model suggested six states of disease progression and two different pathological profiles. One progression profile was mainly determined by the biomarker of FDG and the other by the volume of Hippocampus.

CONCLUSION: The results revealed the existence of various progression profiles of Alzheimers disease, suggesting a new way to evaluate the disease progression.

Table of Contents

List of Figures vi

List of Tables vii

Acknowledgements vii

1 Introduction 1

2 Statistical methods 4

2.1 Latent Transition Models 4

2.2 Statistical Model of Latent Transition Analysis 5

2.3 Parameters Estimation with EM Algorithm 6

2.4 SAS package 7

3 Real data application 8

3.1 ADNI data 8

3.2 Biomarkers Selection 9

4 Results: Latent transition models for ADNI data 13

4.1 Model selection 13

4.2 Latent class and transition probabilities 14

4.3 Two profiles 18

5 Discussion 24

5.1 Two profiles 24

5.2 Presymptomatic biomarkers 25

5.3 Limitation and future study 25

Bibliography 27

A SAS code 32

List of Figures

3.1 Spaghetti plots of selected biomarkers 11

3.2 The distributions of selected biomarkers by clinical results 12

4.1 Transition probabilities 15

4.2 The distributions of FDG and Hippocampus of Normal participants in baseline year and two years later 16

4.3 The distributions of FDG and Hippocampus of H1 participants in baseline year and two years later 16

4.4 The distributions of FDG and Hippocampus of H2 participants in baseline year and two years later 16

4.5 The distributions of FDG and Hippocampus of F1 participants in baseline year and two years later 17

4.6 The distributions of FDG and Hippocampus of F2 participants in baseline year and two years later 17

4.7 The distributions of FDG and Hippocampus of AD participants in baseline year and two years later 17

4.8 Distributions of FDG and Hippocampus by unobserved states at baseline time 19

4.9 Distributions of two profiles based on our model 20

4.10 The distributions of FDG and Hippocampus of clinically normal participants 21

4.11 The distributions of FDG and Hippocampus of MCI participants 22

4.12 The distributions of FDG and Hippocampus of AD participants 23

List of Tables

3.1 Summary Table of ADNI database 9

4.1 Model Selection 13

Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Prof. Wei-Wen Hsu for the support of my MS study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this report. I could not have imagined having a better advisor and mentor for my MS study.

Besides my advisor, I would like to thank the rest of my MS report committee: Prof. James Neill , Prof. Weixing Song, and Prof. Haiyan Wang, for their encouragement, insightful comments, and hard questions.

My sincere thanks also goes to Kristin Oberheide, for offering me the part-time student worker opportunities in her office and leading me working on diverse exciting projects.

Last but not the least, I would like to thank my family: my parents Zhaoyi Gu and Hong Li, for giving birth to me at the first place and supporting me spiritually throughout my life. Besides my family, I would like to thank my friend: Alex McLellan, Maegan McLellan, Betty McLellan, Rudy McLellan, Yeng Xiong, Sharif Mahmood and Richard Opoku-Nsiah.

Chapter 1

Introduction

Alzheimer's Disease (AD) is one type of age-related, neurodegenerative disorder which slowly destroys memory, brain function, and, eventually, leads to loss of independence and total disability (Batsch et al., 2012). In the United States, the prevalence of AD is predicted to be over 14 million by 2050 (Alzheimer's, 2012). AD is also one of the costliest chronic diseases to American society. In 2016, the direct costs to Americans in caring for AD patients were estimated to be 236 billion dollars, with half of the costs borne by Medicare (Alzheimer's, 2016). However, there is no effective treatment so far to slow or reverse the progression of AD.

To develop any potential treatments, the need of an accurate statistical model to assess the progression of AD is becoming increasingly urgent. Modeling patterns of AD progression is an essential component, because it enables advanced understanding of the pathology of AD and helps the development of potential treatments. Recently, this topic is of high interest to biomedical researchers (for example, see Holford, 2014; Johnson et al., 2014; Li et al., 2015; Schmidt-Richberg et al., 2015). Many of these studies often used three clinically diagnostic stages: Normal (NC or pre-symptomatic stage), mild cognitive impairment (MCI or prodromal stage) and AD stage. In NC or pre-symptomatic stage, individuals are clinically normal but may have AD pathology (Johnson et al., 2014; Schmidt-Richberg et al., 2015). The presymptomatic period (NC stage) refers to a pathophysiological process that

is progressing towards developing cognitively and behavioural impairment of AD (Li et al., 2015). In MCI or prodromal stage, patient's cognition is impaired but the daily functions may not be undermined. The AD clinical symptoms such as losing functions in activities appear in AD stage (Alzheimer's, 2012).

However, recent evidence from both genetic and clinical studies revealed that the patterns of disease progression profile could be more complicated than only three clinical stages, suggesting there may exist more states particularly between presymptomatic and prodromal stage. Li et al. (2015) suggested that the pathological process of AD begins before the symptoms appear. Mufson et al. (2016) mentioned that neurofibrillary tangles (NFTs), as a primary marker of AD, could be detected before the AD symptoms emerge. Li et al. (2016) suggested that the gene BACE1, as a sensitive biomarker of early neuropathology, has abnormal expression before the AD symptoms appear. All of these studies indicated that some of AD biomarkers show abnormalities in the presymptomatic stage of AD. Hence, there is a need for a more comprehensive model to evaluate the disease progression and thereby gain new insight into AD.

For this comprehensive model, it is vital to assess the disease states with sensitive biomarkers. Here, the disease state is a general term that refers to any quantifiable variable describing disease progression at a particular point in time (Holford, 2014). It is vital to assess the disease states with sensitive AD biomarkers. Many studies have evaluated the AD progression with various sets of biomarkers in order to find the biomarkers that have high sensitivity to detect AD (Biagioni and Galvin, 2011). For examples, Escudero et al. (2011) suggested using machine learning analysis to assess the pattern of AD progression with Magnetic resonance imaging (MRI) and Cerebrospinal fluid (CSF) biomarkers and Iordanescu et al. (2012) used an automatic plaque segmentation algorithm to classify plaque-related disease states. Madsen et al. (2015) studied the disease states with CSF-tau and Amyloid-β levels. Although the sensitive biomarkers can be selected for the model based on the literature, there are still some questions about AD pathological progression that are under investigation, for example, the transitioning profiles of AD progression are still unclear.

The latent transition models, where the unobserved membership at each time point is

measured with a set of manifest items, could be an ideal method for modeling AD development. Sukkar et al. (2012) proposed a hidden markov model which is also a latent transition model, to monitor the progression of AD with observed biomarkers and compare the hidden states with clinical diagnostic stages. But in their study, the transition probabilities and the profiles of disease progression have not been discussed in detail.

In this study, a latent transition model is employed to model AD progression profiles with a set of biomarkers. The volume of hippocampus and the Fluorodeoxyglucose-PET (FDG) are selected in this study for AD progression modeling. In these selected biomarkers, the volume of hippocampus represents structural or functional AD biomarkers and the Fluorodeoxyglucose-PET (FDG) represents pathological AD biomarkers. Our results indicate that AD progression is not as simple as the clinically diagnostic stages but a more complicated system, suggesting two profiles of AD progression. One of them is primarily characterized by FDG and the other by the volume of hippocampus. This latent transition model provides more details about the pattern of intermediary transitions between states.

This report is organized as follows. In Chapter 2, we briefly introduce the latent transition models and parameters estimation with EM algorithm. In Chapter 3, Alzheimers Disease Neuroimaging Initiative (ADNI) database is introduced as well as the biomarkers selection for modeling disease progression. The analysis results are given in Chapter 4. Finally, some conclusions and discussion of these two pathological profiles are given in Chapter 5.

Chapter 2

Statistical methods

2.1 Latent Transition Models

Modeling the progression of Alzheimer's Disease plays an important role in the development of potential treatments. Many studies of AD progression modeling used the three clinically diagnostic stages. For example, a genetic logistic regression model to predict the progression of AD three clinical stages was presented by Johnson (Johnson et al., 2014). The quantile regression models were employed to describe the conversions from NC to MCI as well as from MCI to AD with the evolution of quantitative biomarkers (Guercio et al., 2015). Schmidt-Richberg et al. (2015) used an apathy evaluation scale to predict the conversion from MCI to AD with the apathy value. However, there is a need for a more comprehensive statistical model which can detect more disease states to evaluate the disease progression.

Latent transition models, as one of the analyses of stage-sequential processes, can be used to assess the progression of Alzheimer's Disease. The latent transition models has been well discussed in the public health field (Chen and Page, 2016). Currently, many studies used latent transition models to evaluate the latent classes and the progression of disease such as AIDS (Dagne, 2016; Sweeting et al., 2010).

In this study, we use the latent transition models to monitor the progression of AD. This method assumes that the latent states at each time point are unobserved, but can be

measured with a set of manifest items (categorized biomarkers). The latent transition model is an extension of latent class models, suggesting it can not only provide the latent classes but also the associated transition probabilities (Chung et al., 2005). Therefore, we can use the latent transition model to assess the transitioning profiles of the disease.

For our real data analysis, we consider only two biomarkers and two time points. The details of biomarkers and data can be found in Chapter 3.

2.2 Statistical Model of Latent Transition Analysis

We first introduce the latent transition models in general. We use the similar notation of Chung's paper (Chung et al., 2005, 2008). Assuming the model with n latent states can be measured at T times, the biomarker responses of ith individual for all the times $t = 1, 2, 3, ..., T$ is:

$$\mathbf{Y}_i = (\boldsymbol{Y}_{i1}, \boldsymbol{Y}_{i2}, \cdots, \boldsymbol{Y}_{iT}),$$

where $\boldsymbol{Y}_{it} = (Y_{i1t}, \cdots, Y_{iMt})$, is a vector of M items (biomarkers) and the biomarker response $\boldsymbol{Y}_{imt} = k$, $k \in \{1, \cdots, J\}$. Let $\mathbf{L}_i = (L_{i1}, L_{i2}, L_{i3}, \cdots, L_{iT})$ be the latent class membership of ith individual at $t = 1, 2, \cdots, T$ and $L_{it} \in \{1, \cdots, s\}$. Simply, L_{it} is ranging from 1 to s, denoting the state of disease. The class variable L_{it} can be measured by $\boldsymbol{Y}_{it}$.

For this model, we assume the following: (1) $Y_{i1t}, \cdots, Y_{iMt}$ are conditionally independent given L_{it}; (2) the sequence of L_{it} constitutes a first-order markov chain for $t = 1, \cdots, T$. Under those assumptions, the complete likelihood function of $\boldsymbol{Y}_i$ and $\mathbf{L}_i$ is given by:

$$P(\boldsymbol{Y}_{i1} = \boldsymbol{y}_{i1}, \cdots, \boldsymbol{Y}_{iT} = \boldsymbol{y}_{iT}, \mathbf{L}_i = \mathbf{l}_i)$$

$$= \delta_{l_{i1}} \tau_{l_{i2}|l_{i1}} \cdots \tau_{l_{iT}|l_{i,T-1}} \prod_{m=1}^{M} \prod_{k=1}^{J} \prod_{t=1}^{T} \rho_{mkt|l_t}^{I(y_{imt}=k)}$$

$$= \delta_{l_{i1}} \prod_{t=2}^{T} \tau_{l_{it}|l_{i,t-1}} \prod_{m=1}^{M} \prod_{k=1}^{J} \prod_{t=1}^{T} \rho_{mkt|l_t}^{I(y_{imt}=k)},$$

where $\delta_{l_{i1}} = P(L_{i1} = l_{i1})$ and $\tau_{l_{it}|l_{i,t-1}} = P(L_{it} = l_{it}|L_{i,t-1} = l_{i,t-1})$ and the $I(y = k)$ as the

indicator function which equals 1 if y equals k and 0 otherwise. The marginal probabilities of each membership at time $t(t > 1)$ can be computed as

$$\delta_{l_{it}} = P(L_{it} = l_{it}) = \sum_{l_{i1}=1}^{s} ... \sum_{l_{i,t-1}=1}^{s} \delta_{l_{i1}} \prod_{t=2}^{T} \tau_{l_{it}|l_{i,t-1}}$$

Therefore, the contribution of the ith individual to the likelihood function of $\boldsymbol{Y}_{i1}, ..., \boldsymbol{Y}_{iT}$ is given by

$$P(\boldsymbol{Y}_{i1} = \boldsymbol{y}_{i1}, ..., \boldsymbol{Y}_{iT} = \boldsymbol{y}_{iT})$$

$$= \sum_{l_{i1}=1}^{s} ... \sum_{l_{iT}=1}^{s} P(\boldsymbol{Y}_{i1} = \boldsymbol{y}_{i1}, ..., \boldsymbol{Y}_{iT} = \boldsymbol{y}_{iT}, \mathbf{L}_i = \mathbf{l}_i)$$

$$= \sum_{l_{i1}=1}^{s} ... \sum_{l_{iT}=1}^{s} \delta_{l_{i1}} \prod_{t=2}^{T} \tau_{l_{it}|l_{i,t-1}} \prod_{m=1}^{M} \prod_{k=1}^{J} \prod_{t=1}^{T} \rho_{mkt|l_t}^{I(y_{imt}=k)}$$

The following sets of parameters are estimated from latent transition models: i) $\boldsymbol{\delta}$ parameters : latent states membership probabilities at baseline time point; ii) $\boldsymbol{\tau}$ parameters : probabilities of transitions between latent states over time; iii) $\boldsymbol{\rho}$: item-response probabilities conditional on latent status membership and time. In addition, the $\boldsymbol{\rho}$ parameters can explain the correspondence between the observed items and the latent states when covariates are included.

In this study, we only consider two biomarkers ($M = 2$) and two time points ($T = 2$).

2.3 Parameters Estimation with EM Algorithm

The parameters of latent transition model can be estimated using an EM algorithm. The expectation-maximization (EM) algorithm is an iterative method which is for finding the maximum likelihood estimation of parameters in statistical models (Dempster et al., 1977). The iteration of EM algorithm alternates between two steps: expectation (E) step and maximization (M) step. The E step is to create the expectation of the log-likelihood function with the current estimate for the parameters. The M step is to computer the parameters

that can maximizing the expected log-likelihood from E step.

For our example ($T = 2$ and $M = 2$), the E-step, we compute the conditional probability that each individual is a member of class l_1 at $t = 1$ and class l_2 at $t = 2$ given the biomarkers $\boldsymbol{y}_i = (y_{i1}, y_{i2})$ of ith subject,

$$\hat{\xi}_{i(l_1,l_2)} = P[L_1 = l_1, L_2 = l_2 | y_1, y_2] = \frac{\delta_{l_1}\tau_{l_2|l_1}\prod_{t=1}^{2}\prod_{m=1}^{2}\prod_{k=1}^{J}\rho_{mkt|l_t}^{I(y_{mt}=k)}}{\sum_{l_1=1}^{s}\cdots\sum_{l_T=1}^{s}\delta_{l_1}\tau_{l_2|l_1}\prod_{t=1}^{2}\prod_{m=1}^{2}\prod_{k=1}^{J}\rho_{mkt|l_t}^{I(y_{mt}=k)}}$$

In the M-step, the updated parameter estimates by

$$\hat{\delta}_{l_1} = \frac{\hat{\eta}_{l_1}^{(1)}}{n}, \quad \hat{\tau}_{l_2|l_1} = \frac{\hat{\eta}_{l_1,l_2}}{\hat{\eta}_{l_1}^{(1)}}, \quad \hat{\rho}_{mkt|l} = \frac{\hat{\eta}_{mkt|l}^{(1)} + \hat{\eta}_{mkt|l}^{(2)}}{\hat{\eta}_{l}^{(1)} + \hat{\eta}_{l}^{(2)}}$$

where $\hat{\eta}_{l_1,l_2} = \sum_i \hat{\xi}_{i(l_1,l_2)}$, $\hat{\eta}_{l_1}^{(1)} = \sum_{l_2}\hat{\eta}_{l_1,l_2}$, $\hat{\eta}_{l_2}^{(2)} = \sum_{l_1}\hat{\eta}_{l_1,l_2}$, $\hat{\eta}_{mkt|l}^{(1)} = \sum_{l_2}\sum_i I(y_{mt} = k)\hat{\xi}_{i(l,l_2)}$ and $\hat{\eta}_{mkt|l}^{(2)} = \sum_{l_1}\sum_i I(y_{mt} = k)\hat{\xi}_{i(l_1,l)}$. Finally, iterating between these two steps provides a sequence of parameter estimates that can converge to maximum of the likelihood function.

2.4 SAS package

The SAS modeling package we used in this study is from methodology center at Pennsylvania State University. The SAS code is given in Appendix.

Chapter 3

Real data application

3.1 ADNI data

In this study, the proposed latent transition model is performed with real data from Alzheimers Disease Neuroimaging Initiative (ADNI). The ADNI database was launched in 2003 as a public-private partnership, which was led by the principal investigator Michael W. Weiner, MD (Liu et al., 2016). The main goal of ADNI is to investigate whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers could be used to measure the progression of Alzheimers disease (AD) (Liu et al., 2016). The development of the ADNI database is in three stages: ADNI1, ADNI GO and ADNI2. The ADNI1, as the earliest stage, is aimed to find out more accurate diagnose method to detect AD in the earliest or pre-diagnosis stage. The biological biomarkers in ADNI1 contain PET (Positron Emission Tomography) and MRI (Magnetic Resonance Imaging), Cerebrospinal fluid (CSF), biomarkers from blood tests, neuropsychological tests and scores from tests of psychiatric function for mood (Grill et al., 2013). These biomarkers have been widely used in disease modeling studies. ADNI GO, as the second stage, was launched from June 2009 to 2011. Based on ADNI 1, ADNI GO aims at developing standardized methods for clinical tests and mostly focus on cognitive decline of AD. The third stage is ADNI2 which started in 2011. It mainly focus on comparing CSF and genetics biomarkers of non-symptomatic

Table 3.1: *Summary Table of ADNI database*

	Normal Control	**MCI**	**AD**
Mean of Age (S.D.)	74.67 (5.75)	73.98 (7.61)	75.06 (7.89)
Gender Ratio (M/F)	213/217	354/225	204/164

population and symptomatic population. Recently, the ADNI database has been used in various studies of AD (Platero and Tobar, 2016).

The data employed in this study are from ADNI1. The longitudinal data performed in this study with the latent transition model contain 2,126 participants from 2005 to 2014 who were examined with repeat visits every 6 months for a period of up to 108 months. The summary statistics of our data is presented in Table 3.1.

3.2 Biomarkers Selection

We use the pathologically sensitive biomarkers for disease progression modeling. The ideal bio-markers should be able to detect specific pathological features of AD. Overall, AD biomarkers can be categorized into following groups: pathological, structural or functional, and clinical. The pathological group of biomarkers such as CSF, Amyloid-β, tau, usually appears in the presymptomatic phase of AD. The structural or functional group reflects neuronal structure and functions appearing in the late presymptomatic or early prodromal phase. The clinical group (behavior, cognition and function) suggests dementia with impairment and loss of function in activities, which usually appears in the prodromal phase of AD (Biagioni and Galvin, 2011). The entire ADNI dataset includes pathological biomarkers such as CSF, tau and structural and functional biomarkers such as the FDG-PET and volume of hippocampus.

In this study, we select the volume of hippocampus and FDG for AD progression modeling. The volume of hippocampus represents structural or functional AD biomarkers and the FDG represents pathological AD biomarkers.

Fluorodeoxyglucose (FDG)-PET:

Positron emission tomography (PET) with 18F-2-deoxy-2-fluoro-D-glucose as a marker (FDG-PET) has been employed in a number of AD studies to examine regional crerbral metabolism

(Wang et al., 2013). Many studies have demonstrated that the decreased regional cerebral metabolism is indicative of who will progress to AD (Alzheimer's, 2016). The FDG, as an indicator of synaptic dysfunction and neurodegeneration in AD, suggests the hypo-metabolism in regional cerebral appearing in the MCI stage and exacerbating in AD stage. Some studies suggest the brain hypo-metabolism is related to the progression of synaptic pathology in AD (Mosconi et al., 2008). The distribution of FDG and the spaghetti plot of FDG for our data are given in this report (Fig. 3.1 and Fig. 3.2).

The Volume of Hippocampus:

The hippocampal atrophy is an indicator of Alzheimer's disease pathology and a potential marker to detect the progression of AD (Schuff et al., 2009). The decreasing of the hippocampal volume appears in the MCI stage and gets worse in the AD stage. The distribution of hippocampal volume and the spaghetti plots of hippocampus for our data are given in this report (Fig. 3.1 and Fig. 3.2).

In addition, gene APOE4 is considered as a covariate for modeling disease progression. Many studies revealed that the expressions of AD-related genes play an important role in the disease progression such as gene APOE4 (Cacciottolo et al., 2016; Koster et al., 2016). Therefore, this gene should be considered as one of covariates for AD progression modeling.

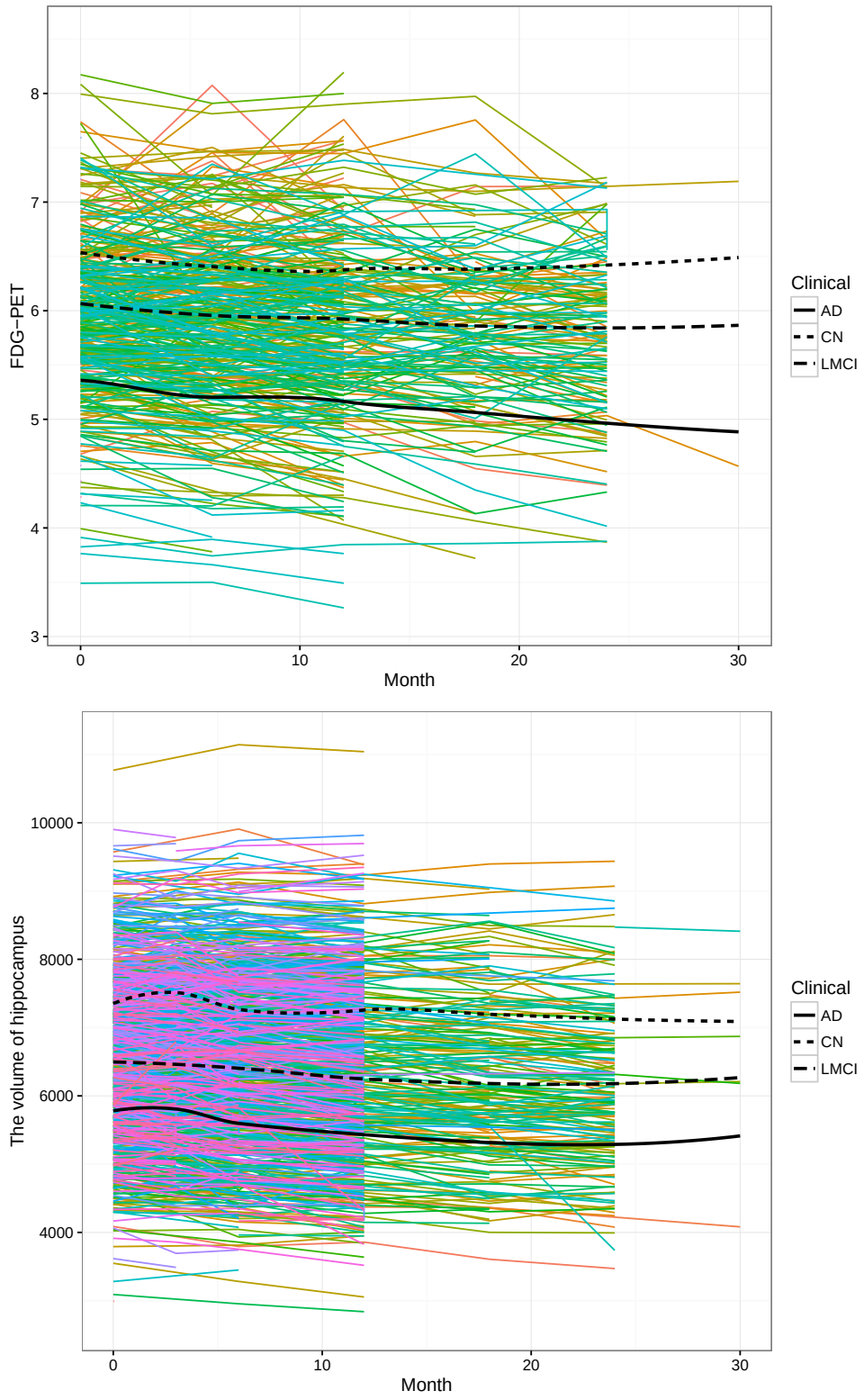

Figure 3.1: *The spaghetti plots of FDG-PET and the volume of hippocampus by the clinical results. AD, Alzheimer's disease; CN, clinically normal; LMCI, Late Mild Cognitive impairment.*

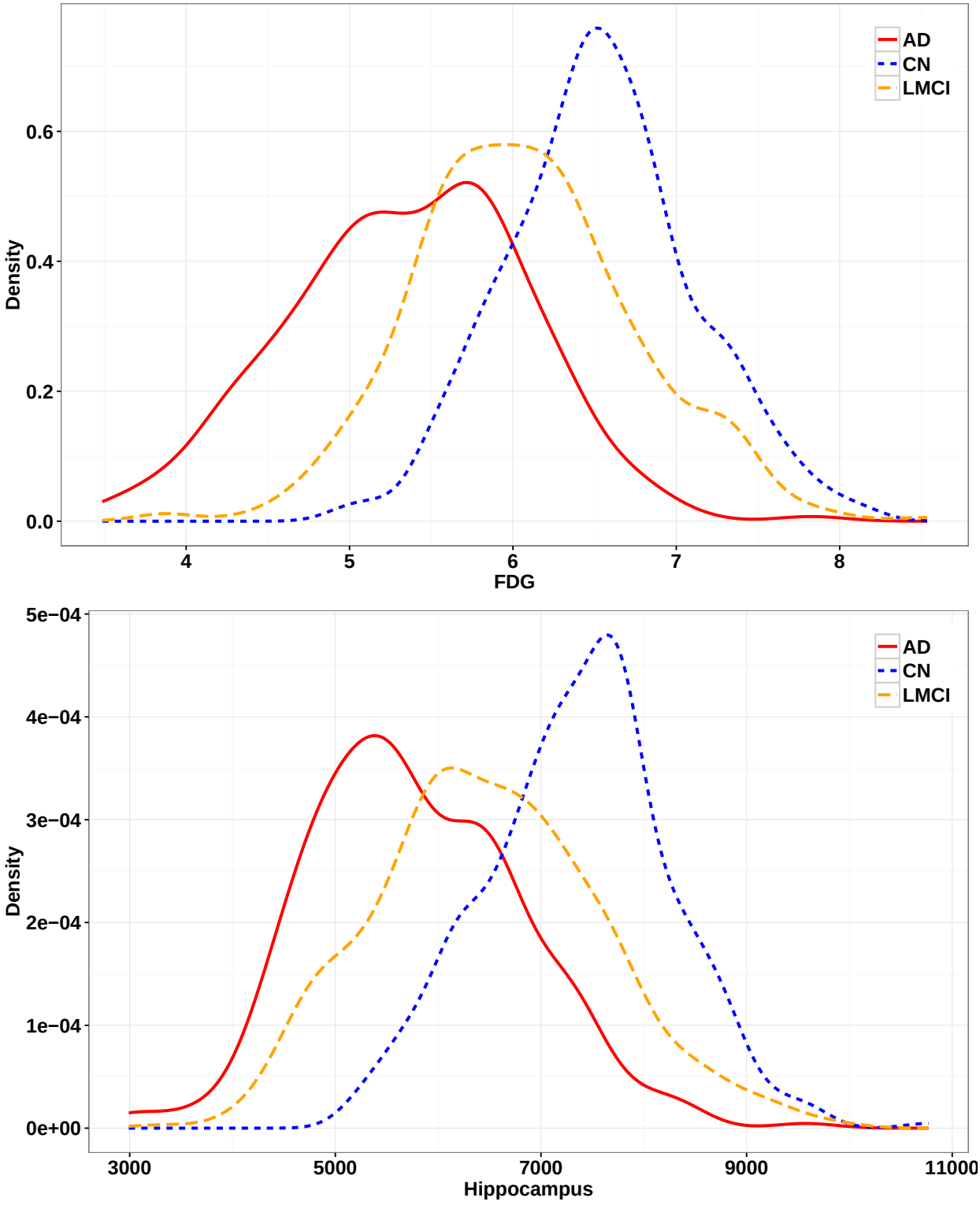

Figure 3.2: *The distributions of selected biomarkers by clinical results.AD, Alzheimer's disease; CN, clinically normal; LMCI, Late Mild Cognitive impairment.*

Chapter 4

Results: Latent transition models for ADNI data

4.1 Model selection

In practice, the number of latent classes of a latent transition model is often unknown, some model selection methods can be used for this purpose. In this study, Bayesian information criterion (BIC) and Akaike information criterion (AIC) are used to determine the number of latent states. BIC and AIC are most commonly used methods of assessing model fit penalized for the number of estimated parameters and the criteria for model selection among a finite set of models. The model with the lowest BIC or AIC is preferred.

Table 4.1: *Model Selection*

State	AIC	BIC
3	5573.20	5822.32
4	4889.06	5313.71
5	4197.14	4842.61
6	**3909.76**	**4821.34**
7	3769.83	4992.83

AIC, Akaike information criterion;
BIC, Bayesian information criterion

The details of AIC and BIC for models selection is given in Table 4.1. Among the five possible models, the model selection criterion, BIC, suggests 6 unobserved disease states based on the ADNI data.

4.2 Latent class and transition probabilities

Using latent transition models, the pattern of disease progression contains six latent states. The latent state of each individual can be determined by the posterior probabilities. Based on the sample characteristics of each states, we label them as Normal, H1, H2, F1, F2 and AD. Normal state suggests that both of FDG and hippocampus are normal (both of hippocampus and FDG with highest values among these six states); H1 state suggests that the hippocampal appears abnormality but the FDG doesn't appear abnormality (FDG with high value and the hippocampus with low value); H2 state represents that the abnormality of hippocampus worse than H1 but the FDG are still normal (FDG with high value but the hippocampus with lower value than H1); F1 suggests that the FDG appearing abnormal but the hippocampus are still normal (FDG with low value but the hippocampus with high value); F2 represents that the abnormality of FDG worse than F1 and the hippocampus is also with low value, but overall, the abnormality of FDG is dominant (FDG with lower value than F1 and the hippocampus with low value, but the abnormality of FDG is dominant); AD means that both of FDG and hippocampus are abnormal (both of FDG and Hippocampus with lowest values).

The latent transition model suggests two different pathological profiles based on the transition probabilities of each disease states in two years. One of them is determined by the volume of hippocampus (blue trajectory). The other is by the FDG (red trajectory) which is characterized by the FDG abnormality. The FDG profile has these following underlying disease states: Normal, F1, F2 and AD. The hippocampus profile has the following disease states: Normal, H1, H2 and AD.

Specifically, the normal state has two different transitioning trajectories. One is converting to H1 state and the other is to F1 state, revealing the existence of two different

pathological profiles. The conversion of participants classified in normal state at baseline time point is presented in Fig.4.2. In Hippocampal progression profile, the participants classified in H1 state at baseline time point only retain or convert to H2 in two years (see Fig.4.3). For the participants in H2 state at the beginning, they only retain or transit to AD state in two years (see Fig.4.4). Similarly, in FDG profile, the participants classified in F1 state at baseline year mostly retain or transit to F2 (see Fig.4.5), and very few of the participants transit to H2 (see Fig.4.1), suggesting there is very few overlap between these two profiles. The participants in F2 state retain or transit to AD state (see Fig.4.6) in two years. All of the participants classified in AD by our model at baseline time point remain in AD state in two years, indicating AD is an irreversible disease (see Fig.4.7).

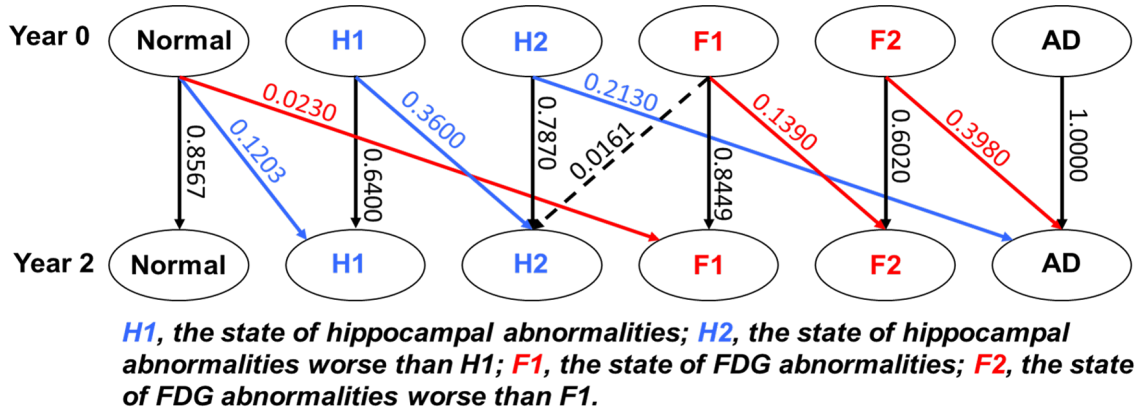

Figure 4.1: *The transition probabilities for our model. The profile determined by the volume of hippocampus is labeled in blue. The profile dominated by the FDG is labeled in red.*

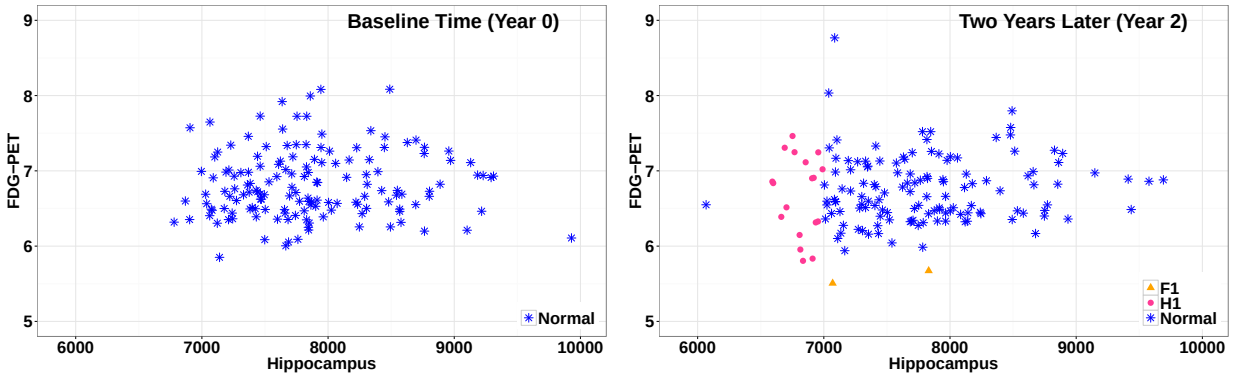

Figure 4.2: *The left graph is the distribution of FDG and hippocampus of participants classified as Normal at baseline time point and the right one is for the same participants after two years. H1, the state of hippocampal abnormality; F1, the state of FDG abnormality.*

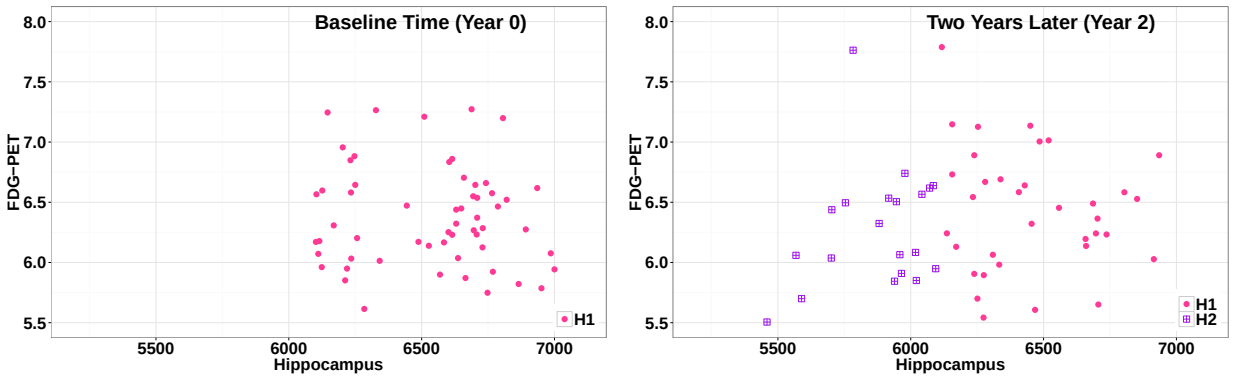

Figure 4.3: *The left graph is the distribution of FDG and hippocampus of participants classified as H1 at baseline time point and the right one is for the same participants after two years. H1, the state of hippocampal abnormality; H2, the state of Hippocampal abnormality worse than H1.*

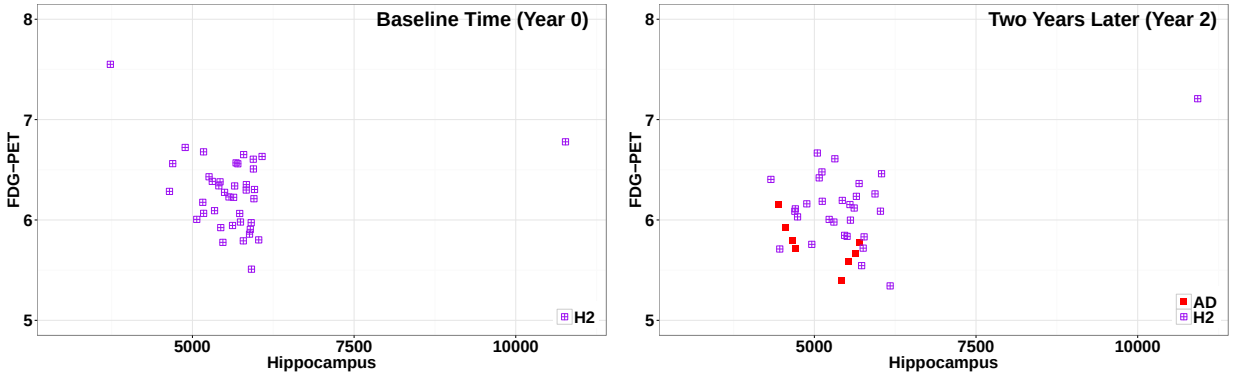

Figure 4.4: *The left graph is the distribution of FDG and hippocampus of participants classified as H2 at baseline time point and the right one is for the same participants after two years. H1, the state of hippocampal abnormality; H2, the state of Hippocampal abnormality worse than H1.*

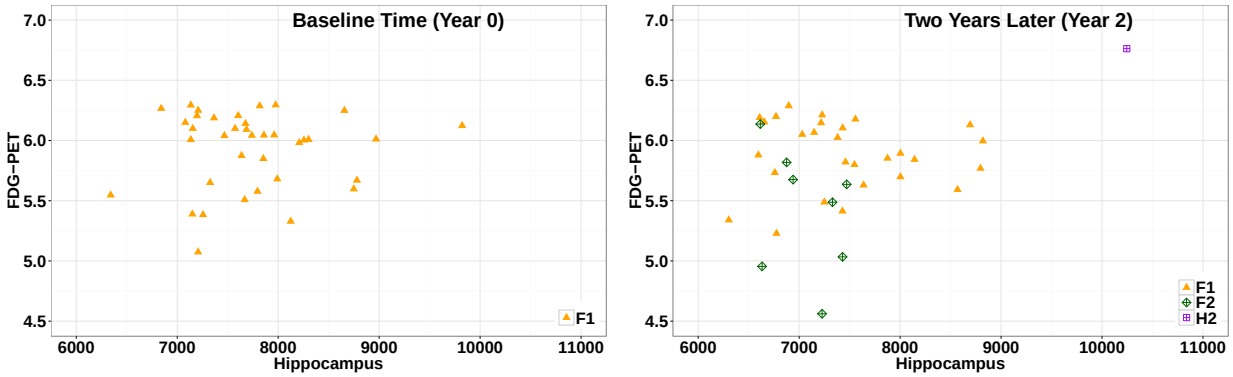

Figure 4.5: *The left graph is the distribution of FDG and hippocampus of participants classified as F1 at baseline time point and the right one is for the same participants after two years. F1, the state of FDG abnormality; F2, the state of FDG abnormality worse than F1.*

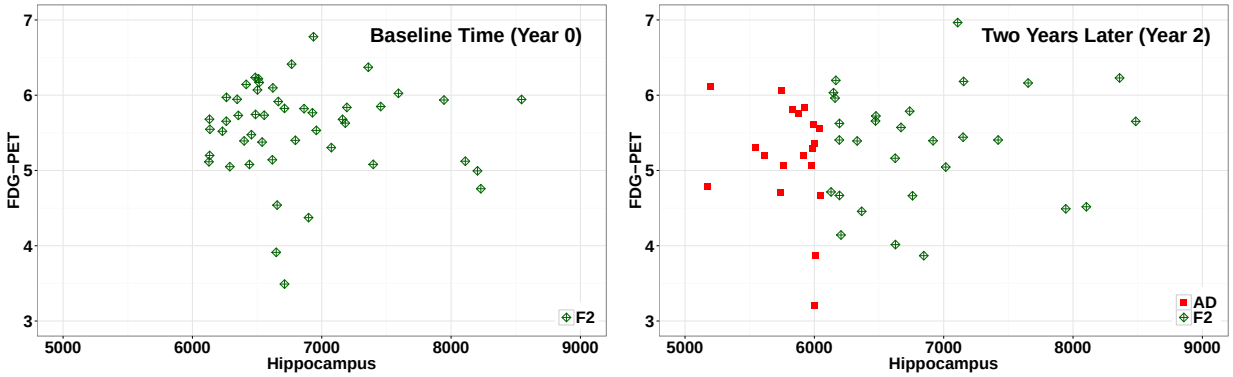

Figure 4.6: *The left graph is the distribution of FDG and hippocampus of participants classified as F2 at baseline time point and the right one is for the same participants after two years. F1, the state of FDG abnormality; F2, the state of FDG abnormality worse than F1.*

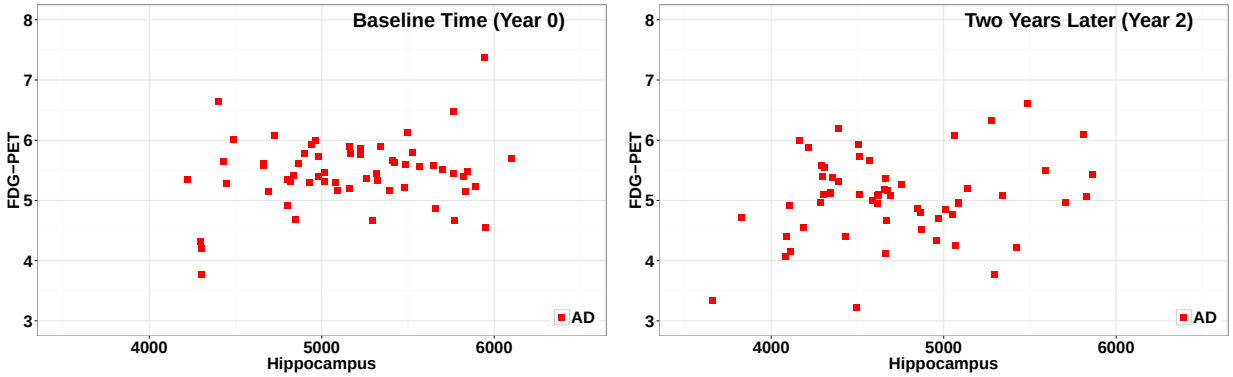

Figure 4.7: *The left graph is the distribution of FDG and hippocampus of participants classified as AD at baseline time point and the right one is for the same participants after two years.*

4.3 Two profiles

Compared the distribution of FDG and Hippocampus by the clinical results, the distribution classified by latent states are clearly distinct (see Fig. 4.8). The decreasing Hippocampal value suggests the abnormality of hippocampus. The H1 state has the volume of hippocampus value decreased but the FDG-PET value retain. The H2 state has the hippocampus value decreases more than H1, indicating worse hippocampal abnormality than H1. Similarly, the decreasing FDG value imply the abnormality of metabolism in cerebral. The F1 state suggests the hypo-metabolism in cerebral and F2 state is worse than F1. The AD state has both lowest hippocampal value and FDG value.

The distributions of the FDG profile and hippocampal profile are clearly distinct which suggest the existing of two progression profiles (see Fig. 4.9).

The participants with normal clinical diagnosis are classified by our model in the states other than normal, indicating the abnormalities of FDG and hippocampus appear in presymptomatic stage (see Fig. 4.10). The participants with MCI clinical diagnosis are classified in other states by our model, suggesting more latent disease state may exist (see Fig. 4.11). Among the participants with AD clinical diagnosis, some of them are classified in the states other than AD by our model, indicating a more complicated disease progression pattern (see Fig. 4.12).

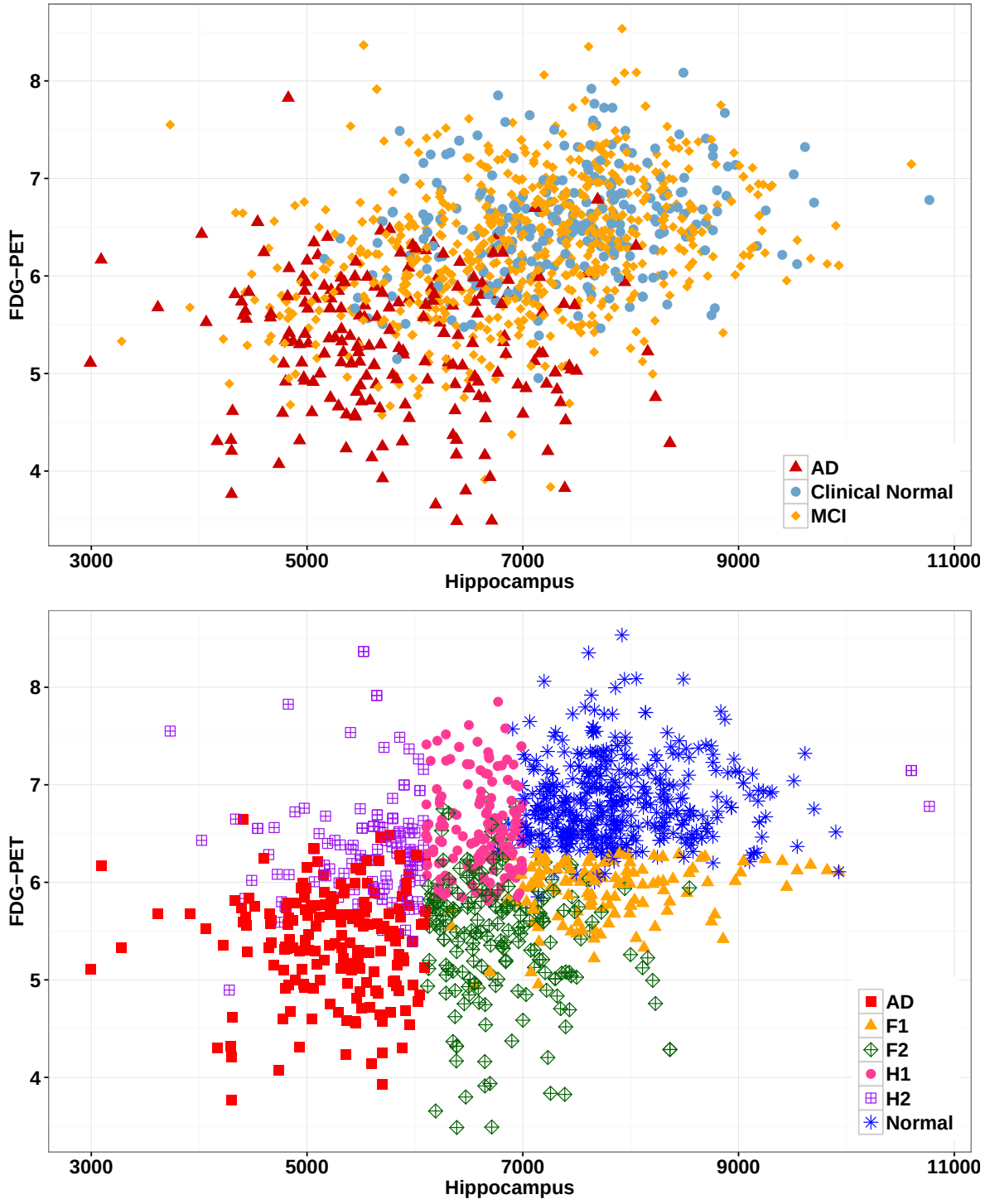

Figure 4.8: *The top graph is the distribution of FDG and hippocampus by clinical results at* ***baseline time (Year 0)****: CN, clinical normal; MCI, Mild Cognitive Impairment; AD, Alzheimer's disease. The bottom graph is the distribution of FDG and hippocampus by our model at* ***baseline time (Year 0)***

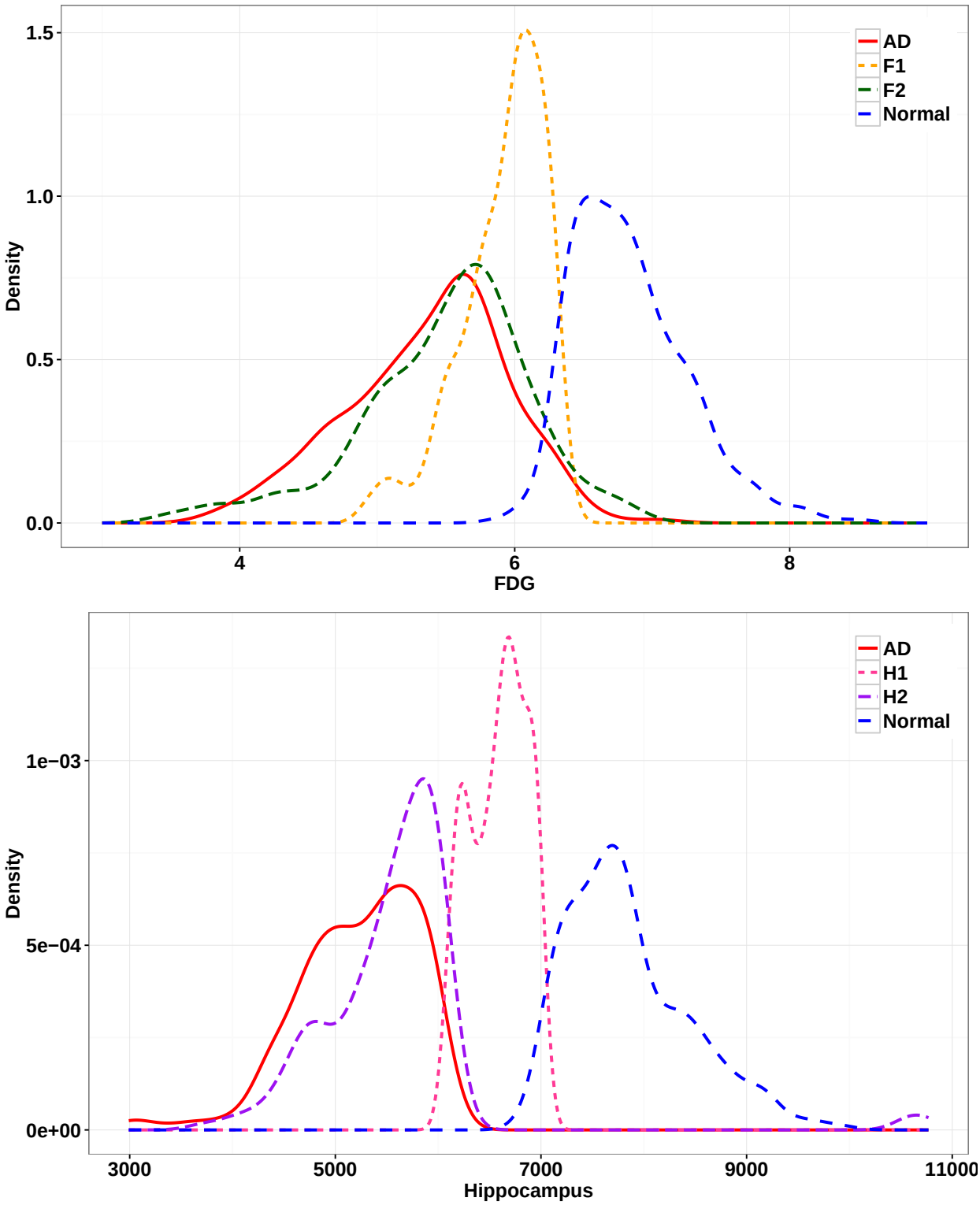

Figure 4.9: *The first graph is the distribution of the FDG profile at* ***baseline time (Year 0)****. The second graph is the distribution of the hippocampal profile at* ***baseline time (Year 0)****. H1, the state of hippocampal abnormality; H2, the state of hippocampal abnormality worse than H1; F1, the state of FDG abnormality; F2, the state of FDG abnormality worse than F1.*

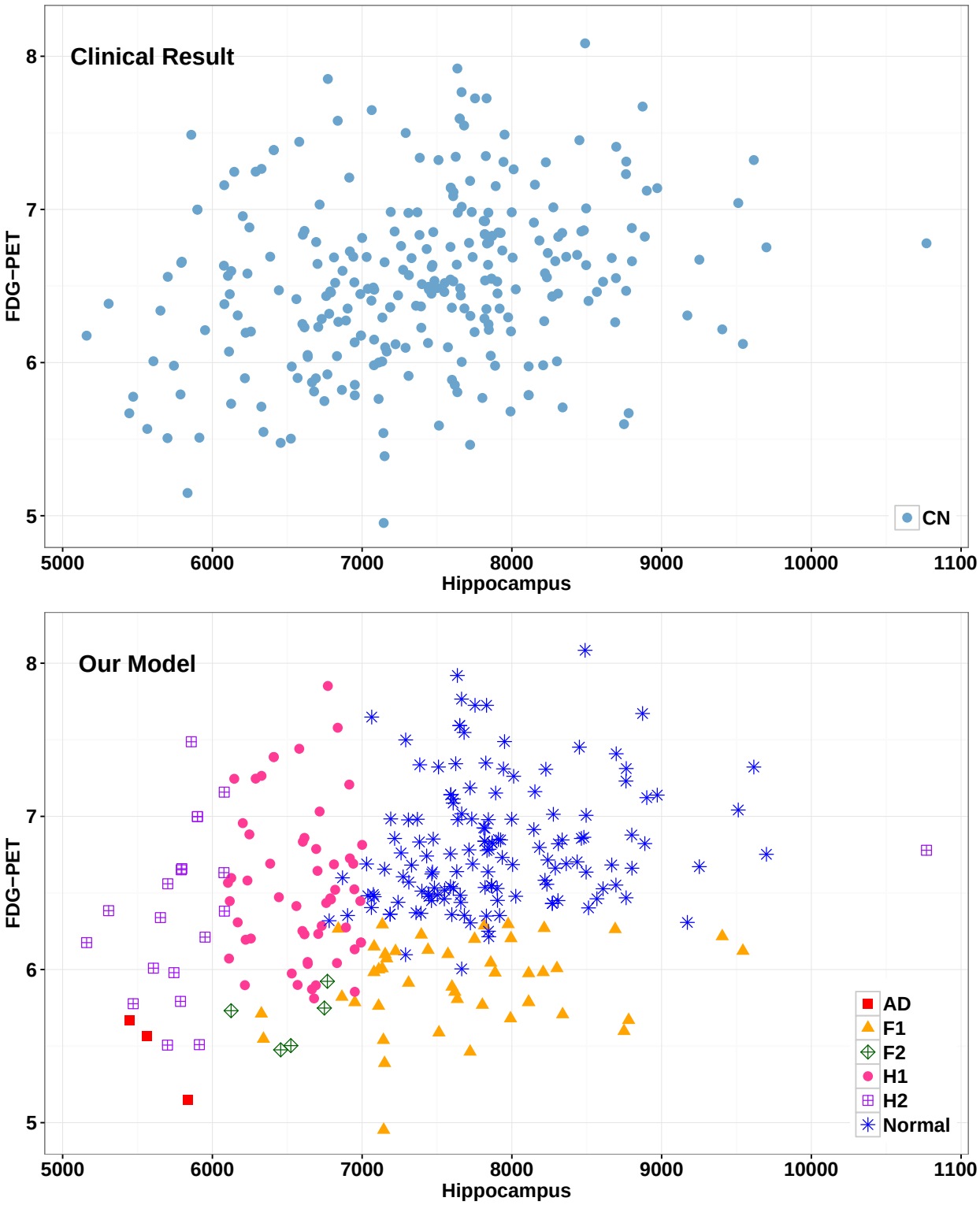

Figure 4.10: *The top graph is the distribution of FDG and hippocampus of clinically normal participants at* ***baseline time (Year 0)****. The bottom graph is the distribution of biomarkers of clinically normal participants by our model at* ***baseline time (Year 0)****. H1, the state of hippocampal abnormality; H2, the state of hippocampal abnormality worse than H1; F1, the state of FDG abnormality; F2, the state of FDG abnormality worse than F1.*

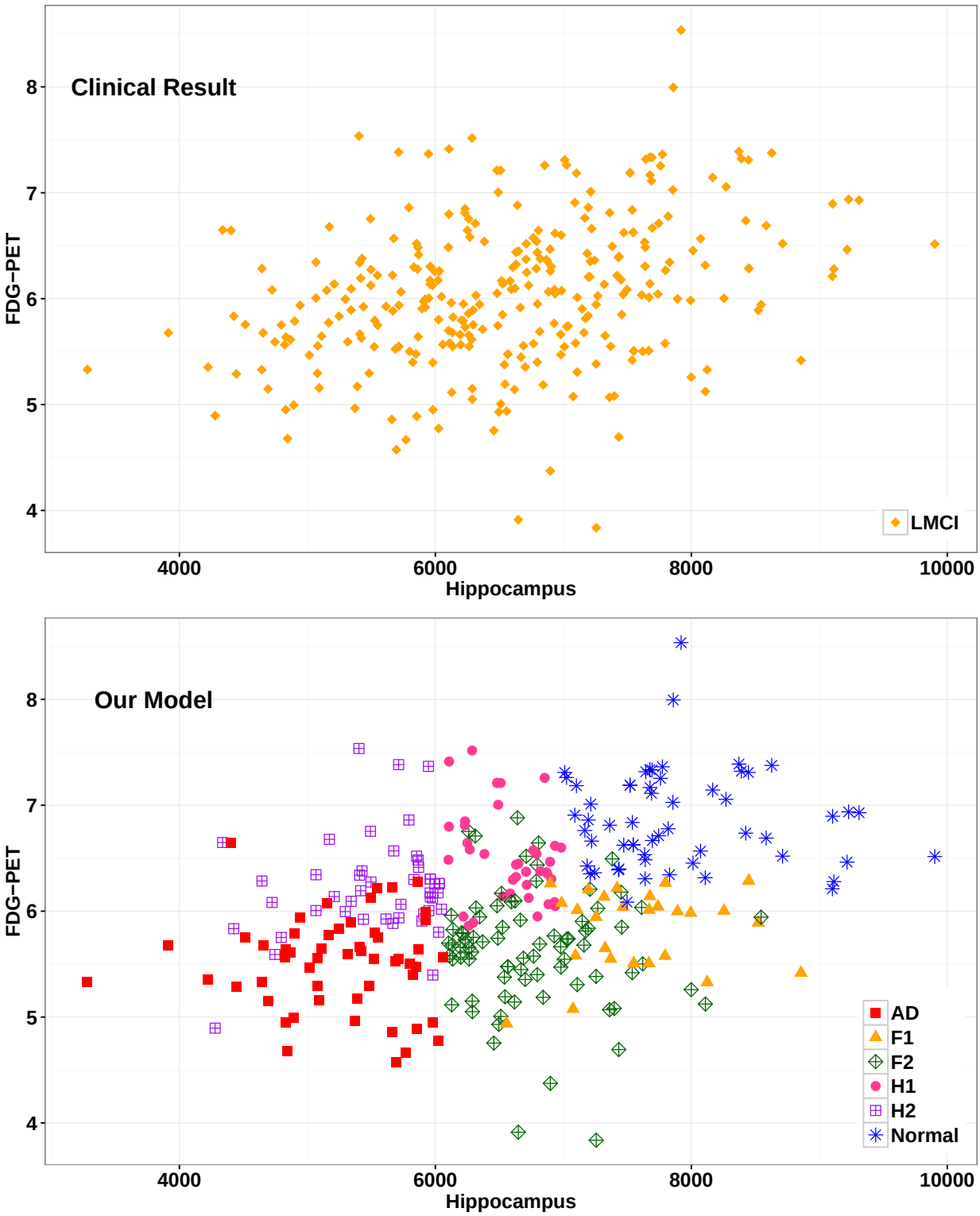

Figure 4.11: *The top graph is the distribution of FDG and hippocampus of MCI participants at* ***baseline time (Year 0)****. The bottom graph is the distribution of biomarkers of MCI participants by our model at* ***baseline time (Year 0)****. H1, the state of hippocampal abnormality; H2, the state of hippocampal abnormality worse than H1; F1, the state of FDG abnormality; F2, the state of FDG abnormality worse than F1.*

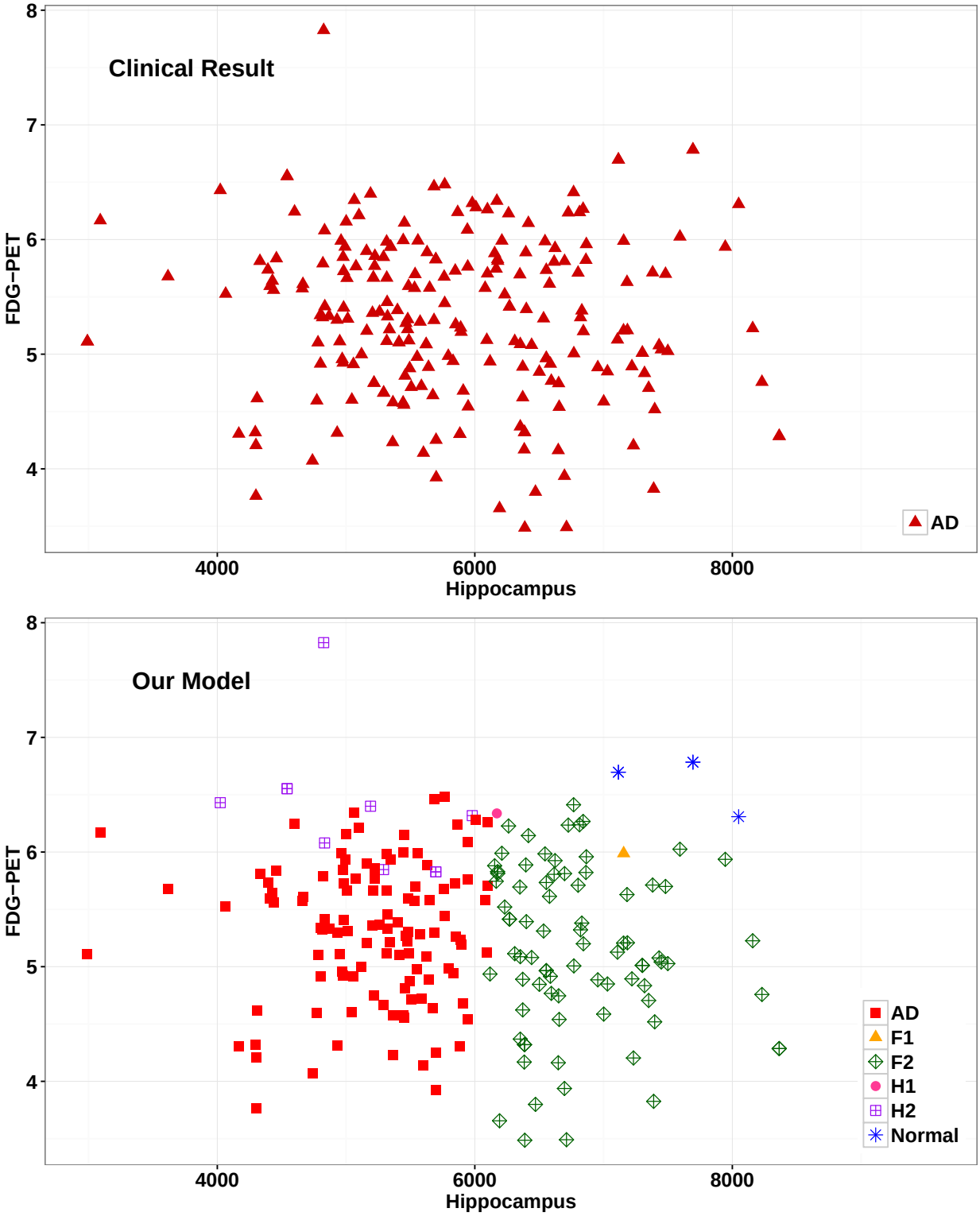

Figure 4.12: *The top graph is the distribution of FDG and hippocampus of AD participants at* ***baseline time (Year 0)****. The bottom graph is the distribution of biomarkers of AD participants by our model at* ***baseline time (Year 0)****. H1, the state of hippocampal abnormality; H2, the state of hippocampal abnormality worse than H1; F1, the state of FDG abnormality; F2, the state of FDG abnormality worse than F1.*

Chapter 5

Discussion

The main findings of our study are as follows: i) the existence of six latent states of disease progression and two different pathological profiles. ii) FDG-PET and the volume of hippocampus could be presymptomatic biomarkers.

5.1 Two profiles

This study revealed the existence of six latent states of disease progression and two different pathological profiles. One progression profile is determined by the biomarker FDG-PET and the other by the volume of hippocampus. The FDG-PET is widely used to investigate the cerebral metabolic rate of glucose and it is a sensitive biomarker to discriminate AD from other neurodegenerative diseases. A low FDG-PET suggests hypo-metabolism in cerebral (Herholz, 2003). Hence, the progression profile determined by FDG-PET is associated with the hypo-metabolism in cerebral. The other profile which is determined by the volume of hippocampus could be described by the hippocampal atrophy. Thereby, the profiles of disease progression could be described by two different pathologies: i) hypo-metabolism in cerebral, ii) hippocampal atrophy.

Although the patterns of structural and metabolic brain alterations in AD are being refined, the exact relationships between hippocampal atrophy and hypometabolism in cerebral

have not been well discussed in the literature. One related study suggested that the cerebral hypo-metabolism could be related to the hippocampal atrophy and the relationship of them depends on inducing factors such as the presence of Amyloid-β (Chtelat et al., 2008). This study indicates the existence of these two different profiles of AD may be associated with Amyloid-β. for the future study, it will be interesting to include Amyloid-β into the model.

5.2 Presymptomatic biomarkers

Since the pathological profiles of FDG and hippocampus suggest that the hypo-metabolism and the hippocampal atrophy appear before the AD symptoms emerge, both of FDG-PET and the volume of hippocampus could be presymptomatic biomarkers. This finding is consistent with the literature. For examples, Mistur et al. (2009) suggested that the hypometabolism occurs in the preclinical stage of AD and the FDG-PET could be a pathological biomarker for early detection of AD. The presymptomatic biomarkers such as FDG-PET are critical for developing potential treatments of AD in presymptomatic or minimally symptomatic populations (Cummings and Jeste, 2007). Andrews et al. (2016) suggested that the volume of hippocampus is one of presymptomatic biomarkers which is related to Amyloid-β. Based on our findings, we may suggest that FDG and the volume of hippocampus can serve as important biomarkersfor the early detection of AD.

5.3 Limitation and future study

Overall, this study revealed the existence of two different progression profiles of Alzheimers disease, suggesting a new way to evaluate the disease progression.

The parameters in our study are estimated by EM algorithm. However, the latent transition analysis may contain some unusual features that can cause difficulties in the EM algorithm especially when with small samples (Chung et al., 2008).

In addition, the model selection methods AIC and BIC can be highly unstable over repeated sampling (Preacher and Merkle, 2012). Thus, selecting model by AIC and BIC

should be more cautious.

The screening test for AD early detection is critical for the prevention of AD. For this, it is essential to identify the most sensitive presymptomatic biomarkers that can detect the early symptoms of AD or clinical impairment. This is the subject of future study.

Bibliography

Assoc. Alzheimer's. 2012 alzheimer's disease facts and figures. *Alzheimers Dement.*, 8(2): 131–168, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/22404854.

Assoc. Alzheimer's. *Alz.org*, 2016. URL http://www.alz.org/facts/.

K. A. Andrews, C. Frost, M. Modat, M. J. Cardoso, C. C. Rowe, V. Villemagne, N. C. Fox, S. Ourselin, and J. M. Schott. Acceleration of hippocampal atrophy rates in asymptomatic amyloidosis. *Neurobiol Aging.*, 39:99–107, 2016. URL https://www.ncbi.nlm.nih.gov/pubmed/26923406.

M. C. Biagioni and J. E. Galvin. Using biomarkers to improve detection of alzheimer's disease. *Neurodegener Dis Manag.*, 1(2):127–139, 2011. URL https://www.ncbi.nlm.nih.gov/pubmed/22076127.

M. Cacciottolo, A. Christensen, A. Moser, J. Liu, C. J. Pike, C. Smith, M. J. LaDu, P. M. Sullivan, T. E. Morgan, E. Dolzhenko, A. Charidimou, L. O. Wahlund, M. K. Wiberg, S. Shams, and G. C. Chiang. The apoe4 allele shows opposite sex bias in microbleeds and alzheimer's disease of humans and mice. *Neurobiol Aging.*, 37:47–57, 2016. URL http://www.ncbi.nlm.nih.gov/pubmed/26686669.

X. Chen and A. Page. Stability and instability of subjective well-being in the transition from adolescence to young adulthood: Longitudinal evidence from 20991 young australians. *PLoS One.*, 11(5):e0156399, 2016. URL http://www.ncbi.nlm.nih.gov/pubmed/27232183.

G. Chtelat, B. Desgranges, B. Landeau, F. Mzenge, J. B. Poline, V. Sayette, F. Viader, F. Eustache, and J. C. Baron. Direct voxel-based comparison between grey matter hy-

pometabolism and atrophy in alzheimer's disease. *Brain.*, 131(Pt 1):60–71, 2008. URL https://www.ncbi.nlm.nih.gov/pubmed/18063588.

H. Chung, Y. Park, and S. T. Lanza. Latent transition analysis with covariates: pubertal timing and substance use behaviours in adolescent females. *Stat Med.*, 24(18):2895–2910, 2005. URL http://www.ncbi.nlm.nih.gov/pubmed/16134129.

H. Chung, S. Lanza, and E. Loken. Latent transition analysis: inference and estimation. *Stat Med.*, 27(11):1834–1854, 2008.

J. L. Cummings and D. V. Jeste. Pharmacotherapy of neuropsychiatric syndromes in neurologic disorders: definitional and regulatory aspects. *Psychopharmacol Bull.*, 40(4):89–98, 2007. URL https://www.ncbi.nlm.nih.gov/pubmed/18227780.

G. Dagne. Bayesian bent-cable growth mixture tobit models for longitudinal data with skewness and detection limit: Application to aids studies. *Stat Med.*, page ahead of print, 2016.

Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

J. Escudero, J. Zajicek, and E. Ifeachor. Selection of subjects for clinical trials in alzheimers disease and mild cognitive impairment with machine learning analysis of mri and csf biomarkers. *Trials.*, 12(Suppl 1):A18, 2011. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3287731/.

J. D. Grill, L. Di, P. H. Lu, C. Lee, J. Ringman, L. G. Apostolova, N. Chow, O. Kohannim, J. L. Cummings, P. M. Thompson, and D. Elashoff. Estimating sample sizes for predementia alzheimer's trials based on the alzheimer's disease neuroimaging initiative. *Neurobiol Aging.*, 34(1):62–72, 2013. URL http://www.ncbi.nlm.nih.gov/pubmed/22503160.

B. J. Guercio, N. J. Donovan, C. E. Munro, S. L. Aghjayan, S. E. Wigman, J. J. Locascio, R. E. Amariglio, D. M. Rentz, K. A. Johnson, R. A. Sperling, and G. A. Marshall. The

apathy evaluation scale: A comparison of subject, informant, and clinician report in cognitively normal elderly and mild cognitive impairment. *J Alzheimers Dis.*, 47(2):421–432, 2015. URL http://www.ncbi.nlm.nih.gov/pubmed/26401564.

K. Herholz. Pet studies in dementia. *Ann Nucl Med.*, 17(2):79–89, 2003. URL https://www.ncbi.nlm.nih.gov/pubmed/12790355.

N. Holford. Clinical pharmacology = disease progression + drug action. *British Journal of Clinical Pharmacology*, 79:18–27, 2014. URL http://onlinelibrary.wiley.com/doi/10.1111/bcp.12170/abstract.

G. Iordanescu, P. N. Venkatasubramanian, and A. M. Wyrwicz. Automatic segmentation of amyloid plaques in mr images using unsupervised support vector machines. *Magn Reson Med.*, 67:1794–1802, 2012. URL http://onlinelibrary.wiley.com/doi/10.1002/mrm.23138/abstract.

P. Johnson, L. Vandewater, W. Wilson, P. Maruff, G. Savage, P. Grahan, L. S. Macaulay, K. A. Ellis, C. Szoeke, R. N. Martins, C. C. Rowe, C. L. Masters, D. Ames, and P. Zhang. Genetic algorithm with logistic regression for prediction of progression to alzheimer's disease. *BMC Bioinformatics*, 15:SUPPLEMENT 16, 2014. http://www.ncbi.nlm.nih.gov/pubmed/25521394.

K. P. Koster, C. Smith, A. C. Valencia-Olvera, G. R. Thatcher, L. M. Tai, and M. J. LaDu. Rexinoids as therapeutics for alzheimer disease: Role of apoe. *Curr Top Med Chem.*, Epub ahead of print, 2016. URL http://www.ncbi.nlm.nih.gov/pubmed/27320328.

C. Li, N. M. Dowling, and R. Chappell. Quantile regression with a change-point model for longitudinal data: An application to the study of cognitive changes in preclinical alzheimer's disease. *Biometrics*, 71(3):625–635, 2015. https://www.ncbi.nlm.nih.gov/pubmed/25892034.

L. Li, J. Luo, D. Chen, J. Tong, L. Zeng, Y. Cao, J. Xiang, X. Luo, J. Shi, H. Wang, and J. Huang. Bace1 in the retina: a sensitive biomarker for monitoring early pathological

changes in alzheimer's disease. *Neural Regen Res.*, 11(3):47–53, 2016. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4829010/.

S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng. Cross-view neuroimage pattern analysis in alzheimer's disease staging. *Front Aging Neurosci.*, 8:23, 2016. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4763344/.

S. K. Madsen, G. Ver-Steeg, A. Mezher, N. Jahanshad, T. M. Nir, X. Hua, B. A. Gutman, A. Galstyan, and P. M. Thompson. Information-theoretic characterization of blood panel predictors for brain atrophy and cognitive decline in the elderly. *Proc IEEE Int Symp Biomed Imaging.*, 2015:980–984, 2015. URL https://www.ncbi.nlm.nih.gov/pubmed/26413208.

R. Mistur, L. Mosconi, S. D. Santi, M. Guzman, Y. Li, W. Tsui, and M. J. Leon. Current challenges for the early detection of alzheimer's disease: Brain imaging and csf studies. *J Clin Neurol.*, 5(4):153–166, 2009. URL https://www.ncbi.nlm.nih.gov/pubmed/20076796.

L. Mosconi, A. Pupi, and M. J. De-Leon. Brain glucose hypometabolism and oxidative stress in preclinical alzheimer's disease. *Ann N Y Acad Sci.*, 1147:180–195, 2008. URL http://www.ncbi.nlm.nih.gov/pubmed/19076441.

E. J. Mufson, M. D. Ikonomovic, S. E. Counts, S. E. Perez, M. Malek-Ahmadi, S. W. Scheff, and S. D. Ginsberg. Molecular and cellular pathophysiology of preclinical alzheimer's disease. *Behav Brain Res.*, 311:54–69, 2016. URL https://www.ncbi.nlm.nih.gov/pubmed/27185734.

C. Platero and CM. Tobar. A fast approach for hippocampal segmentation from t1-mri for predicting progression in alzheimer's disease from elderly controls. *J Neurosci Methods.*, 270:61–75, 2016. URL http://www.ncbi.nlm.nih.gov/pubmed/27328371.

K. Preacher and E. Merkle. The problem of model selection uncertainty in structural equation modeling. *Psychological Methods*, 2012:1–14, 2012.

A. Schmidt-Richberg, R. Guerrero, C. Ledig, H. Molina-Abril, A. F. Frangi, and D. Rueckert. Multi-stage biomarker models for progression estimation in alzheimer's disease. *Inf Process Med Imaging*, 24:387–98, 2015. URL http://www.ncbi.nlm.nih.gov/pubmed/26221689.

N. Schuff, N. Woerner, L. Boreta, T. Kornfield, L. M. Shaw, J. Q. Trojanowski, P. M. Thompson, and C. R. Jack. Mri of hippocampal volume loss in early alzheimer's disease in relation to apoe genotype and biomarkers. *Brain.*, 132(Pt 4):1067–77, 2009. URL http://www.ncbi.nlm.nih.gov/pubmed/19251758.

R. Sukkar, E. Katz, Y. Zhang, D. Raunig, and B. T. Wyman. Disease progression modeling using hidden markov models. *Conf Proc IEEE Eng Med Biol Soc.*, 2012:2845–2848, 2012. URL https://www.ncbi.nlm.nih.gov/pubmed/23366517.

M. Sweeting, V. Farewell, and D. De-Angelis. Multi-state markov models for disease progression in the presence of informative examination times: an application to hepatitis c. *Stat Med.*, 29(11):1161–74, 2010.

N. Wang, H. Zhai, and Y. Lu. Is fluorine-18 fluorodeoxyglucose positron emission tomography useful for the thyroid nodules with indeterminate fine needle aspiration biopsy? a meta-analysis of the literature. *J Otolaryngol Head Neck Surg.*, 42:38, 2013. URL http://www.ncbi.nlm.nih.gov/pubmed/24228840.

Appendix A

SAS code

```
/*LTA with AOPE4 covariates*/
proc lta data=ADfinal ;
title1 'Model1';
nstatus 6;
ntimes 2;
items FDG00ca MMSE00ca Hippocampus00ca
FDG24ca MMSE24ca Hippocampus24ca;
categories 3 3 3;
measurement times;
covariates1 APOE4x ;
reference1 1 ;
seed 592667;
run;
```