Alzheimer's Disease Pathological Phases and Biomarkers: A Statistical Analysis

Executive Summary

In this project, we give comprehensive analysis on Alzheimer's disease data. A model to classify different Alzheimer's phases (Normal, Mild Cognitive Impairment, and Dementia) is developed using proportional odds ratio technique. Risk score formula is proposed and shown that people who had Mild Cognitive Impairment and progressed to Dementia within a year, on average, had higher score. Finally, a regression model to explain variability in Hippocampal volume of patients is discussed.

Introduction

Introduced in 1910, Alzheimer's disease is developed through an irreversible process spanning across several years. Patients typically experience memory loss, impaired cognitive ability, language problems, and stress [1]. Even though over 5.8 million Americans are suffering from it, our capability in accurately screening is still limited especially for Mild Cognitive Impairment (MCI), the early stage of the disease, leading to late treatment [2].

Moreover, the diagnostic process also involves several metrics and subjective components including standardized test (Mini-Mental State Exam MMSE) which is not sensitive to MCI, interviews, and sometimes CT, MRI or PET scan for biomarkers to determine the progress of the disease [3,4,5]. To contribute to this area, three main objectives will be addressed in this project.

1. To build statistical models based on simple neuropsychological tests that could be administered online to predict Alzheimer's phases in patients, and compare their predictive capability with other methods proposed within research community.
2. To construct a risk score metric using the model in the first section in capturing the potential conversion to more severe states of Alzheimer's disease in the future.
3. To investigate the capability of these basic tests in estimating some of the biomarkers that have been believed to play major roles in the progression of the disease [6].

Data and Exploratory Data Analysis

The data was provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) with more than 14,000 records of 115 variables. In this project, we focus on ADNI1 cohort (797 participants). The variables being investigated include demographic information of participants, longitudinal diagnostic results, biomarkers, and Neuropsychological test results [7]. For the diagnosis, each person was classified into one of the three groups: Control Normal (CN), Mild Cognitive Impairment (MCI), and Dementia (AD). The full data dictionary is provided in the appendix section.

The neuropsychological tests explored in this project included:

1. Rey Auditory Verbal Learning Test (RAVLT): an auditory-verbal memory test where high score indicates high rate of forgetting [8]
2. Digit Symbol Substitution Test (DSST): a symbol matching test where low score means impaired coginitive skill [9]
3. Functional Activities Questionnaire (FAQ): a questionnaire for an informant about patient's ability to do daily tasks alone in which high score means being dependent [10]

In the data cleaning process, some negative RAVLT scores are found and the records are removed. Moreover, the unit of height measurement is not consistent across the population noted from clear bimodal distribution. The summary statistics of some key variables are reported in Table 1 below.

Table 1: Summary Statistics of a Subset of Candidate Predictors

	DX: CN (N = 217)	DX: MCI (N = 385)	DX: Dementia (N = 181)
Age			
min	59.9	54.4	55.1
max	89.6	89.3	90.9
mean (sd)	75.98 ± 5.10	74.75 ± 7.36	75.32 ± 7.40
MMSE			
min	25	23	20
max	30	30	28
mean (sd)	29.10 ± 0.99	27.03 ± 1.79	23.40 ± 2.03
RAVLT			
min	0	0	0
max	100	100	100
mean (sd)	36.72 ± 24.87	68.77 ± 30.27	88.32 ± 22.24
DSST			
min	18	5	0
max	80	69	62
mean (sd)	45.84 ± 10.14	36.85 ± 11.12	26.99 ± 12.78
FAQ			
min	0	0	0
max	6	21	30
mean (sd)	0.14 ± 0.61	3.83 ± 4.49	12.92 ± 6.82

For exploratory data analysis, we utilize boxplots (Figure 1) and density plots to examine the possibility of using the variables we have in predicting Alzheimer's disease phases. For RAVLT score, the higher score indicates less ability to recall words after a certain period of time, hence, a decline in cognitive ability. DSST score indicates how fast participants can match the symbols with digits; therefore, higher scores pose less probability of having cognitive problem.

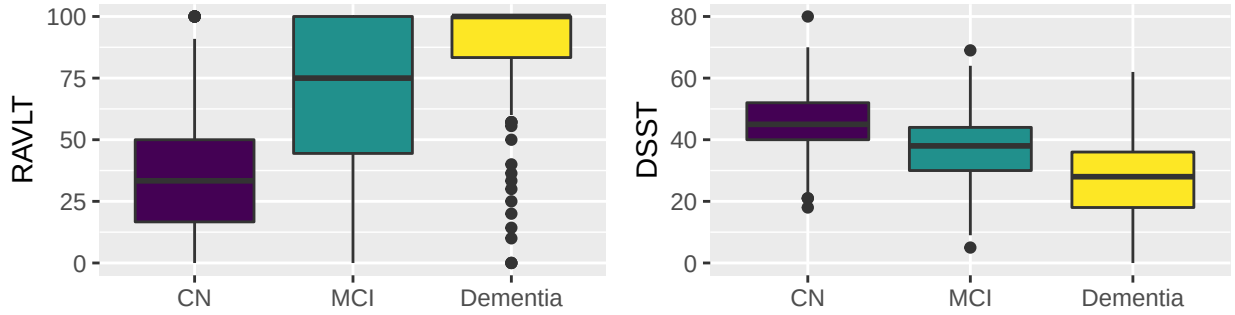

Figure 1: Boxplots showing the distributions of auditory score (RAVLT) and digit substitution score (DSST) of applicants with different Alzheimer's periods

Binned plots were used to study the probability of being in different phases for different predictor bins and are reported in the appendix (Figure 4). We found that the probability trend follows the sigmoid shape; moreover, the probability of being in CN and that of being in CN or MCI are similar in shape indicating that the Proportional odds model may be a suitable option.

Some other notices are the normality of Hippocampal Volume, one of the key biomarkers, including moderate linear association with some of our test scores. For other biomarkers such as Amyloid beta (ABETA), the amino acids causing plaques in brains, and phosphorylated tau (P-tau or PTAU), proteins which if defective can lead to neurons' death, they show weak linear relationship with predictors.

Model Part I: Alzheimer's Phase Prediction

The proportional odds model is used to predict Alzheimer's Phases in participants. We split the data into training and test set (75%:25%). All demographical data and three test variables (RAVLT, DSST, and FAQ) are included. Then, stepwise feature selection using BIC as the criteria is adopted. Finally, interaction terms and transformation are tested in term of the predictive capability.

According to model coefficients (after being exponentiated) in Table 2, the model only depends on three types of tests which are all significant at the 0.05 significance level.

Table 2: Coefficients of the Model to Predict Alzheimer's Phases

term	estimate	std.error	statistic	coefficient_type
DSST	0.954	0.010	-4.859	coefficient
RAVLT	1.024	0.004	6.291	coefficient
sqrt.FAQ	4.167	0.123	11.631	coefficient
CN|MCI	0.628	0.473	-0.982	zeta
MCI|Dementia	135.430	0.586	8.376	zeta

For the model assessment and evaluation, binned plots of the raw residuals versus the fitted probability on the test set is used. For most part, the binned residuals have no clear pattern.

To assess the generalizability, the AUC of the model on the test set is computed. The CN and Dementia cases can be almost perfectly separable (0.999). We also obtained reasonable AUC for separating CN from MCI (0.882) and MCI from Dementia (0.888). We compared our model AUC with that from other researchers that use the same data set and summarized all findings in the Table 3.

Table 3: AUC of classification models to predict different phases of Alzheimer's disease

Sources	Input Data	Techniques	CN/AD	CN/MCI
Fan Li(2018)[10]	MRI Image	ResNet	0.90	0.72
Fan Li(2018)[10]	MRI Image	DenseNet	0.92	0.78
Andres Ortiz(2018)[11]	MRI Image + PET	Ensemble of SVM	0.96	0.82
Bo Cheng(2018)[12]	MRI Image + CFS	SVM	0.98	0.87
Alden L.(2016)[13]	Cognitive tests + Biomarkers	Logistic Regression	0.98	0.88
This project	Cognitive tests	Proportional odds	0.99	0.88

Our model is comparable to those used similar techniques like logistic regression and better than a more complicated techniques like convolutional neural network. This is probably because the small size of the data set cannot accommodate the use of highly flexible models.

As the FAQ score is predictive, we further investigates the subscores of FAQ which target at ten different tasks in daily life such as the ability of applicants in preparing meals [14]. These subscores are incorporated into the modelling strategy and the stepwise feature selection using BIC is used to obtain the final model.

Out of the ten subscores, four of them are included in the final model and including the others doesn't further improve the performance. These tasks are completing forms, preparing meals, remembering appointments, and traveling out of the neighborhood. It is possible that these impairments are relatively easy to observe by relatives. The benefits here is to possibly reduce the number of questions in the FAQ or target at these areas while conducting interview.

Model Part II: Risk Score for Alzheimer's Disease

Unfortunately, rigorous classification modelling cannot be applied because of two reasons. Firstly, not all patients had a consistent follow-up. Secondly, pathologically, the conversion rate for normal people to MCI is

about 6% per year and ADNI has not followed people long enough [15]. Therefore, researchers mostly treated ADNI1 data as a cross-sectional study [13].

Despite these facts, exploring potential characteristics in the converted group is still possible for MCI turning to Dementia. We start by defining the risk score using coefficients from the first model. We posit a hypothesis that people who are relatively more likely to have Dementia and yet was still at MCI might be more likely to convert in the near future. The risk score is reported below.

$$riskscore = -0.047 * DSST + 0.023 * RAVLT + 1.427 * \sqrt{FAQ}$$

From Figure 2 below, investigating the risk score at baseline in people with MCI who turned into Dementia in one year (purple line) and who stayed as MCI (black line) clearly shows that the two groups have different average score. They also showed quite steeper increase in the risk level over time indicating worsen cognitive ability trend. By performing two sample t-test on these two groups, we can confirm the significant difference in the mean risk score at baseline and also the change in the score after one year (p-value < 0.05).

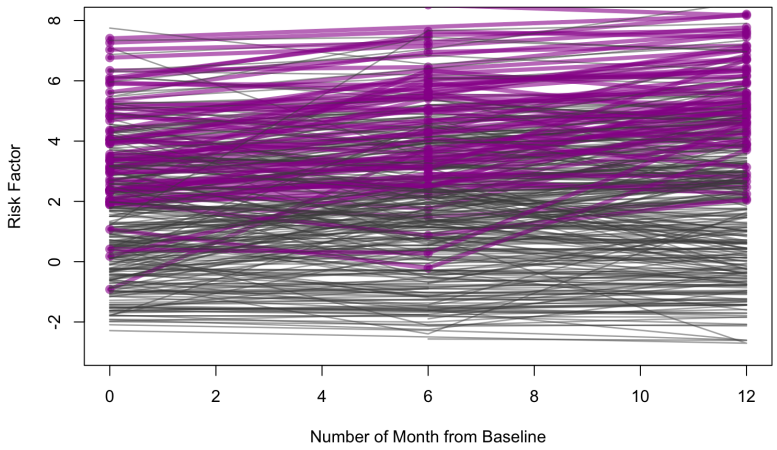

Figure 2: Risk Score in people with MCI who turned (purple) and not turned (black) into AD within 1 year

However, creating a model to predict this conversion results in AUC of only 0.65 in the test set. One reason is the limited data points of conversion events. Moreover, there were some other confounding factors not observed here such as the change in behaviors after being diagnosed, medical treatment, etc. Therefore, more longitudinal study in the more controlled environment is suggested to confirm this. The key takeaway here is that modelling risk score might be the key to the success of early intervention in high-risk patients.

Model Part III: Biomarkers Modelling

There are three biomarkers normally measured in the study of Alzheimer's disease: PTAU, ABETA, and hippocampal volume. The linear regression model considering transformations of test data and demographical information is utilized to capture them. However, as expected from weak linear correlation, we can explain less than 15% of the variability in PTAU and ABETA. One main reason is that these two biomarkers are generated early on in the pathogenesis when all cognitive abilities are still intact [16].

For Hippocampal volume, we obtain r-squared of 0.47 in test set. The model coefficients are reported in Table 4 below. DSST is found to be insignificant in predicting the volume. As expected, as people conditions worsened (higher RAVLT and FAQ), on average, Hippocampal volume will be lower from neuron cell deaths. Sex is a significant predictor as well. On average, male tends to have about 500 mm^3 more hippocampal volume than female of the same conditions. Finally, on average, an increase in one year of age corresponds to 54 mm^3 decrease in volume possibly due to aging.

Table 4: Coefficients of the Regression Model to Predict Hippocampal Volumes

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	11293.421	411.792	27.425	0	10484.569	12102.273
FAQ	-55.450	6.964	-7.963	0	-69.129	-41.772
RAVLT	-11.618	1.312	-8.853	0	-14.196	-9.040
AGE	-54.284	5.504	-9.862	0	-65.096	-43.472
PTGENDERMale	504.639	78.273	6.447	0	350.894	658.384

Diagnostic plots is reported in the appendix (Figure 5). All assumptions of linear regression are satisfied. Even though there is still considerable unexplained variability, we believe that some unobserved additional body measurements could have helped us improve the performance such as head circumference, etc. Hippocampal volume at baseline is also promising. With this information and with a better model, we may be able to reduce the number of times MRI must be administered during treatment.

To end this section, all results found here are consistent with the current understanding of the pathology. As the figure below suggests, PTAU and ABETA are developed first and, hence, are hard to capture with data of cognitive ability [17]. In contrast, the change in brain structure and memory function impairment are overlapped and, thereby, one is found to associate with the other.

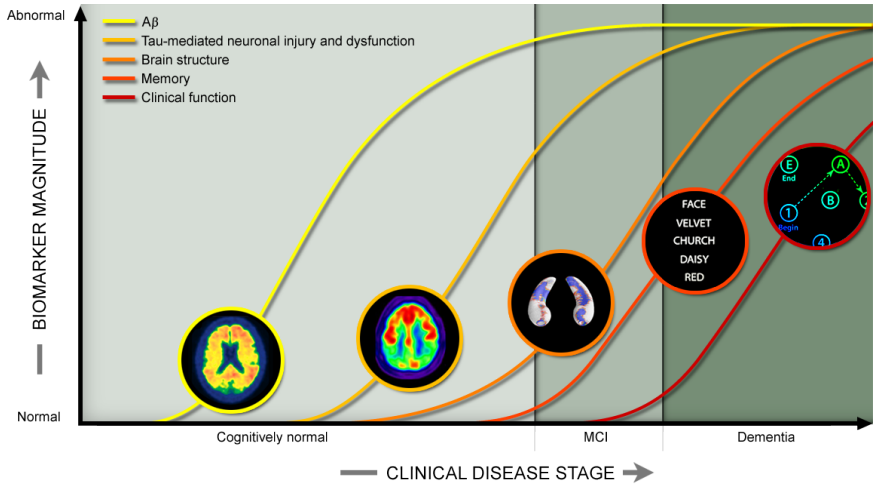

Figure 3: Alzheimer's Biomarkers Magnitude and Clinical Disease Stages

Conclusion

The proportional odds model has capability of classifying people with different Alzheimer's phases at a comparable accuracy with other alternatives in the research community. Key areas of FAQ are identified leading to possible improvement of the interview process. Regarding the future conversion, MCI cohort who turned to Dementia state in one year, had higher risk score compared with the others averagely. Lastly, the regression model on biomarkers can explain some uncertainty in Hippocampal volume.

In this study, the major limitation is that the cohort ADNI1 may be too homogenous demographically (race and age). Moreover, the applicants were not sampled from the true population but rather accepted via enrollment. As some cognitive ability measurements might be dependent on education, and family background, problems in generalization may arise. We also lack some detailed information that could lead to better result in biomarkers modelling such as some body measurements. These are key areas that could be improved in the future.

Appendix

References

[1] Pagel, M., Becker, J., & Coppel, D.B. (1985). Loss of control, self-blame, and depression: an investigation of spouse caregivers of Alzheimer's disease patients. Journal of abnormal psychology, 94 2, 169-82 .

[2] The National Institute on Aging's Alzheimer's and related Dementias Education and Referral (ADEAR) Center

[3] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild Cognitive Impairment: Clinical Characterization and Outcome. Arch Neurol. 1999;56(3):303–308.

[4] Petersen R. C. (2009). Early diagnosis of Alzheimer's disease: is MCI too late?. Current Alzheimer research, 6(4), 324–330.

[5] Vanderstichele, Hugo & Bibl, Mirko & Engelborghs, Sebastiaan & Le Bastard, Nathalie & Lewczuk, Piotr & Molinuevo, Jose & Parnetti, Lucilla & Perret-Liaudet, Armand & Shaw, Leslie & Teunissen, Charlotte & Wouters, Dirk & Blennow, Kaj. (2011). The clinical use of cerebrospinal fluid biomarker testing for Alzheimer's disease diagnosis: A consensus paper from the Alzheimer's Biomarkers Standardization Initiative. Alzheimer's & dementia : the journal of the Alzheimer's Association. 8. 65-73.

[6] Jack, C. R., Jr, Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., ... Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. The Lancet. Neurology, 9(1), 119–128.

[7] http://adni.loni.usc.edu/about/adni1/

[8] Matloubi S, Mohammadzadeh A, Jafari Z, Akbarzadeh Baghban A. Effect of background music on auditoryverbal memory performance. Audiology. 2014:0-.

[9] Patel, T., & Kurdi, M. S. (2015). A comparative study between oral melatonin and oral midazolam on preoperative anxiety, cognitive, and psychomotor functions. Journal of anaesthesiology, clinical pharmacology, 31(1), 37–43.

[10] Li, Fan & Liu, Manhua. (2018). Alzheimer's Disease Diagnosis Based on Multiple Cluster Dense Convolutional Networks. Computerized Medical Imaging and Graphics. 70. 10.1016/j.compmedimag.2018.09.009.

[11] Andres Ortiz, F. Lozano, Juan M. Górriz*, Javier Ramírez, Francisco J. Martínez Murcia and for the Alzheimer's Disease Neuroimaging Initiative, "Discriminative Sparse Features for Alzheimer's Disease Diagnosis Using Multimodal Image Data", Current Alzheimer Research (2018) 15: 67.

[12] Cheng, B., Liu, M., Suk, H. I., Shen, D., Zhang, D., & Alzheimer's Disease Neuroimaging Initiative (2015). Multimodal manifold-regularized transfer learning for MCI conversion prediction. Brain imaging and behavior, 9(4), 913–926.

[13] Gross, A. L., Mungas, D. M., Leoutsakos, J. S., Albert, M. S., & Jones, R. N. (2016). Alzheimer's disease severity, objectively determined and measured. Alzheimer's & dementia (Amsterdam, Netherlands), 4, 159–168.

[14] Pfeffer RI, Kurosaki TT, Harrah CH, Jr, et al. Measurement of functional activities in older adults in the community. J Gerontol. 1982;37:323–329.

[15] Chen, Y., Denny, K. G., Harvey, D., Farias, S. T., Mungas, D., DeCarli, C., & Beckett, L. (2017). Progression from normal cognition to mild cognitive impairment in a diverse clinic-based and community-based elderly cohort. Alzheimer's & dementia : the journal of the Alzheimer's Association, 13(4), 399–405.

[16] Murphy, M. P., & LeVine, H., 3rd (2010). Alzheimer's disease and the amyloid-beta peptide. Journal of Alzheimer's disease : JAD, 19(1), 311–323.

[17] http://adni.loni.usc.edu/study-design/#background-container

Table 5: Data Dictionary

Variable	Description
RID	Applicant ID
DX	Diagnosis (CN/MCI/AD)
AGE	Age in years
PTGENDER	Gender (Male/Female)
VSWEIGHT	Weight (pounds)
VSHEIGHT	Height (unit)
PTEDUCAT	Years of education
Partner	Partner is still alive, if any (1 - yes, 0 - no)
RAVLT	Rey Auditory Verbal Learning Test score
FAQ	Functional Activities Questionnaire average score
DSST	Digit Symbol Substitution Test score
FAQFINAN	Independency in keeping financial records
FAQFORM	Independency in completing business papers
FAQSHOP	Independency in shopping alone
FAQGAME	Independency in playing game of skill
FAQBEVG	Independency in heating water and turning off stove
FAQMEAL	Independency in preparing a balanced meal
FAQEVENT	Ability in keeping track of current events
FAQTV	Ability in understanding tv shows or books
FAQREM	Ability in remembering appointments
FAQTRAVL	Independency in traveling out of the neighborhood
sqrt.FAQ	Squared root of FAQ score
PTAU	phosphorylated tau content (unit)
ABETA	Amyloid beta content (unit)
Hippocampus	Hippocampal volume (mm^3)

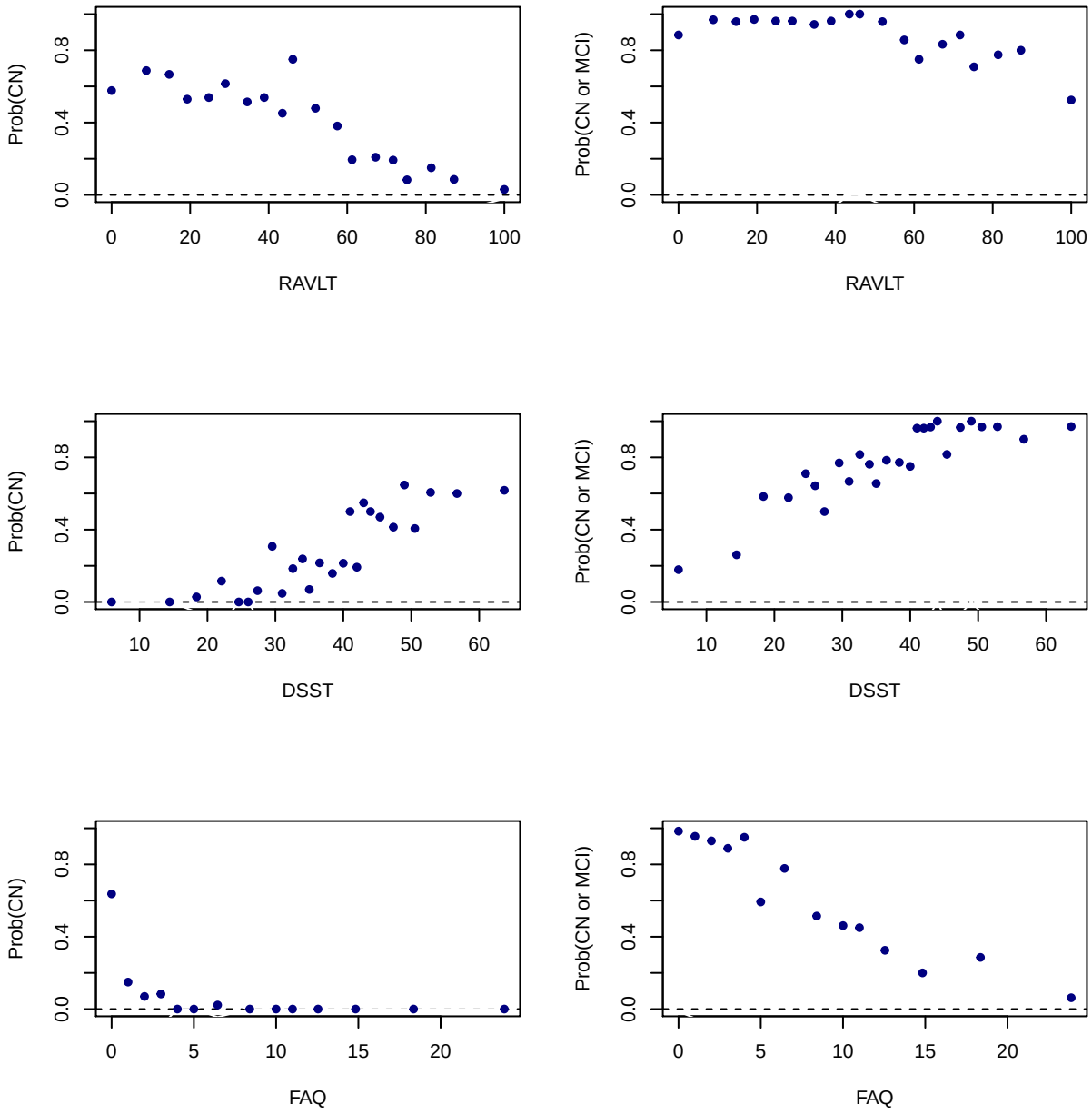

Figure 4: Binned plots showing the average probability of being normal (left) and being MCI at most (right) at different bins of cognitive test results

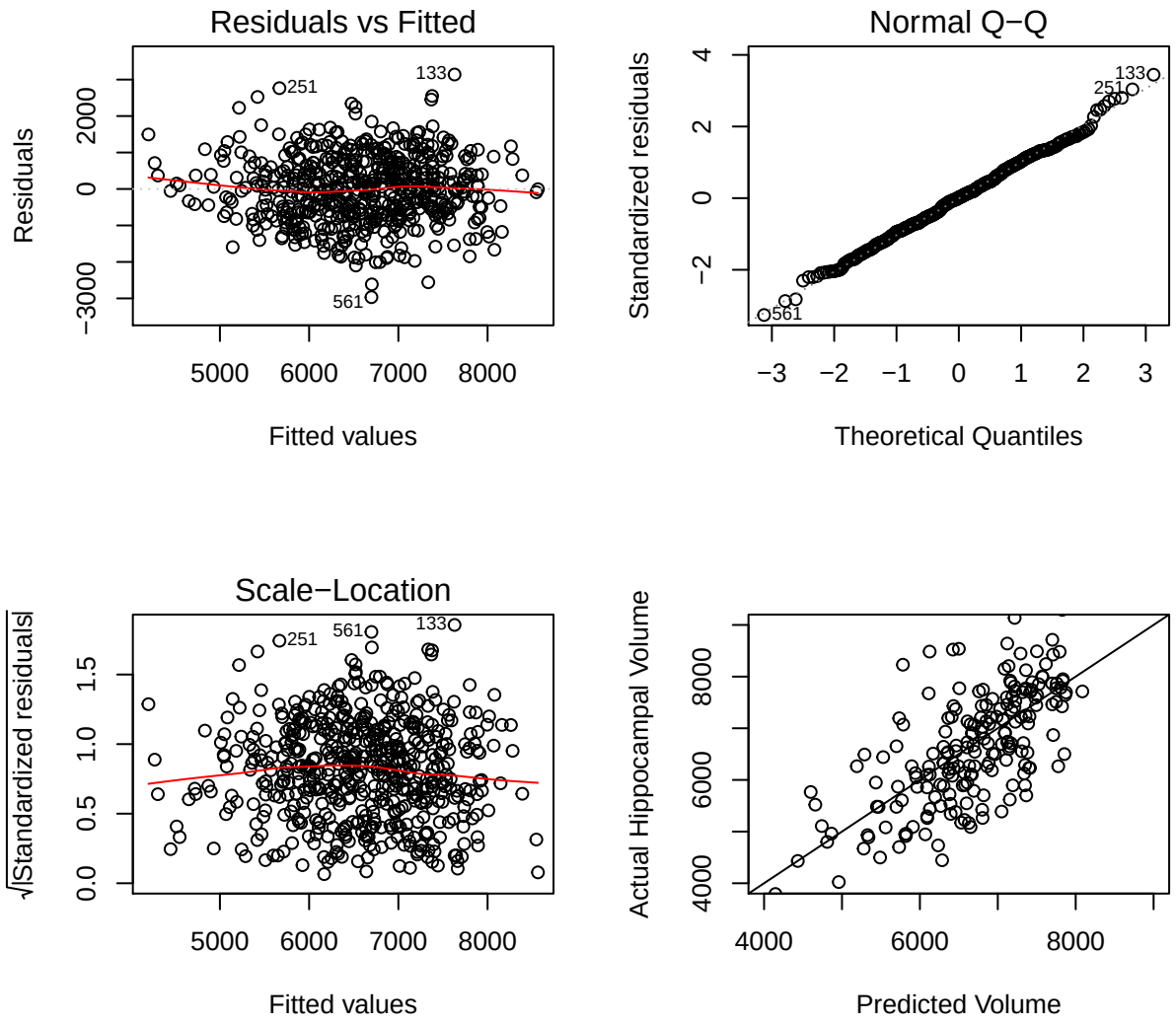

Figure 5: Diagnostic plots of the model to predict hippocampal volume in applicants of ADNI1. The last figure reports the fitting results in the test set