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

Executive Summary

In this project, we gave a comprehensive analysis on Alzheimer's disease data including patient demographical data and coginitive ability test results. A model to classify different Alzheimer's phases (Normal, Mild Cognitive Impairment, and Dementia) was created using proportional odds ratio technique. Risk score for the disease was proposed and shown that people who had Mild Cognitive Impairment and progressed to Dementia within a year tended to have higher risk score at baseline. Lastly, key predictors were used in regression to partially explain variability in Hippocampal volume, a key biomarkers of the disease, of patients at different stages.

Introduction

Formally introduced in 1910, Alzheimer's disease is a fatal disease that develops 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[2], our capability in accurately screening is still limited especially for Mild Cognitive Impairment (MCI), the early stage of the disease, resulting in late treatment.

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

1. To build statistical models based on simple neuropsychological tests that could possibly be administered online, and compare their predictive capability with other diagnostic techniques proposed within research community.
2. To explore the opportunity to use the probability score, obtained from the first part, as a risk factor in capturing the potential conversion to other degrees of Alzheimer's 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 development mechanism of Alzheimer's disease [6].

Data and Exploratory Data Analysis

The permission to access the data used in this project was approved by The Alzheimer's Disease Neuroimaging Initiative (ADNI). Specifically, the analyzed data set corresponds to ADNI1 cohort (797 participants). It includes basic demographic information of participants, longitudinal diagnostic results, vital signs, biomarkers, and Neuropsychological test results [7]. For the diagnosis, each person will be classified into one of the three groups: Control Normal (CN), Mild Cognitive Impairment (MCI), or Dementia. The full data dictionary was provided in the appendix section.

The neuropsychological tests explored in this project included:

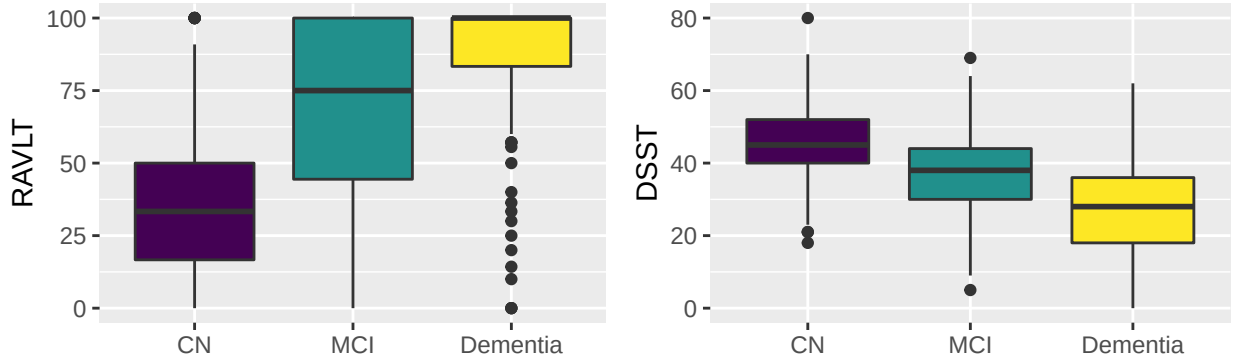
1. Rey Auditory Verbal Learning Test (RAVLT): a short-term auditory-verbal memory test where high score means impaired memory[8]
2. Digit Symbol Substitution Test (DSST): a symbol matching test where low score means impaired coginitive skill[9]
3. Functional Activities Questionnaire (FAQ): a questionnaire 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, one negative value of RAVLT score was found and the record was removed. Moreover, the unit of height measurement was not consistent across the population noted from clear bimodal distribution. Therefore, we avoided using such variables. We also found that not all paticipants had all information recorded at the screening process of the program; consequently, the data measured at baseline

once accepted into the ADNI program will be used. The summary statistics of some key variables are reported below.

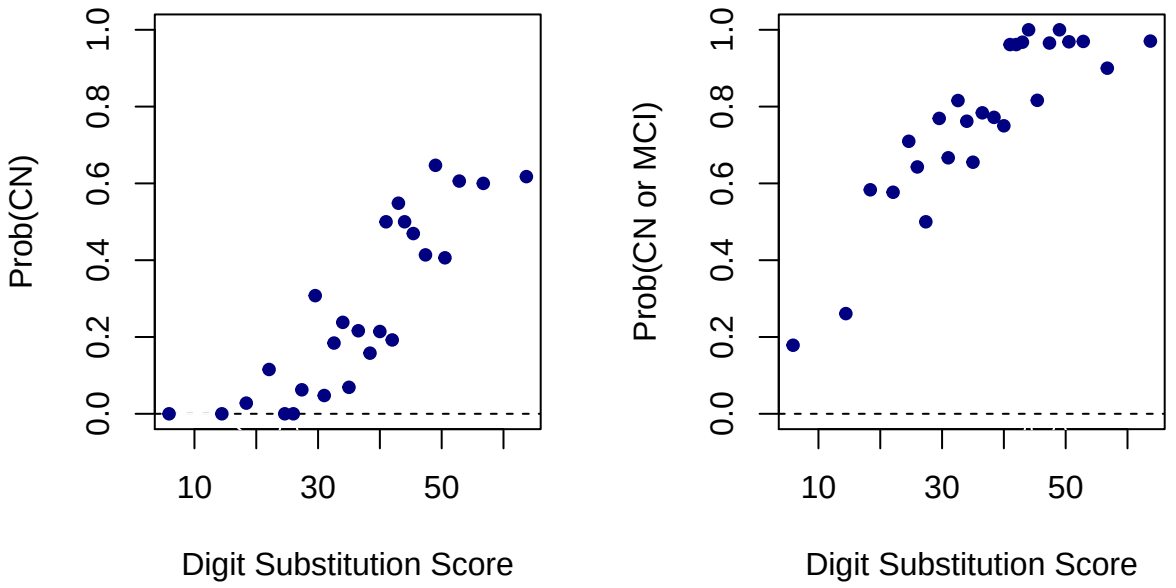
	DX: CN (N = 217)	DX: MCI (N = 385)	DX: Dementia (N = 181)
Age			
min	59.9	54.4	55.1
max	89.6	89.3	90.9
mean (sd)	75.98 ± 5.10	74.75 ± 7.36	75.32 ± 7.40
MMSE			
min	25	23	20
max	30	30	28
mean (sd)	29.10 ± 0.99	27.03 ± 1.79	23.40 ± 2.03
RAVLT			
min	0	0	0
max	100	100	100
mean (sd)	36.72 ± 24.87	68.77 ± 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 utilized boxplots and density plots to gauge the possibility of using the variables we had in predicting Alzheimer's phase of each person in ADNI cohort. Two examples are shown below. 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 and accurate participants can recognize symbols shown to them; therefore, the higher the score a person has, the less likely they are prone to cognitive impairment.



Binned plots were used to study the probability of being diagnosed to different phases for different predictor values. We found that the probability trend follows the sigmoid shape nicely; moreover, the probability of being in CN and that of being in CN or MCI (cumulative probability) are similar in shape with different baseline indicating that the Proportional odds model may be a suitable option. The examples are reported below.

Some other notices are the normality of Hippocampal Volume, one of the key biomarkers that we want to predict, including moderate linear association with some of our test scores (correaltion ranging from 0.3 to 0.5). For other biomarkers such as Amyloid beta (ABETA), the amino acids causing plaques in damaged brains, and phosphorylated tau (P-tau or PTAU), proteins responsible for stabilizing microtubules and if defective or mutated can lead to neurons' death, they show only weak linear relationship with our predictors.



Model

PART I: Alzheimer's Phase Prediction

The proportional odds model was used to predict Alzheimer's Phases in participants as the appropriateness has been shown in EDA process. We started the process by spliting the data into training and test set (75%:25%). To create the model, all demographic data and three test variables (RAVLT, DSST, and FAQ) were included. Then, stepwise feature selection using BIC as the criteria was adopted. Finally, interaction terms and transformation were tested in term of the predictive capability.

According to model coefficients (after being exponentiated) reported below, using BIC, the model only depends on three types of tests which are all significant at the 0.05 significance level.The fact that the model doesn't rely on demographic data like age, education, sex, and family information also increase the easiness of implementation.

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 was used. For most part, the binned residuals have no clear pattern.

After that, to assess the generalizability, the performance of the model in term of AUC on the test set was reported below. The CN and Dementia cases can be almost perfectly separable using our model. We also obtained reasonable AUC for separating CN from MCI (0.882) and MCI from Dementia (0.888) as well. We compared our performance with models from others in the academic community that use the same data set and summarized all findings in the table below.

TATATALTJTJLAT ——

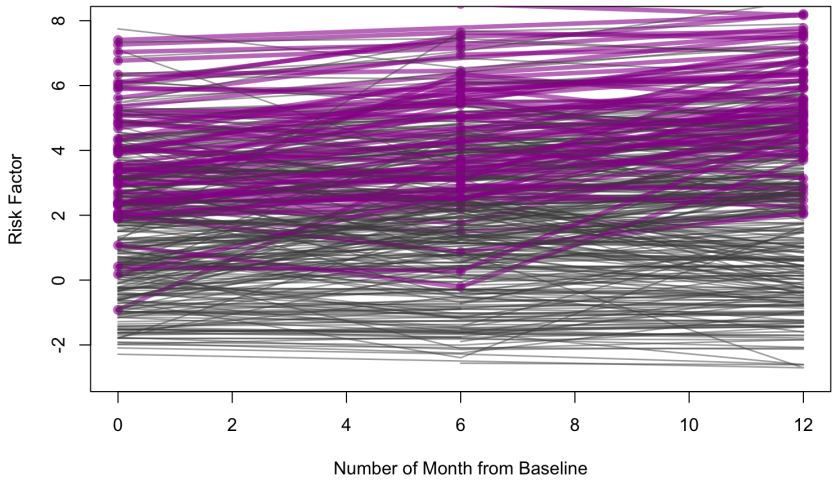
As stated earlier, the FAQ score showed potential in predicting the Alzheimer's phases of ADNI1 participants. But are all components of the test equally important? We addressed this issue by investigating the predictive ability of subscores of FAQ which targeted at ten different tasks in daily life such as the ability of applicants in preparing meals without help or whether they can go shopping alone [10]. These subscores were incorporated into the modelling strategy and the stepwise feature selection using BIC was used to obtain the final model.

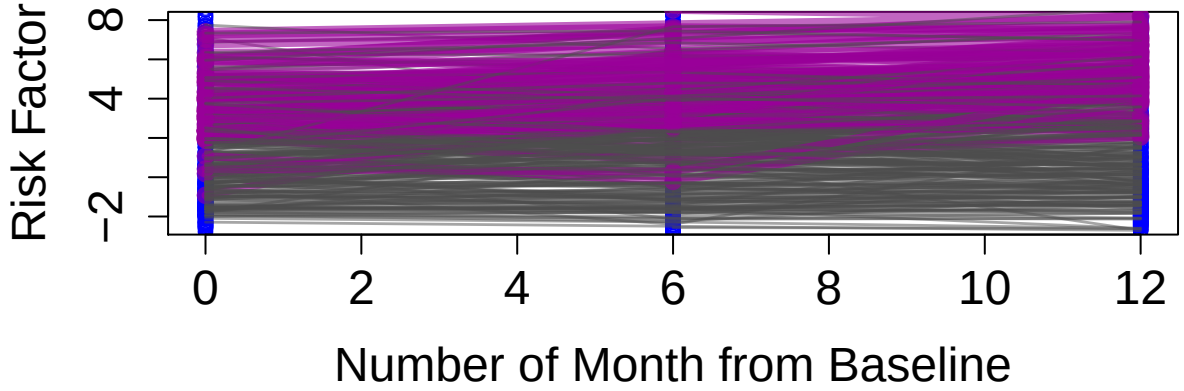
We found that, out of ten areas being observed, four are of particular important in predicting Alzheimer's disease risk. These are the ability to completing forms, preparing meals, remembering appointments, and traveling in the neighborhood alone. There are two main implications here. Possibly, the cognitive ability required to perform these tasks were damaged most from the disease; however, it can be that these tasks

posited a hypothesis that people who were more likely to have Dementia in the current time (described by the model in the previous part) and yet was still diagnosed as MCI might be more likely to turn into Dementia in the near future. The risk score is reported below.

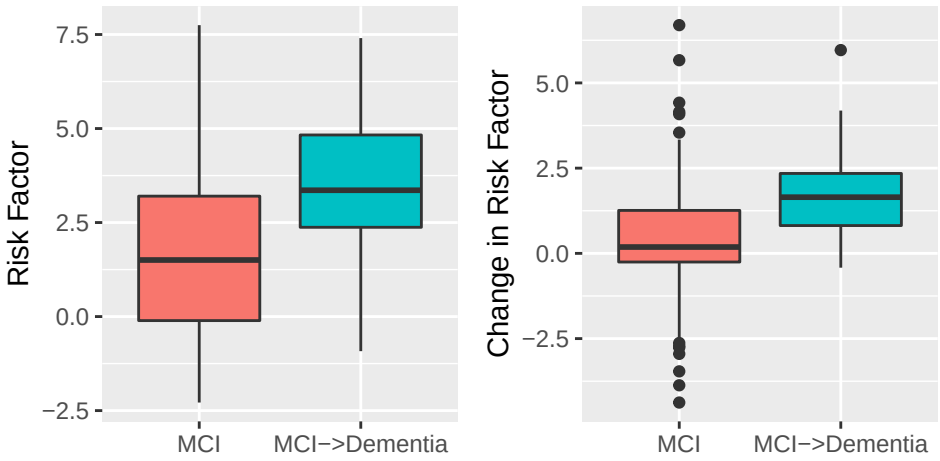
$$riskfactor = -0.047 * DSST + 0.023 * RAVLT + 1.427 * \sqrt{FAQ}$$

Investigating this risk score at baseline in people with MCI who turned into Dementia in one year (Pink line) and who stayed as MCI (Black line) clearly showed that the two groups have different distribution in the score. It was clear that people who progressed in the Alzheimer's development in a year did show higher risk score in the first place. 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 statistically significant difference in the mean risk score at baseline and also the change in the score after one year at 0.05 significance level.





However, creating a model to predict the conversion from MCI to Dementia in one year resulted in AUC of only 0.65 in the test set. Reasons for this are that we don't have enough data points of conversion events. Moreover, there were some other confounding factors such as the change in activities and diets after being diagnosed, medical treatment, etc. Therefore, more longitudinal study in the more controlled environment might be required. One conclusion we can draw here is that Alzheimer's disease phases are not black and white unlike cancer or tumor. The development has its spectrum and how fast the transition from one phase to the others occurs might not be completely random but rather associate with the current conditions. Understanding how different cognitive abilities play indicative roles in the transition might be the key to the success of early intervention in high-risk patients.



PHASE III: Biomarkers Modelling

There are three main biomarkers frequently measured in the study of Alzheimer's: PTAU, ABETA, and hippocampal volume. The linear regression model considering interaction terms and transformations of test data and demographical information was utilized. However, as expected from weak linear correlation observed during the EDA process, 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 [13].

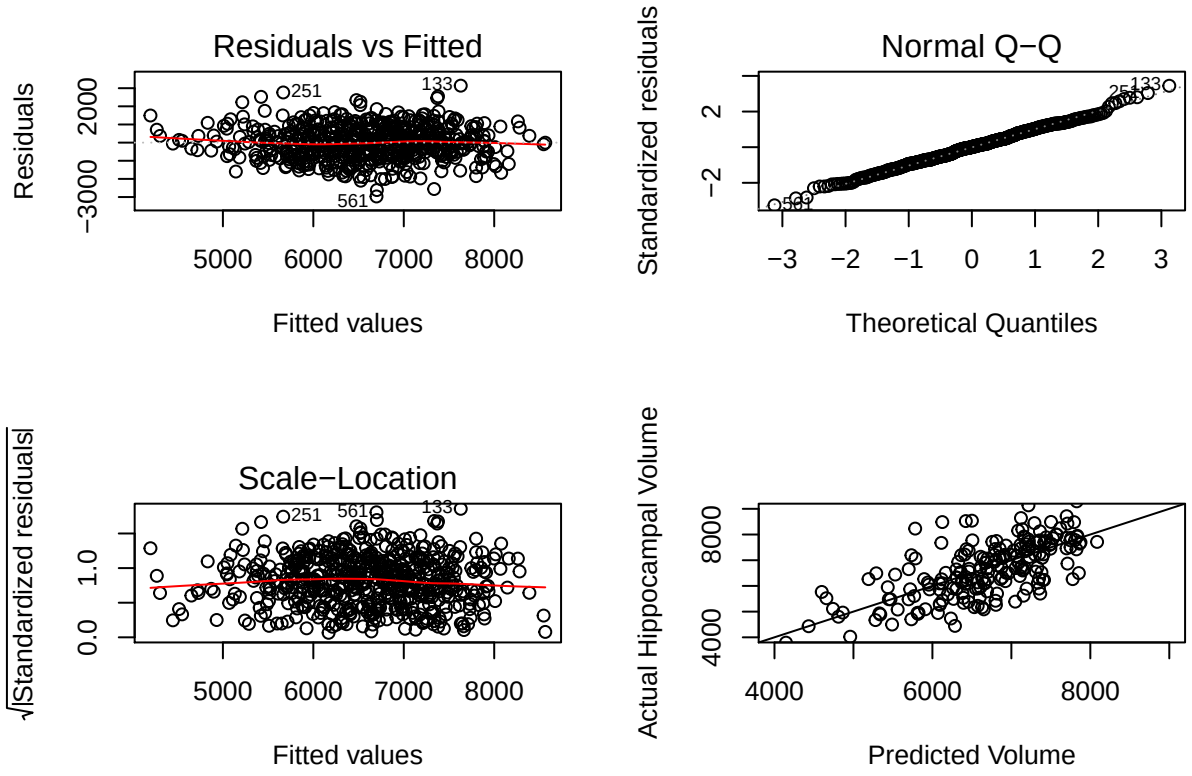
For Hippocampal volume, we obtained positive result as we obtained r-squared of 0.47 in test set. The model coefficients were reported below. DSST was found to be insignificant in predicting the volume and, hence, was removed from the model. As expected, as people conditions were worsened (higher RAVLT and FAQ), on average, Hippocampal volume will decrease due to cell deaths. Sex was found to be significant factor 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 would correspond to 54 mm^3 decrease in volume as a result of aging.

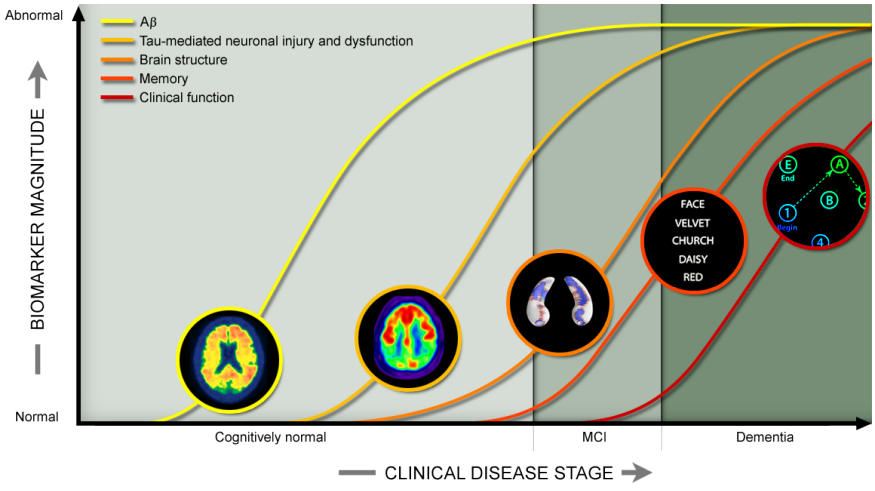
term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	11293.421	411.792	27.425	0	10484.569	12102.273
FAQ	-55.450	6.964	-7.963	0	-69.129	-41.772
RAVLT	-11.618	1.312	-8.853	0	-14.196	-9.040
AGE	-54.284	5.504	-9.862	0	-65.096	-43.472
PTGENDERMale	504.639	78.273	6.447	0	350.894	658.384

For model assessment figures shown below, all assumptions of linear regression were well satisfied and we have no multicollinearity problems. Notice that the last figure reported the fitting result in the test set. Even though there is still too much unexplained variability to make this model commercialized and applicable, the author believe that there are several body measurements missing from the data set that could have helped us improve the model performance such as body height, head circumference, etc. Hippocampal volume at baseline is also promising. If we are able to capture the change in the volume from the baseline, we may be able to reduce the number of times MRI was required during the treatment.

```
## integer(0)
```

To end this section, all results found here are consistent with the current understanding human has about Alzheimer's disease pathology. As the figure below suggests, PTAU and ABETA are developed first and, hence, are hard to capture with data of cognitive ability [14]. In contrast, the change in brain structure change and memory function impairment are largely overlapped in the pathology and, thereby, one is found to associate the other nicely.





\begin{center} \end{center}

Conclusion

A proportional odds model created showed capability of classifying people with different Alzheimer's phases at a comparable accuracy compared with other alternatives in the research community. Key areas of FAQ were identified leading to possible improvement of the tool. Regarding the future conversion, with the risk score we created, MCI cohort who turned to Dementia state within one year, on average, had higher risk score compared with the others and also exhibited higher change in the score within a year. Lastly, the regression model on biomarkers can explain a portion of uncertainty in Hippocampal volume of patients. The general association of cognitive impairment measured by different tests and different biomarkers was align with the pathology of the disease as well.

In this study, the major limitation was 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 and screening. As some cognitive ability measurements were known to be dependent, in term of baseline, on education, and family background, this might pose some problems in generalization. We also lacked some detailed information that could lead to better result in biomarkers modelling such as body measurements. These are key main areas that could be improved in the future.

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] 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. [11] 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. [12] 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. [13] 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. [14] http://adni.loni.usc.edu/study-design/#background-container ## Appendix

Data dictionary * * *