# Longitudinal Study of MRI and Alzheimer's Disease

An In-depth Generalized Linear Model Application

Yicheng Shen & Muyang Shi

3/8/2021

## Introduction

First described in clinical terms in 1906, Alzheimer's disease (AD), also known as senile dementia, is now the most common neurodegenerative disease with no effective cure, affecting more than 20 million cases worldwide (Goedert & Spillantini, 2006). There has been a general scholarly consensus that Alzheimer's disease a irreversible and progressive condition where neurons within the brain stop functioning, lose connection with other neurons and die (Mucke, 2009). It is the usual cause of dementia, a loss of brain function that can adversely impact memory, thinking, language, judgment and behavior.Unfortunately, there is no single test that precisely can determine whether a person has Alzheimer's disease, although an early diagnosis could be made by determining the presence of certain symptoms and ruling out other causes of dementia (Khachaturian, 1985). This involves a careful medical evaluation, including a thorough medical history, mental status testing, a physical and neurological exam, blood tests and brain imaging exams.

Magnetic resonance imaging (MRI), in particular, is currently perceived as the most accurate method that uses a powerful magnetic field, radio frequency pulses and a computer to produce detailed pictures of organs, soft tissues, bone and virtually all other internal body structures. It is capable of detecting brain abnormalities associated with mild cognitive impairment (MCI) and can be used to predict which patients with MCI may eventually develop Alzheimer's disease. In the early stages of Alzheimer's disease, an MRI scan of the brain may be normal. In later stages, MRI often shows a decrease in the size of different areas of the brain (mainly affecting the temporal and parietal lobes).

Existing research in the area of fighting against AD has already achieved substantial progress in understanding the what are the significant predictors of AD. The clinical experts have frequently employed longitudinal brain imaging in studying normal and diseased aging of brain. Previous studies have found that the volumes of the brain structures associated with memory tend to atrophy at a high rate in mild cognitive

impairment and early AD (Killiany et al., 2000). However, later research has also revealed the universality of brain atrophy within a greater population. Dr. Fotenos and his team, for example, have discovered that the rate of whole-brain volume decline evolves at a near 0.5% per year even in nondemented older adults (2005). Other scholars have focused on predicting the individual AD based on scores and evaluations from more traditional methods like cognitive assessment questionnaires and personal backgrounds (Bhagwat, et al., 2019). In recent years, more sophisticated machine learning methods have also been applied to classify AD with improved accuracy, with its validity requiring more examination (Fulton, et al., 2019).

Accurately diagnosing AD early could facilitate long term family planning and reduce costs from delaying long-term care. Unfortunately, the current academic absence of a general consensus in what are the significant factors in revealing the existence of AD has called for an in-depth exploratory research into this aspect. Funded by the Washington University Alzheimer's Disease Research Center, the Open Access Series of Imaging Studies (OASIS) is a program aimed at making MRI data sets of the brain freely available to the scientific community (Marcus et al., 2010). Our project, based upon a longitudinal MRI data set collected by OASIS for 150 individuals older than 60 years old, is going to examine whether MRI statistics are reliable indicators of patients' AD diagnosis. The results of our analysis, pointing out several significant indicators of AD-related dementia, have contributed extensively to the present field of research.

## Methods

Our proposed model built in the subsequent section is evaluated primarily based upon a raw data set obtained from the Open Access Series of Imaging Studies (OASIS). This data set consists of a collection of 373 observations for 150 volunteer patients aged from 60 to 98. The majority of these patients received MRI test more than once. The data set contains 62 male and 88 female subjects who are all right-handed. Each observation is the information collected from a single visit of one patient, with several times of visits grouped under each patient.

The data also includes age, education level (`EDUC`) and socio-economic status (`SES`) of the subjects. Moreover, some other medical statistics in the data set are gathered from specific MRI test results, including intracranial volumes (`eTIV`) and brain volumes (`nWBV`) of the subjects. The two groups of subjects are demented and non-demented in which the subject is diagnosed as AD or not, respectively.

The raw data set identifies subjects that develop the AD during the tests are grouped as converted. However, we are more interested in a subject's health condition at each visit. Therefore, we adjusted these converted subjects' conditions based on their clinical dementia rating (`CDR`) and classified them into either demented or nondemented status. The definition and range of each variable is listed in the table below.

Table 1: Variables in the OASIS Dataset

| Variable | Description | Range |
|---|---|---|
| Group | Mental Status (Response Variable) | Demented/Non-demented |
| CDR | Clinical dementia rating | 0, 0.5, 1 or 2 |
| Subject.ID | Unique Identifier for individual subjects | 1 - 150 |
| Visit | Number of visit of a subject | 1 - 5 |
| MR.Delay | Delay of visit by a subject since first visit (Number of days) | 0 - 2639 |
| Sex | Gender of the subject | Male/Female |
| Age | Age of a subject at the time of test | 60 - 98 |
| EDUC | Years of education of a subject | 6 - 23 |
| SES | Socio-economic status by Hollingshead's index of social position | 1 - 5 |
| MMSE | Mini Mental State Examination value by questionnaire | 4 - 30 |
| eTIV | Estimated total intracranial volume | 1106 - 2004 |
| nWBV | Normalized whole brain volume as percent of all voxels | 0.644 - 0.837 |
| ASF | Atlas Scale Factor; volume scaling factor for brain size | 0.876 - 1.587 |

The variables considered by this model fit properly as components of a hierarchical generalized linear model (GLM). We have recognized that there are two levels embedded in the data set. The first level is a subject's single MRI visit, with variables such as `MMSE`, `eTIV`, and `nWBV` at this level. The second level is a subject, containing individual characteristics such as `Sex`, `EDUC`, and `SES`.

# Results

The fields of neuroscience and cognitive biology are relatively obscure and unfamiliar for new researchers to fathom. Therefore, the exploratory data analysis (EDA) is a key procedure in previewing the breadth and complexity of our data set and potential models.

Table 2 below is a summary of subjects' basic demographic information. The distribution of `Age` is right-skewed, with over 75% of subjects under 80 years old. In terms of the education background, the majority of the subjects had received from 10 to 20 years of education, with highest being 23 years and the lowest being 6 years. The largest socio-economic status group is `SES` of 2, which indicates approximately upper-middle class of the society. More detailed distribution graphs of other individual variables could be found in our appendix.

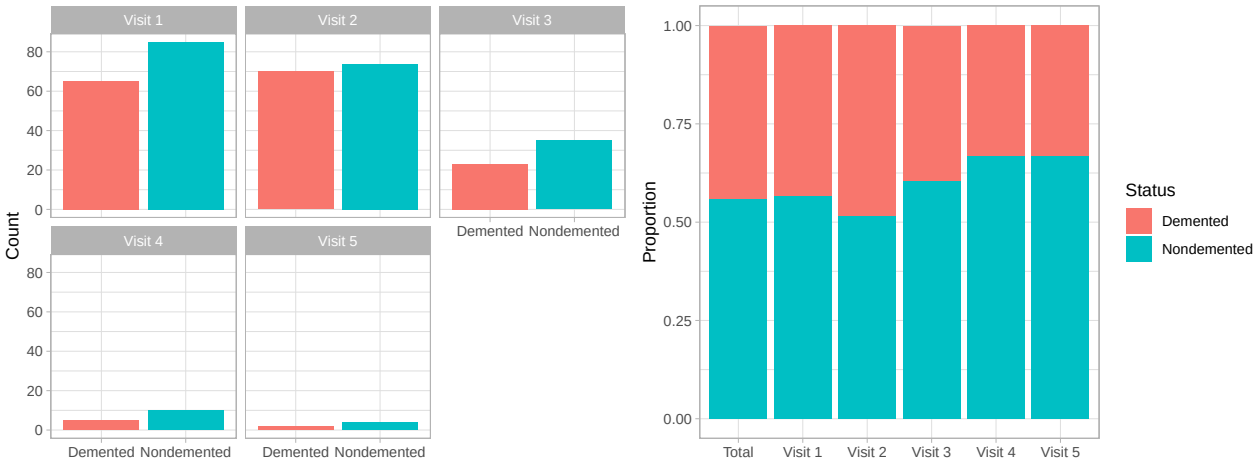
The main response in our study, denoted as `Group`, is a categorical variable with two levels, classifying subjects as either demented or nondemented based on their clinical dementia rating. Higher value of `CDR` would suggest more serious conditions of the demented patients. Each of 150 subjects attended their first MRI visit, with fewer of them coming back for more follow-up tests (only six of them finished all 5 MRI tests). The plot below suggests that, among the sample, there are more subjects tested as nondemented than

Table 2: Demographic Information of Subjects

| Age | Count | Sex | Count | Educ | Count | SES | Count |
|---|---|---|---|---|---|---|---|
| From 60 to 70 | 42 | Male | 62 | Within 10 Years | 5 | 1 | 33 |
| From 70 to 80 | 71 | Female | 88 | From 10 to 15 years | 78 | 2 | 42 |
| From 80 to 90 | 34 | | | Over 15 Years | 67 | 3 | 34 |
| Older than 90 | 3 | | | | | 4 | 30 |
| | | | | | | 5 | 3 |

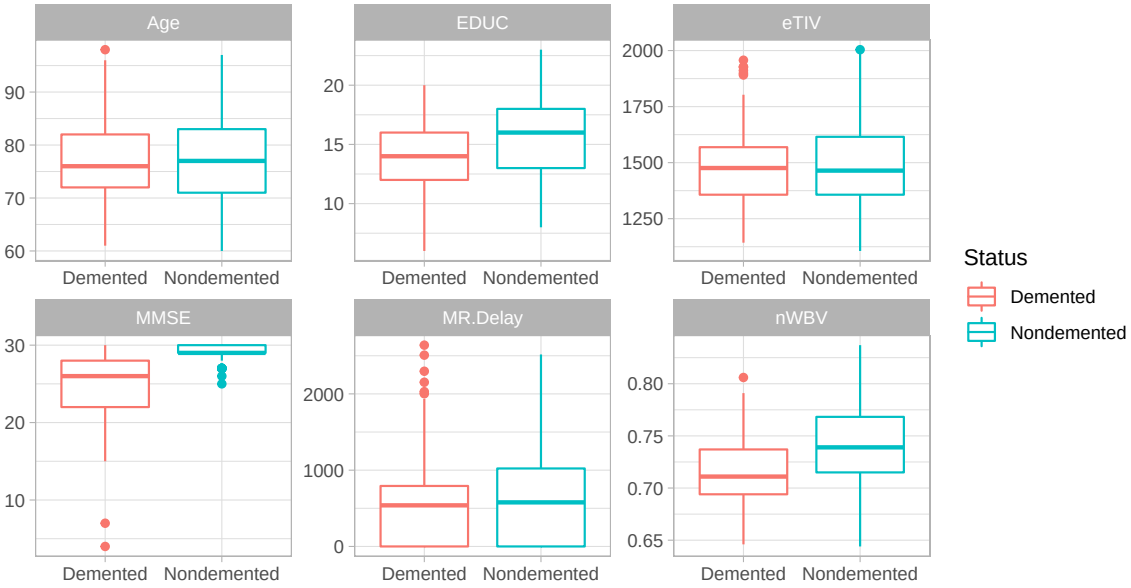
demented during each visit. Overall, there are 72 subjects who remained undemented throughout the whole study. 64 of the subjects were characterized as demented at the time of their initial visits and remained so for subsequent scans, including 51 individuals with mild to moderate Alzheimer's disease. Another 14 subjects were characterized as nondemented at the time of their initial visit and were subsequently characterized as demented at a later visit.

Figure 1. Subject's Dementia Status by Times of MRI Visit



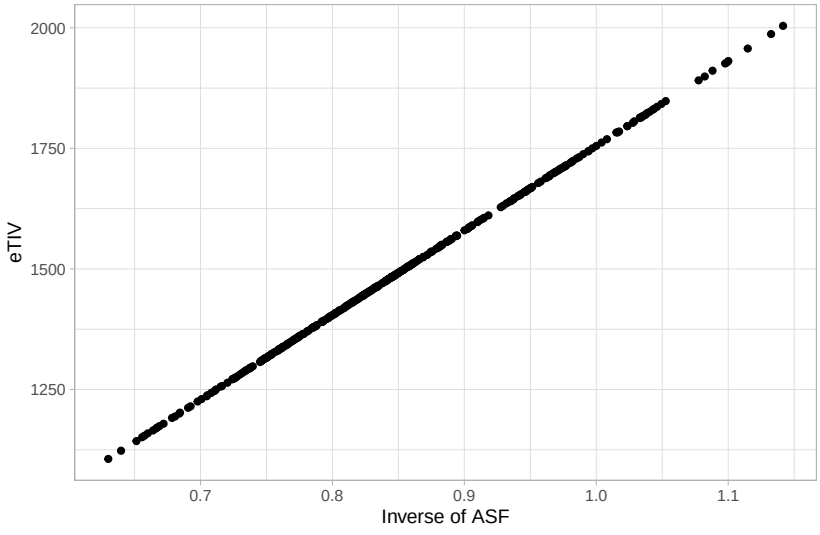

Next, we began to carefully examine each of the explanatory variables and its effects on dementia. The first series of variables that we selected are numeric variables. This category includes age of a subject at the time of test (`Age`), years of education of a subject (`EDUC`), days of delay since first visit (`MR.Delay`), mini mental state examination value by questionnaire (`MMSE`), estimated total intracranial volume (`eTIV`), and normalized whole brain volume as percent of all voxels (`nWBV`). Among this six variables, `MMSE` shows particularly strong potential association with the dementia status, with nondemented individuals scoring exceptionally higher in filling out the questionnaires than the demented group. Other variables with moderately strong effects are `EDUC` and `nWBV`, which suggests that nondemented subjects tend to have richer educational background and larger whole brain volumes. Interestingly, ages among the two groups do not seem to differ, with both groups approximately normally distributed across the whole range.

Figure 2. Main Numeric Explantory Variables of Interest



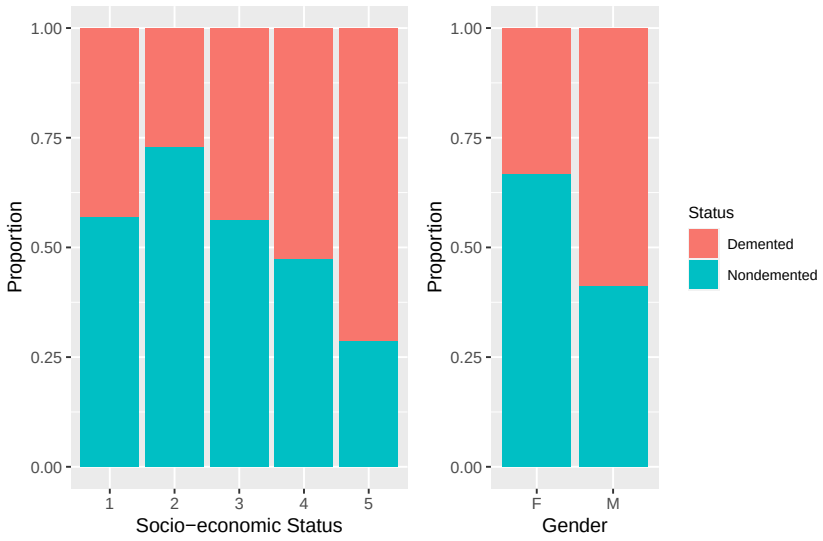

We also noticed an abonoramlity here that the Atlas Scale Factor, denoted as `ASF`, has a strong linear correlation with `eTIV`. Therefore, our subsequent model only incorporated `eTIV` to avoid modeling variables with repeated information. The clinical dementia rating, `CDR`, is also omitted in the model construction process because of its direct correlation with dementia status: The rationale of determining a demented subject is closely related to his or her `CDR` result. A `CDR` score of 0 will automatically decide the subject as nondemented, whereas a `CDR` score from 0.5 to 2 indicates mild to severe dementia conditions. Unfortunately, our data set has very few cases of subjects beyond mild dementia, therefore all non-zero `CDR` results are classified as dementia rather than a diversified record of dementia status.

Figure 3. Atlas Scale Factor & Estimated total intracranial volume



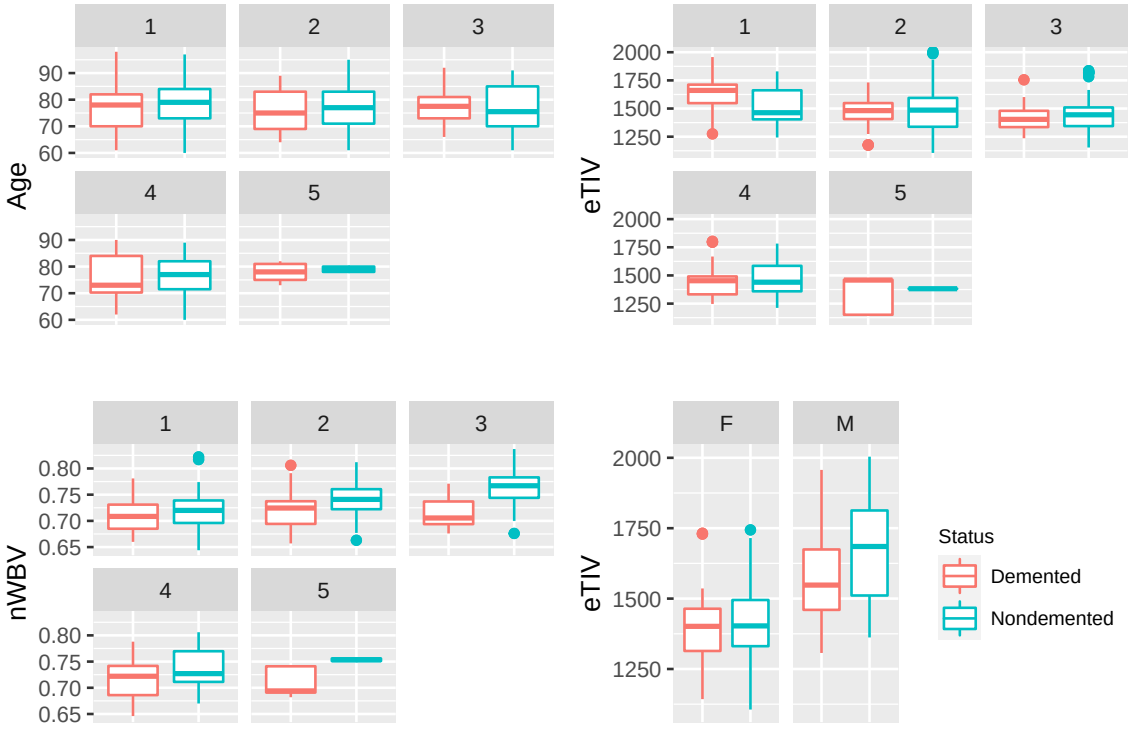
Two categorical variables in the data set are subject's gender (`Sex`) and socio-economic status (`SES`). Our EDA plots show that both of them are potentially significant indicator of dementia. For example, subjects with a higher `SES` score, which means unsatisfied socio-economic status, have higher proportion of suffering from dementia than their counterparts with lower SES. In terms of gender difference, the data suggests that males are faced up with a much higher risk of becoming dementia than females.

Figure 4. Categorical Variables




With eight initial variables taken into considerations, we also pondered upon the possible interactions between these variables. In this section, we mainly visualize the interactions between categorical and numeric variables. We identified several potentially strong interactions. The socio-economic status (`SES`) is particularly noteworthy in having strong interactions and affecting the influences of several other variables on the response. The sex of the subject, which already show disparity of dementia status between two gender groups, could also have interactions with estimated total intracranial volume (`eTIV`). In specific, we found that male subjects' `eTIV` is much larger than females, and the difference between demented and nondemented group is also wider for males. More interaction plots could be found in our appendix. These interactions would be candidates that receive further examination in the next section.

Figure 5. Interaction Terms




## Model Building Process

After the exploratory data analysis, we began to model the logarithmic odds of a subject being diagnosed as demented. The first model we fitted was a hierarchical generalized linear model version of the unconditional means model, with the first level being the visits (the tests) and the second level being the subject, and the `Group` variable being the response (a binary indicator of whether or not a patient is demented). We found the global intercept of this model to be -0.8733, which translates to an odd of 0.42 of being diagnosed as having Alzheimer Disease/demented. The patient-level random intercept explains a variance of 34.26 of the variability in the data.

We then could proceed with a hierarchical binary logistic regression with two levels (visit and patient/subject levels, same as the above), using the `Group` (the binary indicator) as response. From the exploratory data analysis, we would like to include the following explanatory variables into consideration when building the model: `MR.Delay` (delay of visit by a subject since the first visit), `MMSE` (Mini Mental State Examination value assessed through questionnaire), `eTIV` (estimated total intracranial volume), `nWBV` (normalized whole brain volume as percent of all voxels), `Sex` , `Age` , `EDUC` (years of education of a subject), `SES` (socio-economic status assessed by Hollingshead's index of social position), as well as the interactions between `Sex` and `Age` with `MMSE`,`eTIV`, `nWBV`, `EDUC`, `SES`, and the interactions between `SES` and `MMSE`, `Age`,

eTIV, nWBV, and EDUC, plus the random intercept at the subject level.

However, this model would not converge, even after we had it run for ten times more iterations. Therefore, we decided to build the hierarchical binary logistic model from the ground up – start with a simple model with variables of interests without the interaction terms (so as to find a model that would converge relatively easily), select the significant variables, and selectively add-in interaction terms that would keep the model converge and is/are significant through a forward selection process using Log-Likelihood Ratio Test (LRT).

To do so, we first found one hierarchical binary logistic model that converges, with variables MMSE (Mini Mental State Examination value assessed through questionnaire), eTIV (estimated total intracranial volume), nWBV (normalized whole brain volume as percent of all voxels), Age, and Sex. Then we went through a series of Likelihood Ratio Tests to determine the addition of interaction terms among those that we would like to include (interactions between Sex and Age, MMSE,eTIV, nWBV, and interactions between Age and MMSE, eTIV, and nWBV). In the end, we found the interaction between Age and MMSE to be significant and meaningful. Note that, to help the model to converge, we had re-scaled (normalized) the quantitative variables such as MMSE, eTIV, nWBV, and Age. Therefore, the hierarchical model formulation is:

Let $Y_{ij} = 1$ indicates the $i$th subject is diagnosed as demented in the $j$th visit, and 0 otherwise.

Let $\pi_{ij} = P(Y_{ij} = 1)$.

**Level 1**:

$$\log(\frac{\pi_{ij}}{1-\pi_{ij}}) = a_i + b \times MMSE + c \times eTIV + d \times nWBV + e \times Age + f \times MMSE : Age$$

**Level 2**:

$$a_i = a_0 + a_1 \times SexMale + u_i$$

where $u_i \sim N(0, \sigma_1^2)$.

The estimated coefficients are shown in Table 3 Estimation of Coefficients below.

Beside the above model (let's call it model1), we have also found another hierarchical binary regression model that converges, with variables MMSE (Mini Mental State Examination value assessed through questionnaire), nWBV (normalized whole brain volume as percent of all voxels), Age, Sex, and EDUC (years of education of a subject). Following a similar forward selection process using LRT, we still found the interaction term between Age and MMSE to be significant and meaningful. Similar to the above mentioned re-scaling (normalizing) process, we rescaled the quantitative variables to help with model convergence. Therefore, the hierarchical model formulation for this model 2 is:

**Level 1**:

$$\log(\frac{\pi_{ij}}{1-\pi_{ij}}) = a_i + b \times MMSE + c \times nWBV + d \times Age + e \times MMSE : Age$$

**Level 2**:

$$a_i = a_0 + a_1 \times SexMale + a_2 \times EDUC + u_i$$

where $u_i \sim N(0, \sigma_1^2)$.

Now we put estimated coefficients side by side with those from model 1 and created the table below.

## Estimates

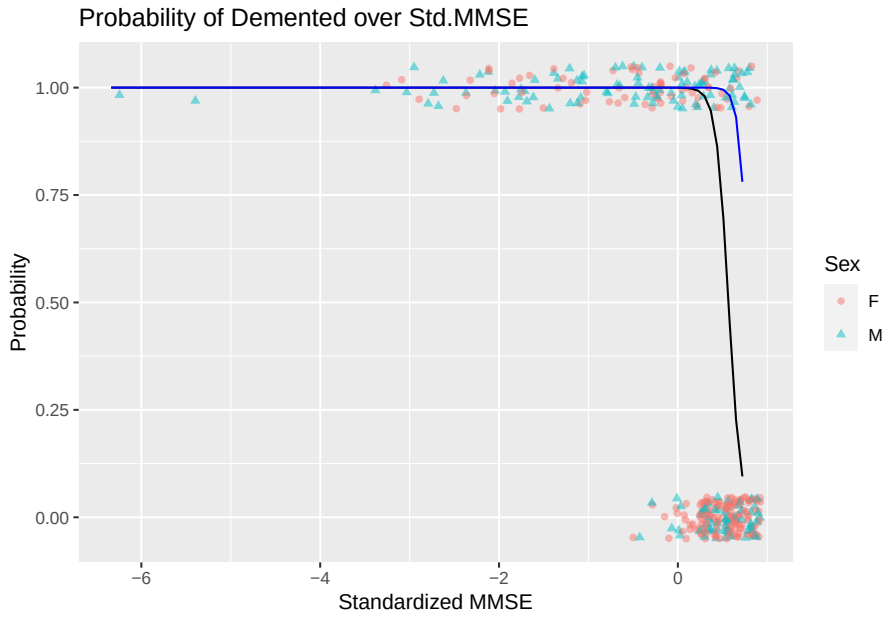
**Table 3. Estimation of Coefficients**

| Generalized Linear Mixed Effect Regression Results | | |
|---|---|---|
| | Dependent variable: | |
| | Dementia Stauts | |
| | (1) | (2) |
| MMSE | -14.504*** | -18.940*** |
| | (-22.785, -6.223) | (-28.740, -9.141) |
| eTIV | -4.190** | |
| | (-7.654, -0.726) | |
| nWBV | -8.402** | -10.958*** |
| | (-15.378, -1.425) | (-18.037, -3.879) |
| Age | -1.222 | -0.960 |
| | (-3.719, 1.275) | (-4.376, 2.457) |
| SexM | 12.363** | 16.498** |
| | (0.756, 23.971) | (2.196, 30.800) |
| EDUC | | -2.424** |
| | | (-4.464, -0.385) |
| MMSE:Age | -7.777*** | -10.848*** |
| | (-13.435, -2.119) | (-18.014, -3.682) |
| Constant | -4.150* | 33.824** |
| | (-8.925, 0.625) | (4.722, 62.926) |
| Observations | 371 | 371 |
| Log Likelihood | -99.522 | -99.233 |
| Akaike Inf. Crit. | 215.045 | 214.465 |
| Bayesian Inf. Crit. | 246.375 | 245.795 |
| Note: | *p<0.1; **p<0.05; ***p<0.01 | |

And for the random effect portion, $\sigma_1$ in model 1 is estimated to be 15.04, and $\sigma_1$ in model 2 is estimated to be 24.84. However, the second model with `EDUC` has a slightly better AIC (214 v.s 215). We kept both models because they have similar scores in model comparison tests such as AIC and BIC. Furthermore, model 1 focuses more the `Visit` level data (through the inclusion of the term `eTIV`), whereas model 2 puts

more weight on each subject by including `EDUC`, number of years of education a person has.

The figure 6 below shows the probability of being diagnosed as demented over (scaled/normalized) `MMSE`, with the other quantitative variables fixed at their mean values from the data set (observation), and set `Sex` to be Male. The black line represents the predicted probability estimated by the first model, and the blue line represents the predicted probability curve estimated by the second model.




# Discussion

Interpreting the estimated coefficients of the model can be straight-forward, and we use our first model to illustrate the process of interpreting the meaning of estimated coefficients of significant variables. Take `MMSE`, the Mini Mental State Examination value assessed through questionnaire, as an example. Every 3.68 increase in the score of the Mini Mental State Examination for a 77-year-old woman while holding the other variables constant is associated with a decrease in the odds of her being demented by a factor of approximately 0.0000005, with a 95% confidence interval of $(1 \times 10^{-10}, 0.002)$. In fact, the `MMSE` variable is actually one of the most important explanatory variable – for its big (in absolute value) estimated coefficient and small p-value.

Similarly, per 176 units of measure increase in `eTIV` (estimated total intracranial volume) is associated with a decrease in the odds of a subject being diagnosed as demented by a factor of roughly 0.015 with a 95% confidence interval of $(0.00047, 0.484)$, while holding the other variables constant. For a factor variable like `Sex`, being a man is associated with a increase of a factor of 233982 on the odds of being diagnosed demented

(with a 95% confidence interval of $(2.13, 2.6 \times 10^{10})$!

The other variables can be interpreted in a similar fashion. We note that for the variable `EDUC` (number of years of education) in model, we can interpret it as every 2.88 years increase in education would be associated with a decrease in the odds of being demented by a factor of 0.089 with a 95% confidence interval of (0.0115, 0.68).

Note the negative coefficients of the interaction term between `MMSE` and `Age` in both models: it indicates that the older a person gets, the stronger the association between a higher `MMSE` score and better cognitive conditions (being nondemented).

Comprehensively, take subject OAS2-0001 for example, at age 90, if he gets a similar performance in the evaluations and examinations as his last visit (an average `MMSE` score of 27, estimated total intracranial volume of 2004, normalized whole brain volume of 0.681, and 14 years of education, the estimated odds of him being demented predicted by model 1 would be 0.045, and that estimated by model 2 to be 0.14.

There are several limitations and inadequacy of our models that should be acknowledged. One of the biggest issues of applying hierarchical generalized linear models (GLM) on the OASIS data set is that the data does not fully satisfy the model independence assumption. We recognize that once a subject begins to suffer from dementia and is diagnosed as AD during one of the visits, it is highly likely that his or her subsequent diagnostic results will remain as a demented AD patient because it has been well-established that AD is an irreversible process and has no practical cure (Aditya, et al., 2017). Therefore, a patient's dementia status after a positive diagnosis is strongly correlated, thus violating GLM's key assumption which assumes independence among observations under the same group. The consequence of this violation implies that the variance of a true GLM model is larger than our current one, therefore the confidence intervals that we estimated above is narrower than the truth.

The potential solution would be to eliminate any subsequent observations once the subject is confirmed demented, thus keeping only independent cases. However, this would cause a substantial loss of available data for our analysis and causes insufficiency of our model. Ideally, we could also improve this study by incorporating a larger data set with more independent observations.

Furthermore, another direction for the future improvements of our model is to consider a zero-inflated Poisson model with `CDR` as the response variable. Currently the majority of the demented patients in the data set are diagnosed as mild or medium demented with their `CDR` as either 0.5 or 1. If OASIS could provide us with a larger sample with patients ranging from healthy to severe symptoms, with `CDR` as 0, 1, 2, 3... As a result, we could divide our data into two populations: The first has subjects who remain healthy and clear-minded, and the latter includes subjects who became demented. Under this circumstance, the zero-truncated Poisson model should be able to offer us with much more fine-grained results.

By exploring this MRI data set, our goal is to promote better assessments and future discoveries in basic and clinical Neuroscience. According to the structure of the data, we employed a two-level hierarchical generalzied linear model. Our finding points out to the significance of several factors and measurements that are closely related to severity of Alzheimer's disease. As the result of our study, we found that both individual backgrounds including age, gender and education, and MRI test statistics including estimated total intracranial volume and normalized whole brain volume as percent of all voxels could be considered as important factors when diagnosing AD-related dementia. However, the most significant and influential indicator when assessing a subject's dementia status is the result from Mini Mental State Examination value by questionnaire (MMSE). A higher MMSE score, especially for older aged people, is strongly associated with a better and functional mental and cognitive condition.

# Reference


Aditya, C. R., & Pande, M. S. (2017). Devising an interpretable calibrated scale to quantitatively assess the dementia stage of subjects with alzheimer's disease: A machine learning approach. Informatics in Medicine Unlocked, 6, 28-35.

Bhagwat, N., Pipitone, J., Voineskos, A. N., Chakravarty, M. M., & Alzheimer's Disease Neuroimaging Initiative. (2019). An artificial neural network model for clinical score prediction in Alzheimer disease using structural neuroimaging measures. Journal of psychiatry & neuroscience: JPN, 44(4), 246.

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. Neurology, 64(6), 1032-1039.

Fulton, L. V., Dolezel, D., Harrop, J., Yan, Y., & Fulton, C. P. (2019). Classification of Alzheimer's disease with and without imagery using gradient boosted machines and ResNet-50. Brain sciences, 9(9), 212.

Goedert, M., & Spillantini, M. G. (2006). A century of Alzheimer's disease. science, 314(5800), 777-781.

Khachaturian, Z. S. (1985). Diagnosis of Alzheimer's disease. Archives of neurology, 42(11), 1097-1105.

Killiany, R. J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., & Albert, M. S. (2000). Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society, 47(4), 430-439.

Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2010). Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. Journal of cognitive neuroscience, 22(12), 2677-2684.

Mucke, L. (2009). Alzheimer's disease. Nature, 461(7266), 895-897.