SBRMIC008_EDA

Michael Seebregts

2025-03-08

1: Introduction

Alzheimer's disease, which accounts for an estimated 60% - 80% of dementia cases, is particularly significant for people 65 and over (Alzheimer's Association, 2024). The US census bureau approximated that 910 000 people aged 65 or older developed Alzheimer's in 2011 with an incidence rate of 0.4% for people aged 65-74, 3.2% for people aged 75 - 84 and 7.6% for people aged 86 and older. These incidence rates are projected to double by 2050, due to the increasing number of people aged 65 in the US (Alzheimer's Association, 2024). Women are affected more by Alzheimers than men, with almost two-thirds of Americans with Alzheimers being women. Age is one of the greatest risk factors for Alzheimers, and women live on average longer than men. Studies have found that Alzheimers does not affect women more than men (Alzheimer's Association, 2024). Therefore the higher prevalence of Alzheimers is due to living longer.

With alzheimers being so prevalent, being able to accurately predict early signs of Alzheimer's will allow early prevention steps to be taken in order to potentially slow or stop the progression of this disease.

Alzheimers is a type of brain disease that is characterized by damage to the neurons in the brain, particularly those in the hippocampus which is responsible for thinking, talking and walking (Alzheimer's Association, 2024). Alzheimers is believed to begin over 20 years before the symptoms appear. As the disease progresses, there are significant changes to the brain. These changes include the accumulation of the protein beta-amyloid outside neurons, twisted strands of protein tau inside neurons, as well as inflammation and atrophy of brain tissue (Alzheimer's Association, 2024).

Structural MRI markers are a part of the clinical assessment in order to determine if someone has Alzheimer's (Frisoni GB, et al, 2010). These MRI markers provide people with more accurate and precise diagnosis, and allow for the progression of Alzheimers to be measured. According to current research, particularly powerful markers are the rates of whole brain and hippocampal atrophy (Frisoni GB, et al, 2010).

This paper explores whether multivariate techniques can be applied to Magnetic Image Resonance images on individuals in order to effectively predict Alzheimers. Based on the current research and information, the inflammation and atrophy of brain tissue may be recognisable by a sufficiently informed model.

2: Data Description

The data set consists of 400 observations with 9 predictors along with the 3 brain scans and 1 predictor, being clinical dementia rating. The patient information is as follows:

1. SESSION ID
2. AGE - age of patients
3. M/F - whether patient is male or female
4. HAND - dominant hand of the patient, in this case all patients are right handed
5. EDUC - education level of the patient, education level ranges from 1, no formal education, to 5 being a graduate or professional degree
6. SES - socio-economic status of the patient, SES ranges from 1, very low socioeconomic status, to 5, very high socioeconomic status
7. CDR - clinical dementia rating of the patient, the CDR scores have a range of 0, no dementia, 0.5, mild dementia, 1, moderate dementia and 2, severe dementia
8. MMSE - mini mental state examination (test scoring their memory, attention and problem solving). A MMSE score of 0-17 indicates severe cognitive impairment while 28 - 30 indicates normal cognitive function.
9. eTIV - estimated total volume of the skull
10. ASF - factor used to normalize brain measurements
11. nWBV - total brain volume, normalized for intracranial volume The brain scans themselves have been normalized already so that brain scans for the front all have the same pixels, with the same being done for the side and top scans.

The brain scans themselves have been normalized already so that brain scans for the front all have the same pixels, with the same being done for the side and top scans.

This data was collected by each subject being aged between 18 - 96. For each subject, 3 or 4 individual T1-weighted MRI scans were obtained in a single scan session. All the subjects were right handed and are a mixture of men and women. The data is currently hosted on WashU Medicine, titled OASIS1 Dataset (Marus, et al, 2010).

3: Analysis Approach

The response variable is the CDR rating which will either be 0, 0.5 or 1, after combining levels 1 and 2. The main predictor will be the MRI scans of the individuals, but other predictors such as age and MMSE score may be used as well. The MRI scans for the 58th observation is:
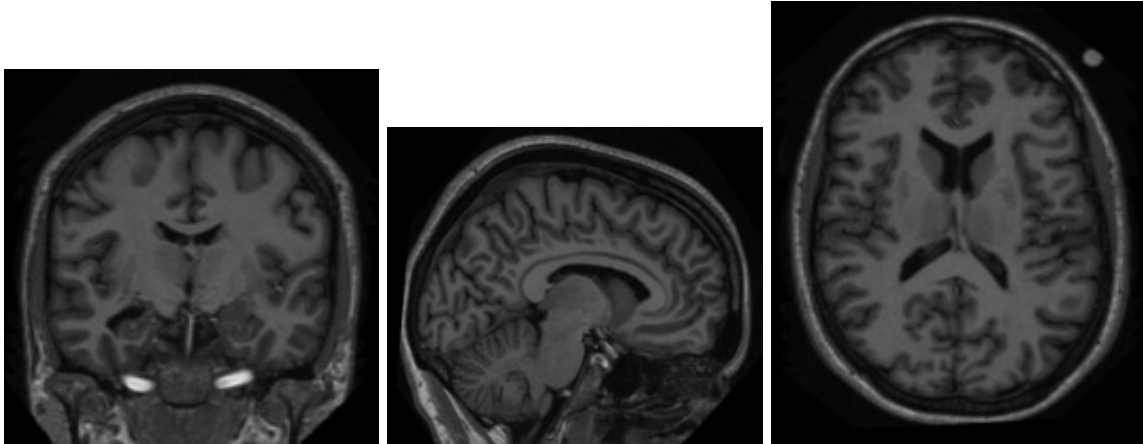

Figure 1: Front, Side and Top MRI Scans for Patient 58

Each individual has these same 3 MRI scans (Marus, et al, 2010).

Table 1: Table of number of indiviuals with different levels of CDR

CDR = NA	CDR = 0	CDR = 0.5	CDR = 1	CDR = 2
176	130	67	25	2

We can see that 176 of the patients do not have a CDR rating. These observations will need to be thrown out, leaving 224 observations.

For a CDR level of 2, there are only 2 observations in this category. This lack of data may not allow a model to effectively predict this level. CDR level of 2 will be therefore be combined with CDR level of 1.

Table 2: Table of the probability of indiviuals having different levels of CDR

CDR = 0	CDR = 0.5	CDR = 1
0.5803571	0.2991071	0.1205357

Can see that around 60% of observations have no Alzheimers, 30% have mild Alzheimers and 10% have moderate Alzheimers.

Table 3: Table of number of male and female patients

Males	Females
74	150

3

Table 4: Table of number of male and female patients with or without alzheimers diagnosis

Males with Alzheimers	Males without Alzheimers	Females with Alzheimers	Females without Alzheimers
38	36	56	94

two-thirds of the women have Alzheimers in some form while half of men have Alzheimers in some form. This is in line with what we expect, considering women live longer on average and are thus at greater risk of developing Alzheimers
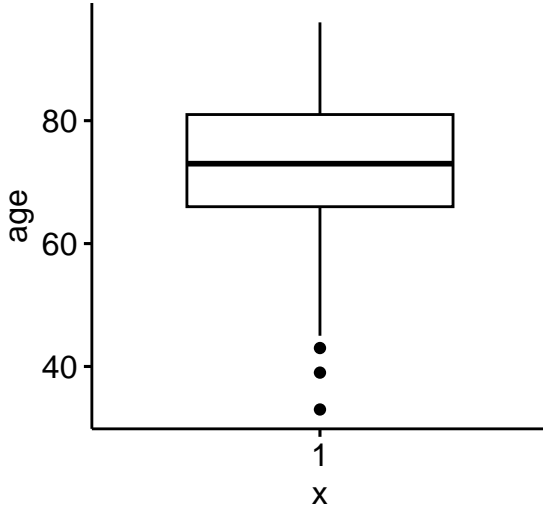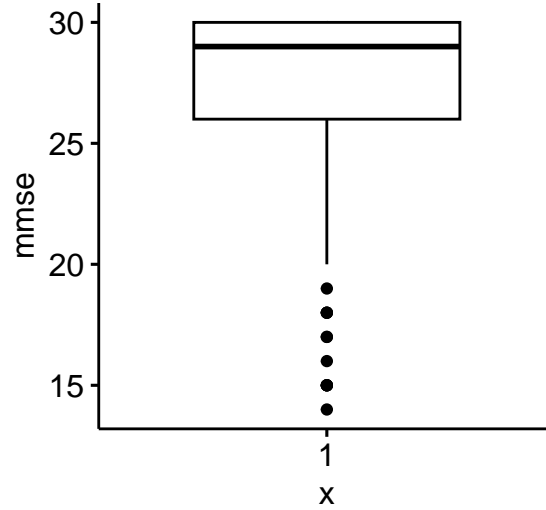

Figure 2: Boxplot of Age

Figure 3: Boxplot of MMSE scores

Can see that the average age is around 70, with most observations falling between high 60's and 80s. Also have a few young outliers. People aged 65+ are at the greatest risk of developing Alzheimers and thus have a high risk pool of observations.

The MMSE Scores are quite high, with only a few outliers falling in the 13-17 range which indicates severe cognitive impairment. The people falling between 20-25 also show not perfect cognitive abilities and may indicate some level of Dementia.

References

Alzheimer's Association. 2024. *2024 Alzheimer's disease facts and figures.* Available:https://www.alz.org/getmedia/76e51bb6-c003-4d84-8019-e0779d8c4e8d/alzheimers-facts-and-figures.pdf. [2024, 7 March]

Frisoni G., Fox N., Jack C Jr., Scheltens P., Thompson P. 2010. *The clinical use of structural MRI in Alzheimer disease.* Nat Rev Neurol. 6(2):67-77. doi: 10.1038/nrneurol.2009.215.

Marcus, D, Wang T, Parker J., Csernasky J., Morris J., Buckner R.2007. *Open Accesss Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young Middle Aged, Nondemented, and Demented Older Adults.* Journal of Cognitive Neuroscience, 19, 1498 - 1507. doi: 10.1162/jocn.2007.19.9.1498.