Statistical Theory Project: Analysis of Dementia Factors

Zehavit Barda - 212074223
Lea Meiraz - 208848150

September 2024

Abstract

Dementia is a neurodegenerative condition marked by cognitive decline, making early diagnosis crucial. This study analyzes various neuroanatomical and demographic features from the OASIS-2 dataset to identify markers associated with dementia. Among the features examined, normalized whole brain volume (nWBV) and estimated total intracranial volume (eTIV) showed significant associations with dementia diagnosis.

A two-sample t-test revealed a significant reduction in nWBV for demented subjects compared to nondemented ones (t-statistic = -3.34, p-value = 0.001). A Wilcoxon Signed-Rank test showed significant increases in eTIV between visits for demented individuals (p-value = 0.0009). Additionally, Spearman's correlation indicated a negative association between nWBV and dementia diagnosis (coefficient = -0.26, p-value = 0.0013).

While these findings suggest nWBV and eTIV as potential markers of dementia, the small sample size and limited follow-up duration restrict the generalizability of the results. Further research is needed to validate these findings in larger populations.

1 Introduction

Dementia is a growing concern, especially as the population ages. Understanding how brain structures change over time can provide valuable insights into the early detection of this condition. Longitudinal studies that track these changes are essential, as they allow researchers to observe patterns and identify risk factors for dementia.

Magnetic resonance imaging (MRI) is a key tool in studying brain changes. For example, a recent study [2] followed 653 healthy individuals and found significant age-related changes in brain volumes, including atrophy in key areas associated with cognitive function. This highlights the importance of long-term research to better understand the aging brain and its link to neurodegenerative diseases.

In addition to traditional analysis methods, machine learning approaches, such as support vector machines (SVM), have been employed using data from the Open Access Series of Imaging Studies (OASIS-2). Although one study [1] achieved an accuracy rate of approximately 68.75% in predicting dementia, the key takeaway is that the algorithm's ability to make predictions underscores the existence of a significant relationship between the selected features and the diagnosis of dementia. This finding reinforces the importance of exploring these relationships through statistical analysis in our research.

In this project, we will investigate the factors associated with dementia using the OASIS-2 dataset. Our goal is to analyze how various brain metrics and cognitive scores change over time and to assess their significance in predicting dementia. By applying robust statistical methods, we aim to provide insights that could improve early diagnosis and intervention strategies.

2 Results

2.1 Normalized Whole Brain Volume (nWBV) Results

Difference Between Groups (T-test)

Using a two-sample t-test, we found a significant difference in normalized whole brain volume (nWBV) between the demented and nondemented groups, with a t-statistic of -3.34 and a p-value of 0.001. This suggests that reduced nWBV is associated with dementia diagnosis.

Descriptive Statistics:

Table 1: Descriptive Statistics for nWBV by Group

Group	Mean nWBV	Standard Deviation
Demented	0.72	0.03
Nondemented	0.74	0.03

Although the statistical analysis demonstrates a significant difference in nWBV between the demented and nondemented groups, it is important to note that the actual difference in mean values is relatively small.

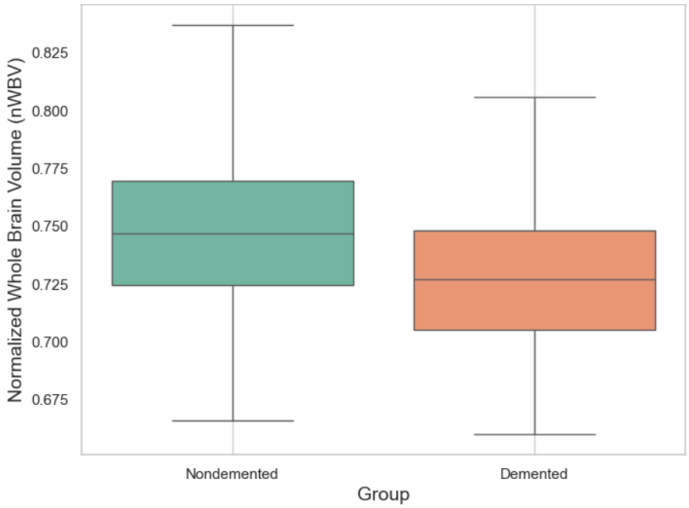

Figure 1: Boxplot of normalized whole brain volume (nWBV) for demented and nondemented groups. The boxplot illustrates the median, interquartile range, and potential outliers.

Spearman Correlation Analysis

We conducted a Spearman correlation analysis to assess the relationship between the features and dementia diagnosis. The analysis revealed a statistically significant negative correlation (Spearman correlation coefficient: -0.26, p-value: 0.0013), between normalized whole brain volume (nWBV) and dementia diagnosis. Even though the correlation coefficient is relatively low, these results are indicating that lower nWBV might be associated with dementia. This finding supports the results of our previous tests and suggests that nWBV could be a meaningful feature in understanding dementia diagnosis.

2.2 Estimated Total Intracranial Volume (eTIV) Results

Change Between Visits (Wilcoxon Signed-Rank Test)

To evaluate changes in estimated total intracranial volume (eTIV) between visits in demented and nondemented subjects, we performed a Wilcoxon Signed-Rank Test. The results indicated a statistically significant increase in eTIV from the first visit to the second visit in demented individuals, with a test statistic of 797.5 and a p-value of 0.0009. In the nondemented group there was no indication of a significant change in eTIV across visits.

Descriptive Statistics:

Table 2: Descriptive Statistics for eTIV by Visit in Demented Subjects

Visit	Mean eTIV	Standard Deviation
Visit 1	1473.05	164.43
Visit 2	1485.09	166.65

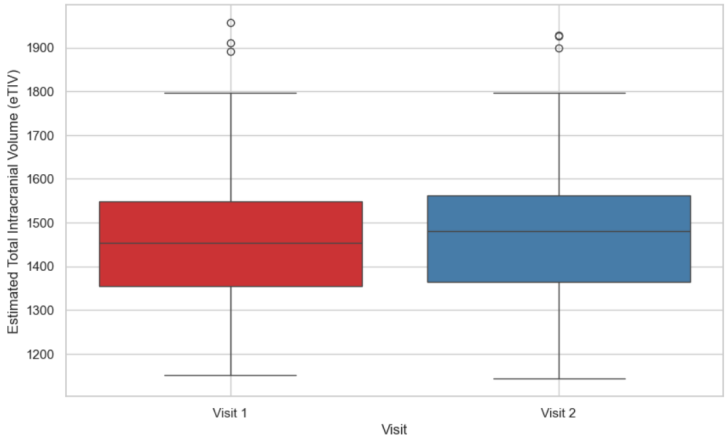

Figure 2: Boxplot of estimated total intracranial volume (eTIV) for demented subjects at Visits 1 and 2. The plot illustrates the median, interquartile range, and potential outliers.

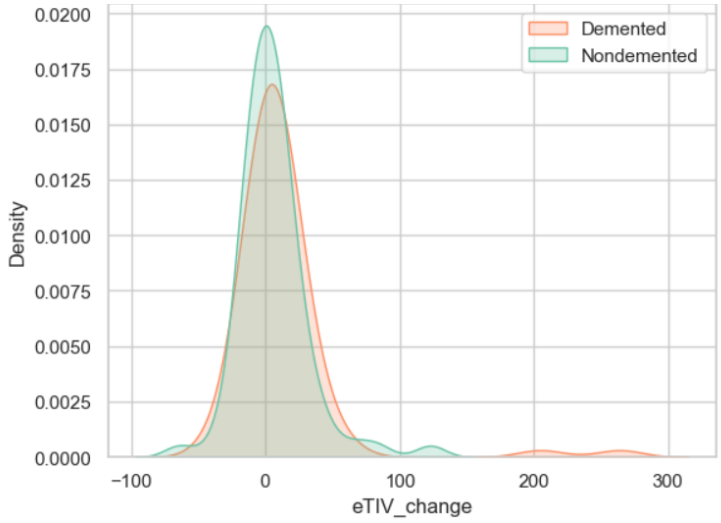

Figure 3: Kernel Density Estimate (KDE) of the changes in estimated total intracranial volume (eTIV) for both demented and nondemented groups..

3 Methods

3.1 Data Collection

The dataset used in this study was sourced from the OASIS-2 dataset, consisting of data for 150 subjects (373 samples), aged 60-95, categorized into two groups: demented and non-demented individuals. Each subject had multiple visits recorded, capturing features including:

- Demographic information (e.g., age, gender)
- Cognitive scores (e.g., *Mini-Mental State Examination (MMSE)*, *Clinical Dementia Rating (CDR)*)
- Brain imaging metrics from MRI scans (e.g., *normalized whole brain volume (nWBV)*, *atlas scaling factor (ASF)*, *estimated total intracranial volume (eTIV)*)

The dataset includes subjects with up to five recorded visits; however, the majority of subjects had data available for only two visits. To maintain consistency in the analysis, the first two visits were selected for each subject. In cases where data for the second visit was unavailable, data from the third visit was used, provided the MRI delay (i.e., the number of days since the first visit) was within a reasonable range.

Fourteen subjects initially belonged to the "Converted" group, meaning they were not diagnosed with dementia at the beginning of the study but received a dementia diagnosis in subsequent visits. For the purpose of analysis, these subjects were reclassified into the Demented group to reflect their eventual diagnosis. This reclassification allowed for a more consistent focus on dementia progression rather than the transition phase, providing clearer insights into the factors associated with dementia development.

3.2 Statistical tests

3.2.1 Kolmogorov-Smirnov Test

In this study, the KS test was used to evaluate the normality of features across groups.

Mathematical Definition The test compares the empirical cumulative distribution function (ECDF) of the sample, $F_n(x)$, with the cumulative distribution function (CDF) of the reference distribution, $F(x)$. The test statistic D is the maximum absolute difference between the two:

$$D = \sup_x |F_n(x) - F(x)|$$

A p-value is calculated to determine whether to reject the null hypothesis that the sample follows the reference distribution.

3.2.2 Levene's Test

We use Levene's test when performing parametric tests like the t-test, which assume homogeneity of variances between groups.

Mathematical Definition Levene's test evaluates whether the variance of a variable is equal across different groups by transforming the data and testing the null hypothesis that all group variances are equal. Let y_{ij} represent the value of the j-th observation in the i-th group, and $\bar{y}_i$ be the group median (or mean). The transformed variable is $z_{ij} = |y_{ij} - \bar{y}_i|$.

The test statistic W is given by:

$$W = \frac{(N-k)}{(k-1)} \cdot \frac{\sum_{i=1}^{k} N_i (Z_i - \bar{Z})^2}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - \bar{Z}_i)^2}$$

where k is the number of groups, N is the total number of observations, and Z represents the transformed data. A p-value is calculated to determine whether to reject the null hypothesis of equal variances.

3.2.3 Independent t-test

In this study, the t-test was used to compare the means of two independent groups, specifically between demented and non-demented subjects. The test assumes normality and equal variances between groups. We used a two-sided t-test, which checks for differences in both directions, testing whether the means are significantly different from each other, regardless of which group has a higher mean.

Mathematical Definition Given two groups with sample sizes n_1 and n_2, and sample means $\bar{x}_1$ and $\bar{x}_2$, the independent t-test evaluates whether the means of these groups differ significantly. The test statistic t is calculated as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where s_p^2 is the pooled variance, defined as:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Here, s_1^2 and s_2^2 are the sample variances for the two groups. The degrees of freedom for the test are $n_1 + n_2 - 2$.

The null hypothesis is that the population means are equal ($H_0 : \mu_1 = \mu_2$), and the two-sided test assesses whether the means of the two groups differ significantly, either direction. A p-value is calculated to determine whether to reject this hypothesis.

3.2.4 Mann-Whitney U Test

In this study, the Mann-Whitney U test was used as a non-parametric alternative to the t-test in cases where the assumption of normality did not hold. This test evaluates whether there is a significant difference between two independent groups based solely on the rank order of the data, without considering the actual magnitude of differences between values.

Mathematical Definition Given two independent samples, the Mann-Whitney U test ranks all the data points from both groups together and compares their ranks. The test statistic U is calculated as:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

$$U = \min(U_1, U_2)$$

where: n_1 and n_2 are the sample sizes of the two groups, R_1 and R_2 are the sum of the ranks for groups 1 and 2, respectively, U_1 and U_2 represent the U statistics for the two groups.

This test compares the rank order of the values, meaning that it only considers the relative positions of the data points and not the actual difference between values. The null hypothesis is that the two groups come from the same distribution ($H_0 : F_1(x) = F_2(x)$). A p-value is calculated to determine whether to reject this hypothesis.

3.2.5 Wilcoxon Signed-Rank Test

In this study, the Wilcoxon Signed-Rank Test was used to evaluate differences between paired observations, specifically focusing on changes across visits for the same subjects.

The null hypothesis (H_0) is that there is no difference in the paired observations between visits, while the alternative hypothesis (H_1) posits that there is a significant difference. The test evaluates the distribution of differences between paired samples, focusing on the ranks of these differences rather than their actual values, similar to the Mann-Whitney U Test.

To calculate the Wilcoxon statistic, the following steps are taken:

1. For each pair of observations, compute the difference $d_i = X_i - Y_i$ between the two values.

2. Rank the absolute values of these differences, ignoring any pairs where the difference is zero (if there is a tie, the mean rank is assigned).

3. Assign the sign of the difference d_i to the corresponding ranks.

4. Calculate the sum of the ranks for the positive and negative differences separately.
5. The test statistic is the smaller of these two sums.

A p-value is then calculated to determine whether to reject the null hypothesis.

3.2.6 Spearman Correlation

In this study, Spearman correlation was used to evaluate the strength and direction of association between ranked variables. Unlike Pearson correlation, which measures linear relationships, Spearman correlation assesses how well the relationship between two variables can be described using a monotonic function.

The Spearman correlation coefficient, denoted as ρ, is calculated by first ranking the data points for each variable, then applying the formula:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

where d_i represents the difference between the ranks of each pair of observations, and n is the number of observations. A coefficient close to +1 indicates a strong positive association, while a coefficient close to -1 indicates a strong negative association. A coefficient around 0 suggests no association.

Additionally, exploratory tests such as the Chi-square test were conducted to examine the association between categorical variables (e.g., demographic factors) and dementia diagnosis. However, these tests did not yield statistically significant results and are not further elaborated in the results section.

3.3 Multiple Comparison Correction: Dunn-Sidak

In this study, we performed multiple statistical tests. To account for the increased risk of Type I errors due to the repeated testing, we applied the Dunn-Sidak correction.

The correction adjusts the significance level α based on the number of independent tests performed. The adjusted significance level is calculated using the following formula:

$$\alpha_{\text{corrected}} = 1 - (1 - \alpha)^{\frac{1}{n}}$$

Where α is the desired significance level (in our case 0.05), and n is the number of independent tests.

4 Discussion

4.1 Conclusions

This study provides evidence for the association between normalized whole brain volume (nWBV), estimated total intracranial volume (eTIV), and dementia diagnosis. The analysis showed that demented individuals had significantly lower nWBV compared to nondemented subjects, as well as a significant change in eTIV between visits. These findings suggest that both metrics may play a role in understanding the progression of dementia.

While these results offer statistical insight, the practical significance of some findings, particularly the small differences in nWBV between groups, raises questions about their clinical utility. The Spearman correlation further supports the relationship between brain volume and dementia, yet more research is needed to establish these metrics as reliable diagnostic tools.

4.2 Limitations

Despite these significsnt results, the study is not without limitations. First, the relatively small sample size of 150 subjects may limit the generalizability of the findings. The lack of statistical power could prevent the detection of smaller but meaningful differences, and the results could be influenced by outliers.

Second, the limited longitudinal data, with most subjects only completing two visits, affects our ability to assess long-term trends. More extensive follow-up data would help in understanding how nWBV and eTIV evolve over time in relation to dementia progression.

Third, the issue of multiple comparisons presents a challenge. Despite applying the Dunn-Sidak correction, there is still a risk of false positives, which could affect the reliability of some significant results.

Additionally, the study focused on a narrow set of features, and it is likely that other factors, such as genetic markers or more detailed cognitive assessments, would provide a more comprehensive understanding of dementia. Expanding the feature set in future research would allow for a more holistic view of the condition.

Finally, the reliance on statistical tests alone limits the predictive power of the study. Integrating machine learning models or other advanced techniques in future research could uncover more complex patterns and improve the accuracy of dementia diagnosis.

Code Availability

The code for this project can be accessed at GitHub repository.

References

[1] Gopi Battineni, Nalini Chintalapudi, and Francesco Amenta. "Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)". In: *Informatics in Medicine Unlocked* 16 (2019), p. 100200.

[2] Shohei Fujita et al. "Characterization of brain volume changes in aging individuals with normal cognition using serial magnetic resonance imaging". In: *JAMA Network Open* 6.6 (2023), e2318153–e2318153.