Module_1:

Team Members:

Angelina Leung Esther Akinade

Project Title:

Assessing the correlation between the presence of the APOE4 gene and the severity of Alzheimer's disease measured using Thal score.

Project Goal:

We want to see if there is a correlation between APOE4 and Thal score. First, we want to determine which patients were diagnosed with Alzheimer's disease. Then, we want to separate them into two groups: one that has APOE4 present and one that does not. Next, we will determine the average Thal score in those two groups. We will assess the correlation between APOE4 and Thal score using statistical analysis with a t-test. Lastly, we will use linear regression to determine if there is a correlation between Thal score and age of death.

Disease Background:

- Prevalence & incidence
    - Alzheimer’s Disease has an incidence risk affecting 1 in 9 (or 6.5 million) Americans that are over 65 years old. AD has an annual prevalence of 10.7% in that same age group [1].
- Economic Burden
    - In 2021, costs associated with AD in the US was close to $340 billion [2]. These costs came from treatments, caregiving, and so on.
- Risk factors (genetic, lifestyle)
    - There are many contributing factors for Alzheimer’s. Risk factors can stem from a combination of genetics and lifestyle choices. One of the greatest factors is age. As a person grows older, their likelihood of getting Alzheimer’s

increases, especially after turning 65 years old [3]. Alzheimer's can also be tied to Down syndrome. It can also be related to the genotype of the apolipoprotein E (APOE) gene. This is a gene involved in transporting cholesterol through the bloodstream. There are different forms of APOE, but APOE4 tends to be found in patients with Alzheimer's [3]. Additionally, family history plays a role in risk of AD. As for lifestyle factors, poor exercise and diet habits can contribute to AD progression. Lower physical activity and an unhealthy diet with excessive alcohol and nicotine consumption can worsen people's condition [3].

- Societal determinants
    - Societal determinants can play a big role in the progression of Alzheimer's in a patient. There's a trend between patients with lower education levels having worse cognitive decline [4]. Access to health care is also a key factor. Alzheimer's is tied to health conditions such as diabetes and other chronic diseases. With limited health care access, these conditions go undetected or untreated, resulting in greater risk of AD. Even isolation can increase risk of premature death in patients [4].
- Symptoms
    - Common symptoms of AD include memory loss, difficulty planning and completing tasks, spatial disorientation, comprehension problems, trouble speaking and writing, poor judgment, and abnormal changes in behavior and mood [5].
- Diagnosis
    - There is no one set test that can tell a provider if a patient has Alzheimer's or not. A diagnosis will come from a variety of tests, usually starting with a typical health assessment that checks for medical history, vital signs, and lab work for blood and urine [6]. Doctors will also do neurological exams to assess brain function through reflexes and strength tests. Cognitive tests can also be used to evaluate a patient's memory and processing skills. Because isolation and depression can be a risk factor for Alzheimer's, mood assessments will also be done. Lastly, tests for blood,

cerebrospinal fluid, and brain imaging can reveal abnormal growths, fluid build-up, and tau and beta-amyloid levels that may indicate Alzheimer's disease [6].

- Standard of care treatments (& reimbursement)
    - Treatments generally either address the biological mechanisms involved in Alzheimer's or the cognitive symptoms. Two commonly prescribed drugs are lecanemab and donanemab. They both aim to reduce beta-amyloid in the brain, which is what the plaques associated with Alzheimer's are made of [1]. They are administered intravenously on a routine basis. Side effects include allergic reactions, headaches, and amyloid-related imaging abnormalities (ARIA). Both of these drugs are fairly new and their effects are still being monitored. Donepezil, rivastigmine, galantamine, memantine, and memantine and donepezil together, are used to address cognitive symptoms [1]. They have different effects on regulating neurotransmitters involved in Alzheimer's. Alternative drugs target other symptoms such as insomnia and agitation.
    - Additional treatments include exercising and exposure to art. These types of treatments aim to improve the patient's quality of life [1].
    - There are government programs and long-term care insurance plans that cover some Alzheimer's costs, but if a person doesn't already have a long-term care plan, they cannot apply for one after being diagnosed with Alzheimer's [7].
- Disease progression & prognosis
    - There are three stages of Alzheimer's disease. It starts mild, then progresses to moderate and severe. During the mild stage, a person is still able to carry out most of their normal functions [8]. They may have some trouble remembering small things, such as names and words. The moderate stage usually lasts the longest of the three stages. During this stage, a person may have more irrational behavior and need some assistance for tasks. Still, the person can carry out most normal functions with help. In the severe stage is where a person will require constant care. They may struggle to communicate and to carry out physical tasks.

They would not have much control over their minds and bodies [8].
    - After being diagnosed, a person could have a few months to plenty of years left to live [8]. The prognosis depends on many factors, such as the age when diagnosed, family history, if there are other diseases affecting the patient, and so on. In general, patients older than 65 tend to have four to eight years after being diagnosed [8].
- Continuum of care providers
    - Patients with Alzheimer's may require a large network of care providers. Often, family, friends, residential care providers can assist a person day-to-day [1]. There are also health care providers such as nurses, primary care physicians, neurologists, geriatricians, psychologists, and psychiatrists to help with the physical and mental tolls that can come with Alzheimer's [1].
- Biological mechanisms (anatomy, organ physiology, cell & molecular physiology)
    - The biological mechanisms primarily affected by Alzheimer's disease at the anatomical level are the hippocampus, cerebral cortex, amygdala, basal ganglia, and the frontal and parietal lobes. Damage to the hippocampus is typically one of the first symptoms of Alzheimer's. Degeneration to the cerebral cortex leads to impairment in reasoning, problem-solving, and language [9]. Additionally, the impairment of the basal ganglia causes difficulty in movement and coordination. At the cellular and molecular levels, Alzheimer's disease causes the loss of synapses and neurons because amyloid-beta plaques form between neurons, and tau tangles form within neurons [10].
- Clinical Trials/next-gen therapies
    - Next-gen therapies include drugs similar to lecanemab and donanemab, which aim to reduce amyloid beta deposits in the brain. Clinical trials are also being done with gene therapy, anti-inflammatory drugs, biomarkers, systolic blood pressure, and so on [11].

Data-Set:

The data was obtained from a study done by Gabitto et al., in October 2024 [12]. They used a system to organize different types of biological data related to Alzheimer's and related pathologies. These techniques are called "multiomics, spatial genomics, and reference atlases" [12]. Gabitto et al., also used quantitative neuropathology to categorize donors based on severity of their conditions [12]. There were a total of 84 donor subjects; 33 were male, and 51 were female. The metadata came from the University of Washington BioRepository and Integrated Neuropathology (BRaIN) Laboratory from the Kaiser Permanente Washington Health Research Institute ACT Study and University of Washington ADRC. A limitation with this data set is that there was no randomization [12]. These participants chose to be involved, and the study included them all.

The metadata csv gives the donor ID, the study the donor is from, age at death, sex, race, education level, years of education, APOE genotype, cognitive status (dementia/no dementia), age of onset cognitive symptoms, age of dementia diagnosis, head injury, if the patient has had neuroimaging, if the patient has had other cognitive disorders, CASI score, MMSE score, MOCA score, PMI, frozen tissue type, type of imaging, brain weight and pH, AD neuropathological change, Thal, Braak, CERAD, CAA, highest lewy body disease, total microinfarcts, atherosclerosis, RIN, and whether or not a patient is severely affected.

Data Analyis:

(Describe how you analyzed the data. This is where you should intersperse your Python code so that anyone reading this can run your code to perform the analysis that you did, generate your figures, etc.)

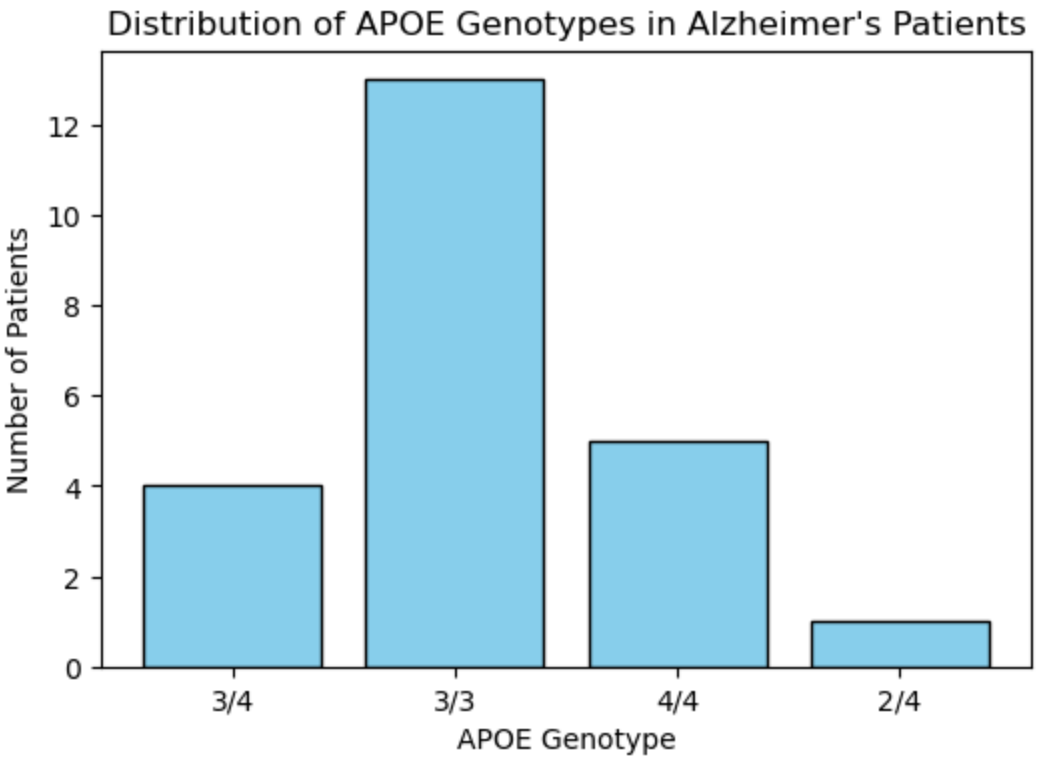
First, we decide to see the difference in APOE Genotype in the participants that has alzheimers. We also wanted to create a list only with participant with only alzhemiers with their APOE Genotype and Thal score. We were also able to do a Student-T-test that compares the thal score in donor with APOE4 Present and Not Present. The results are t-statistic: 4.411, p-value: 0.0002. This means that the difference between the thal score of donors with APOE4 Present and Not Present is statistically significant.

In [9]:
```
# this code was imported from ChatGPT, which helps separate all the patients t
import pandas as pd


df = pd.read_csv("NO DATE GENOTYPE Metadata.csv")
```

```python
id_col = "Donor ID"
dx_col = "Consensus Clinical Dx (choice=Alzheimers disease)"
thal_col = "Thal"
apoe_col = "APOE Genotype"


# Filter Alzheimer's patients from consensus clinical dx
alz_df = df[df[dx_col] == "Checked"]


# Select only ID, Thal, and APOE genotype
alz_selected = alz_df[[id_col, thal_col, apoe_col]]


# Convert to a list of tuples (ID, Thal, APOE)
alz_list = alz_selected.to_records(index=False).tolist()


print("\nList of Alzheimer's patients (ID, Thal, APOE):")
print(alz_list)


# this code is separating the different type of apoe genotype and counting how
from collections import Counter


apoe_list = [item[2] for item in alz_list]


# Count occurrences of each genotype
apoe_counts = Counter(apoe_list)


print("Number of subjects per APOE genotype (Alzheimer's patients):")
print(apoe_counts)
# this code is separating the different thal score genotype and counting how m


thal_scores = [thal for _, thal, _ in alz_list]


# Count frequency of each Thal score
thal_counts = Counter(thal_scores)


print("\nCounts of each Thal score among Alzheimer's patients:")
print(thal_counts)


# this is creating a graph of how many APOE genotype there are and how many al
```

```python
import matplotlib.pyplot as plt


apoe_genotypes = [apoe for _, _, apoe in alz_list]


# Count frequency of each genotype
apoe_counts = Counter(apoe_genotypes)


# Plot
plt.figure(figsize=(6,4))
plt.bar(apoe_counts.keys(), apoe_counts.values(), color="skyblue", edgecolor="
plt.xlabel("APOE Genotype")
plt.ylabel("Number of Patients")
plt.title("Distribution of APOE Genotypes in Alzheimer's Patients")
plt.show()
```

```
List of Alzheimer's patients (ID, Thal, APOE):
[('H20.33.004', 'Thal 5', '3/4'), ('H20.33.011', 'Thal 5', '3/4'), ('H20.33.01
5', 'Thal 3', '3/3'), ('H20.33.017', 'Thal 4', '3/3'), ('H20.33.018', 'Thal 5',
'3/4'), ('H20.33.020', 'Thal 5', '4/4'), ('H20.33.026', 'Thal 4', '4/4'), ('H2
0.33.028', 'Thal 4', '3/3'), ('H20.33.029', 'Thal 4', '3/3'), ('H20.33.031', 'T
hal 4', '3/3'), ('H20.33.037', 'Thal 5', '3/3'), ('H20.33.045', 'Thal 5', '4/
4'), ('H20.33.046', 'Thal 5', '3/3'), ('H21.33.005', 'Thal 3', '3/3'), ('H21.3
3.007', 'Thal 4', '3/3'), ('H21.33.008', 'Thal 4', '3/3'), ('H21.33.009', 'Thal
5', '4/4'), ('H21.33.027', 'Thal 5', '3/4'), ('H21.33.029', 'Thal 5', '2/4'),
('H21.33.039', 'Thal 4', '3/3'), ('H21.33.042', 'Thal 5', '4/4'), ('H21.33.04
4', 'Thal 3', '3/3'), ('H21.33.046', 'Thal 4', '3/3')]
Number of subjects per APOE genotype (Alzheimer's patients):
Counter({'3/3': 13, '4/4': 5, '3/4': 4, '2/4': 1})

Counts of each Thal score among Alzheimer's patients:
Counter({'Thal 5': 11, 'Thal 4': 9, 'Thal 3': 3})
```



In [10]:

```python
import pandas as pd
import matplotlib.pyplot as plt
from collections import Counter


class Patient:
    def __init__(self, donor_id, thal, apoe):
        self.donor_id = donor_id
        self.thal = thal
        self.apoe = apoe

    def __repr__(self):
        return f"Patient(ID={self.donor_id}, Thal={self.thal}, APOE={self.apoe


class AlzheimerCohort:
    def __init__(self, file_path,
                 id_col="Donor ID",
                 dx_col="Consensus Clinical Dx (choice=Alzheimers disease)",
                 thal_col="Thal",
                 apoe_col="APOE Genotype"):
        # Load data
        df = pd.read_csv(file_path)

        # Filter Alzheimer's patients
        alz_df = df[df[dx_col] == "Checked"]

        # Create Patient objects for each row
        self.patients = [
```

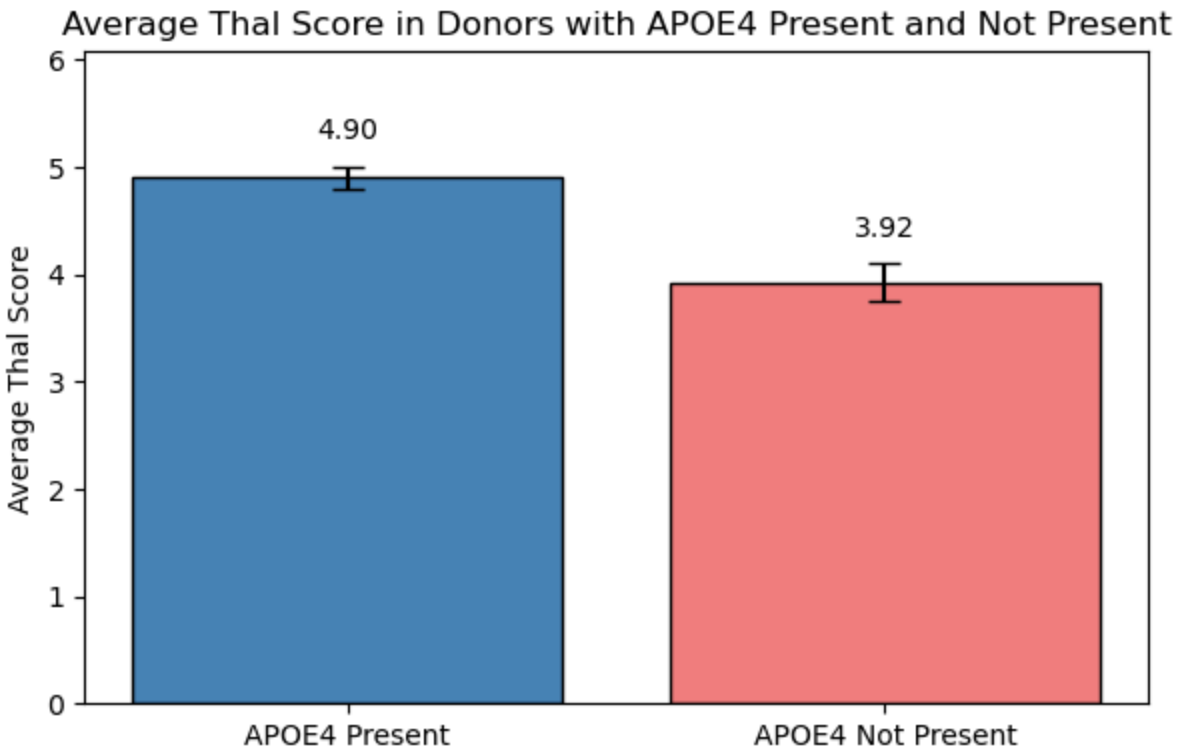
```
            Patient(row[id_col], row[thal_col], row[apoe_col])
            for _, row in alz_df.iterrows()
        ]

    def apoe_counts(self):
        """Count patients per APOE genotype."""
        apoe_list = [p.apoe for p in self.patients]
        return Counter(apoe_list)

    def thal_counts(self):
        """Count patients per Thal score."""
        thal_list = [p.thal for p in self.patients]
        return Counter(thal_list)

    def plot_apoe_distribution(self):
        """Plot APOE genotype distribution among patients."""
        counts = self.apoe_counts()
        plt.figure(figsize=(6,4))
        plt.bar(counts.keys(), counts.values(),
                color="skyblue", edgecolor="black")
        plt.xlabel("APOE Genotype")
        plt.ylabel("Number of Patients")
        plt.title("Distribution of APOE Genotypes in Alzheimer's Patients")
        plt.show()
```

In [ ]:

```
# this code was created with help from ChatGPT, which helps separate all the p
# also creates two groups: one where APOE4 is present, one where it's not pres
# we take the average Thal score for each group and compare that in the bar gr


import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import ttest_ind


# Load data
df = pd.read_csv("NO DATE GENOTYPE Metadata.csv")


# Columns
id_col = "Donor ID"
dx_col = "Consensus Clinical Dx (choice=Alzheimers disease)"
thal_col = "Thal"
apoe_col = "APOE Genotype"




# Get rid of "Thal" before each Thal score listed in the metadata -> makes the
df[thal_col] = df[thal_col].str.extract(r'(\d+)')
df[thal_col] = pd.to_numeric(df[thal_col], errors="coerce")
```

```
# Filter Alzheimer's patients
alz_df = df[df[dx_col] == "Checked"]


# Split into groups based on if APOE4 is present or not
thal_present = alz_df[alz_df[apoe_col].str.contains("4")][thal_col].dropna().a
thal_notpresent = alz_df[~alz_df[apoe_col].str.contains("4")][thal_col].dropna


# t-test, assuming the data has an approximately normal distribution
t_stat, p_val = ttest_ind(thal_present, thal_notpresent, nan_policy="omit")
print(f"t-statistic: {t_stat:.3f}, p-value: {p_val:.4f}")
if p_val < 0.05:
    print("There is a significant difference between the Thal score in patient


# Getting average Thal score for each group and error bars for bar graph
avg_scores = [np.mean(thal_present), np.mean(thal_notpresent)]
sems = [
    np.std(thal_present, ddof=1) / np.sqrt(len(thal_present)),
    np.std(thal_notpresent, ddof=1) / np.sqrt(len(thal_notpresent))
]


# Bar labels
groups = ["APOE4 Present", "APOE4 Not Present"]


plt.figure(figsize=(6,4))
bars = plt.bar(groups, avg_scores, yerr=sems, capsize=6,
               color=["steelblue", "lightcoral"], edgecolor="black")


plt.ylabel("Average Thal Score")
plt.title("Average Thal Score in Donors with APOE4 Present and Not Present")


# Label bars with averages above error bars
for bar, sem, avg in zip(bars, sems, avg_scores):
    plt.text(bar.get_x() + bar.get_width()/2, avg + sem + 0.2,
             f"{avg:.2f}", ha="center", va="bottom")


plt.ylim(0, max(avg_scores) + max(sems) + 1)
plt.tight_layout()
plt.show()
```

```
t-statistic: 4.411, p-value: 0.0002
There is a significant difference between the Thal score in patients with APOE4
and in patients without APOE4.
```



In [1]:

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import ttest_ind

class Patient:
    def __init__(self, file_path, id_col="Donor ID",
                 dx_col="Consensus Clinical Dx (choice=Alzheimers disease)",
                 thal_col="Thal", apoe_col="APOE Genotype"):
        # Load data
        self.df = pd.read_csv(file_path)
        self.id_col = id_col
        self.dx_col = dx_col
        self.thal_col = thal_col
        self.apoe_col = apoe_col

        # Clean up Thal column (extract numeric score)
        self.df[self.thal_col] = (
            self.df[self.thal_col].str.extract(r'(\d+)')
        )
        self.df[self.thal_col] = pd.to_numeric(self.df[self.thal_col], errors=

        # Filter Alzheimer's patients
        self.alz_df = self.df[self.df[self.dx_col] == "Checked"]

    def split_groups(self):
        """Split patients into APOE4 present vs not present groups."""
        thal_present = (
            self.alz_df[self.alz_df[self.apoe_col].str.contains("4", na=False)
```

```
            [self.thal_col].dropna().astype(float).values
        )
        thal_notpresent = (
            self.alz_df[~self.alz_df[self.apoe_col].str.contains("4", na=False
            [self.thal_col].dropna().astype(float).values
        )
        return thal_present, thal_notpresent

    def t_test(self):
        """Perform t-test on Thal scores between APOE4 groups."""
        thal_present, thal_notpresent = self.split_groups()
        t_stat, p_val = ttest_ind(thal_present, thal_notpresent, nan_policy="o

        print(f"t-statistic: {t_stat:.3f}, p-value: {p_val:.4f}")
        if p_val < 0.05:
            print("Significant difference found between groups.")
        else:
            print("No significant difference found.")
        return t_stat, p_val

    def plot_scores(self):
        """Plot average Thal score for each group with error bars."""
        thal_present, thal_notpresent = self.split_groups()
        avg_scores = [np.mean(thal_present), np.mean(thal_notpresent)]
        sems = [
            np.std(thal_present, ddof=1) / np.sqrt(len(thal_present)),
            np.std(thal_notpresent, ddof=1) / np.sqrt(len(thal_notpresent))
        ]
        groups = ["APOE4 Present", "APOE4 Not Present"]

        plt.figure(figsize=(6,4))
        bars = plt.bar(groups, avg_scores, yerr=sems, capsize=6,
                       color=["steelblue", "lightcoral"], edgecolor="black")

        plt.ylabel("Average Thal Score")
        plt.title("Average Thal Score in Donors with APOE4 Present and Not Pre

        # Add labels above bars
        for bar, sem, avg in zip(bars, sems, avg_scores):
            plt.text(bar.get_x() + bar.get_width()/2, avg + sem + 0.2,
                     f"{avg:.2f}", ha="center", va="bottom")

        plt.ylim(0, max(avg_scores) + max(sems) + 1)
        plt.tight_layout()
        plt.show()
```

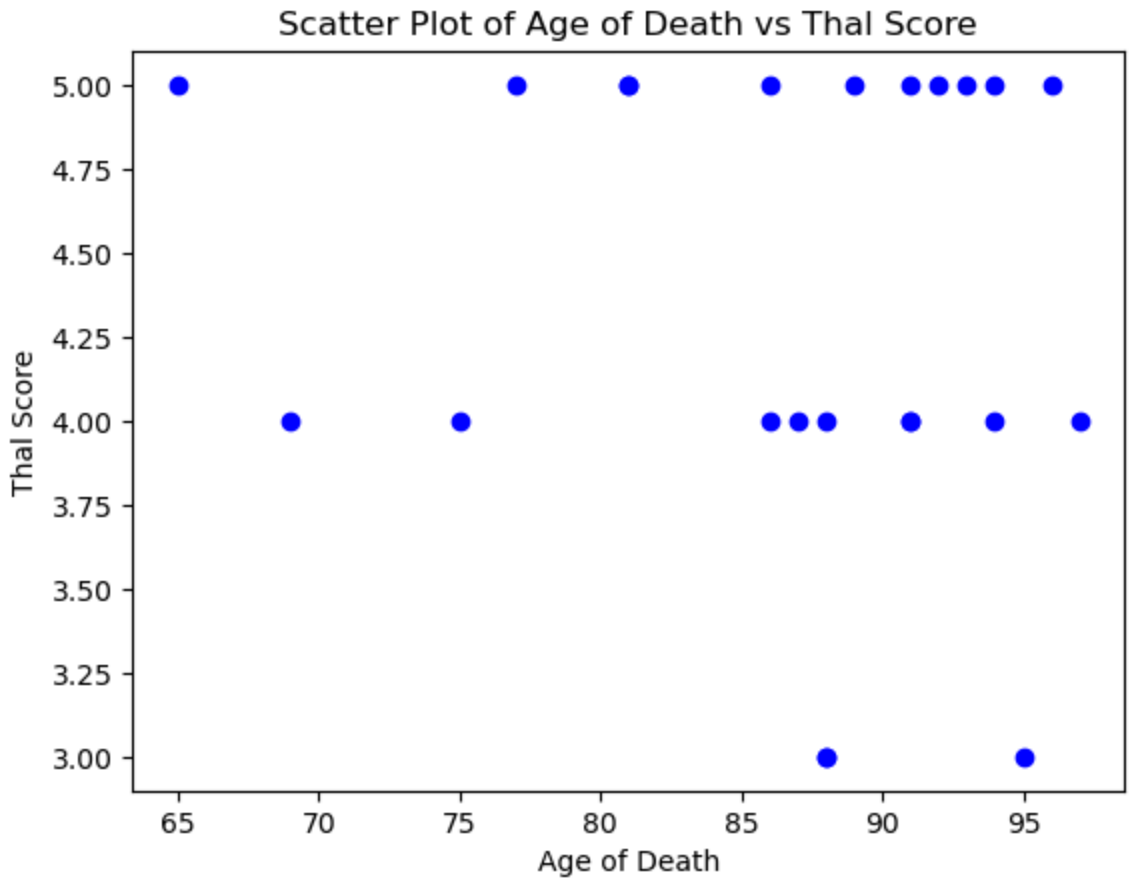
In [2]:

```
# making a scatter plot for the age of death and the thal score for patients f

# 1. making a list of the patient id, age of death, thal score, apoe genotype
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

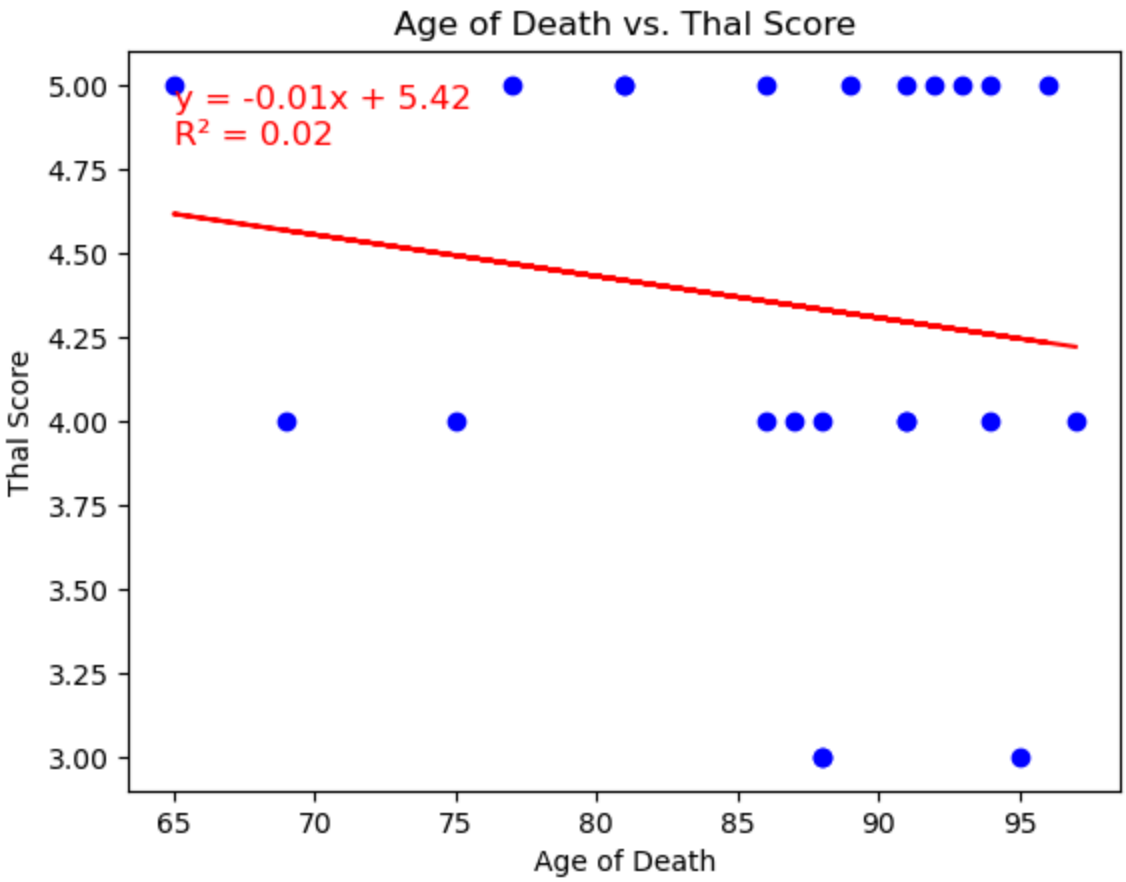
```
df = pd.read_csv("NO DATE GENOTYPE Metadata.csv")

id_col = "Donor ID"
dx_col = "Consensus Clinical Dx (choice=Alzheimers disease)"
thal_col = "Thal"
apoe_col = "APOE Genotype"
age_death = "Age at Death"

# Filter Alzheimer's patients from consensus clinical dx
alz_df = df[df[dx_col] == "Checked"]

# Select only ID, Thal, and APOE genotype
alz_selected = alz_df[[id_col, thal_col, apoe_col,age_death]]

# Convert to a list of tuples (ID, Thal, APOE, Age of death)
alz_list = alz_selected.to_records(index=False).tolist()



print("\nList of Alzheimer's patients (ID, Thal, APOE, Age at Death):")
print(alz_list)

# 2. getting the patients death_age_list and Thal score

death_age_list = []
Thal_score = []

for patient in alz_list:
    death_age_list.append(patient[3])


for patient in alz_list:
    Thal_score.append(int(patient[1].split()[1]))

print(death_age_list)
print(Thal_score)

# 3. making the graph for the scatter plot using the death_age_list and the Th
x = [death_age_list]
y = [Thal_score]

plt.scatter(x, y, color='blue')
plt.xlabel('Age of Death')
plt.ylabel('Thal Score')
plt.title('Scatter Plot of Age of Death vs Thal Score')
plt.show()

# 4. export data to a .csv file
# Create a DataFrame
df = pd.DataFrame({'Age of Death': death_age_list,'Thal Score': Thal_score})

# Write to CSV
```

```
df.to_csv('patient_data.csv', index=False)
print("CSV file 'patient_data.csv' has been created.")

# 5. Loading thr libaries for a linear regression
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score

# 6. loading the data set for a linear regression
df = pd.read_csv("patient_data.csv")

# 7. updating the variable names to match Exactly your .csv file header
x = df["Age of Death"].values.reshape(-1, 1)
y = df["Thal Score"].values

# 8. Perform the linear regression
model = LinearRegression()
model.fit(x, y)

slope = model.coef_[0]
intercept = model.intercept_
r2 = model.score(x, y)

# 9. Make scatterplot with the equation for the linear regression and the R-sq
plt.scatter(x, y, color='blue')
plt.plot(x, model.predict(x), color="red")
# Annotate equation
equation = f"y = {slope:.2f}x + {intercept:.2f}\nR² = {r2:.2f}"
plt.text(x.min(), y.max(), equation, color="red", fontsize=12, verticalalignme

# Annotate scatterplot with labels and title
plt.xlabel("Age of Death")
plt.ylabel("Thal Score")
plt.title("Age of Death vs. Thal Score")
plt.show()
```

```
List of Alzheimer's patients (ID, Thal, APOE, Age at Death):
[('H20.33.004', 'Thal 5', '3/4', 86), ('H20.33.011', 'Thal 5', '3/4', 93), ('H2
0.33.015', 'Thal 3', '3/3', 88), ('H20.33.017', 'Thal 4', '3/3', 69), ('H20.3
3.018', 'Thal 5', '3/4', 81), ('H20.33.020', 'Thal 5', '4/4', 81), ('H20.33.02
6', 'Thal 4', '4/4', 75), ('H20.33.028', 'Thal 4', '3/3', 94), ('H20.33.029',
'Thal 4', '3/3', 91), ('H20.33.031', 'Thal 4', '3/3', 87), ('H20.33.037', 'Thal
5', '3/3', 96), ('H20.33.045', 'Thal 5', '4/4', 77), ('H20.33.046', 'Thal 5',
'3/3', 94), ('H21.33.005', 'Thal 3', '3/3', 95), ('H21.33.007', 'Thal 4', '3/
3', 86), ('H21.33.008', 'Thal 4', '3/3', 91), ('H21.33.009', 'Thal 5', '4/4', 6
5), ('H21.33.027', 'Thal 5', '3/4', 92), ('H21.33.029', 'Thal 5', '2/4', 89),
('H21.33.039', 'Thal 4', '3/3', 88), ('H21.33.042', 'Thal 5', '4/4', 91), ('H2
1.33.044', 'Thal 3', '3/3', 88), ('H21.33.046', 'Thal 4', '3/3', 97)]
[86, 93, 88, 69, 81, 81, 75, 94, 91, 87, 96, 77, 94, 95, 86, 91, 65, 92, 89, 8
8, 91, 88, 97]
[5, 5, 3, 4, 5, 5, 4, 4, 4, 4, 5, 5, 5, 3, 4, 4, 5, 5, 5, 4, 5, 3, 4]
```




```
CSV file 'patient_data.csv' has been created.
```



Verify and validate your analysis:

Our results showed a correlation between the presence of APOE4 and Thal score. This is largely consistent with published literature in the field. Qian et al., found that APOE genotype does affect the rate of cognitive decline related to Alzheimer's [13]. Instead of using Thal score as an indicator for cognitive decline, Qian et al., used Clinical Dementia Rating scale Sum of Boxes and Mini-Mental State Examination to assess AD neuropathological changes [13]. Raulin et al., researched the impact of APOE genotype on amyloid accumulation, which is what Thal score is based on [14]. Raulin et al., also found that APOE4 increases the risk of Alzheimer's [14]. There are not many studies that directly relate APOE genotype and Thal score, which indicates possible weakness in the original research question and opens up possibilities for future work mentioned later.

Conclusions and Ethical Implications:

We wanted to determine whether or not the presence of APOE4 is correlated to Thal score, which categorizes the spread of beta-amyloid plaques in the brain. We

performed a t-test on average Thal score in donors with and without APOE4, and found the p-value to be 0.0002, which is less than 0.05. Thus, we can reject the null hypothesis and conclude there is correlation between the presence of APOE4 and Thal score.

We also wanted to assess the effect of having a higher Thal score by plotting age of death vs Thal score. After performing linear regression and calculating $R^2 = 0.02$, we can conclude there is not a significant correlation between the two factors.

While we did find correlation between the presence of APOE4 and Thal score, we cannot recommend judging Thal score based on the presence of APOE4. Thal score also should not be used to assess the length/quality of a patient's life, especially because Thal score is discrete and would not accurately represent the various ways Alzheimer's can present in a patient. Many factors should be considered when judging the severity of Alzheimer's.

Limitations and Future Work:

Limitations such as sample size should be taken into consideration when understanding the results. We limited the donor pool for analyzing the effect of APOE4 to patients diagnosed with Alzheimer's. This makes the sample pool consist of 20 donors, which is not enough to draw a strong conclusion from the data results.

Future work may include determining if there is a correlation between APOE genotype and amyloid beta 40/42. Thal score is based on the amount of amyloid beta plaques that accumulate in the brain, so it would be better to perform an analysis with a variable with greater range rather than with a discrete variable like Thal score.

Additionally, APOE genotype for each donor was originally given in pairs: some patients had both copies of APOE4, while others only had one copy or none. For this study, two groups were formed based on if any APOE4 was present versus not present at all. For a more comprehensive study in the future, groups should be split up by number of copies of APOE4. The relationship between the number of copies of APOE4 and amyloid beta buildup would be assessed. There has also been discussion in the field about how APOE genotypes affect different races, so future work could create groups based on race with the same genotype throughout, and compare the amyloid beta buildup among those groups.

Team Notes:

hyperlink to notes: https://docs.google.com/document/d/1Fz930gJWZ8tk_qww5iujnogwaSvxO-55bxGLeK9NVIY/edit?usp=sharing

Questions for TA:

- We don't have any more questions.

References

[1] “Alzheimer’s Association 2025 Alzheimer’s Disease Facts and Figures”.

[2] “Who is Eligible for the New Alzheimer’s Drug Leqembi?” Accessed: Sep. 28, 2025. [Online]. Available: https://www.ncoa.org/article/leqembi-and-alzheimers-disease-the-benefits-side-effects-and-what-it-costs/

[3] “What Causes Alzheimer’s (Genetic Risks & Lifestyle Factors).” Accessed: Sep. 28, 2025. [Online]. Available: https://optoceutics.com/what-causes-alzheimers-disease-genetics-lifestyle-factors/

[4] CDC, “Non-Medical Factors that Affect Alzheimer’s Disease and Related Dementias Risk,” Alzheimer’s Disease and Dementia. Accessed: Sep. 28, 2025. [Online]. Available: https://www.cdc.gov/alzheimers-dementia/php/sdoh/index.html

[5] “10 Early Signs and Symptoms of Alzheimer’s & Dementia | alz.org,” Alzheimer’s Association. Accessed: Sep. 28, 2025. [Online]. Available: https://www.alz.org/alzheimers-dementia/10_signs

[6] “Diagnosis,” Alzheimer’s Association. Accessed: Sep. 28, 2025. [Online]. Available: https://www.alz.org/alzheimers-dementia/diagnosis

[7] “Insurance | Alzheimer’s Association,” Alzheimer’s Association. Accessed: Sep. 29, 2025. [Online]. Available: https://www.alz.org/help-support/caregiving/financial-legal-planning/insurance

[8] “Alzheimer’s Stages - Early, Middle, Late Dementia Symptoms | alz.org,” Alzheimer’s Association. Accessed: Sep. 28, 2025. [Online]. Available: https://www.alz.org/alzheimers-dementia/stages

[9] Alzra, “What part of the brain does Alzheimer’s affect?,” Alzheimer’s Research Association. Accessed: Sep. 30, 2025. [Online]. Available: https://www.alzra.org/

blog/what-part-of-the-brain-does-alzheimers-affect/

[10] “New Protein Interaction Map Sheds Light on How Brain Cell Communication Breaks Down in Alzheimer’s Disease | Mount Sinai - New York,” Mount Sinai Health System. Accessed: Sep. 30, 2025. [Online]. Available: https://www.mountsinai.org/about/newsroom/2025/new-protein-interaction-map-sheds-light-on-how-brain-cell-communication-breaks-down-in-alzheimers-disease

[11] “NIA-Funded Active Alzheimer’s and Related Dementias Clinical Trials and Studies,” National Institute on Aging. Accessed: Sep. 30, 2025. [Online]. Available: https://www.nia.nih.gov/research/ongoing-AD-trials

[12] M. I. Gabitto et al., “Integrated multimodal cell atlas of Alzheimer’s disease,” Nat. Neurosci., vol. 27, no. 12, pp. 2366–2383, Dec. 2024, doi: 10.1038/s41593-024-01774-5.

[13] J. Qian, Y. Zhang, R. A. Betensky, B. T. Hyman, and A. Serrano-Pozo, “Neuropathology-Independent Association Between APOE Genotype and Cognitive Decline Rate in the Normal Aging-Early Alzheimer Continuum,” Neurol. Genet., vol. 9, no. 1, p. e200055, 2023, doi: 10.1212/NXG.0000000000200055.

[14] A.-C. Raulin et al., “Impact of APOE on amyloid and tau accumulation in argyrophilic grain disease and Alzheimer’s disease,” Acta Neuropathol. Commun., vol. 12, no. 1, p. 25, Feb. 2024, doi: 10.1186/s40478-024-01731-0.