**Section 1.**

1. What is the effect of acupuncture on chronic headaches?

A visual assessment of severity before and after acupuncture and control was conducted first to try visually identify any differences.

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

In the control group on the left we can see a small difference in the pk scores before and after treatment. The patients in the control group received the standard intervention which appear to have been successful for some. On the right we can see that the pk values trended towards lower values and at a higher rate after acupuncture as an intervention. There is a reduction in both groups but it is more pronounced in the acupuncture group.

Chart, histogram

Description automatically generated

The post intervention severity was compared between the groups and we can see the frequency of lower severity is higher in the intervention arm.

The data was plotted on qqplots against a normal distribution to guide the use of statistical test.

|  |  |
| --- | --- |
|  |  |
|  |  |

The data failed to fit the qqnorm plots so a non-parametric Wilcoxon test was used. Both groups came out as significant however the treatment group was much more significant (control: p = 1.253e-06, treatment: p = 2.494e-14). Both groups received an intervention so improvement could be seen regardless but there was a stronger reduction in severity score when acupuncture was used.

A logistic regression was then completed on the response category and group category. Response is binary with a >35% improvement from baseline considered a success, while group is 0 for control and 1 for intervention. ‘Group’ was used as the explanatory variable and ‘response’ as the response variable in the linear model. The coefficient for the "group" variable is estimated as 0.9091, which means that the log-odds of success are expected to increase by 0.9091 for each unit increase in the "group" variable (i.e., from control to intervention group). Indicating a higher success with the treatment group (p = 0.000155). While a significant decrease in severity was seen in two groups it was much greater in the intervention group and the regression identified a strong relationship between acupuncture and a classified success based on response variable.

1. Does it work the same way for migraine and other types of chronic headaches?

Chart, histogram

Description automatically generated Chart, histogram

Description automatically generated

The density plots above show us that in the treatment arm the pk5 response is spread out more with a decrease in the lower values and a slight increase in mid and upper values, overall there isn’t that much difference in the shape of the graph besides a lower peal. There was no significant difference in the medians (p = 0.05191).

In the control arm we see a dramatic decrease in the frequency of lower pk values and the first peak is broader. Uppers values also decrease with no pk value ~>50, a t-test returned a non-significant result indicating no difference between the groups (p= 0.5541).Thus, neither acupuncture nor the standard of care significantly reduced the severity of headaches.

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

For the migrane arm we see a similar graph to when the cohorts weren’t separated. The reported pk values decreased after accupunture while in the control arm there was also a small decrease but not as dramatic. A t-test on both these groups were significant (treatment: p= 1.055e-13, Control: p= 1.243e-06) indicating that accupunture and standard of care have an affect on migraine but neither has an effect on chronic headaches.

|  |  |  |
| --- | --- | --- |
|  | response.0 | response.1 |
| 0 | 8 | 1 |
| 1 | 66 | 86 |

A chi-square test was completed on the above table which was a count of the response in non-migraine vs migraine in the treatment arm alone. The p value was significant (p= 0.0206) indicating a difference in the expected responses in migraine and headache. It is clear from the numbers that acupuncture was not considered a success with headaches.

1. What is the effect of gender on overall efficacy of acupuncture?

Chart

Description automatically generatedChart

Description automatically generated

We can see in both groups that before the intervention with acupuncture that both groups reported higher severity at baseline compared to 1 year follow up. Males also had a small second peak at higher values while females had a relatively smooth tale. Both groups reported decreased severity after the use of acupuncture although the curve for males was quite irregular it was still a significant difference in the medians (p-value = 0.001084). The difference in medians were very significant for females (p-value < 7.592e-12).

|  |  |  |
| --- | --- | --- |
|  | response.0 | response.1 |
| Sex 0 | 14 | 14 |
| Sex 1 | 60 | 73 |

However, while both males and females responded to treatment there was no significant difference between the genders in the response variable despite the numbers (chisquare test: p = 0.7925). Gender did not influence treatment response.

1. Does acupuncture influence general health and packs of medications beside headache overall?

Chart, histogram

Description automatically generated

The SF36 scale ratings, which are an assessment of general health, were very similar before and after treatment with acupuncture. There was a slight decrease in mid-range values and slightly higher in upper values but no drastic differences. There was no significant difference in the medians of the two variables (p = 0.09145). Acupuncture did not have an effect on reported quality of life despite significantly reducing the severity of headaches. Perhaps this was because pain meds may have been effectively managing the pain symptoms at baseline were consumed despite the higher severity and thus maintaining general health.

Diagram

Description automatically generated

On the other hand there was a drastic change in the frequency of medicine packs taken before and after treatment with acupuncture that was significant (p= 2.2e-16).

In conclusion acupuncture is a viable alternative to the current standard of care for chronic migraine suffers.

**Section 2:**

First the inclusion criteria was defined:

# Creating the inclusion criteria  
age\_range = (40, 80)  
ICD\_10 = ['G30']  
ATC = ['N06D']

Next exclusion criteria was defined:

# Creating exclusion criteria  
ICD\_10\_NA = ['F01', 'F02', 'F03']

A theoretical dataset was saved into the environment. This would contain the data to be filtered in order to create the RWE data frame with the criteria of interest. This data can come from sources such as electronic health records, registries or biobanks that contain real world populations which are hopefully diverse.

# Load patient data  
patient\_data = pd.read\_csv('sample\_data.csv')#theoretical DF

The chosen exclusion and exclusion criteria were used to filter the data frame and saved to a new dataset. It was assumed that the columns would be AGE and ICD\_10 that would contain the filtering criteria. Only the datapoints in the dataset that met the criteria will be included in the final dataframe.

# Filtering via criteria

patient\_data = patient\_data[(patient\_data['AGE'] >= age\_range[0]) & (patient\_data['AGE'] <= age\_range[1])]#upper and lower age

patient\_data = patient\_data[(patient\_data['ICD\_10'].isin(ICD\_10))] #filtering based on alzehimers ICD code

patient\_data = patient\_data[(patient\_data['ATC'].isin(ATC))]#filtering based on drugs

patient\_data = patient\_data[~patient\_data['ICD\_10'].isin(ICD\_10\_NA)] #Filtering out non aplicable ICD-10 codes

ICD codes are a centralised classification of diseases in a codified manner. They are used in the context of diseases, symptoms and procedures and are maintained by the WHO.

# Save RWE dataframe as Data00  
Data00 = patient\_data

The dataframe was then saved as the desired name Data00. This dataframe can be used to inform about risks and benefits of interventions and can give a more generalised interpretation.

**Section 3.**

1. **Build a decision tree – an *rpart* tree and report confusion table, balanced accuracy as well as the tree diagram?**

Diagram

Description automatically generated with medium confidence

|  |  |  |
| --- | --- | --- |
|  | B | M |
| B | 79 | 10 |
| M | 5 | 48 |
| Sensitivity | 79/79+10 = 88.76% | |
| Specificity | 48/(48+5) = 90.57% | |

In the decision tree the majority of splits in the first node were based on the variable perimeter worst. This variable gave us the most information gain and lower the entropy of the model before the next split. At each node the variable that contributed to the most information gain is displayed and the split value i.e., if true it goes to the left. At the leaf nodes we have the output/classification of the datapoints. We want the bins to be as pure as possible meaning only data of one type being in each bin or as minimal as possible. This is so we can be confident in correctly identifying if the tumours are malignant or benign based on the morphology data.

Graphical user interface, text

Description automatically generated

Lets take node 4 for example. That is a datpoint were classified as benign in this leaf there is a 98% probability that this was the correct call. The p(node) also shows how much of the data is in this leaf i.e., just over half. This could be an indication of a relationship between the values of perimeter\_worst and concace.points\_worst being less than the cut off points and tumours being accurately classified as benign.

A balance in the sensitivity and specificity are needed when deciding on acceptability. High sensitivity and low specificity would lead to high false negatives and overtreatment of those patients with benign tumours. While the opposite would lead to false negatives and thus delay diagnosis and treatment. In terms of cancer screening, we ideally want to keep the false negatives low thus we would prioritise sensitivity at the risk of overtreating some patients. This tree has a good balance on the two but if we were to prioritise sensitivity, we may have to adjust some of the decision criteria to improve the sensitivity.

1. **Build a random forest with at least 10 trees and report confusion table and balanced accuracy. Is it different compared to the one from the tree above? Compare and interpret the confusion tables from both models – the decision tree and the random forest.**

|  |  |  |
| --- | --- | --- |
| Ntree= 10 | | |
|  | B | M |
| B | 82 | 7 |
| M | 8 | 45 |
| Sensitivity | 82/(82+7) = 92.13% | |
| Specificity | 45/(45+8) = 84.9% | |

|  |  |  |
| --- | --- | --- |
| Ntree= 500 | | |
|  | B | M |
| B | 83 | 6 |
| M | 6 | 47 |
| Sensitivity | 83/(83+6) = 93.25% | |
| Specificity | 47/(47+6) = 88.67% | |

Two random forests were built, one with ten trees and the other with 500. The accuracy of the forest didn’t really improve drastically with the addition of more decision trees. However, the sensitivity has increased from the singular tree which is something that would be considered when deciding on a model for cancer screening. The improved sensitivity could be from stricter criteria in the splits.

1. **Find important features from both models. Compare and interpret on the similarities/differences of important features/covariates from the tree (generated from Q1 of this section) and random forests (generated from Q1 of this section)?**

In the tree model the variables importance were listed as:

Text

Description automatically generated with medium confidenceTimeline

Description automatically generated with medium confidence

While both models use different methods of calculating the importance both are based on the Gini index. In both perimeter\_worst appeared as one of the top 3 variables in the decisions indicating that this variable may have significant prediction abilities in these circumstances.

Index:

Q1.

Organizing the data and creating density plots

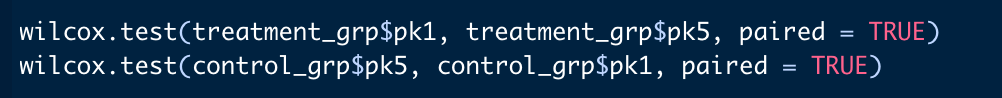


Testing for normality

Text

Description automatically generated

Statistical tests

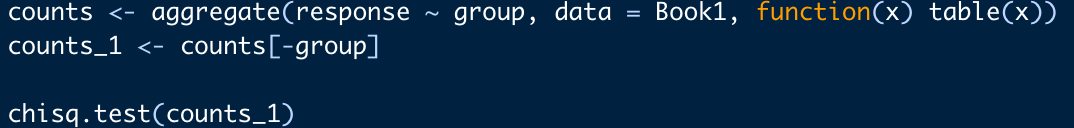


Text

Description automatically generated

The same code was used to create the other plots using the other variables.

Creating count of variables and chisquare test. This was adapted for each question as needed.



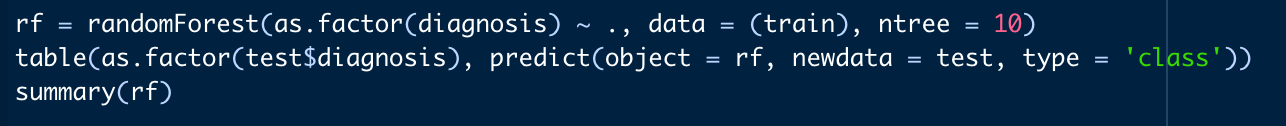
Q3.1

Separating the data. Creating and testing the decision tree.

Graphical user interface, text

Description automatically generated

Creating and testing random forest.



Testing the importance of the decision tree and random forest.

Graphical user interface, application

Description automatically generated

Summary output for decision tree. Text

Description automatically generated with medium confidence