Part A:

Question 1:

The prototypical patient within this study was a married, employed female aged 52. They could drive and had received on average 14 years of education up to a High School Level and had been ill with MS for 16 years.

Bar charts were created to represent the relationship between employment type and illness duration. As illness duration increases, the number of patients in different employment changes (see Figure 1 and 2). For example, as a general trend, as illness duration increases past twelve years, the number of patients employed reduces. Additionally, within this dataset, there were very few students, and these students had a relatively low illness duration. No patients in the first 4 years of illness duration gave disability as their employment status. However, it became more common after this, and until around 30 years of illness, the most common types were employed or disability. Retirement did not appear to be influenced by illness duration, as there were retired patients throughout the whole graph.

Figure 1

A bar chart showing illness duration and the number of patients in different types of employment.

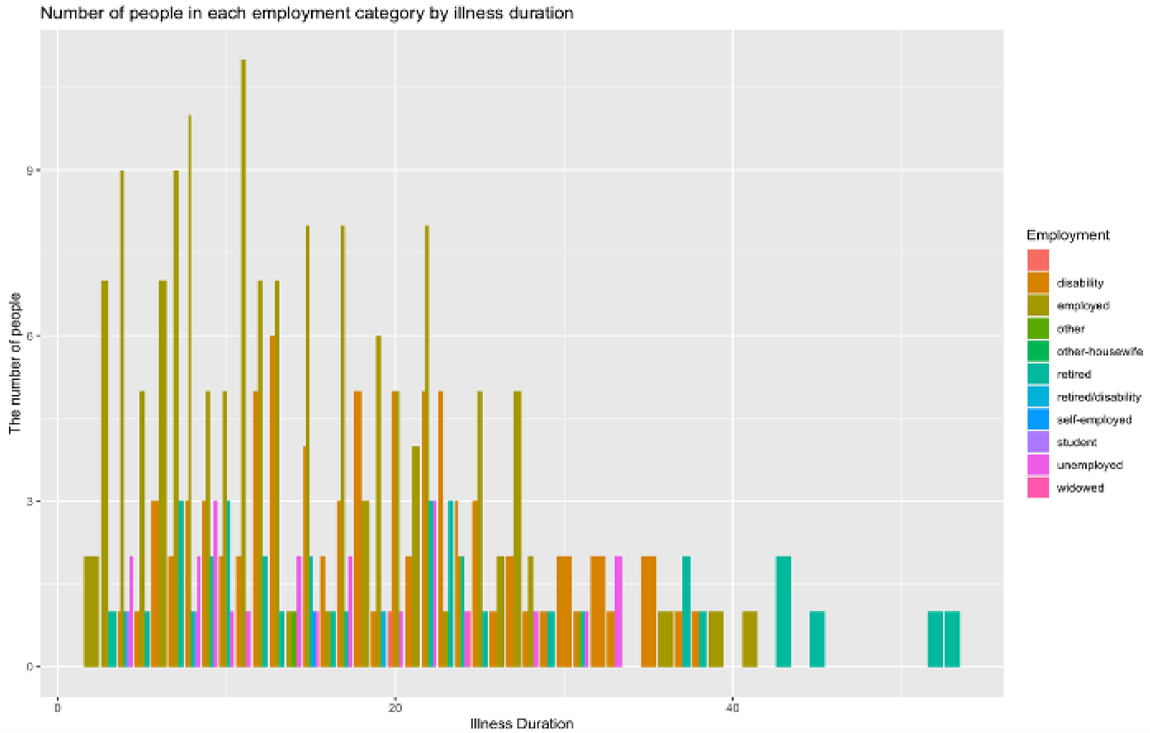
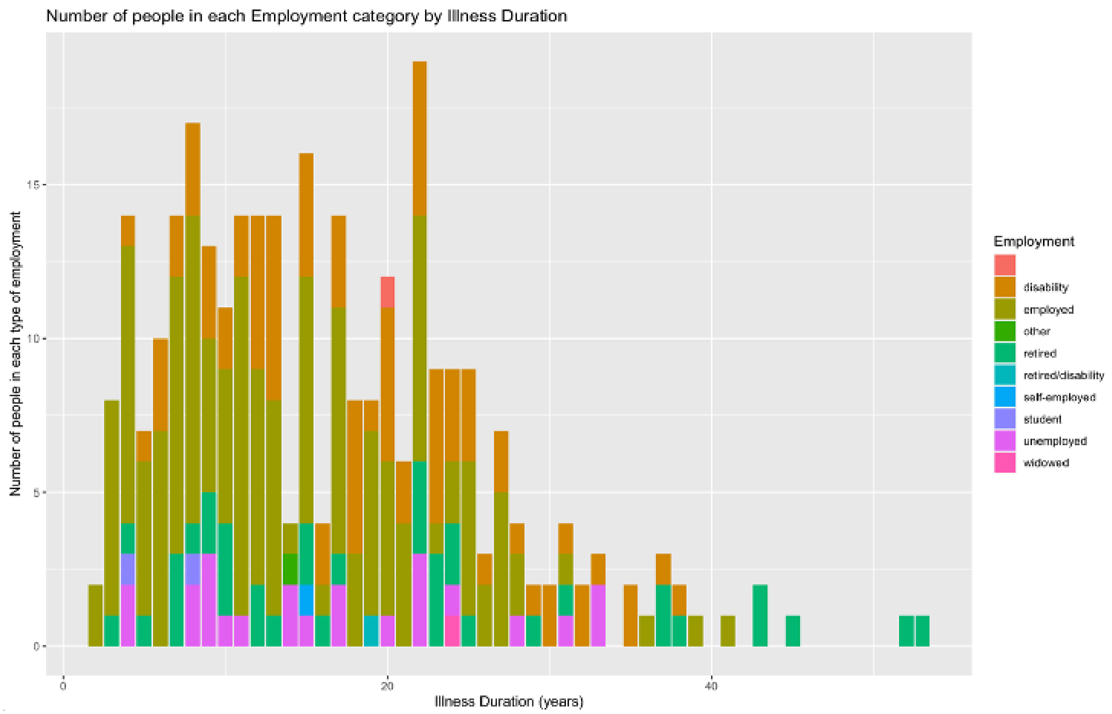


Figure 2

A stacked bar chart showing illness duration and the number of patients in different types of employment.



Question 2: Decision making and cognitive ability

T-tests were conducted to test the alternative hypotheses that patients who share decision making with their physician differ in their cognitive abilities from patients who believe one party (the physician or the patient) should make medical decisions. Each hypothesis focuses on the impact of shared or solo decision-making on a different cognitive ability: memory, attention, and visuospatial ability.

For each hypothesis, an investigation of outliers was conducted, as well as testing of normality and equal variance assumptions before the t-test (Appendix A). For all three tests, the assumption of equal variance was confirmed; however, the data were not normally distributed. I decided to go ahead with the three t-tests despite the violation of normality. Despite non-normal distribution, T-tests work well when the data has more than 25 observations per group and no extreme outliers influencing test results (le Cessie et al., 2020). As my data fits these conditions, I conducted the t-tests. Additionally, the parametric test has an advantage over the non-parametric equivalents which only provide a P-value, as the t-test also produces the observed mean differences, with a 95% confidence interval (CI) (le Cessie et al., 2020).

Decision making and memory ability:

A Welch Two Sample t-test was conducted on the data with a 95% confidence interval (CI) for the mean differences in memory scores between the groups. It was found that the group of patients who share their decision making with their physician had a significantly higher memory score (M=96.18) than those who believed one party should make decisions (*M*=88.42), (*t*(51.84)=2.51,*p*=.018) with a difference of 7.75 (95% CI, 1.41 to 14.10).

Therefore, the alternative hypothesis, which states that the true difference in means between the groups is not equal to 0, is supported by the data.

Decision making and attention ability:

A Welch Two Sample t-test was conducted on the data with a 95% confidence interval (CI) for the mean differences in attention scores between the groups. It was found that the group of patients who share their decision making with their physician had a significantly higher attention score (*M*=97.58) than those who believed one party should make decisions (*M*=91.34), (*t*(57.45)=2.35,*p*=.022) with a difference of 6.23 (95% CI, 0.92 to 11.54).

Therefore, the alternative hypothesis, which states that the true difference in means between the groups is not equal to 0, is supported by the data.

Decision making and visuo-spatial ability:

A Welch Two Sample t-test was conducted on the data with a 95% confidence interval (CI) for the mean differences in visuospatial ability between the groups. It was found that the group of patients who share their decision making with their physician had a significantly higher attention score (*M*=104.68) than those who believed one party should make decisions (*M*=95.86), (*t*(53.15)=2.76,*p*=.008) with a difference of 8.82 (95% CI, 2.40 to 15.25).

Therefore, the alternative hypothesis, which states that the true difference in means between the groups is not equal to 0, is supported by the data.

Overall conclusion:

Therefore, the results of the three t-tests suggest that findings generalise across all three cognitive abilities, suggesting that patients who share the decision-making with their physician have better cognitive abilities in memory, attention, and visuospatial than patients who believe one party should make the decision.

Question 3: Linear modelling

A linear model was created to understand whether disease management can predict illness perception. See appendix B, for full assumption test results, and justifications for linear regression analysis.

Linear regression established that disease management could significantly predict illness perception, $F(1, 337) = 40.77$, $p < .001$ and disease management accounted for 10.5% of the explained variability in illness perception. See table 1 for full results. The regression equation was: Predicted illness perception = -1.2697 + 65.69 x (managing meds score), suggesting that for each unit increase in Managing_Meds, the expected value of Brief_Illness decreases by 1.2847 units (see Figure 3).

Table 1

Regression of disease management model predicting illness perception

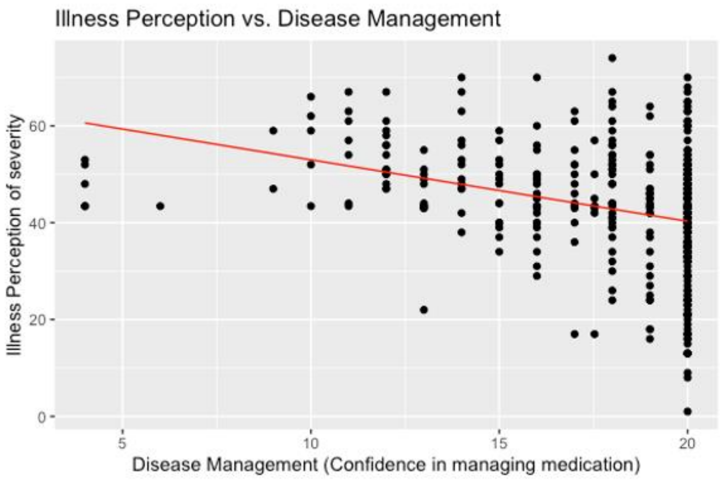
	Illness perception		
	Estimates	*p*-value	95%CI
Intercept	65.69	<.001***	[58.71, 72.66]
Managing Meds	-1.27	<.001***	[-1.66, -0.88]

***p<.001, **p<.01, p*<.05

Observations 339
R^2/Adjusted R^2 0.108 / 0.105

Figure 3

Negative relationship between illness perception and disease management. Diagonal line represents best regression fit to data points.



Does shared decision-making alter the strength of this relationship?

A multiple regression analysis was conducted to predict illness perception from disease management and shared decision making. The model was significant, $F(2,336) = 20.87$, $p < .001$, and explained 11% of the variance in illness perception. Disease management was a significant predictor of illness perception, β = -1.28, t(336) = -6.44, p < .001. However, shared decision-making was not a significant predictor, β = -1.95, t(336) = -

0.99, p = .323. See table 2 for full results. Therefore, shared decision-making did not alter the strength of the relationship between illness perception and disease management.

Table 2

Regression of disease management and shared decision model predicting illness perception

	Illness perception		
	Estimates	*p*-value	95%CI
Intercept	66.20	<.001***	[59.15, 73.25]
Managing Meds	-1.28	<.001***	[-1.68, -0.89]
Shared DM	-1.95	0.323	[-5.83,1.93]

***p<.001, **p<.01, p*<.05

Observations 339
R^2/Adjusted R^2 0.111 / 0.105

Question 4

A correlation matrix was conducted to infer the relationship between illness perception, cognitive abilities, and mental health (see Figure 4).

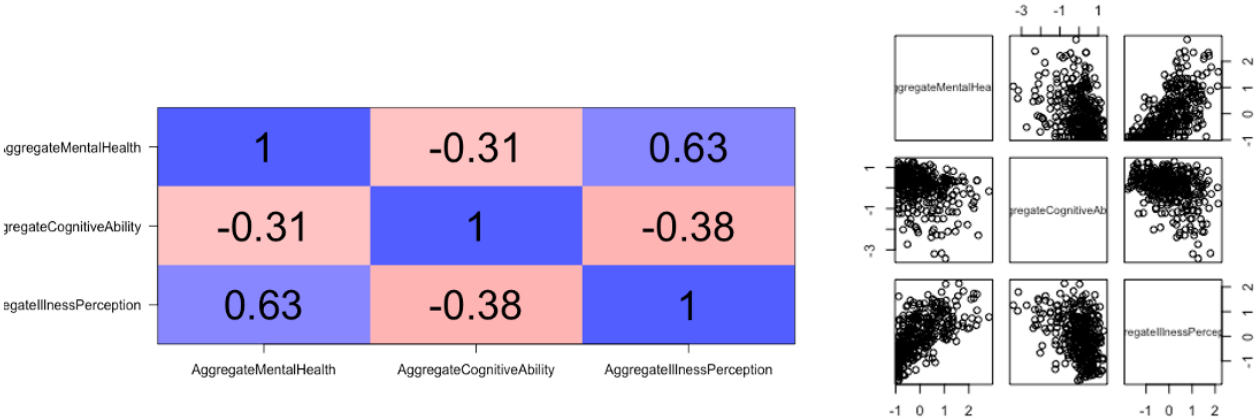
Scales were created to aggregate the scores of multiple indicators into the general terms.

- A higher score in illness perception represents worse perceived severity of illness
- A higher score in mental health represents poorer mental health

- A higher score in cognitive ability represents better cognitive ability

Figure 4

Correlation matrix showing the relationship between illness perception, cognitive ability, and mental health



Pearson’s correlation analysis was conducted, and all correlations were significant. The strongest correlation was between mental health and illness perception, with a correlation of 0.63; stating that, as illness perception increased, mental health declined. There were negative correlations between cognitive ability and mental health (-0.38) and cognitive ability and illness perception (-0.31). This demonstrated that cognitive ability and illness perception declined as mental health declined.

Appendix A

Assumption test results for question 2

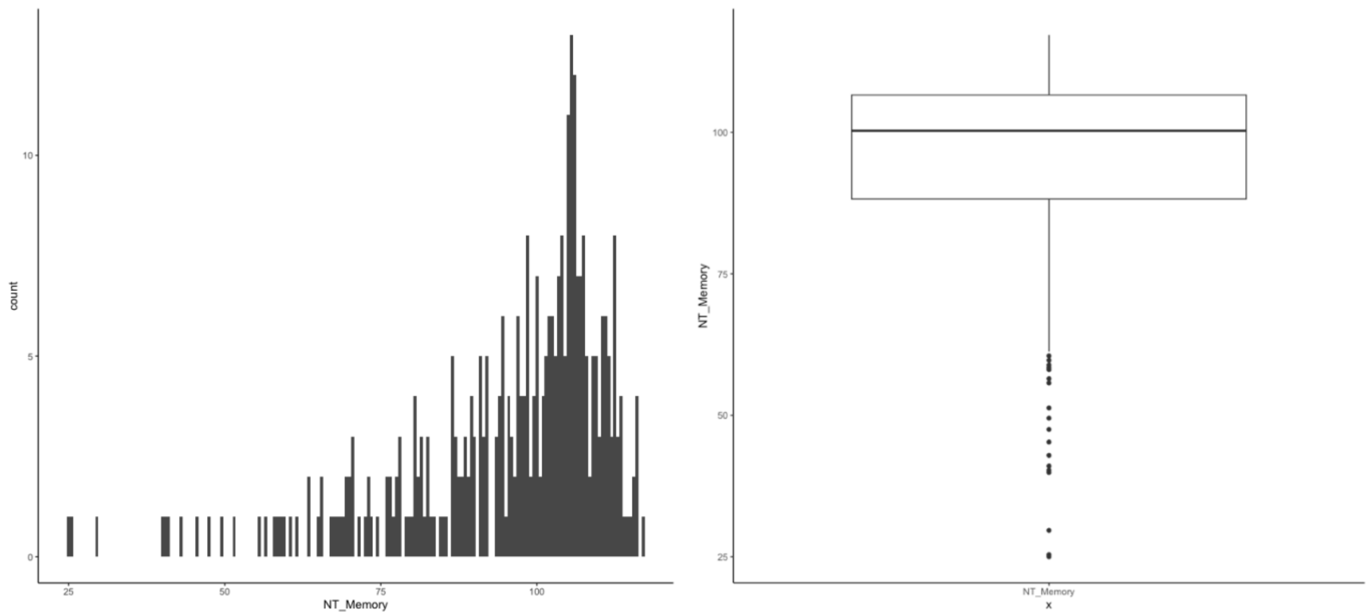
Decision making and memory t-test:

Group 1 and group 2 refer to the decision-making preferences. Group 1 are the patients who prefer to share decision making with their doctor. Group 2 are the patients who prefer one party (patient or doctor) to make these decisions.

Identifying outliers:

Figure 5

A histogram and boxplot of memory scores



Visual inspection of the histogram and the boxplot suggests there may be observations lower and higher than all other observations (see Figure 5). Percentiles were used to determine outliers; potential outliers were defined as observations that lie outside the interval formed by the 2.5th and 97.5th percentiles. From this, 18 outliers were identified. As the data was not collected by myself, I cannot be sure of the source of the outliers. Therefore, I cannot justify removing them because of data entry or measurement errors. However, removing these outliers influenced the significance of the results of the Welch two-sample t-test ($t(51.84)=2.51, p=.018$ including outliers, $t(49.13)=2.1.91, p=.063$ excluding outliers). Therefore, the outliers were removed before the t-test was conducted, and the report presents the t-test with the outliers excluded.

Assumption of normality:

In a T-test, the residuals should be normally distributed (Kim & Park, 2019). First, I visually inspected the data; the histograms and Q-Q plots suggested the residuals were not normally distributed when outliers were included and excluded from the data (see Figure 6 and 7).

Figure 6

Histograms and Q-Q Plots of Memory scores including outliers

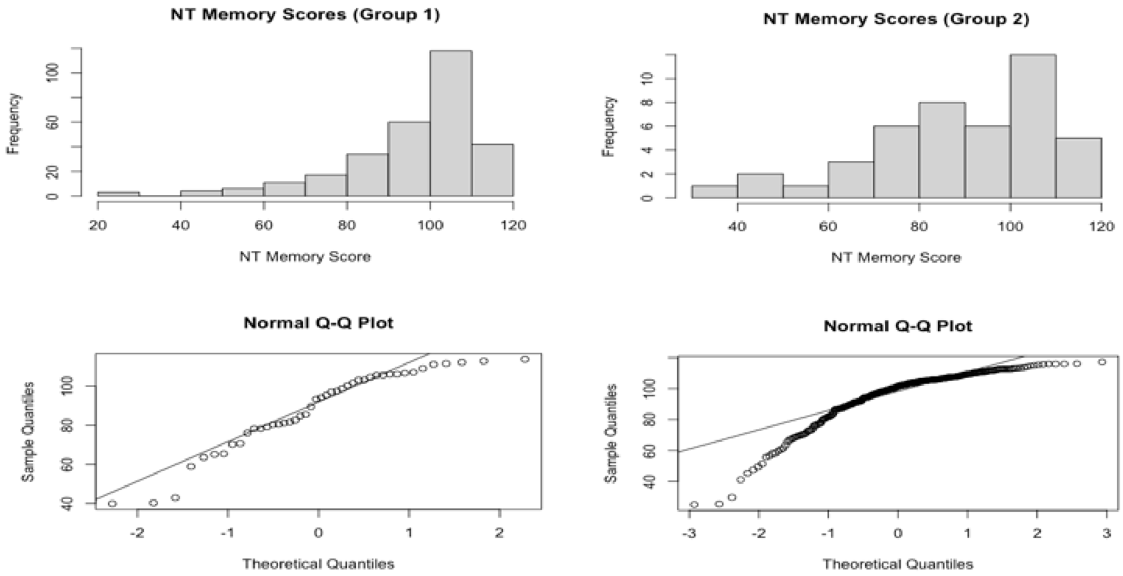
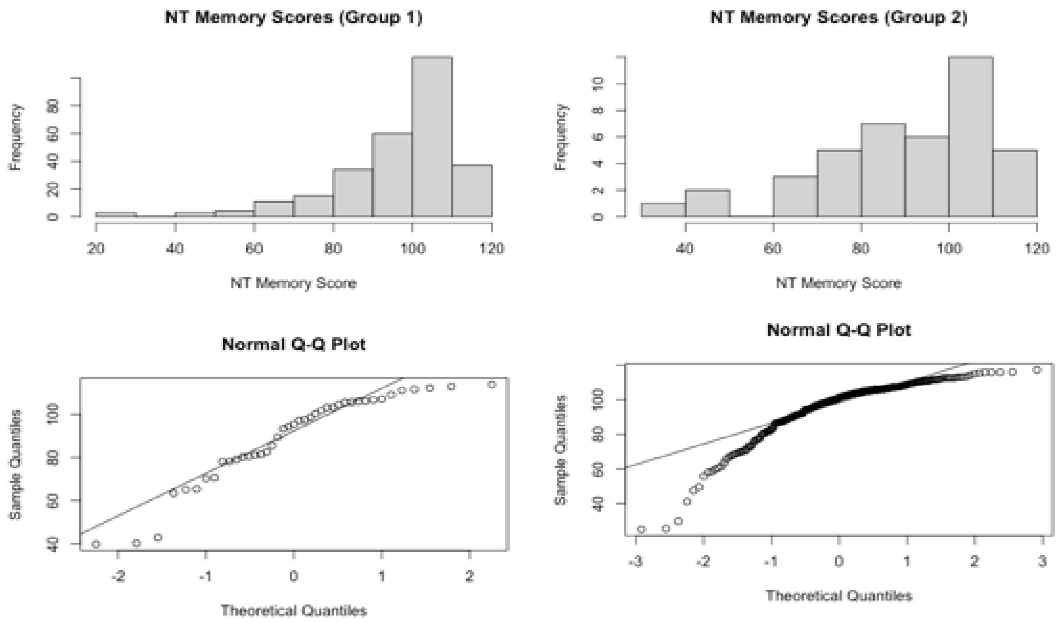


Figure 7

Histograms and Q-Q Plots of Memory scores excluding outliers



Shapiro-Wilk tests gave further information about the normality by testing the null hypothesis that residuals fit a normal distribution. The results of this test suggest the residuals are not normally distributed for the shared decision group, *W* = .847, *p* < .001, and for the solo decision group, *W* = .915, *p* = .003, when outliers are included. The low p-values indicate that the residuals' distribution is unlikely to follow a normal distribution. Similarly, excluding outliers, the Shapiro-Wilk test finds that the data remains non-normally distributed for the shared decision group, *W* = .844, *p* < .001, and for the solo decision group, *W* = .896, *p* = .001.

Assumption of equal variance:

First, the box plot of the visuospatial scores was visually inspected (see Figure 8 and 9). The variance in the box plot and whiskers did not appear very different.

Figure 8

Box plot of Memory scores including outliers

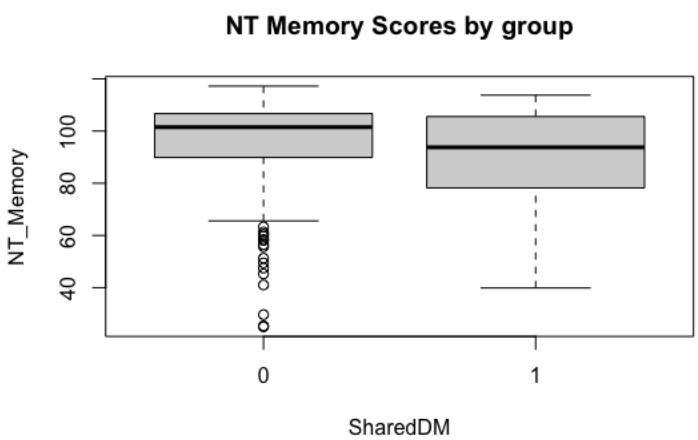
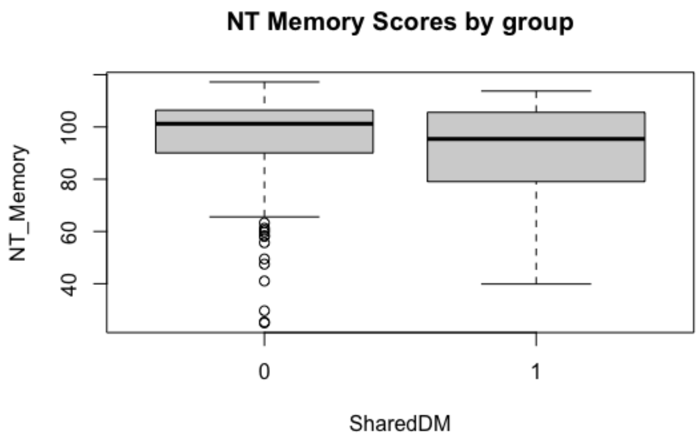


Figure 9

Box plot of Memory scores excluding outliers



To test the assumption of equal variances (Kim & Park, 2019), I conducted Levene's test using the Memory scores from Group 1 and Group 2. The results indicated that the variances of the two groups were not significantly different, $F(294, 43) = 0.66$, $p = 0.054$, with a 95% confidence interval ranging from 0.40 to 1.01. Therefore, the assumption of equal variance was met.

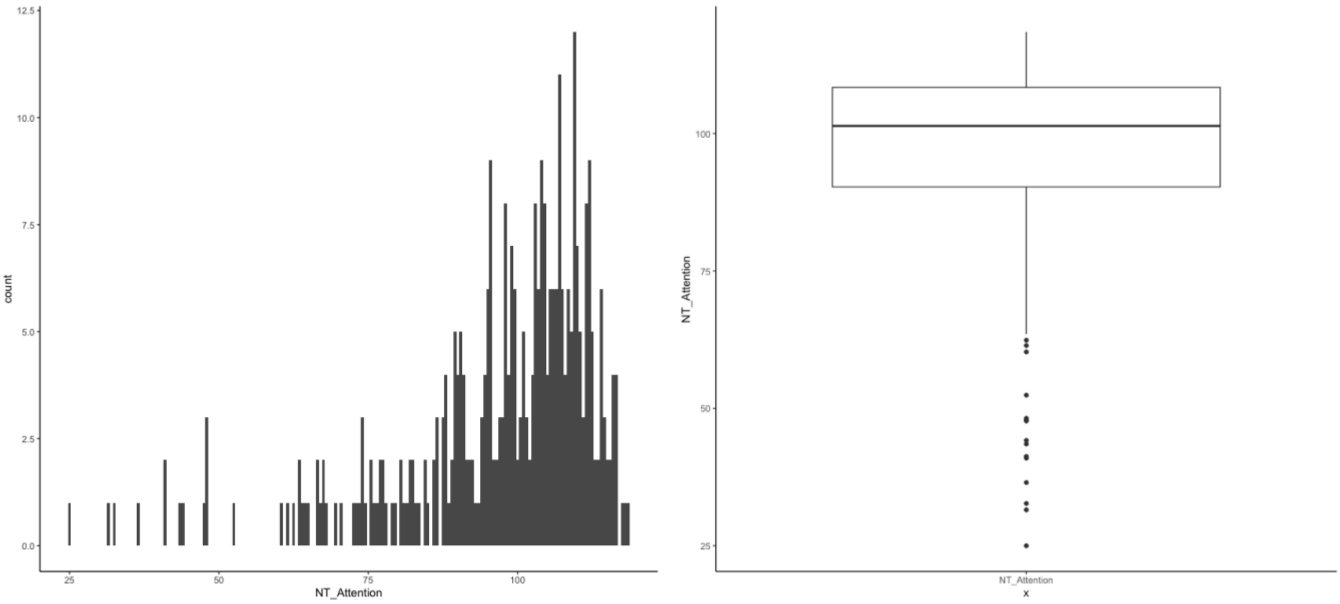
When outliers were excluded, the results indicated that the assumption of equal variance was violated. The results indicated that the variances of the two groups were significantly different, $F(281, 40) = 0.61$, $p = 0.02$, with a 95% confidence interval ranging from 0.36 to 1.01.

Decision making and attention:

Identifying outliers:

Figure 10

A histogram and boxplot of attention scores



Visual inspection of the histogram and the boxplot suggested that a few observations were lower than all other observations (see Figure 10). I then used the percentiles to determine outliers. Outliers were points that lie outside the 2.5th and 97.5th percentiles. From this, 18 outliers were identified. As I did not collect the data myself, I cannot be sure of the source of the outliers; therefore, I cannot justify removing them because of data entry or measurement errors. Secondly, removing these outliers did not influence the significance of the results of the Welch two-sample t-test, and they had minimal influence on the results. Excluding the cases did not signficantly alter the attention scores for each group, shared decision making group (*M*=97.58 including outliers, and

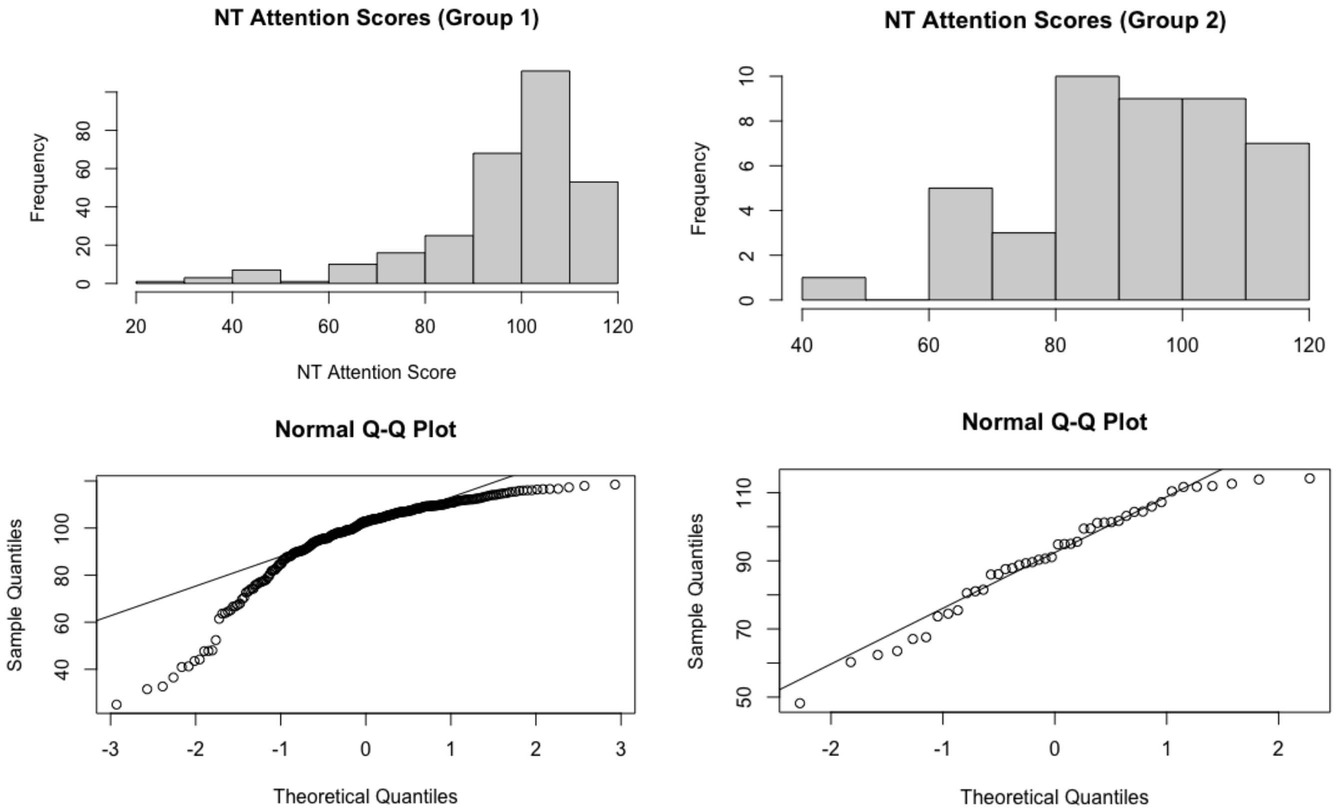
M=98.88 excluding outliers) and solo decision making group (*M*=91.34 including and excluding outliers). Nor did it impact the significance of the t-test ($t(57.45)=2.35, p=.022$ including outliers, $t(52.10)=2.91, p=.005$ excluding outliers). Therefore, the outliers were not removed before the t-test was conducted.

Assumption of normality:

The data and the residuals should be normally distributed (Kim & Park, 2019). First, plots were made visually inspect the data; the Q-Q plots suggested that the residuals were not normally distributed (see Figure 11).

Figure 11

Histogram and Q-Q plot for the attention scores of each decision making group (including outliers)



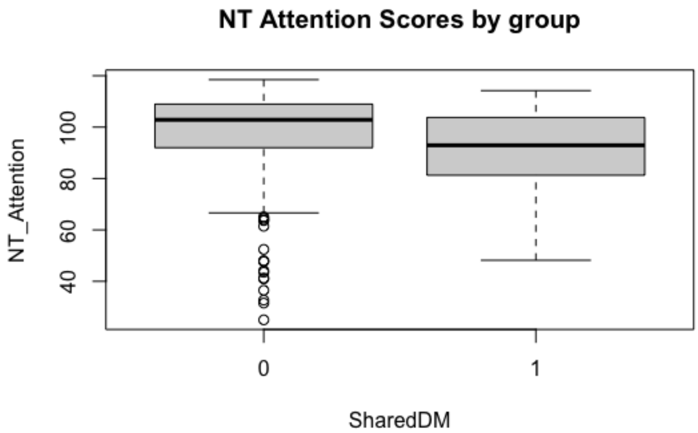
Shapiro-Wilk tests give further information about the normality by testing the null hypothesis that residuals fit a normal distribution. The results of this test suggest the residuals were not normally distributed for the shared decision group, $W = .826$, $p < .001$. The low p-values indicate that the residuals are unlikely to be normal distributed. However, the assumption of normality distribution is met for the solo decision group, $W = .950$, $p = .053$.

Assumption of equal variance:

First, the box plot of the visuospatial scores was visually inspected (see Figure 12). The variance in the box plot and whiskers appears similar. However, a statistical test is required to be sure.

Figure 12

Box plot of Attention scores including outliers



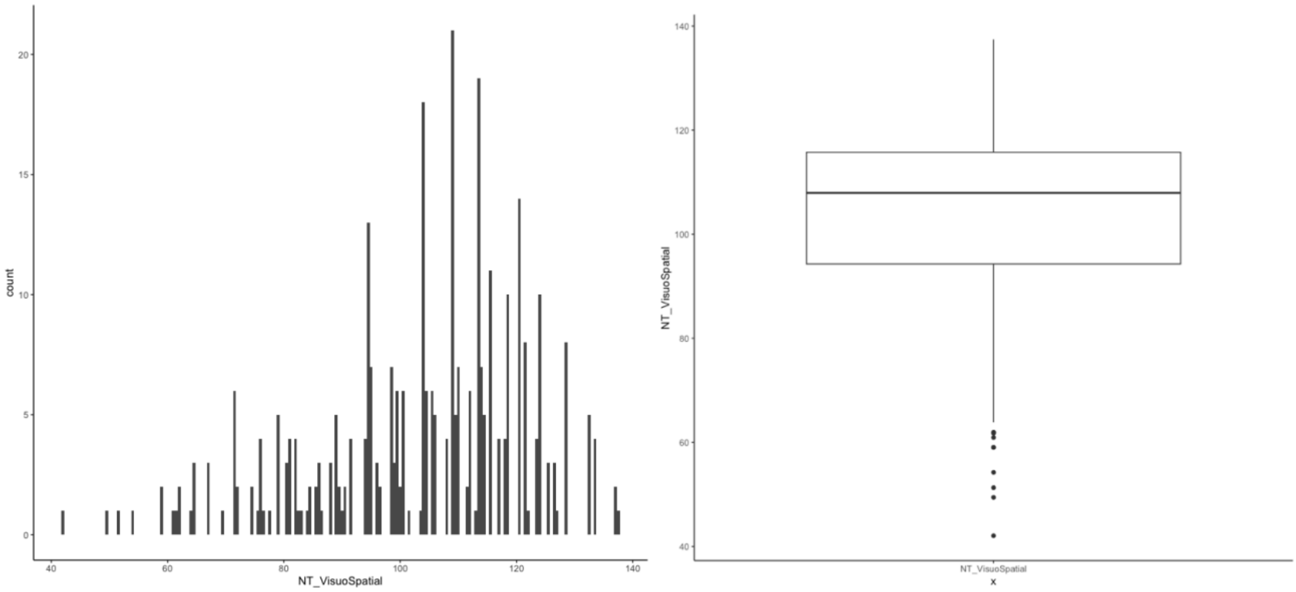
To test the assumption of equal variances (Kim & Park, 2019) I conducted Levene's test using the Attention scores from Group 1 and Group 2. The results indicated that the variances of the two groups were not significantly different, $F(294, 43) = 1.06$, $p = 0.85$, with a 95% confidence interval ranging from 0.65 to 1.61. Therefore, the assumption of equal variance was met.

Decision making and visuo-spatial ability:

Identifying outliers:

Figure 13

A histogram and boxplot of visuo-spatial scores



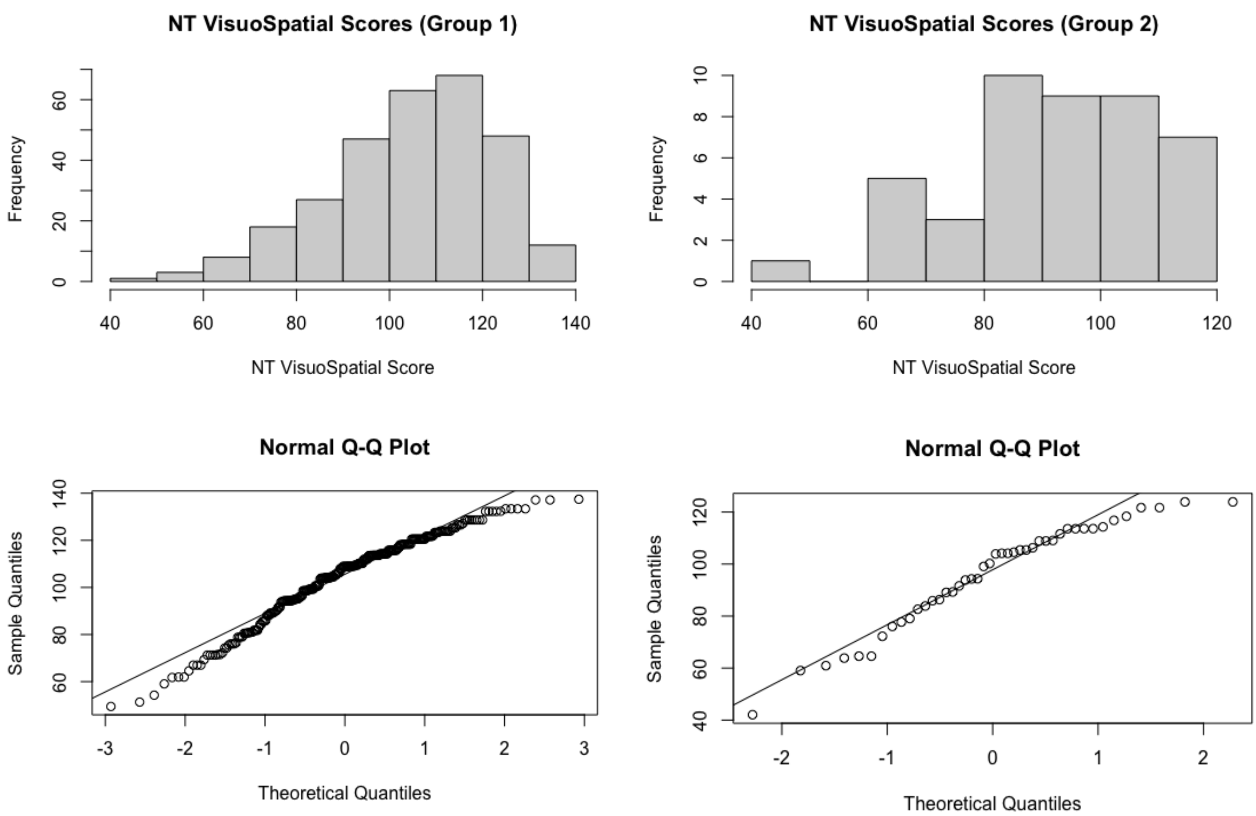
Visual inspection of the histogram and the boxplot suggests there may be observations lower and higher than all other observations (see Figure 13). I then used the percentiles to determine outliers, I considered the potential outliers to be all observations that lie outside the interval formed by the 2.5 and 97.5. From this, 18 outliers were identified. As the data was not collected by myself, I cannot be sure of the source of the outliers; therefore, I cannot justify removing them because of data entry or measurement errors. Secondly, removing these outliers did not influence the significance of the Welch two-sample t-test, and they had minimal influence on the results. Excluding the cases did not significantly alter the attention scores for each group, shared decision-making group (*M*=97.58 including outliers, and *M*=98.88 excluding outliers) and solo decision-making group (*M*=91.34 including and excluding outliers). Nor did it impact the significance of the t-test (*t*(57.45)=2.35,*p*=.022 including outliers, *t*(52.10)=2.91,*p*=.005 excluding outliers). Therefore, the outliers were not removed before the t-test was conducted.

Assumption of normality:

The data and the residuals should be normally distributed (Kim & Park, 2019). First, plots were visually inspected, and the Q-Q plots suggest the residuals were not normally distributed (see Figure 14).

Figure 14

Histograms and Q-Q plot for the VisuoSpatial scores of each decision making group (including outliers)



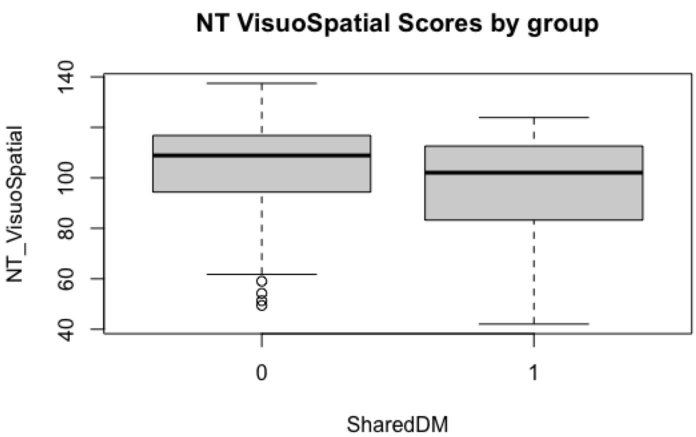
Shapiro-Wilk tests give further information about the normality by testing the null hypothesis that residuals fit a normal distribution. The results of this test suggest the residuals were not normally distributed for the shared decision group, $W = .963$, $p < .001$, and for the solo decision group, $W = .941$, $p = .027$. The low p-values indicate that the residuals were not normally distributed.

Assumption of equal variance:

First, the box plot of the visuo-spatial scores was visually inspected (see Figure 15). The variance in the box plot and whiskers appears similar. However, a statistical test is required to be sure.

Figure 15

Box plot of Visuospatial scores including outliers



To test the assumption of equal variances (Kim & Park, 2019), I conducted Levene's test using the VisuoSpatial scores from Group 1 and Group 2. The results indicated that the variances of the two groups were not significantly different, $F(294, 43) = 0.76$, $p = 0.19$, with a 95% confidence interval ranging from 0.46 to 1.15. Therefore, the assumption of equal variance was met.

Appendix B

Assumption testing for question 3

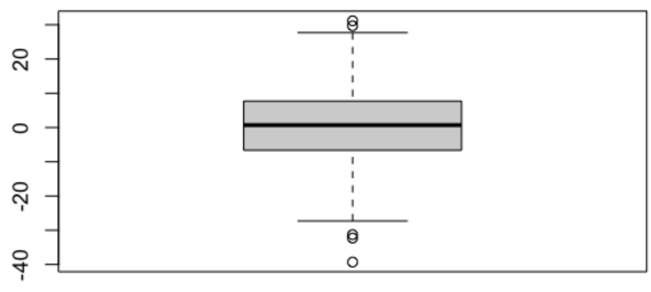
Model 1: single linear regression predicting illness perception from disease management

Identifying outliers

Outliers were identified using a boxplot (see Figure 16). Outliers were defined as participants with brief_illness scores higher than [the upper quartile + (1.5 x the inter-quartile range)] or lower than [the lower quartile - (1.5 x the inter-quartile range)]. Three participants were outliers (participants 7, 111, and 173). However, these outliers did not significantly affect the first or second linear regression results nor whether assumptions were met and therefore they were not removed from the data.

Figure 16

A box plot of brief_illness scores



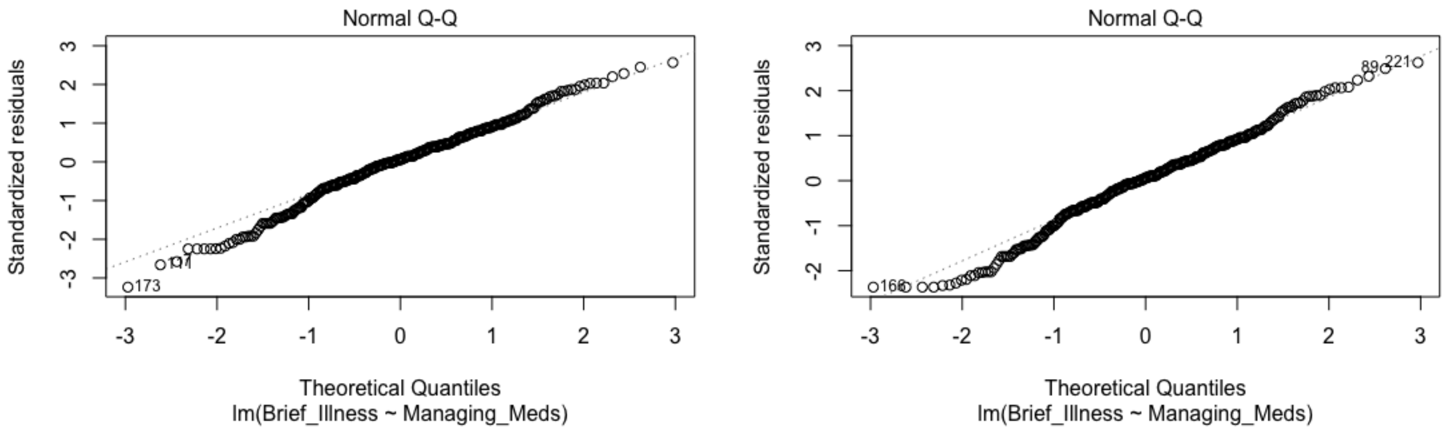
Several assumptions need to be met when conducting a linear regression (Uyanık & Güler, 2013). Assumption tests were conducted, including and excluding outliers, to determine their impact on the assumptions.

Normal distribution:

In linear regression, residuals should be normally distributed (Uyanık & Güler, 2013). Q-Q plots of standardised residuals against theoretical quantiles of a normal distribution are used to indicate whether this assumption is met. The Q-Q plots of the models (with and without outliers) showed that the data was fairly normally distributed; however, there were deviations at the extremes of the scale, where the points moved away from the diagonal line (see Figure 17).

Figure 17

Q-Q plots of the multiple linear regression when outliers are included and excluded respectively



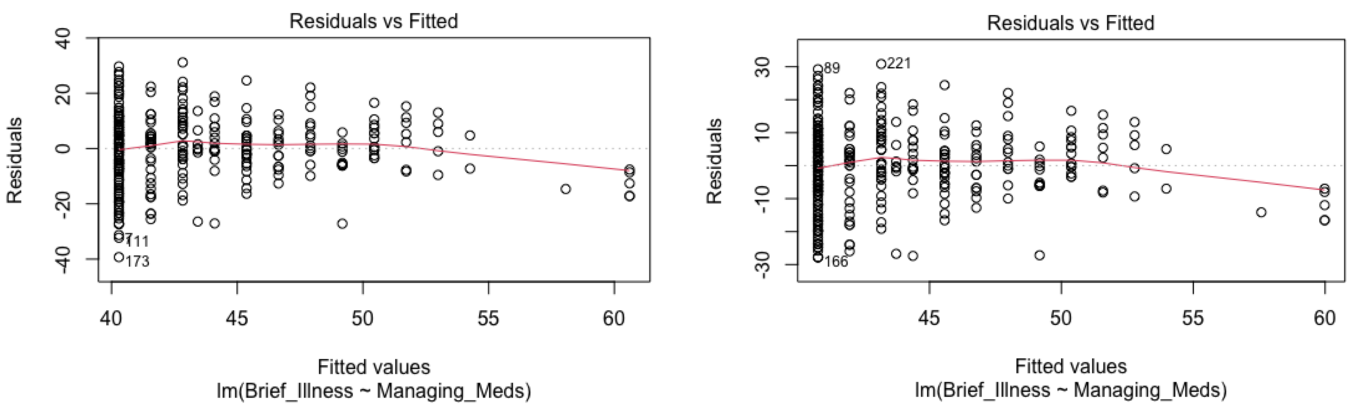
Shapiro-Wilk tests analyse the null hypothesis that the residuals fit a normal distribution. This test found that the residuals were not normally distributed for the model when outliers were included, W = .988, p = .009, and excluded, W = .988, p = .008. Low p-values indicate that the residuals are not normally distributed.

Linearity

Another assumption of linear regression is that there is a linear relationship between the predictors and the response variable (Uyanık & Güler, 2013). The assumption of linearity was tested for the model (see Figure 18). For data including and excluding outliers, no linear relationship was detected; however, the scatterplot also did not suggest a non-linear relationship between the variables. Therefore, a linear model was still conducted (Howitt & Cramer, 2011).

Figure 18

Scatter plots of the linear regression when outliers are included and excluded respectively

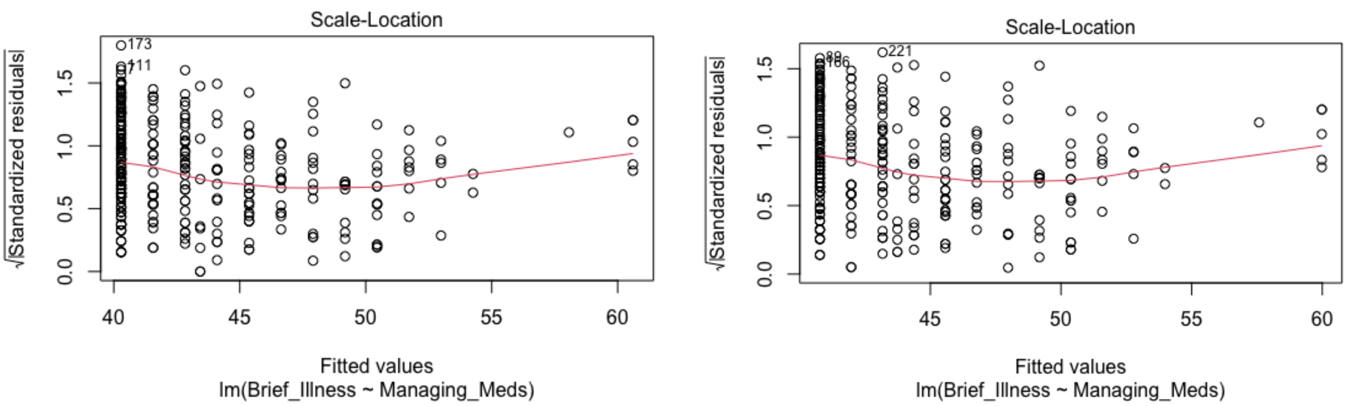


Homoscedasticity/Homogeneity of residual variance

The assumption of homoscedasticity implies equal variances of residuals for all conditions (Tranmer et al., 2020). First, scale-location plots were conducted for the multiple linear regression, including, and excluding the outliers (see Figure 19). In these plots, the pattern does not appear random, as you would expect; however, there are roughly equal plots above and below the centre line.

Figure 19

Scale-location plot of the linear model including and excluding outliers, respectively



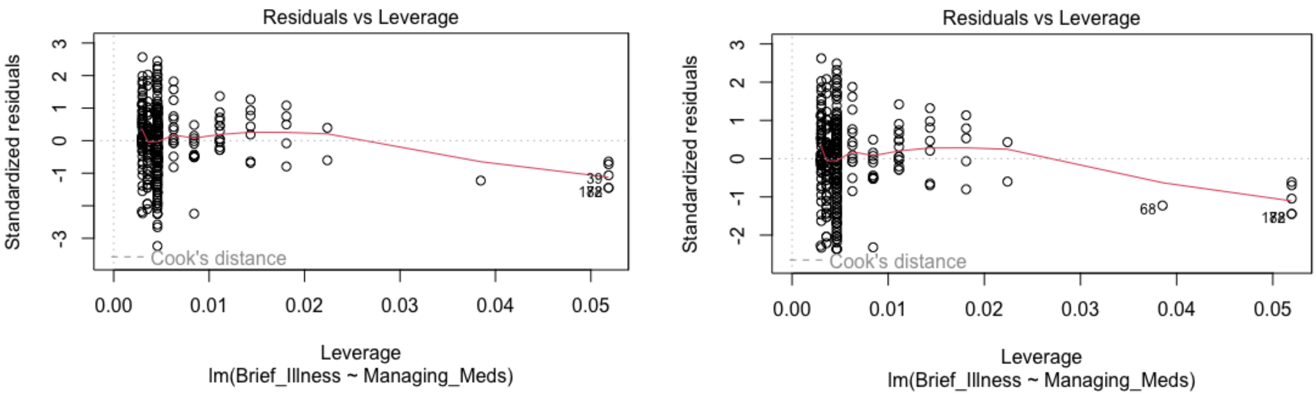
This was further investigated with the Fligner-Killeen test. Results of this test found that the variances of residuals are not equal across the fitted values for the dataset, including outliers (Fligner-Killeen(15) = 31.10, p = .008) and excluding outliers (Fligner-Killeen(16) = 30.28, p = .017). Therefore, the assumption was violated.

Testing for other influential data points

I investigated if there were any influential data points in the model after outliers had been excluded. The residuals vs. leverage plots show that all standardised residuals were between +3 and -3, and none of the points have a high residual, indicating no outlying residual values (Zhang, 2016) (see Figure 20).

Figure 20

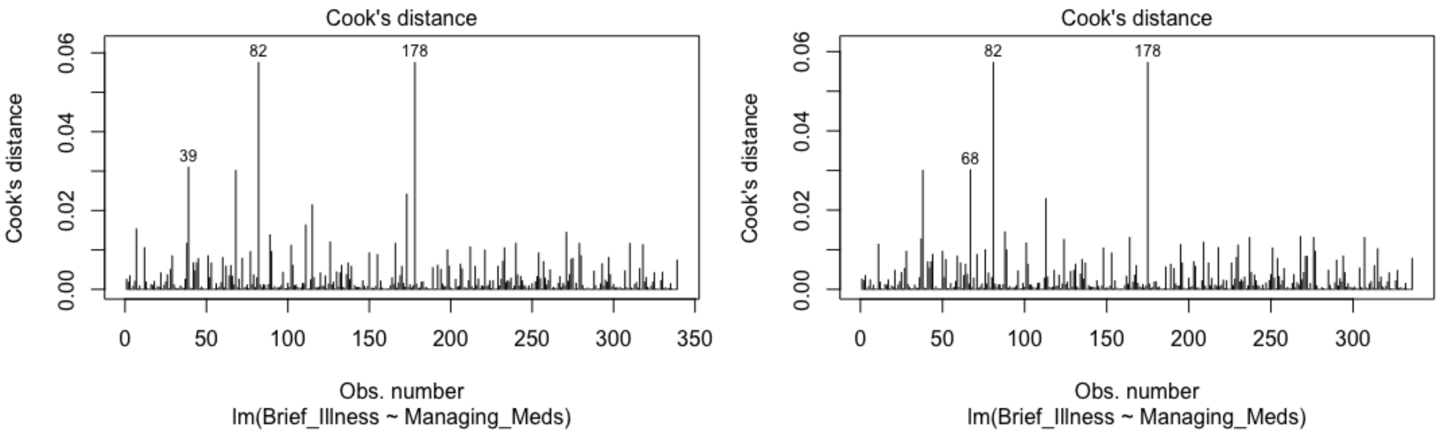
Residuals vs leverage plot for the linear regression analysis with and without outliers



Cook's distance was used to investigate the assumption further (see Figure 21). Results found that Cook's distance was below 0.5 for all points and therefore, not considered influential (PennState, 2018).

Figure 21

Cooks distance plot for the linear regression analysis with and without outliers



Why did I conduct a linear regression if the assumptions were violated?

Despite the assumption of normality was violated, I still decided to conduct the linear model. This is because central limit theorem states that if a sample is larger than above 30, then the samples' means will be normally distributed (Pek et al., 2018). Consequently, in large samples like those found in this data, normal distribution is not necessary for regression (Pek et al., 2019). Additionally, although no linear relationship was detected, the scatterplot also did not suggest a non-linear relationship, therefore, the linear model was still appropriate (Howitt & Cramer, 2008). Lastly, the assumption of equal variance was violated; however, unless the heteroscedasticity is pronounced, it's effects will not be severe (Stat Guide, 2022), therefore, I continued with the linear regression.

Assumption testing for model 2: the multiple linear regression

The assumptions for the multiple linear regression are similar to those of the single regression above, however they also assume multicollinearity.

Outliers:

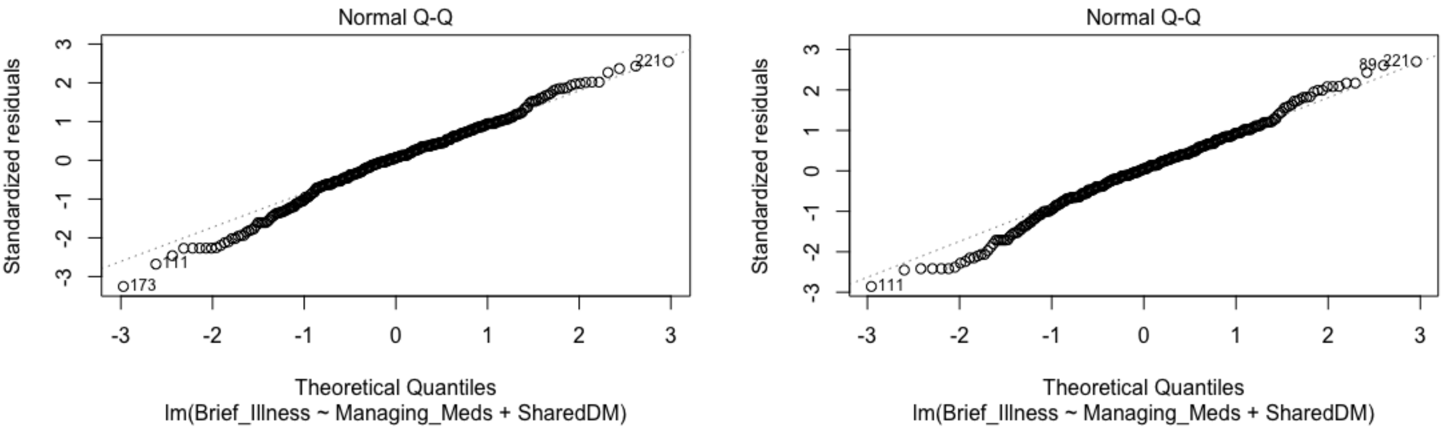
I defined outliers as observations with a Cook's distance greater than 4 divided by the number of rows in the data frame (Statology, 2019). However, once these outliers were removed, they did not influence the results of the model significantly, or the results of the assumption tests. Additionally, as I did not conduct the research myself, I cannot know if the outliers occurred due to methodological or data input errors. Therefore, I kept these outliers in.

Normal distribution:

In a linear regression, residuals should be normally distributed (Uyanık & Güler, 2013). Q-Q plots of standardised residuals against theoretical quantiles of a normal distribution were used to indicate whether this assumption is met. The Q-Q plots of the models (with and without outliers) showed that the data was quite normally distributed, however, there were deviations at the extremes of the scale, where the points move away from the line (see Figure 22).

Figure 22

Q-Q plots of the multiple linear regression when outliers are included and excluded respectively



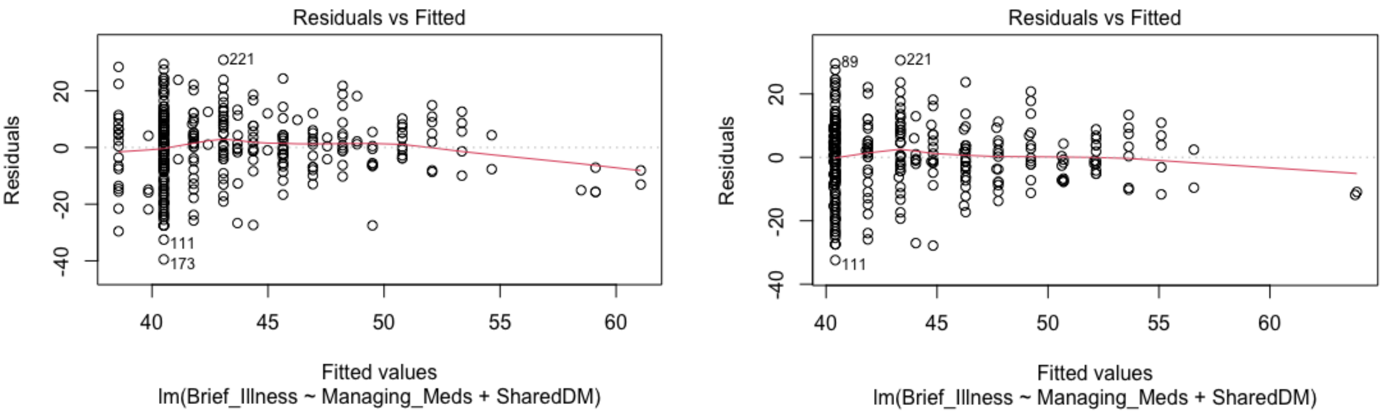
Shapiro-Wilk tests analyse the null hypothesis that the residuals fit a normal distribution. This test found that the residuals were not normally distributed for model when outliers were included, *W* = .989, *p* = .014, and excluded, *W* = .990, *p* = .022. Low p-values indicate that the residuals do not fit a normal distribution.

Linearity

Another assumption of linear regression is that there is a linear relationship between the predictors and the response variable (Uyanık & Güler, 2013). The assumption of linearity was tested in the multiple linear regression model (see Figure 23). No linear relationship was detected; however, the scatterplot did not suggest a non-linear relationship between the variables. Therefore, a linear model was still conducted (Howitt & Cramer, 2008).

Figure 23

Scatter plots of the multiple linear regression model when outliers are included and excluded respectively

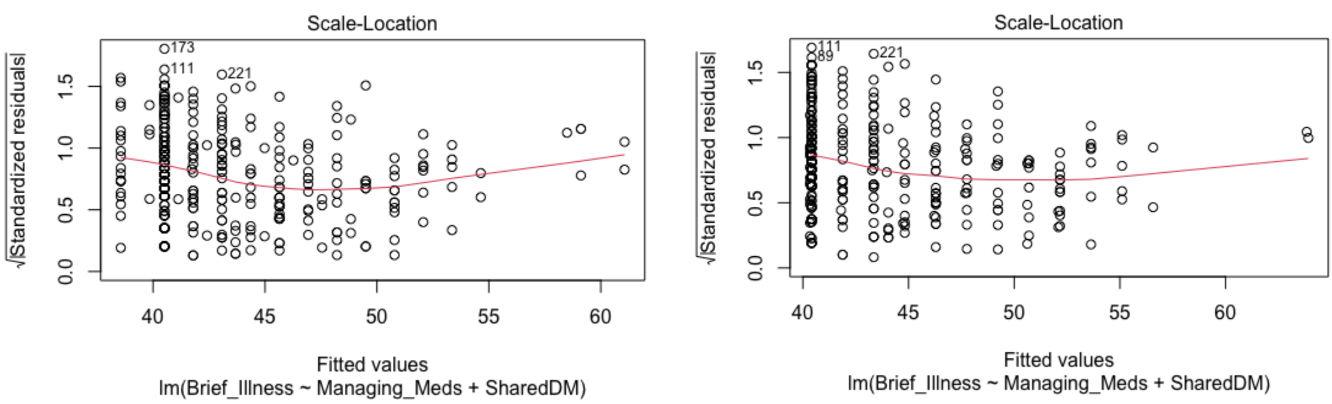


Homogeneity:

First, a scale-location plot was conducted for the multiple linear regression, including and excluding the outliers, respectively. In these plots, the pattern does not appear random (see Figure 24). However, there are roughly equal plots above and below the centre line.

Figure 24

Scale-location plot of the linear model for data with and without outliers



The results of this test found that the variances of residuals were equal across the fitted values for the dataset, including outliers (Fligner-Killeen(26) = 37.23, p = .070). However, when excluding outliers, the assumption of equal variances was violated (Fligner-Killeen(27) = 30.63, p = .29).

Multicollinearity

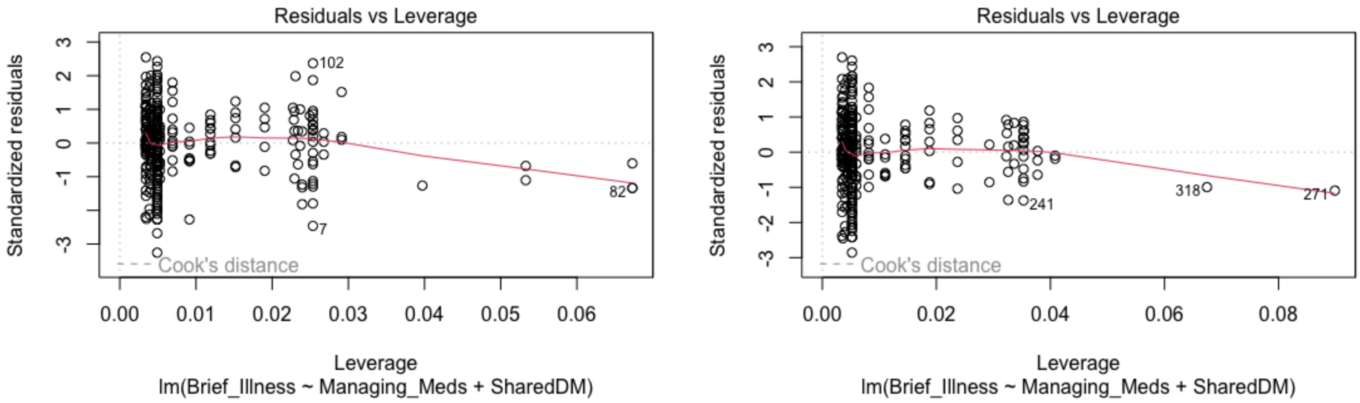
The assumption of collinearity suggests that the independent variables in a model should not be correlated (Tranmer et al., 2020). The VIF values for disease management and shared decision-making are close to 1. Therefore, multicollinearity is not a problem in the model when outliers are included (*VIF* = 1.005) and excluded (*VIF* = 1.006).

Checking for influential data points:

I investigated if there were any influential data points in the model after outliers had been excluded. The residuals vs leverage plots (see **Figure 25**) show that all standardised residuals were between +3 and -3, and none of the points have a high residual, indicating no outlying residual values (Zhang, 2016).

Figure 25

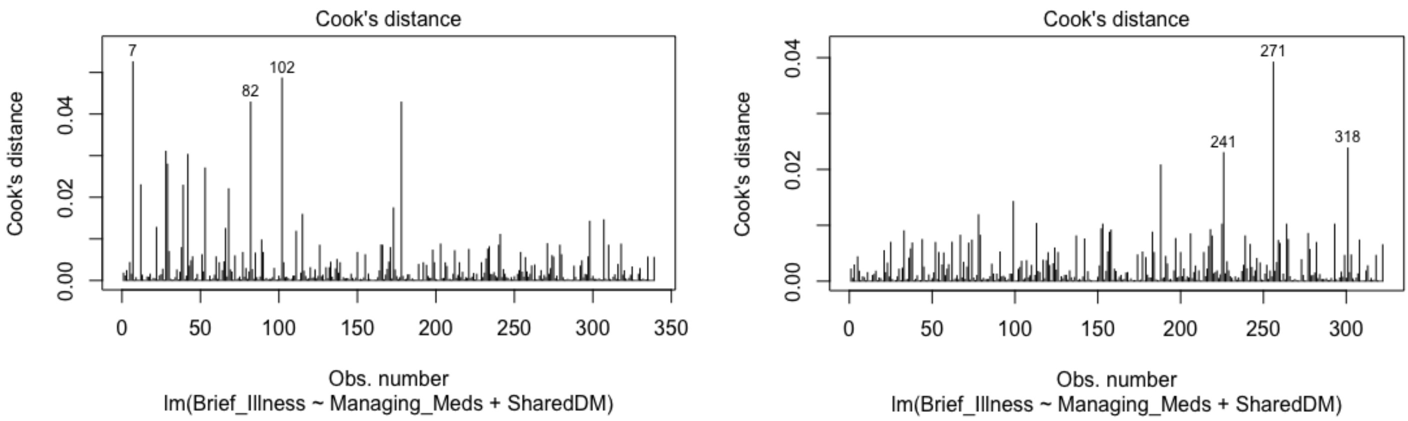
Residuals vs leverage plot for the multiple linear regression analysis with and without outliers



Cook’s distance (see Figure 26) was used to investigate this further. Results found that Cook’s distance was below 0.5 for all points and therefore not considered influential (PennState, 2018).

Figure 26

Cooks distance plot for the multiple linear regression analysis with and without outliers



Part B: Predicting academic performance in students

Academic performance is correlated with stress and sleep quality (Ahrberg et al., 2012). This may be due to sleep deprivation harming cognitive ability (Norbury & Evans, 2019), while stress may cause students to be distracted or unable to perform to their best ability (Frazier et al., 2018). Additionally, depression is also correlated with low academic performance (Singh & Masih, 2022). These issues are essential to investigate as depression has become more prevalent among higher-education students, while students also sleep less than the recommended amount and have high stress levels (Lund et al., 2010; Asif et al., 2020; Beiter et al., 2015). Since these factors can influence academic performance, I hypothesise that students' academic performance can be predicted based on their stress level, depression level and average sleep duration, using data from the Student Life Publication (Wang et al., 2014).

A mean sleep variable was created to test the hypothesis, and the ordinal independent variables (stress and depression) were coded as dummy variables (Appendix C). Outliers were detected, and linear regression assumptions were tested and explained (Appendix C).

Multiple linear regression was conducted to examine the relationship between GPA scores and the predictor variables (depression, stress, and sleep duration). Results are shown in Table 3.

The regression equation was insignificant, suggesting that there was not a significant relationship between the predictor variables and GPA ((Adjusted $R^2 = 0.281$, F(6, 21)=1.37, p=.273. The F value is relatively low (F=1.367), suggesting the model is not a good fit. The multiple R-squared value indicates that the predictors can explain about 28% of the variance in GPA scores. The model's predictive ability may be weak, as the difference between the multiple R-squared and adjusted R-squared is large, indicating that the model may be overfitting. Overall, the model may not be very useful for predicting students' GPA scores because there was no relationship between GPA scores and two predictors, sleep duration and stress levels.

However, the model found that higher GPA scores could be predicted by having no depression or mild depression. On average, GPA scores are expected to increase by 1.14 points for students with no depression compared to students with some level of depression, holding other variables constant. Similarly, on average, GPA scores are expected to increase by 1.03 points for students with mild depression compared to students with some level of depression, holding other variables constant.

Table 3.
Multiple linear regression results

Predictor	**Beta**	**t-value**	**p-value**
(Intercept)	2.4597§	2.75	.0118*
no_depression	1.1349	2.14	.0441*
mild_depression	1.0337	2.12	.0465*
moderate_depression	.6491	1.39	.1789
meansleep	.0518	.44	.6646
low_stress	-.3629	-.84	.4088
medium_stress	-.4650	-1.38	.1816
Adjusted R^2		.075	
F		1.367	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix C

Measures of variables

Mean Sleep Duration:

The mean sleep duration variable was created from the sleep hours variable, which was the number of hours participants slept the previous night. Mean sleep was calculated by averaging the sleep hours for each participant over the 55 days (*M*= 7.19, *SD* = 0.88, Range = 4.90-9.29).

Grade point average:

GPA is calculated on a 4-point scale, from 0 (a fail) to 4 (an A grade). GPA was measured cumulatively to assess the academic achievement of each student (*M*= 3.44, *SD* = 0.40, Range = 2.40-3.95).

Dummy variables:

Although there are some issues with using dummy variables and considerations should be made (Blankmeyer, 2006), I decided to compute dummy variables as it allowed me to conduct the linear regression model with the independent variables for depression and stress that were ordinal measures. It is important to use dummy variables instead of treating the ordinal variables as continuous (Agresti, 2012), because the difference between points on a scale may not be equal. This is important for the measures for depression and stress as they were created by accumulating the scores from Likert Scale questionnaires. For example, when asked to rate levels of depression from 1 to 7, the difference between 1 and 2 may not be equal to the difference between 5 and 6. Creating these dummy variables allows the relationship to be studied more accurately (Winship & D. Mare, 1981). The number of patients in each of the dummy variables is shown in Table 4.

Depression:

Depression was measured in the dataset by the Patient Health Questionnaire, which creates scores between 0 and 30. A score from 0-5 represents no depression, 5-9 represents mild depression, 10-14 represents moderate depression, and 15-30 represents severe depression (Bernard & Daros, 2018). Thus, dummy variables were created using these thresholds.

Stress:

Stress was measured using the Perceived Stress Scale, an accumulation of Likert scale ratings, with higher scores representing higher stress. Within each Likert scale question, a score of zero represented never feeling stressed, while four represented feeling stressed often. Therefore, dummy variables were created by splitting each group equally into three. Scores of 0-14 represented low stress, 15-24 represented medium stress, and 25-35 represented high stress.

Table 4

The number of cases in each of the dummy variables

Dummy variable	Number of students
No depression	23
Mild depression	16
Moderate depression	6
Severe depression	1
Low stress	16
Medium stress	24
High stress	6

Appendix D

Assumption testing and outlier detection:

Outliers:

Figure 27

A boxplot to show the residuals of the multiple regression model

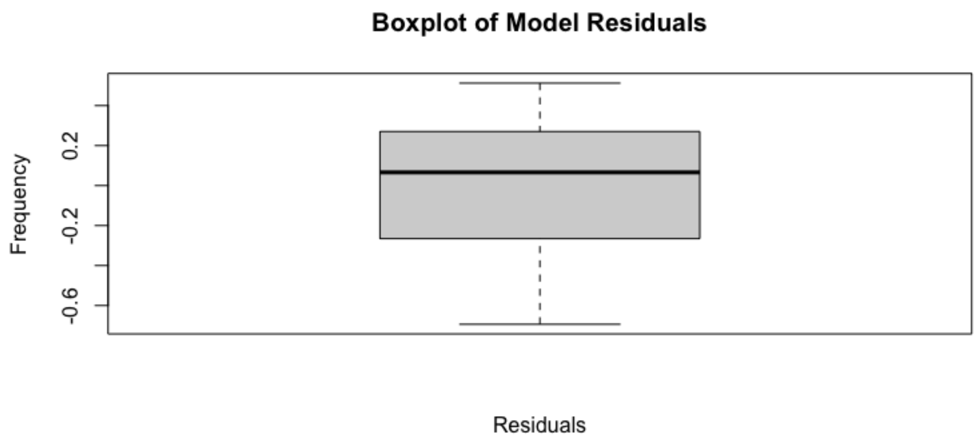
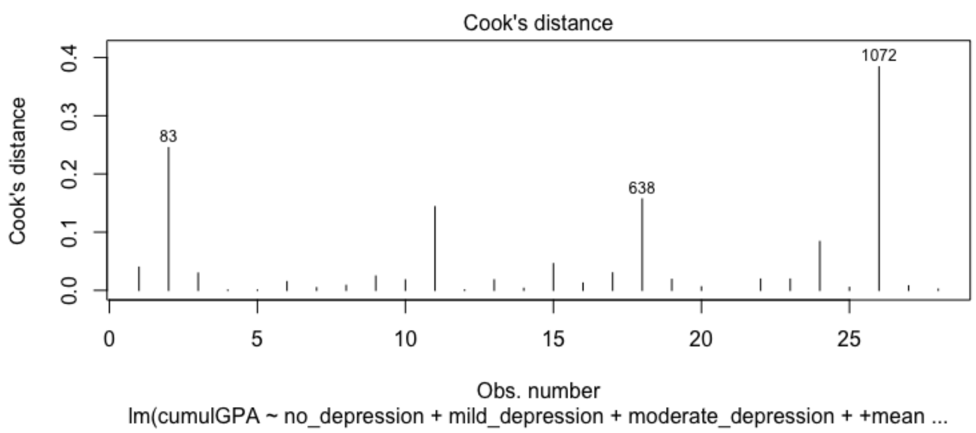


Figure 28

A Cook's distance graph of the multiple regression model



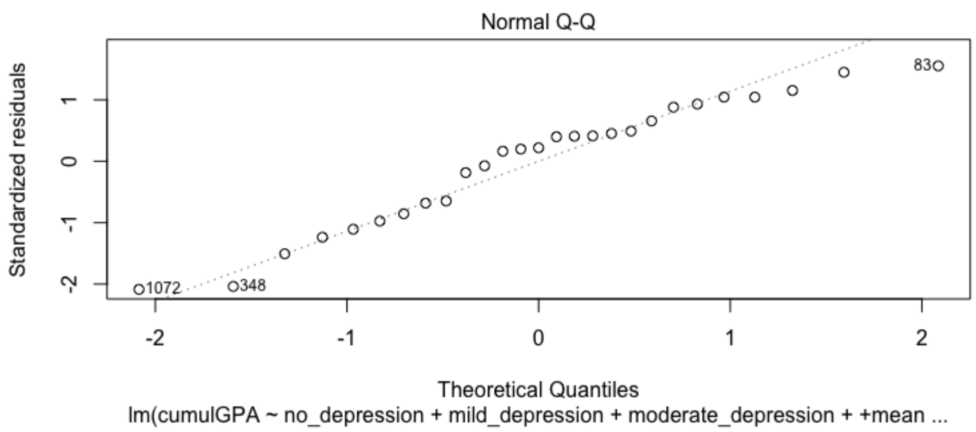
Visual inspection of a box plot of the residuals suggests no outliers in the data (see Figure 27). Categorical data cannot have outliers by definition; therefore, the only variable that could have outliers is mean sleep duration. However, the Cooks distance graph (see Figure 28) showed that all points were significantly below 1. Therefore, all the points were considered to have an insignificant impact, and no outliers were detected.

Assumption of normality:

In linear regression, residuals should be normally distributed. Q-Q plots of standardised residuals against theoretical quantiles of a normal distribution are used to indicate whether this assumption is met. The Q-Q plot of the model showed that the data was fairly normally distributed as the plots were close to the line (see Figure 29).

Figure 29

A Q-Q plot showing the quantiles of the sample data to the quantiles of a theoretical distribution.



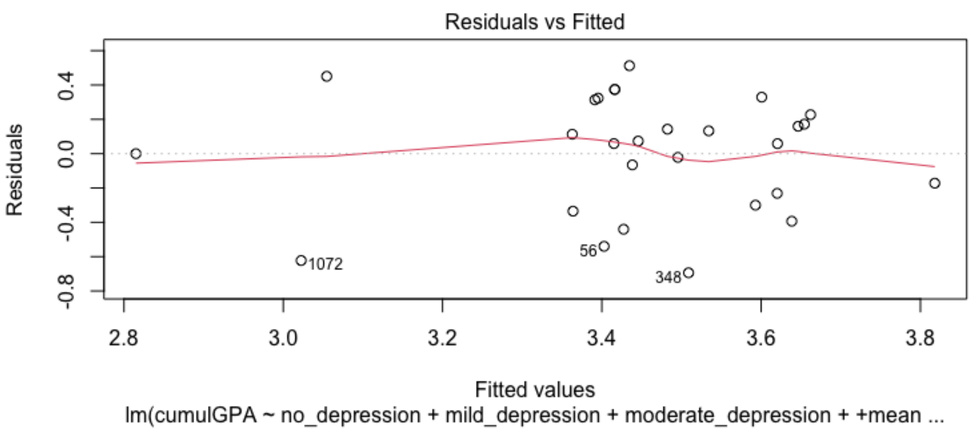
This was investigated further, and a Shapiro-Wilk test was used to analyse the null hypothesis that the residuals fit a normal distribution. This test found that the residuals may be normally distributed W= .950, p = .192. As the p-value is greater than 0.05, then the null hypothesis cannot be rejected, and the variable may be normally distributed.

Assumption of homoscedasticity

I could not conduct the Levene test, as the model uses dummy variables that are not independent; therefore, I assessed residuals vs fitted model (see Figure 30). The residuals have a fairly equal variance across the fitted values. Although the residuals are not equally distributed throughout the graph, there is no strong evidence of residuals fanning out. This suggests that the variances of the error terms are equal; it is likely that the assumption of equal variance has not been violated.

Figure 30

A residuals vs fitted plot of the residuals in the model



Assumption of linearity:

The assumption of linearity states that the effect of each predictor on the outcome should not be non-linear. For a dummy variable with multiple categories, this assumption is trivially met, since the line of best fit connects the conditional means of the categories (Hardy, 1993).

Assumption of multicollinearity:

VIF tests were conducted to assess multicollinearity. The variance inflation factors (VIFs) for the predictor variables were as follows: no_depression (VIF = 13.58), mild_depression (VIF = 10.65), moderate_depression (VIF = 5.19), meansleep (VIF = 1.90), low_stress (VIF = 7.86), and medium_stress (VIF = 5.48). These VIF scores suggest moderate to severe multicollinearity in the model. However, this is expected as the models use indicator dummy variables. This is especially the case because the number of cases in the reference categories (severe depression and high stress) is small, the indicator variables will necessarily have high VIFs, even if the categorical variable is not associated with other variables in the regression model (Hardy, 1993).

References

Agresti, A. (2012). *Analysis of ordinal categorical data*. John Wiley & Sons.

Ahrberg, K., Dresler, M., Niedermaier, S., Steiger, A., & Genzel, L. (2012). The interaction between sleep quality and academic performance. *Journal of Psychiatric Research*, *46*(12), 1618–1622. https://doi.org/10.1016/j.jpsychires.2012.09.008

Alotaibi, A. D., Alosaimi, F. M., Alajlan, A. A., & Bin Abdulrahman, K. A. (2020). The relationship between sleep quality, stress, and academic performance among medical students. *Journal of Family & Community Medicine*, *27*(1), 23–28. https://doi.org/10.4103/jfcm.JFCM_132_19

Asif, S., Muddassar, A., Shahzad, T. Z., Raouf, M., & Pervaiz, T. (2020). Frequency of depression, anxiety and stress among university students. *Pakistan Journal of Medical Sciences*, *36*(5). https://doi.org/10.12669/pjms.36.5.1873

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The Prevalence and Correlates of depression, anxiety, and stress in a sample of college students. *Journal of Affective Disorders*, *173*(1), 90–96. https://doi.org/10.1016/j.jad.2014.10.054

Bernard, K., & Daros, A. (2018). *Relationships between daily sleep ratings, well-being, and academic success over the course of a college semester: A StudentLife study*.

Blankmeyer, E. (2006). *How robust is linear regression with dummy variables ?* https://files.eric.ed.gov/fulltext/ED493866.pdf

Frazier, P., Gabriel, A., Merians, A., & Lust, K. (2018). Understanding stress as an impediment to academic performance. *Journal of American College Health*, *67*(6), 562–570. https://doi.org/10.1080/07448481.2018.1499649

Hardy, M. (1993). *Regression with dummy variables*. SAGE Publications, Inc. https://doi.org/10.4135/9781412985628

Howitt, D., & Cramer, D. (2008). *Introduction to research methods in psychology*. Prentice Hall.

Howitt, D., & Cramer, D. (2011). *Introduction to research methods in psychology third edition introduction to research methods in psychology*. https://aishwaryajaiswal.com/wp-content/uploads/2022/01/Introduction-to-research-Methods-in-Psychology-3rd-ed.-D.-Howitt-D.-Cramer-Pearson-2011-BBS.pdf

Kim, T. K., & Park, J. H. (2019). More about the basic assumptions of t-test: Normality and sample size. *Korean Journal of Anesthesiology*, *72*(4), 331–335. https://doi.org/10.4097/kja.d.18.00292

le Cessie, S., Goeman, J. J., & Dekkers, O. M. (2020). Who is afraid of non-normal data? Choosing between parametric and non-parametric tests. *European Journal of Endocrinology*, *182*(2), E1–E3. https://doi.org/10.1530/eje-19-0922

Lund, H. G., Reider, B. D., Whiting, A. B., & Prichard, J. R. (2010). Sleep patterns and predictors of disturbed sleep in a large population of college students. *Journal of Adolescent Health*, *46*(2), 124–132. https://doi.org/10.1016/j.jadohealth.2009.06.016

Norbury, R., & Evans, S. (2019). Time to think: Subjective sleep quality, trait anxiety and university start time. *Psychiatry Research*, *271*, 214–219. https://doi.org/10.1016/j.psychres.2018.11.054

Pek, J., Wong, O., & Wong, A. (2019). Data transformations for inference with linear regression: Clarifications and recommendations. *Practical Assessment, Research, and Evaluation*, *22*(1). https://doi.org/10.7275/2w3n-0f07

Pek, J., Wong, O., & Wong, A. C. M. (2018). How to address non-normality: A taxonomy of approaches, reviewed, and illustrated. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02104

PennState. (2018). *9.5 - Identifying influential data points | STAT 462*. Online.stat.psu.edu. https://online.stat.psu.edu/stat462/node/173/

Samanez-Larkin, G. (2019). Sleep and academic performance in college students. *Osf.io*. https://osf.io/5egfq/

Singh, R., & Masih, P. (2022). An Overview of the Relation between Depression - ProQuest. *IAHRW International Journal of Social Sciences Review*, *10*(3). https://www.proquest.com/docview/2723857532/44B659BB5B14403PQ/1?accountid=17230

Stat Guide. (2022). *Does your data violate linear regression assumptions?* Www.quality-Control-Plan.com. https://www.quality-control-plan.com/StatGuide/linreg_ass_viol.htm#Nonconstant%20variance%20of%20Y

Statology. (2019, May 11). *How to identify influential data points using cook's distance - Statology*. Statology. https://www.statology.org/how-to-identify-influential-data-points-using-cooks-distance/

Tranmer, M., Murphy, J., Elliot, M., & Pampaka, M. (2020). *Multiple linear regression (2 nd edition)*. http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/multiple-linear-regression.pdf

Uyanık, G. K., & Güler, N. (2013). A study on multiple linear regression analysis. *Procedia - Social and Behavioral Sciences*, *106*(1), 234–240. https://doi.org/10.1016/j.sbspro.2013.12.027

Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., & Ben-Zeev, D. (2014). StudentLife. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp ’14 Adjunct*. https://doi.org/10.1145/2632048.2632054

Winship, C., & D. Mare, R. (1981). Regression model with ordinal variables. *American Sociological Review*, *49*, 512–525. https://scholar.harvard.edu/files/cwinship/files/asr_1984.pdf

Wolfson, A. R., & Carskadon, M. A. (2003). Understanding adolescent’s sleep patterns and school performance: A critical appraisal. *Sleep Medicine Reviews*, *7*(6), 491–506. https://doi.org/10.1016/s1087-0792(03)90003-7

Zhang, Z. (2016). Residuals and regression diagnostics: Focusing on logistic regression. *Annals of Translational Medicine*, *4*(10), 195–195. https://doi.org/10.21037/atm.2016.03.36