

Postgraduate Science Framework

Department of Science

School of Health and Life Sciences

**NKAEPE EDEM BASSEY-ASUQUO**

**C2766519**

**DATA ANALYSIS**

**SCI4019-N**

**ACADEMIC LEVEL 7**

**20 CREDITS**

**FINAL ASSESSMENT**

**SEMESTER 1**

**2023 / 24**

**19TH JANUARY, 2024**

**TABLE OF CONTENTS**

**TITLE PAGE1**

**TABLE OF CONTENT2**

**DATA SET 1 (BABIES)5**

**ABSTRACT5**

**1.0 INTRODUCTION AND BACKGROUND 5**

**1.1 OBJECTIVES 5**

**2.0 METHODS 5**

**3.0 RESULTS 6**

**3.1 DESCRIPTIVE ANALYSIS 7**

**3.2 EXPLORATORY ANALYSIS 9**

**3.1 INDEPENDENT T-TEST 14**

**4.0 DISCUSSION15**

**5.0 CONCLUSION 17**

**DATA SET 2 (SURGERY)18**

**ABSTRACT18**

**1.0 INTRODUCTION AND BACKGROUND 18**

**1.1 RESEARCH QUESTIONS18**

**1.2 OBJECTIVES 19**

**2.0 METHODS 19**

**3.0 RESULTS 20**

**3.1 DESCRIPTIVE ANALYSIS 21**

**3.2 EXPLORATORY ANALYSIS 24**

**3.3. NON-PARAMETRIC TEST (MANN U-TEST) 27**

**4.0 DISCUSSION28**

**5.0 CONCLUSION 29**

**DATA SET 3 (WEIGHT)30**

**ABSTRACT30**

**1.0 INTRODUCTION AND BACKGROUND 30**

**1.1 RESEARCH QUESTIONS30**

**1.2 OBJECTIVES 30**

**2.0 METHODS 30**

**3.0 RESULTS 31**

**3.1 DESCRIPTIVE ANALYSIS 31**

**3.2 EXPLORATORY ANALYSIS 33**

**3.3 ANALYSIS OF VARIANCE 40**

**4.0 DISCUSSION43**

**5.0 CONCLUSION 44**

**DATA SET 4(HD\_GENES)45**

**ABSTRACT45**

**1.0 INTRODUCTION AND BACKGROUND 45**

**1.1 RESEARCH QUESTIONS45**

**1.2 OBJECTIVES 45**

**2.0 METHODS 45**

**3.0 RESULTS 46**

**3.1 DESCRIPTIVE ANALYSIS 46**

**3.2 EXPLORATORY ANALYSIS 58**

**3.1 INDEPENDENT T-TEST 98**

**4.0 DISCUSSION101**

**5.0 CONCLUSION 103**

**ABSTRACT**

The aim of this research was to examine possible variations in length, weight, and head size among male and female full-term infants. Although descriptive statistics and graphical displays (histograms, boxplots) indicated some potential differences, especially in head size, inferential statistics showed that only head size differed significantly between male and female infants. There were no significant differences in weight and length, despite some inconsistencies between descriptive and inferential statistics. This implies that while gender may influence these measurements slightly, they are not statistically significant.

**1.0 INTRODUCTION AND BACKGROUND**

The study utilizes data from a population study, ensuring a potentially representative sample of 256 (119 boys and 137 girls) term-born babies. By analyzing and comparing birth size measurements between genders, the study seeks to contribute to our understanding of potential sex-based variations in early development. Specifically, the study aims to answer the following questions:

* Do boys have longer legs than girls?
* Are boys heavier than girls?
* Are the heads of boys bigger than those of girls?

H0: The lengths of the male and female newborns do not significantly differ from one another (P>=0.05).

HA: There is a statistically significant difference in the length of the male and female newborns P<0.05.

H0: The weight of the male and female newborns does not significantly differ from one another.

HA: There is a statistically significant difference in the weight of male and female newborns.

H0: The head circumference of male and female newborns does not significantly differ from one another.

HA: There is a statistically significant difference in the head circumference of male and female newborns.

* 1. **OBJECTIVES**

This study aims to investigate potential sex differences in birth size measurements (length, weight, and head circumference) at 1 month of age in a group of 256 term-born babies.

1. **METHODS**

**Sample:** 119 males and 137 females A total of 256 babies.

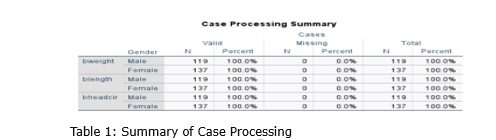
**Data Collection:** Data was imported from an Excel file called "babies.xlsx” into SPSS.

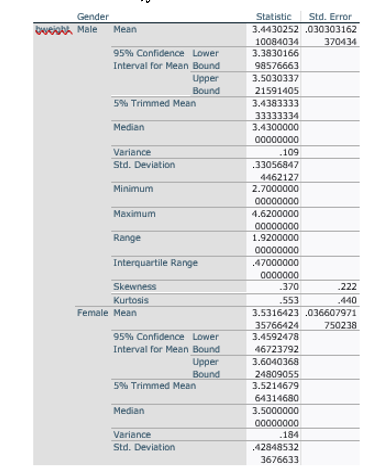
**Independent T-test** will be used because all babies appear only once in their group and the effect size after calculations will be compared to the effect size table.

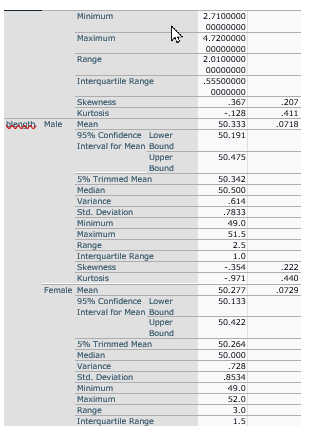
|  |  |
| --- | --- |
|  |  |

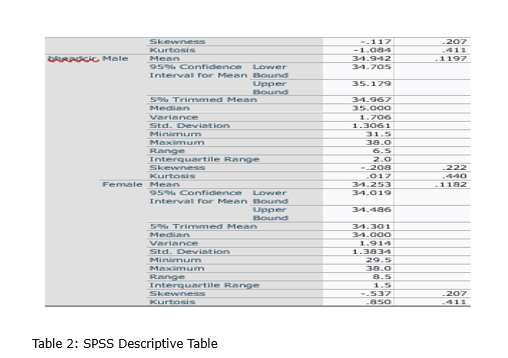
1. **RESULTS**

**3.1 DESCRIPTIVE ANALYSIS (TABLES FROM SPSS)**

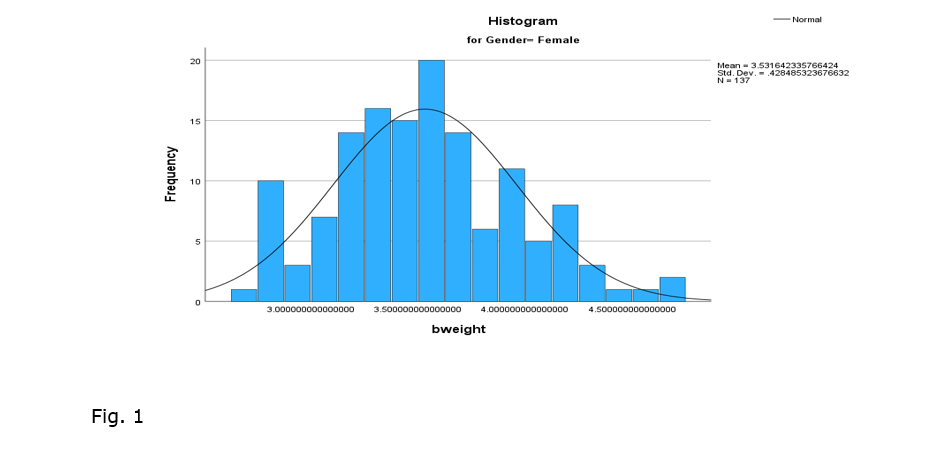
****

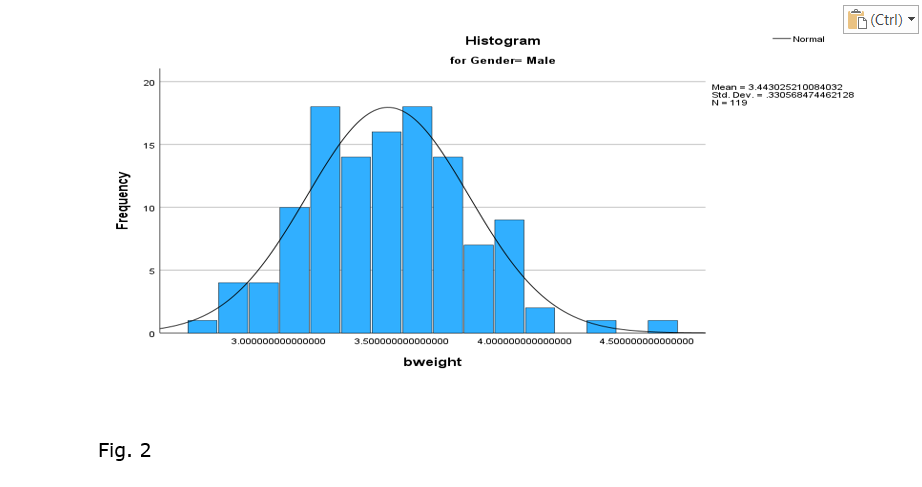


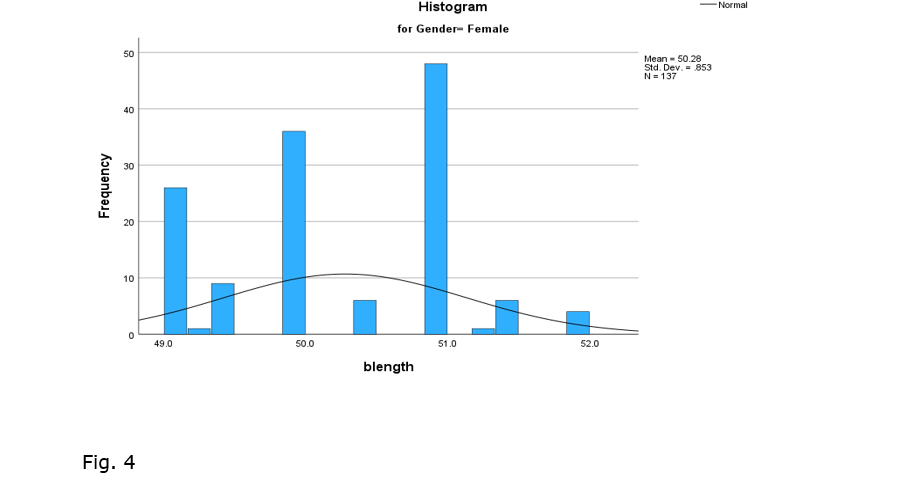
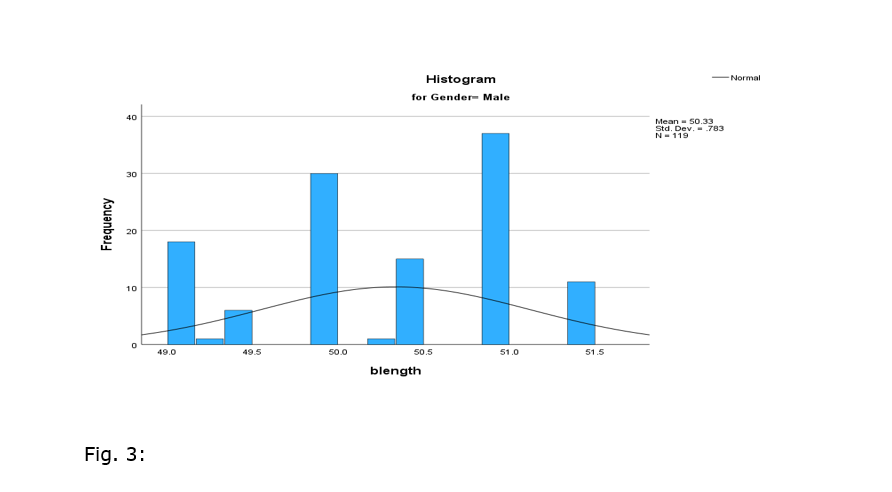


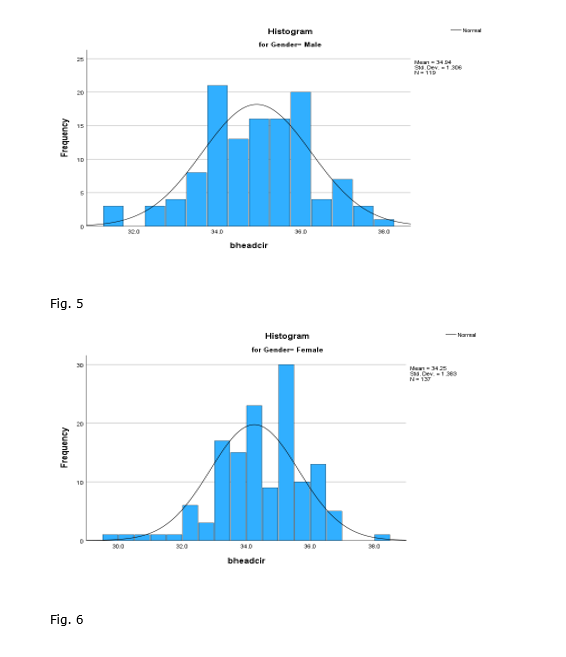
****

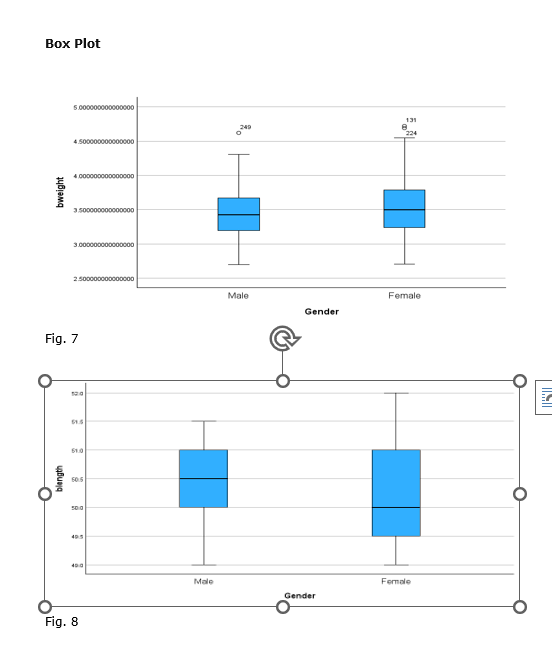
**3.2 Exploratory Analysis (Graphs)**

****

****

****

****

****

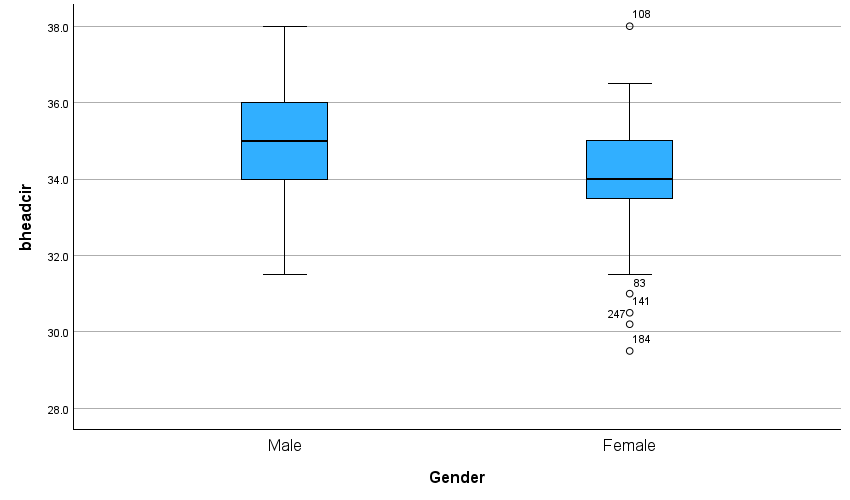
****

Fig. 9

**3.3 Independent T-Test**

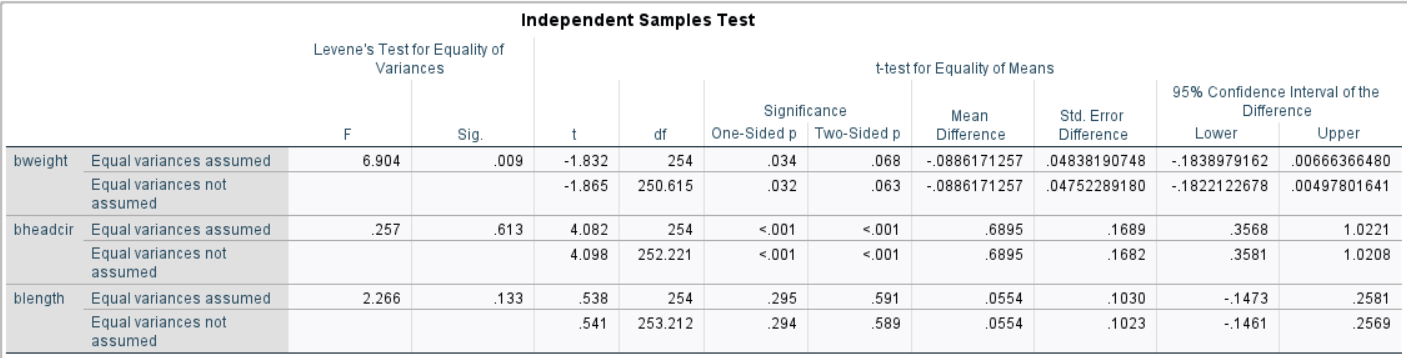
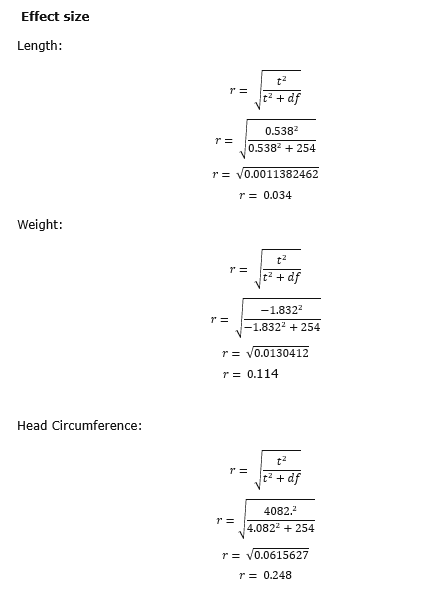
****

Table 4

****

1. **DISCUSSION**

This study compared birth size measurements (length, weight, head circumference) between male and female newborns. Mean values for length and weight appear normally distributed based on skewness and kurtosis values. (Table 2). However, visual exploration through histograms (Fig.1-6) and boxplots (Fig.7-9) suggests some deviations from normality. Table 2 demonstrates that the data is normally distributed. Male babies had length (=50.33, SE = 0.72, median = 50.50, ẟ=0.78), compared to the female babies (= 50.27, SE = 0.73, median = 50.00, ẟ=0.85) r = 0.034. Female babies had weight (=3.53, SE = 0.37, median = 3.50, ẟ=0.43), compared to the male babies (= 3.44, SE = 0.30, median = 3.44, ẟ=0.33, r = 0.114, whereas the male newborns' heads had larger circumferences, (=34.94, SE = 0.12, median = 34.00, ẟ=1.38), compared to the male babies (= 34.25, SE = 0.12, median = 35.00, ẟ=1.31), r = 0.248. Male newborns had slightly higher average head circumference, while females had slightly lower averages for weight and length. These differences, however, were all considered small effect sizes with values. A few outlier values were identified in both groups, particularly for length and head circumference in females. Birth weight distributions for both genders passed normality tests, while length and head circumference for both genders significantly deviated from normality.  Independent T-tests with (the difference, 0.056, BCa 95% CI [-0.14,0.26], (254.212) = 0.54, p = 0.588) and weight (difference, -0.09, BCa 95% CI [-0.19, -0.00], (250.615) = -1.865) p = 0.068; revealed significant differences only in head circumference between genders. No significant differences were found for length or weight. Therefore, while head circumference showed a statistically significant difference between male and female newborns, the magnitude of the difference was small and the distributions for other measurements (length and weight) were not statistically different despite some deviations from normality.

1. **CONCLUSION**

The results indicate that there is no significant difference in leg length or weight between newborn boys and girls, but that boys have larger head circumferences than girls which corresponds with *Allen, P., Bennett, K. et. Al., 2018* and *Agyemang, C. O et. al.,2019*. The study also shows that the birth measurements of term infants are influenced by a combination of trends and variability, which require further exploration with bigger samples and adjustment for possible confounding factors.

**ABSTRACT**

This study investigated the length of stay in hospital for newborn babies with and without infections. Analyzing descriptive statistics and visual explorations revealed deviations from normality in both groups: **longer mean and variance for infected babies, positive skewness and kurtosis, and non-linearity in Q-Q plots.** Statistical tests confirmed these observations, with Kolmogorov-Smirnov and Shapiro-Wilk tests rejecting normality for both groups (p < 0.05). Furthermore, a Mann-Whitney U-test showed a statistically significant difference in length of stay between infected and uninfected babies (p < 0.05). However, the effect size was small (0.25), suggesting a **modest difference in length of stay between the two groups despite the statistical significance.**

**1.0 INTRODUCTION AND BACKGROUND**

The aim of this study was to examine the relationship between infection and length of hospital stay in infants who had surgery. The sample consisted of 132 infants who had surgery at a specific hospital, of whom 52 had infections, 80 did not have infections, and 9 had missing data. The research question was, does infection increase the duration of hospital stay in infants who had surgery?

**H0**: The duration of stay for the two infant groups does not differ much.

**HA:** The duration of stay for the two groups of infants differs significantly.

The number of days an infant spends in the hospital is measured by a continuous variable called the length of stay. An infant’s infection status is determined by a categorical variable called "infection".

**1.2 Objective:**

This study aims to explore the possible link between hospital-acquired infections and length of stay in infants who had surgery. The study will examine how having an infection during their stay affects how long they stay in the hospital compared to those without infections.

**2.0 Methods**

**Sample:** A total of 141 babies, 52 babies with infection and 80 babies with no infection with 9 missing values.

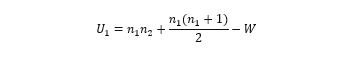
**Data Collection:** The data was imported from a spreadsheet “Surgery.xlsx”. with 141 babies who underwent surgery at a hospital. The file contained details on whether the babies had infection (52) or not (80).

An independent non-parametric sample Mann-Whitney U test will be employed to conduct the test because the groups are independent and the data distribution is non-normal.

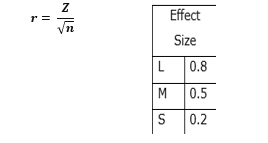
Effect size will also be calculated to measure the magnitude of the difference between the groups.

**U-static:**

Use the sum of ranks from Table 3 to get the value of R1 and R2. Use the sums of ranks to calculate the U-static using the formula for each group.



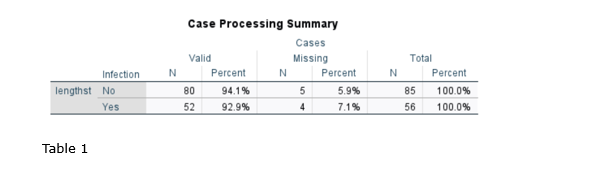
**Effect Size:**

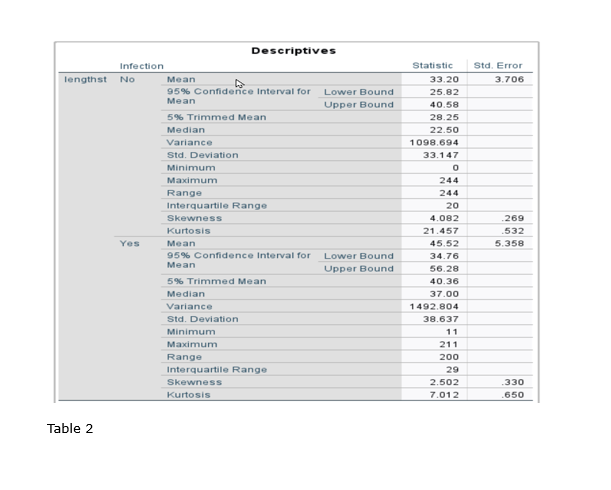
****

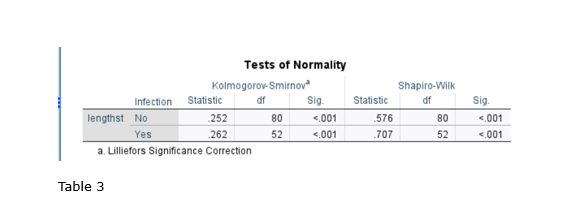
1. **Result:**

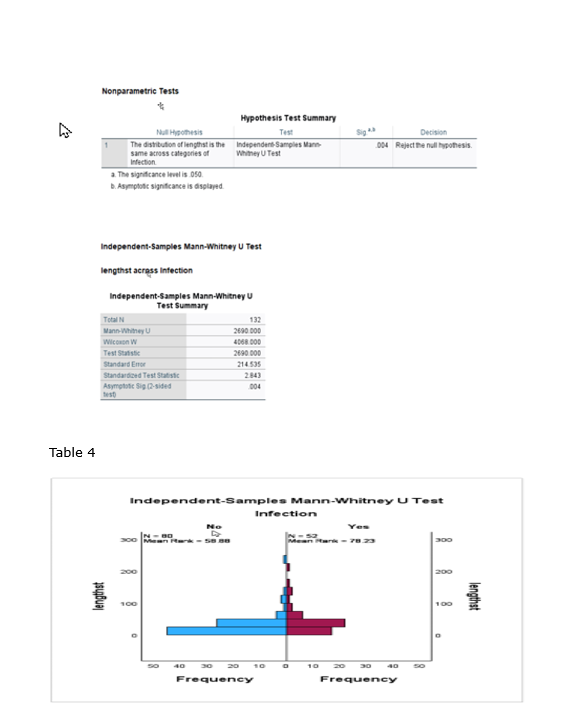
Table 1 is a table about case summary that shows the valid number of samples for the infected and non-infected babies and also number of missing values with their percentages.

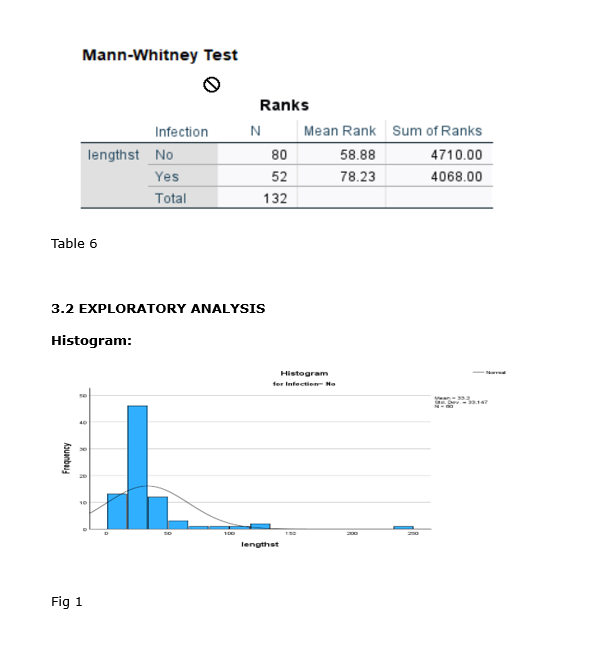
**3.1 DESCRIPTIVE ANALYSIS**

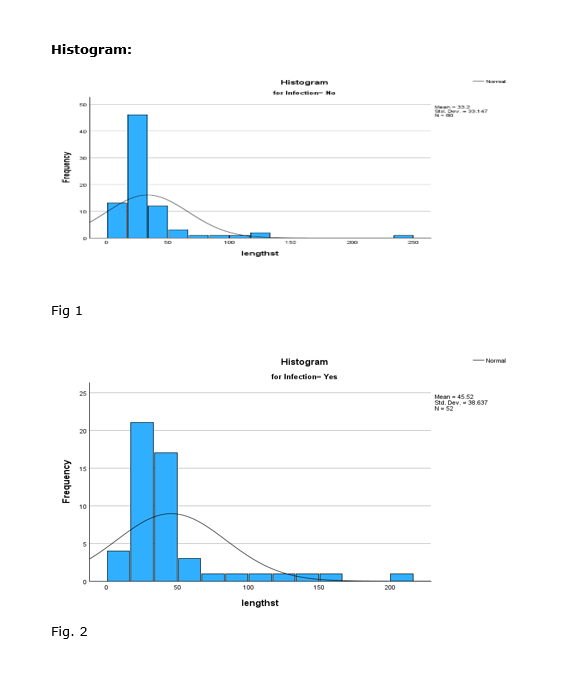


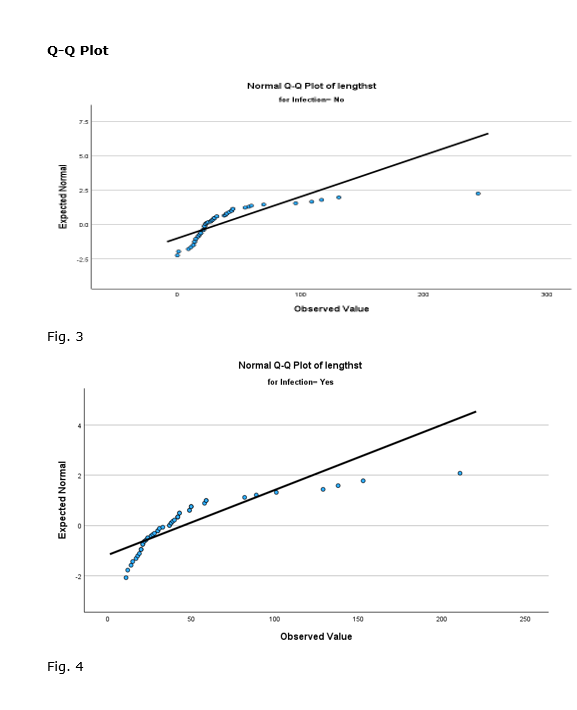


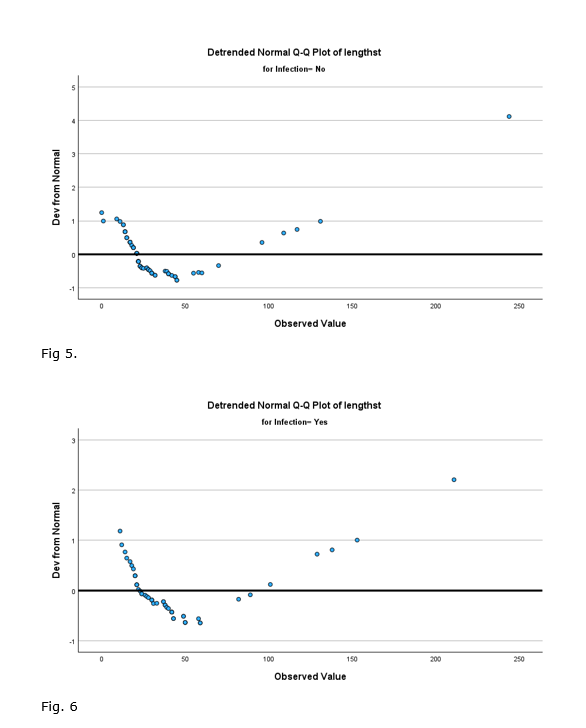


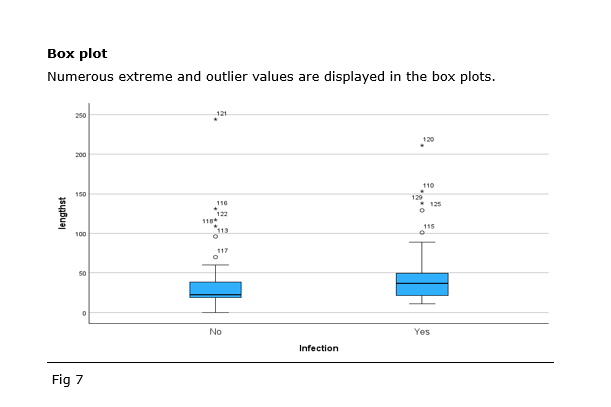




****

****



****

**3.3 NON -PARAMETRIC TEST (Mann-Whitney-U-Test)**

Mann-Whitney U test:

**Mann-Whitney-U-Test Effect Size**

Using this data, the Effect Size can then be computed using the following equation.

**4.0 DISCUSSION**

As shown in Table 2, infected babies have a longer average stay than uninfected babies. The statistics for infants without infection are (x ̅ =33.20, median =22.50 and Variance= 1098.694) and for infants with infection are (x ̅= 45.52, median=37.00 and Variance =1492.804), indicating that the variances are not equal. The data are not normally distributed, as evidenced by the high skewness and kurtosis values, which are all above 2. Table 3 shows that both Kolmogorov–Smirnov and Shapiro–Wilk tests of normality have P values less than 0.05, rejecting the null hypothesis of normality for both groups.

The result in table 4, with a p-value of 0.004, clearly shows that the distributions of the two groups are significantly different, thus rejecting the null hypothesis of no difference in the length of stay between infected and uninfected babies. The exploratory data analysis also confirmed the non-normal distribution and positive skewness of both groups, as seen in the histograms in figure 1 and 2. The box plot showed the variation in the central tendency of the two groups. The Q-Q plots for each group in Fig. 3-6 were noticeably curved and deviated from the normality line. The length of stay was significantly different between infected and uninfected babies (p < 0.05), as supported by the descriptive and exploratory analysis.

The Effect Size in the Mann-Whitney-U-Test was calculated using the Standardized test statistic z and the number of pairs n. The Effect Size is 0.25, which is a small effect. Therefore, there is a small difference between the group of infected babies and the group of uninfected babies based on their length of stay.

**5.0 CONCLUSION:**

There's a significant difference in the duration of stay between the two infant groups. These findings indicate that while babies with infections tend to stay in hospital longer on average, the magnitude of this difference is relatively small which is similar to *Field, A., 2013*. Future research may explore specific factors contributing to this difference and potential interventions to reduce length of stay for both infected and uninfected newborns.

**ABSTRACT**

The research aims to examine the normality and homogeneity of variance of birth weights for babies with different numbers of siblings (none, one, two, or three and more). The data summary and normality tests indicate that babies with one sibling show some deviation from normality, while the other groups are normal. This is also supported by the data visualization, which shows that babies with one sibling have non-normal distribution in histograms and QQ plots. The ANOVA test shows significant differences in mean weights among the groups, but the post-hoc tests indicate that only the babies with no siblings are significantly different from the rest. The number of siblings accounts for only 13% of the variation in weight, indicating a small effect size. The research suggests that the number of siblings has a small impact on birth weight, and babies with one sibling have slightly different weight patterns than the other groups.

**1.0 INTRODUCTION AND BACKGROUND**

Parity, has long been suspected to influence various aspects of children's development, including weight. This study delves into this potential relationship by analyzing the birth weight of 550 term babies at 1 month of age, categorized by their birth order (parity). The research question is: Do the weights of these babies demonstrate a statistically significant relationship with their birth order?

H0: There is no significant difference in mean weight between groups defined by parity.

HA: There is a significant difference in mean weight between groups defined by parity.

Weight is continuous variable while parity is the factor.

**1.2 Objective:**

To analyze if the weight of babies after 1 month of birth is related to their parity using descriptive statistics, Exploratory analysis and one-way ANOVA.

**2.0 METHODS**

**Sample:** A total of 550 term babies, parity 180 single ton, 192 with 1 sibling,116 with 2 siblings and 62 with 3 or more siblings.

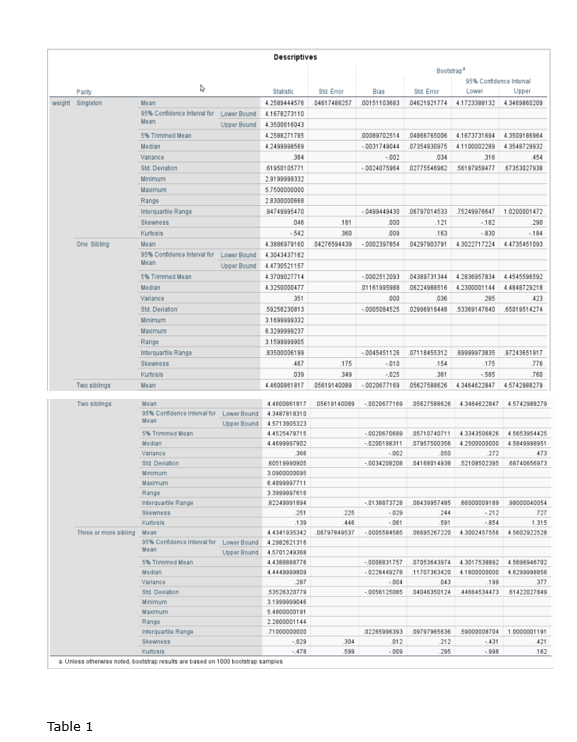
**Data Collection:** Data from 550 babies that were imported from an excel file “weight.xlsx”.

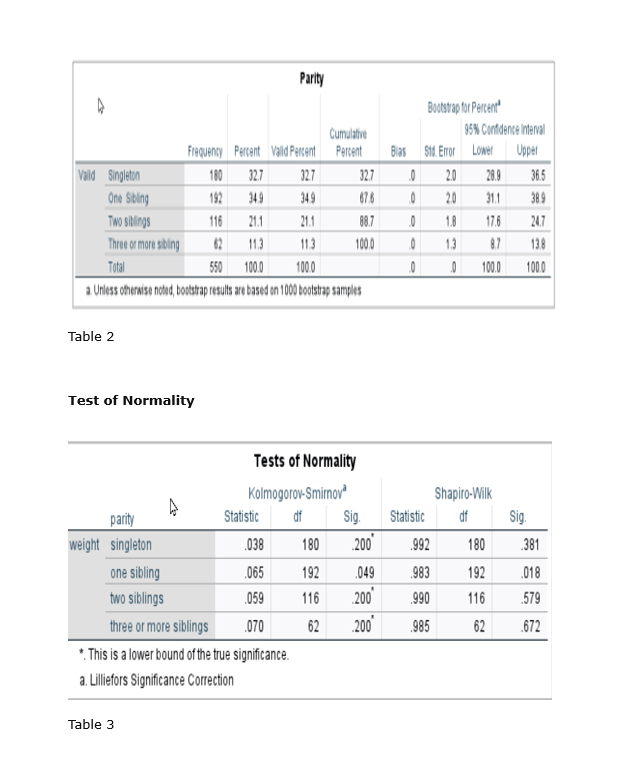
A one-way ANOVA test will be used to test the differences in the mean weight of babies across four groups based on the number of siblings they because it is more convenient and reliable than comparing each group pair by pair, which could lead to problems with multiple testing.

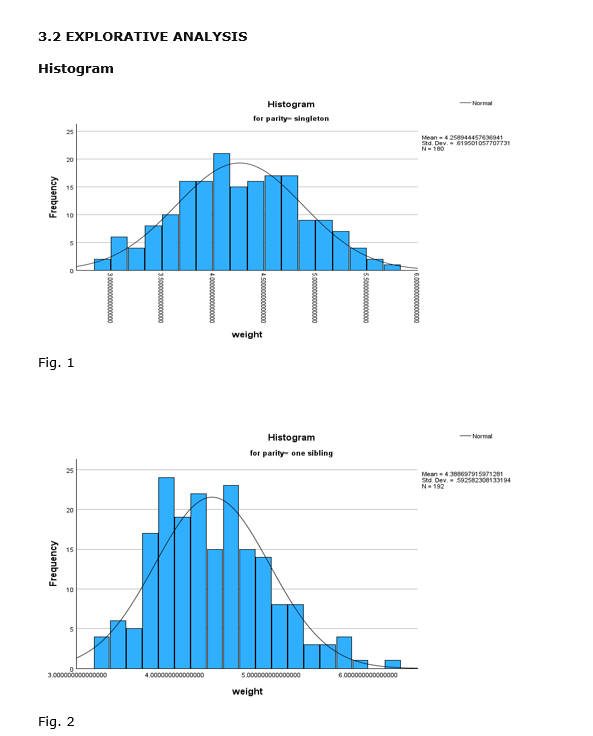
Effect size will also be calculated.

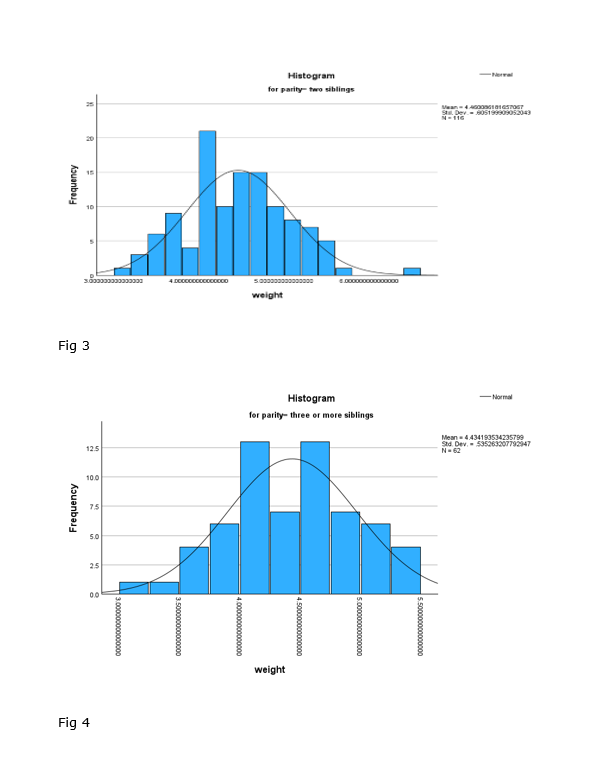
**3.0 RESULTS**

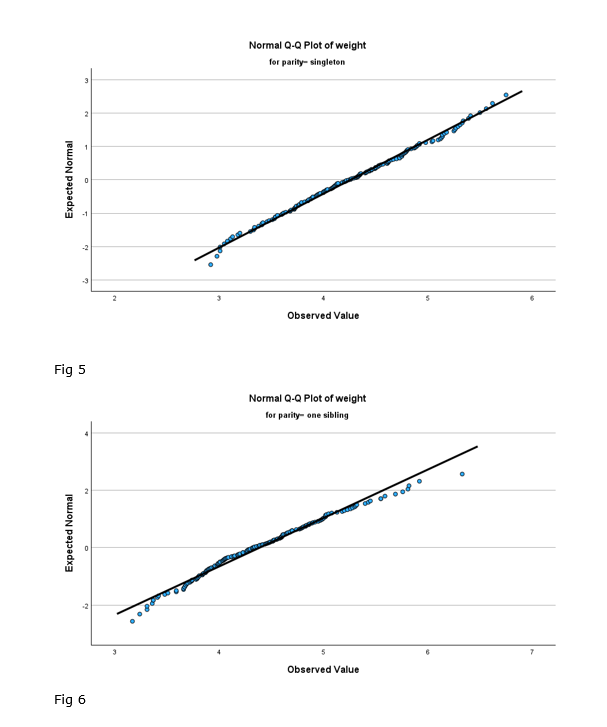
**3.1 DESCRIPTIVE ANALYSIS**





****





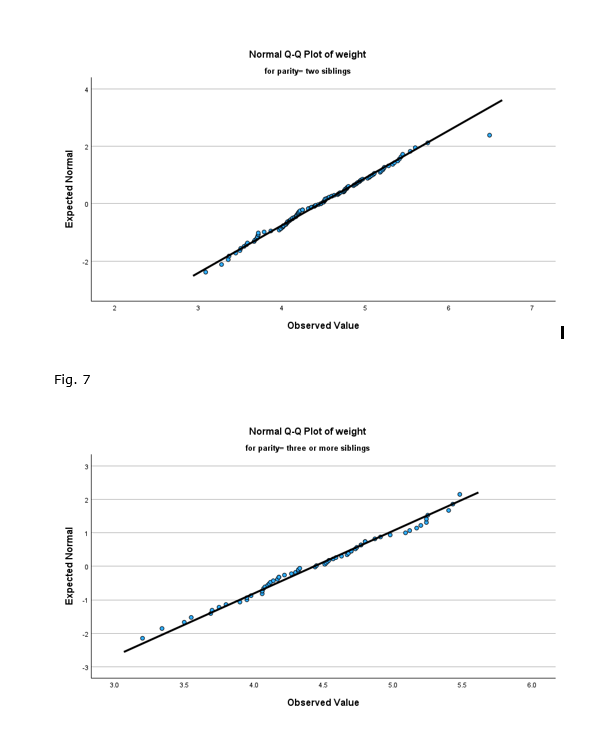
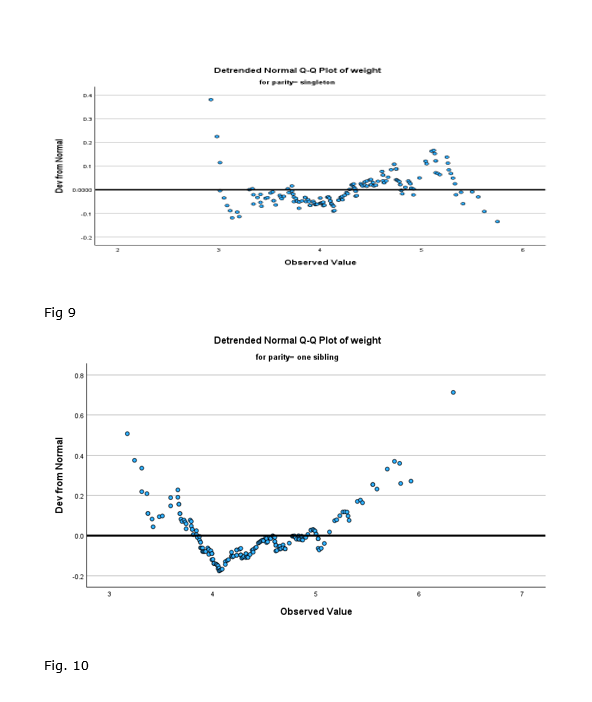
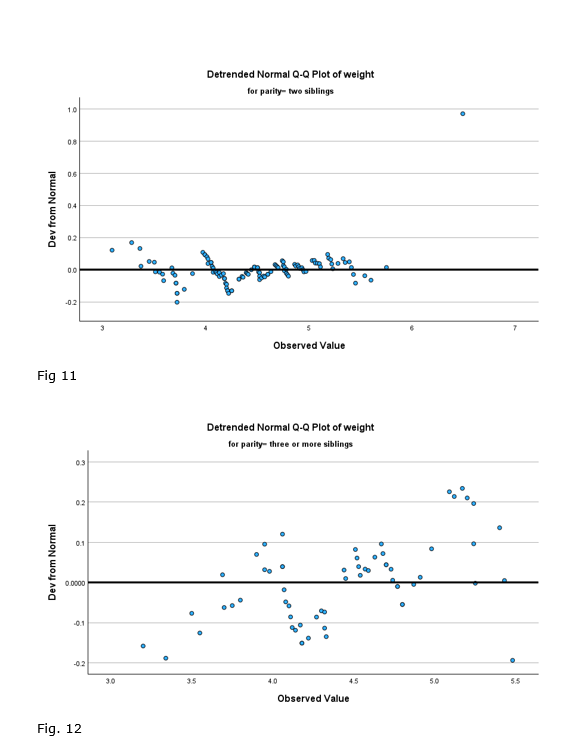
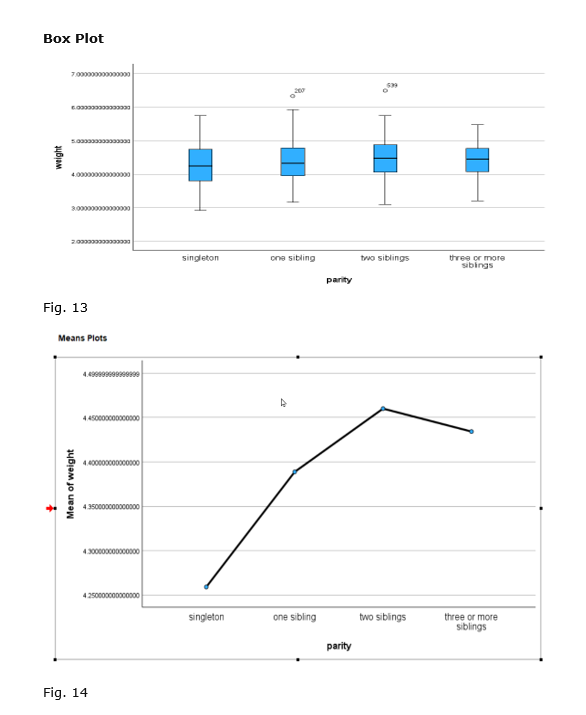
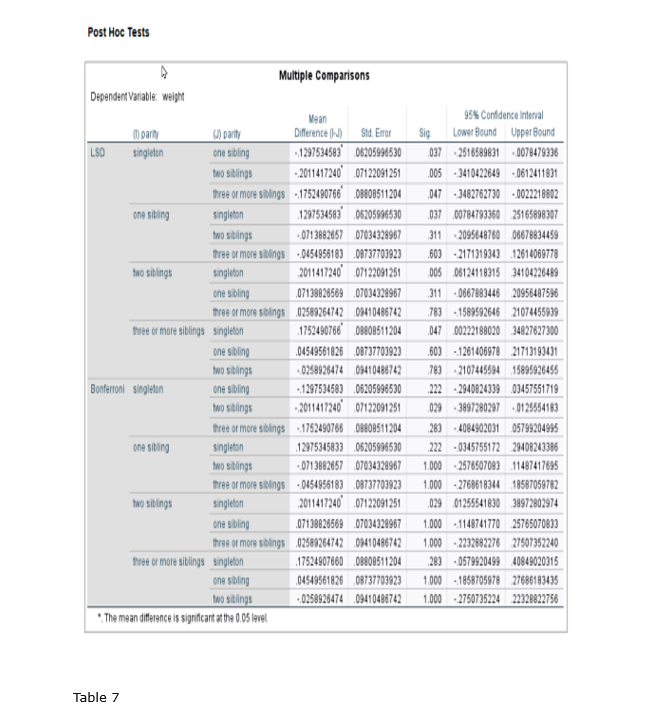


Fig. 8

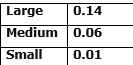






****

**EFFECT SIZE**

****

**4.0 DISCUSSION**

The descriptive table 1 indicates that the weight, means, and medians within each group are roughly equal, and the skewness and kurtosis values fall within the range of -1 to +1, indicating that the data are almost normally distributed. The Variance in the groups were 0.384, 0.351, 0.366, and 0.287 respectively.

P>0.05 indicate that the data for singletons, babies with two siblings, and newborns with three or more siblings adhere to normality, according to the Kolmogorov-Smirnov and Shapiro-Wilk statistics in the Tests of Normality table 3. Based on these tests, babies with one sibling do not seem to fit into a normal distribution because the P values of 0.049 and 0.018 <0.05.

Explorative Analysis fig. 2 shows that the distribution of babies with one sibling exhibits slightly scattered tails, deviating from a perfect bell-shaped curve. The normal Q–Q plots shown in Fig. 5-8 displays normal Q-Q graphs with little variations at the extremities but Fig 6 deviates slightly from normality at both extremes. Despite the fact that in histogram Fig. 4 is not typically bell shaped, the normal Q–Q plot suggests that this distribution roughly follows the normal bell curve.

Fig. 13 shows that there are two outlier values, one in the group of babies with one sibling and one in the group of babies with two siblings.

In table 4, the mean P value is 0.590 hence p> 0.05, indicates that the variances of the groups are not significantly different from one another.

Table 5 shows that the F value greater (3.239) and is significant at P=0.022. Therefore, H0 is rejected and this means that the mean population values of the four parity groups differ significantly from one another.

According to the calculated effect size and in table 6 parity only accounts for 13% of the variation in weight. This represents a small effect size.

According to the Post Hoc test, LSD showed that Singletons had significantly different mean weights compared to all other parity groups and finally, when two siblings was compared with three or more siblings, it also showed that they were not significantly different.

**5.0 CONCLUSION**

These findings suggest that parity has a statistically significant but modest impact on birth weight, with singletons generally differing significantly compared to babies with siblings. Further research could explore potential mechanisms underlying this association and factors that might influence weight variation within parity groups*.( Pallant, J., 2020.* )

**ABSTRACT**

This study investigated the expression levels of ten genes in patients with Huntington's disease (HDs) compared to healthy controls. Analyzing means, medians, standard deviations, and normality curves revealed that some genes (DDR1, SLC38A2, CCNB1) showed higher average expression in HDs, while others (XG, CRYM, SCIN) had lower averages. Control patients generally displayed greater variability in expression than HDs, though exceptions existed (THBS3, SLC38A2). Independent T-tests confirmed significant differences in expression for DDR1, SLC38A2, CCNB1, SCIN, SMAD7, and APOH. Impact of these differences varied, with DDR1 exhibiting the largest effect (large), followed by moderate effects for THBS3, SLC38A2, CCNB1, and SMAD7.These findings suggest that altered expression of certain genes, particularly DDR1, might be involved in Huntington's disease pathogenesis. Further research is needed to elucidate the functional significance of these changes and their potential as diagnostic or therapeutic targets.

**1.0 INTRODUCTION AND BACKGROUND**

HD causes progressive brain damage. The genetic cause of HD is clear, but the molecular mechanisms that underlie its effects are not well understood. Changes in gene expression may offer clues for potential therapies. The study will use a dataset of 20 tissue samples from healthy and HD individuals to find genes that have significant expression differences in HD, and to investigate how these genes may contribute to disease progression.

**1.1 RESEARCH AND QUESTION**

Are the genes significantly dysregulated in HD?

H0: There is no significant difference in the mean expression levels of these genes across the two groups.

HA: There is a significant difference in the mean expression levels of these genes across the two groups.

**1.2 Objective:**

To analyze the gene expression of 10 healthy control and 10 patients with HD disease using descriptive statistics, Exploratory analysis and independent t-test.

**2.0 METHODS**

**Sample:** A total of 20 tissue sample population size 10 control samples and 10 HD samples.

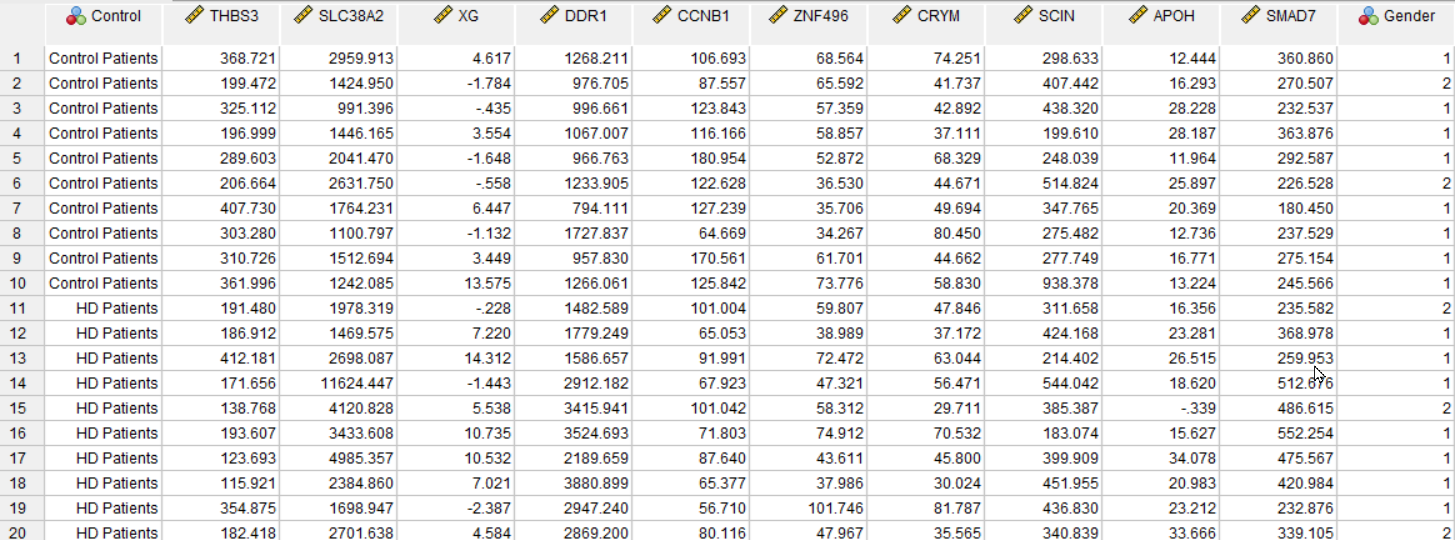
**Data Collection:** Data from a population sample of 20 tissue samples, which are stored in the spreadsheet HD\_Genes.xlsx will be used.

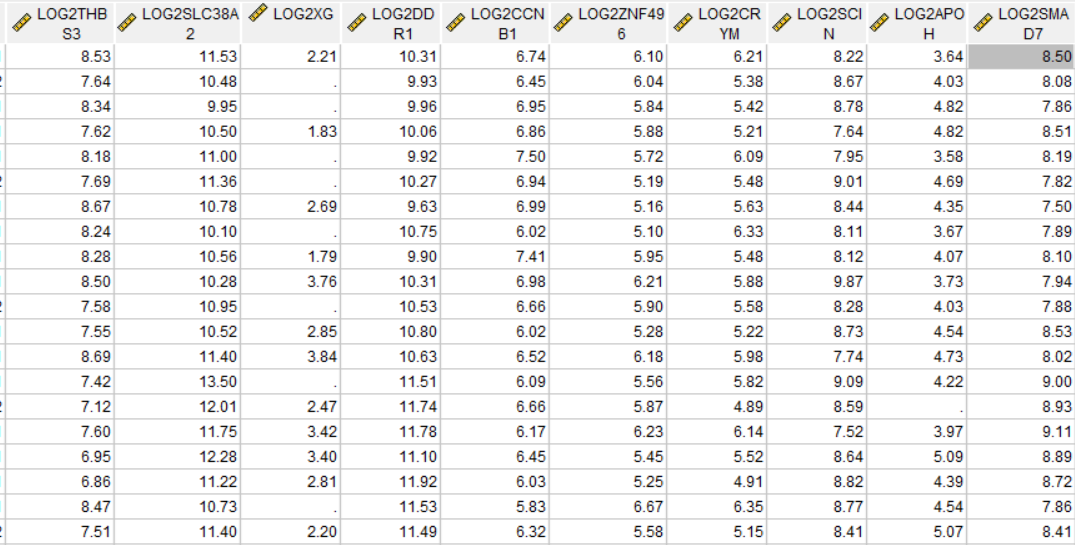
**An independent t-test** was used in the study to compare the mean expression level of each gene between the two groups: Controls and HD and effect size will also be calculated.

To normalize the expression values and enhance the analysis, first log transformation will be applied, a common technique in gene expression studies. Subsequently, SPSS statistical methods to compare expression levels between the control and HD groups, identifying genes with significant dysregulation.

**3.0 RESULTS**

**3.1 DESCRIPTIVE ANALYSIS**



****

**Table 1a**

**Explore**

|  |  |  |
| --- | --- | --- |
| **Notes** | | |
| Output Created | | 13-JAN-2024 14:00:55 |
| Comments | |  |
| Input | Data | C:\Users\C2563749\OneDrive - Teesside University\Data analytics\ICA HD genes.sav |
| Active Dataset | DataSet1 |
| Filter | <none> |
| Weight | <none> |
| Split File | <none> |
| N of Rows in Working Data File | 20 |
| Missing Value Handling | Definition of Missing | User-defined missing values for dependent variables are treated as missing. |
| Cases Used | Statistics are based on cases with no missing values for any dependent variable or factor used. |
| Syntax | | EXAMINE VARIABLES=logTHBS3 logSLC38A2 logXG logDDR1 logCCNB1 logZNF496 logCRYM logSCIN logAPOH logSMAD7 BY Control /PLOT BOXPLOT HISTOGRAM NPPLOT /COMPARE GROUPS /STATISTICS DESCRIPTIVES /CINTERVAL 95 /MISSING LISTWISE /NOTOTAL. |
| Resources | Processor Time | 00:00:07.78 |
| Elapsed Time | 00:00:07.43 |

**Control**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Case Processing Summary** | | | | | | | |
|  | Control | Cases | | | | | |
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| logTHBS3 | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logSLC38A2 | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logXG | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logDDR1 | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logCCNB1 | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logZNF496 | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logCRYM | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logSCIN | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logAPOH | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| logSMAD7 | Control | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |
| Hd disease | 10 | 100.0% | 0 | 0.0% | 10 | 100.0% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Descriptives** | | | | | |
|  | Control | | | Statistic | Std. Error |
| logTHBS3 | Control | Mean | | 8.1697648 | .12228570 |
| 95% Confidence Interval for Mean | Lower Bound | 7.8931353 |  |
| Upper Bound | 8.4463942 |  |
| 5% Trimmed Mean | | 8.1723211 |  |
| Median | | 8.2620029 |  |
| Variance | | .150 |  |
| Std. Deviation | | .38670134 |  |
| Minimum | | 7.62204 |  |
| Maximum | | 8.67147 |  |
| Range | | 1.04943 |  |
| Interquartile Range | | .82810 |  |
| Skewness | | -.499 | .687 |
| Kurtosis | | -1.277 | 1.334 |
| Hd disease | Mean | | 7.5741184 | .18808810 |
| 95% Confidence Interval for Mean | Lower Bound | 7.1486336 |  |
| Upper Bound | 7.9996032 |  |
| 5% Trimmed Mean | | 7.5521242 |  |
| Median | | 7.5286598 |  |
| Variance | | .354 |  |
| Std. Deviation | | .59478679 |  |
| Minimum | | 6.85700 |  |
| Maximum | | 8.68713 |  |
| Range | | 1.83014 |  |
| Interquartile Range | | .74048 |  |
| Skewness | | .917 | .687 |
| Kurtosis | | .268 | 1.334 |
| logSLC38A2 | Control | Mean | | 10.6547177 | .16338410 |
| 95% Confidence Interval for Mean | Lower Bound | 10.2851172 |  |
| Upper Bound | 11.0243182 |  |
| 5% Trimmed Mean | | 10.6449832 |  |
| Median | | 10.5304605 |  |
| Variance | | .267 |  |
| Std. Deviation | | .51666591 |  |
| Minimum | | 9.95332 |  |
| Maximum | | 11.53134 |  |
| Range | | 1.57802 |  |
| Interquartile Range | | .85200 |  |
| Skewness | | .497 | .687 |
| Kurtosis | | -.634 | 1.334 |
| Hd disease | Mean | | 11.5761280 | .27580897 |
| 95% Confidence Interval for Mean | Lower Bound | 10.9522048 |  |
| Upper Bound | 12.2000513 |  |
| 5% Trimmed Mean | | 11.5275835 |  |
| Median | | 11.3986699 |  |
| Variance | | .761 |  |
| Std. Deviation | | .87218456 |  |
| Minimum | | 10.52118 |  |
| Maximum | | 13.50487 |  |
| Range | | 2.98369 |  |
| Interquartile Range | | 1.18226 |  |
| Skewness | | 1.158 | .687 |
| Kurtosis | | 1.713 | 1.334 |
| logXG | Control | Mean | | 1.7281286 | .60073877 |
| 95% Confidence Interval for Mean | Lower Bound | .3691631 |  |
| Upper Bound | 3.0870941 |  |
| 5% Trimmed Mean | | 1.6554584 |  |
| Median | | 1.3938091 |  |
| Variance | | 3.609 |  |
| Std. Deviation | | 1.89970280 |  |
| Minimum | | .00144 |  |
| Maximum | | 4.76288 |  |
| Range | | 4.76144 |  |
| Interquartile Range | | 3.32593 |  |
| Skewness | | .314 | .687 |
| Kurtosis | | -1.800 | 1.334 |
| Hd disease | Mean | | 2.7994079 | .62873028 |
| 95% Confidence Interval for Mean | Lower Bound | 1.3771212 |  |
| Upper Bound | 4.2216946 |  |
| 5% Trimmed Mean | | 2.8415313 |  |
| Median | | 3.6405208 |  |
| Variance | | 3.953 |  |
| Std. Deviation | | 1.98821972 |  |
| Minimum | | .00144 |  |
| Maximum | | 4.83915 |  |
| Range | | 4.83771 |  |
| Interquartile Range | | 4.40215 |  |
| Skewness | | -.818 | .687 |
| Kurtosis | | -1.306 | 1.334 |
| logDDR1 | Control | Mean | | 11.1044410 | .09901160 |
| 95% Confidence Interval for Mean | Lower Bound | 10.8804612 |  |
| Upper Bound | 11.3284209 |  |
| 5% Trimmed Mean | | 11.0944929 |  |
| Median | | 11.0101565 |  |
| Variance | | .098 |  |
| Std. Deviation | | .31310218 |  |
| Minimum | | 10.63320 |  |
| Maximum | | 11.75475 |  |
| Range | | 1.12155 |  |
| Interquartile Range | | .39307 |  |
| Skewness | | .763 | .687 |
| Kurtosis | | 1.031 | 1.334 |
| Hd disease | Mean | | 12.3022212 | .15877952 |
| 95% Confidence Interval for Mean | Lower Bound | 11.9430369 |  |
| Upper Bound | 12.6614054 |  |
| 5% Trimmed Mean | | 12.3104636 |  |
| Median | | 12.4971588 |  |
| Variance | | .252 |  |
| Std. Deviation | | .50210493 |  |
| Minimum | | 11.53390 |  |
| Maximum | | 12.92218 |  |
| Range | | 1.38827 |  |
| Interquartile Range | | .99364 |  |
| Skewness | | -.476 | .687 |
| Kurtosis | | -1.385 | 1.334 |
| logCCNB1 | Control | Mean | | 7.8835522 | .13544622 |
| 95% Confidence Interval for Mean | Lower Bound | 7.5771515 |  |
| Upper Bound | 8.1899528 |  |
| 5% Trimmed Mean | | 7.8975868 |  |
| Median | | 7.9452566 |  |
| Variance | | .183 |  |
| Std. Deviation | | .42831857 |  |
| Minimum | | 7.01500 |  |
| Maximum | | 8.49948 |  |
| Range | | 1.48448 |  |
| Interquartile Range | | .43105 |  |
| Skewness | | -.624 | .687 |
| Kurtosis | | 1.027 | 1.334 |
| Hd disease | Mean | | 7.2749623 | .09206233 |
| 95% Confidence Interval for Mean | Lower Bound | 7.0667029 |  |
| Upper Bound | 7.4832218 |  |
| 5% Trimmed Mean | | 7.2786058 |  |
| Median | | 7.2449954 |  |
| Variance | | .085 |  |
| Std. Deviation | | .29112666 |  |
| Minimum | | 6.82553 |  |
| Maximum | | 7.65881 |  |
| Range | | .83328 |  |
| Interquartile Range | | .52821 |  |
| Skewness | | .027 | .687 |
| Kurtosis | | -1.365 | 1.334 |
| logZNF496 | Control | Mean | | 6.7180503 | .13132105 |
| 95% Confidence Interval for Mean | Lower Bound | 6.4209815 |  |
| Upper Bound | 7.0151192 |  |
| 5% Trimmed Mean | | 6.7253989 |  |
| Median | | 6.8605450 |  |
| Variance | | .172 |  |
| Std. Deviation | | .41527363 |  |
| Minimum | | 6.09875 |  |
| Maximum | | 7.20508 |  |
| Range | | 1.10633 |  |
| Interquartile Range | | .86865 |  |
| Skewness | | -.656 | .687 |
| Kurtosis | | -1.303 | 1.334 |
| Hd disease | Mean | | 6.7970659 | .14504954 |
| 95% Confidence Interval for Mean | Lower Bound | 6.4689410 |  |
| Upper Bound | 7.1251907 |  |
| 5% Trimmed Mean | | 6.7791718 |  |
| Median | | 6.7248456 |  |
| Variance | | .210 |  |
| Std. Deviation | | .45868693 |  |
| Minimum | | 6.24740 |  |
| Maximum | | 7.66883 |  |
| Range | | 1.42143 |  |
| Interquartile Range | | .78508 |  |
| Skewness | | .624 | .687 |
| Kurtosis | | -.329 | 1.334 |
| logCRYM | Control | Mean | | 6.7134174 | .12296387 |
| 95% Confidence Interval for Mean | Lower Bound | 6.4352538 |  |
| Upper Bound | 6.9915810 |  |
| 5% Trimmed Mean | | 6.7069196 |  |
| Median | | 6.5581332 |  |
| Variance | | .151 |  |
| Std. Deviation | | .38884591 |  |
| Minimum | | 6.21377 |  |
| Maximum | | 7.33002 |  |
| Range | | 1.11625 |  |
| Interquartile Range | | .71161 |  |
| Skewness | | .494 | .687 |
| Kurtosis | | -1.317 | 1.334 |
| Hd disease | Mean | | 6.5558833 | .16212200 |
| 95% Confidence Interval for Mean | Lower Bound | 6.1891379 |  |
| Upper Bound | 6.9226288 |  |
| 5% Trimmed Mean | | 6.5483857 |  |
| Median | | 6.5488011 |  |
| Variance | | .263 |  |
| Std. Deviation | | .51267479 |  |
| Minimum | | 5.89293 |  |
| Maximum | | 7.35380 |  |
| Range | | 1.46087 |  |
| Interquartile Range | | .92747 |  |
| Skewness | | .137 | .687 |
| Kurtosis | | -1.282 | 1.334 |
| logSCIN | Control | Mean | | 9.4811367 | .20180025 |
| 95% Confidence Interval for Mean | Lower Bound | 9.0246328 |  |
| Upper Bound | 9.9376406 |  |
| 5% Trimmed Mean | | 9.4504260 |  |
| Median | | 9.3320994 |  |
| Variance | | .407 |  |
| Std. Deviation | | .63814842 |  |
| Minimum | | 8.64104 |  |
| Maximum | | 10.87403 |  |
| Range | | 2.23299 |  |
| Interquartile Range | | .76590 |  |
| Skewness | | 1.082 | .687 |
| Kurtosis | | 1.539 | 1.334 |
| Hd disease | Mean | | 9.4597946 | .15560366 |
| 95% Confidence Interval for Mean | Lower Bound | 9.1077947 |  |
| Upper Bound | 9.8117945 |  |
| 5% Trimmed Mean | | 9.4773352 |  |
| Median | | 9.6168460 |  |
| Variance | | .242 |  |
| Std. Deviation | | .49206198 |  |
| Minimum | | 8.51628 |  |
| Maximum | | 10.08757 |  |
| Range | | 1.57129 |  |
| Interquartile Range | | .63430 |  |
| Skewness | | -1.002 | .687 |
| Kurtosis | | .292 | 1.334 |
| logAPOH | Control | Mean | | 5.1387040 | .15798773 |
| 95% Confidence Interval for Mean | Lower Bound | 4.7813109 |  |
| Upper Bound | 5.4960971 |  |
| 5% Trimmed Mean | | 5.1319109 |  |
| Median | | 5.0470386 |  |
| Variance | | .250 |  |
| Std. Deviation | | .49960108 |  |
| Minimum | | 4.58063 |  |
| Maximum | | 5.81906 |  |
| Range | | 1.23843 |  |
| Interquartile Range | | 1.06280 |  |
| Skewness | | .371 | .687 |
| Kurtosis | | -1.674 | 1.334 |
| Hd disease | Mean | | 4.9579693 | .56395944 |
| 95% Confidence Interval for Mean | Lower Bound | 3.6822044 |  |
| Upper Bound | 6.2337342 |  |
| 5% Trimmed Mean | | 5.1703988 |  |
| Median | | 5.4639740 |  |
| Variance | | 3.181 |  |
| Std. Deviation | | 1.78339635 |  |
| Minimum | | .00144 |  |
| Maximum | | 6.09077 |  |
| Range | | 6.08933 |  |
| Interquartile Range | | .79955 |  |
| Skewness | | -2.889 | .687 |
| Kurtosis | | 8.765 | 1.334 |
| logSMAD7 | Control | Mean | | 9.0391182 | .09785955 |
| 95% Confidence Interval for Mean | Lower Bound | 8.8177445 |  |
| Upper Bound | 9.2604919 |  |
| 5% Trimmed Mean | | 9.0433114 |  |
| Median | | 9.0097446 |  |
| Variance | | .096 |  |
| Std. Deviation | | .30945907 |  |
| Minimum | | 8.49546 |  |
| Maximum | | 9.50730 |  |
| Range | | 1.01185 |  |
| Interquartile Range | | .41649 |  |
| Skewness | | .099 | .687 |
| Kurtosis | | .016 | 1.334 |
| Hd disease | Mean | | 9.5347445 | .14991649 |
| 95% Confidence Interval for Mean | Lower Bound | 9.1956098 |  |
| Upper Bound | 9.8738792 |  |
| 5% Trimmed Mean | | 9.5401269 |  |
| Median | | 9.6225063 |  |
| Variance | | .225 |  |
| Std. Deviation | | .47407757 |  |
| Minimum | | 8.86342 |  |
| Maximum | | 10.10919 |  |
| Range | | 1.24577 |  |
| Interquartile Range | | .95885 |  |
| Skewness | | -.412 | .687 |
| Kurtosis | | -1.514 | 1.334 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tests of Normality** | | | | | | | |
|  | Control | Kolmogorov-Smirnova | | | Shapiro-Wilk | | |
| Statistic | df | Sig. | Statistic | df | Sig. |
| logTHBS3 | Control | .208 | 10 | .200\* | .881 | 10 | .134 |
| Hd disease | .285 | 10 | .021 | .880 | 10 | .131 |
| logSLC38A2 | Control | .171 | 10 | .200\* | .953 | 10 | .703 |
| Hd disease | .180 | 10 | .200\* | .923 | 10 | .386 |
| logXG | Control | .318 | 10 | .005 | .798 | 10 | .014 |
| Hd disease | .279 | 10 | .026 | .780 | 10 | .008 |
| logDDR1 | Control | .177 | 10 | .200\* | .929 | 10 | .440 |
| Hd disease | .243 | 10 | .096 | .902 | 10 | .230 |
| logCCNB1 | Control | .201 | 10 | .200\* | .927 | 10 | .421 |
| Hd disease | .146 | 10 | .200\* | .934 | 10 | .486 |
| logZNF496 | Control | .217 | 10 | .199 | .858 | 10 | .073 |
| Hd disease | .179 | 10 | .200\* | .942 | 10 | .571 |
| logCRYM | Control | .225 | 10 | .165 | .910 | 10 | .281 |
| Hd disease | .146 | 10 | .200\* | .947 | 10 | .639 |
| logSCIN | Control | .158 | 10 | .200\* | .932 | 10 | .466 |
| Hd disease | .204 | 10 | .200\* | .900 | 10 | .221 |
| logAPOH | Control | .196 | 10 | .200\* | .868 | 10 | .094 |
| Hd disease | .402 | 10 | <.001 | .574 | 10 | <.001 |
| logSMAD7 | Control | .143 | 10 | .200\* | .949 | 10 | .651 |
| Hd disease | .175 | 10 | .200\* | .895 | 10 | .191 |
| \*. This is a lower bound of the true significance. | | | | | | | |
| a. Lilliefors Significance Correction | | | | | | | |

**logTHBS3**

**Histograms**

img.emf

img.emf**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logSLC38A2**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logXG**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logDDR1**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logCCNB1**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logZNF496**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logCRYMHistograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf**logSCIN**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**logAPOH**

**Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf**logSMAD7Histograms**

img.emf

img.emf

**Normal Q-Q Plots**

img.emf

img.emf

**Detrended Normal Q-Q Plots**

img.emf

img.emf

img.emf

**T-Test**

|  |
| --- |
|  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Group Statistics** | | | | | |
|  | Control | N | Mean | Std. Deviation | Std. Error Mean |
| logTHBS3 | Control | 10 | 8.1697648 | .38670134 | .12228570 |
| Hd disease | 10 | 7.5741184 | .59478679 | .18808810 |
| logSLC38A2 | Control | 10 | 10.6547177 | .51666591 | .16338410 |
| Hd disease | 10 | 11.5761280 | .87218456 | .27580897 |
| logXG | Control | 10 | 1.7281286 | 1.89970280 | .60073877 |
| Hd disease | 10 | 2.7994079 | 1.98821972 | .62873028 |
| logDDR1 | Control | 10 | 11.1044410 | .31310218 | .09901160 |
| Hd disease | 10 | 12.3022212 | .50210493 | .15877952 |
| logCCNB1 | Control | 10 | 7.8835522 | .42831857 | .13544622 |
| Hd disease | 10 | 7.2749623 | .29112666 | .09206233 |
| logZNF496 | Control | 10 | 6.7180503 | .41527363 | .13132105 |
| Hd disease | 10 | 6.7970659 | .45868693 | .14504954 |
| logCRYM | Control | 10 | 6.7134174 | .38884591 | .12296387 |
| Hd disease | 10 | 6.5558833 | .51267479 | .16212200 |
| logSCIN | Control | 10 | 9.4811367 | .63814842 | .20180025 |
| Hd disease | 10 | 9.4597946 | .49206198 | .15560366 |
| logAPOH | Control | 10 | 5.1387040 | .49960108 | .15798773 |
| Hd disease | 10 | 4.9579693 | 1.78339635 | .56395944 |
| logSMAD7 | Control | 10 | 9.0391182 | .30945907 | .09785955 |
| Hd disease | 10 | 9.5347445 | .47407757 | .14991649 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Independent Samples Test** | | | | | |
|  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | |
| F | Sig. | t | df |
|
| logTHBS3 | Equal variances assumed | .444 | .514 | 2.655 | 18 |
| Equal variances not assumed |  |  | 2.655 | 15.455 |
| logSLC38A2 | Equal variances assumed | 1.498 | .237 | -2.874 | 18 |
| Equal variances not assumed |  |  | -2.874 | 14.624 |
| logXG | Equal variances assumed | .021 | .888 | -1.232 | 18 |
| Equal variances not assumed |  |  | -1.232 | 17.963 |
| logDDR1 | Equal variances assumed | 4.395 | .050 | -6.401 | 18 |
| Equal variances not assumed |  |  | -6.401 | 15.080 |
| logCCNB1 | Equal variances assumed | .198 | .662 | 3.716 | 18 |
| Equal variances not assumed |  |  | 3.716 | 15.853 |
| logZNF496 | Equal variances assumed | .092 | .765 | -.404 | 18 |
| Equal variances not assumed |  |  | -.404 | 17.825 |
| logCRYM | Equal variances assumed | .764 | .394 | .774 | 18 |
| Equal variances not assumed |  |  | .774 | 16.780 |
| logSCIN | Equal variances assumed | .463 | .505 | .084 | 18 |
| Equal variances not assumed |  |  | .084 | 16.907 |
| logAPOH | Equal variances assumed | 1.501 | .236 | .309 | 18 |
| Equal variances not assumed |  |  | .309 | 10.404 |
| logSMAD7 | Equal variances assumed | 2.954 | .103 | -2.768 | 18 |
| Equal variances not assumed |  |  | -2.768 | 15.491 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Independent Samples Test** | | | | | |
|  | | t-test for Equality of Means | | | |
| Significance | | Mean Difference | Std. Error Difference |
| One-Sided p | Two-Sided p |
| logTHBS3 | Equal variances assumed | .008 | .016 | .59564637 | .22434555 |
| Equal variances not assumed | .009 | .018 | .59564637 | .22434555 |
| logSLC38A2 | Equal variances assumed | .005 | .010 | -.92141029 | .32056974 |
| Equal variances not assumed | .006 | .012 | -.92141029 | .32056974 |
| logXG | Equal variances assumed | .117 | .234 | -1.07127928 | .86959119 |
| Equal variances not assumed | .117 | .234 | -1.07127928 | .86959119 |
| logDDR1 | Equal variances assumed | <.001 | <.001 | -1.19778011 | .18712090 |
| Equal variances not assumed | <.001 | <.001 | -1.19778011 | .18712090 |
| logCCNB1 | Equal variances assumed | <.001 | .002 | .60858984 | .16377165 |
| Equal variances not assumed | <.001 | .002 | .60858984 | .16377165 |
| logZNF496 | Equal variances assumed | .346 | .691 | -.07901551 | .19566448 |
| Equal variances not assumed | .346 | .691 | -.07901551 | .19566448 |
| logCRYM | Equal variances assumed | .224 | .449 | .15753409 | .20347889 |
| Equal variances not assumed | .225 | .450 | .15753409 | .20347889 |
| logSCIN | Equal variances assumed | .467 | .934 | .02134208 | .25482512 |
| Equal variances not assumed | .467 | .934 | .02134208 | .25482512 |
| logAPOH | Equal variances assumed | .381 | .761 | .18073466 | .58567088 |
| Equal variances not assumed | .382 | .764 | .18073466 | .58567088 |
| logSMAD7 | Equal variances assumed | .006 | .013 | -.49562627 | .17902918 |
| Equal variances not assumed | .007 | .014 | -.49562627 | .17902918 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Independent Samples Test** | | | |
|  | | t-test for Equality of Means | |
| 95% Confidence Interval of the Difference | |
| Lower | Upper |
| logTHBS3 | Equal variances assumed | .12431387 | 1.06697888 |
| Equal variances not assumed | .11868884 | 1.07260391 |
| logSLC38A2 | Equal variances assumed | -1.59490232 | -.24791827 |
| Equal variances not assumed | -1.60622206 | -.23659853 |
| logXG | Equal variances assumed | -2.89822258 | .75566401 |
| Equal variances not assumed | -2.89849386 | .75593529 |
| logDDR1 | Equal variances assumed | -1.59090655 | -.80465368 |
| Equal variances not assumed | -1.59643469 | -.79912553 |
| logCCNB1 | Equal variances assumed | .26451838 | .95266131 |
| Equal variances not assumed | .26114782 | .95603187 |
| logZNF496 | Equal variances assumed | -.49009133 | .33206031 |
| Equal variances not assumed | -.49038092 | .33234990 |
| logCRYM | Equal variances assumed | -.26995919 | .58502737 |
| Equal variances not assumed | -.27219774 | .58726592 |
| logSCIN | Equal variances assumed | -.51402562 | .55670979 |
| Equal variances not assumed | -.51651738 | .55920154 |
| logAPOH | Equal variances assumed | -1.04971420 | 1.41118352 |
| Equal variances not assumed | -1.11738576 | 1.47885508 |
| logSMAD7 | Equal variances assumed | -.87175261 | -.11949993 |
| Equal variances not assumed | -.87616678 | -.11508577 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Independent Samples Effect Sizes** | | | | | |
|  | | Standardizera | Point Estimate | 95% Confidence Interval | |
| Lower | Upper |
| logTHBS3 | Cohen's d | .50165190 | 1.187 | .216 | 2.131 |
| Hedges' correction | .52383832 | 1.137 | .207 | 2.041 |
| Glass's delta | .59478679 | 1.001 | -.008 | 1.968 |
| logSLC38A2 | Cohen's d | .71681572 | -1.285 | -2.242 | -.300 |
| Hedges' correction | .74851814 | -1.231 | -2.147 | -.288 |
| Glass's delta | .87218456 | -1.056 | -2.034 | -.035 |
| logXG | Cohen's d | 1.94446502 | -.551 | -1.438 | .351 |
| Hedges' correction | 2.03046235 | -.528 | -1.377 | .336 |
| Glass's delta | 1.98821972 | -.539 | -1.435 | .385 |
| logDDR1 | Cohen's d | .41841506 | -2.863 | -4.120 | -1.567 |
| Hedges' correction | .43692019 | -2.741 | -3.946 | -1.501 |
| Glass's delta | .50210493 | -2.386 | -3.755 | -.967 |
| logCCNB1 | Cohen's d | .36620454 | 1.662 | .616 | 2.674 |
| Hedges' correction | .38240057 | 1.591 | .590 | 2.561 |
| Glass's delta | .29112666 | 2.090 | .772 | 3.358 |
| logZNF496 | Cohen's d | .43751908 | -.181 | -1.057 | .700 |
| Hedges' correction | .45686912 | -.173 | -1.012 | .671 |
| Glass's delta | .45868693 | -.172 | -1.048 | .712 |
| logCRYM | Cohen's d | .45499263 | .346 | -.542 | 1.225 |
| Hedges' correction | .47511546 | .332 | -.519 | 1.173 |
| Glass's delta | .51267479 | .307 | -.589 | 1.187 |
| logSCIN | Cohen's d | .56980628 | .037 | -.840 | .914 |
| Hedges' correction | .59500695 | .036 | -.804 | .875 |
| Glass's delta | .49206198 | .043 | -.835 | .919 |
| logAPOH | Cohen's d | 1.30959990 | .138 | -.742 | 1.014 |
| Hedges' correction | 1.36751922 | .132 | -.710 | .971 |
| Glass's delta | 1.78339635 | .101 | -.779 | .976 |
| logSMAD7 | Cohen's d | .40032141 | -1.238 | -2.188 | -.260 |
| Hedges' correction | .41802631 | -1.186 | -2.095 | -.249 |
| Glass's delta | .47407757 | -1.045 | -2.021 | -.026 |

**Nonparametric Tests**

|  |  |  |  |
| --- | --- | --- | --- |
| **Hypothesis Test Summary** | | | |
|  | Null Hypothesis | Test | Sig.a,b |
| 1 | The distribution of logXG is the same across categories of Control. | Independent-Samples Mann-Whitney U Test | .165c |

|  |  |
| --- | --- |
| **Hypothesis Test Summary** | |
|  | Decision |
| 1 | Retain the null hypothesis. |
|  |  | |  |  |
|  |  | |  |  |
|  |  | |  |  |

**Independent-Samples Mann-Whitney U Test**

**logXG across Control**

|  |  |
| --- | --- |
| **Independent-Samples Mann-Whitney U Test Summary** | |
| Total N | 20 |
| Mann-Whitney U | 68.500 |
| Wilcoxon W | 123.500 |
| Test Statistic | 68.500 |
| Standard Error | 12.804 |
| Standardized Test Statistic | 1.445 |
| Asymptotic Sig.(2-sided test) | .149 |
| Exact Sig.(2-sided test) | .165 |

img.emf

img.emf

img.emf

**SHOW**

img.emf

**4.0 DISCUSSION**

The log2 transformation was applied to the data in this study because it is a widely used and standard method in many tools and software for gene expression analysis. It allows for an easy interpretation of fold change, where a log2 difference of 1 means a 2-fold change, 2 means a 4-fold change, and so on. Five genes (THBS3, SLC38A, DDR1, CCNB1, SMAD7) were statistically analyzed in Huntington's disease (HD) and showed significant differential expression of 0.016, 0.010, <0.01 ,0.02, and 0.013 respectively, and effect sizes of 1.187, 1.285, 2.863, 1.662, and 1.238 (Cohen's) respectively. Using a significance level of 0.05, these genes exhibit a small but significant dysregulation in HD. This dysregulation, marked by moderate effect sizes (Cohen's d ~ 1.2-2.9), indicates a minor but statistically meaningful impact on gene activity. On the other hand, XG, ZNF496, CRYM, SCIN, and APOH did not show significant results, suggesting their possible lack of role in HD pathogenesis. These results support the null hypothesis rejection, confirming the alteration of specific genes in HD.

**5.0 CONCLUSION:**

Out of the 10 genes examined, dysregulated (up- or down-regulated) THBS3, SLC38A2, CCNB1, SMAD7, and DDR1 were shown to be associated with the pathological state of Huntington disease, but the remaining five were not. The findings suggest that altered expression of certain genes, particularly DDR1, might play a role in Huntington's disease. Further research is needed to explore the functional significance of these differences and their potential use as diagnostic or therapeutic targets and pave the way for improved disease management.

**REFERENCES**

Allen, P., Bennett, K. and Heritage, B., 2018. *SPSS statistics: A practical guide with student resource access 12 months*. Cengage AU.

Agyemang, C. O., Owusu-Agyei, S., Amoyaw, Y., Koram, K. A., & Osei, K. (2019). Sex differences in neonatal anthropometry in a large African cohort. PLoS One, 14(3), e0211058. <https://doi.org/10.1371/journal.pone.0211058>

Field, A., 2013. *Discovering statistics using IBM SPSS statistics*. sage.

Pallant, J., 2020. *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. McGraw-hill education (UK).