Pakorn Sagulkoo 6380064520

# Effect of Cognitive Behavioural and Family Treatment on Weight Changes in Anorexia nervosa

# Outline

- Introduction
- Data analysis by R
- Regression analysis by R
- Discussion and Conclusion

# What is anorexia nervosa?

- Serious psychiatric eating disorder
- Usually occurs in adolescent and adult female
- Clinical features
  - Severe weight loss
  - Malnutrition
  - Amenorrhea
  - Electrolyte and acid-base imbalance

# What is anorexia nervosa?

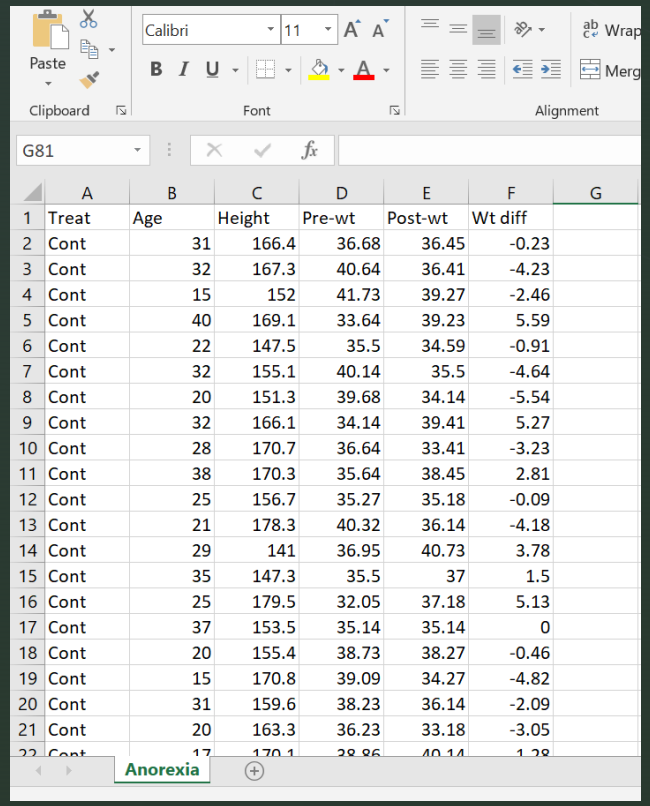
- Weight loss taken place by
  - diminished food intake caused by fear of overweight perception
- Treatment : psychiatrically pharmacological medication
- Poor prognosis
  - Patients usually deny eating
  - Causes severe malnutrition
  - End up with death from severe electrolyte and acid-base disturbance.

# Study Method

- Randomized control trial
- Separate patients into 3 groups
  - Control
  - Cognitive behavioural therapy (CBT)
  - Family treatment (FT)
- Sample size of each groups are 30 patients
- Patients are measure the weight before and after treatment
- Duration of treatment is 30 days

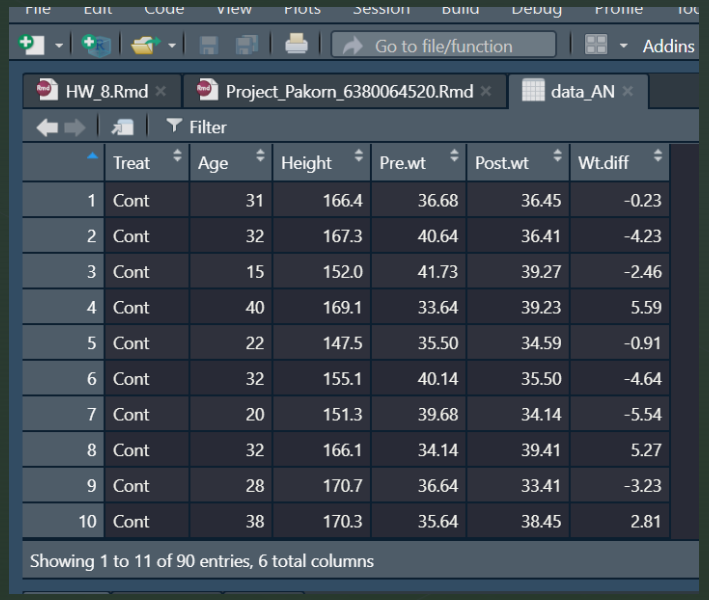
# Data Analysis

- Import data frame from pivot table in Excel



| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | Treat | Age | Height | Pre-wt | Post-wt | Wt diff | |
| 2 | Cont | 31 | 166.4 | 36.68 | 36.45 | -0.23 | |
| 3 | Cont | 32 | 167.3 | 40.64 | 36.41 | -4.23 | |
| 4 | Cont | 15 | 152 | 41.73 | 39.27 | -2.46 | |
| 5 | Cont | 40 | 169.1 | 33.64 | 39.23 | 5.59 | |
| 6 | Cont | 22 | 147.5 | 35.5 | 34.59 | -0.91 | |
| 7 | Cont | 32 | 155.1 | 40.14 | 35.5 | -4.64 | |
| 8 | Cont | 20 | 151.3 | 39.68 | 34.14 | -5.54 | |
| 9 | Cont | 32 | 166.1 | 34.14 | 39.41 | 5.27 | |
| 10 | Cont | 28 | 170.7 | 36.64 | 33.41 | -3.23 | |
| 11 | Cont | 38 | 170.3 | 35.64 | 38.45 | 2.81 | |
| 12 | Cont | 25 | 156.7 | 35.27 | 35.18 | -0.09 | |
| 13 | Cont | 21 | 178.3 | 40.32 | 36.14 | -4.18 | |
| 14 | Cont | 29 | 141 | 36.95 | 40.73 | 3.78 | |
| 15 | Cont | 35 | 147.3 | 35.5 | 37 | 1.5 | |
| 16 | Cont | 25 | 179.5 | 32.05 | 37.18 | 5.13 | |
| 17 | Cont | 37 | 153.5 | 35.14 | 35.14 | 0 | |
| 18 | Cont | 20 | 155.4 | 38.73 | 38.27 | -0.46 | |
| 19 | Cont | 15 | 170.8 | 39.09 | 34.27 | -4.82 | |
| 20 | Cont | 31 | 159.6 | 38.23 | 36.14 | -2.09 | |
| 21 | Cont | 20 | 163.3 | 36.23 | 33.18 | -3.05 | |

```
data_AN <- read.csv("D:\\Bioinfor MSc\\Stat for Bioinfo\\Project\\Pivot table
\\Anorexia.csv")
attach(data_AN)
head(data_AN)
```



| | Treat | Age | Height | Pre.wt | Post.wt | Wt.diff |
|---|---|---|---|---|---|---|
| 1 | Cont | 31 | 166.4 | 36.68 | 36.45 | -0.23 |
| 2 | Cont | 32 | 167.3 | 40.64 | 36.41 | -4.23 |
| 3 | Cont | 15 | 152.0 | 41.73 | 39.27 | -2.46 |
| 4 | Cont | 40 | 169.1 | 33.64 | 39.23 | 5.59 |
| 5 | Cont | 22 | 147.5 | 35.50 | 34.59 | -0.91 |
| 6 | Cont | 32 | 155.1 | 40.14 | 35.50 | -4.64 |
| 7 | Cont | 20 | 151.3 | 39.68 | 34.14 | -5.54 |
| 8 | Cont | 32 | 166.1 | 34.14 | 39.41 | 5.27 |
| 9 | Cont | 28 | 170.7 | 36.64 | 33.41 | -3.23 |
| 10 | Cont | 38 | 170.3 | 35.64 | 38.45 | 2.81 |

Showing 1 to 11 of 90 entries, 6 total columns

# Data Analysis

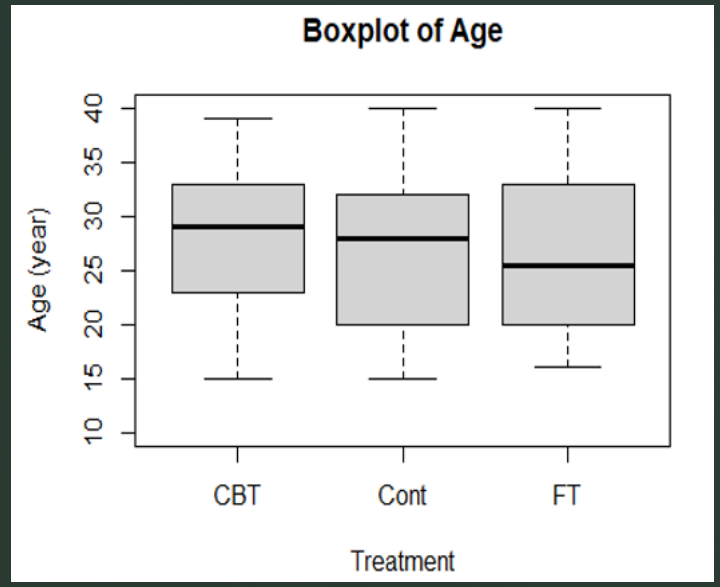
- General summary of dataframe

```
summary(data_AN)

##     Treat                  Age              Height          Pre.wt
##  Length:90          Min.   :15.00   Min.   :141.0   Min.   :31.82
##  Class :character   1st Qu.:21.00   1st Qu.:156.0   1st Qu.:36.02
##  Mode  :character   Median :28.00   Median :165.5   Median :37.52
##                     Mean   :27.21   Mean   :163.0   Mean   :37.36
##                     3rd Qu.:32.75   3rd Qu.:170.4   3rd Qu.:39.03
##                     Max.   :40.00   Max.   :179.5   Max.   :43.14
##     Post.wt         Wt.diff
##  Min.   :32.41   Min.   :-5.540
##  1st Qu.:36.21   1st Qu.:-0.595
##  Median :38.62   Median : 1.150
##  Mean   :39.03   Mean   : 1.668
##  3rd Qu.:41.63   3rd Qu.: 4.228
##  Max.   :48.90   Max.   :14.880
```

# Data Analysis

- Test randomization
  - Age

```
names(data_AN)

## [1] "Treat"   "Age"     "Height"  "Pre.wt"  "Post.wt" "Wt.diff"

plot(factor(Treat),Age,xlab="Treatment",ylab="Age (year)", main="Boxplot of A
ge",ylim=c(10,40))
```



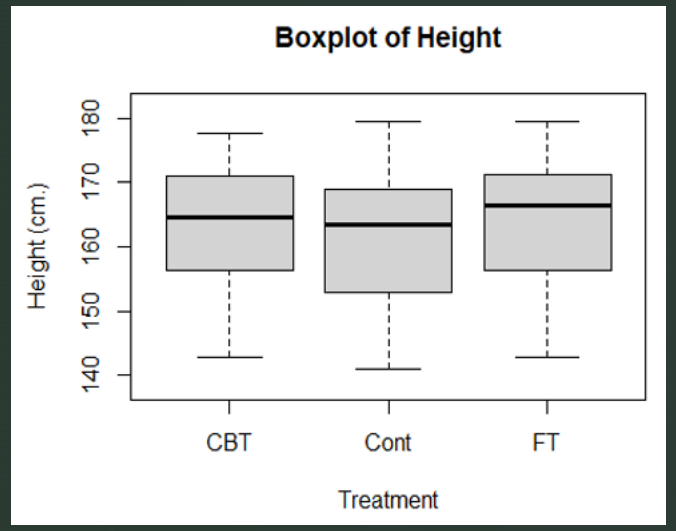

```
summary(aov(Age~Treat))

##             Df Sum Sq Mean Sq F value Pr(>F)
## Treat        2     38   19.03   0.367  0.694
## Residuals   87   4513   51.87
```

  - Because p-value is more than 0.05
    - there are no differences in the mean of age in 3 groups
  - Confirm randomization again by height approach.

# Data Analysis

- Test randomization
  - Height

```
plot(factor(Treat),Height,xlab="Treatment",ylab="Height (cm.)", main="Boxplot
of Height",ylim=c(138,182))
```



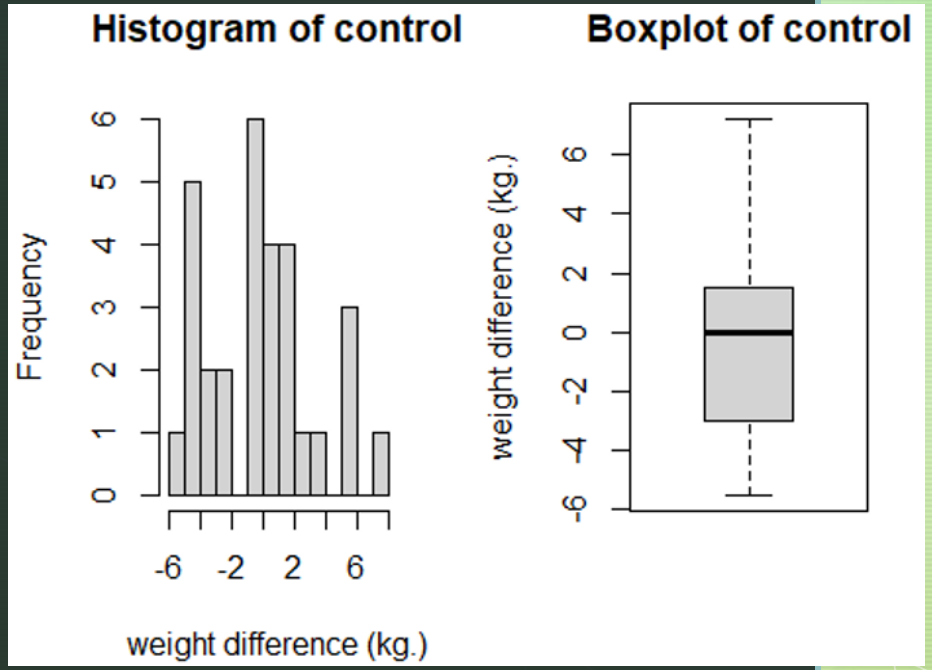

```
summary(aov(Height~Treat))

##             Df Sum Sq Mean Sq F value Pr(>F)
## Treat        2    122   61.05   0.615  0.543
## Residuals   87   8632   99.22
```

  - p-value of ANOVA of height is also more than 0.05.
  - It can be concluded that patients are randomly grouped

# Data Analysis

- Test normality
  - Control group

```
data_ctrl <- data_AN[1:30,6]
summary(data_ctrl)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -5.540  -2.902  -0.025  -0.103   1.445   7.230

par(mfrow=c(1,2))
hist(data_ctrl,breaks=(-6:8),xlab="weight difference (kg.)",main="Histogram o
f control")
boxplot(data_ctrl,ylab="weight difference (kg.)",main="Boxplot of control")
```

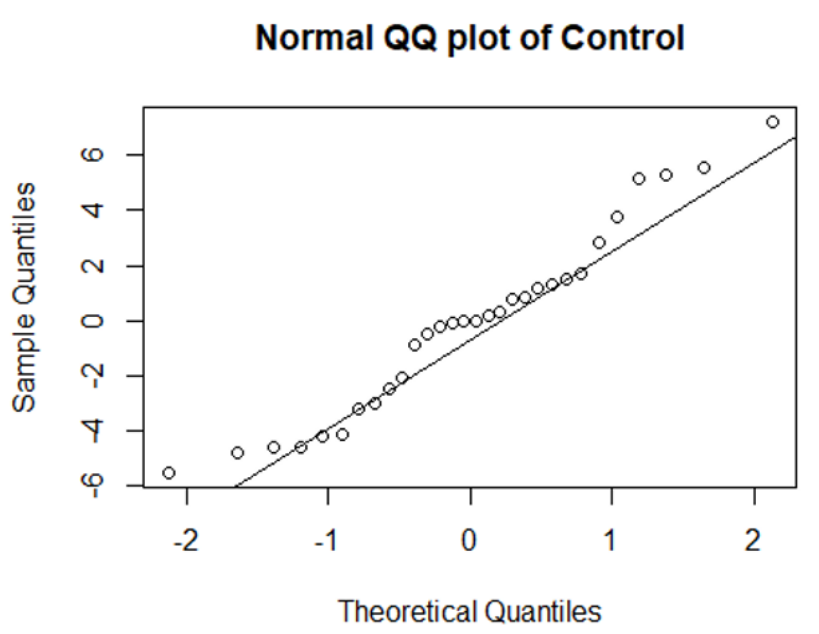



  - Histogram of control group looks like normal distribution
  - As a result, it is tested by QQ plot.

# Data Analysis

- Test normality
  - Control group

```
qqnorm(data_ctrl,main="Normal QQ plot of Control")
qqline(data_ctrl)
```



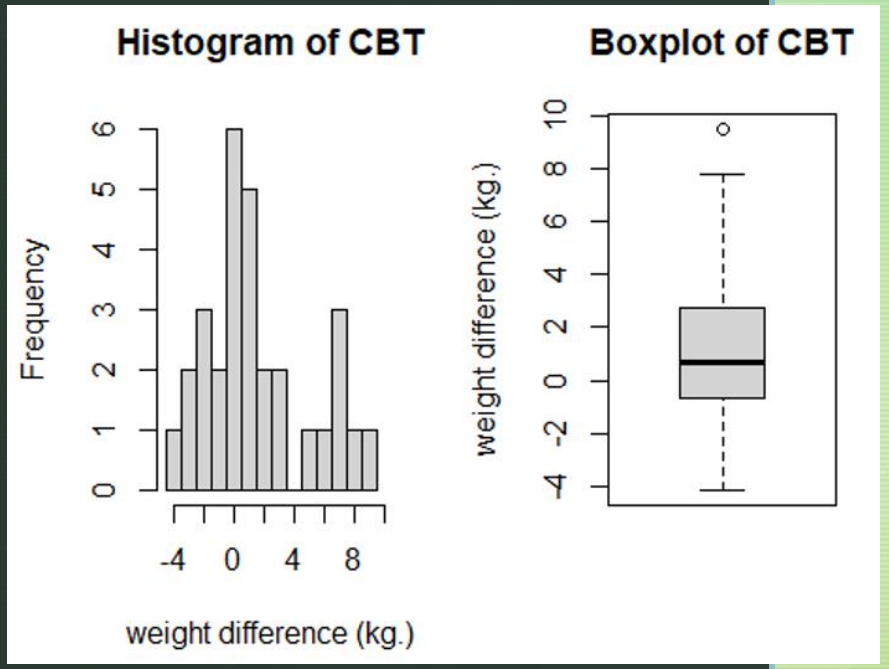

```
shapiro.test(data_ctrl)

##
##  Shapiro-Wilk normality test
##
## data:  data_ctrl
## W = 0.95881, p-value = 0.2888

library(nortest)
ad.test(data_ctrl)

##
##  Anderson-Darling normality test
##
## data:  data_ctrl
## A = 0.41386, p-value = 0.3162
```

- QQ plot looks like normal distribution
  - Confirmed by two statistical tests (Shapiro-Wilk and Anderson test)
- No statistical significance in both normality tests.
  - Can conclude that data of weight change in control is normal distribution.

# Data Analysis

- Test normality
  - CBT group

```
data_cbt <- data_AN[31:60,6]
summary(data_cbt)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -4.140  -0.570   0.685   1.499   2.708   9.500

par(mfrow=c(1,2))
hist(data_cbt,breaks=(-4.5:10),xlab="weight difference (kg.)",main="Histogram
of CBT")
boxplot(data_cbt,ylab="weight difference (kg.)",main="Boxplot of CBT")
```

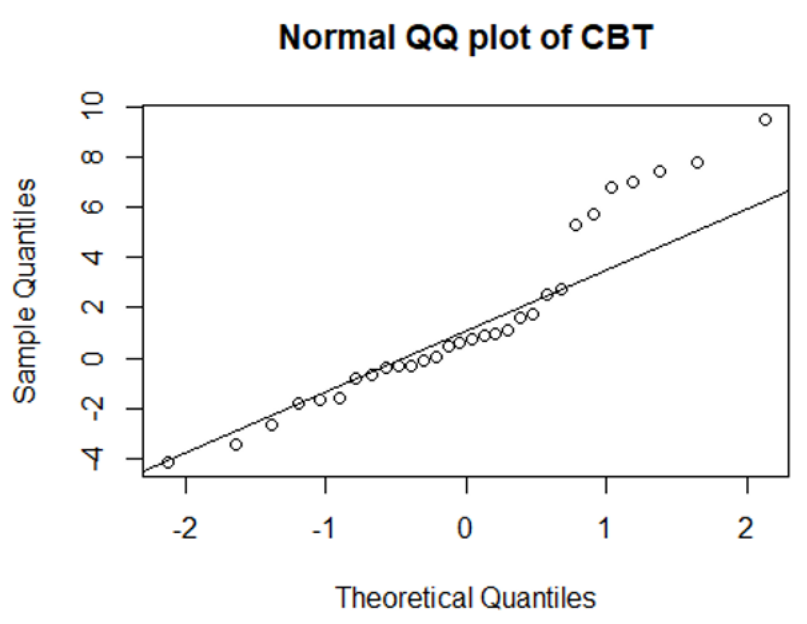



  - Histogram of CBT is not normal distribution (right skewed pattern)
  - Boxplot shows one upper outlier
    - Outlier cannot be deleted because weight before and after treatment are measured accurately.

# Data Analysis

- Test normality
  - CBT group

```
qqnorm(data_cbt,main="Normal QQ plot of CBT")
qqline(data_cbt)
```



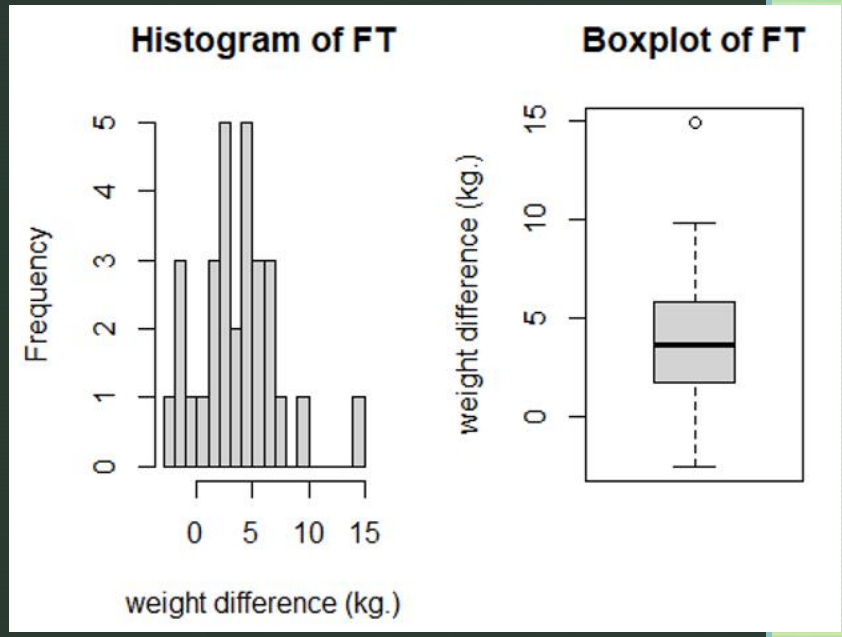

```
shapiro.test(data_cbt)

##
##  Shapiro-Wilk normality test
##
## data:  data_cbt
## W = 0.92155, p-value = 0.02943

ad.test(data_cbt)

##
##  Anderson-Darling normality test
##
## data:  data_cbt
## A = 1.0383, p-value = 0.008431
```

- There is statistical significance in both normality test
- QQ plot and both normality tests tell weight change in CBT group is not normal distribution

# Data Analysis

- Test normality
  - FT group

```
data_ft <- data_AN[61:90,6]
summary(data_ft)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -2.480   1.790   3.675   3.608   5.690  14.880

par(mfrow=c(1,2))
hist(data_ft,breaks=(-3:15.5),xlab="weight difference (kg.)",main="Histogram
of FT")
boxplot(data_ft,ylab="weight difference (kg.)",main="Boxplot of FT")
```

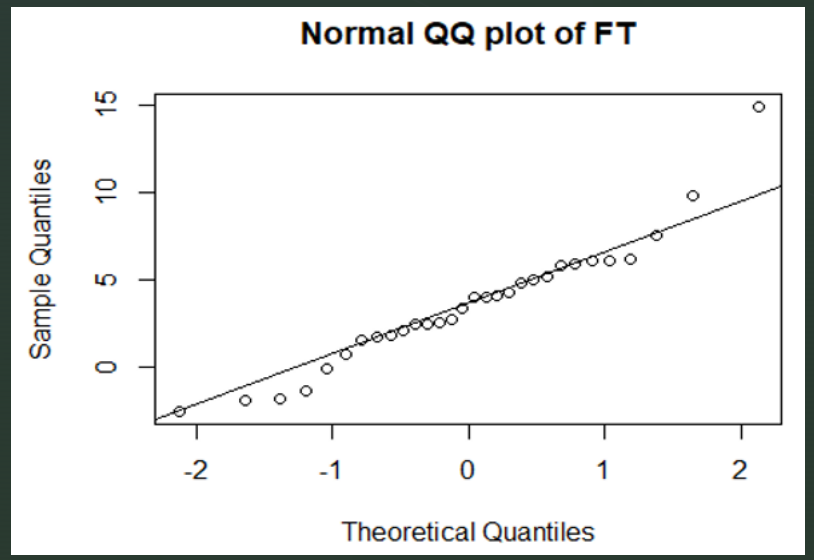



  - Histogram of FT is normal distribution when exclude outlier on right
  - Boxplot shows one upper outlier
    - Nevertheless, outlier cannot be deleted because weight before and after treatment are measured accurately

# Data Analysis

- Test normality
  - FT group

```
qqnorm(data_ft,main="Normal QQ plot of FT")
qqline(data_ft)
```



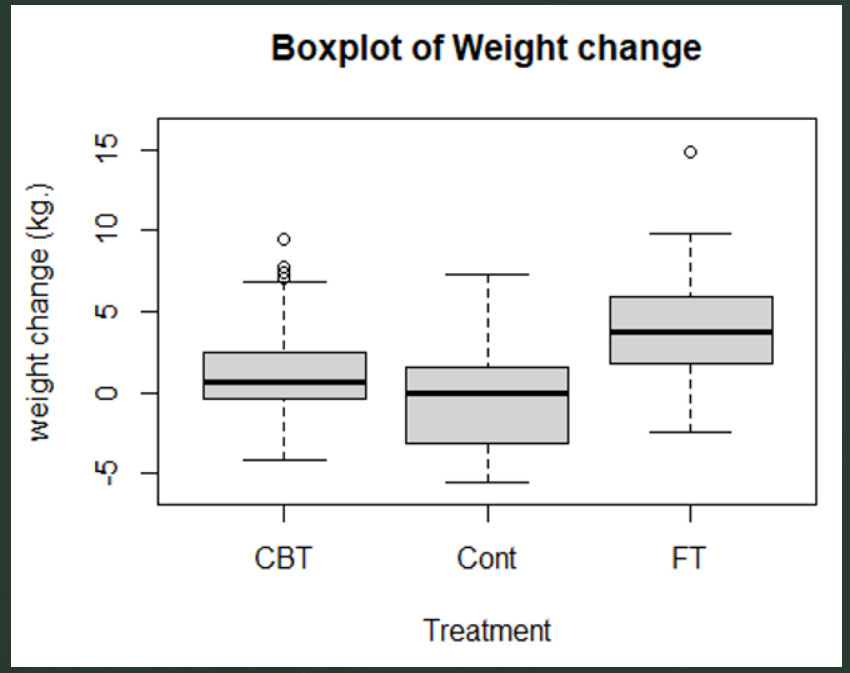

```
shapiro.test(data_ft)

##
##  Shapiro-Wilk normality test
##
## data:  data_ft
## W = 0.94048, p-value = 0.09373

ad.test(data_ft)

##
##  Anderson-Darling normality test
##
## data:  data_ft
## A = 0.44742, p-value = 0.2615
```

- QQ plot and both normality tests provide information th at weight change in FT group is normal distribution

# Data Analysis

- Statistical analysis
  - Because data of weight difference of three groups are not normal distribution,
    - ANOVA cannot be used.
    - Kruskal-Willis test, non parametric test, should be used

```
plot(factor(Treat),Wt.diff,xlab="Treatment",ylab="weight change (kg.)", main=
"Boxplot of Weight change",ylim=c(-6,16))
```



# Data Analysis

- Statistical analysis

```
kruskal.test(Wt.diff ~ Treat, data = data_AN)

##
##  Kruskal-Wallis rank sum test
##
## data:  Wt.diff by Treat
## Kruskal-Wallis chi-squared = 14.645, df = 2, p-value = 0.0006605
```

  - Test shows statistical significance in experiments
  - Then, pairwise-comparison is performed

# Data Analysis

- Statistical analysis

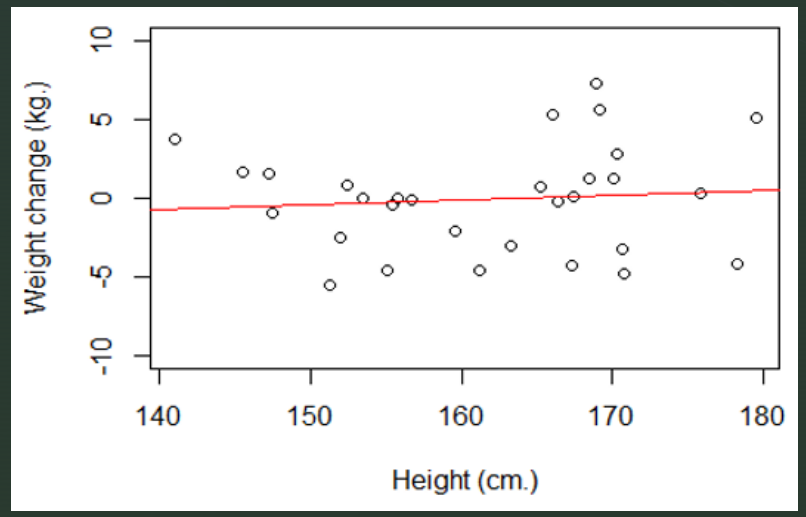
```
pairwise.wilcox.test(Wt.diff,Treat,p.adj= "BH")

##  Pairwise comparisons using Wilcoxon rank sum test with continuity correct
ion
##
## data:  Wt.diff and Treat
##
##      CBT   Cont
## Cont 0.130 -
## FT   0.013 0.001
##
## P value adjustment method: BH
```

  - Pairwise-comparison test gives information that Family treatment has impact on weight changes more than control group in AN patients in statistically significant level
  - However, there are no statistically significant changes in weight between cognitive behavioural therapy and control group

# Regression Analysis

- Analyse relationship between weight changes, and age and height in each group of patients in experiment
- Control group

```
df_ctrl <- data_AN[1:30,]
plot(df_ctrl$Wt.diff ~ df_ctrl$Age,xlab="Age (year)",ylab="Weight change (kg.
)",ylim=range(-10:10))
abline(lm(df_ctrl$Wt.diff ~ df_ctrl$Age),col="red")
```



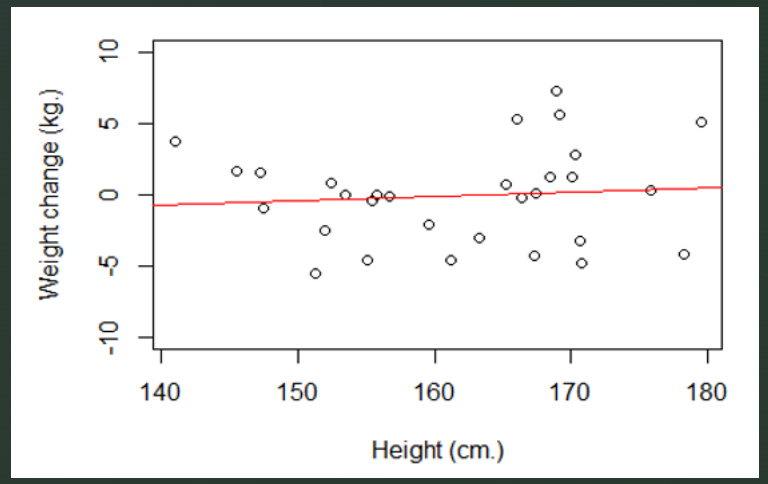

```
summary(lm(df_ctrl$Wt.diff ~ df_ctrl$Height))

##
## Call:
## lm(formula = df_ctrl$Wt.diff ~ df_ctrl$Height)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1291 -2.7624  0.1187  1.8213  7.1215
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -4.87573   10.19614  -0.478    0.636
## df_ctrl$Height  0.02951    0.06292   0.469    0.643
##
## Residual standard error: 3.433 on 28 degrees of freedom
## Multiple R-squared:  0.007794,   Adjusted R-squared:  -0.02764
## F-statistic: 0.2199 on 1 and 28 DF,  p-value: 0.6427
```

- Because R-squared is close to 0, linear model is not fit to data between weight change and height in control group.

# Regression Analysis

- Control group

```
plot(df_ctrl$Wt.diff ~ df_ctrl$Height,xlab="Height (cm.)",ylab="Weight change
(kg.)",ylim=range(-10:10))
abline(lm(df_ctrl$Wt.diff ~ df_ctrl$Height),col="red")
```




```
summary(lm(df_ctrl$Wt.diff ~ df_ctrl$Height))

##
## Call:
## lm(formula = df_ctrl$Wt.diff ~ df_ctrl$Height)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1291 -2.7624  0.1187  1.8213  7.1215
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -4.87573   10.19614  -0.478    0.636
## df_ctrl$Height  0.02951    0.06292   0.469    0.643
##
## Residual standard error: 3.433 on 28 degrees of freedom
## Multiple R-squared:  0.007794,   Adjusted R-squared:  -0.02764
## F-statistic: 0.2199 on 1 and 28 DF,  p-value: 0.6427
```

- Because R-squared is close to 0
- Linear model is not fit to data between weight change and height in control group.

# Regression Analysis

- Control group

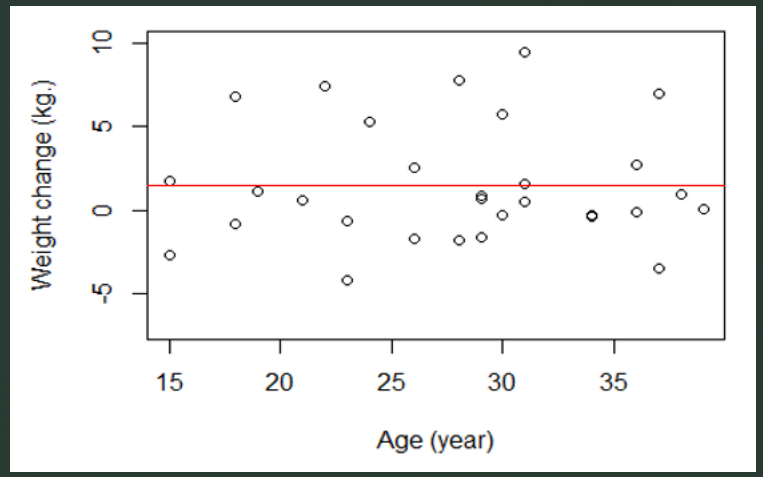
```
summary(lm(df_ctrl$Wt.diff ~ df_ctrl$Age + df_ctrl$Height))

##
## Call:
## lm(formula = df_ctrl$Wt.diff ~ df_ctrl$Age + df_ctrl$Height)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1834 -2.3607  0.1808  1.1927  8.0076
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -11.89570   10.27291  -1.158   0.2570
## df_ctrl$Age      0.16398    0.08084   2.028   0.0525 .
## df_ctrl$Height   0.04544    0.06021   0.755   0.4569
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.256 on 27 degrees of freedom
## Multiple R-squared:  0.139,  Adjusted R-squared:  0.07522
## F-statistic: 2.179 on 2 and 27 DF,  p-value: 0.1326
```

- In multiple linear regression (between weight change and age and height), there are weak correlations in model (r square is 0.14)

# Regression Analysis

- CBT group

```
df_cbt <- data_AN[31:60,]
plot(df_cbt$Wt.diff ~ df_cbt$Age,xlab="Age (year)",ylab="Weight change (kg.)"
,ylim=range(-7:10))
abline(lm(df_cbt$Wt.diff ~ df_cbt$Age),col="red")
```



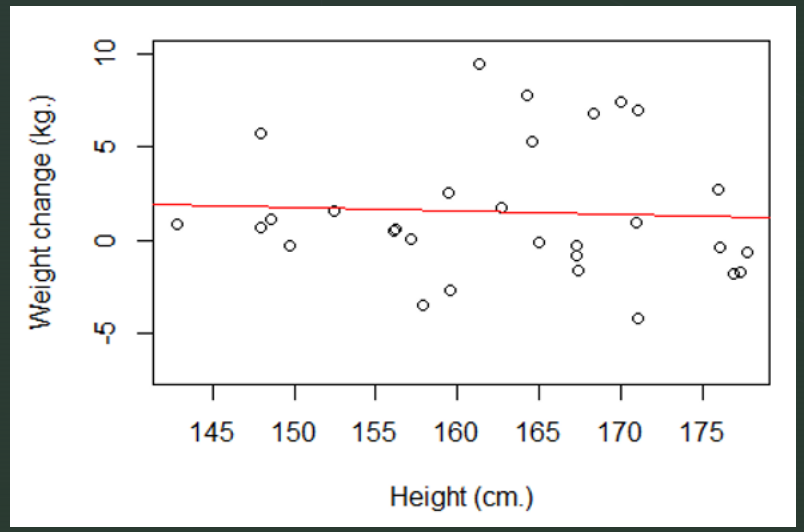

```
summary(lm(df_cbt$Wt.diff ~ df_cbt$Age))

##
## Call:
## lm(formula = df_cbt$Wt.diff ~ df_cbt$Age)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.6257 -2.0633 -0.8063  1.1926  7.9920
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.421580   2.769284   0.513    0.612
## df_cbt$Age  0.002787   0.096412   0.029    0.977
##
## Residual standard error: 3.606 on 28 degrees of freedom
## Multiple R-squared:  2.984e-05,  Adjusted R-squared:  -0.03568
## F-statistic: 0.0008355 on 1 and 28 DF,  p-value: 0.9771
```

- Because R-squared is close to 0
- Linear model is not fit to data between weight change and age in CBT group.

# Regression Analysis

- CBT group

```
plot(df_cbt$Wt.diff ~ df_cbt$Height,xlab="Height (cm.)",ylab="Weight change (
kg.)",ylim=range(-7:10))
abline(lm(df_cbt$Wt.diff ~ df_cbt$Height),col="red")
```




```
summary(lm(df_cbt$Wt.diff ~ df_cbt$Height))

##
## Call:
## lm(formula = df_cbt$Wt.diff ~ df_cbt$Height)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.485 -2.022 -1.040  1.377  7.967
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.65438   11.08223   0.420    0.678
## df_cbt$Height -0.01935    0.06786  -0.285    0.778
##
## Residual standard error: 3.601 on 28 degrees of freedom
## Multiple R-squared:  0.002896,   Adjusted R-squared:  -0.03271
## F-statistic: 0.08134 on 1 and 28 DF,  p-value: 0.7776
```

- Because R-squared is close to 0
- Linear model is not fit to data between weight change and height in CBT group.

# Regression Analysis

- CBT group

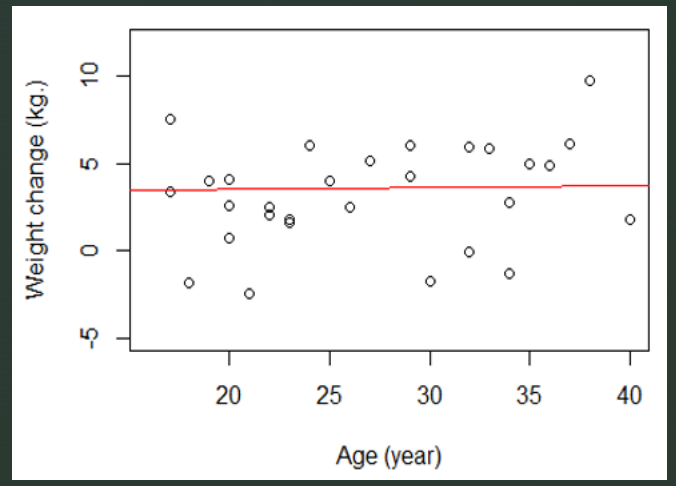
```
summary(lm(df_cbt$Wt.diff ~ df_cbt$Age + df_cbt$Height))

##
## Call:
## lm(formula = df_cbt$Wt.diff ~ df_cbt$Age + df_cbt$Height)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.466 -2.024 -1.046  1.357  7.956
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.565070  11.528676   0.396    0.695
## df_cbt$Age     0.003718   0.098094   0.038    0.970
## df_cbt$Height -0.019442   0.069142  -0.281    0.781
##
## Residual standard error: 3.667 on 27 degrees of freedom
## Multiple R-squared:  0.00295,    Adjusted R-squared:  -0.07091
## F-statistic: 0.03994 on 2 and 27 DF,  p-value: 0.9609
```

- In multiple linear regression (between weight change and age and height), there are weak correlations in model (r square is 0.003)

# Regression Analysis

- FT group

```
df_ft <- data_AN[61:90,]
plot(df_ft$Wt.diff ~ df_ft$Age,xlab="Age (year)",ylab="Weight change (kg.)",y
lim=range(-5:12))
abline(lm(df_ft$Wt.diff ~ df_ft$Age),col="red")
```



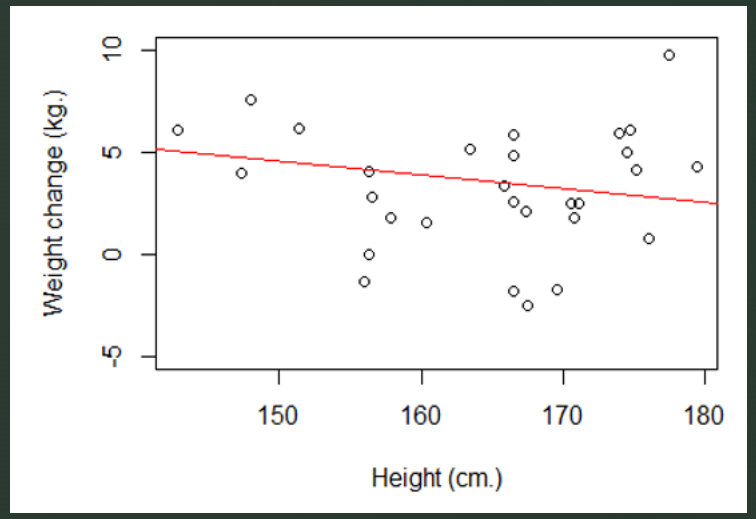

```
summary(lm(df_ft$Wt.diff ~ df_ft$Age))

##
## Call:
## lm(formula = df_ft$Wt.diff ~ df_ft$Age)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.0291 -1.9124  0.1259  2.0317 11.3829
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.33079    2.62274   1.270    0.215
## df_ft$Age    0.01040    0.09523   0.109    0.914
##
## Residual standard error: 3.66 on 28 degrees of freedom
## Multiple R-squared:  0.0004255,  Adjusted R-squared:  -0.03527
## F-statistic: 0.01192 on 1 and 28 DF,  p-value: 0.9138
```

- Because R-squared is close to 0
- Linear model is not fit to data between weight change and age in FT group.

# Regression Analysis

- FT group

```
plot(df_ft$Wt.diff ~ df_ft$Height,xlab="Height (cm.)",ylab="Weight change (kg
.)",ylim=range(-5:10))
abline(lm(df_ft$Wt.diff ~ df_ft$Height),col="red")
```




```
summary(lm(df_ft$Wt.diff ~ df_ft$Height))

##
## Call:
## lm(formula = df_ft$Wt.diff ~ df_ft$Height)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.8651 -1.8713 -0.3912  1.7061 10.2856
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  14.63848   11.03692   1.326    0.195
## df_ft$Height -0.06718    0.06710  -1.001    0.325
##
## Residual standard error: 3.597 on 28 degrees of freedom
## Multiple R-squared:  0.03456,    Adjusted R-squared:  8.423e-05
## F-statistic: 1.002 on 1 and 28 DF,  p-value: 0.3253
```

- Because R-squared is close to 0
- Linear model is not fit to data between weight change and height in FT group.

# Regression Analysis

- FT group

```
summary(lm(df_ft$Wt.diff ~ df_ft$Age + df_ft$Height))

##
## Call:
## lm(formula = df_ft$Wt.diff ~ df_ft$Age + df_ft$Height)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.7305 -1.8126 -0.2373  1.6240 10.4964
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  14.37392   11.28357   1.274    0.214
## df_ft$Age     0.02267    0.09598   0.236    0.815
## df_ft$Height -0.06925    0.06882  -1.006    0.323
##
## Residual standard error: 3.659 on 27 degrees of freedom
## Multiple R-squared:  0.03656,    Adjusted R-squared:  -0.03481
## F-statistic: 0.5122 on 2 and 27 DF,  p-value: 0.6049
```

- In multiple linear regression (between weight change and age and height), there are weak correlations in model (r square is 0.037)

# Conclusion

- Patients are grouped randomly into 3 groups (control, CBT, and FT) because one-way ANOVA of age and height in groups is not statistically significant.
- Statistical analysis reveals that family treatment give result of weight change better than control but cognitive behavioural therapy does not give result better than control.
- In linear regression analysis, there are not any strong relationship between weight change, and age and height of patients in 3 groups.