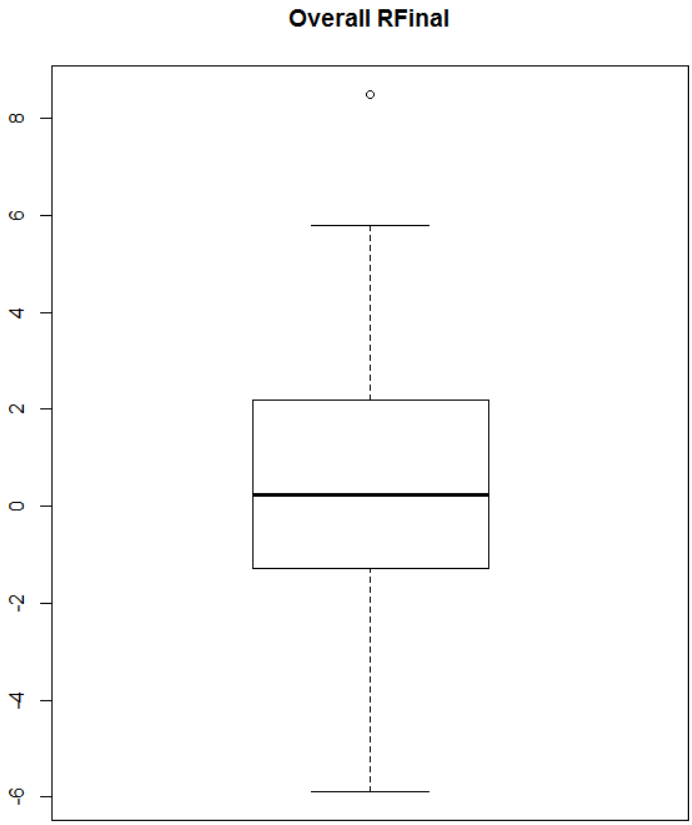
Drugs' Efficiency Comparison in Clinical Trail

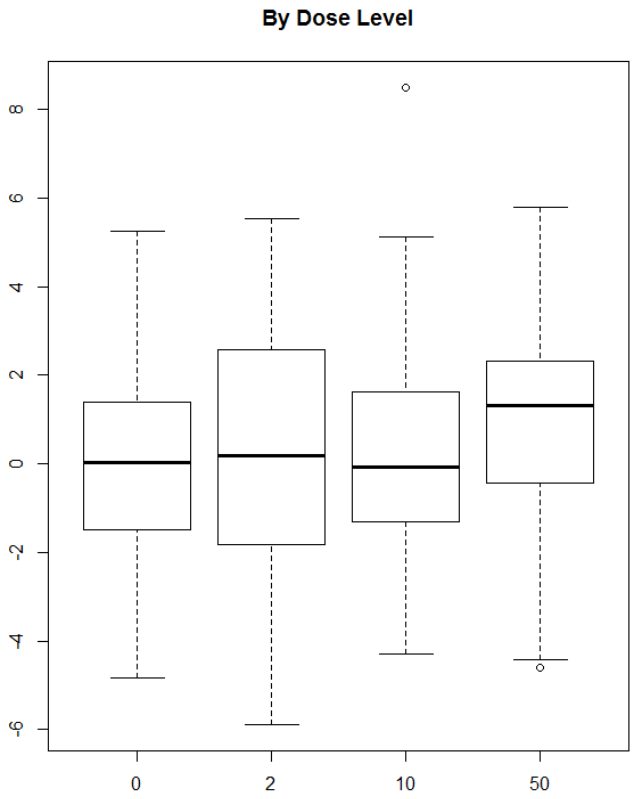
Methodology

One-way ANOVA for linear contrast testing, Tukey HSD test and multivariate T- test.
Because there is just one effect (treatment effect) when we ignore covariates and only consider RFinal, the treatment effect is fixed and the patients are randomly chosen, fit the assumption of independence of observations**,** which means that there is no relationship between the observations in each group or between the groups themselves.
For example, there must be different participants in each group with no participant being in more than one group. And we simply delete all of the missing value.

Assumption check

1. Outlier check





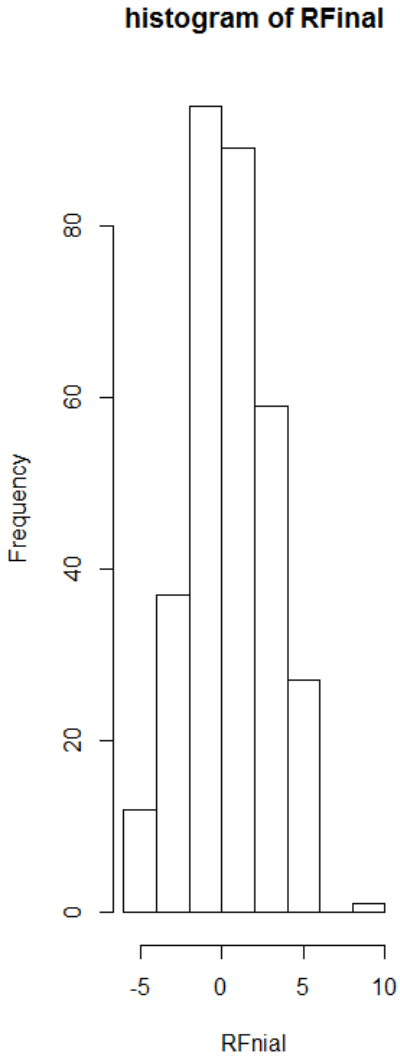
According to the boxplot of each treatment, there is an extreme value in dose 10. Tested by Chi-square test of outliers

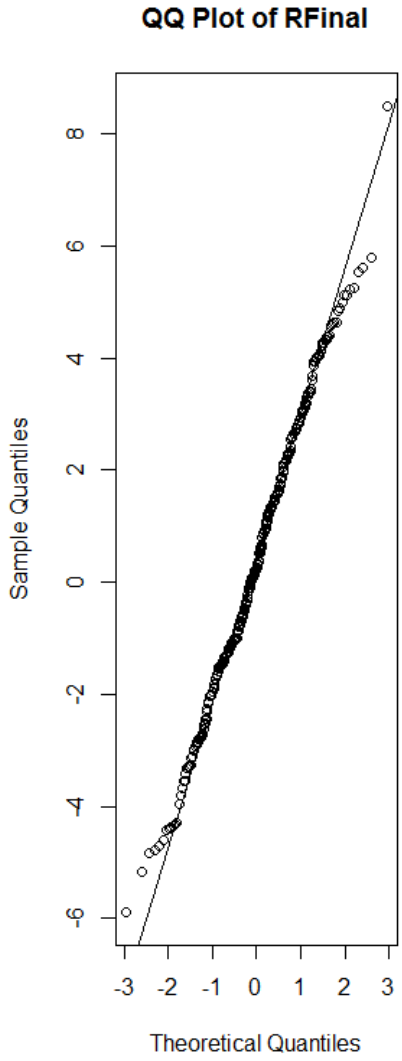
	Test Statistics	P-value	Outlier Value
Dose 10	12.43	0.0004226	8.5
All Dose	10.919	0.000952	8.5

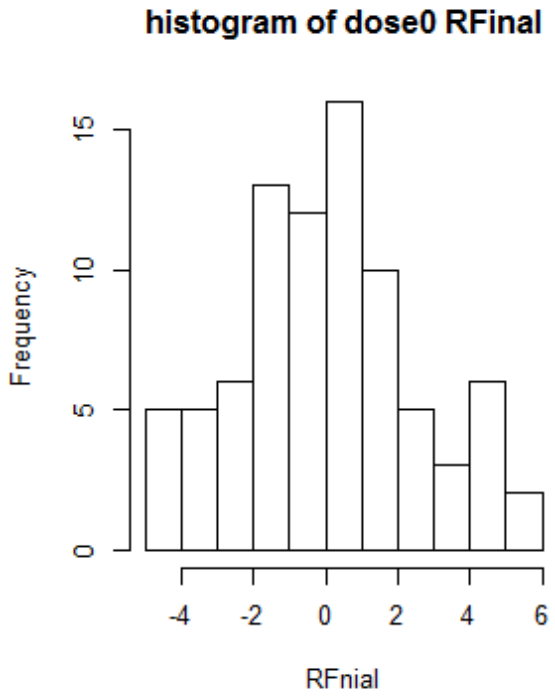
Thus, the point 8.5 is an outlie, moreover, in the latter test of normality, the outlier value 8.5 may influence the normality of the data, in order to fit the assumption of one-way ANOVA, we should delete the outlier.

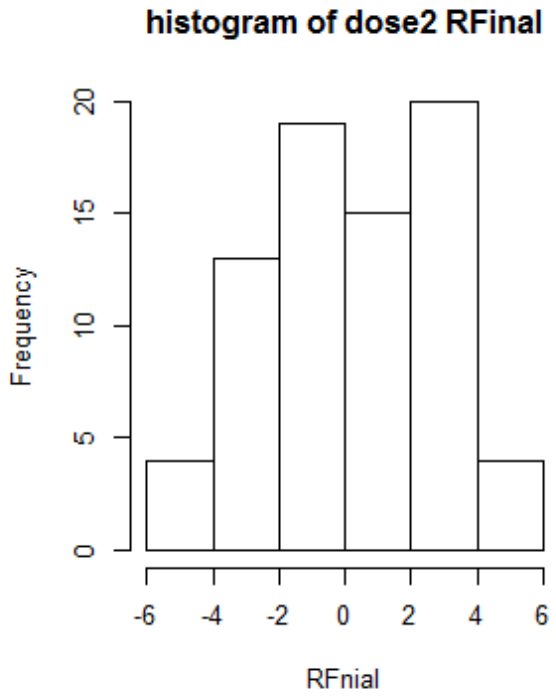
2. normality of response

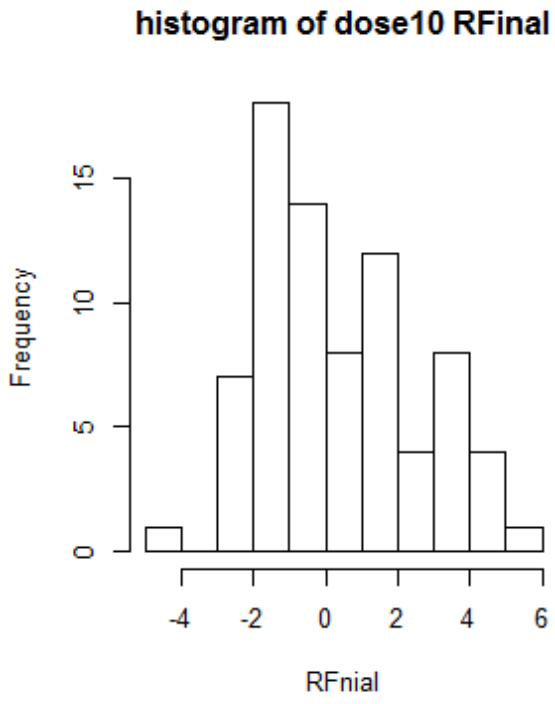
according to the histogram and qq-plot of RFinal, the distribution alike normal distribution.

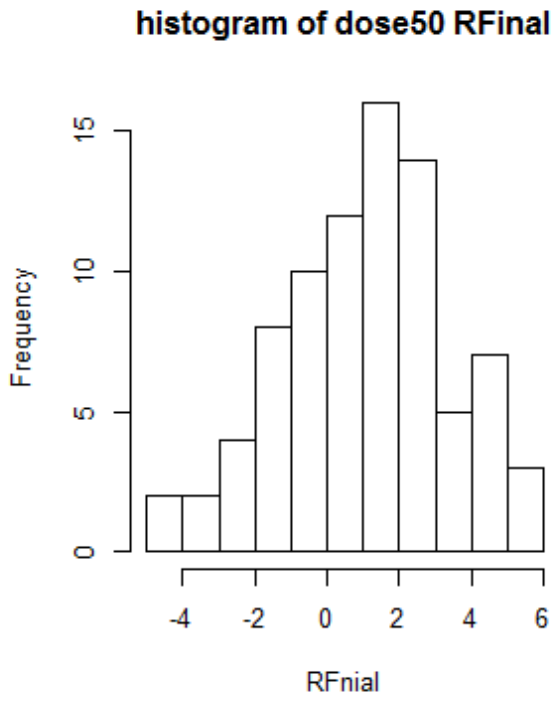

histogram of RFinal
Frequency
0
20
40
60
80
-5
0
5
10
RFnial


QQ Plot of RFinal
Sample Quantiles
-6
-4
-2
0
2
4
6
8
-3
-2
-1
0
1
2
3
Theoretical Quantiles

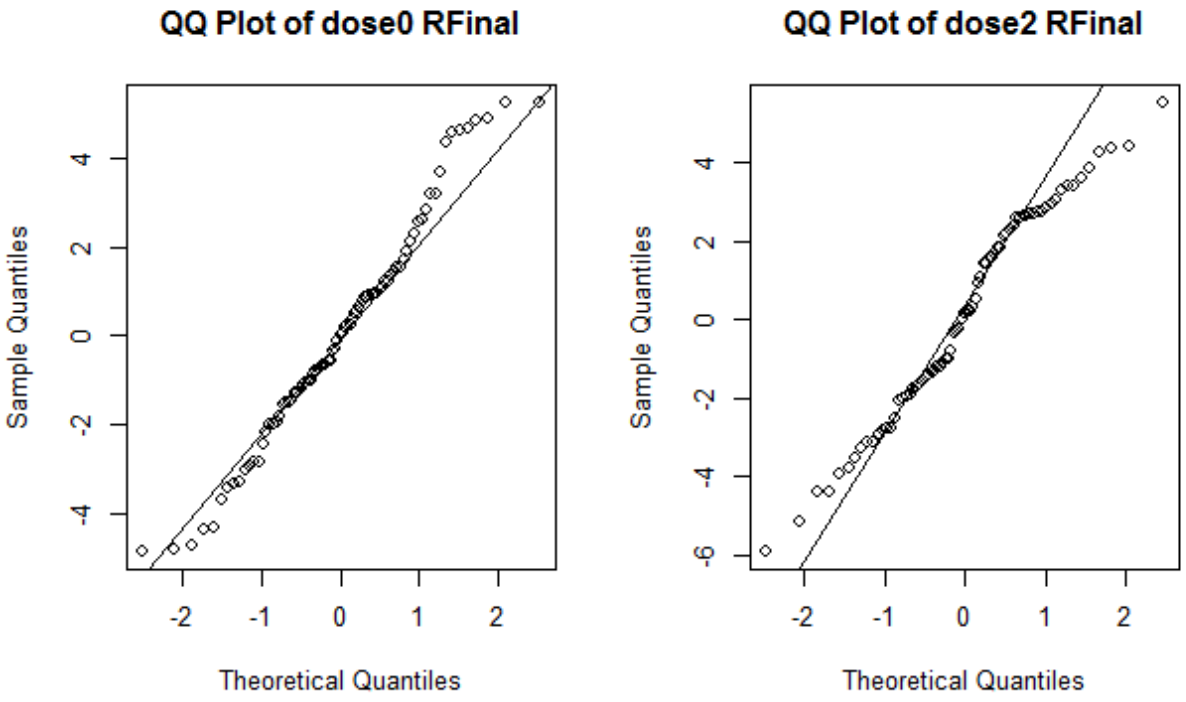

histogram of dose0 RFinal
Frequency
0
5
10
15
-4
-2
0
2
4
6
RFnial


histogram of dose2 RFinal
Frequency
0
5
10
15
20
-6
-4
-2
0
2
4
6
RFnial

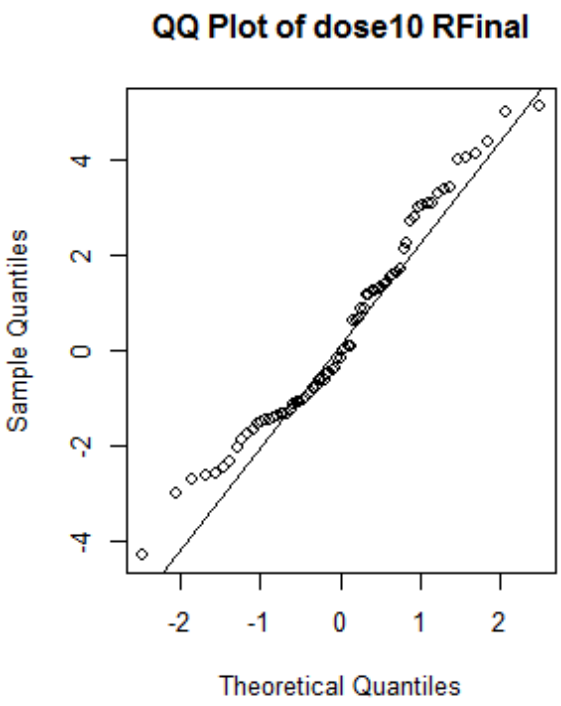

histogram of dose10 RFinal
Frequency
0
5
10
15
-4
-2
0
2
4
6
RFnial

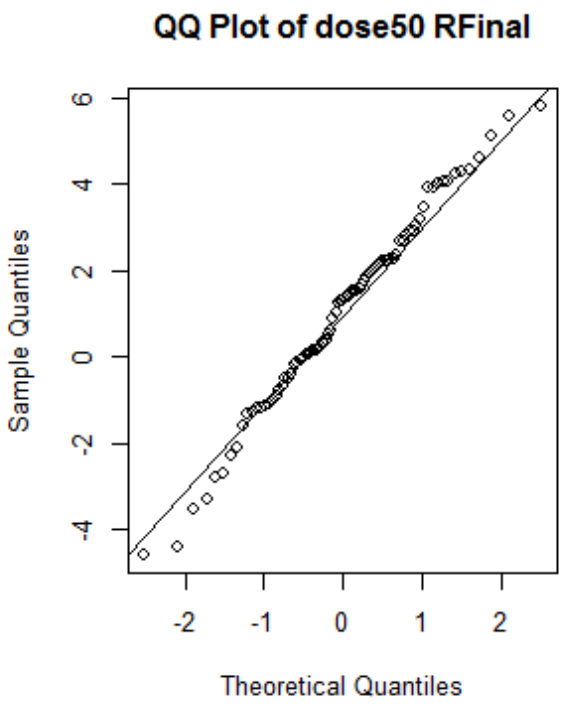

histogram of dose50 RFinal
Frequency
0
5
10
15
-4
-2
0
2
4
6
RFnial

qq plot:


QQ Plot of dose0 RFinal
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2
-4
-2
0
2
4

QQ Plot of dose2 RFinal
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2
-6
-4
-2
0
2
4


QQ Plot of dose10 RFinal
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2
-4
-2
0
2
4


QQ Plot of dose50 RFinal
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2
-4
-2
0
2
4
6

approximately normally distributed

	Over all	Dose 0	Dose 2	Dose 10	Dose 50
Shapiro-Wilk normality test P-value	0.06929	0.1868	0.08635	0.03657	0.6799
Test statistics	0.99164	0.9788	0.97132	0.96588	0.98858

The one-way ANOVA is considered a robust test against the normality assumption. This means that it tolerates violations to its normality assumption rather well when the sample size is not too small.

3. homogeneity of variance

	Test statistics	df	p-value
Bartlett test	4.9894	3	0.1726
levene.test	2.5284		0.05738

(a)

Model conduct

One-way ANOVA model F test.

It is a linear contrast problem and $\sum ci = 0$, One way ANOVA is applied. If no empty cells, contrast and estimate work the same way as with balanced data. In most cases, tests of differences in means are equivalent with balanced and unbalanced data.

	Df	SS	MS	F	P-Value	Significant
Dose	3	50.7	16.89	2.969	0.0321	*
(-4,-4,2,6)	1	45.2	45.21	7.947	0.00512	**
(-1,-1,-1,3)	1	5.3	5.27	0.926	0.33665	
Residuals	314	1786.2	5.69			

Thus, for the Hypothesis test H0: $c1 * \mu 1 + c2 * \mu 2 + c3 * \mu 3 + c4 * \mu 4 = 0$;

verse H1: $c1 * \mu 1 + c2 * \mu 2 + c3 * \mu 3 + c4 * \mu 4 \neq 0$,

the contrast (-1, -1, -1, 3) is not significant, we do not reject H0.

However, the (-4, -4, 2, 6) contrast is significant and we reject the null hypothesis. The alternative hypothesis of interest is H1: $-4 * \mu 1 - 4 * \mu 2 + 2 * \mu 3 + 6 * \mu 4 > 0$.

Thus, we conduct a Tukey HSD to determine which groups differ at 95% confidence interval. Though ANOVA can tell whether groups in the sample differ, it cannot tell which groups differ. If the results of ANOVA are positive in the sense that they state there is a significant difference among the groups, the obvious question becomes: Which groups in this sample differ significantly? It is not likely that all groups differ when compared to each other, only that a handful have significant differences. Tukey's HSD can clarify which groups among the sample in specific have significant differences.

The interest is to test $-4 * \mu 1 - 4 * \mu 2 + 2 * \mu 3 + 6 * \mu 4 > 0$.

With the table obtained from the Tukey HSD test in 95% confidence interval, we can conclude that the contrast is larger than 0 because the sum of difference is positive and the 95% confidence interval does not contain 0.

Tukey multiple comparisons of means 95% family-wise confidence level				
	diff	lwr	upr	p
2 vs 50	3.692556	-0.2331892	7.6183009	0.29475739
0 vs 10	0.599227	2.5486623	-1.3502082	-1.71435967
0 vs 50	1.989398	3.9019317	0.0768635	-0.07583587
contrast sum	6.28118	6.217405	6.344956	-1.495438

The contrast (-1, -1, -1, 3) equals 0, that is the mean of Dose 0, Dose2 and Dose 10 is the same as Dose 50, and the contrast (-4, -4, 2, 6) is larger than 0, in other words, Dose 0 and Dose 2 smaller than Dose 10 and Dose 50.

T test:
Now exam if there is indeed no difference in the second contrast by t test. Linear regression is applied to test the mean contrast. The result is quite different from the ANOVA test, where the ANOVA test has a result only the contrast (-4, -4, 2, 6) significant, but the t-test shows that both contracts are significant.

Linear Hypotheses:					
Dose	estimate	se	t	Pr(>|t|)	significant
(-4,-4,2,6)	5.821	2.952	1.972	0.0741	.
(-1,-1,-1,3)	2.498	1.031	2.423	0.025	*

For the contract that compares the difference between the mean of Dose 0, Dose 2, Dose 10 and Dose 50, we conduct a Tukey HSD test. With the table obtained from the Tukey HSD test in 95% confidence interval, we can conclude that the the mean of Dose 0, Dose 2, Dose 10 is smaller than Dose 50 and the 95% confidence interval does not contain 0.

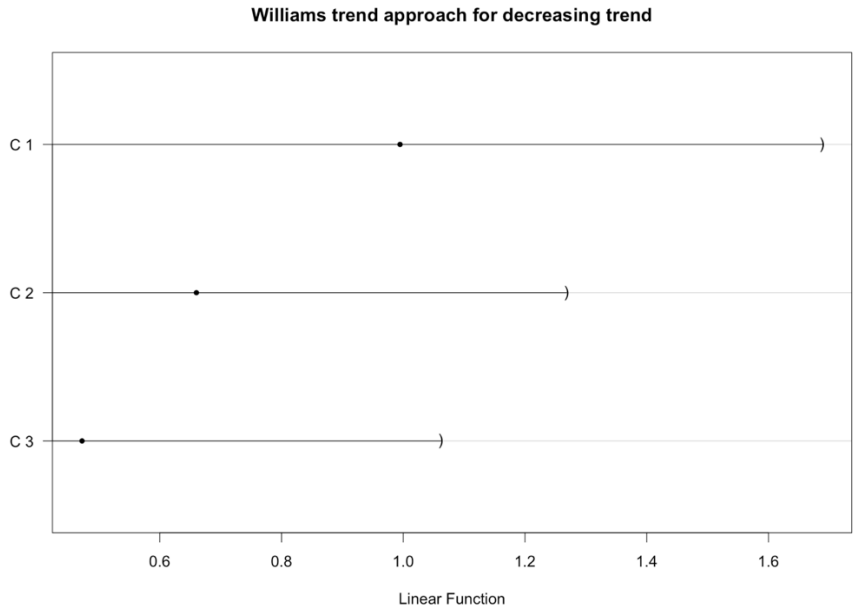
Tukey multiple comparisons of means 95% family-wise confidence level				
	diff	lwr	upr	p
50 vs 2	0.923139	-0.05829731	1.90457522	0.07368935
50 vs 10	0.6950853	-0.27963235	1.6698029	0.25573882
50 vs 0	0.9946988	1.95096584	0.03843175	-0.03791793
sum of contrast	2.612923	1.6130362	3.6128099	0.2915102

To sum up, both contracts larger than 0, the output difference between ANOVA test and T-test is probably due to type of SS in ANOVA. The sample size is unbiased and there may be some bias when we use type I SS instead of type III SS.

(b). Trend test

Original Williams test:

Linear Hypotheses: Original Williams test:				
	Estimate	SE	t value	Pr(<t)
C1>=0	0.9947	0.3665	2.714	0.999
C2>=0	0.6602	0.3214	2.054	0.991
C3>=0	0.4723	0.3116	1.516	0.967



From the output above, it is easy to tell that even the lowest p-value among the tested contrasts is 0.967, suggesting that there is a no evidence for a monotonic increasing trend in the response. The plot of Williams approach shows that the trend is not increasing likely.

Extended William test:

Linear Hypotheses: extended Williams test				
	Estimate	SE	t value	Pr(<t)
C1>=0	0.07156	0.37998	0.188	0.694
C2>=0	0.18709	0.32552	0.575	0.818
C3>=0	0.47233	0.30454	1.551	0.972

Lowest P-value = 0.694, there is no evidence of monotonic increasing trend.

(c). Contrast Matrix

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0.4934211 & 0.5065789 & 0 \\ -1 & 0.3191489 & 0.3276596 & 0.3531915 \end{pmatrix}$$

Select two contrasts:
Dose 0 vs Dose 2 & Dose 10: (-1,0.49,0.51,0)
Dose 0 vs mean of Dose 2, Dose 10 and Dose 50: (-1,0.32,0.33,0.35)

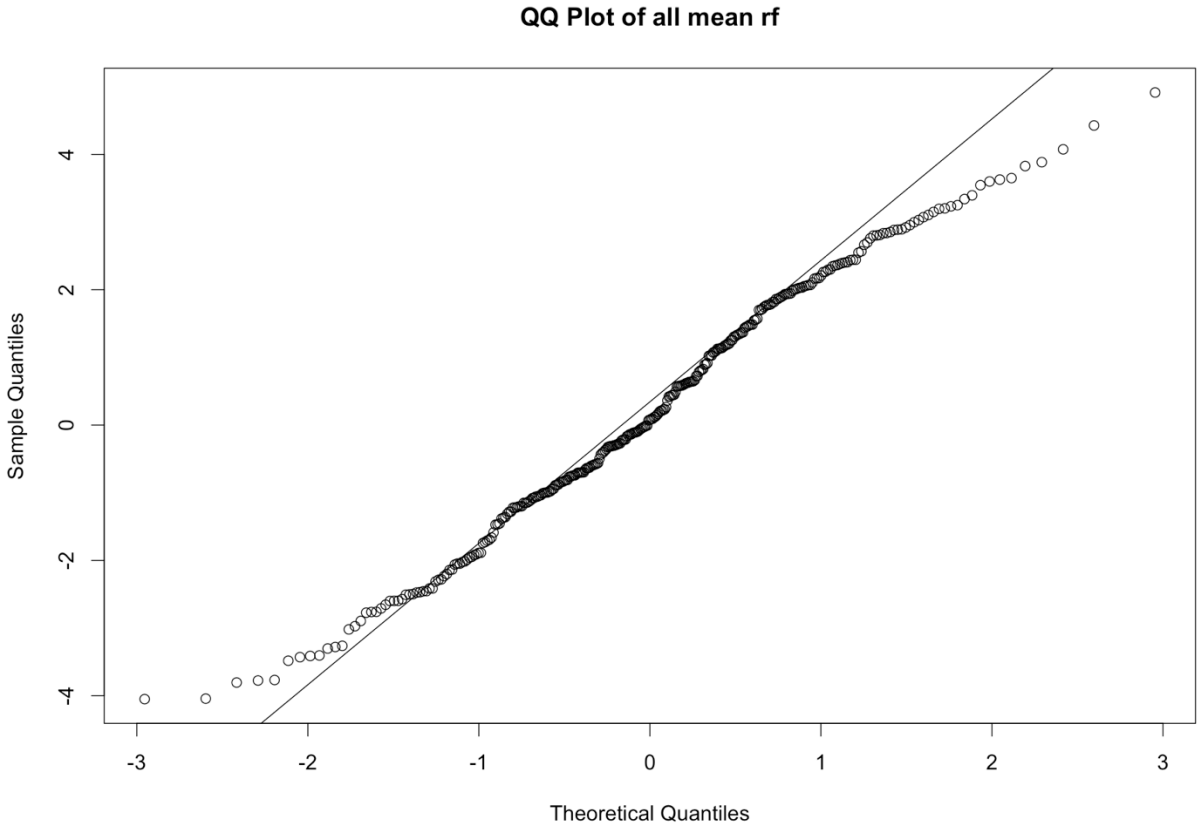
For the max type contrast statistics, T = max{T1,T2} , where Tq and T2 are jointly multivariate t distribution with p =2, denote the final test statistics of the extended Williams test, this test is closely related to the original Williams test but more powerful because it takes the unbalanced sample sizes of all groups into account and the variance estimate is not restricted to the sample size in group 1 and the last group, that is group Dose 0 and group Dose 50. The sum of each row in the contrast matrix are 0, thus, the test is unbiased as (b), each single test consist of comparison between the Dose 0 and the weighted average over the Dose 2, Dose 10 and Dose 50 group, the test performs the same type of contrast as the test in (b).
we can obtain a global test of trend by combining the individual statistics to an overall test result. Since every single test is the most powerful test in its own contrast, and a combination of the most powerful test will be the most sensitive at the alternative region. Thus, the test statistics is similar as (b) but only more sensitive.

P-value:
Since T1, T2 follow a multivariate t distribution with p=2 and df=v, we can apply this distribution to find out the P-value
The p-value is calculated under the null hypothesis H0 : $\mu 1 = \mu 2 = \mu 3$.
The alternative hypothesis is $\mu 1 < \mu 2 < \mu 3 < \mu 4$, $\mu 1 < \mu 4$.
The rejection region: RR = {Tv > Tobs}
Calculate the observed multivariate t test statistics (Tobs) under the null hypothesis $\mu 1 = \mu 2 = \mu 3 = \mu 4$ from the sample given, that is the Dose data, then generate R=10000 numbers of random multivariate, calculate the the percentage of the randomly generated multivariate t test statistic greater than the observed multivariate t test statistics in the 10000 trails.
That is P-value = (count (Tv > Tobs))/10000.

Case of more than two contrasts:
Test statistics: T = Max {T1, T2, T3…Tk-1}
T1, T2, T3 … Tk-1 follows jointly multivariate t distribution with p = k-1 and degree of freedom (df) = v. the test performs the same type of contrast as the test in (b). Since we have more test that is the most powerful test in its own contrast, the combined test will be much more sensitive at the alternative region than the test will less contrasts. Thus, the test statistics is similar as (b) but more sensitive than the test with only two contrasts.
P-value:
T1, T2, T3…Tk-1 follow a multivariate t distribution with p = k-1 and degree of freedom (df) = v
The p-value is calculated under the null hypothesis H0 : $\mu 1 = \mu 2 = \mu 3$. The alternative hypothesis is $\mu 1 < \mu 2 < \mu 3 < \mu 4$, $\mu 1 < \mu 4$.
The rejection region: RR ={Tv > Tobs}
Calculate the observed multivariate t test statistics (Tobs) under the null hypothesis $\mu 1 = \mu 2 = \mu 3 = \mu 4$ from the sample given, that is the Dose data, then generate R=10000 numbers of random multivariate t distribution, calculate the the percentage of the randomly generated multivariate t test statistic greater than the observed multivariate t test statistics in the 10000 trails.
That is P-value = (count (Tv > Tobs))/10000.

(d). Covariate included

Since there are three responses, we propose to reduce the multiplicity problem by combining the four visits into one variable of interest, thus, the primary response variable is the average visit.
normality check on the response and the covariate

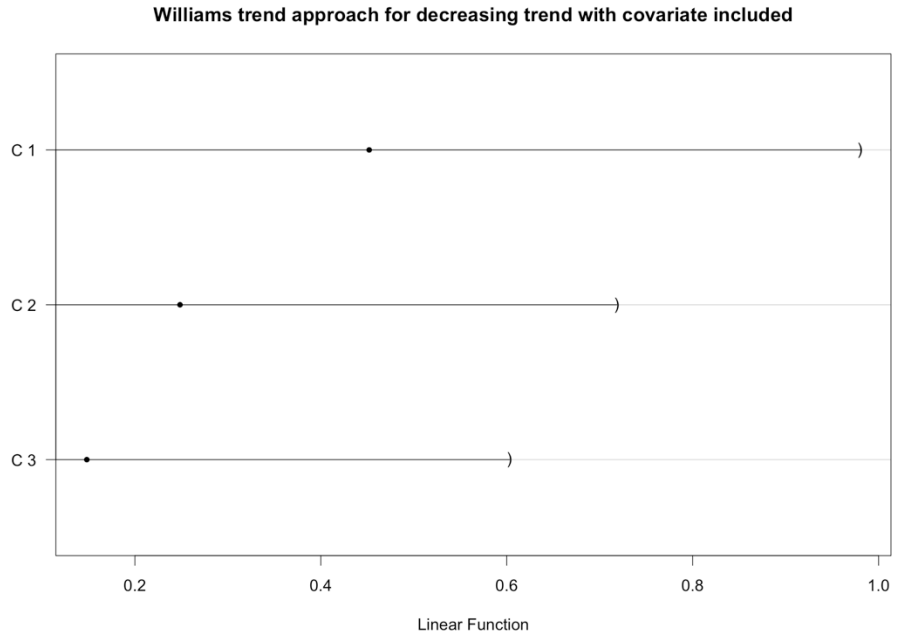


Test for homoscedasticity

	Test statistics	df	p-value
Bartlett test	3.4638	3	0.3255
levene.test	2.404		0.0675

The response variable meets the assumption of homoscedasticity.

rf.mean~Dose+Bvalue1+ Bvalue2				
Linear Hypotheses: original Williams test				
	Estimate	SE	t value	Pr(<t)
C1>=0	0.452	0.2789	1.621	0.975
C2>=0	0.2485	0.2481	1.002	0.907
C3>=0	0.1483	0.2402	0.617	0.823



Extended Williams test:
Dt Williams=

$$\begin{pmatrix} 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0.4901961 & 0.5098039 & 0 & 0 & 0 \\ 0 & -1 & 0.3177966 & 0.3305085 & 0.3516949 & 0 & 0 \end{pmatrix}$$

rf.mean~Dose+Bvalue1+ Bvalue2				
Linear Hypotheses: extended Williams test				
	Estimate	SE	t value	Pr(<t)
C1>=0	-0.06695	0.29934	-0.224	0.532
C2>=0	-0.01652	0.25585	-0.065	0.597
C3>=0	0.14827	0.23934	0.619	0.829

None of these two tests has a lowest P-value smaller than 0.05, Thus, there is no positive treatment effect when covariates are included.