To Oil or Not to Oil: An Investigation into Agrabathi and Old Wive's Tales

STA2005S

Dhiya Tathiah and Keeran Moodley
Revised on: September 11, 2024

Abstract

A quatitative analysis of the burn time of Agrabathi when covered in various common oils found in Indian households.

Table of contents

1 Introduction

There is an old wives tale which hails from the ancient Indian subcontinent and has been told for generations; applying oil to an incense stick will cause it to burn faster. This quick burn time ensured that the smoke created a sacred atmosphere, a well scented home and most importantly, carried your family's prayers to the almighty deities above. While there can be no doubt about the role of incense sticks (or Agrabathi) in the cultural and spiritual settings of an Indian home, doubt remained about the effectiveness of dipping these sticks into the rich, often fragranced, oils. This study seeks to unravel the truth behind this age-old belief, offering modern families the wisdom to discern whether investing in extra oils for Agrabathi truly enhances both their connection to the divine and the speed in which the fragrance emanates throughout their homes. By providing this knowledge, this study aims to empower families to make informed decisions and potentially save them from unnecessary expenses if these treatments are found not to significantly extend the burn time of incense sticks.

This study explores the influence of different oil treatments on the burn time of incense sticks through the application of a randomized block design (RBD).

2 Objectives

The objective of the study is to determine if the treatments have an effect on the burn time of Agrabathi. The study considers three levels to the treatment factor: a control of no oil, castor oil, and coconut oil. The study will examine whether these commonly used oils differ from each other (comparison of castor and coconut oil) and whether there is a difference between the oils and the control (comparing the average response of the oils to the average response of the control). By blocking for the different brands of incense sticks, we can more confidently deduce the differences between treatments, as the blocks contain homogeneous units. Once the incense sticks are lit and the smoke clears, the last burning stick will reveal whether oils truly influence the burn time.

Formally this study will test the following hypothesis:

> ℹ H_0: The application of different oils has no effect on the burn time of Agrabathi
> H_A: The application of at least one of the oils has an effect on the burn time of Agrabathi

Additionally the following two comparisons of means will be conducted:

> ℹ L_1: Effect of sandalwood oil is equal to the effect of coconut oil.
> L_2: The effect of no oil is equal to the average effect of applying the oils

3 Design and Procedure

This experiment will employ a randomised block design with a single treatment factor - application of oil - of three levels, viz., control (no oil), coconut oil, and castor oil. The experiment will block for heterogeneity of experimental units arising from the use of different brands of Agrabathi viz., Hem, Malarani and Tulasi.

The factor levels have been selected as they are oils commonly used in Indian households across the world and are the de facto choices during day to day use. The brands of Agrabathi from which the experimental units are drawn from represent easily found and widely exported brands.

By acknowledging the differences that might arise due to manufacturing or material quality, blocking for brands serves as a crucial step in isolating the true effects of the oil treatments. By structuring the experiment in this way, the study aims to minimize the impact of confounding variables, allowing for the attribution of any variations in burn time more confidently to the treatments themselves rather than to the intrinsic characteristics of the incense sticks.

A pilot study will be conducted to assess the viability of the experimental procedure which is outlined below:

1. Select experimental units from each brand of Agrabathi
2. Randomly assign treatments to the units within each block

3. Apply the relevant treatment in the form of coating the sticks of Agrabathi in the appropriate oil ensuring that there is even and consistent covering
4. Light the Agrabathi sticks at their tip and place them in a sheltered area to burn
5. Record the time taken of the Agrabathi to completely burn

Precise details about the randomisation procedure will be discussed in **Link to the relevant section**.

To reduce variance in the experiment due to external factors several steps will be taken to ensure that the experimental conditions will be kept consistent:

1. The Agrabathi will be burnt in the same area to prevent confounding due to location
2. The Agrabathi will be sheltered from wind and sunlight to prevent confounding due to increased airflow over the flaming tip and increased energy due to the sunlight
3. The blocks will be burnt at 10 minute intervals from each other to reduce confounding due to time. The interval is given to allow for the experimenters to set up and light the Agrabathi. This also allows for the majority of the Agrabathi in each group to burn concurrently to further reduce confounding due to time as well as increase the efficiency of the experiment.

The response variable is the time taken for the Agrabathi to burn given in seconds. The measurement of this was achieved via online stopwatch websites and the data was then manually transcribed.

4 Randomisation

Randomisation took place within each block of 3 experimental units (EUs). The procedure was as follows:

1. Label the EUs 1-3
2. Generate three random numbers between 1 and 1000 and iteratively assign them to the EUs (first generated number to EU 1, etc.)
3. Sort the random numbers in ascending order
4. Assign the treatments to the EUs using this ordered list, i.e., the EU corresponding to the lowest random number will be assigned the control treatment of no oil, the second number will get the coconut oil treatment and the largest number will get the castor oil treatment
5. Repeat 1-4 for all three blocks
6. Repeat 1-5 for every replication of the experiment

A sample randomisation for a singly replicated experiment is given below:

Table 1: Sample Randomisation

	1	2	3
Hem	A	C	B
Malarani	A	B	C
Tulasi	A	C	B

Where A,B, and C correspond to the treatment of no oil, coconut oil and castor oil respectively. The full randomisation used in given in the Appendix.

5 Pilot study

The pilot study was run with 18 experimental units and blocks were replicated twice.

Several difficulties were experienced while conducting the pilot study. Due to the large volume of smoke produced by the Agrabathi as it burnt, the experiment had to be conducted outdoors. This made it difficult to control for environmental factors such as wind, humidity, and sunlight. Additionally it was difficult to determine exactly when the Agrabathi stopped burning and thus there are slight non-systematic errors in the measurements of the burn times due to experimental error.

The original data is provided in the appendix. A basic descriptive analysis was conducted to analyse the data:

Table 2: Basic descriptive statistics

	Median	Mean	SD
Control	2243.44	1951.10	549.10
Coconut Oil	2780.95	2642.46	390.76
Castor Oil	2835.09	2712.11	321.78

The grand mean is 2435.23 and grand sample standard deviation is 537.57. From Table 1, one notes some differences in the means across the three treatments. The control group shows the lowest mean burn time but displays the highest standard deviation out of all the treatments. This may be due to the heterogeneity of experimental units. The oil treatments show smaller standard deviations which may be indicative of a treatment effect. Additionally all three treatments display a positive skew. These insights suggest a need for more data to test for significant effects.

6 Data collection and Assumptions

The full experiment was run with 30 replications per block. This took place using the same experimental and randomisation procedure as outlined above. The original data is given in the Appendix.

Normality tests were then conducted to justify the assumptions which will be made (discussed in the next section) in the model. **Add in all of Dhiya's analysis here and typeset**

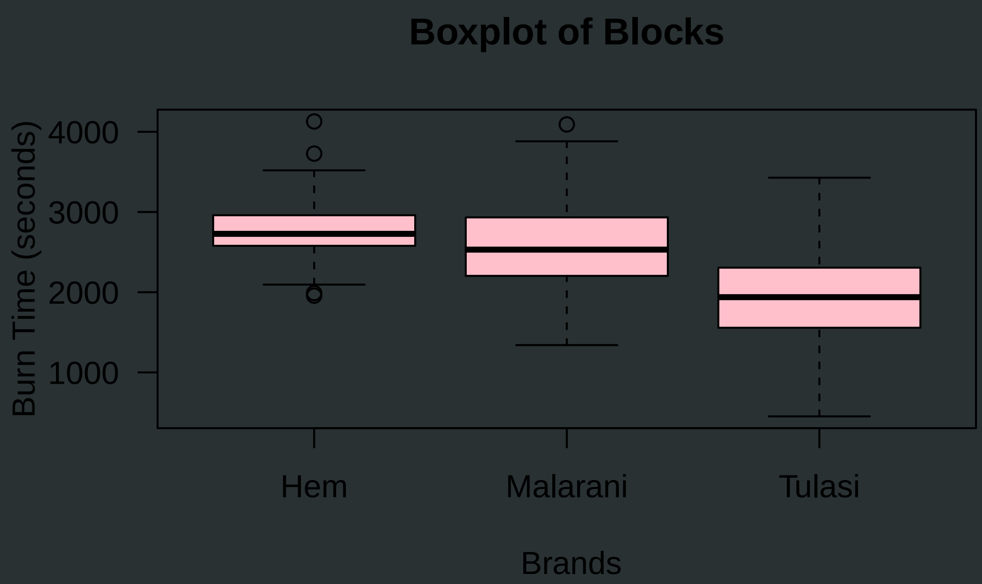


Figure 1: Box plot of brands versus burn time

In Figure 1, there is an equal proportion of the inner quartile range both above and below the mean and the tails of the boxplots are symmetrical, these properties are evidence that the data follows a normal distribution. In a randomized block design, it's important to check if the data within each block follows a normal distribution. This helps ensure that the residuals of the ANOVA model are likely to be normally distributed, which is a key assumption for the validity of the F-tests.

Non-normality may indicate potential issues like outliers or skewed data, which could violate ANOVA's assumptions and lead to inaccurate conclusions. It is visually clear from Figure 1 that the data is normal. The whiskers on the box plots are approximately equal across the blocks, yet slightly shorter for 'Hem' than 'Malarani' or 'Tulasi'. Due to an approximate average length of whiskers across the blocks, one can conclude equal variances. A Levene test is nevertheless performed to ensure accuracy since there is some disparity within the 'Hem' block of this exploratory analysis.

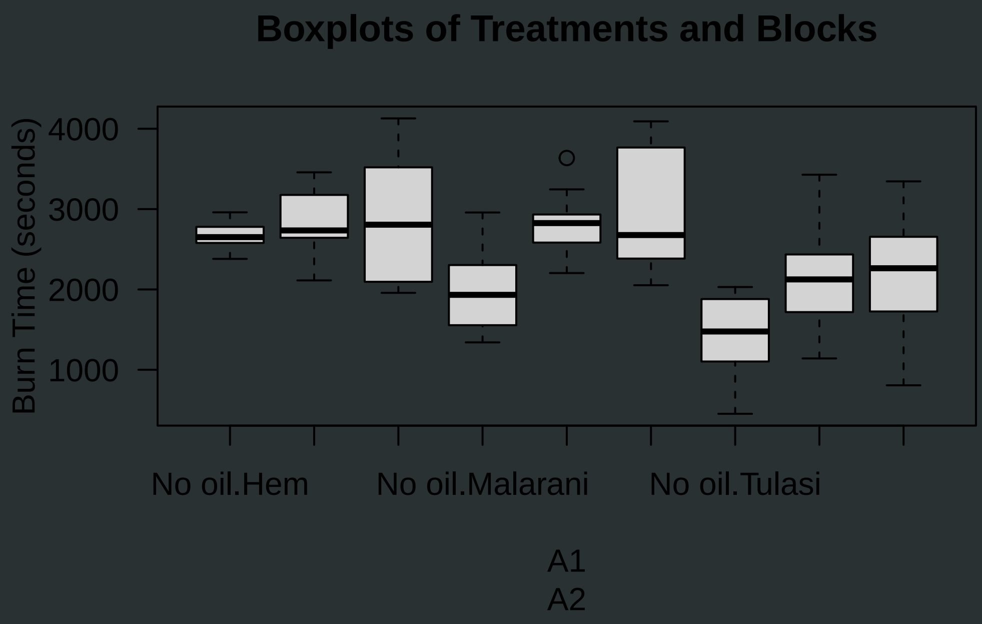


Figure 2: Box plot of brands and treatments versus burn time

Figure 2 delves into the treatments within the blocks. Within block 1:

1. Treatment 1: The burn times are clustered, a small interquartile range and short, symmetric tails.
2. Treatment 2: The median is approximately the same as treatment 1 (control) yet a larger variability is evident. The lower tail seems to be slightly longer than the upper tail indicating a possible left skew (even though very small difference in tail length).
3. Treatment 3: High variability, clear positive skew due to a longer upper tail. The median is marginally above the median for Treatment 1 & 2. The large spread suggests increased burn time under this treatment.

With Block 2:

1. Treatment 1: Significantly lower median compared to other treatments in block 2. There is a positive skew due to a longer upper tail. There is a much lower median present compared to the same treatment in block 1.
2. Treatment 2: There is an outlier present in this data, suggesting that we have an observation which is quite far from the rest of the data. The median is the highest in this block even though the data is grouped quite closely (smallest IQR of block 2). The entire distribution is shifted upward relative to treatment 1 in this block yet treatment 2's lower quartile overlaps with the upper quartile of treatment 1.
3. Treatment 3: High Variability and is potentially right skewed.

Within Block 3:

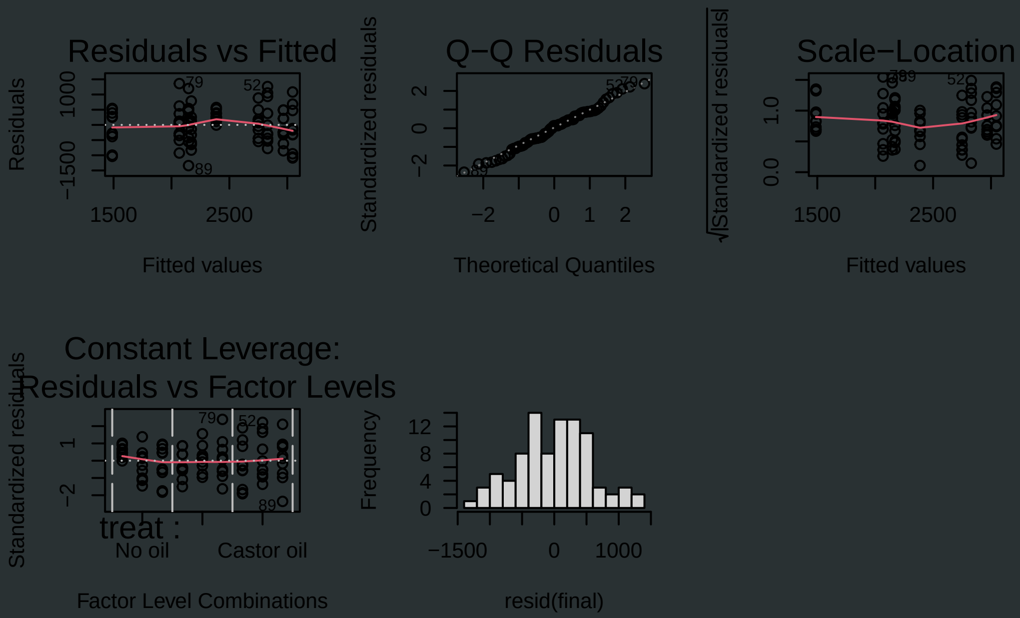
1. Treatment 1: Lowest median in this block and overall blocks, there is a clear left-skew due to asymmetric tail length.
2. Treatment 2: Symmetric tails and median which appears to be in the center of the IQR, larger variability compared to Treatment 2 in the other 2 blocks (longer tails).
3. Treatment 3: Highest median in this block with a symmetrical box plot (central median and approximately equal tail length).

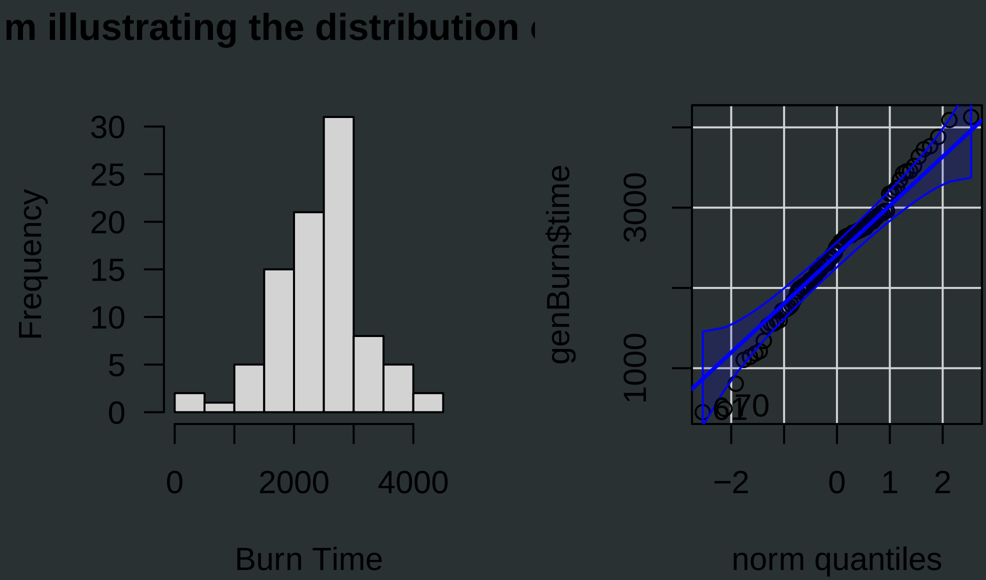
The overall impression of this more in-depth analysis of the box plots does seem to indicate some inconsistencies which do not represent an accurate normal distribution with homoscedasticity. This is due to asymmetric tail lengths, non-central medians and tail lengths of treatments which are not consistent throughout blocks. While there are these indications of a stray from a normal distribution, an inspection of the Shapiro Wilk normality test will test if this is a result of significant non-normality or simple randomness within the data.

The null hypothesis for this test would be that the data is normally distributed while the alternate hypothesis would be that the data follows a different distribution. Since our p-value is quite large $p = 0.5643 > 0.5$, therefore there is insufficient evidence to reject the null hypothesis, thus the distribution of the reponse variable is normal.

```
    Shapiro-Wilk normality test

data:  resid(final)
W = 0.99055, p-value = 0.7704
```






```
[1] 61 70
```

The variability among blocks and treatments will also be further explored by means of a Levene test. This is to establish homoscedasticity of the treatment groups. The null hypothesis of this test is that the population variances between the groups is equal. The alternate hypothesis is that there is a difference in the variance of the population underlykng each treatment group. The Levene test fails to reject the null hypothesis ($p = 0.44$). This is to say that the variance between groups is equal and the assumption of homoscedasticity is met.

Above (histogram and QQplot): (Yet our Levene Test is secure so there is a homoscedasticity) The Q-Q plot suggests that the central portion of the burn time data follows a normal distribution reasonably well, as the points around the middle of the plot lie close to the straight line. However, there are notable deviations at both ends, with the lower tail showing lighter-than-expected behavior (fewer extreme small values) and the upper tail indicating heavier-than-expected values (more extreme large values). These deviations from the line, particularly at the tails, suggest that the data does not fully adhere to the assumption of normality. The plot shows a blue confidence band. Points falling within this band suggest acceptable deviations from normality. However, several points in the tails fall outside the band, strengthening the evidence of non-normality in the extremes. The histogram of burn time data appears to have a slight right skew. Most of the burn time values are concentrated between 1,500 and 3,500, with a peak frequency around 2,500. There are a few observations with higher burn times, reaching up to 4,000, which contribute to the right tail. This skewness aligns with what is seen in the Q-Q plot, where the points in the upper quantiles deviate above the line, indicating heavier tails on the right side. While the overall distribution of burn times does not seem to be drastically skewed, the deviations in the tail could suggest a slight departure from normality, as indicated by both the histogram and the Q-Q plot.

Fitted Vs residual and Scale Location plot: There is no funnel shape (no widening or narrowing or the spread of residuals in data) present in the Fitted vs residual plot, therefore no indication of heteroscedasticity. There is no discernible pattern regarding the residuals around the horizontal 0 line. Scale Location has a slightly decreasing red line, it suggests that the spread of residuals (i.e., the variability of the residuals) decreases as the fitted values increase. This indicates that the model might exhibit heteroscedasticity, where the variance of the residuals is not constant across the range of the fitted values. In this case, the residuals show less variability at higher fitted values compared to lower fitted values. Levene test: The P value =0.4419 indicating that there is homogeneity present in the variances.

The histogram of the residuals show a bell-shaped curve which is a well-known characteristic of a normal distribution. There is asymmetry present, indicating there are more positive residuals than negative ones (prediction is lower than our observed value) which signals that our assumption of normality might be violated. Yet the Shapiro Test confirms normality.

Interaction plot: (We can use blocking on x axis or as the lines yet give different plots) *Still uncertain whether should be included or not **Question for tutors. Since there is no overlap or intersection between the lines and they are parallel -> no interaction. OR: interaction.

7 Model

This study will employ the following model for the data:

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}$$

Where $\varepsilon \sim N(0, \sigma^2)$ is the error term, $1 \leq i \leq 3$ indexes the treatments, $1 \leq j \leq 3$ indexes the blocks. Additionally we employ the sum to zero constraint such that $\sum_{i=1}^{3} \alpha_i = \sum_{j=1}^{3} \beta_j = \sum_{ij} (\alpha\beta)_{ij} = 0$. Additionally we assume an interactive model and thus include the possibility of interaction between block and treatment effects.

Each of these terms are interpreted as follows:

1. μ is the overall mean
2. α_i is the main effect of the i-th level of the treatment factor (oil application)
3. β_j is the main effect of of the j-th level of the blocking factor (brand)
4. $(\alpha\beta)_{ij}$ is the interaction between the i-th level of the treatment and the j-th level of the blocking factor.

This model will make the following assumptions:

1. Homoscedasticity
2. Normally distributed error terms
3. Independent observations

These assumptions will be justified and assessed during the model checking stage of the paper.

8 Outline of Analysis

The analysis of results will aim to provide conclusive evidence to support or reject the a priori hypothesis of this study as well as evaluate the root causes of these results my means of analysing the contrasts of interest.

ANOVA Based on the above model an ANOVA was performed to determine if the treatments have a significant effect on the burn time of Agrabathi.

Table 3: Analysis of Variance

	DF	SS	MSS	F	P(>F)
Block	2	13198564	6599282.1	19.272	1.00e-07
Treatment	2	7760310	3880155.2	11.331	4.34e-05
Error	85	29106829	342433.3	NA	NA

Blocks: The F-value for blocks is 19.27 with a very low p-value ($p = 1.00 \times 10^{-7}$), indicating that the variability between blocks is highly significant and thus it was reasonable to block for the brands of Agrabathi used. Treatments: The F-value for treatments is 11.33 with a low p-value ($p = 4.34 \times 10^{-5}$), suggesting that the differences between treatments are also significant. This is to say that the application of one of the treatments has an effect on the burn time of Agrabathi. At this stage it is unclear which treatment (or even the control) is responsible for this conclusion. The specific cause of this effect is analysed more deeply by the contrasts considered in the following section. Significance of Block Effect: The significant p-value for blocks suggests that differences between blocks are substantial, meaning that the block effect is crucial in explaining the variability in burn time. Significance of Treatment Effect: The significant p-value for treatments indicates that the treatments have a meaningful impact on burn time, suggesting that different treatments result in different burn times. Residuals: The residuals provide insight into the unexplained variation after accounting for block and treatment effects. Although this variation is present, it's relatively small compared to the explained variance from blocks and treatments. Since the p value<0.05 for both treatment and block effect, we can reject the null hypothesis and conclude that different oils (treatments) and brands (blocking factors) affect the burn time of incense sticks.

9 Contrasts

This study examines 2 planned contrasts, viz., if there is a difference in the burn time of Agrabathi when no oil is applied versus when oil is applied and if there is difference in the burn time of Agrabathi when coconut oil is applied as opposed to castor oil.

To account for these comparisons this study sets a maximum allowable experiment-wise type I error rate of 5. The comparisons are then corrected via the Bonferroni method to ensure this limit is upheld.

Formally we consider the following contrasts:

$$L_1 = \mu_{Coconut} - \mu_{Castor}$$
$$L_2 = \mu_{No\ Oil} - \frac{1}{2}(\mu_{Coconut} + \mu_{Castor})$$

Note that L_1 and L_2 are orthogonal contrasts and thus partition the sum square treatment. The table below summarises the analysis of the contrasts.

Table 4: Analysis of Contrasts

	DF	SS	MSS	F	P(>F)
Treatment	2	7760310.4	3880155.2	11.331	0.0000
—L1	1	94562.5	94562.5	0.276	0.6006
—L2	1	7665747.9	7665747.9	22.386	0.0000
Error	2	13198564.2	6599282.1	19.272	0.0000
Total	85	29106828.9	342433.3	NA	NA

This reveals an interesting nuance to the data. In the previous section we concluded that there is indeed an effect induced by the treatments. Contrast L_2 compares the effect of no oil to the effect of applying oil and this shows that there is a statistically significant difference between them with $p = 8.79 \times 10^{-6}$. Conversely contrast L_1 shows little effect ($p = 0.60$) which is to say that there is little difference between the types of oil applied.

Bonferroni corrected confidence intervals for these contrasts are now constructed to ensure that the conclusions drawn above are valid and not simply type I errors. This study will permit a tolerance of $\alpha = 5\%$ for type I errors. This is to say that the conclusions are drawn with 95% confidence. The choice of Bonferroni's correction was made as only a priori contrasts are considered and there is small number of them. Had the study performed a post hoc analysis and made all pairwise-comparisons more sophisticated methods such as Tukey's or Sheffe's would have been selected.

Table 5: Bonferroni corrected confidence intervals

	Lower Bound	Upper Bound	Point Estimate
Contrast L1	-189.905	110.507	-39.699
Contrast L2	-586.178	-239.292	-412.735

Based on these intervals the conclusions drawn by the initial analysis of the p-values is correct as the CI for L1 contains zero thus there is no significant difference between the oils. Similarly the CI for L2 does not contain zero and thus there is a difference between the application of oil versus no oil. **talk about correctio n**

10 Conclusion

both thursday 12pm M202

11 Appendix

The original randomisation used is given below:

Table 6: Randomisation within blocks

Block	1	2	3
Malarani	B	C	A
Tulasi	C	B	A
Tulasi	C	B	A
Tulasi	C	B	A
Malarani	A	C	B
Tulasi	C	B	A
Tulasi	B	A	C
Malarani	C	B	A
Tulasi	A	B	C
Malarani	B	A	C
Malarani	C	B	A
Tulasi	A	B	C
Malarani	C	A	B
Hem	A	C	B
Malarani	C	A	B
Hem	C	A	B
Malarani	A	C	B
Tulasi	C	A	B
Tulasi	A	B	C
Malarani	A	C	B
Hem	A	B	C
Tulasi	C	B	A
Tulasi	B	C	A
Tulasi	A	B	C
Hem	A	C	B
Tulasi	A	C	B
Malarani	A	B	C
Tulasi	A	B	C
Malarani	B	A	C
Hem	B	C	A
Malarani	A	C	B
Tulasi	A	B	C
Hem	C	A	B
Tulasi	C	B	A
Tulasi	A	C	B
Tulasi	A	B	C
Hem	C	B	A
Hem	B	A	C
Malarani	A	B	C
Tulasi	B	A	C
Malarani	A	B	C
Malarani	C	B	A
Hem	B	A	C
Malarani	B	A	C
Hem	C	A	B
Hem	C	A	B

Malarani	B	A	C
Malarani	A	B	C
Tulasi	C	B	A
Malarani	B	C	A
Hem	C	B	A
Hem	B	C	A
Hem	C	B	A
Tulasi	A	B	C
Tulasi	C	A	B
Tulasi	A	C	B
Hem	C	A	B
Hem	A	C	B
Malarani	A	C	B
Tulasi	A	B	C
Tulasi	C	A	B
Hem	A	B	C
Malarani	A	B	C
Hem	A	B	C
Malarani	B	C	A
Hem	B	A	C
Hem	C	B	A
Malarani	B	C	A
Malarani	A	B	C
Hem	B	A	C
Hem	C	A	B
Hem	B	A	C
Hem	C	A	B
Malarani	A	B	C
Malarani	A	B	C
Malarani	C	B	A
Tulasi	C	A	B
Malarani	A	C	B
Hem	A	B	C
Malarani	C	A	B
Hem	A	C	B
Hem	B	A	C
Malarani	C	A	B
Hem	B	C	A
Hem	A	B	C
Tulasi	A	B	C
Tulasi	B	A	C
Hem	B	C	A
Tulasi	A	B	C
Tulasi	B	A	C

The full data (after sorting) for the experiment is given below:

Table 7: Randomisation within blocks

block	treat	time
Hem	No oil	2650.5227
Hem	No oil	2778.7920
Hem	No oil	2959.5394

Hem	No oil	2633.4956
Hem	No oil	2904.9172
Hem	No oil	2578.9041
Hem	No oil	2651.6279
Hem	No oil	2502.9739
Hem	No oil	2380.3736
Hem	No oil	2751.1898
Hem	Coconut oil	3451.0439
Hem	Coconut oil	2642.6727
Hem	Coconut oil	2347.5647
Hem	Coconut oil	3457.3515
Hem	Coconut oil	2694.2535
Hem	Coconut oil	2721.5697
Hem	Coconut oil	3173.9402
Hem	Coconut oil	3176.1767
Hem	Coconut oil	2112.3550
Hem	Coconut oil	2747.0112
Hem	Castor oil	2927.3282
Hem	Castor oil	3727.8999
Hem	Castor oil	2733.7080
Hem	Castor oil	2719.9791
Hem	Castor oil	4128.9509
Hem	Castor oil	2877.9955
Hem	Castor oil	1957.2929
Hem	Castor oil	1989.6532
Hem	Castor oil	3519.4677
Hem	Castor oil	2095.1505
Malarani	No oil	2249.2034
Malarani	No oil	2957.8254
Malarani	No oil	1595.5402
Malarani	No oil	2427.3412
Malarani	No oil	1524.0672
Malarani	No oil	2303.4432
Malarani	No oil	1554.8915
Malarani	No oil	2024.1093
Malarani	No oil	1341.1688
Malarani	No oil	1841.5132
Malarani	Coconut oil	2826.8532
Malarani	Coconut oil	3245.2903
Malarani	Coconut oil	2824.5364
Malarani	Coconut oil	2583.3981
Malarani	Coconut oil	2203.6966
Malarani	Coconut oil	2305.7980
Malarani	Coconut oil	2860.9482
Malarani	Coconut oil	2932.7839
Malarani	Coconut oil	3635.8458
Malarani	Coconut oil	2706.5840
Malarani	Castor oil	2526.1670
Malarani	Castor oil	4091.8076
Malarani	Castor oil	2817.5176
Malarani	Castor oil	2297.8197
Malarani	Castor oil	3208.6125

Malarani	Castor oil	2052.4874
Malarani	Castor oil	3881.3805
Malarani	Castor oil	2384.1233
Malarani	Castor oil	2536.1593
Malarani	Castor oil	3766.5081
Tulasi	No oil	452.1814
Tulasi	No oil	1103.2669
Tulasi	No oil	1211.8502
Tulasi	No oil	1880.1949
Tulasi	No oil	2030.0183
Tulasi	No oil	1996.7522
Tulasi	No oil	1760.2772
Tulasi	No oil	1741.0815
Tulasi	No oil	1182.2138
Tulasi	No oil	494.3818
Tulasi	Coconut oil	1141.3884
Tulasi	Coconut oil	2689.8000
Tulasi	Coconut oil	1787.2440
Tulasi	Coconut oil	1718.2321
Tulasi	Coconut oil	2188.9874
Tulasi	Coconut oil	1557.4480
Tulasi	Coconut oil	2435.4088
Tulasi	Coconut oil	2143.0334
Tulasi	Coconut oil	3427.8434
Tulasi	Coconut oil	2106.6795
Tulasi	Castor oil	1592.9635
Tulasi	Castor oil	2306.2427
Tulasi	Castor oil	1725.6686
Tulasi	Castor oil	2598.4327
Tulasi	Castor oil	2221.3615
Tulasi	Castor oil	2686.8291
Tulasi	Castor oil	3345.1062
Tulasi	Castor oil	2049.2998
Tulasi	Castor oil	806.6389
Tulasi	Castor oil	2655.1505

Normality tests: Fitted vs residuals: Should test for linearity, unequal error variances and outliers. The data should be randomly located around the horizontal line, there should be no extreme outliers, and the residual spread should be roughly constant against the fitted values to suggest homoscedasticity. Our graph does not show any extreme funnel shape but there is a widening as we move right through the plot indicating a potential loss of homoscedasticity.
QQ plot: The most important test for normality, the data is basically supposed to stay on the 45 degree line but ours deviates into almost an S shape which indicates skewness. Since this is the most important test of normality and the Shapiro test is very accurate for small data sizes such as ours, this does not inspire confidence in our normal assumption which appears to be violated. Scale Location: This plot checks for homoscedasticity. The red line should be horizontal if the variance is constant. Our red line decreases, showing that variance is likely not constant. Constant Leverage: Similar to our fitted vs residuals, should be scattered around 0 with no discernible pattern. Our final histogram: No clear bell shape, there seems to be a positive skew. There is no symmetry in this graph.