# Luxuriant's efficacy analysis for hair growth

MT5763 Assignment 2

SID: 210034046


## Abstract

To analyse the effects of *Luxuriant* for treating male pattern baldness, we have conducted a trial with 4 groups, with patients suffering from baldness. Three of the groups were given a specific brand's treatment: *Luxuriant*, *BaldBeGone* and *SkinheadnNoMore*, and the 4th group was the placebo group. The results of this trial show that, although the Luxuriant group showed significantly more growth than the placebo group, the rivals *BaldBeGone* and *SkinheadNoMore* heavily outperform our product.


## Introduction

A clinical trial was conducted, for the purpose of testing *Luxuriant's* efficacy for treating male pattern baldness. The trial consisted of patients suffering from male pattern baldness, divided into 4 groups, each treated with one of the three brands, *Luxuriant*, *BaldBeGone* and *SkinheadNoMore* or the placebo group with no effective treatment. All the participants were shaved, then measured for their hair growth after a month.

The study's research questions were:

1. *Is Luxuriant causing growth?*
2. *How does Luxuriant's effectiveness compares with other brands?*
3. *Does age effect hair growth?*

## Analysis Description

Firstly, we analysed if the treatment actually caused growth. To this end, a comparison of Luxuriant's and placebo group's hair growth was carried out. We had to check the null hypothesis that the treatment and non-treatment growth were exactly the same. The sample was not independent, being collected from the same group of people before and after the treatment. This left us with two alternatives, either the paired t-test or the non parametric tests. The QQ-plot (Figure 1) showed deviation from the straight line for the placebo group's data. Further analysis of the data using Shapiro-Wilk (Figure 2), shows that the data is not normally distributed, therefore violating one of the assumptions for a t-test. To this end, we finally conducted a non-parametric Wilcoxon test (Figure 3).

Secondly, we conducted a comparative analysis of our brand with BaldBeGone and SkinheadNoMore. For this purpose, an Anova analysis (Figure 4) was carried out, testing the null hypothesis that all the 3 brands' growth means are the same. In the case that the null hypothesis is rejected, further analysis is required. To that end, a Tukey's HSD test (Figure 5) was used to see the pairwise difference between the 3 brands.

Lastly, we checked the effect of age on hair growth, using linear models (Figure 6), fitted by the data collected from the trials. We used the brands as a factor variable, the hair growth as response and age as an explanatory variable.

## Results

The *Luxuriant vs placebo* comparison, using the Wilcoxon test (Figure 3) showed that there is a significant difference between the the two means (of hair growth) of the two groups. Furthermore, it is evident from the mean scores (luxuriant=144.7>placebo=56.3), luxuriant caused significant growth among patients.

The *Luxuriant vs other brands* comparison, using the Analysis Of Variance (ANOVA, Figure 4), which rejects the null hypothesis, this means there was at least one group that had a different growth mean (higher or lower). The test showed that at least one brand's hair growth was significantly different. To check the difference between Luxuriant and other brands' hair growth efficacy, we did a pairwise test of their mean growth rates, using Tukey's HSD (Figure 5). Test showed our product to be heavily under-performing in comparison to its rivals, BaldBeGone was 17.4 mm, and SkinheadNoMore had 5.741 mm higher mean growth.

The *Effect of age on hair growth*, assessed using the linear model (Figure 6), showed no significant relationship. This is evident from the graph below in Figure 7, which is almost a flat line. The parameter value was -0.067 and the parameter's p-value 0.59, proves no significant effect of age on hair growth.

## Conclusions

The trial conducted showed that *Luxuriant* is an effective treatment, causing hair growth among its users. However, it is not as effective as its competitors, *BaldBeGone* and *SkinheadNoMore*. BaldBeGone is the most effective for hair growth in the market. The goal of the company with this product should be to be increase the efficacy of our product beyond the level of BaldBeGone. This requires a lot of improvement as it caused 17.4 mm more hair growth. The 2nd best was *SkinheadNoMore*, having 5.741 mm more hair growth per month than our product. Our initial target should be to at least make our product more effective than SkinheadNomore, before pursuing the final goal, becoming the most effective product providers.

## Appendix

The code for the analysis.

```
# Importing baldy.csv data.
PROC import out= baldy
dbms=csv replace
file="/home/u59691081/sasuser.v94/SasFolder/New Folder3/Baldy.csv";
getnames=yes;
datarow=2;
run;
DATA baldy_mm;
SET baldy;
Luxuriant=luxuriant*25.4;
placebo=placebo*25.4;
baldbegone=baldbegone*25.4;
skinheadnomore=skinheadnomore*25.4;
RUN;
# Checking contents.
proc contents data=baldy_mm;
run;
# Checking numeric variables.
proc univariate data=najmi050.baldy_mm;
run;

# Task 1: Comparing Luxuriant with placebo.
```

```
#Extracting subset of required data.
data LuxuriantVsPlacebo;
set baldy_mm;
keep luxuriant placebo;
run;

#Double transposing, to get factor variable.

#First transpose converts data into two rows of brands, each having their own 100 columns of hair growt
proc transpose data=luxuriantvsplacebo out=temp;
by rows:;
var luxuriant placebo;
run;
#Second  transpose gets data into two columns, brand names as factor variable and 2nd column hair growt
#Therefore, each row is a subject that has used a particular brand and his hair growth.
proc transpose data=temp out=want;
   by _name_;
   var col:;
run;

#Saving data and renaming the columns to brand and growth_mm.

data LuxuriantVsPlacebo;
set work.want;
rename _name_=LUXorPlacebo col1=growth_mm;
run;

#Conducting a two sample ttest to check difference between the two groups' growth.

proc ttest alpha=0.05 data=luxuriantvsplacebo sided=2;
class LUXorPlacebo;
var growth_mm;
run;

#The F-test for equality variance (p<0.0001) shows that there is enough evidence to reject the null hyp
#the variance of the two groups is equal.
#Therefore, options are to either use the un-pooled Satterthwaite method,
#provided the normality assumption holds or go for a non-parametric test.
#Although, the QQplot for placebo shows deviation from the normal straight line,
#to confirm our suspicion we conduct a normality test.
#Additionally, a regular t-test would not be valid to begin with, as the sample given is a dependent sa
proc univariate data=luxuriantvsplacebo normaltest;
run;

#According to Shapiro-wilk test, the assumption that data is normal, doesn't
#satisfy, p-value<0.0001.
#To this end, we use non-parametric tests to test growths are equal.
proc npar1way wilcoxon data=luxuriantvsplacebo;
class luxorplacebo;
var growth_mm;
exact;
run;
```

```
#The p-value <0.0001, lower than the 0.05 significance level.
#Therefore, we conclude that the placebo and luxuriant have significantly
#different growth rates. It is evident from the mean scores (luxuriant=144.7>placebo=56.3), luxuriant
#is causing more growth, in comparison to placebo.

#Task 2: Luxuriant vs all other brands.

#Extracting subset of data for analysis.

Data LUXVSALLBRANDS;
set baldy_mm;
keep luxuriant baldbegone skinheadnomore;
run;
#Double transposing, to get it into a factor variable form.

#1st transpose converts the data into 3 rows, each corresponding to brands,
#with each having 100 columns.
proc transpose data=luxvsallbrands out=temp;
by rows:;
var luxuriant baldbegone skinheadnomore;
run;
#Sorting data according to brand.
proc sort data= temp;
by _name_;
run;
#Tranposing data again to get 400 rows and 2 columns.
#1st column is a factor variable with brand names and the 2nd column
#contains the values of growth in mm of a subject that used that brand.
proc transpose data=work.temp out=need;
by _name_;
var col:;
run;
#Transferring data to luxvsallbrands in the permanent library.
data luxvsallbrands;
set need;
run;
#Renaming columns with brand and growth_mm.
data luxvsallbrands;
set luxvsallbrands;
rename _name_=Brand col1=Growth_mm;
run;

#Conducting an anova analysis to test whether there is difference in mean
#in any one of the 3 brand's growth, followed by a Tukey's HSD to analyse pairwise differences.
proc anova data=luxvsallbrands;
class brand;
model growth_mm=brand;
means brand/tukey cldiff;
run;
#P-value<0.0001 for anova shows that our null hypothesis should be rejected and there is a significant d
#between the growth rate of atleast one brand compared to the others.
#From the Tukey's HSD it is clear that that there are significant differences between all the brands. We
```

```
#Task3: analysing the effect of age on hair growth
#Dividing into two separate data sets, one for growth data, one for age data.

#Extracting growth data set.
data Growths_set;
set baldy_mm;
keep luxuriant placebo baldbegone skinheadnomore;
run;

#Double transposing growth set, to get into factor variable form for analysis.
#First transpose converts 4 rows of brands, with 100 columns of growth data.
proc transpose data=growths_set out=temp;
by rows:;
var luxuriant placebo baldbegone skinheadnomore;
run;
#Sorting the data by brand name.
proc sort data=temp;
by _name_;
run;
#2nd transpose to get into factor variable form, 1st column is brand names and 2nd is growth in mm.
proc transpose data=temp out=need;
by _name_;
var col:;
run;
#Saving data and renaming columns to brand and growth_mm.
data Growths_set;
set need;
rename _name_=brand col1=growth_mm;
run;

#Extracting 2nd set, with ages and agebrand.
data age_set;
set baldy_mm;
drop luxuriant placebo baldbegone skinheadnomore;
run;

#Renaming the columns to brand names.
data age_set;
set age_set;
rename ageluxuriant=luxuriant ageplacebo=placebo agebaldbegone=baldbegone ageskinheadnomore=skinheadnom
run;

#Double transposing again to get factor variable form.

#1st transpose to get 4 rows with 100 columns of ages.
proc transpose data=age_set out=temp;
by rows:;
var luxuriant placebo baldbegone skinheadnomore;
run;

#Sorting the data set by brand-name.
proc sort data=temp;
by _name_;
```

```
run;

#2nd transpose to get factor variable form.
proc transpose data=temp out=need;
by _name_;
var col:;
run;

#Saving data and renaming columns to brand and ages.
data ages_set;
set need;
rename _name_=brand col1=ages;
run;

#Merging the two data sets side by side.
data Merged_Growth_Ages;
merge growths_set ages_set;
run;

#Fitting a linear model similar to R's: growth~ages.
proc glm data=merged_growth_ages;
class brand;
model growth_mm=ages/ solution clparm;
run;
#From this analysis we conclude that, age doesn't effect the growth. As it
#is evident from a very small parameter estimate (-0.067). Furthermore, the p-value for the t-test 0.59
```

# Figures of analysis and results

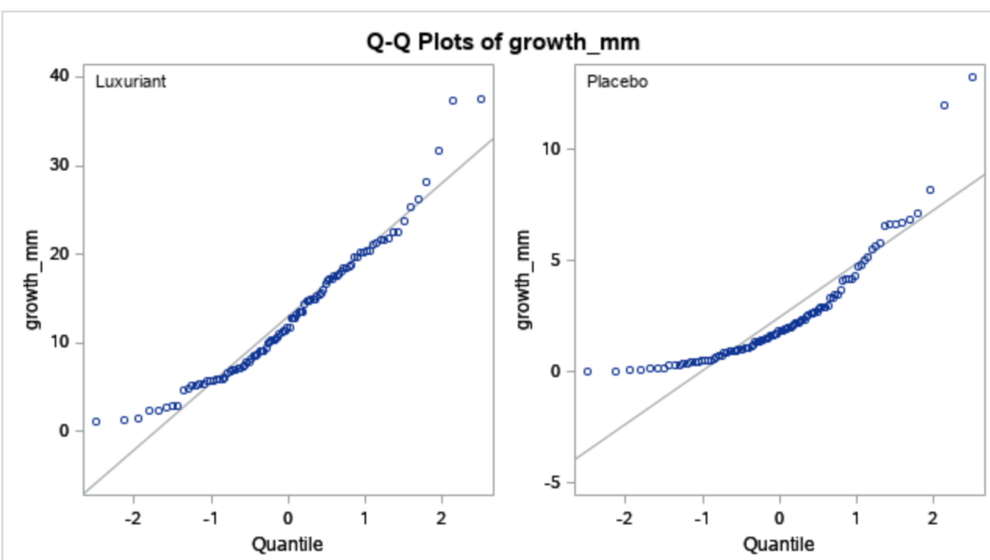


Figure 1: QQ plot: shows placebo deviates from straight line.

| Tests for Normality | | | | |
|---|---|---|---|---|
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.853309 | **Pr < W** | <0.0001 |
| **Kolmogorov-Smirnov** | **D** | 0.157506 | **Pr > D** | <0.0100 |
| **Cramer-von Mises** | **W-Sq** | 1.596124 | **Pr > W-Sq** | <0.0050 |
| **Anderson-Darling** | **A-Sq** | 9.156043 | **Pr > A-Sq** | <0.0050 |

Figure 2: Shapiro-Wilk Test: shows not normally distributed.

**The NPAR1WAY Procedure**

| Wilcoxon Scores (Rank Sums) for Variable growth_mm Classified by Variable LUXorPlacebo | | | | | |
|---|---|---|---|---|---|
| **LUXorPlacebo** | **N** | **Sum of Scores** | **Expected Under H0** | **Std Dev Under H0** | **Mean Score** |
| **Luxuriant** | 100 | 14470.0 | 10050.0 | 409.267639 | 144.70 |
| **Placebo** | 100 | 5630.0 | 10050.0 | 409.267639 | 56.30 |

<table>
<tr><th colspan="8">Wilcoxon Two-Sample Test</th></tr>
<tr><th rowspan="2">Statistic (S)</th><th rowspan="2">Z</th><th rowspan="2">Pr > Z</th><th rowspan="2">Pr > |Z|</th><th colspan="2">t Approximation</th><th colspan="2">Exact</th></tr>
<tr><th>Pr > Z</th><th>Pr > |Z|</th><th>Pr >= S</th><th>Pr >= |S-Mean|</th></tr>
<tr><td>14470.00</td><td>10.7986</td><td><.0001</td><td><.0001</td><td><.0001</td><td><.0001</td><td><.0001</td><td><.0001</td></tr>
<tr><td colspan="8">Z includes a continuity correction of 0.5.</td></tr>
</table>

Figure 3: Wilcox Test results: showing significant difference between Luxuriant and placebo.

**The ANOVA Procedure**

**Dependent Variable: Growth_mm**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 2 | 29072.59683 | 14536.29841 | 76.78 | <.0001 |
| Error | 297 | 56227.33491 | 189.31763 | | |
| Corrected Total | 299 | 85299.93174 | | | |

| R-Square | Coeff Var | Root MSE | Growth_mm Mean |
|---|---|---|---|
| 0.340828 | 60.79950 | 13.75927 | 22.63057 |

| Source | DF | Anova SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Brand | 2 | 29072.59683 | 14536.29841 | 76.78 | <.0001 |

Figure 4: ANOVA: Checking significant difference in atleast one group's growth rate.

**The ANOVA Procedure**

**Tukey's Studentized Range (HSD) Test for Growth_mm**

**Note:** This test controls the Type I experimentwise error rate.

| | |
|---|---|
| **Alpha** | 0.05 |
| **Error Degrees of Freedom** | 297 |
| **Error Mean Square** | 189.3176 |
| **Critical Value of Studentized Range** | 3.33122 |
| **Minimum Significant Difference** | 4.5835 |

| Comparisons significant at the 0.05 level are indicated by ***. | | | | |
|---|---|---|---|---|
| **Brand Comparison** | **Difference Between Means** | **Simultaneous 95% Confidence Limits** | | |
| **BaldBeGone - Skinheadnomore** | 17.412 | 12.829 | 21.996 | *** |
| **BaldBeGone - Luxuriant** | 23.153 | 18.569 | 27.736 | *** |
| **Skinheadnomore - BaldBeGone** | -17.412 | -21.996 | -12.829 | *** |
| **Skinheadnomore - Luxuriant** | 5.741 | 1.157 | 10.324 | *** |
| **Luxuriant - BaldBeGone** | -23.153 | -27.736 | -18.569 | *** |
| **Luxuriant - Skinheadnomore** | -5.741 | -10.324 | -1.157 | *** |

Figure 5: Tukey's HSD: showing pairwise differences.

**The GLM Procedure**

**Dependent Variable: growth_mm**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 84.0643 | 84.0643 | 0.29 | 0.5919 |
| Error | 398 | 116227.4252 | 292.0287 | | |
| Corrected Total | 399 | 116311.4895 | | | |

| R-Square | Coeff Var | Root MSE | growth_mm Mean |
|---|---|---|---|
| 0.000723 | 97.12774 | 17.08885 | 17.59420 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| ages | 1 | 84.06431320 | 84.06431320 | 0.29 | 0.5919 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| ages | 1 | 84.06431320 | 84.06431320 | 0.29 | 0.5919 |

| Parameter | Estimate | Standard Error | t Value | Pr > \|t\| | 95% Confidence Limits | |
|---|---|---|---|---|---|---|
| Intercept | 20.94823772 | 6.30949127 | 3.32 | 0.0010 | 8.54414179 | 33.35233365 |
| ages | -0.06712778 | 0.12511495 | -0.54 | 0.5919 | -0.31309656 | 0.17884100 |

Figure 6: Glm results.

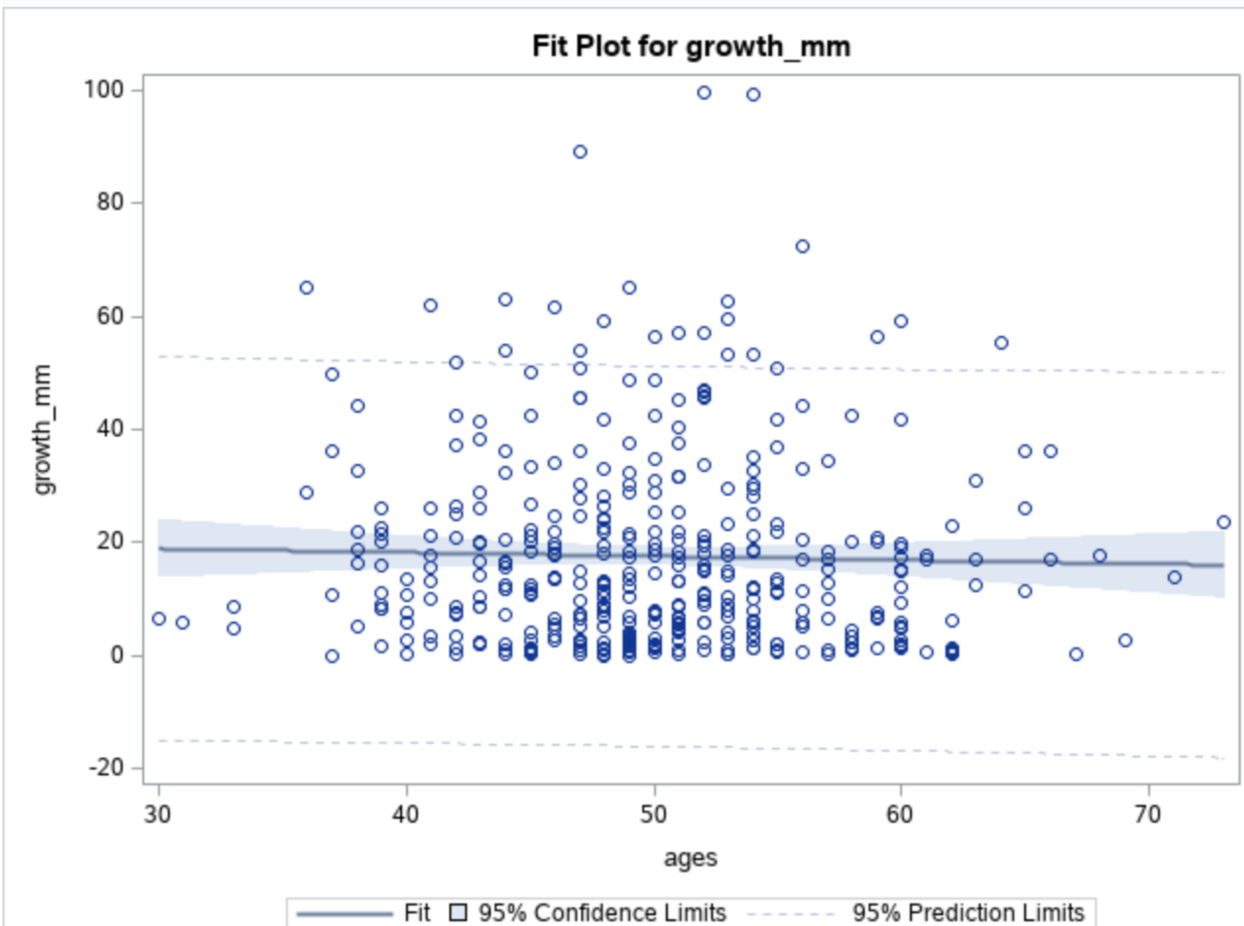


Figure 7: Age showing no relationship.