

Using the Length of Odontoblasts (Teeth Cells) of Guinea Pigs as Indicator of Vitamin C Intake

by: Mohammed K. Barakat

July 1, 2015

Overview

This report presents a study of the effect of Vitamin C intake on the odontoblasts (teeth cells) length carried out on 60 guinea pigs. The response (cell length) is measured at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

Data

The data used in this analysis (ToothGrowth) already exists in the R datasets package. You can load the data from R as shown in the following code chunk.

```
library(datasets)
data("ToothGrowth")
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
```

Basic Summary of the Data

ToothGrowth data used in the analysis consists of 60 observations with 3 variables, namely:

1. len: length of the odontoblast cell in microns
2. supp: supplement, which is the source of vitamin C intake of two types: orange juice (OJ), and ascorbic acid (VC).
3. dose: dose levels in milligrams of the supplement provided to the guinea pigs.

The experiment was carried out on 60 guinea pigs divided into 6 groups. Each group received specific type of supplement and dose after which lengths of odontoblasts were measured and recorded for each pig in the group. Below is a snapshot of the raw data as well as the structure of it.

```
head(ToothGrowth)
##      len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
str(ToothGrowth)
```

```
## 'data.frame': 60 obs. of 3 variables:
## $ len : num 4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
## $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
## $ dose: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

Exploratory Data Analysis

Two Exploratory Data Analysis tools were used at the beginning of the analysis in order to understand the data properties and to find possible patterns in the data; namely, Summary of data, and the Boxplot graph.

- **Summary of data**

Data was summarized in two levels by type of supplement (OJ and VC). The first summary level of orange juice supplement (OJ) showed that data was collected on 30 guinea pigs of which each 10 pigs received one of three levels of doses (0.5, 1.0, and 2.0 mg). The responses- lengths of odontoblasts- through the 30 pigs ranged from 8.20 to 30.90 microns with median and mean of 22.70 and 20.66 microns, respectively.

```
summary(subset(ToothGrowth, supp=="OJ"))
```

```
##      len      supp  dose
## Min.   : 8.20   OJ:30  0.5:10
## 1st Qu.:15.53  VC: 0   1  :10
## Median :22.70           2  :10
## Mean   :20.66
## 3rd Qu.:25.73
## Max.   :30.90
```

The second summary level of ascorbic acid supplement (VC) showed that data was collected on the other 30 guinea pigs of which each 10 pigs received the same levels of doses of that of the OJ supplement (0.5, 1.0, and 2.0 mg), but the lengths of odontoblasts through the 30 pigs ranged from 4.20 to 33.90 microns with median and mean of 16.50 and 16.96 microns, respectively.

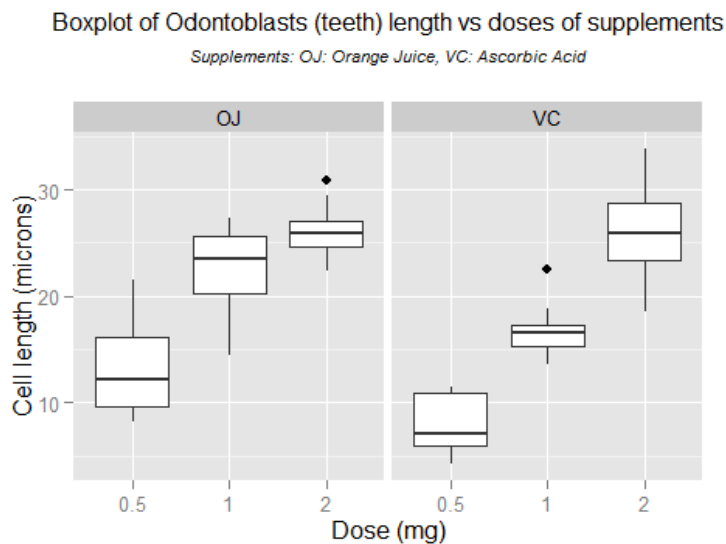
```
summary(subset(ToothGrowth, supp=="VC"))
```

```
##      len      supp  dose
## Min.   : 4.20   OJ: 0   0.5:10
## 1st Qu.:11.20  VC:30   1  :10
## Median :16.50           2  :10
## Mean   :16.96
## 3rd Qu.:23.10
## Max.   :33.90
```

- **Boxplot of data**

In order to see the variation of cell lengths across different doses for each supplement type a Boxplot (Box-and-Whisker plot) was used for each supplement type for the entire data observations.

The boxplot revealed interesting information about variation in cell lengths amongst doses of same supplement and between same doses of different supplements. It is clear that increase in cell length from dose 0.5 to 1.0 mg of OJ is more than that from 1.0 to 2.0 mg. Whereas nearly equal level of increase in length can be seen in the VC supplement from 0.5 to 1.0 to 2.0 mg.



Besides, 2.0 mg in OJ and 1.0 mg in VC exhibited much less variability than other doses in both supplements. This can be seen with the small height of the 1st to 3rd quartiles of the respective boxplots. The plots showed some skewness in data as in doses of 0.5 and 1.0 mg of VC with a couple of outlier points in doses 2.0 (OJ) and 1.0 (VC). (For code, see Appendix: EDA-Boxplot-R Code)

Hypothesis Testing and Confidence Intervals of cell lengths within the same supplement

In order to provide a statistical evidence of whether there is a significant difference in odontoblasts mean lengths when changing dose levels of vitamin C using the same supplement a Hypothesis Testing (HT) was done at different levels of doses in each supplement.

Testing significance of length difference amongst different doses when using orange juice (OJ)

The odontoblasts length data related to orange juice was tested for all pigs receiving different levels of doses. The null hypothesis assumed no change in cell lengths amongst

different doses where doses are tested in pairs. I.e. we assumed the null hypothesis of no difference in means of cell lengths amongst each pair of doses. Whereas the alternative hypothesis is that the means of cell lengths are not equal amongst each pair of doses.

- $H_0: \mu_1 = \mu_2$
- $H_a: \mu_1 \neq \mu_2$

Considering all combinations of doses under the same supplement we end up doing 3 Hypothesis tests for each of the OJ and VC supplements.

Table.1 below shows the results of Hypothesis Testing for cell lengths under each combination of the 3 levels of doses using the orange juice supplement. The table shows the confidence intervals of the mean length as well as the p-Value that is used to decide on the significance of difference in mean length between each pair of doses.

The HT results shows that the p-values for the three combinations (0.5 vs 1.0, 0.5 vs 2.0, and 1.0 vs 2.0) were (8.36e-05, 3.4e-07, and 0.0374, respectively). All p-values fall in the rejection region of the hypothesis below the α level of 0.05. Hence, the null hypothesis of no change between mean lengths of each pair of such doses is rejected. Which means that there's a statistical evidence that giving different doses of vitamin C through orange juice affects a change in the odontoblasts mean length. Moreover, the p-values shows a higher change in mean length between comparing 0.5 vs 2.0 mg and 0.5 vs 1.0 mg. This means that the higher the dose is, the longer the odontoblasts.

Table.1

| ## | supp | dose | CI.LL | CI.UL | p.Value |
|------|------|------------|--------|-------|----------|
| ## 1 | OJ | 0.5 vs 1.0 | -13.41 | -5.53 | 8.36e-05 |
| ## 2 | OJ | 0.5 vs 2.0 | -16.28 | -9.38 | 3.4e-07 |
| ## 3 | OJ | 1.0 vs 2.0 | -6.50 | -0.22 | 0.0374 |

(For R code, see Appendix: Orange Juice (OJ) Supplement-R code)

Testing significance of length difference amongst different doses when using ascorbic acid (VC)

Similarly, the odontoblasts length data related to ascorbic acid was tested for all pigs receiving different levels of doses. The null hypothesis assumed no change in cell lengths amongst different doses where doses are tested in pairs. Whereas the alternative hypothesis is that the means of cell lengths are not equal amongst each pair of doses.

- $H_0: \mu_1 = \mu_2$
- $H_a: \mu_1 \neq \mu_2$

Table.2 below shows the results of Hypothesis Testing for cell lengths under each combination of the 3 levels of doses for the ascorbic acid supplement. The table shows the confidence intervals of the mean length as well as the p-Value that is used to decide on the significance of difference in mean length between each pair of doses.

The HT results shows that the p-values for the three combinations (0.5 vs 1.0, 0.5 vs 2.0, and 1.0 vs 2.0) were (6.49e-07, 4.96e-09, and 3.4e-05, respectively). All p-values fall in the rejection region of the hypothesis below the α level of 0.05. Hence, the null hypothesis of no change between mean lengths of each pair of such doses is rejected. Which means that there's a statistical evidence that giving different doses of vitamin C through ascorbic acid affects a change in the odontoblasts mean length.

Table.2

| ## | supp | dose | CI.LL | CI.UL | p.Value |
|------|------|------------|--------|--------|----------|
| ## 1 | VC | 0.5 vs 1.0 | -11.26 | -6.32 | 6.49e-07 |
| ## 2 | VC | 0.5 vs 2.0 | -21.83 | -14.49 | 4.96e-09 |
| ## 3 | VC | 1.0 vs 2.0 | -12.97 | -5.77 | 3.4e-05 |

(For R code, see Appendix: Ascorbic Acid (VC) Supplement-R code)

Hypothesis Testing and Confidence Intervals of cell lengths between different supplements

In order to provide a statistical evidence of whether there is a significant difference in odontoblasts lengths when changing the supplement type for the same level of dose a Hypothesis Testing (HT) was done at different supplements for the same dose.

The odontoblasts length data related to each dose was tested for all pigs receiving this dose using either orange juice or ascorbic acid. The null hypothesis assumed no change in cell lengths amongst different supplement types for the same dose where supplements are tested in pairs. I.e. we assumed the null hypothesis of no difference in means of cell lengths amongst each pair of supplements. Whereas the alternative hypothesis is that the means of cell lengths are not equal amongst each pair of supplements for the same dose.

- $H_0: \mu_1 = \mu_2$
- $H_a: \mu_1 \neq \mu_2$

Table.3 below shows the results of Hypothesis Testing for cell lengths for each one of the doses under different supplements. The table shows the confidence intervals of the mean length as well as the p-Values.

The HT results shows that the p-values for the three combinations (OJ vs VC at 0.5 mg, OJ vs VC at 1.0 mg, and OJ vs VC at 2.0 mg) were (0.0053, 0.000781, and 0.964, respectively). The first two values fall in the rejection region of the hypothesis below the α level of 0.05. Whereas the third value is higher than α level. Hence, the null hypothesis of no change between mean length between OJ and VC supplements is rejected for doses 0.5 and 1.0 mg doses. Whereas we cannot reject the null hypothesis of dose 2.0 mg. Which means that there's a statistical evidence of difference between giving the same dose using two different supplements for doses 0.5 and 1.0 mg. While there's no enough evidence to prove difference in odontoblasts length when a 2.0 mg is given using different supplements.

Table.3

| ## | supp | dose | CI.LL | CI.UL | p.Value |
|------|----------|------|-------|-------|----------|
| ## 1 | OJ vs VC | 0.5 | 1.77 | 8.73 | 0.0053 |
| ## 2 | OJ vs VC | 1.0 | 2.84 | 9.02 | 0.000781 |
| ## 3 | OJ vs VC | 2.0 | -3.72 | 3.56 | 0.964 |

(For R code, see Appendix: Hypothesis Testing and Confidence Intervals of cell lengths between different supplements-R code)

Assumptions and Conclusion

List of assumptions used in the analysis

- The experiment was carried out on 60 guinea pigs divided in 6 groups of 10's. Each group is independent of other groups and have been tested for receiving one level of dose by one type of supplement as follows:
 1. 0.5 mg does by orange Juice: 10 pigs
 2. 1.0 mg does by orange Juice: 10 pigs
 3. 2.0 mg does by orange Juice: 10 pigs
 4. 0.5 mg dose by ascorbic Acid: 10 pigs
 5. 1.0 mg dose by ascorbic Acid: 10 pigs
 6. 2.0 mg dose by ascorbic Acid: 10 pigs
- Doses are given to pigs for equal periods of time enough to compare their odontoblasts lengths fairly.
- Experiment pigs are equally healthy and comparable to be tested in different conditions of dose level and supplement type.
- 0.05 Probability of Type I error (α) was assumed in performing the hypothesis testing. This gives a 95% confidence interval of mean length.
- Null hypotheses assumed equal length means. Whereas alternative hypotheses assumed not equal length means.
- Independent Groups t-test type was used rather than the Paired t-test (paired=FALSE) since the 6 groups of pigs are independent of each other. I.e. each group was tested once (one dose level with one supplement type).
- Constant variance across the 6 groups is assumed (var.equal=TRUE).

Conclusion

- There's a statistical evidence that giving different doses of vitamin C through orange juice affects a change in the odontoblasts mean length. Moreover, a higher change in mean length is observed in 0.5 vs 2.0 mg than the change observed in 0.5 vs 1.0 mg. This means that the higher the dose is, the longer the odontoblasts. This conclusion is also true when giving different doses of vitamin C through ascorbic acid (VC).

- When comparing cell lengths for each one of the doses under different supplements there's a statistical evidence that the mean length changes when we give the same dose using both different supplements. This is true for doses 0.5 and 1.0 mg. While there's no enough evidence to prove difference in odontoblasts mean length when a 2.0 mg dose is given using different supplements.

APPENDIX

EDA-Boxplot-R Code

```
library(ggplot2)
g<-ggplot(ToothGrowth,aes(dose,len))+geom_boxplot(aes(factor(dose),len))+face
t_grid(.~supp)+
  labs(x="Dose (mg)",y="Cell length (microns)")+theme(plot.title = elem
ent_text(size = 12, face = "bold", colour = "black", vjust = +1))+ggtitle(exp
ression(atop("Boxplot of Odontoblasts (teeth) length vs doses of supplements"
,atop(italic("Supplements: OJ: Orange Juice, VC: Ascorbic Acid")))))
g
```

Hypothesis Testing and Confidence Intervals of cell lengths within the same supplement

Orange Juice (OJ) Supplement-R code

First, subset the data into required groups to study the HT. Data is subsetted in pairs of doses for OJ supplement. E.g. "gOJ0.5v1.0" subset refers to the group of observations of 0.5 and 1.0 mg doses of OJ supplement.

```
gOJ0.5v1.0<-subset(ToothGrowth,supp %in% "OJ" & dose %in% c(0.5,1.0))
gOJ0.5v2.0<-subset(ToothGrowth,supp %in% "OJ" & dose %in% c(0.5,2.0))
gOJ1.0v2.0<-subset(ToothGrowth,supp %in% "OJ" & dose %in% c(1.0,2.0))
```

Second, run the t-test on each subset of data to compare mean lengths (μ_1 , μ_2) at two levels of doses to get the confidence interval and p-value then decide on the significance of difference in means.

$H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$ where (μ_1 is mean length at 0.5 mg and μ_2 is mean length at 1.0 mg)

```
g1ConfInt<-round(as.vector(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=g
OJ0.5v1.0)$conf.int),2)
g1pVal<-format.pval(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=gOJ0.5v1
.0)$p.value,3)
```

The above hypothesis test is repeated for the other two subsets where μ_1 is mean length at 0.5 mg and μ_2 is mean length at 2.0 mg for the second subset, and μ_1 is mean length at 1.0 mg and μ_2 is mean length at 2.0 mg for the third subset.

```
g2ConfInt<-round(as.vector(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=g
OJ0.5v2.0)$conf.int),2)
g2pVal<-format.pval(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=gOJ0.5v2
.0)$p.value,3)

g3ConfInt<-round(as.vector(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=g
OJ1.0v2.0)$conf.int),2)
g3pVal<-format.pval(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=gOJ1.0v2
.0)$p.value,3)
```


Combine HT results (CI and p-value) of the orange juice subsets in one dataset.

Table.1-R code

```
data.frame(supp=rep("OJ",3),dose=c("0.5 vs 1.0","0.5 vs 2.0","1.0 vs 2.0"),CI
.LL=c(g1ConfInt[1],g2ConfInt[1],g3ConfInt[1]),CI.UL=c(g1ConfInt[2],g2ConfInt[
2],g3ConfInt[2]),p.Value=c(g1pVal,g2pVal,g3pVal))
```

Ascorbic Acid (VC) Supplement-R code

Repeat the previous steps using the subset of ascorbic acid. That is:

- Subset the data into required groups to study the HT on VC supplement
- Run the t-test on the subset of data to compare mean lengths (μ_1 , μ_2) at two levels of doses to get the confidence interval and p-value.
- Combine HT results (CI and p-value) of the ascorbic acid subsets in one dataset.
- Decide on the significance of difference in means.

```
gVC0.5v1.0<-subset(ToothGrowth,supp %in% "VC" & dose %in% c(0.5,1.0))
gVC0.5v2.0<-subset(ToothGrowth,supp %in% "VC" & dose %in% c(0.5,2.0))
gVC1.0v2.0<-subset(ToothGrowth,supp %in% "VC" & dose %in% c(1.0,2.0))
```

$H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$ where (μ_1 is mean length at 0.5 mg and μ_2 is mean length at 1.0 mg)

```
g4ConfInt<-round(as.vector(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=g
VC0.5v1.0)$conf.int),2)
g4pVal<-format.pval(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=gVC0.5v1
.0)$p.value,3)
```

Again, the above hypothesis test is repeated for the other two subsets where μ_1 is mean length at 0.5 mg and μ_2 is mean length at 2.0 mg for the second subset, and μ_1 is mean length at 1.0 mg and μ_2 is mean length at 2.0 mg for the third subset. Code is not shown for limitation of space in Appendix.

```
g5ConfInt<-round(as.vector(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=g
VC0.5v2.0)$conf.int),2)
g5pVal<-format.pval(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=gVC0.5v2
.0)$p.value,3)

g6ConfInt<-round(as.vector(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=g
VC1.0v2.0)$conf.int),2)
g6pVal<-format.pval(t.test(len~dose,paired=FALSE,var.equal=TRUE,data=gVC1.0v2
.0)$p.value,3)
```

Combine HT results (CI and p-value) of the ascorbic acid subsets in one dataset.

Table.2-R code

```
data.frame(supp=rep("VC",3),dose=c("0.5 vs 1.0","0.5 vs 2.0","1.0 vs 2.0"),CI
.LL=c(g4ConfInt[1],g5ConfInt[1],g6ConfInt[1]),CI.UL=c(g4ConfInt[2],g5ConfInt[
2],g6ConfInt[2]),p.Value=c(g4pVal,g5pVal,g6pVal))
```

Hypothesis Testing and Confidence Intervals of cell lengths between different supplements-R code

First, subset the data into required groups to study the HT. Data is subsetted in pairs of supplements for each dose. E.g. "g0.5OJvsVC" subset refers to the group of observations of OJ and VC supplements at the 0.5 mg dose.

```
g0.5OJvsVC<-subset(ToothGrowth,dose ==0.5)
g1.0OJvsVC<-subset(ToothGrowth,dose ==1.0)
g2.0OJvsVC<-subset(ToothGrowth,dose ==2.0)
```

Second, run the t-test on each subset of data to compare means of length (μ_1 , μ_2) at both supplement types to get the confidence interval and p-value then decide on the significance of difference in means.

$H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$ where (μ_1 is mean length using OJ and μ_2 is mean length using VC, both at 0.5 mg dose).

```
g7ConfInt<-round(as.vector(t.test(len~supp,paired=FALSE,var.equal=TRUE,data=g0.5OJvsVC)$conf.int),2)
g7pVal<-format.pval(t.test(len~supp,paired=FALSE,var.equal=TRUE,data=g0.5OJvsVC)$p.value,3)
```

The above hypothesis test is repeated for the other two subsets where μ_1 is mean length using OJ and μ_2 is mean length using VC, both at 1.0 mg dose for the second subset, and μ_1 is mean length using OJ and μ_2 is mean length using VC, both at 2.0 mg dose for the third subset. Code is not shown for limitation of space in Appendix.

```
g8ConfInt<-round(as.vector(t.test(len~supp,paired=FALSE,var.equal=TRUE,data=g1.0OJvsVC)$conf.int),2)
g8pVal<-format.pval(t.test(len~supp,paired=FALSE,var.equal=TRUE,data=g1.0OJvsVC)$p.value,3)
g9ConfInt<-round(as.vector(t.test(len~supp,paired=FALSE,var.equal=TRUE,data=g2.0OJvsVC)$conf.int),2)
g9pVal<-format.pval(t.test(len~supp,paired=FALSE,var.equal=TRUE,data=g2.0OJvsVC)$p.value,3)
```

Combine HT results (CI and p-value) of the doses subsets in one dataset.

Table.3-R code

```
data.frame(supp=rep("OJ vs VC",3),dose=c(0.5,1.0,2.0),CI.LL=c(g7ConfInt[1],g8ConfInt[1],g9ConfInt[1]),CI.UL=c(g7ConfInt[2],g8ConfInt[2],g9ConfInt[2]),p.Value=c(g7pVal,g8pVal,g9pVal))
```