ToothGrowth data: Effect of vitamin C on odontoblasts length of guinea pigs

Fernando Flores

December 20th, 2015

Synopsis

The current report aims to investigate the effect of vitamin C on the length of odontoblasts (cells responsible for tooth growth). The source data provides the measurements done in an experiment where 60 guinea pigs received one of three dose levels of vitamin C (0,5, 1 and 2 mg/day) by one of two delivery methods (orange juice or ascorbic acid). Through Student's t statistic and confidence intervals we research the relationship between the cells growth by the two delivery methods and the three dose.

Exploratory Data Analysis

Useful information about the dataset can be found in:

1. [ToothGrowth docs](#)
2. [Crampton, E. W. (1947) The growth of the odontoblast of the incisor teeth as a criterion of vitamin C intake of the guinea pig. The Journal of Nutrition](#)

The summary of the source data is the following:

```
##       len          supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000

## '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: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Figure 1. Summary of the source data.

The data is organized in 60 observations of 3 columns, where each observation corresponds to the measurement of the odontoblast cells of one of 60 guinea pigs.

```
##     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
```

Figure 2. First observations of the dataset.

For a better manipulation of the information, the values in column `supp` are modified with a more verbose description, keeping in mind that `OJ = Orange juice` and `VC = Vitamin C` (ascorbic acid).

The plots in *Figure 3* and *Figure 4* reports useful information to establish hypothesis of how the measured length of the odontoblast cells varies by supplement and dose.
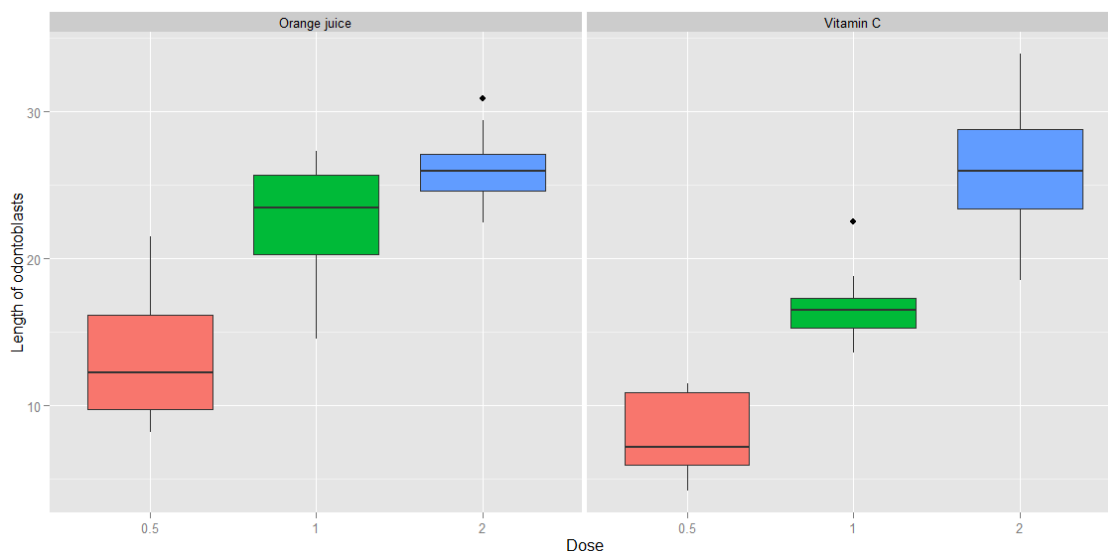

Figure 3. Boxplots of the measured length of the odontoblasts of 60 guinea pigs, by supplement and dose.
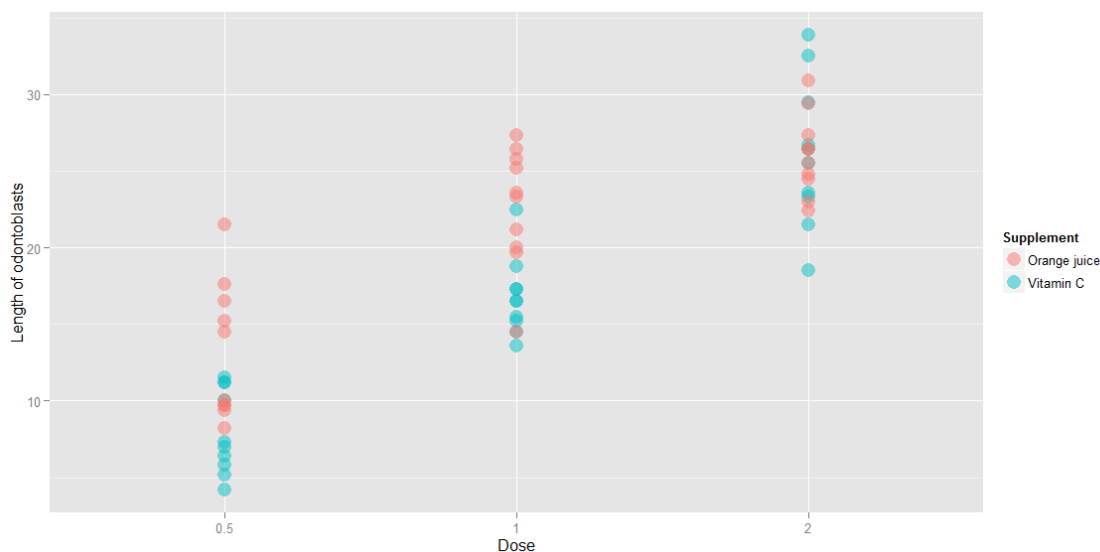

Figure 4. Scatterplot of the measured length of the odontoblasts of 60 guinea pigs, by supplement and dose.

Confidence Interval and statistic tests

Looking at the data summary and plots after the exploratory data analysis, there are potential questions of interest to research by comparing the measured length of odontoblasts by delivery method and dose.

The Student's t statistic is used here to try to answer those questions.

```
##                       statType intervalMin intervalMax tValue
## 1 Orange juice, 0.5 vs 1 mg/day      -13.42       -5.52  -5.05
## 2 Orange juice, 0.5 vs 2 mg/day      -16.34       -9.32  -7.82
## 3   Orange juice, 1 vs 2 mg/day       -6.53       -0.19  -2.25
## 4    Vitamin C, 0.5 vs 1 mg/day      -11.27       -6.31  -7.46
## 5    Vitamin C, 0.5 vs 2 mg/day       -21.9      -14.42 -10.39
## 6      Vitamin C, 1 vs 2 mg/day      -13.05       -5.69  -5.47
## 7     Dose 0.5 mg/day, OJ vs VC        1.72        8.78   3.17
## 8       Dose 1 mg/day, OJ vs VC         2.8        9.06   4.03
## 9       Dose 2 mg/day, OJ vs VC        -3.8        3.64  -0.05
```

Figure 5. Statistics of combined comparisons by supplement and dose. Columns shows the stat comparison, confidence intervals (min and max values), and the t statistic.

The first six rows of the *Figure 5* researchs the variation in dose given a specific delivery method.

For `Orange juice`, comparing the dose between 0.5 vs 1 mg/day, 1 vs 2 mg/day and 0.5 vs 2 mg/day, all the confidence intervals are below zero, concluding that at a lower dose, the odontoblast cells grows less. By looking at the t value of each stat, the test for 0.5 vs 1 mg/day reports a bigger growth proportional to the dose variation.

For `Vitamin C`, comparing the dose between 0.5 vs 1 mg/day, 1 vs 2 mg/day and 0.5 vs 2 mg/day, all the confidence intervals are below zero, concluding that at a lower dose, the odontoblast cells grows less. By looking at the t value of each stat, the test for 1 vs 2 mg/day reports a bigger growth proportional to the dose variation.

The last three rows of the figure researchs the variation in delivery method given a specific dose.

For a dose of 0.5 or 1 mg/day, comparing the supplement `Orange juice` vs `Vitamin C`, the confidence intervals are above zero, concluding that the cells grows more for the first delivery method.

Finally, for a dose of 2 mg/day, comparing the supplement `Orange juice` vs `Vitamin C`, the confidence interval includes zero which doesn't give a significant statistical conclusion between the delivery methods.

Assumptions

- According to the reference documentation the measurements are unrelated to each other, so the observations used in the statistic analysis can't be paired.
- There is no enough information about the variances of the samples used for the statistics, meaning the t tests were executed for unequal variances.
- 95% confidence intervals are used for the tests.

Conclusions

- For larger dose, the odontoblast cells grows larger, independent of the delivery method.
- When the delivery method is `Orange juice`, the growth of odontoblast cells is proportionally larger when the dose is between 0.5 and 1 mg/day.
- When the delivery method is `Vitamin C`, the growth of odontoblast cells is proportionally larger when the dose is between 1 and 2 mg/day.
- Comparing the delivery methods, `Orange juice` provides more growth of odontoblast cells than `Vitamin C` when the dose is 0.5 or 1 mg/day. For a dose of 2 mg/day, there is no statistical difference observed regarding delivery method.

Appendix

A1. Code for data loading

```
library(datasets)
data(ToothGrowth)
```

A2. Code for data summary

```
dimTG <- dim(ToothGrowth)
headTG <- head(ToothGrowth)

summary(ToothGrowth)
str(ToothGrowth)
```

A3. Code for improvement of the data labels

```
## For plot reasons, change the factor labels
levels(ToothGrowth$supp)[1] <- "Orange juice"  ## "OJ" in the original
levels(ToothGrowth$supp)[2] <- "Vitamin C"  ## "VC" in the original
```

A4. Code for boxplot

```
library(ggplot2)
boxplotTG <- ggplot(ToothGrowth,
                    aes(x = as.factor(dose), y = len)) +
    geom_boxplot(aes(fill = as.factor(dose))) +
    labs(list(x = "Dose",
              y = "Length of odontoblasts")) +
    theme(legend.position = "none") +
    facet_grid(. ~ supp)
boxplotTG
```

A5. Code for scatterplot

```
scatterplotTG <- ggplot(ToothGrowth,
                        aes(x = as.factor(dose), y = len)) +
    geom_point(aes(colour = supp), size = 5, alpha = 0.5) +
    labs(list(x = "Dose",
              y = "Length of odontoblasts")) +
    labs(colour = "Supplement")
scatterplotTG
```

A6. Code for statistics

```
### supp OJ, dose 0.5 vs 1
OJ_0.5_1 <- subset(ToothGrowth, supp == "Orange juice" & dose %in% c(0.5, 1))
test_OJ_0.5_1 <-
    t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = OJ_0.5_1)

### supp OJ, dose 0.5 vs 2
OJ_0.5_2 <- subset(ToothGrowth, supp == "Orange juice" & dose %in% c(0.5, 2))
test_OJ_0.5_2 <-
    t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = OJ_0.5_2)
```

```r
### supp OJ, dose 1 vs 2
OJ_1_2 <- subset(ToothGrowth, supp == "Orange juice" & dose %in% c(1, 2))
test_OJ_1_2 <-
    t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = OJ_1_2)


### supp VC, dose 0.5 vs 1
VC_0.5_1 <- subset(ToothGrowth, supp == "Vitamin C" & dose %in% c(0.5, 1))
test_VC_0.5_1 <-
    t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = VC_0.5_1)


### supp VC, dose 0.5 vs 2
VC_0.5_2 <- subset(ToothGrowth, supp == "Vitamin C" & dose %in% c(0.5, 2))
test_VC_0.5_2 <-
    t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = VC_0.5_2)


### supp VC, dose 1 vs 2
VC_1_2 <- subset(ToothGrowth, supp == "Vitamin C" & dose %in% c(1, 2))
test_VC_1_2 <-
    t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = VC_1_2)



### dose 0.5, supp OJ vs VC
TG0.5_VC_OJ <- subset(ToothGrowth, dose == 0.5)
test_TG0.5_VC_OJ <-
    t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = TG0.5_VC_OJ)


### dose 1, supp OJ vs VC
TG1_VC_OJ <- subset(ToothGrowth, dose == 1)
test_TG1_VC_OJ <-
    t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = TG1_VC_OJ)


### dose 2, supp OJ vs VC
TG2_VC_OJ <- subset(ToothGrowth, dose == 2)
test_TG2_VC_OJ <-
    t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = TG2_VC_OJ)

## data frame with statistics
summaryStats <-
    data.frame(statType = "Orange juice, 0.5 vs 1 mg/day",
               intervalMin = round(test_OJ_0.5_1$conf.int[1], 2),
               intervalMax = round(test_OJ_0.5_1$conf.int[2], 2),
               tValue = round(as.numeric(test_OJ_0.5_1$statistic), 2),
               stringsAsFactors = FALSE)
summaryStats <- rbind(summaryStats,
                      c("Orange juice, 0.5 vs 2 mg/day",
                        round(test_OJ_0.5_2$conf.int[1], 2),
                        round(test_OJ_0.5_2$conf.int[2], 2),
                        round(as.numeric(test_OJ_0.5_2$statistic), 2)),
                      c("Orange juice, 1 vs 2 mg/day",
                        round(test_OJ_1_2$conf.int[1], 2),
                        round(test_OJ_1_2$conf.int[2], 2),
                        round(as.numeric(test_OJ_1_2$statistic), 2)),
                      c("Vitamin C, 0.5 vs 1 mg/day",
                        round(test_VC_0.5_1$conf.int[1], 2),
```

```
                    round(test_VC_0.5_1$conf.int[2], 2),
                    round(as.numeric(test_VC_0.5_1$statistic), 2)),
                  c("Vitamin C, 0.5 vs 2 mg/day",
                    round(test_VC_0.5_2$conf.int[1], 2),
                    round(test_VC_0.5_2$conf.int[2], 2),
                    round(as.numeric(test_VC_0.5_2$statistic), 2)),
                  c("Vitamin C, 1 vs 2 mg/day",
                    round(test_VC_1_2$conf.int[1], 2),
                    round(test_VC_1_2$conf.int[2], 2),
                    round(as.numeric(test_VC_1_2$statistic), 2)),
                  c("Dose 0.5 mg/day, OJ vs VC",
                    round(test_TG0.5_VC_OJ$conf.int[1], 2),
                    round(test_TG0.5_VC_OJ$conf.int[2], 2),
                    round(as.numeric(test_TG0.5_VC_OJ$statistic), 2)),
                  c("Dose 1 mg/day, OJ vs VC",
                    round(test_TG1_VC_OJ$conf.int[1], 2),
                    round(test_TG1_VC_OJ$conf.int[2], 2),
                    round(as.numeric(test_TG1_VC_OJ$statistic), 2)),
                  c("Dose 2 mg/day, OJ vs VC",
                    round(test_TG2_VC_OJ$conf.int[1], 2),
                    round(test_TG2_VC_OJ$conf.int[2], 2),
                    round(as.numeric(test_TG2_VC_OJ$statistic), 2))
                  )

summaryStats
```