Comparison of Parasite Treatment in Two Rabbit Species

Andrew Mullins

August 16, 2015

Summary

The UGA Veterinary Clinic is interested in determining the best way to treat a particular parasitic infection in two species of rabbits. The investigators have identified a number of medications and materials that can be used in combination to treat the disease and would like to know which works best for each species. To answer this question, they designed an experiment consisting of 300 rabbits (evenly divided between species) who were infected with the parasite and treated using various combinations of medication and materials. The rabbits were then housed for a month before a measurement was taken that quantified the amount of remaining infection.

The data as provided by the researchers had a fairly problematic structure. The biggest issue was the large number of zero scores, which made up 105 of the 300 observations in the dataset. These scores are somewhat inherent to the design of this experiment, as a zero score response indicates that a treatment was not effective at treating the infection. Thus any ineffective treatment (including a control group) was scored entirely as zero. These zero values created an extreme deviation from the assumptions required for ANOVA. To address this issue, I removed the control group and one particularly problematic factor level. This left me with a balanced design that I was able to analyze. I also performed a smaller subanalysis on the factor level I had removed.

To analyze the remaining data, I fit several ANOVA models. These models were able to answer a few of the researchers' initial questions. First, there is in fact a significant difference between the species in terms of what treatment combinations are most effective. Secondly, the different levels of medicine and material do interact with each other. Having determined this, I used multiple comparison techniques to identify the particular treatment combinations that were most effective.

The multiple comparison techniques indicated that for each species there were several combinations of medicine and material that were equally effective at treating the infection. These results were found to be at least consistent with what I saw in the subanalysis, although considering the reduced scope of that work it is hard to put much weight on this result.

Ultimately, I was able to recommend to the investigators that Material 1 in combination with Medicine 5 is the best and most commercially viable treatment option. This choice is one of the most effective combinations across both species. Additionally, contextual information provided by the investigators indicate that Material 1 is cheaper and more commonly available than other materials that do not perform significantly better. I was also able to make a recommendation for future experimentation that would avoid some of the issues with the existing design.

Introduction

The UGA Veterinary Clinic is interested in determining how to best treat rabbits who have become infected with a certain parasite. In particular, they want to determine which of several combinations of different medications and application materials will most effectively stop the spread of the parasite. To investigate this problem, the researchers randomly selected 300 uninfected rabbits (150 from each of two different species) and exposed them to the parasite.

The investigators have 5 levels of medicine and 6 levels of application material they are interested in testing. Each combination of medicine and material was applied to 5 rabbits from each species. In tabular form, the experiment looks like this:

Experimental Design

	Medicine 1	Medicine 2	Medicine 3	Medicine 4	Medicine 5
Material 1	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2
Material 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2
Material 3	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2
Material 4	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2
Material 5	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2
Material 6	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2	5 × Species 1 5 × Species 2

There are a number of relevant pieces of contextual information about the medicine and materials that are worth noting. First, Medicine 1 is a control that should have no effect on the infection. The researchers included this as a baseline to measure from. Additionally, Materials 4 and 5 represent combinations of other materials (1-2 and 1-3 respectively). Material 1 is cheap and commonly used, while Materials 2 and 3 are more expensive and rarely used. Finally, Material 6 is a new material that has similar properties to Material 1.

The researchers treated each rabbit with the assigned medicine-material combination, then waited a month before measuring the results. To determine how well the treatment worked, the researchers recorded the circular area around the initial infection point that was parasite-free. The observed scores range from 0 to 44, with lower scores indicating the researcher had to go less distance to find a parasite. In other words, lower scores represent that a treatment did not work as well as a higher score.

There are a number of questions posed by the investigators. First, they want to know which combination of treatment is the best for each species. Secondly, they're curious if there are multiple combinations that can be effective and third, they want to know if there

is an interaction between the medicine and material used to apply it.

Exploratory Data Analysis

As a preliminary step, I first verified that the data was complete and as the researchers had specified. Having confirmed that there were no missing values, I created some basic summary statistics and a histogram for the response variable, Effect. These can be found in the appendix (page 9) but I will discuss their implications here. The histogram revealed an extremely large concentration of observations in the lowest bin. Further investigation revealed that 105 of the 300 rabbits had Effect scores of 0. In one sense, this is encouraging, because it means there are likely cases where treatments were completely ineffective at fighting the infection. Considering the presence of a control treatment (Medicine 1), we might even expect to see a large number of zero values.

Unfortunately, the presence of these zero scores has another implication for our analysis. The design specified by the investigators (three factors with discrete levels) would naturally lend itself to analysis of variance, but these zero values represent a huge violation of our assumption of normality and constant variance of error. In some situations we can address these violations through transformation, but in this case any transformation we could apply would just map the zeroes to some other value and create the same problem we have now.

A natural course of action is to attempt to determine why these zero scores are occuring. Luckily, there seems to be a very clear pattern. Here are two tables (one for each species) showing the proportion of non-zero scores for each medicine-material combination. A '0' indicates that all the rabbits in a particular cell were scored zero while a '1' indicates that all of the rabbits received a non-zero score.

Non-Zero Score Proportions (Material by Medicine)

	1	2	3	4	5
1	0	1	1	1	1
2	0	1	0	1	1
3	0	1	0	1	1
4	0	1	0	1	1
5	0	1	0	1	1
6	0	1	1	1	1

Species 1

	1	2	3	4	5
1	0	1	0	1	1
2	0	1	0	1	1
3	0	1	0	1	1
4	0	1	0	1	1
5	0	1	0	1	1
6	0	1	1	1	1

Species 2

There are two obvious takeaways here. First, the baseline group universally scored zero. This makes sense and is consistent with the researchers' belief that the control should have no effect on the parasite. Secondly, Medicine 3 appears to be another source of many of the zero values, but only when combined with certain materials that vary by species.

As previously mentioned, this issue cannot be addressed through transformation. However, continuing to include this data will lead to major issues. In particular, the fact that all

of the zero values fully comprise various treatment combinations means that we will almost certainly have heteroscedasticity (the variance in error will be zero for those groups and presumably something larger for the others).

There are three initially obvious possible ways to address this issue. One possibility is to remove all the treatment combinations that resulted in entirely zero score rabbits under the assumption that these combinations are ineffective. The main advantage of doing this is that it completely solves our violation of assumptions. However, it has a disadvantage (aside from just the fact that we're losing data): we will no longer have a balanced design. In fact, we end up with something even worse - a design where we have no observations at all for certain combinations of treatments. This presents significant challenges to determining whether or not those combinations and their corresponding main factors are significant.

A second possibility is to remove the control medication while leaving in the Medicine 3 observations of zero. This approach represents a bit of a compromise. On the one hand, it potentially does not do enough to address the issues with our model assumptions. On the other hand, we still have a balanced design and all the advantages that entails.

Finally, a third possibility is to completely remove the control medication and **all** Medicine 3 entries. This is by far the most destructive of the three options. I'd be throwing away 40% of the data provided by the investigators. In exchange for this, I would be fixing both the assumptions of the model I plan to use and avoid creating missing treatment combinations.

Ultimately, I settled on the third method. While this throws away a lot of the data, this is not as bad as it initially looks. For starters, no matter how I proceeded I would almost certainly need to get rid of the baseline group (20% of the original data), so the 40% figure is somewhat inflated. Secondly, even after removing these values I will still have 180 rabbits left in a $2 \times 3 \times 6$ design with 5 replications, more than enough to perform a fruitful analysis. To compensate for the removal of any information about Medicine 3, I also performed a short separate analysis of the cells for which I have useable data.

After removing the zero values, the distribution of Effect looks much more reasonable (see page 10). While still not totally normal, some amount of deviation from normality in the response distribution should be fine as long as the residuals appear normal. This can be verified after a model has been fit through standard diagnostics and residual analysis.

Analysis

Having prepared the data for analysis as described above, I fit a three-way ANOVA model with interactions. Here is the resulting ANOVA table:

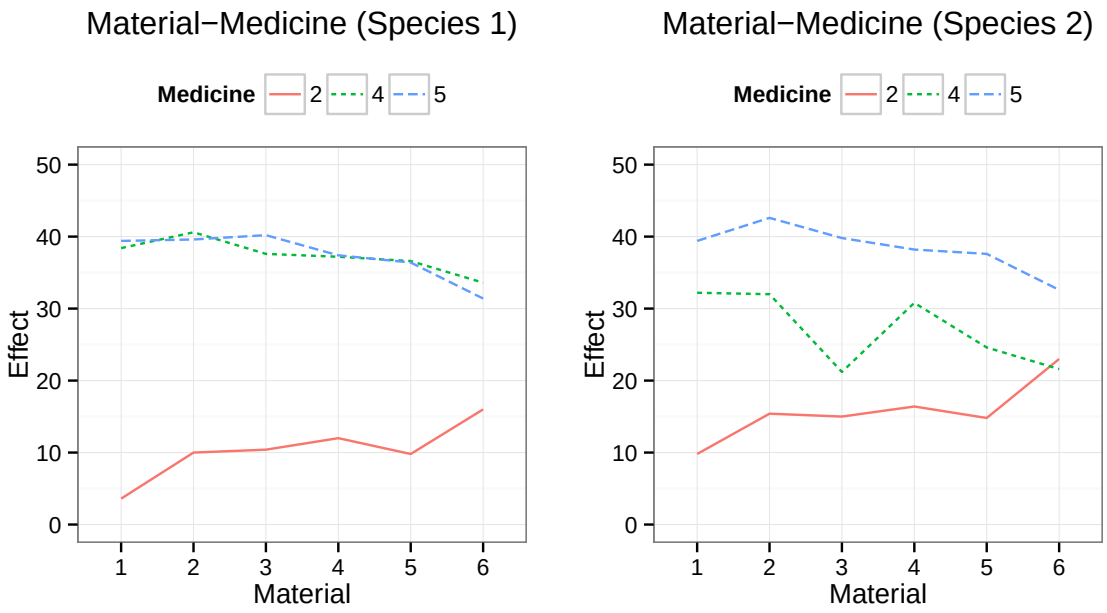
Three-way ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Medicine	2	20373.03	10186.52	1580.67	0.0000
Material	5	291.80	58.36	9.06	0.0000
Species	1	74.76	74.76	11.60	0.0009
Medicine:Material	10	1686.17	168.62	26.16	0.0000
Medicine:Species	2	1963.14	981.57	152.31	0.0000
Material:Species	5	90.58	18.12	2.81	0.0187
Medicine:Material:Species	10	134.32	13.43	2.08	0.0293
Residuals	144	928.00	6.44		

All three main effects, their interactions and the three-way interaction were found to be significant at $\alpha = .05$. The most important finding here is that we have a significant three-way interaction. This implies that the effect of particular material-medicine combinations is different between the species. Before I investigated this, I verified the assumptions of my model. Performing an analysis on the residuals (see page 12), it appears that our assumptions are fairly well met, likely due to removing the problematic levels of the medicine factor. Were we still having an issue, we would be able to easily see the violation of these assumptions in the distribution of residuals by levels of the medicine factor and in the normal Q-Q plot.

In the ANOVA table above, there is evidence that the effect of the medicine-material interaction is different between the two species. Here are two plots visualizing this relationship:

Interaction Plots



There is a clearly different pattern between the two species. From the plot, we can see that in the first species Medicine 4 and Medicine 5 seem to be about equally effective (and superior to Medicine 2). However, in Species 2 Medicine 5 is more effective than Medicine 4, and the gap between Medicine 4 and Medicine 2 is not so clear. A significant three-way interaction where the two factor levels look as different as we are seeing here suggest an investigation into the simpler two-way interaction effects that the researchers are primarily interested in. To do so, I treated each species as though it were a separate two-way layout.

Species 1

The ANOVA table for the Species 1 two-way layout can be found in the appendix (page 12). Unsurprisingly, both main effects and their interaction remained significant. This implies that at least one pair of medicine and material had a greater effect on the parasite than the other. To determine which combination was most effective, I used the Tukey method for multiple comparisons. A full table showing the treatment estimates and their groupings can be found on appendix page 13. Two rows that share one or more grouping numbers are not significantly different from each other at $\alpha = .05$ after applying Tukey's HSD adjustment. These results are consistent with the interaction plot. The combination with the highest

estimated effect is Material 2-Medicine 4, but this is not significantly different from a number of other treatment combinations. In fact, any of Materials 1-5 in combination with Medicine 4 or 5 appear to work equally well.

Species 2

A two-way ANOVA performed on the Species 2 data found similar results to Species 1. Again, a table of estimated group means can be found in the appendix (page 14). The results here are also similar to what we might guess from looking at the interaction plot. For Species 2, it appears that any Material from 1-5 in combination with Medicine 5 will be the most effective treatment for the parasitic infection.

Notes on Medicine 3

Due to structural issues with the data, I made the decision to drop Medicine 3 from my larger analysis. In the interest of completion, here are a few notes on a smaller analysis I've performed on the removed data. To perform this analysis, instead of removing Medicine 3, I removed the materials that had zero cells in combination with Medicine 3. Having established in my larger model that there is a significant interaction between species and medicine-material combination, I further split the removed data along species. This left me with a two-way layout (for Species 1) and a one-way layout (for Species 2). Following similar steps to the larger analysis, I fit these ANOVA models and performed the Tukey method for multiple comparisons. The tables can be seen in the appendix (page 14).

While these results are obviously more limited than what I was able to find in the larger analysis, they seem to reinforce some of the results found there. In Species 1, this smaller analysis identified Material 1 in combination with Medicine 4 or 5 as significantly more effective than the other combinations. This is consistent with the results found in the larger model, although that model was also able to consider Materials 2, 3, 4 and 5. In Species 2 (which only had data for Material 6), Medicine 5 was identified as significantly more effective than the other medicines. Again, while limited, this result is consistent with the results in the larger model.

Conclusion

One major obstacle in completing this analysis was the structural problem created by the presence of numerous zero entries in the response variable. In future studies, steps should be taken to avoid this situation. One possibility is, instead of using the existing scoring system (which will inherently create zero entries whenever a treatment does not work), the researchers should measure infection in some other way that more directly quantifies the degree of parasitic infestation on the rabbit. For example, the researchers could choose a specific radius around the infection region and count the number of parasites (assuming such a count is possible to obtain). Under this measurement system, zero values would only be created when the parasite was completely eradicated. The response would also have a

direct physical interpretation. For example, we would be able to estimate the decrease in parasite density associated with a particular treatment combination.

The investigators' primary question is what Medicine-Material combination is the most effective treatment for each species. For Species 1, the treatment with the greatest estimated effect is Material 2 in combination with Medicine 4. For Species 2, the most effective treatment is Material 2 in combination with Medicine 5. However, in each of these cases there are a number of different combinations that are not significantly different from the so-called 'best' treatments. These are summarized in the following tables (any combination of Material and Medicine should be equally effective on the specified species):

Most Effective Treatment Combinations by Species

Material	Medicine
1, 2, 3, 4, 5	4, 5

Species 1

Material	Medicine
1, 2, 3, 4, 5	5

Species 2

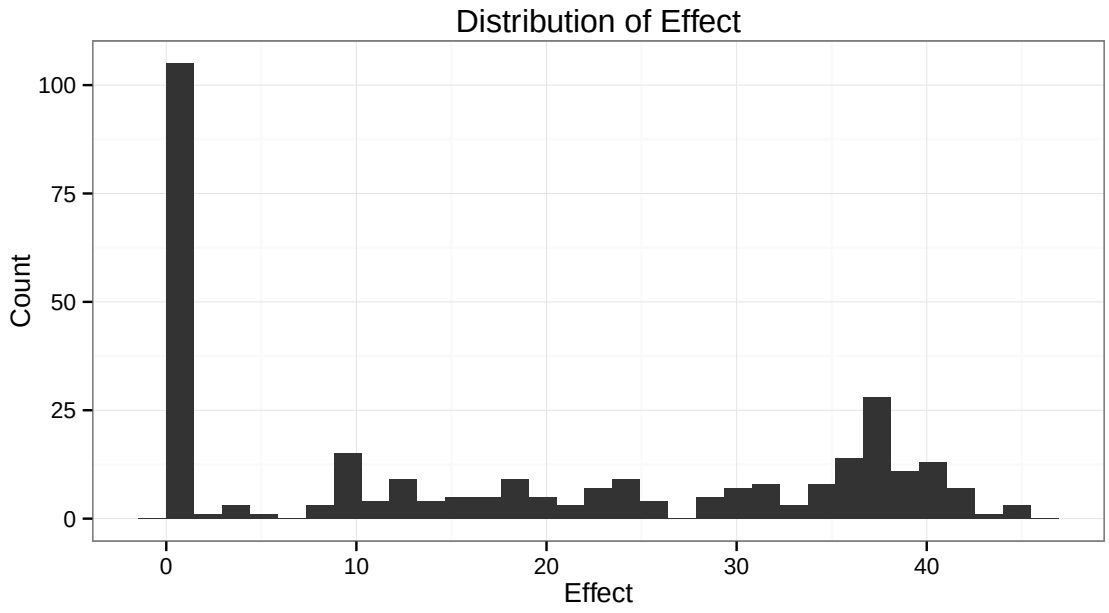
A few comments that may be of interest to the investigators. First, while Material 1 is cheap and common it also appears to be an effective choice. Interestingly, Material 6 never ended up being as effective as Material 1 despite their similar chemical properties. It is hard to know what to think of Materials 4 and 5. These materials are created through combining Material 1 with one of the other (more expensive) materials. Considering that they were not significantly better than Material 1, it is hard to imagine that they would be preferred over the cheaper option. In terms of medicine, Medicine 5 is both effective and versatile as it is one of the most effective medications on either rabbit species. While not as conclusive, the results I found in my smaller analysis of Medicine 3 data is consistent with these results.

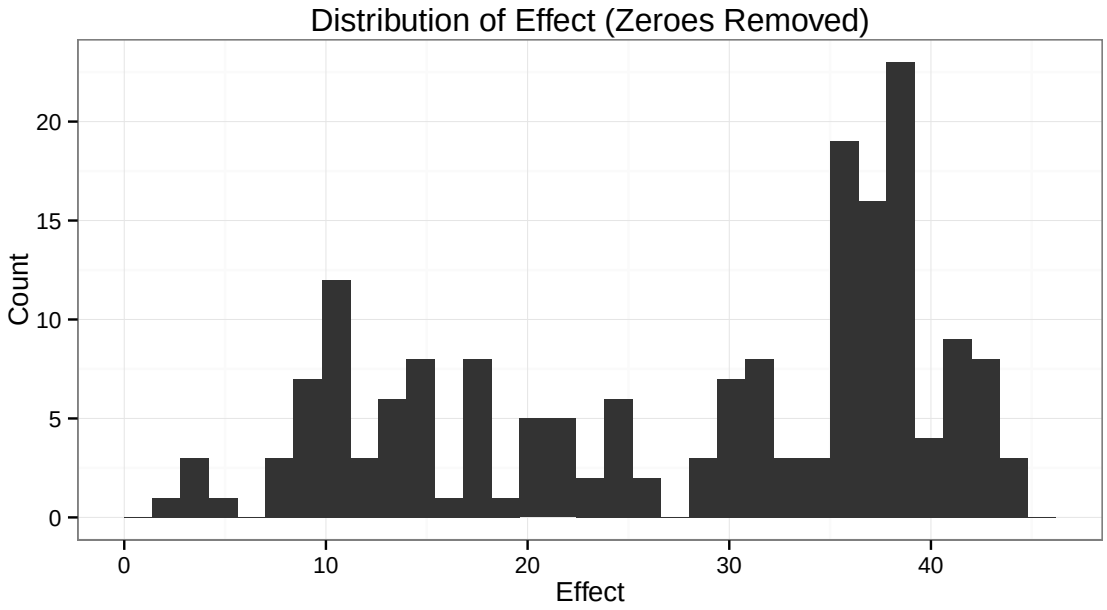
Were I the investigators and interested in the best and most commercially viable treatment for this parasite, I would recommend Medicine 5 in combination with Material 1. This combination has a few advantages. First, it is cheap and commonly available. Second, it was consistently one of the most effective treatment combinations. And finally, unlike some other treatment combinations it is one of the most effective choices for both rabbit species. It is important to keep in mind the limitations of the data and the steps I've had to take to work around them. That being said, I am still confident in my results and the method I used to find them. To confirm these results, the researchers might consider a smaller study focusing on less treatment combinations that utilizes my suggestions to generate a more robust dataset.

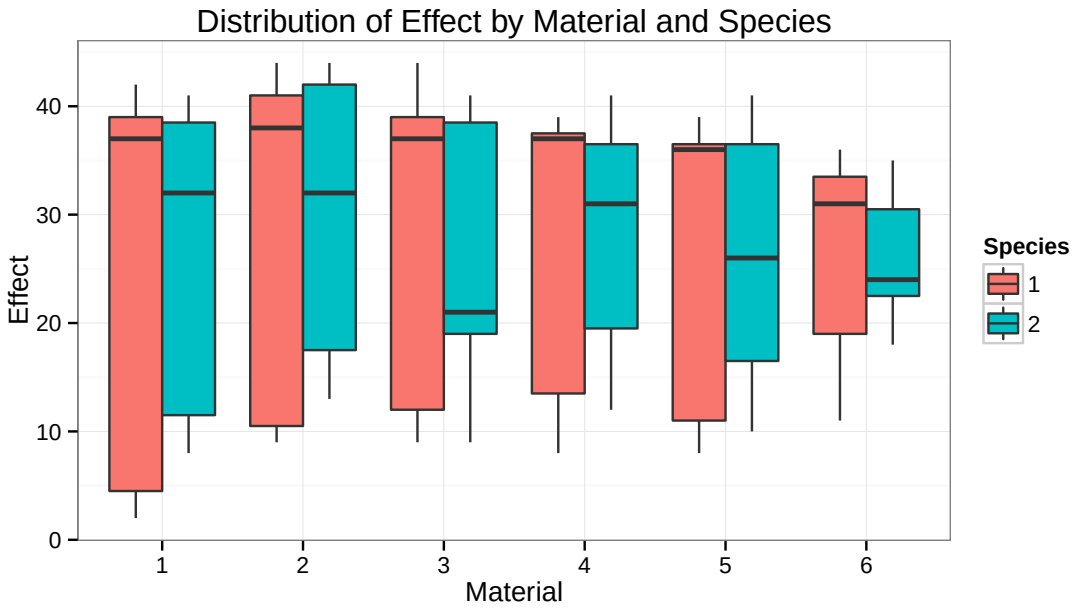
Appendix - Tables and Charts

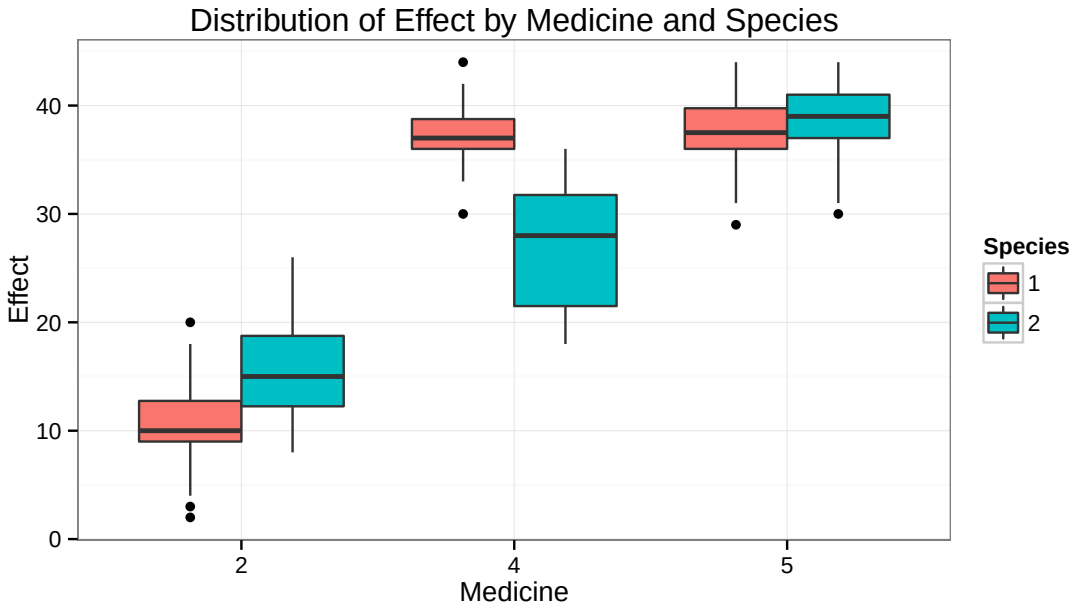
Summary of Effect Variable

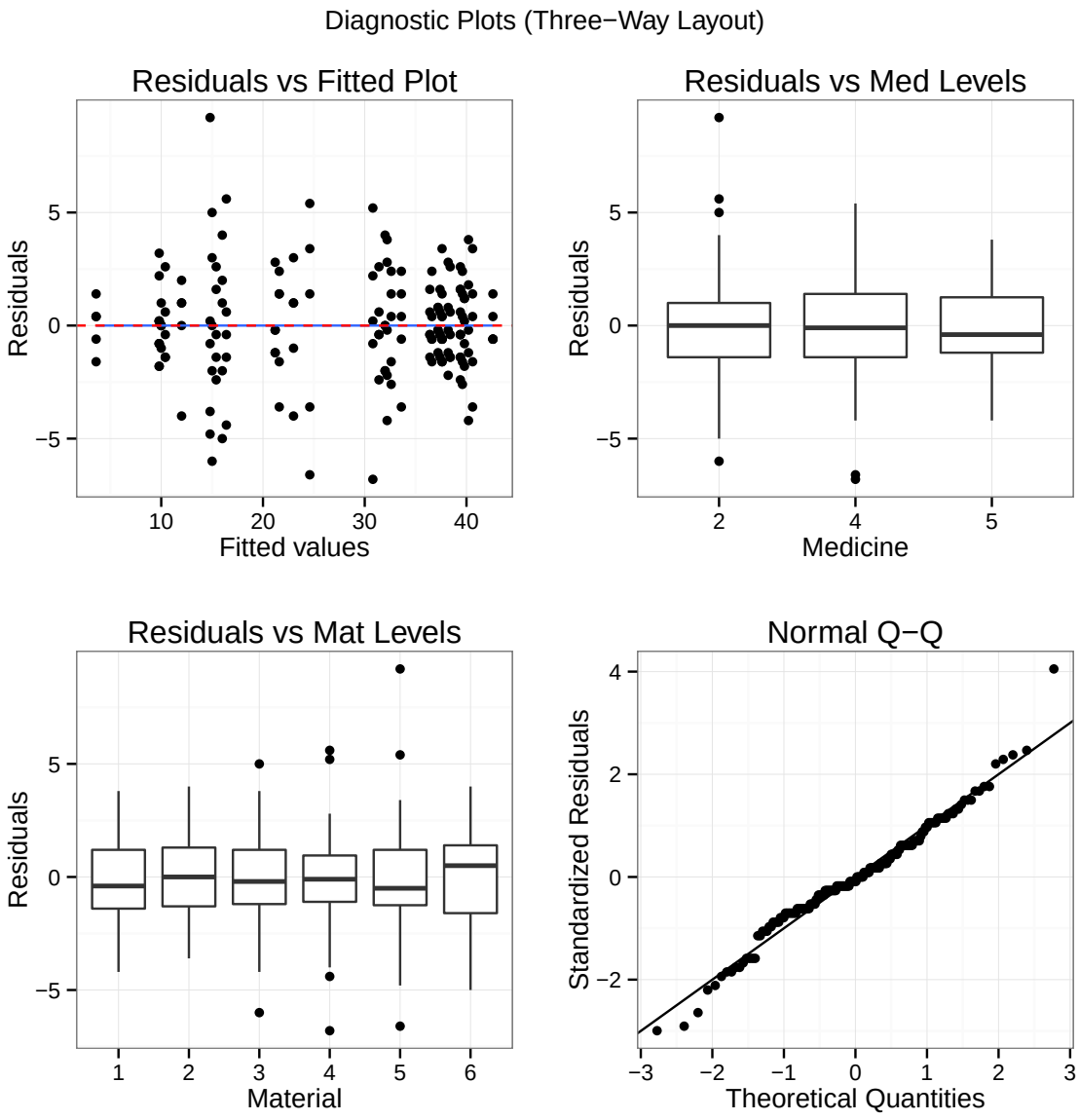
Statistic	N	Mean	St. Dev.	Min	Max
Effect	300	17.743	16.046	0	44




Distribution of Effect (Zeroes Removed)
Count
Effect
0
5
10
15
20
0
10
20
30
40


Distribution of Effect by Material and Species
Effect
Material
Species
1
2
0
10
20
30
40
1
2
3
4
5
6


Distribution of Effect by Medicine and Species
Effect
0
10
20
30
40
2
4
5
Medicine
Species
1
2



Two-way ANOVA (Species 1)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Material	5	122.72	24.54	6.48	0.0001
Medicine	2	14652.16	7326.08	1933.57	0.0000
Material:Medicine	10	680.64	68.06	17.96	0.0000
Residuals	72	272.80	3.79		

Two-way ANOVA (Species 2)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Material	5	259.66	51.93	5.71	0.0002
Medicine	2	7684.02	3842.01	422.20	0.0000
Material:Medicine	10	1139.84	113.98	12.53	0.0000
Residuals	72	655.20	9.10		

Grouped Tukey HSD (Species 1)

Material	Medicine	prediction	Group
1	2	3.60	1
5	2	9.80	2
2	2	10.00	2
3	2	10.40	2
4	2	12.00	23
6	2	16.00	3
6	5	31.40	4
6	4	33.60	45
5	5	36.40	56
5	4	36.60	56
4	4	37.20	56
4	5	37.40	56
3	4	37.60	56
1	4	38.40	6
1	5	39.40	6
2	5	39.60	6
3	5	40.20	6
2	4	40.60	6

Grouped Tukey HSD (Species 2)

Material	Medicine	prediction	Group
1	2	9.80	1
5	2	14.80	12
3	2	15.00	12
2	2	15.40	12
4	2	16.40	123
3	4	21.20	234
6	4	21.60	234
6	2	23.00	34
5	4	24.60	45
4	4	30.80	56
2	4	32.00	67
1	4	32.20	67
6	5	32.60	678
5	5	37.60	6789
4	5	38.20	789
1	5	39.40	89
3	5	39.80	9
2	5	42.60	9

Grouped Tukey HSD (Species 1, Medicine 3 Analysis)

Material	Medicine	prediction	Group
1	2	3.60	1
6	2	16.00	2
6	3	18.00	2
1	3	24.80	3
6	5	31.40	4
6	4	33.60	4
1	4	38.40	5
1	5	39.40	5

Grouped Tukey HSD (Species 2, Medicine 3 Analysis)

Medicine	prediction	Group
4	21.60	1
2	23.00	1
3	24.60	1
5	32.60	2

Appendix - R Code

```
library(ggplot2)
library(dplyr)
library(stargazer)
library(xtable)
library(grid)
library(gridExtra)
library(lsmeans)

# Plot for diagnostics
diagplot <- function(model) {
  p1 <- ggplot(model, aes(.fitted, .resid)) + geom_point() +
    stat_smooth(method="loess", se = FALSE) +
    geom_hline(yintercept=0, col="red", linetype = "dashed") +
    xlab("Fitted values") +
    ylab("Residuals") +
    ggtitle("Residuals vs Fitted Plot") +
    theme_bw()

  p2 <- ggplot(model, aes(y=.resid, x=Medicine)) +
    geom_boxplot() +
    xlab("Medicine") +
    ylab("Residuals") +
    ggtitle("Residuals vs Med Levels") + theme_bw()

  p3 <- ggplot(model, aes(y=.resid, x=Material)) +
    geom_boxplot() +
    xlab("Material") +
    ylab("Residuals") +
    ggtitle("Residuals vs Mat Levels") + theme_bw()

  p4 <- ggplot(model, aes(qqnorm(.stdresid)[[1]], .stdresid)) +
    geom_point(na.rm = TRUE) +
    geom_abline(aes(qqline(.stdresid))) +
    xlab("Theoretical Quantities") +
    ylab("Standardized Residuals") +
    ggtitle("Normal Q-Q") + theme_bw()

  return(list(rvfPlot=p1, rvf1plot=p2, rvf2plot=p3, qqplot=p4))
}
```

```r
# Load data and fix some issues
infected <- read.table("infected.txt", header = TRUE,
                       colClasses = c("integer","factor","factor","factor",
                                      "integer")) %>%
  # Replace '12' with '0' as per Dr. Jaxk Reeves email
  mutate(Effect = ifelse(Sample == 217, 0, Effect)) %>%
  mutate(treated = ifelse(Effect!=0,TRUE,FALSE))

species1 <- infected %>%
  filter(Species == 1)

species2 <- infected %>%
  filter(Species == 2)

# Proportion table for zero values, species 1
species1Prop <- species1 %>%
  group_by(Material, Medicine) %>%
  summarise(prop = mean(treated))
print(xtable(xtabs(prop~Material+Medicine, data = species1Prop),
             digits = 0),
      floating=FALSE)

# Proportion table for zero values, species 2
species2Prop <- species2 %>%
  group_by(Material, Medicine) %>%
  summarise(prop = mean(treated))
print(xtable(xtabs(prop~Material+Medicine, data = species2Prop),
             digits = 0),
      floating=FALSE)

# Delete baseline and medicine 3
infectedClean <- infected %>%
  filter(!(Medicine %in% c(1,3))) %>%
  mutate(Medicine = droplevels(Medicine))

# Three-way ANOVA
infecmod <- lm(Effect~Medicine*Material*Species, data = infectedClean)
print(xtable(summary(aov(infecmod)),
             caption = c("Three-way ANOVA")), table.placement = "H",
      caption.placement = "top")
```

```r
# Plot second-order interaction for each species
species1clean <- species1 %>%
  filter(!(Medicine %in% c(1,3)))

species2clean <- species2 %>%
  filter(!(Medicine %in% c(1,3)))
intp1 <- ggplot(species1clean, aes(x=Material, y = Effect,
                                   group = Medicine,
                                   color = Medicine,
                                   lty = Medicine)) +
  stat_summary(fun.y = mean, geom = "line") +
  ggtitle("Material-Medicine (Species 1)") +
  ylim(0,50) + theme_bw() + theme(legend.position = "top")
intp2 <- ggplot(species2clean, aes(x=Material, y = Effect,
                                   group = Medicine,
                                   color = Medicine,
                                   lty =Medicine)) +
  stat_summary(fun.y = mean, geom = "line") +
  ggtitle("Material-Medicine (Species 2)") +
  ylim(0,50) + theme_bw() + theme(legend.position = "top")
intplots <- list(intp1, intp2)
do.call(grid.arrange, c(intplots, ncol = 2))

# ANOVA for Species 1 two-way
spec1 <- lm(Effect~Material*Medicine, data = species1clean)
print(xtable(summary(aov(spec1)),
             caption = c("Two-way ANOVA (Species 1)")),
      table.placement = "H",
      caption.placement = "top")

# ANOVA for Species 2 two-way
spec2 <- lm(Effect~Material*Medicine, data = species2clean)
print(xtable(summary(aov(spec2)),
             caption = c("Two-way ANOVA (Species 2)")),
      table.placement = "H",
      caption.placement = "top")

# Summary statistics for response
stargazer(infected, keep = c(5), title = c("Summary of Effect Variable"),
          table.placement = "H")

# Distribution of response (pre-clean)
```

```
qplot(infected$Effect, main = "Distribution of Effect", xlab = "Effect",
      ylab = "Count") + theme_bw()

# Distribution of response (clean)
qplot(infectedClean$Effect, main = "Distribution of Effect (Zeroes Removed)",
      xlab = "Effect",
      ylab = "Count") + theme_bw()

# Boxplots showing response by material and species (clean)
ggplot(infectedClean, aes(y= Effect, x = Material)) +
  geom_boxplot(aes(fill = Species)) +
  ggtitle("Distribution of Effect by Material and Species") + theme_bw()

# Boxplots showing response by medicine and species (clean)
ggplot(infectedClean, aes(y= Effect, x = Medicine)) +
  geom_boxplot(aes(fill = Species)) +
  ggtitle("Distribution of Effect by Medicine and Species") + theme_bw()

# Diagnostics for three-way ANOVA
spec1diag <- diagplot(infecmod)
do.call(grid.arrange, c(spec1diag,
                        main = "Diagnostic Plots (Three-Way Layout)",
                        ncol = 2))

# Tukey for Species 1
refgrid1 <- ref.grid(spec1)
cld1 <- cld(refgrid1) %>%
  select(Material, Medicine, prediction, .group) %>%
  rename(Group = .group)
print(xtable(cld1,
             caption = c("Grouped Tukey HSD (Species 1)")),
      table.placement = 'H',
      caption.placement = "top",
      include.rownames = FALSE)

# Tukey for Species 2
refgrid2 <- ref.grid(spec2)
cld2 <- cld(refgrid2) %>%
  select(Material, Medicine, prediction, .group) %>%
  rename(Group = .group)
print(xtable(cld2,
             caption = c("Grouped Tukey HSD (Species 2)")),
```

```
      table.placement = 'H',
      caption.placement = "top",
      include.rownames = FALSE)

# Subanalysis of medicine 3 for both species
infectedmed3 <- infected %>%
  filter(Medicine != 1) %>%
  filter(!(Material %in% c(2,3,4,5)))

# This can be two-way ANOVA on Medicines 2-5 and Materials 1 and 6
spec1med3 <- infectedmed3 %>%
  filter(Species == 1)

spec1mod3 <- lm(Effect~Material*Medicine, data = spec1med3)
#summary(aov(spec1mod3))

refgridmod31 <- ref.grid(spec1mod3)
cld3 <- cld(refgridmod31) %>%
  select(Material, Medicine, prediction, .group) %>%
  rename(Group = .group)
print(xtable(cld3,
             caption = c("Grouped Tukey HSD (Species 1, Medicine 3 Analysis)")),
      table.placement = 'H',
      caption.placement = "top",
      include.rownames = FALSE)

# This will just be one-way ANOVA on Medicines 2-5
spec2med3 <- infectedmed3 %>%
  filter(Species == 2) %>%
  filter(Material != 1)

spec2mod3 <- lm(Effect~Medicine, data = spec2med3)
#summary(aov(spec1mod3))

refgridmod32 <- ref.grid(spec2mod3)
cld4 <- cld(refgridmod32) %>%
  select(Medicine, prediction, .group) %>%
  rename(Group = .group)
print(xtable(cld4,
             caption = c("Grouped Tukey HSD (Species 2, Medicine 3 Analysis)")),
      table.placement = 'H',
      caption.placement = "top",
```

```
      include.rownames = FALSE)
```