# Report to investigate the prevalence of parasite infections in cattle and the efficacy of three anthelmintics

Student 202078239

## 1 Introduction

The health impact and economic losses of cattle from worm parasite infections is a significant issue in the farming industry. Worm eggs are passed in manure and hatch producing larvae that contaminate the pastures that cattle graze. Conditions on farms vary, so both pasture management and targeted anthelmintic treatment are required to address the increasing concern over anthelmintic resistance.[1] There are seasonal variations in parasite levels, with egg counts typically dropping in the autumn as larvae enter a state of hypobiosis, when larvae development pauses. Typically little to no transmission occurs when cattle are housed over the winter.[2] Despite anthelmintics having been used in livestock since the 1960s, it is still unclear as to what threshold level of parasite infection requires treatment.[2] This report investigates the prevalence of parasite infections in cattle on a particular farm and the effect of treating them with three different anthelmintics.

## 2 Method

The anthelmintics used in this study are fenbendazole (FBZ), ivermectin (IVM) and doramectin (DOR). All are broad spectrum treatments, active against most species of roundworms and lungworms. FBZ is from a group called benzimidazoles and most, including FBZ are active against tapeworms. DOR is a derivative of IVM and both are from a group called macrocyclic lactones (ML). ML treatments are active against ectoparasites but do not work against tapeworms.[3] Three groups of 31 cattle were randomly assigned to each anthelmintic treatment group. Dosages were based on the individual animal's weight, measured via weight tape, which is considered to be fairly accurate.[4] Faecal samples were taken from the cattle on day 0, before treatment had been administered and at 14 and 21 days post treatment. The groups of cattle were grazed on different fields on the farm to avoid cross contamination, however there is no information regarding parasite levels in the fields prior to the study. The response variable in the study is faecal egg counts (FECs), recorded in eggs per gram of faeces (epg), using microscopic methods. The exact preparation for examination and egg counting is unknown but in previous studies, using a double centrifuge method, a recovery rate of 62.5% was expected and hence a correction factor of 1.6 was used to account for egg loss.[2] Without further information on the preparation of the FECs we cannot make any appropriate adjustment to the data, but the potential for egg loss during collection and counting should be noted.

Data recorded from the study included the FECs, treatment type, days after treatment that the FECs were collected on and animal ID. I note that there is a missing data point for the day 0 measurement of subject number 33 in the DOR treatment group. It is possible that the faecal sample size was not large enough. This means that the standard errors will not be the same for each group so a linear model that can deal with unbalanced data must be used. The prevalence of parasite infections will be investigated using data taken on day 0, before treatment. A model will then be fitted to include the main factors of interest: A main interest in the experiment is time after treatment, which we must include as a fixed factor. Another key interest is the effect of the treatments, so treatment group must also be a fixed factor in our statistical analysis. The primary interest in the study is the interaction of time and treatment to evaluate which anthelmintics work best over time. We are not interested in how the treatments affect particular individuals, but we include Subject ID as a random effect nested in treatment to calculate how much variation is associated with repeated measurements and how much is associated with differences between the cattle. A repeated measures design (RMD) is therefore used with a mixed effects model. A RMD is more efficient than when individual subjects are considered at

each time point. Subjects can show large variations, so using the same individual at each time point improves the time effect estimates and a RMD effectively incorporates individual subject as a blocking factor. The linear effects model for the measured response, $X_{ijk}$ is

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + \epsilon_{ijk} \tag{1}$$

where $\mu$ is the grand mean, assumed constant for all treatment combinations tested, $\alpha_i$ is the main effect of level $i$ of factor A (Treatment), $\beta_{j(i)}$ is the effect of level $j$ of factor B (Animal ID) nested within level $i$ of factor A, $\gamma_k$ is the main effect of level $k$ of factor C (Days), $(\alpha\gamma)_{ik}$ is the interaction effects of A and C and $\epsilon_{ijk}$ are the error terms.

Any fixed effects, including the interaction term, must sum up to zero and error terms must be independent and normally distributed. It is also assumed that errors within a subject are equally correlated and have equal variance representing the compound symmetry property.

Statistical analyses will be implemented using R code. Summary statistics will be produced and described with the aid of boxplots. A fixed effects model will be used initially to check that the model assumptions are valid. The random effect will then be incorporated into a linear mixed effects model and results from hypothesis testing analysed using lme() from the nlme library. Since DOR and IVM are chemically similar drugs and FBZ is different, the lme function's default constraint will be used, which sets the first treatment effect to be zero and the code will be set up to ensure that FBZ is the first treatment level. This will make it easier to compare the groups of anthelmintics.

Students's t-tests will be used to evaluate whether calculated effects are statistically equal to zero (null hypothesis), or if any of the effects are different from zero (alternative hypothesis). Similarly, Fisher's F-test will be used to calculate the F-ratios, defined as the ratio of the mean square effect and the mean square error. If the null hypothesis is true, all effects equal zero, which means that the mean square of the effect and the mean square of the error are estimates of the same variance, $\sigma^2$. Their ratio can therefore be used to compare the estimates of variance. Under the null hypothesis, the ratio follows an F distribution on the factor effect's degrees of freedom and the error degrees of freedom. The F-value calculated for the model is compared with the value from the F-distribution. For both tests a significance value of 5% is used.

## 3 Results

The summary statistics in table 1 show that the missing data point is for the DOR group at 0 days. The FECs range from 2 to 290epg, with a median of 79.5epg. The $\sqrt{FECs}$ values are taken for the boxplots as the variations are very small and difficult to see. However, unless transformation of FECs is required to validate the model assumptions, analysis will continue with untransformed data for ease of interpretation.

Table 1: Summary statistics

| Treatment | AnimalID | Day | FECs | cow |
|---|---|---|---|---|
| FBZ:93 | 1 : 3 | 0 :92 | Min. : 2.00 | FBZ-1 : 3 |
| DOR:92 | 2 : 3 | 14:93 | 1st Qu.: 17.25 | FBZ-2 : 3 |
| IVM:93 | 3 : 3 | 21:93 | Median : 79.50 | FBZ-3 : 3 |
| | 4 : 3 | | Mean : 99.33 | FBZ-4 : 3 |
| | 5 : 3 | | 3rd Qu.:179.50 | FBZ-5 : 3 |
| | 6 : 3 | | Max. :290.00 | FBZ-6 : 3 |
| | (Other):260 | | | (Other):260 |

Figure 1 shows the $\sqrt{FECs}$ values for each treatment at Day 0, 14 and 21. The range of values for each Treatment:Day group is very small suggesting that all cattle within a group are affected similarly by parasite infections and by the treatment given. Looking at the first three boxplots in the graph it appears that cattle in the IVM group have the highest levels of FECs before starting treatment, followed by the DOR group and the FBZ group have the lowest levels

of FECs pre-treatment. FBZ appears to reduce the FECs most successfully, keeping FECs low at day 21. DOR reduces the FECs the least and numbers rise again by day 21. IVM reduces FECs by the largest number between day 0 and 14, however FECs increase to the highest number by day 21, showing good short term effect only. There are a few outlying data points shown as circles on the plot, representing subjects with unusually high or low FECs compared to the rest of their group. An example outlier is the minimum FECs measured of 2epg found in the FBZ group.

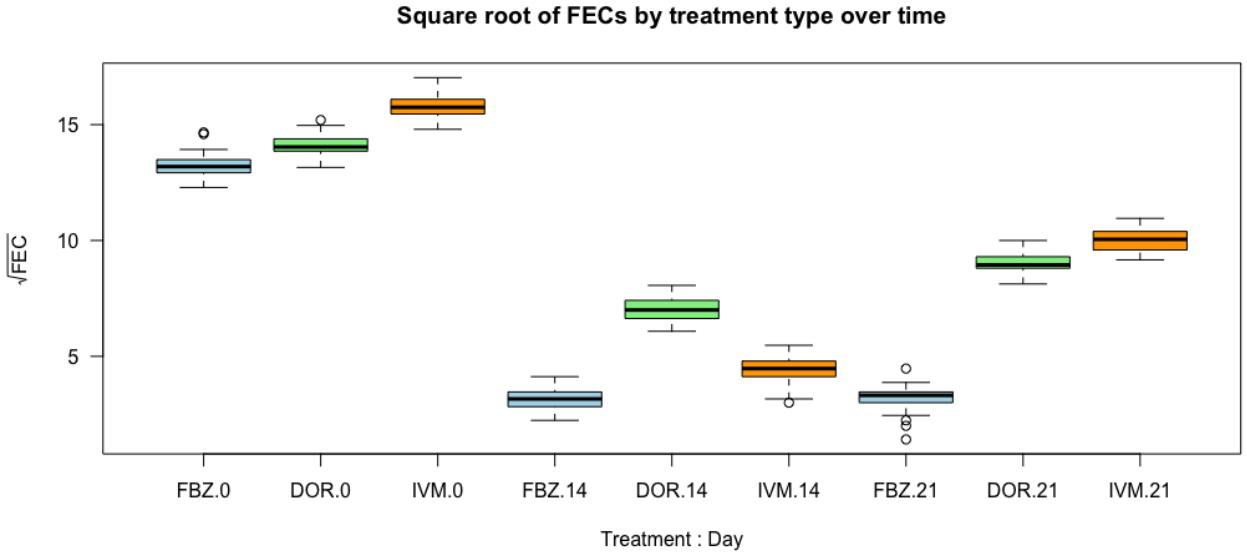


Figure 1: Boxplots showing the mean $\sqrt{FEC}$ levels in the different Treatment:Day groups

Next, we fit a fixed linear model to check model assumptions. The top three plots of figure 2 show that the residuals are normally distributed, with a mean of zero. A small number of outliers can be seen in the QQ-plot but the rest of the residuals follow the line very closely. The versus fits plot shows the residuals are well scattered with a mean of zero and the time series plot shows no pattern that might suggest temporal dependency of the residuals. All model assumptions are considered to be valid so we now add the random factor of Subject ID into the model for further analysis.

Table 2 shows the ANOVA table for the linear mixed effects model. The interaction between Treatment and Day is a 'within subject' test with $F = 220.637$ and a p-value of $< 0.0001$. This is below the 5% significance level and suggests that this interaction is highly significant on the resulting FECs. The main effects of Treatment and Day are also significant with p-values of $< 0.0001$, but the interaction effect is the most important to analyse so an interaction plot is produced and shown in figure 3.

Table 2: ANOVA table for the linear mixed effects model

| | numDF | denDF | F-value | p-value |
|---|---|---|---|---|
| (Intercept) | 1 | 179 | 23462.954 | <.0001 |
| Treatment | 2 | 90 | 727.823 | <.0001 |
| Day | 2 | 179 | 8777.763 | <.0001 |
| Treatment:Day | 4 | 179 | 220.637 | <.0001 |

Figure 3 shows that on day 0 the FECs for the treatment groups are roughly 175, 200 and 250 epg showing significant variation in the baseline counts pre-treatment, so any effect of treatment should take this into account. This highlights that prevalence of worm parasites differs between herds and that it isn't necessarily appropriate to treat all herds with the same dosage or anthelmintic type. Lack of parallel lines in an interaction plot represent that levels

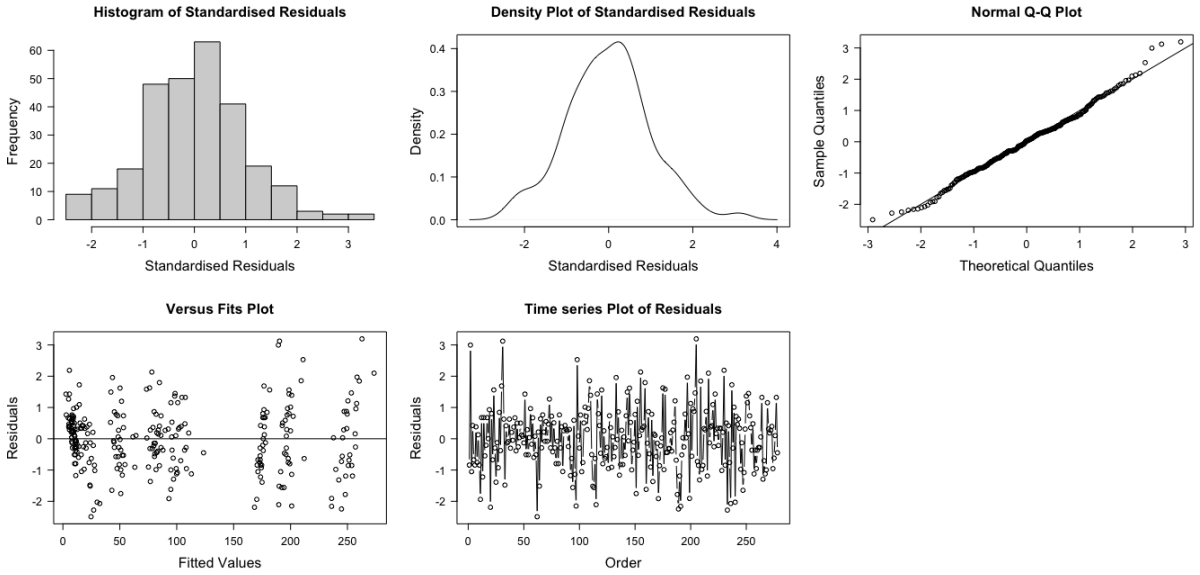


Figure 2: Residual plots to determine whether model assumptions are valid

of one factor do not have the same effect across the range of levels of the second factor. From day 0 to 14, IVM reduces its mean FECs at the highest rate of all treatment groups, followed by FBZ which reduces FECs at a slightly slower rate. The slope for DOR between days 0 and 14 is much less steep showing that DOR is reducing FECs at the slowest rate. The interaction effect is greatest at day 21 where the biggest effect is seen between FBZ and IVM. At 21 days IVM now has the highest level of FECs and is more similar to DOR which is also high. FBZ is the only treatment to maintain FECs at low levels at 21 days. The treatment that appears to work best over 21 days is FBZ, however if FECs must be reduced urgently, IVM may be the best short term treatment. The interaction effects are quantitative, meaning the direction of the effect is the same for all treatments and only the magnitude of the effects differ. Testing main effects with a quantitative interaction can sometimes be helpful, but it is tricky to isolate the intrinsic interaction from analysis of the main effects, leading to results that may not be entirely representative.

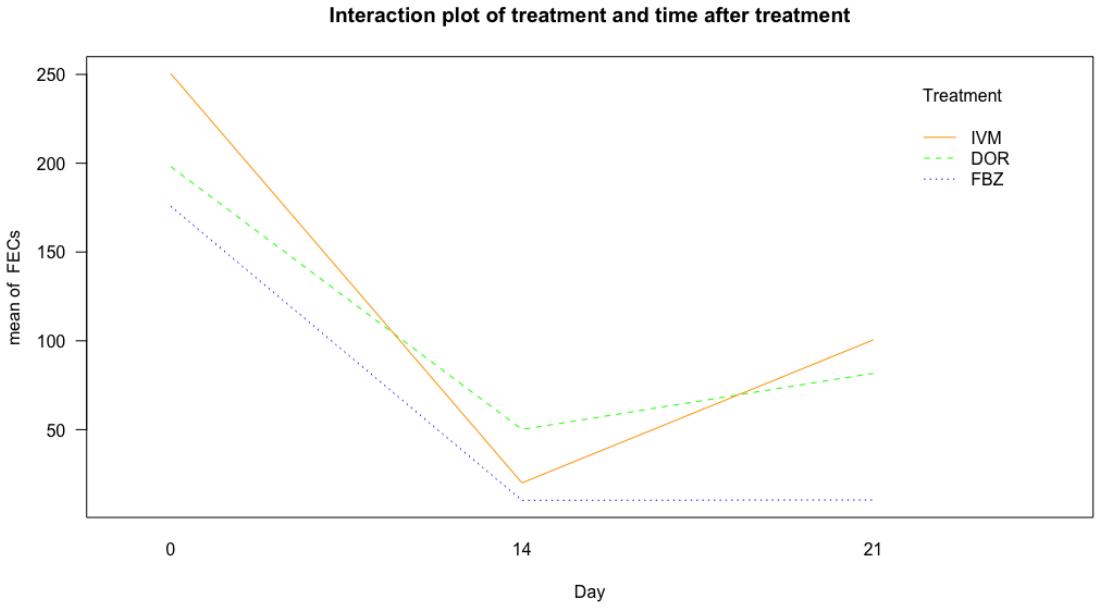


Figure 3: Graph to show how the interaction of treatment and time affects mean FECs

The values of the mean FECs that we may be interested in from the interaction plot can be found using the emmeans function from the emmeans library. It calculates TukeyHSD pairwise comparisons but accounts for missing values or unbalanced data. TukeyHSD uses a modification of the Bonferroni adjustment which accounts for the different p-values that should be used when doing simultaneous hypothesis tests. In Table 3 we can see the FEC means and their associated 95% confidence intervals for each treatment group by each day and the pairwise contrasts of the treatments by day. For example, we can see that the mean FBZ FECs reduces from 175.8epg on day 0 to 10.2epg on day 14 and increases slightly to 10.4epg on day 21. The confidence interval shows that the value of 175.8epg may be as low as 172.15 or as high as 179.4. Percentage of FECs post treatments can be seen in appendix A for interest.

The contrasts show the mean differences of FECs between the treatment pairs, according to each day. For example FBZ – DOR on day 14 gives a difference of -39.94epg. This is negative as FBZ has a lower FEC than DOR on day 14.

The emmeans function was used again for pairwise comparisons of time according to treatment group. The table can be found in appendix B and the point of interest is that all p-values are $< 0.05$, the significance level, indicating statistically non-zero effects, except for the comparison of days 14 and 21 for FBZ, which has a p-value of 0.9941, indicating that there is no significant difference in mean values of FECs for FBZ between days 14 and 21.

Table 3: Mean FECs and pairwise comparisons

```
$emmeans
Day = 0:
 Treatment emmean   SE df lower.CL upper.CL
 FBZ        175.8 1.83 92   172.15    179.4
 DOR        198.0 1.86 90   194.36    201.7
 IVM        250.4 1.83 90   246.76    254.0

Day = 14:
 Treatment emmean   SE df lower.CL upper.CL
 FBZ         10.2 1.83 92     6.53     13.8
 DOR         50.1 1.83 90    46.47     53.7
 IVM         20.0 1.83 90    16.37     23.6

Day = 21:
 Treatment emmean   SE df lower.CL upper.CL
 FBZ         10.4 1.83 92     6.79     14.0
 DOR         81.7 1.83 90    78.08     85.3
 IVM        100.5 1.83 90    96.89    104.1

Degrees-of-freedom method: containment
Confidence level used: 0.95
```

```
$contrasts
Day = 0:
 contrast  estimate   SE df t.ratio p.value
 FBZ - DOR   -22.27 2.60 90  -8.553 <.0001
 FBZ - IVM   -74.61 2.58 90 -28.890 <.0001
 DOR - IVM   -52.34 2.60 90 -20.101 <.0001

Day = 14:
 contrast  estimate   SE df t.ratio p.value
 FBZ - DOR   -39.94 2.58 90 -15.463 <.0001
 FBZ - IVM    -9.84 2.58 90  -3.810 0.0007
 DOR - IVM    30.10 2.58 90  11.653 <.0001

Day = 21:
 contrast  estimate   SE df t.ratio p.value
 FBZ - DOR   -71.29 2.58 90 -27.604 <.0001
 FBZ - IVM   -90.10 2.58 90 -34.885 <.0001
 DOR - IVM   -18.81 2.58 90  -7.282 <.0001

Degrees-of-freedom method: containment
P value adjustment: tukey method for comparing
a family of 3 estimates
```

Having analysed the fixed effects, it is useful to look at how much variation is associated with differences between the cattle. It was mentioned earlier that the spread shown in the boxplots for each Treatment:Day interaction group is very small, indicating that between cattle variation is very low. The calculated variation of the random effects can be found in the summary of the lme function, however the summary output shown for the same linear mixed model using the lmer function from the lme4 library displays this variance in a clearer format and is shown in table 4. This shows the variance between the cattle to be 6.761epg. The residual variance of 96.625epg comes from the 'within cattle' variance due to the repeated measurements. The total random variance is the sum of these two values, 103.386epg. The proportion of this total random variance associated with 'between cattle' variance is found to be 6.761/103.386=0.0654, about 6.5%. Since most, 93.5% of the random variance is associated with 'within cattle' variance, then we do not need to worry too much about 'between cattle' variation and can therefore save resources in future experiments.

Table 4: Random Effects

```
Random effects:
 Groups   Name        Variance Std.Dev.
 cow      (Intercept)  6.761   2.60
 Residual             96.625   9.83
Number of obs: 278, groups:  cow, 93
```

The coefficient of determination, $R^2$ measures the 'goodness of fit' of the model and is calculated to be 0.988. This means that 98.8% of the variability in the predicted FECs is explained by the model indicating that the model works very well.

# 4 Conclusion

The data analysed in this report indicates that there are significant differences in the prevalence of parasites in different herds of cattle. Our data also points to FBZ being the preferred treatment to use. The effect of time after treatment has been found to be highly significant with FECs reducing significantly 14 days after treatment and FECs rising again between 14 and 21 days. A desirable anthelmintic treatment not only needs to reduce FECs quickly, but also maintain FECs at low levels for as long as possible. A longer study would be useful to properly investigate the residual activity of FBZ. The study found very little variation between the cattle and that most random variation is due to the repeated observations.

The statistical part of the study was set up to compare FBZ with DOR and IVM. The main difference between these two groups is that whilst FBZ maintains low levels of FECs at 21 days, the ML treatment groups have increased levels of FECs, seen clearly in the large difference between the treatment means at 21 days. Compared to DOR, IVM was found to have a greater effect in decreasing FECs up to day 14, however, its residual activity dropped off quicker than DOR between days 14 and 21.

Previous studies have shown ML treatments to have a *longer* duration of activity and have highlighted that FBZ does *not* have good residual activity.[2] Whilst this appears to contradict our results, these studies are in strong agreement with treatment withdrawal times for meat and milk entering the food chain. These withdrawal times are commensurate with resistance activity of the anthelmintics.[3] Why does our data seemingly contradict these results?

An important consideration in choice of anthelmintic is the type of parasites that are infecting the cattle. FBZ targets tapeworms, whilst DOR and IVM do not, so if tapeworm was a major contributor to the parasite infections of these cattle, it would explain why FBZ appears to have been the most successful treatment. However, this doesn't explain the effect of IVM treatment which until day 14 works very well. Perhaps the initial potency kills more types of parasites than the drug is marketed for. Different anthelmintics may treat different stages of parasites more or less effectively than others too. A limitation to our experiment is the unknown loss of FECs prior to counting which may be significant, although one would expect the loss to be consistent over the treatment groups. There is also uncertainty as to whether FECs are representative of worm counts within an animal, so other measures of parasite infection could be investigated. Application of treatment is another variable that affects how long a drug is active for, with injections typically lasting longer.[3] No information is given regarding the method of treatment application in this study. Since parasite activity in grazing pastures is a source of infection, it may be considered useful to include the field in which the cattle are grazing as a random factor in the model. Following the results of this study, coproculture methods should be used to determine what parasites are prevalent in an area, so targeted treatment can be used. Including a placebo treatment in a future study would also help to monitor natural fluctuations.

On a final note, with parasites' growing resistance to anthelmintic drugs, perhaps it would be beneficial to turn our attention to studying better pasture management strategies as a more integral part of parasite management in cattle.

## References


[1] Farmer's Guardian. *Pasture management as a parasite control strategy.* URL: https://www.fginsight.com/beattheparasites/sponsored-beat-the-parasites-2019-articles/pasture-management-as-a-parasite-control-strategy-85622.

[2] Kaley G. Mackie et al. "Efficacy of fenbendazole and ivermectin in treating gastrointestinal nematode infections in an Ontario cow-calf herd". *The Canadian Veterinary Journal* 60.11 (2019), pp. 1213–1219.

[3] *Cattle Parasite Control Guide - A comprehensive list of products for the control of internal and external parasites of cattle.* 2014. URL: https://www.cattleparasites.org.uk/app/uploads/2018/04/COWS_cattle_parasite_control_guide.pdf.

[4] Kesang Wangchuk, Jigme Wangdi and Mindu Mindu. "Comparison and reliability of techniques toestimate live cattle body weight". *Journal of Applied Animal Research* 46.1 (2018), pp. 349–352.


## Appendix A: Percentage of original FEC values after treatment

| | Day 0 | Day 14 | Day 21 |
|---|---|---|---|
| FBZ | 100% | 6% | 6% |
| DOR | 100% | 25% | 41% |
| IVM | 100% | 8% | 40% |

## Appendix B: Pairwise comparisons of time according to treatment group

```
$contrasts
Treatment = FBZ:
 contrast estimate   SE  df t.ratio p.value
 0 - 14    165.613 2.50 179  66.331 <.0001
 0 - 21    165.355 2.50 179  66.228 <.0001
 14 - 21    -0.258 2.50 179  -0.103 0.9941
```

## Appendix C: R code used for analysis

```
library(lme4)
library(nlme)
library(emmeans)

datafull <- read.csv("ExpDesign-Project-Data.csv", header=T)
data <- subset(datafull, Student!=202078239 | Student %in% c(NA)) #My data
data <- data[,1:4]
summary(data)

#Organise data appropriately
#3 treatments: Put FBZ first as DOR and IVM are made similarly
data$Treatment <- factor(data$Treatment, levels = c("FBZ", "DOR", "IVM"))
data$AnimalID <- factor(data$AnimalID)  #31 cows/treatment
data$Day <- factor(data$Day)  #3 times to collect data
data$cow <- as.factor(paste(data$Treatment, data$AnimalID, sep="-"))
data$cow <- factor(data$cow, levels=c("FBZ-1", "FBZ-2", "FBZ-3", "FBZ-4", "FBZ-5",
                                      "FBZ-6", "FBZ-7", "FBZ-8", "FBZ-9", "FBZ-10",
                                      "FBZ-11", "FBZ-12", "FBZ-13", "FBZ-14", "FBZ-15",
                                      "FBZ-16", "FBZ-17", "FBZ-18", "FBZ-19", "FBZ-20",
                                      "FBZ-21", "FBZ-22", "FBZ-23", "FBZ-24", "FBZ-25",
                                      "FBZ-26", "FBZ-27", "FBZ-28", "FBZ-29", "FBZ-30",
                                      "FBZ-31",
```

```
                                    "DOR-32", "DOR-33", "DOR-34", "DOR-35", "DOR-36", "DOR-37",
                                    "DOR-38", "DOR-39", "DOR-40", "DOR-41", "DOR-42", "DOR-43",
                                    "DOR-44", "DOR-45", "DOR-46", "DOR-47", "DOR-48", "DOR-49",
                                    "DOR-50", "DOR-51", "DOR-52", "DOR-53", "DOR-54", "DOR-55",
                                    "DOR-56", "DOR-57", "DOR-58", "DOR-59", "DOR-60", "DOR-61",
                                    "DOR-62",

                                    "IVM-63", "IVM-64", "IVM-65", "IVM-66", "IVM-67", "IVM-68",
                                    "IVM-69", "IVM-70", "IVM-71", "IVM-72", "IVM-73", "IVM-74",
                                    "IVM-75", "IVM-76", "IVM-77", "IVM-78", "IVM-79", "IVM-80",
                                    "IVM-81", "IVM-82", "IVM-83", "IVM-84", "IVM-85", "IVM-86",
                                    "IVM-87", "IVM-88", "IVM-89", "IVM-90", "IVM-91", "IVM-92",
                                    "IVM-93"))
summary(data)

#boxplots to visualise data
par(mfrow=c(1,1))
boxplot(sqrt(FECs)~Treatment:Day, data, col=(c("lightblue","lightgreen", "orange")), las=1,
        main="Square root of FECs by treatment type over time", ylab=expression(sqrt(FEC)))

#Fit fixed effects model to check model assumptions
z.lm <- lm(FECs~Treatment + cow + Day + Treatment:Day, data)

### Residual plots ###
par(mfrow=c(2,3))
stan_res <- (residuals(z.lm)-mean(residuals(z.lm)))/sd(residuals(z.lm))
#Histogram of resids
hist(stan_res, main = "Histogram of Standardised Residuals", xlab="Standardised Residuals",
     ylab = "Frequency", cex.lab=1.5, cex.main=1.5, cex.axis=1.2, las=1)
#Density plot of resids
plot(density(stan_res), main = "Density Plot of Standardised Residuals",
     ylab="Density", xlab="Standardised Residuals", cex.lab=1.5, cex.main=1.5, cex.axis=1.2, las=1)
#QQ plot
qqnorm(stan_res, cex.lab=1.5, cex.main=1.5, cex.axis=1.2, las=1)
abline(a=0, b=1)  #and its line
#Versus fit of resids
plot(stan_res ~ fitted(z.lm), main="Versus Fits Plot", xlab="Fitted Values",
     ylab="Residuals", cex.lab=1.5, cex.main=1.5, cex.axis=1.2, las=1)
abline(a=0, b=0)
#Temporal Plot - needed as observations taken over time
plot(stan_res, type="b", main = "Time series Plot of Residuals",
     ylab="Residuals", xlab="Order", cex.lab=1.5, cex.main=1.5, cex.axis=1.2, las=1)

#Missing data point so unbalanced --> use lmer/lme
z.lmer <- lmer(FECs ~ Treatment*Day + (1|cow), data)
summary(z.lmer)
anova(z.lmer)

#Can look at lme output also
z.lme <- lme(FECs ~ Treatment + Day + Day:Treatment, random = ~ 1|cow, data, method = "REML")
summary(z.lme)
anova(z.lme)

#interaction plot
par(mfrow=c(1,1))
with(data, interaction.plot(Day, Treatment, FECs, las=1,
                            main="Interaction plot of treatment and time after treatment",
                            col = c("blue2", "green", "orange")))
```

```
#Means and pairwise info
emmeans(z.lme, pairwise~Treatment|Day, adjust="Tukey")
emmeans(z.lme, pairwise~Day|Treatment, adjust="Tukey")

#Calculate R^2
resid <- residuals(z.lme)  #Obtain residuals
mean_FECs <- mean(data$FECs)

#The percentage of variance explained by both fixed and random effects
(R2_both <- 1 - sum(resid^2)/sum((data$FECs - mean_FECs)^2))
```