Havi Fecundity Analysis

Contents

Load Libraries

```
pacman::p_load(ggplot2, readxl, ggbeeswarm, readr, dplyr, formatR, tidyr, tidyverse,
    devtools, cowplot, knitr, emmeans, ggpmisc, lme4, lmerTest, RColorBrewer, viridis,
    install = FALSE)
```

Load Data

```
aphid_count <- read_excel("data/fecundity_assay_count.xlsx") %>%
    filter(alate_alive_r1 != "control") %>%
    filter(block != 1) %>%
    mutate(nymphs_r1 = as.numeric(nymphs_r1), nymphs_r2 = as.numeric(nymphs_r2))
```

```
## Warning: There were 2 warnings in `mutate()`.
## The first warning was:
## i In argument: `nymphs_r1 = as.numeric(nymphs_r1)`.
## Caused by warning:
## ! NAs introduced by coercion
## i Run `dplyr::last_dplyr_warnings()` to see the 1 remaining warning.
```

```
# filter by block
aphid_count_block2 <- aphid_count %>%
    filter(block == 2)
aphid_count_block3 <- aphid_count %>%
    filter(block == 3)
```

```
# === Pivot to long format ===
aphid_long <- aphid_count %>%
```

```
    pivot_longer(cols = c(starts_with("nymphs_r"), starts_with("alate_alive_r")),
        names_to = c(".value", "round"), names_pattern = "(.*)_r(\\d)")
```

Survival

Survival Calculations

```
# === Survival summary by block and round ===
survival_summary <- aphid_long %>%
  group_by(treatment, block, round) %>%  # Group data by treatment condition, experimental block, and r
  summarise(
    # Count how many alates were alive in each group (handling NAs just in case)
    alive_n = sum(alate_alive == "Yes", na.rm = TRUE),
    # Count how many alates were dead in each group
    dead_n  = sum(alate_alive == "No",  na.rm = TRUE),
    # Compute total number of alates observed (alive + dead)
    total   = alive_n + dead_n,
    # Calculate percent survival per block per round (0-100%)
    survival_pct = 100 * alive_n / total,
    # Drop grouping structure after summarising (avoids nested data frames in later steps)
    .groups = "drop"
  )

# === Average survival across blocks ===
survival_summary_avg <- survival_summary %>%
  group_by(treatment, round) %>%  # Now group just by treatment and round (collapsed across blocks)
  summarise(
    # Compute the mean survival percentage across all blocks in each treatment/round
    mean_survival = mean(survival_pct, na.rm = TRUE),
    # Calculate standard error: standard deviation divided by square root of the number of blocks
    se_survival   = sd(survival_pct, na.rm = TRUE) / sqrt(n()),
    # Drop grouping again for a clean summary table
    .groups = "drop"
  )
```

Survival Plots

```
# Filter by block
survival_summary_no_b4 <- survival_summary %>%
    filter(block != 4)

# === All Blocks ===
ggplot(data = survival_summary_no_b4, aes(x = interaction(treatment, round), y = survival_pct,
    fill = treatment)) + geom_col(color = "black", size = 0.3) + facet_wrap(~block) +
    labs(x = "Treatment", y = "Alate Survival %", title = "Aphid Survival on #220 or buffer plants - Bl
    scale_fill_manual(values = c("#546214", "#844e27")) + ylim(0, 100) + theme_minimal() +
    theme(legend.position = "right", panel.background = element_rect(fill = "white"),
        panel.grid = element_blank(), panel.grid.minor.x = element_blank(), plot.title = element_text(h
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
```

Aphid Survival on #220 or buffer plants – Block 2
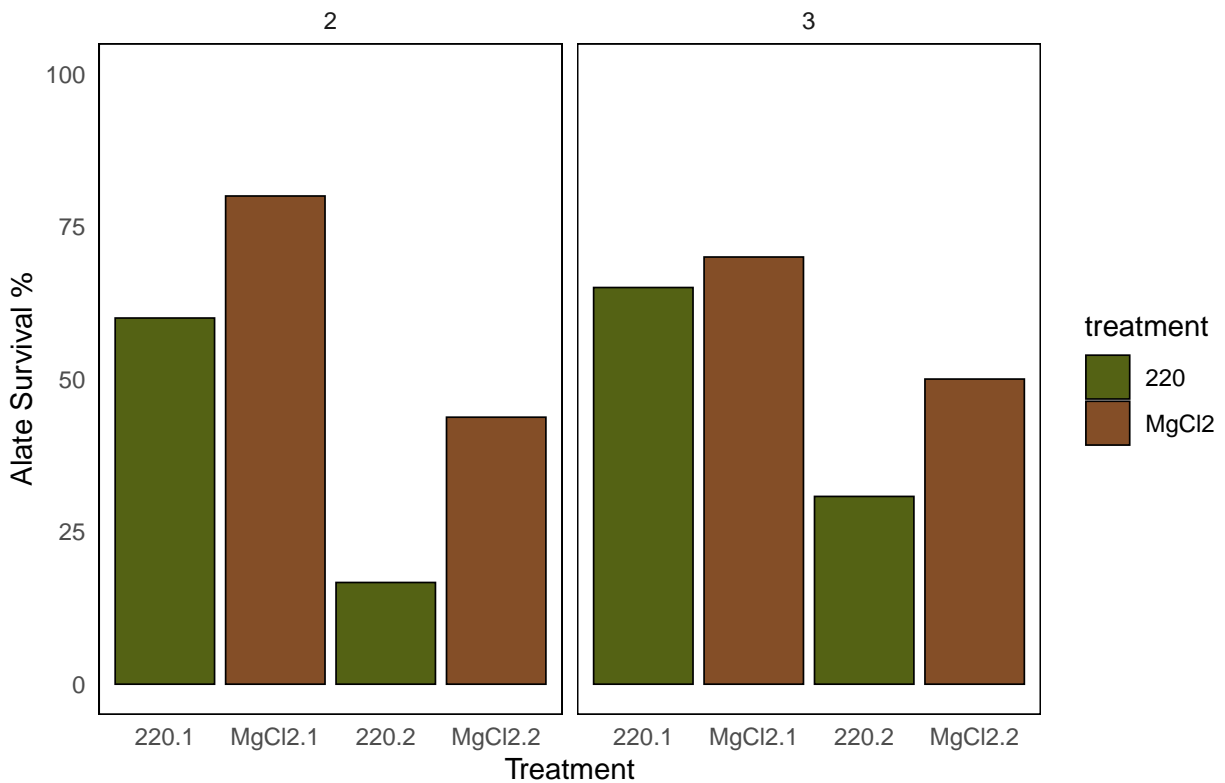

```
# === All Blocks ===
ggplot(survival_summary_avg, aes(x = interaction(treatment, round), y = mean_survival,
    fill = treatment)) + geom_col(color = "black", size = 0.3, position = position_dodge(width = 0.9)) +
    geom_errorbar(aes(ymin = mean_survival - se_survival, ymax = mean_survival +
        se_survival), width = 0.2, color = "black", position = position_dodge(width = 0.9)) +
    labs(x = "Treatment × Round", y = "Mean Alate Survival (%)", title = "Mean Aphid Survival Across Blo
    scale_fill_manual(values = c("#546214", "#844e27")) + ylim(0, 100) + theme_minimal() +
    theme(legend.position = "right", panel.background = element_rect(fill = "white"),
        panel.grid = element_blank(), plot.title = element_text(hjust = 0.5))
```

Mean Aphid Survival Across Blocks
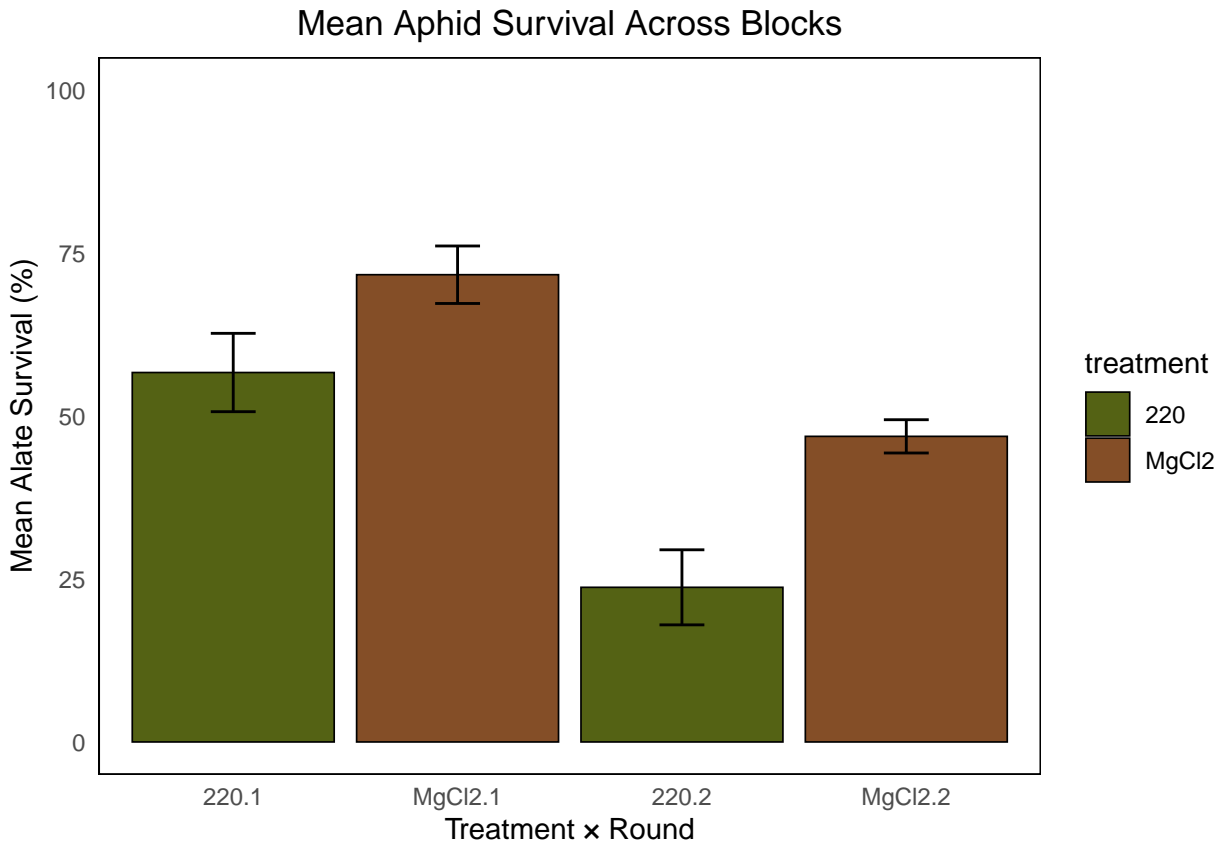
Fecundity

Fecundity Calculations

```r
# === Step 1: Calculate fecundity summary by block and round ===
fecundity_by_block <- aphid_long %>%
  # Group data by treatment condition, block (replicate), and round (e.g. R1 vs R2)
  group_by(treatment, block, round) %>%
  # Summarise the fecundity statistics within each block × treatment × round group
  summarise(
    mean_fecundity = mean(nymphs, na.rm = TRUE), # Mean number of nymphs per alate
    se_nymphs = sd(nymphs, na.rm = TRUE) / sqrt(n()), # Standard error = SD / sqrt(n blocks)

    .groups = "drop" # Ungroup to simplify downstream summarization
  )

# === Step 2: Aggregate fecundity across blocks for each treatment and round ===
fecundity_by_treatment <- aphid_long %>%
  # Now group by treatment and round only (collapse across blocks)
  group_by(treatment, round) %>%
  # Summarise the average of block means and calculate standard error
  summarise(
    mean_nymphs = mean(nymphs, na.rm = TRUE), # Mean of block-level means
    se_nymphs = sd(nymphs, na.rm = TRUE) / sqrt(n()), # Standard error = SD / sqrt(n blocks)
    .groups = "drop" # Drop grouping for a flat summary table
  )
```
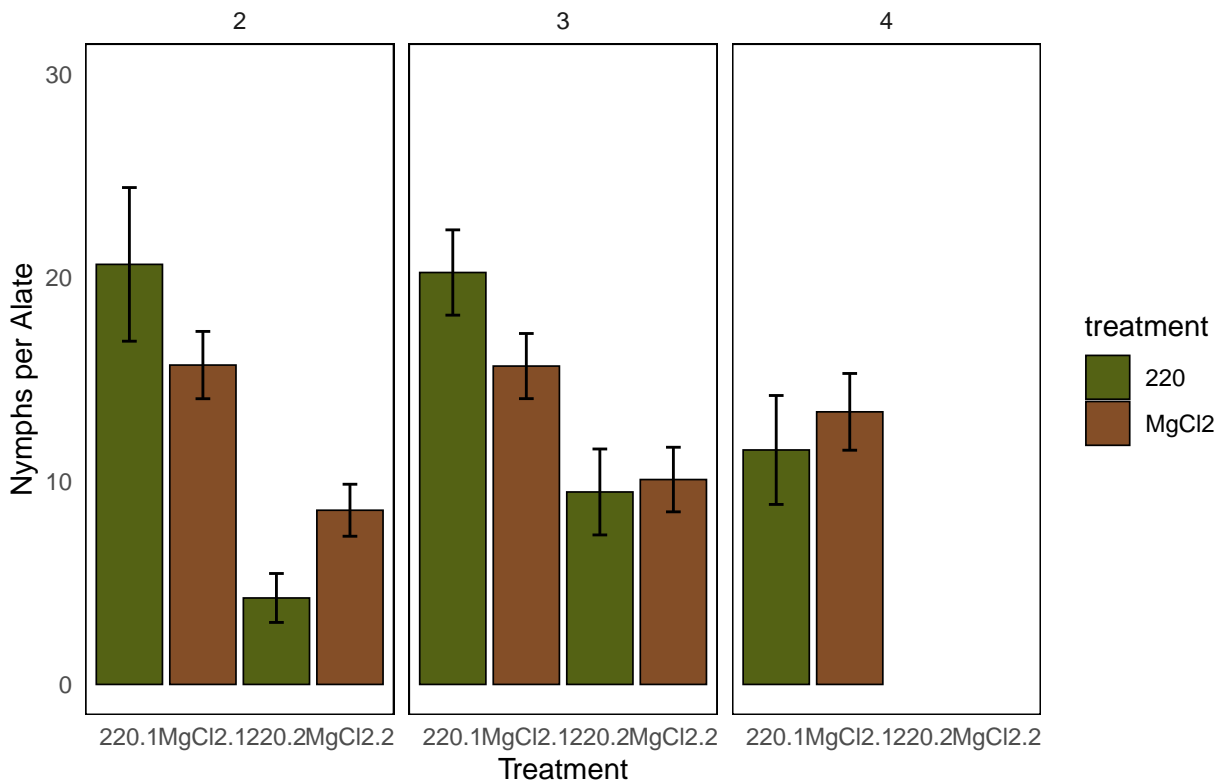
Fecundity Plots

```r
# by block
ggplot(data = fecundity_by_block, aes(x = interaction(treatment, round), y = mean_fecundity,
    fill = treatment)) + geom_col(color = "black", size = 0.3) + facet_wrap(~block) +
    geom_errorbar(aes(ymin = mean_fecundity - se_nymphs, ymax = mean_fecundity +
        se_nymphs), width = 0.2, color = "black", position = position_dodge(width = 0.8)) +
    labs(x = "Treatment", y = "Nymphs per Alate", title = "Aphid Fecundity on #220 or buffer plants - B
    scale_fill_manual(values = c("#546214", "#844e27")) + ylim(0, 30) + theme_minimal() +
    theme(legend.position = "right", panel.background = element_rect(fill = "white"),
        panel.grid = element_blank(), panel.grid.minor.x = element_blank(), plot.title = element_text(h
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_col()`).
```


Aphid Fecundity on #220 or buffer plants – Block 2

```r
# === both blocks ====
ggplot(data = fecundity_by_treatment, aes(x = interaction(treatment, round), y = mean_nymphs,
    fill = treatment)) + geom_col(color = "black", size = 0.3) + geom_errorbar(aes(ymin = mean_nymphs -
    se_nymphs, ymax = mean_nymphs + se_nymphs), width = 0.2, color = "black", position = position_dodge
    labs(x = "Treatment", y = "Nymphs per Alate", title = "Aphid Fecundity on #220 or buffer plants - B
    scale_fill_manual(values = c("#546214", "#844e27")) + ylim(0, 30) + theme_minimal() +
    theme(legend.position = "right", panel.background = element_rect(fill = "white"),
        panel.grid = element_blank(), panel.grid.minor.x = element_blank(), plot.title = element_text(h
```

Aphid Fecundity on #220 or buffer plants – Block 2
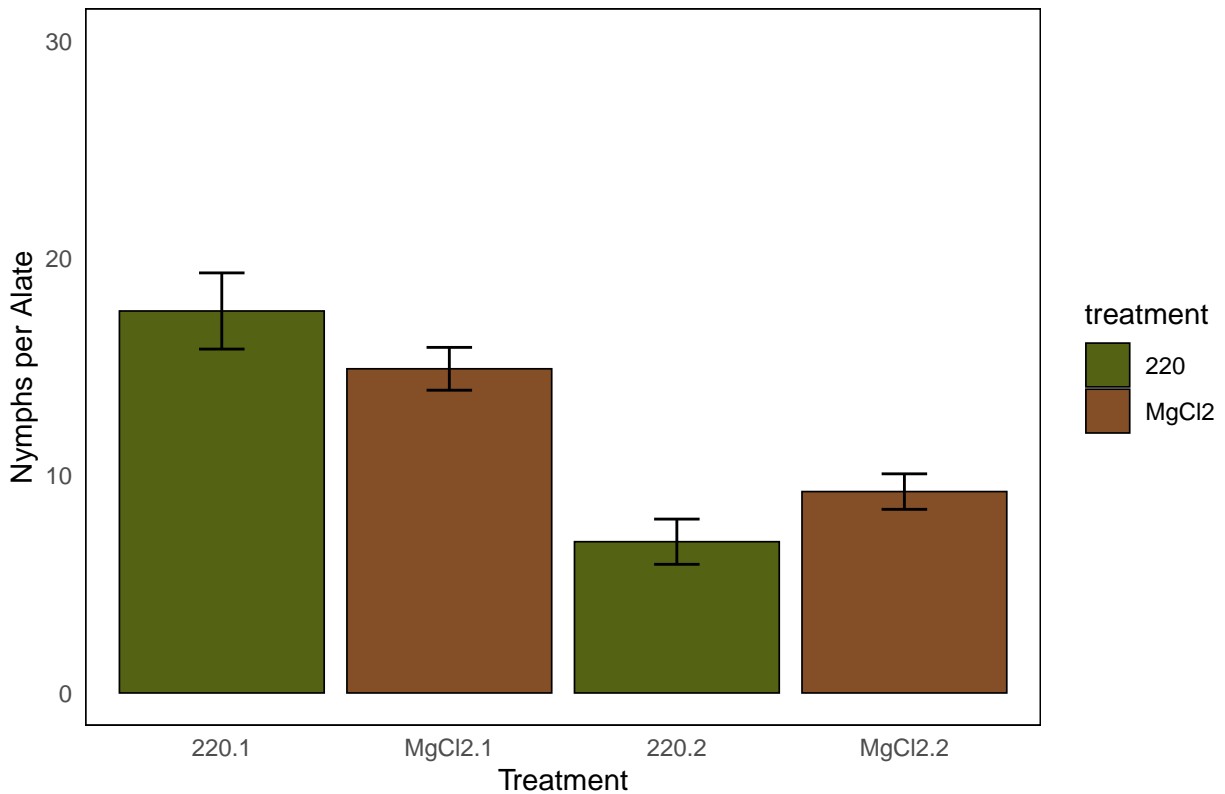
Statistics

alate death

```r
# filter by round
surv_r1 <- survival_summary %>%
    filter(round == 1)
surv_r2 <- survival_summary %>%
    filter(round == 2)

# === Round 1 ===

ch_r1 <- glm(cbind(alive_n, dead_n) ~ treatment, family = binomial, data = surv_r1)

summary(ch_r1)
```

```
##
## Call:
## glm(formula = cbind(alive_n, dead_n) ~ treatment, family = binomial,
##     data = surv_r1)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.2683     0.2605   1.030   0.3031
## treatmentMgCl2  0.6597     0.3872   1.704   0.0884 .
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 5.8964  on 5  degrees of freedom
## Residual deviance: 2.9443  on 4  degrees of freedom
## AIC: 26.951
##
## Number of Fisher Scoring iterations: 4
```

```r
# Pairwise comparisons (log-odds, back-transformed to probability)
emmeans(ch_r1, ~treatment, type = "response")
```

```
##  treatment  prob     SE  df asymp.LCL asymp.UCL
##  220       0.567 0.0640 Inf     0.440     0.685
##  MgCl2     0.717 0.0582 Inf     0.591     0.816
##
## Confidence level used: 0.95
## Intervals are back-transformed from the logit scale
```

```r
# === Round 2 ===

ch_r2 <- glm(cbind(alive_n, dead_n) ~ treatment, family = binomial, data = surv_r2)

summary(ch_r2)
```

```
##
## Call:
## glm(formula = cbind(alive_n, dead_n) ~ treatment, family = binomial,
##     data = surv_r2)
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.1527     0.4683  -2.461   0.0138 *
## treatmentMgCl2  1.0191     0.5943   1.715   0.0864 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 3.90306  on 3  degrees of freedom
## Residual deviance: 0.80946  on 2  degrees of freedom
## AIC: 16.514
##
## Number of Fisher Scoring iterations: 4
```

```r
emmeans(ch_r2, ~treatment, type = "response")
```

```
##  treatment  prob     SE  df asymp.LCL asymp.UCL
##  220       0.240 0.0854 Inf     0.112     0.442
##  MgCl2     0.467 0.0911 Inf     0.299     0.642
##
## Confidence level used: 0.95
## Intervals are back-transformed from the logit scale
```

Stats on alate fecundity

```r
# filter by round
aphid_count_r1_alive_only <- aphid_count %>%
    filter(block != 4, alate_alive_r1 == "Yes")  # Only include live alates in R1
aphid_count_r2_alive_only <- aphid_count %>%
    filter(block != 4, alate_alive_r2 == "Yes")  # Only include live alates in R2

# Exclude round 4
aphid_count_alive_and_dead <- aphid_count %>%
    filter(block != 4)

# Only include aphids that had kids
aphid_count_alive_and_dead_yes_nymphs_r1 <- aphid_count %>%
    filter(block != 4, nymphs_r1 != 0)
aphid_count_alive_and_dead_yes_nymphs_r2 <- aphid_count %>%
    filter(block != 4, nymphs_r2 != 0)

# === Round 1: Alive only ===

ch_fec_r1_alive <- lm(nymphs_r1 ~ treatment, data = aphid_count_r1_alive_only)

summary(ch_fec_r1_alive)
```

```
##
## Call:
## lm(formula = nymphs_r1 ~ treatment, data = aphid_count_r1_alive_only)
##
## Residuals:
##     Min     1Q Median     3Q    Max
##   -21.0   -5.2   -1.2    4.4   29.0
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       24.000      1.835  13.080  < 2e-16 ***
## treatmentMgCl2    -6.800      2.484  -2.737  0.00842 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.174 on 53 degrees of freedom
## Multiple R-squared:  0.1238, Adjusted R-squared:  0.1073
## F-statistic: 7.492 on 1 and 53 DF,  p-value: 0.008418
```

```r
emmeans(ch_fec_r1_alive, ~treatment)
```

```
##  treatment emmean   SE df lower.CL upper.CL
##  220         24.0 1.83 53     20.3     27.7
##  MgCl2       17.2 1.67 53     13.8     20.6
##
## Confidence level used: 0.95
```

```r
# === Round 1: Alive and dead ===

ch_fec_r1 <- lm(nymphs_r1 ~ treatment, data = aphid_count_alive_and_dead)
```

```r
summary(ch_fec_r1)
```

```
##
## Call:
## lm(formula = nymphs_r1 ~ treatment, data = aphid_count_alive_and_dead)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.450  -5.869  -0.062   5.800  32.550
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     20.450      1.708  11.970   <2e-16 ***
## treatmentMgCl2  -4.775      2.416  -1.976   0.0517 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.81 on 78 degrees of freedom
## Multiple R-squared:  0.04769,    Adjusted R-squared:  0.03548
## F-statistic: 3.906 on 1 and 78 DF,  p-value: 0.05165
```

```r
emmeans(ch_fec_r1, ~treatment)
```

```
##  treatment emmean   SE df lower.CL upper.CL
##  220         20.4 1.71 78     17.0     23.9
##  MgCl2       15.7 1.71 78     12.3     19.1
##
## Confidence level used: 0.95
```

```r
# === Round 1: Alive and dead, yes kids ===

ch_fec_r1 <- lm(nymphs_r1 ~ treatment, data = aphid_count_alive_and_dead_yes_nymphs_r1)

summary(ch_fec_r1)
```

```
##
## Call:
## lm(formula = nymphs_r1 ~ treatment, data = aphid_count_alive_and_dead_yes_nymphs_r1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.108  -5.804  -0.500   4.892  30.892
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     22.108      1.636  13.510   <2e-16 ***
## treatmentMgCl2  -5.608      2.299  -2.439   0.0171 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.954 on 73 degrees of freedom
## Multiple R-squared:  0.07537,    Adjusted R-squared:  0.0627
## F-statistic: 5.951 on 1 and 73 DF,  p-value: 0.01714
```

```
emmeans(ch_fec_r1, ~treatment)
```

```
##  treatment emmean   SE df lower.CL upper.CL
##  220         22.1 1.64 73     18.8     25.4
##  MgCl2       16.5 1.61 73     13.3     19.7
##
## Confidence level used: 0.95
```

```
# === Round 2: Alive only ===
```

```
ch_fec_r2_alive <- lm(nymphs_r2 ~ treatment, data = aphid_count_r2_alive_only)
```

```
summary(ch_fec_r2_alive)
```

```
##
## Call:
## lm(formula = nymphs_r2 ~ treatment, data = aphid_count_r2_alive_only)
##
## Residuals:
##     Min     1Q  Median     3Q     Max
## -8.7857 -4.5060  0.2143  4.3333 12.2143
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     18.667      2.509   7.441 6.77e-07 ***
## treatmentMgCl2  -5.881      2.998  -1.961   0.0655 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.145 on 18 degrees of freedom
## Multiple R-squared:  0.1761, Adjusted R-squared:  0.1303
## F-statistic: 3.847 on 1 and 18 DF,  p-value: 0.0655
```

```
emmeans(ch_fec_r2_alive, ~treatment)
```

```
##  treatment emmean   SE df lower.CL upper.CL
##  220         18.7 2.51 18    13.40     23.9
##  MgCl2       12.8 1.64 18     9.34     16.2
##
## Confidence level used: 0.95
```

```
# === Round 2: Alive and dead ===
```

```
ch_fec_r2 <- lm(nymphs_r2 ~ treatment, data = aphid_count_alive_and_dead)
```

```
summary(ch_fec_r2)
```

```
##
## Call:
## lm(formula = nymphs_r2 ~ treatment, data = aphid_count_alive_and_dead)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.267 -6.113 -1.960  4.733 20.040
##
## Coefficients:
```

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)       6.960       1.433   4.857  1.1e-05 ***
## treatmentMgCl2    2.307       1.940   1.189     0.24
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.165 on 53 degrees of freedom
##   (25 observations deleted due to missingness)
## Multiple R-squared:  0.02597,    Adjusted R-squared:  0.007596
## F-statistic: 1.413 on 1 and 53 DF,  p-value: 0.2398
```

```
emmeans(ch_fec_r2, ~treatment)
```

```
##  treatment emmean   SE df lower.CL upper.CL
##  220         6.96 1.43 53     4.09     9.83
##  MgCl2       9.27 1.31 53     6.64    11.89
##
## Confidence level used: 0.95
```

```
# === Round 2: Alive and dead, yes kids ===
```

```
ch_fec_r1 <- lm(nymphs_r2 ~ treatment, data = aphid_count_alive_and_dead_yes_nymphs_r2)
```

```
summary(ch_fec_r1)
```

```
##
## Call:
## lm(formula = nymphs_r2 ~ treatment, data = aphid_count_alive_and_dead_yes_nymphs_r2)
##
## Residuals:
##     Min     1Q Median     3Q    Max
## -9.235 -5.235 -1.235  4.071 16.765
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    10.2353     1.6403   6.240 1.64e-07 ***
## treatmentMgCl2 -0.3067     2.0795  -0.147    0.883
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.763 on 43 degrees of freedom
## Multiple R-squared:  0.0005057, Adjusted R-squared:  -0.02274
## F-statistic: 0.02176 on 1 and 43 DF,  p-value: 0.8834
```

```
emmeans(ch_fec_r1, ~treatment)
```

```
##  treatment emmean   SE df lower.CL upper.CL
##  220        10.24 1.64 43     6.93     13.5
##  MgCl2       9.93 1.28 43     7.35     12.5
##
## Confidence level used: 0.95
```