Relative Brain Mass

Kyle Emerson

12/13/2022

## R Markdown

This RMarkdown will contain start to finish analysis of Relative Brain Mass information compiled from the Emerson 1 Microbial Experiment (Aim 2). It is updated to be more streamlined and easier to knit the output and be placed into my data notebook. Note: This is updated for Manuscript Writing and all analysis was done in RStudio

**Relative Brain Mass**

library(lme4)

## Warning: package 'lme4' was built under R version 4.1.3

## Loading required package: Matrix

## Warning: package 'Matrix' was built under R version 4.1.3

library(car)

## Warning: package 'car' was built under R version 4.1.3

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.1.3

library(readr)

## Warning: package 'readr' was built under R version 4.1.3

library(moments)

## Warning: package 'moments' was built under R version 4.1.3

library(psych)

## Warning: package 'psych' was built under R version 4.1.3

##   
## Attaching package: 'psych'

## The following object is masked from 'package:car':  
##   
## logit

library(pastecs)  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.1.3

##   
## Attaching package: 'ggplot2'

## The following objects are masked from 'package:psych':  
##   
## %+%, alpha

library(ggbiplot)

## Loading required package: plyr

## Warning: package 'plyr' was built under R version 4.1.3

## Loading required package: scales

## Warning: package 'scales' was built under R version 4.1.3

##   
## Attaching package: 'scales'

## The following objects are masked from 'package:psych':  
##   
## alpha, rescale

## The following object is masked from 'package:readr':  
##   
## col\_factor

## Loading required package: grid

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v tibble 3.1.7 v dplyr 1.0.9  
## v tidyr 1.2.0 v stringr 1.4.0  
## v purrr 0.3.4 v forcats 0.5.1

## Warning: package 'tibble' was built under R version 4.1.3

## Warning: package 'tidyr' was built under R version 4.1.3

## Warning: package 'dplyr' was built under R version 4.1.3

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x ggplot2::%+%() masks psych::%+%()  
## x scales::alpha() masks ggplot2::alpha(), psych::alpha()  
## x dplyr::arrange() masks plyr::arrange()  
## x scales::col\_factor() masks readr::col\_factor()  
## x purrr::compact() masks plyr::compact()  
## x dplyr::count() masks plyr::count()  
## x purrr::discard() masks scales::discard()  
## x tidyr::expand() masks Matrix::expand()  
## x tidyr::extract() masks pastecs::extract()  
## x dplyr::failwith() masks plyr::failwith()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::first() masks pastecs::first()  
## x dplyr::id() masks plyr::id()  
## x dplyr::lag() masks stats::lag()  
## x dplyr::last() masks pastecs::last()  
## x dplyr::mutate() masks plyr::mutate()  
## x tidyr::pack() masks Matrix::pack()  
## x dplyr::recode() masks car::recode()  
## x dplyr::rename() masks plyr::rename()  
## x purrr::some() masks car::some()  
## x dplyr::summarise() masks plyr::summarise()  
## x dplyr::summarize() masks plyr::summarize()  
## x tidyr::unpack() masks Matrix::unpack()

library(modelbased)

## Warning: package 'modelbased' was built under R version 4.1.3

library(dplyr)  
  
file.choose()

## [1] "C:\\R\\Emerson-Microbial-1\\RMarkdowns\\Knitted\\EM1-Relative Brain Mass.docx"

df <- read.csv("C:\\Users\\kjeme\\OneDrive\\Desktop\\Woodley Lab\\Aim 2 - Emerson Microbial 1\\Emersion Microbial Experiment 1 (2021)\\Emerson Microbial Exp1\_Batch1.csv")  
df$Microbial\_Trtmt = factor(df$Microbial\_Trtmt)  
df$Replicate = factor(df$Replicate)  
  
shapiro.test(df$Log\_BrainMass)

##   
## Shapiro-Wilk normality test  
##   
## data: df$Log\_BrainMass  
## W = 0.9606, p-value = 0.1756

Shapiro-Wilk test to ensure normality of Brain Mass. Data is normal.

leveneTest(df$Log\_BrainMass, df$Microbial\_Trtmt, center = mean, na.rm = TRUE)

## Levene's Test for Homogeneity of Variance (center = mean: TRUE)  
## Df F value Pr(>F)  
## group 1 0.9094 0.3463  
## 38

Levenes test for homogeneity of variance. Data is homoscedastic.

anovaBrainMass <- aov(Log\_BrainMass~Microbial\_Trtmt\*Log\_BodyMass, data = df)  
summary(anovaBrainMass, type = "III")

## Df Sum Sq Mean Sq F value Pr(>F)   
## Microbial\_Trtmt 1 0.05096 0.05096 22.743 3.04e-05 \*\*\*  
## Log\_BodyMass 1 0.03621 0.03621 16.161 0.000284 \*\*\*  
## Microbial\_Trtmt:Log\_BodyMass 1 0.00586 0.00586 2.617 0.114475   
## Residuals 36 0.08067 0.00224   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## 32 observations deleted due to missingness

Based on P-value, shows that slopes are homogenous across treatment combos when factoring in body mass. This serves as our ANCOVA (univariate analysis) which took body mass into account when developing our relative brain mass values. By confirming there was no significant interactions between our treatments, body mass and brain mass, we can confirm that the slopes of lines for this trait were parallel across both treatment groups. Now that we have passed our assumptions and our homogeneity of slopes, I need to calculate my MAV.

anovaBrainMass$residuals

## 1 4 7 9 10   
## -0.0756676758 -0.1402419809 -0.0283306490 -0.0256066829 -0.0136985488   
## 12 13 15 16 19   
## 0.0271241596 -0.0842971443 -0.0107925203 0.0357431707 -0.0613924854   
## 21 22 24 25 31   
## -0.0007293586 -0.0099861200 -0.0110562497 0.0137450125 -0.0069298316   
## 33 34 37 38 39   
## 0.0161641621 0.0317423279 -0.0208847892 -0.0014363210 0.0254032276   
## 40 43 44 45 49   
## -0.0100873996 -0.0007463521 -0.0639329769 0.0574995970 0.0857608218   
## 50 52 53 55 56   
## 0.0030920149 0.0468480304 0.0144045905 0.0466883514 -0.0044588660   
## 57 58 59 61 62   
## 0.0186146495 0.0318935735 0.0249544265 -0.0293236748 0.0520690293   
## 63 67 68 70 71   
## -0.0037020152 0.0243773516 0.0748171254 0.0418860278 -0.0695260079

The Anova calculated our residuals from this statistical model that we can add to our EMMs, which we will calculate now.

model <- lm(Log\_BrainMass ~ Microbial\_Trtmt\*Log\_BodyMass, data = df)  
means\_complex <- estimate\_means(model)

## We selected `at = c("Microbial\_Trtmt")`.

## NOTE: Results may be misleading due to involvement in interactions

means\_complex$Mean

## [1] -1.779171 -1.748328

These are our estimated marginal means for our Natural and Autoclaved pond water treatment groups. We will add our residuals that we calculated to these values, which are located in a new data frame within the environment. Each residual from each individual will be added to the EMM corresponding to their treatment group. These are approximately the same EMM output we got from SPSS (Analyze -> GLM -> Univariate)

residuals <- anovaBrainMass$residuals

Turned residuals into separate vector to add to the original data frame when you are doing this the first time.

# ID <-df$ID  
# Log\_BrainMass <- df$Log\_BrainMass  
# Microbial\_Trtmt <- df$Microbial\_Trtmt  
# Replicate <- df$Replicate  
# df2 <- cbind(ID, Log\_BrainMass)  
# df2 <- cbind(df2, Microbial\_Trtmt)  
# df2 <- cbind(df2, Replicate)  
# w = complete.cases(df2)  
# df2 = df2[w,]  
# df2 <- cbind(df2, residuals)  
# write.table(df2, file = "EM1dfwithresiduals.csv", sep = ",")

If you are doing this for the first time, this will give you a new data frame CSV file containing your individuals and their residuals. Once you have your residuals correctly assigned to each tadpole, add them to the original master CSV file. Now that our new residuals (Residuals\_New) have been added to the original master CSV file, we are going to Mass adjust those variables. To do so, I am going to add a new column with EMM next to each column. Then, I am going to add a third column representing our mass adjusted brain mass next to the EMM column I will add the Residuals\_New to the EMM based on treatment group and have the resulting number be our MA\_Brain Mass, in the original master CSV file. NOTE: If a tadpole did not have its mass included for any reason (RNALater, damage, etc), it will not have a residual to add to an EMM. I will leave these sections in the excel file as NA, so they will not be incorporated as 0’s into the data analysis.

MABM\_glmm <- glmer(MA\_BrainMass~Microbial\_Trtmt + (1|Replicate), data = df, family = "gaussian")

## Warning in glmer(MA\_BrainMass ~ Microbial\_Trtmt + (1 | Replicate), data = df, :  
## calling glmer() with family=gaussian (identity link) as a shortcut to lmer() is  
## deprecated; please call lmer() directly

## boundary (singular) fit: see help('isSingular')

Anova(MABM\_glmm)

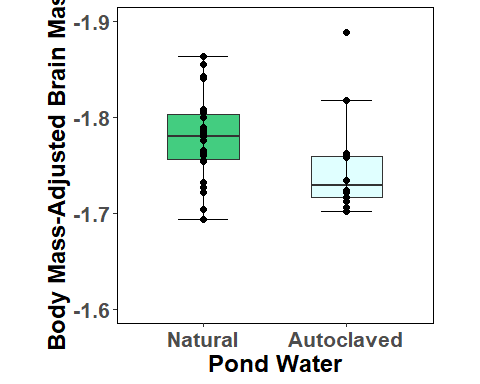
## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: MA\_BrainMass  
## Chisq Df Pr(>Chisq)   
## Microbial\_Trtmt 4.0778 1 0.04345 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Our Anova shows a significant effect of microbial treatment on our mass adjusted relative brain mass values.

Pond\_Water <- c("Natural", "Autoclaved")  
  
ggplot(df, aes(x= Microbial\_Trtmt, y = MA\_BrainMass, fill = Microbial\_Trtmt)) +  
 geom\_boxplot(width = 0.5, outlier.colour = "transparent") +  
 coord\_cartesian(ylim = c(-1.6, -1.9)) +  
 annotate(x = 1, xend = 2, y = .12, yend = .12, geom = "segment") +  
 annotate(x=1.5, y=.128, label = "p < 0.05", geom = "text", size = 5.2) +  
 theme\_classic() +  
 stat\_boxplot(geom = "errorbar", width = .35) +  
 geom\_point(size = 2)+  
 scale\_fill\_manual(values = c("seagreen3", "lightcyan"),  
 name = "Pond Water",  
 labels = c("Natural", "Autoclaved")) +  
 labs(x = "Pond Water", y = "Body Mass-Adjusted Brain Mass") +  
 scale\_x\_discrete(labels = Pond\_Water) +  
 theme(panel.background = element\_rect(fill = "white", colour = "black")) +  
 theme(aspect.ratio = 1) +  
 theme(axis.text = element\_text(face = "bold", size = 16)) +  
 theme(axis.title = element\_text(face = "bold", size = 18)) +  
 theme(legend.position = "none")

## Warning: Removed 32 rows containing non-finite values (stat\_boxplot).  
## Removed 32 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 32 rows containing missing values (geom\_point).

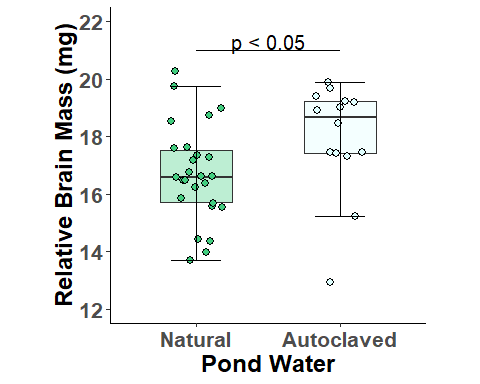


Now, to make the final graph, I use the inverse log excel function to turn our MA\_BrainMass into tangible values. Because these values are in grams, I will multiply by 1000 so our Relative\_BrainMass column is in mg.

ggplot(df, aes(x= Microbial\_Trtmt, y = Relative\_BrainMass, fill = Microbial\_Trtmt)) +  
 geom\_boxplot(width = 0.5, outlier.colour = "transparent", alpha =0.35) +  
 geom\_jitter(width = .2, size = 2.5, shape = 21, color = "black") +  
 coord\_cartesian(ylim = c(12, 22)) +  
 scale\_y\_continuous(breaks = seq(12, 22, 2)) +  
 annotate(x = 1, xend = 2, y = 21, yend = 21, geom = "segment") +  
 annotate(x=1.5, y= 21.3, label = "p < 0.05", geom = "text", size = 5.2) +  
 theme\_classic() +  
 stat\_boxplot(geom = "errorbar", width = .35) +  
 scale\_fill\_manual(values = c("seagreen3", "lightcyan"),  
 name = "Pond Water",  
 labels = c("Natural", "Autoclaved")) +  
 labs(x = "Pond Water", y = "Relative Brain Mass (mg)") +  
 scale\_x\_discrete(labels = Pond\_Water) +  
 theme(aspect.ratio = 1) +  
 theme(axis.text = element\_text(face = "bold", size = 16)) +  
 theme(axis.title = element\_text(face = "bold", size = 18)) +  
 theme(legend.position = "none")

## Warning: Removed 32 rows containing non-finite values (stat\_boxplot).  
## Removed 32 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 32 rows containing missing values (geom\_point).



This is our completed relative brain mass figure and analysis!