Behavioral Tolerance, Vigor, and resilience

David R. Clark, Jason Walsman, Rachael Kramp, and Jessica Stephenson

2024-03-25

Contents

1 Introduction

1.1 Overall summary

1.2 Quick experimental summary

1.3 Load in packages needed for analysis

```
# Clear the working environment
rm(list = ls())
# Visualization
library(ggplot2)
library(visreg)
source("http://highstat.com/Books/BGS/GAMM/RCodeP2/HighstatLibV6.R")
# (generalized) Linear mixed modeling
library(lme4)
library(glmmTMB)
library(lmodel2)
# Statistcal analysis reporting and model validation
library(performance)
library(car)
library(lmtest)
library(DHARMa)
# Data wrangling
library(dplyr)
library(plyr)
library(tidyverse)
library(tidylog)
library(splancs)
```

1.4 Load in dataframe

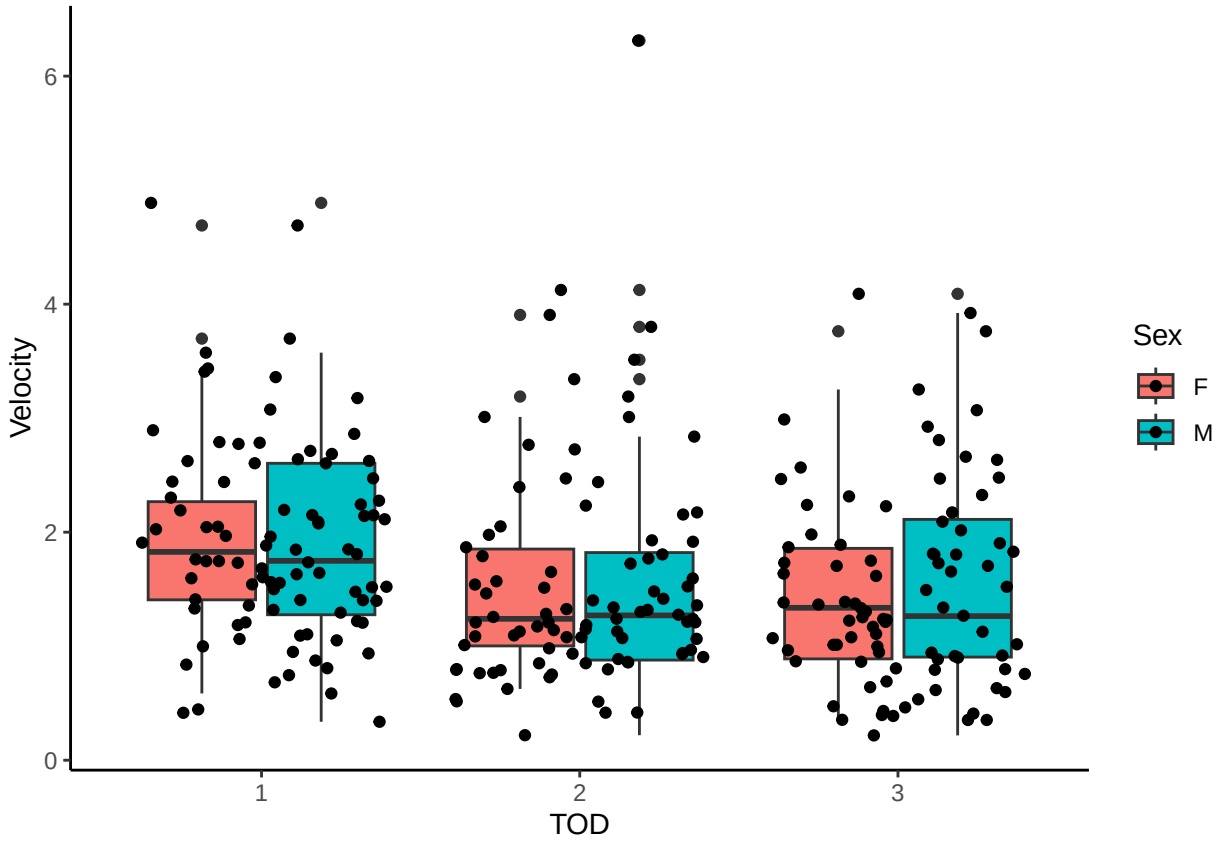
```
# Loading in data set
IndBehav <- read_csv("VIEBehavior_20210913_V2.csv", col_types = cols(fishID = col_character(),
    Infection = col_character(), InfDate = col_character(), BehavGroup = col_character(),
    dayofinf = col_character(), CountDay = col_character(), Behavdate = col_character(),
    TOD = col_character(), Duration = col_double(), Velocity = col_double(), Distance = col_double()))

# Filtering Data set so we can perform our analysis
IndBehav1 <- IndBehav %>%
    # Removing columns we dont need
select(-c(Blind.ID, InfLength, dayofinf, Countintial, Notes, InfWeight))
```

```
## select: dropped 6 variables (Blind.ID, InfWeight, InfLength, dayofinf, Countintial, ...)
```

```
# Quickly checking whether individuals have stress response from a new
# environment.

# Filter to only before measurements
IndBehav1_Before <- IndBehav1 %>%
    filter(TrialTime == "Before")
```

```
## filter: removed 828 rows (75%), 276 rows remaining
```

```
# plot the Activity by time of day

ggplot(IndBehav1_Before, aes(x = TOD, y = Velocity, fill = Sex)) + geom_boxplot(aes(fill = Sex)) +
    geom_jitter() + theme_classic()
```

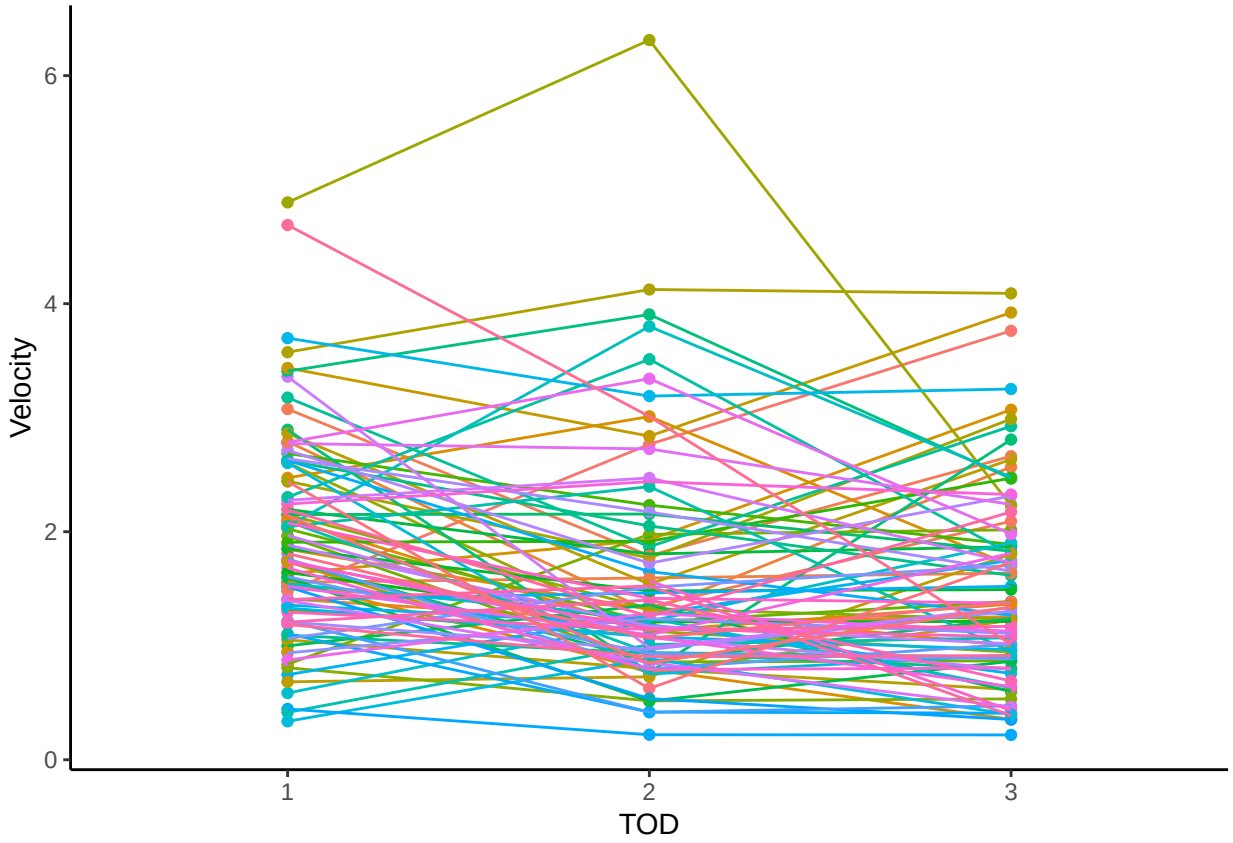



```
# Quick Anova to verify that the difference is significantly different

anova(lm(Velocity ~ TOD * Sex, IndBehav1_Before))
```

```
## Analysis of Variance Table
##
## Response: Velocity
##            Df  Sum Sq Mean Sq F value   Pr(>F)
## TOD         2   9.570  4.7852  5.7966 0.003438 **
## Sex         1   0.010  0.0097  0.0117 0.913883
## TOD:Sex     2   0.148  0.0740  0.0896 0.914307
```

```
## Residuals 264 217.939  0.8255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# We do see a significant TOD effect so lets more specifically look at each
# individual
ggplot(IndBehav1_Before, aes(x = TOD, y = Velocity, color = fishID)) + geom_point() +
    geom_line(aes(group = fishID)) + theme_classic() + theme(legend.position = "none")
```




```
# Dropping the first video due to these TOD effects
IndBehav1 <- IndBehav1 %>%
    filter(TOD != "1" | TrialTime != "Before")
```

```
## filter: removed 92 rows (8%), 1,012 rows remaining
```

1.5 Quickly check for stress response when individauls are

1.6 Calculate metrics needed for further analyses

1.6.1 Calculating all body condition metrics

```
##### This code is just to examine how many fish cleared their infection
##### (13)#### IndBehavRec<- IndBehav1 %>% filter(Recov==1)%>% filter(Fish==1)
```

```
##### This code is just to examine how many fish died during their infection
##### (2)##### IndBehavDead<- IndBehav1 %>% filter(Died==1)%>% filter(Fish==1)

FemaleOnly <- IndBehav1 %>%
    filter(Sex == "F")
```

```
## filter: removed 495 rows (49%), 517 rows remaining
```

```
MaleOnly <- IndBehav1 %>%
    filter(Sex == "M")
```

```
## filter: removed 517 rows (51%), 495 rows remaining
```

```
# Calculating SMI metrics Preinfection SMI#####

# PreinfectionSMI for females To calculate the SMI we take the OLS slope
lmodel2(log(as.numeric(PreWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = FemaleOnly)
```

```
## RMA was not requested: it will not be computed.

## No permutation test will be performed

##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(PreWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = FemaleOnly)
##
## n = 517   r = 0.9254634   r-square = 0.8564825
## Parametric P-values:   2-tailed = 3.036124e-219    1-tailed = 1.518062e-219
## Angle between the two OLS regression lines = 2.596854 degrees
##
## Regression results
##   Method Intercept    Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -3.553804 2.865432        70.76165               NA
## 2     MA -4.874770 3.299656        73.13995               NA
## 3    SMA -4.255871 3.096214        72.10083               NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS      -3.862846       -3.244762   2.763890    2.966975
## 2     MA      -5.242439       -4.530218   3.186396    3.420515
## 3    SMA      -4.569842       -3.952028   2.996335    3.199421
##
## Eigenvalues: 0.08510999 0.001071282
##
## H statistic used for computing C.I. of MA: 9.675165e-05
```

```
# Take the median length from summary
summary(FemaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.60   18.40   20.20   20.03   21.20   24.20
```

```
# Adding PreSMI to Dataframe
FemaleOnly <- FemaleOnly %>%
    group_by(fishID) %>%
    mutate(PreSMI = PreWeight * ((20.2/PreLength)^0.4621474))
```

```
## group_by: one grouping variable (fishID)
```

```
## mutate (grouped): new variable 'PreSMI' (double) with 47 unique values and 0% NA
```

```
# PreinfectionSMI for males
lmodel2(log(as.numeric(PreWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = MaleOnly)
```

```
## RMA was not requested: it will not be computed.
```

```
## No permutation test will be performed
```

```
##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(PreWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = MaleOnly)
##
## n = 495   r = 0.751645   r-square = 0.5649703
## Parametric P-values:   2-tailed = 3.745691e-91    1-tailed = 1.872846e-91
## Angle between the two OLS regression lines = 9.502639 degrees
##
## Regression results
##   Method Intercept    Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -2.184579 2.359450        67.03143               NA
## 2     MA -6.768959 4.002268        75.97140               NA
## 3    SMA -4.360093 3.139048        72.32979               NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS      -2.695974       -1.673184   2.176240    2.542660
## 2     MA      -7.704742       -5.960666   3.712615    4.337607
## 3    SMA      -4.886259       -3.863742   2.961180    3.327600
##
## Eigenvalues: 0.04339483 0.00170565
##
## H statistic used for computing C.I. of MA: 0.0003334766
```

```
# Take the median length from summary
summary(MaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   13.70   14.50   15.30   15.32   16.00   18.50
```

```r
# Adding PreSMI to Dataframe
MaleOnly <- MaleOnly %>%
    group_by(fishID) %>%
    mutate(PreSMI = PreWeight * ((15.3/PreLength)^0.2489888))
```

```
## group_by: one grouping variable (fishID)
```

```
## mutate (grouped): new variable 'PreSMI' (double) with 44 unique values and 0% NA
```

```r
##### Late Infection SMI####

# Late infection SMI for females Take SMA slope from this model
lmodel2(log(as.numeric(LateWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = FemaleOnly)
```

```
## RMA was not requested: it will not be computed.
```

```
## No permutation test will be performed
```

```
##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(LateWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = FemaleOnly)
##
## n = 462   r = 0.7677924   r-square = 0.5895052
## Parametric P-values:   2-tailed = 5.561381e-91    1-tailed = 2.78069e-91
## Angle between the two OLS regression lines = 10.22187 degrees
##
## Regression results
##   Method Intercept    Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -0.931036 1.978491        63.18634                NA
## 2     MA -4.543783 3.166554        72.47385                NA
## 3    SMA -2.750589 2.576857        68.79028                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS      -1.391251      -0.4708205   1.827219    2.129763
## 2     MA      -5.335219      -3.8552645   2.940133    3.426820
## 3    SMA      -3.224076      -2.3040815   2.430021    2.732564
##
## Eigenvalues: 0.06381766 0.003295778
##
## H statistic used for computing C.I. of MA: 0.0004820595
```

```r
# Take the median length from summary
summary(FemaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.60   18.40   20.20   20.03   21.20   24.20
```

```r
# Adding PreSMI to Dataframe
FemaleOnly <- FemaleOnly %>%
    group_by(fishID) %>%
    mutate(LateSMI = PreWeight * ((20.1/PreLength)^0.3687192))
```

```
## group_by: one grouping variable (fishID)

## mutate (grouped): new variable 'LateSMI' (double) with 47 unique values and 0% NA
```

```r
# Late Infection SMI for males Take OLS slope from this model
lmodel2(log(as.numeric(LateWeight) + 1) ~ log(as.numeric(PreLength) + 1), data = MaleOnly)
```

```
## RMA was not requested: it will not be computed.

## No permutation test will be performed

##
## Model II regression
##
## Call: lmodel2(formula = log(as.numeric(LateWeight) + 1) ~
## log(as.numeric(PreLength) + 1), data = MaleOnly)
##
## n = 418   r = 0.2491056   r-square = 0.06205362
## Parametric P-values:   2-tailed = 2.479029e-07    1-tailed = 1.239514e-07
## Angle between the two OLS regression lines = 38.0442 degrees
##
## Regression results
##   Method  Intercept     Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS   1.319746  1.144385        48.85201                NA
## 2     MA -44.595472 17.624781        86.75262                NA
## 3    SMA  -8.290986  4.593975        77.71963                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS       0.124854        2.514638  0.7155953    1.573175
## 2     MA     -73.972720      -31.197743 12.8159205   28.169186
## 3    SMA      -9.541247       -7.151987  4.1851529    5.042732
##
## Eigenvalues: 0.07219651 0.003188958
##
## H statistic used for computing C.I. of MA: 0.000449061
```

```r
# Take the median length from summary
summary(MaleOnly$PreLength)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   13.70   14.50   15.30   15.32   16.00   18.50
```

```r
# Adding PreSMI to Dataframe
MaleOnly <- MaleOnly %>%
    group_by(fishID) %>%
    mutate(LateSMI = PreWeight * ((15.3/PreLength)^0.45174853))
```

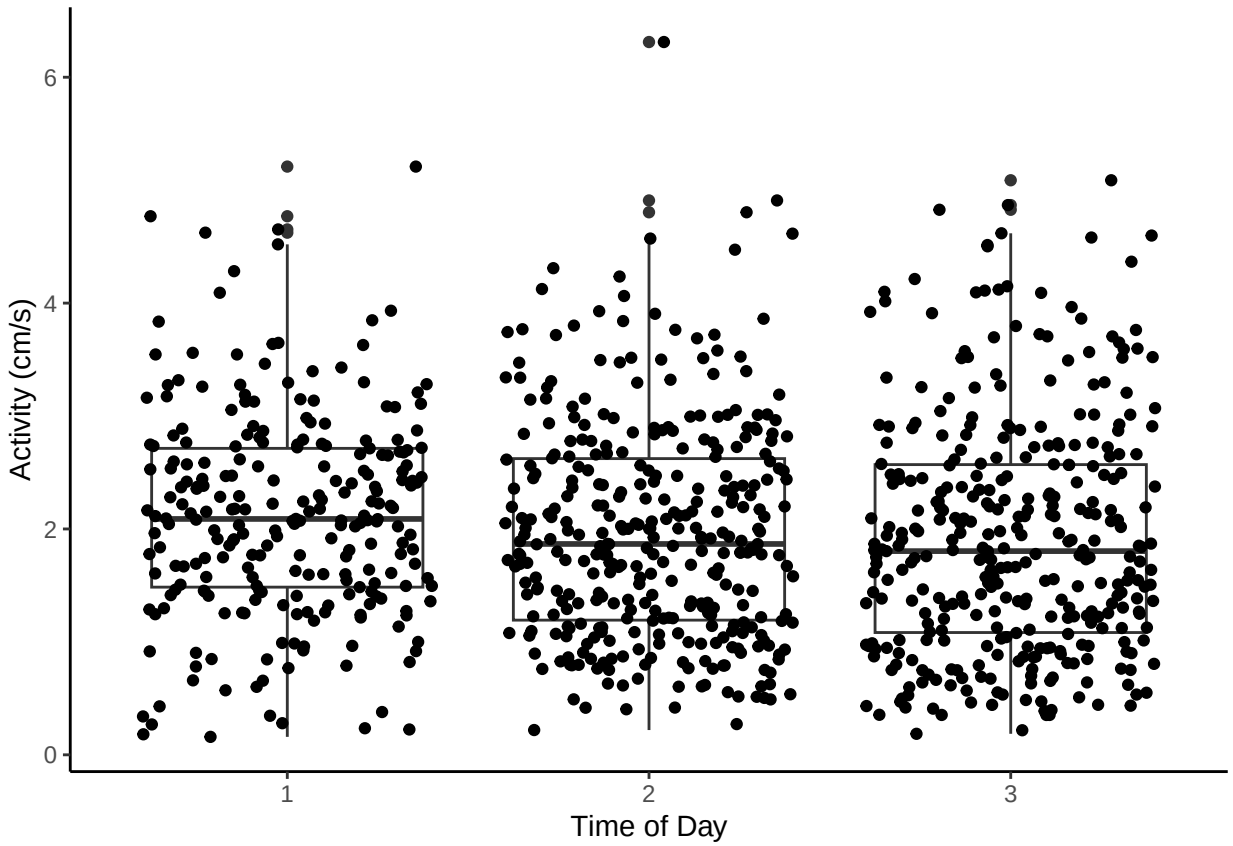
```
## group_by: one grouping variable (fishID)

## mutate (grouped): new variable 'LateSMI' (double) with 44 unique values and 0% NA
```

```
#####

# Bringing the male and female dataframes together to create our overall
# dataframe
IndBehav2 <- rbind(FemaleOnly, MaleOnly)


# View(IndBehav2)
```

1.6.2 Checking that fish activity is not different at different times of day.

We recorded the activity of the fish at 3 time periods over the course of the day. Once in the monring between 0900-1100, once in the afternoon 1200-1400, 1500-1700. Therefore we wanted to confirm that there is no consistent Time of day effect before we collapse the behavioral measurements into means for each day.

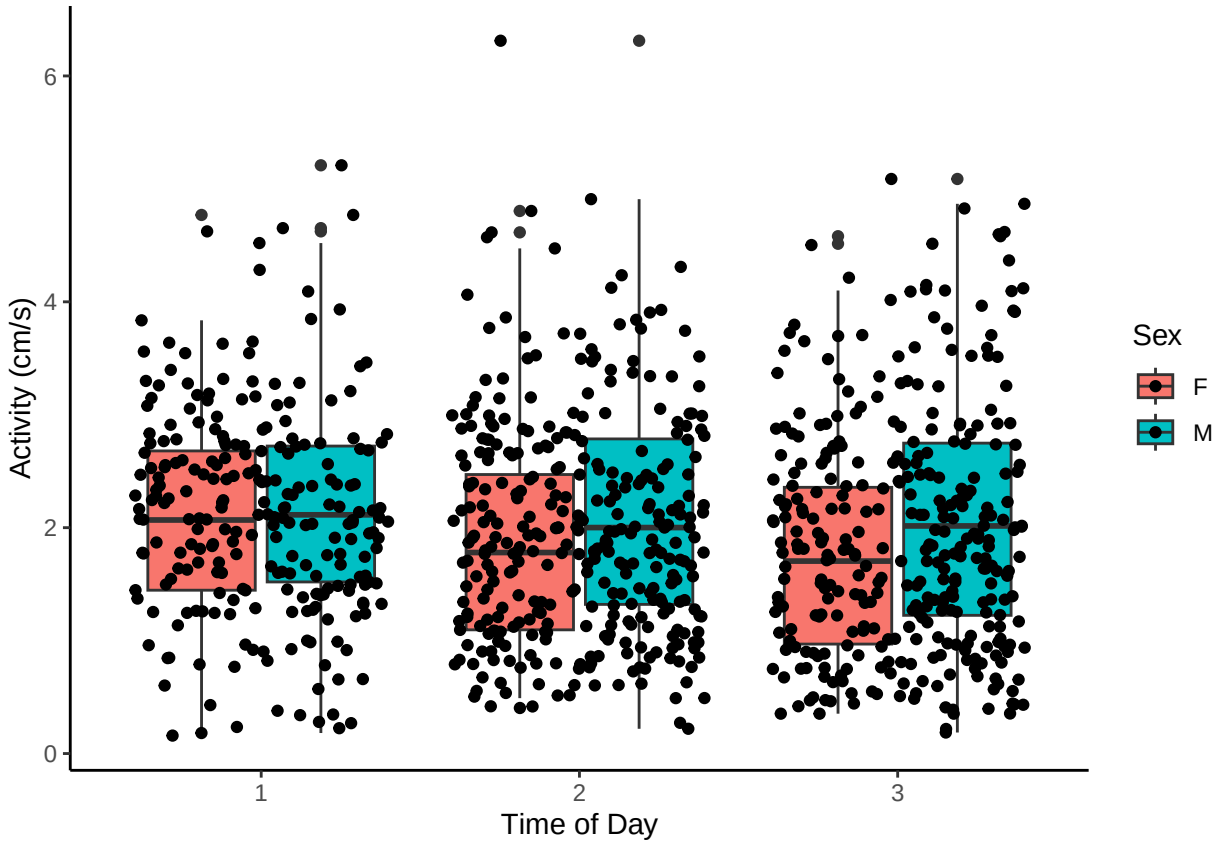
```
##### First visually look at the TOD effect for behavior####
ggplot(IndBehav2, aes(TOD, Velocity)) + geom_boxplot() + geom_jitter() + xlab("Time of Day") +
    ylab("Activity (cm/s)") + theme_classic()
```



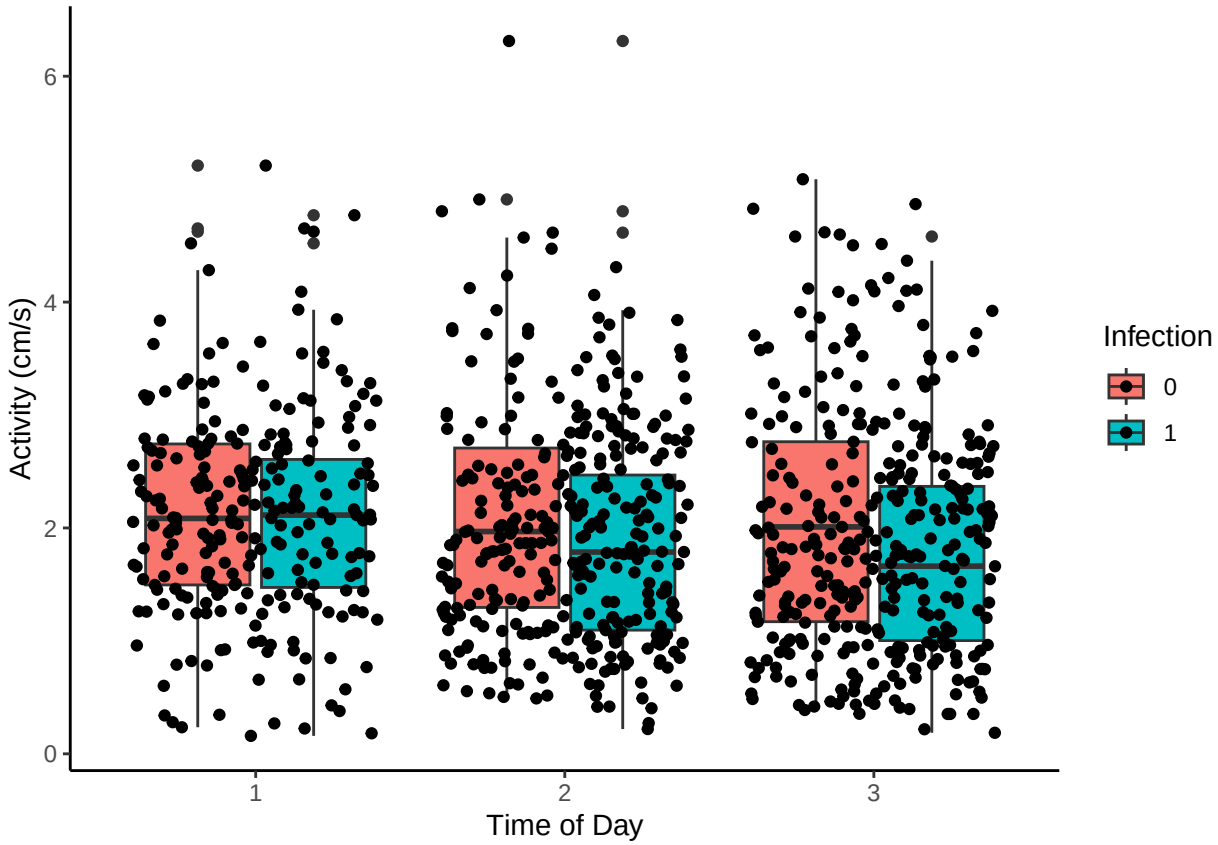
```
#### Verify that this is not statistically significant#####
anova(lm(Velocity ~ TOD, IndBehav2))
```

```
## Analysis of Variance Table
##
## Response: Velocity
##            Df Sum Sq Mean Sq F value  Pr(>F)
## TOD         2   5.55 2.77265  2.8353 0.05921 .
## Residuals 927 906.51 0.97789
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##### Lets also confirm this does not change by sex####
ggplot(IndBehav2, aes(TOD, Velocity, fill = Sex)) + geom_boxplot() + geom_jitter() +
    xlab("Time of Day") + ylab("Activity (cm/s)") + theme_classic()
```




```
##### Lets also confirm this does n0t change by Infection Status####
ggplot(IndBehav2, aes(TOD, Velocity, fill = Infection)) + geom_boxplot() + geom_jitter() +
    xlab("Time of Day") + ylab("Activity (cm/s)") + theme_classic()
```



Visually there does not seem to be a distinguished pattern overall and an anova confirms it. It also does not seem to visually differ by sex or infection status. Therefore we can collapse the three different time points within a day to a mean activity and variance of activity across each time point, Before, Early, Late, and Later.

1.6.3 Calculating the behavioral tolerance metrics needed for analyses

We are calculating behavioral tolerance through two metrics. First is a measure of behavioral tolerance using linear mixed models using slope random effects for fishID. This allows us to extract each individuals change in activity with parasite burden and therefore allow us to calculate the tolerance (fishID slope) for each individual. This also allows us to use the random effect intercept for fishID as a measure of behavioral vigor. The second way we could calculate these metrics are a measure of change in behavior between two points of behavior (often referred to as point tolerance, CITATION FOR THIS). This is calculated in the second half of the code as the change in behavior between pre-infection and early, late, and later points of infection.

```
# Calulating the behavioral tolerance metrics were interested in for our
# analysis

# Random slope from random effects model Random effects model with random slope
# for each fish by worm burden prior to measuring its activity. This slope is
# the tolerance an individual has to changing its behavior as infection
# increases. The slope of the behavior is the individuals behavioral vigor, or
# its pre-infection behavior#####
BehavTolLM <- lmer(Velocity ~ (Wormbf | fishID), IndBehav2)
summary(BehavTolLM)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Velocity ~ (Wormbf | fishID)
##    Data: IndBehav2
##
## REML criterion at convergence: 1254
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7050 -0.6143 -0.0137  0.5728  4.4818
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  fishID   (Intercept) 5.583e-01 0.747194
##           Wormbf      1.112e-05 0.003334 -0.41
##  Residual             5.583e-01 0.747187
## Number of obs: 493, groups:  fishID, 48
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   1.9022     0.1076   17.67
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 5.11298 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

```
##### Extracting the intercept and slope from the linear mixed model####
betaTol <- data.frame(coef(BehavTolLM)$fishID)
##### Renaming the slope and intercept to align with our metrics#####
betaTol <- betaTol %>%
    rename(BehavTol = Wormbf, BehavVig = X.Intercept.)
```

```
## rename: renamed 2 variables (BehavTol, BehavVig)
```

```
##### Creating a fishID column for combination with future dataframe####
betaTol <- cbind(fishID = rownames(betaTol), betaTol)

# Calculating the mean and variance in activity per day, per fish

# We need to split the dataframes into the differnet time periods to calculate
# averages since mutate isnt splitting by TrialTime.

##### creating a dataframe to calculate the mean and variance of activity for
##### before measurements.####

InBehavBefore <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Before") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining
```

```
## filter (grouped): removed 750 rows (81%), 180 rows remaining

## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)

## mutate (grouped): new variable 'AvgVel' (double) with 91 unique values and 0% NA

##                   new variable 'VarVel' (double) with 90 unique values and 1% NA

## distinct (grouped): removed 89 rows (49%), 91 rows remaining
```

```
##### creating a dataframe to calculate the mean and variance of activity for
##### early measurements.#####

InBehavEarly <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Early") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining

## filter (grouped): removed 670 rows (72%), 260 rows remaining

## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)

## mutate (grouped): new variable 'AvgVel' (double) with 87 unique values and 0% NA

##                   new variable 'VarVel' (double) with 87 unique values and 0% NA

## distinct (grouped): removed 173 rows (67%), 87 rows remaining
```

```
##### creating a dataframe to calculate the mean and variance of activity for
##### Late measurements. #####

InBehavLate <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Late") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining

## filter (grouped): removed 671 rows (72%), 259 rows remaining
```

```
## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)

## mutate (grouped): new variable 'AvgVel' (double) with 87 unique values and 0% NA

##                   new variable 'VarVel' (double) with 87 unique values and 0% NA

## distinct (grouped): removed 172 rows (66%), 87 rows remaining
```

```
##### creating a dataframe to calculate the mean and variance of activity for
##### Later measurements.####

InBehavLater <- IndBehav2 %>%
    drop_na(Velocity) %>%
    filter(TrialTime == "Later") %>%
    select(-c(InfDiff, Fish, Behavdate, Duration, Distance, comp3, Totworm, rate05,
        rate012, rate512, rate1216, rate18end, TOD, day, linf, maxlinf, maxworm,
        peakday, Ddate, Rdate, CountDay)) %>%
    mutate(AvgVel = mean(Velocity), VarVel = var(Velocity)) %>%
    distinct(fishID, .keep_all = TRUE)
```

```
## drop_na (grouped): removed 82 rows (8%), 930 rows remaining

## filter (grouped): removed 699 rows (75%), 231 rows remaining

## select: dropped 21 variables (Fish, CountDay, day, Totworm, comp3, ...)

## mutate (grouped): new variable 'AvgVel' (double) with 79 unique values and 0% NA

##                   new variable 'VarVel' (double) with 79 unique values and 0% NA

## distinct (grouped): removed 152 rows (66%), 79 rows remaining
```

```
#### Adding these dataframes together to create one large one####

IndBehav3 <- rbind(InBehavBefore, InBehavEarly, InBehavLate, InBehavLater)


##### Pivoting wider so I can subtract activity changes between periods to get
##### change in behavior####
IndBehav4 <- IndBehav3 %>%
    group_by(TrialTime, fishID, add = TRUE) %>%
    select(-Velocity) %>%
    pivot_wider(names_from = TrialTime, values_from = c(Wormbf, Ratebf, AvgVel, VarVel))
```

```
## group_by: 2 grouping variables (fishID, TrialTime)

## select: dropped one variable (Velocity)

## pivot_wider: reorganized (Wormbf, TrialTime, Ratebf, AvgVel, VarVel) into (Wormbf_Before, Wormbf_Earl
```

```r
##### subset down to infected individuals only and calculate point tolerance
##### metrics for each timeframe. #####
IndBehav4Inf <- IndBehav4 %>%
    filter(Infection == 1) %>%
    mutate(EChBe = AvgVel_Early - AvgVel_Before, LChBe = AvgVel_Late - AvgVel_Before,
        LtrChBe = AvgVel_Later - AvgVel_Before, ERatebf = (Wormbf_Early - Wormbf_Before)/6,
        LRatebf = (Wormbf_Late - Wormbf_Early)/6, LtrRatebf = (Wormbf_Later - Wormbf_Late)/6) %>%
    pivot_longer(cols = starts_with("Wormbf"), names_to = "TrialTime", names_prefix = "Wormbf_",
        values_to = "Wormbf") %>%
    select(-c(AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later, Ratebf_Before,
        Ratebf_Early, Ratebf_Late, Ratebf_Later)) %>%
    filter(fishID != c(44, 118, 139, 94))
```

```
## filter (grouped): removed 39 rows (43%), 52 rows remaining

## mutate (grouped): new variable 'EChBe' (double) with 49 unique values and 8% NA

##                   new variable 'LChBe' (double) with 49 unique values and 8% NA

##                   new variable 'LtrChBe' (double) with 42 unique values and 21% NA

##                   new variable 'ERatebf' (double) with 23 unique values and 10% NA

##                   new variable 'LRatebf' (double) with 42 unique values and 10% NA

##                   new variable 'LtrRatebf' (double) with 37 unique values and 23% NA

## pivot_longer: reorganized (Wormbf_Before, Wormbf_Early, Wormbf_Late, Wormbf_Later) into (TrialTime, W

## select: dropped 8 variables (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later, AvgVel_Before,

## filter (grouped): removed 4 rows (2%), 204 rows remaining
```

```r
##### Make two dataframes with pivot longer for wormbf and activity then merge
##### them together on fishID and TrialTime####
IndBehav4InfV <- IndBehav4 %>%
    filter(Infection == 1) %>%
    pivot_longer(cols = starts_with("AvgVel"), names_to = "TrialTime", names_prefix = "AvgVel_",
        values_to = "AvgVel") %>%
    select(fishID, AvgVel, TrialTime) %>%
    filter(fishID != c(44, 118, 139, 94))
```

```
## filter (grouped): removed 39 rows (43%), 52 rows remaining

## pivot_longer: reorganized (AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later) into (TrialTime, A

## select: dropped 25 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

## filter (grouped): removed 4 rows (2%), 204 rows remaining
```

```
##### Making the growth rate columns into one column for infected
##### individuals####
IndBehavGRInf <- IndBehav4 %>%
    filter(Infection == 1) %>%
    pivot_longer(cols = starts_with("Ratebf"), names_to = "TrialTime", names_prefix = "Ratebf_",
        values_to = "Ratebf") %>%
    select(fishID, TrialTime, Ratebf)
```

```
## filter (grouped): removed 39 rows (43%), 52 rows remaining
```

```
## pivot_longer: reorganized (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later) into (TrialTime, R
```

```
## select: dropped 25 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)
```

```
##### Making the new growth rate columns into one column for infected
##### individuals#####
IndBehavNGRInf <- IndBehav4Inf %>%
    select(fishID, ERatebf, LRatebf, LtrRatebf) %>%
    rename(NRatebf_Early = ERatebf, NRatebf_Late = LRatebf, NRatebf_Later = LtrRatebf) %>%
    add_column(NRatebf_Before = 0) %>%
    pivot_longer(cols = starts_with("NRatebf"), names_to = "TrialTime", names_prefix = "NRatebf_",
        values_to = "NRatebf") %>%
    select(fishID, TrialTime, NRatebf) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

```
## select: dropped 22 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)
```

```
## rename: renamed 3 variables (NRatebf_Early, NRatebf_Late, NRatebf_Later)
```

```
## pivot_longer: reorganized (NRatebf_Early, NRatebf_Late, NRatebf_Later, NRatebf_Before) into (TrialTi
```

```
## select: no changes
```

```
## distinct (grouped): removed 608 rows (75%), 208 rows remaining
```

```
##### Making the Variance columns into one column for infected individuals#####
IndBehavVarInf <- IndBehav4Inf %>%
    select(-TrialTime) %>%
    pivot_longer(cols = starts_with("VarVel"), names_to = "TrialTime", names_prefix = "VarVel_",
        values_to = "VarVel") %>%
    select(fishID, TrialTime, VarVel) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

```
## select: dropped one variable (TrialTime)
```

```
## pivot_longer: reorganized (VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later) into (TrialTime, V
```

```
## select: dropped 20 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)
```

```
## distinct (grouped): removed 608 rows (75%), 208 rows remaining
```

```
##### Merging the tolerance and vigor metrics ####
IndBehav4Inf1 <- merge(IndBehav4Inf, betaTol, by.x = "fishID", by.y = "fishID", .keep = all)

##### Merge the two dataframes together based on the fishID and TrialTime to
##### get all variables we want into one dataframe#####
IndBehav5Inf <- merge(IndBehav4InfV, IndBehav4Inf1, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Inf2 <- merge(IndBehav5Inf, IndBehavGRInf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Inf3 <- merge(IndBehav5Inf2, IndBehavNGRInf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Inf4 <- merge(IndBehav5Inf3, IndBehavVarInf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
# Subset down to uninfected individuals and calculate their point tolerance
# metrics Creating a dataframe for calculating the tolerance metrics and then
# pivot longer by Wormbf####
IndBehav4Unf <- IndBehav4 %>%
    filter(Infection == 0) %>%
    mutate(EChBe = AvgVel_Early - AvgVel_Before, LChBe = AvgVel_Late - AvgVel_Before,
        LtrChBe = AvgVel_Later - AvgVel_Before, NRatebf_Early = (Wormbf_Early - Wormbf_Before)/5,
        NRatebf_Late = (Wormbf_Late - Wormbf_Early)/5, NRatebf_Later = (Wormbf_Later -
            Wormbf_Late)/5, ) %>%
    pivot_longer(cols = starts_with(c("Wormbf")), names_to = "TrialTime", names_prefix = c("Wormbf_"),
        values_to = "Wormbf") %>%
    select(-c(AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later, Ratebf_Before,
        Ratebf_Early, Ratebf_Late, Ratebf_Later))
```

```
## filter (grouped): removed 52 rows (57%), 39 rows remaining

## mutate (grouped): new variable 'EChBe' (double) with 39 unique values and 0% NA

##                   new variable 'LChBe' (double) with 39 unique values and 0% NA

##                   new variable 'LtrChBe' (double) with 39 unique values and 3% NA

##                   new variable 'NRatebf_Early' (double) with one unique value and 100% NA

##                   new variable 'NRatebf_Late' (double) with one unique value and 100% NA

##                   new variable 'NRatebf_Later' (double) with one unique value and 100% NA

## pivot_longer: reorganized (Wormbf_Before, Wormbf_Early, Wormbf_Late, Wormbf_Later) into (TrialTime, W

## select: dropped 8 variables (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later, AvgVel_Before,
```

```
##### Creating a second dataframe where we pivot longer for velocity####
IndBehav4UnfV <- IndBehav4 %>%
    filter(Infection == 0) %>%
    pivot_longer(cols = starts_with("AvgVel"), names_to = "TrialTime", names_prefix = "AvgVel_",
        values_to = "AvgVel") %>%
    select(fishID, TrialTime, AvgVel)
```

```
## filter (grouped): removed 52 rows (57%), 39 rows remaining

## pivot_longer: reorganized (AvgVel_Before, AvgVel_Early, AvgVel_Late, AvgVel_Later) into (TrialTime, A

## select: dropped 25 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)
```

```
##### Making the growth rate columns into one column for uninfected
##### individuals#####
IndBehavGRUnf <- IndBehav4 %>%
    filter(Infection == 0) %>%
    pivot_longer(cols = starts_with("Ratebf"), names_to = "TrialTime", names_prefix = "Ratebf_",
        values_to = "Ratebf") %>%
    select(fishID, TrialTime, Ratebf)
```

```
## filter (grouped): removed 52 rows (57%), 39 rows remaining

## pivot_longer: reorganized (Ratebf_Before, Ratebf_Early, Ratebf_Late, Ratebf_Later) into (TrialTime, R

## select: dropped 25 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)
```

```
##### Making the new growth rate columns into one column for infected
##### individuals#####
IndBehavNGRUnf <- IndBehav4Unf %>%
    select(-TrialTime) %>%
    add_column(NRatebf_Before = 0) %>%
    pivot_longer(cols = starts_with("NRatebf"), names_to = "TrialTime", names_prefix = "NRatebf_",
        values_to = "NRatebf") %>%
    select(fishID, TrialTime, NRatebf) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

```
## select: dropped one variable (TrialTime)

## pivot_longer: reorganized (NRatebf_Early, NRatebf_Late, NRatebf_Later, NRatebf_Before) into (TrialTi

## select: dropped 21 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

## distinct (grouped): removed 468 rows (75%), 156 rows remaining
```

```
##### Making the Variance columns into one column for infected individuals#####
IndBehavVarUnf <- IndBehav4Unf %>%
    select(-TrialTime) %>%
    pivot_longer(cols = starts_with("VarVel"), names_to = "TrialTime", names_prefix = "VarVel_",
        values_to = "VarVel") %>%
    select(fishID, TrialTime, VarVel) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

```
## select: dropped one variable (TrialTime)

## pivot_longer: reorganized (VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later) into (TrialTime, V

## select: dropped 20 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

## distinct (grouped): removed 468 rows (75%), 156 rows remaining
```

```r
##### Merge the different datasets together#####
IndBehav5Unf <- merge(IndBehav4UnfV, IndBehav4Unf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Unf2 <- merge(IndBehav5Unf, IndBehavGRUnf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Unf3 <- merge(IndBehav5Unf2, IndBehavNGRUnf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)
IndBehav5Unf4 <- merge(IndBehav5Unf3, IndBehavVarUnf, by.x = c("fishID", "TrialTime"),
    by.y = c("fishID", "TrialTime"), all = TRUE)

# Creating the BehavTol and BehavVig columns for the uninfected and calculating
# the metric for uninfected individual so we can combine these dataframes
# together
IndBehav5Unf4 <- IndBehav5Unf4 %>%
    group_by(fishID) %>%
    add_column(BehavTol = 0, BehavVig = 0)
```

```
## group_by: one grouping variable (fishID)
```

```r
# Rbinding our infected and uninfected dataframes together into one overall
# dataframe
IndBehav5 <- rbind(IndBehav5Unf4, IndBehav5Inf4)

## Pivot longer for Change in behavior
IndBehav5ChBe <- IndBehav5 %>%
    select(-TrialTime) %>%
    rename(ChBehav_Early = EChBe, ChBehav_Late = LChBe, ChBehav_Later = LtrChBe) %>%
    add_column(ChBehav_Before = 0) %>%
    pivot_longer(cols = starts_with("ChBehav"), names_to = "TrialTime", names_prefix = "ChBehav_",
        values_to = "ChBehav") %>%
    select(fishID, TrialTime, ChBehav) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE)
```

```
## select: dropped one variable (TrialTime)

## rename: renamed 3 variables (ChBehav_Early, ChBehav_Late, ChBehav_Later)

## pivot_longer: reorganized (ChBehav_Early, ChBehav_Late, ChBehav_Later, ChBehav_Before) into (TrialTi

## select: dropped 30 variables (AvgVel, Sex, PreWeight, PreLength, Treatment, ...)

## distinct (grouped): removed 1,092 rows (75%), 364 rows remaining
```

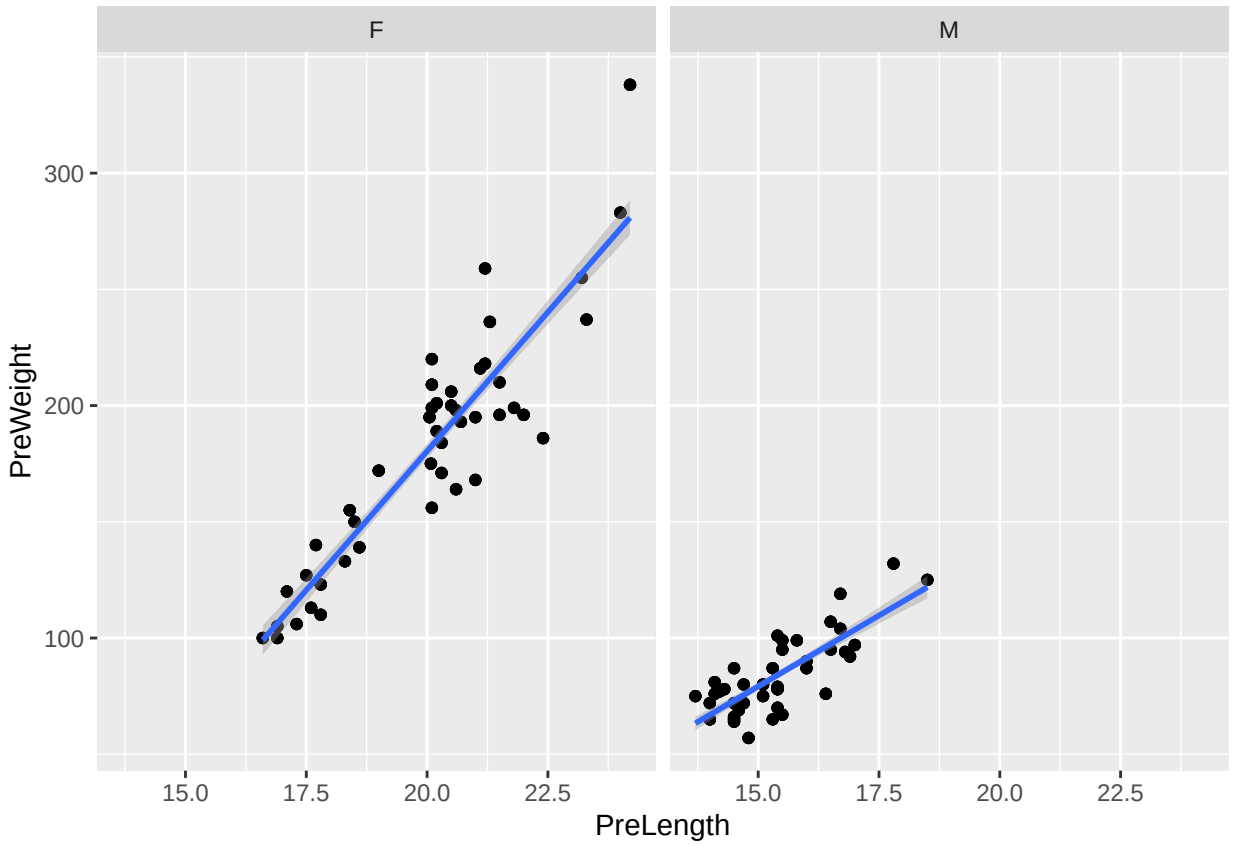
```r
# Merge dataframes together for one large dataframe
IndBehav6 <- merge(IndBehav5, IndBehav5ChBe, by.x = c("fishID", "TrialTime"), by.y = c("fishID",
    "TrialTime"), all = TRUE)

# Removing some of the columns we dont want in the dataframe Create a point
# tolerance metric for each timepoint
IndBehav6 <- IndBehav6 %>%
    select(-c(ERatebf, LRatebf, LtrRatebf, NRatebf_Early, NRatebf_Late, NRatebf_Later,
```

```
        VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later)) %>%
    mutate(PTol = ChBehav/Wormbf) %>%
    mutate(PTol = replace(PTol, PTol == NaN, 0)) %>%
    mutate(PTol = replace(PTol, PTol == -Inf, 0)) %>%
    mutate(PTol = replace(PTol, PTol == Inf, 0)) %>%
    select(-c(EChBe, LChBe, LtrChBe)) %>%
    distinct(fishID, TrialTime, .keep_all = TRUE) %>%
    drop_na(Sex)
```

```
## select: dropped 10 variables (VarVel_Before, VarVel_Early, VarVel_Late, VarVel_Later, NRatebf_Early,

## mutate: new variable 'PTol' (double) with 123 unique values and 63% NA

## mutate: no changes

## mutate: changed 8 values (2%) of 'PTol' (0 new NA)

## mutate: changed 7 values (2%) of 'PTol' (0 new NA)

## select: dropped 3 variables (EChBe, LChBe, LtrChBe)

## distinct: no rows removed

## drop_na: removed 17 rows (5%), 347 rows remaining
```

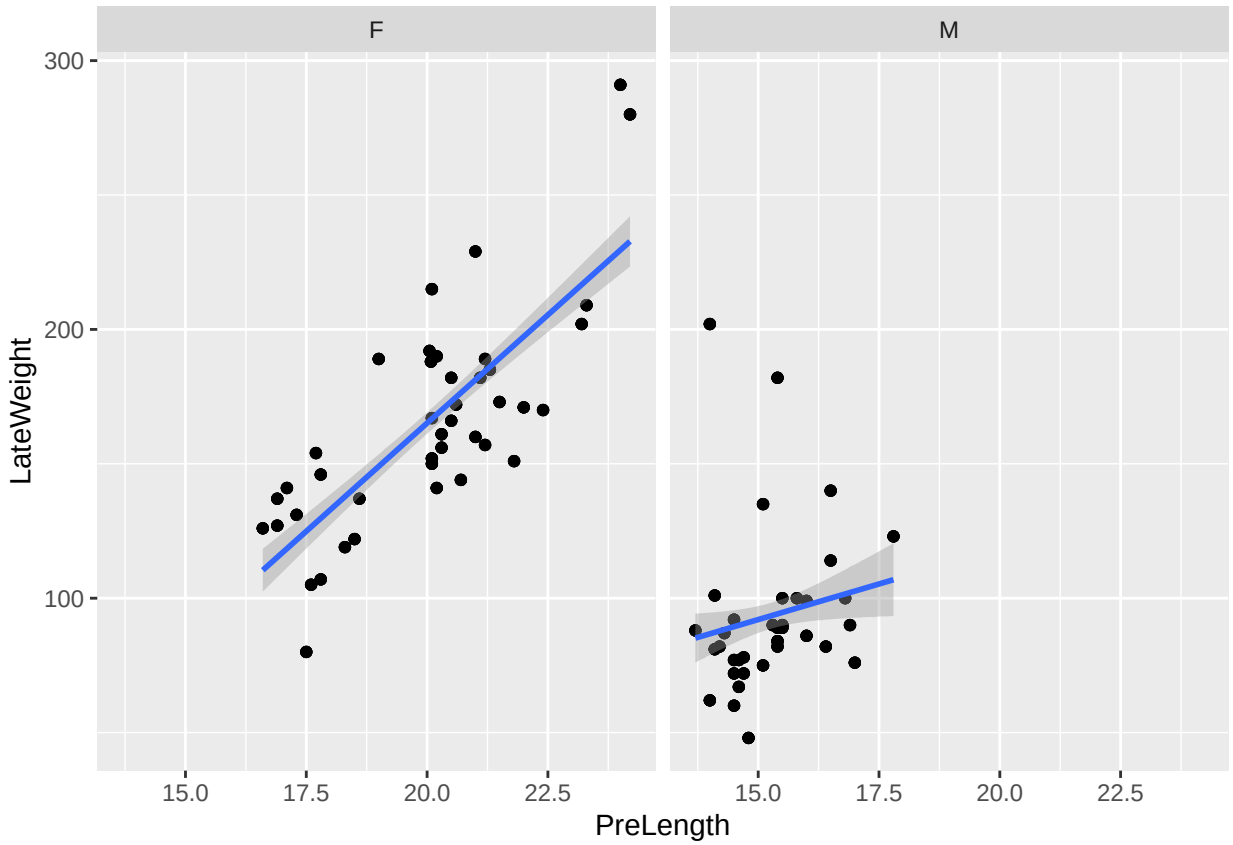
```
# Remove dataframes used for data wrangling and managing rm(c(IndBehav,
# IndBehav1,IndBehav2,IndBehav3,IndBehav3Inf1,IndBehav4,IndBehav4Inf,IndBehav4Inf1,IndBehav4InfV,IndBeh

# Setting some characters as factors
IndBehav6$TrialTime <- as.factor(IndBehav6$TrialTime)
IndBehav6$Sex <- as.factor(IndBehav6$Sex)
IndBehav6$Infection <- as.factor(IndBehav6$Infection)
IndBehav6$ContInf <- as.factor(IndBehav6$ContInf)
IndBehav6$Died <- as.factor(IndBehav6$Died)
IndBehav6$Treatment <- as.factor(IndBehav6$Treatment)

ggplot(IndBehav6, aes(x = PreLength, y = PreWeight)) + geom_point() + geom_smooth(method = "lm") +
    facet_wrap(~Sex)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
ggplot(IndBehav6, aes(x = PreLength, y = LateWeight)) + geom_point() + geom_smooth(method = "lm") +
    facet_wrap(~Sex)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# #Exporting dataframe for saving
write.csv(IndBehav6, "IndividualBehaviors_20240603.csv")
```

Now we have to decided how best to calculate whether an individual has recovered. In the experiment only had 13 fish recover from infection totally, therefore we are thinking of another metric that indicates a fish is in recovery during infection.

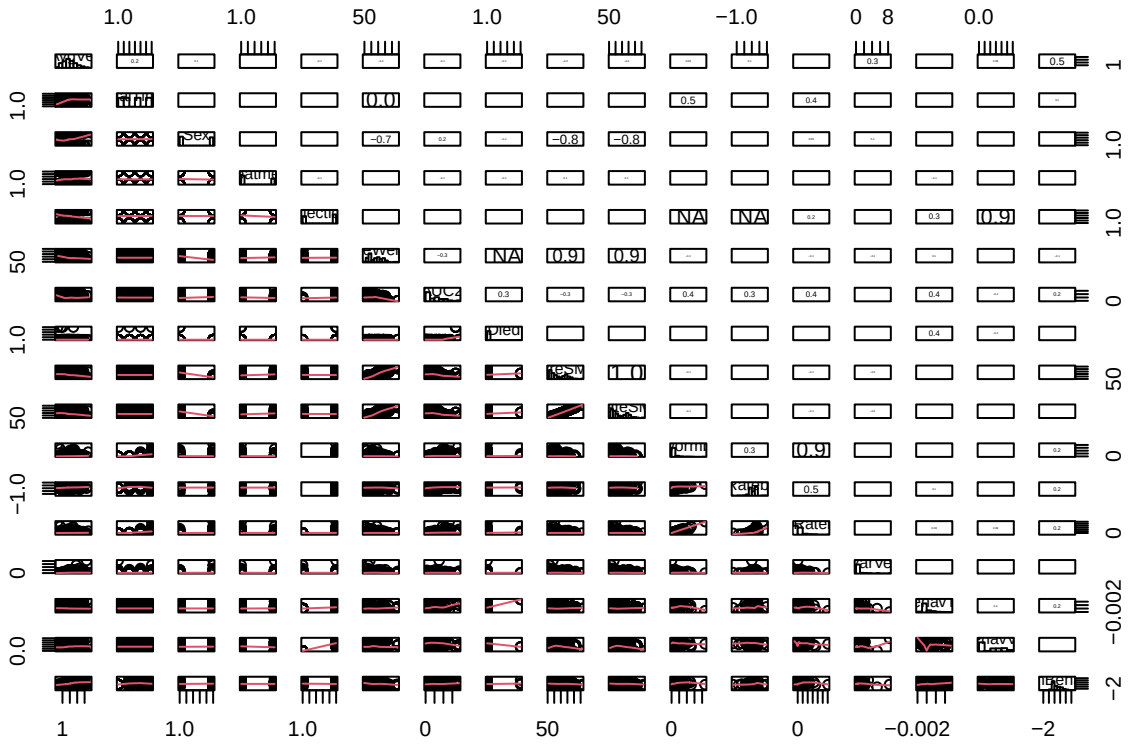
1.7 Description of data, structure, and type

fishID: The individual ID of each fish used in the trial.
TrialTime: The point of infection where each behavior trial happened. Early - early infection (6 days), Late - late in infection (12 days), Later - later in infection (18 days)
AvgVel: The average activity from three separate behavioral trials of each fish for each of the Trial Times. (cm/s)
Sex: The sex of the individual. F - female, M - male
PreWeight: The weight of the individual prior to their fist behavior trial and pre-infection. $(grams)$
PreLength: The length of the individual prior to their first behavior trial and pre-infection (mm)
Treatment: What treatment the fish received prior to their first behavior and pre-infection. VIE - visible implant elastomer implant, UNTOUCHED - control individual, received no injection nor implant (mm)
Infection: Whether or not the individual was infected with Gyrodactylus turnbulli. 1 - infected, 0 - uninfected
LateWeight: The weight of the individual after their final behavior trial and after infection. $(grams)$
wormJump: The number of worms that jumped from the donor fish to the trial fish during manual infections.
AUC2: The area under the curve of infection over the total infection trajectory for each individual.
RecovPeriod: The time frame in which an individual started recovering from infection. This is calculated

by looking at the growth rate of the worms between each count and see when the worm growth rate was decreasing overall.
ContInf: Whether or not the individual was controling infection. We qualified controling infection by having a negative growth rate post peak infection. 1 - Controlled, 0 - Uncontrolled
Died: Whether or not the individual Died from infection during the experimental trial. 1 - Died, 0 - did not died
PreSMI: The body condition of the individual prior to their first behavioral trial and pre infection. (mm/g)
LateSMI: The body condition of the individual after to their last behavioral trial and after infection. (mm/g)
Wormbf:The number of worms on the fish prior to each trial.
Ratebf:The rate of growth of worms on the fish prior to each trial from the worm count immediately prior to the trial.
NRatebf:The rate of growth of worms on the fish prior to each trial calculated as the change in worms between each time point. (For Example, From before to early).
VarVel:The variance in activity from the separate behavioral trials during each time point. ($(cm/s)^2$)
BehavTol:The behavioral tolerance of each individual calculated as the slope of a random effect model where fishID is the random effect term. ($cm/s/worms$)
BehavVig:The behavioral Vigor of each individual calculated as the activity of each individual's pre-infection measurements
ChBehav:The change in activity behavior between trial times. (cm/s)
PTol:Point tolerance metric calculated by dividing the change in behavior between time points by the worms on the fish before the trial ($cm/s/worms$)

1.8 Visualizing the relationships between variables in our dataset

```
pairs(~AvgVel + TrialTime + Sex + Treatment + Infection + LateWeight + AUC2 + Died +
    PreSMI + LateSMI + Wormbf + Ratebf + NRatebf + VarVel + BehavTol + BehavVig +
    ChBehav, lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav6)
```



1.9 Visualize some patterns in the raw data

```
# Load in dataset from above. This bit of code is meant to save time so people
# dont have to rerun the entire data parsing and calculating step above. Code
# will be saved for reproducibility.
IndBehav7 <- read_csv("IndividualBehaviors_20240501.csv")
```

```
## Rows: 347 Columns: 27
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (3): TrialTime, Sex, Treatment
## dbl (23): fishID, BehavGroup, AvgVel, BehavVig, VarvelBef, PreWeight, PreLen...
## lgl  (1): ContrPeriod
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# Setting some of the factors back to factors
IndBehav7$fishID <- as.factor(IndBehav7$fishID)
IndBehav7$TrialTime <- as.factor(IndBehav7$TrialTime)
IndBehav7$Sex <- as.factor(IndBehav7$Sex)
IndBehav7$Infection <- as.factor(IndBehav7$Infection)
IndBehav7$ContInf <- as.factor(IndBehav7$ContInf)
IndBehav7$Died <- as.factor(IndBehav7$Died)
IndBehav7$Treatment <- as.factor(IndBehav7$Treatment)


# Calculating tissue tolerance for each individual.
```

```
IndBehav7 <- IndBehav7 %>%
    mutate(ChSMI = LateSMI - PreSMI) %>%
    mutate(TisTol = ChSMI/Totworm)
```

```
## mutate: new variable 'ChSMI' (double) with 85 unique values and 0% NA
## mutate: new variable 'TisTol' (double) with 47 unique values and 46% NA
```

```
# Subsetting down to female and males only to scale their SMI Males
IndBehavM <- IndBehav7 %>%
    filter(Sex == "M") %>%
    mutate(ScPSMI = scale(PreSMI), ScLSMI = scale(LateSMI))
```

```
## filter: removed 180 rows (52%), 167 rows remaining
## mutate: new variable 'ScPSMI' (double) with 41 unique values and 0% NA
##         new variable 'ScLSMI' (double) with 41 unique values and 0% NA
```

```
# Females
IndBehavF <- IndBehav7 %>%
    filter(Sex == "F") %>%
    mutate(ScPSMI = scale(PreSMI), ScLSMI = scale(LateSMI))
```

```
## filter: removed 167 rows (48%), 180 rows remaining
## mutate: new variable 'ScPSMI' (double) with 45 unique values and 0% NA
##         new variable 'ScLSMI' (double) with 45 unique values and 0% NA
```

```
# Combine the two separate dataframes together
IndBehav8 <- rbind(IndBehavF, IndBehavM)

# Scaling some variables to make them biologically comparable and better for
# model fitting

IndBehav8 <- IndBehav8 %>%
    mutate(ScVarvelBef = c(scale(VarvelBef)), ScNRatebf = c(scale(NRatebf)), ScBehavTol = c(scale(Behav
        ScTotworm = c(scale(Totworm)), ScAUC = c(scale(AUC2)), ScBehavVig = c(scale(BehavVig)),
        ScChSMI = c(scale(ChSMI)), ScTisTol = c(scale(TisTol)))
```

```
## mutate: new variable 'ScVarvelBef' (double) with 86 unique values and 2% NA
##         new variable 'ScNRatebf' (double) with 85 unique values and 37% NA
##         new variable 'ScBehavTol' (double) with 49 unique values and 0% NA
##         new variable 'ScTotworm' (double) with 47 unique values and 46% NA
##         new variable 'ScAUC' (double) with 49 unique values and 43% NA
##         new variable 'ScBehavVig' (double) with 87 unique values and 1% NA
##         new variable 'ScChSMI' (double) with 85 unique values and 0% NA
##         new variable 'ScTisTol' (double) with 47 unique values and 46% NA
```

```
IndBehav8$fishID <- as.character(IndBehav8$fishID)
```

2 Calculating athe total area of a polygon based on the number of worms and the avergae activity of each individual

This metric is used as a combination as the tolerance (change in behavior) and resistance (the change in parasites over the course of infection).

2.1 What hypotheses we want to test with these data and what data we can use to test them?

For uninfected and infected individuals only:

Do infected and uninfected individuals differ in their average and variation in activity?

Is there sexual variation in the average and variation in activity?

Do we see differences in the change in behavior and between infected and uninfected indivdiauls?

For infected individuals only:

Does activity predict the intensity of infection?

Is there sexual variation in host behavioral tolerance and behavioral vigor?

Do host behavioral tolerance, tissue tolerance, and behavioral vigor correlate together and does this differ by sex?

Do hosts with lower behavioral tolerance have higher behavior resilience (i.e. how hosts behave once they have started to clear infection/have cleared infection)

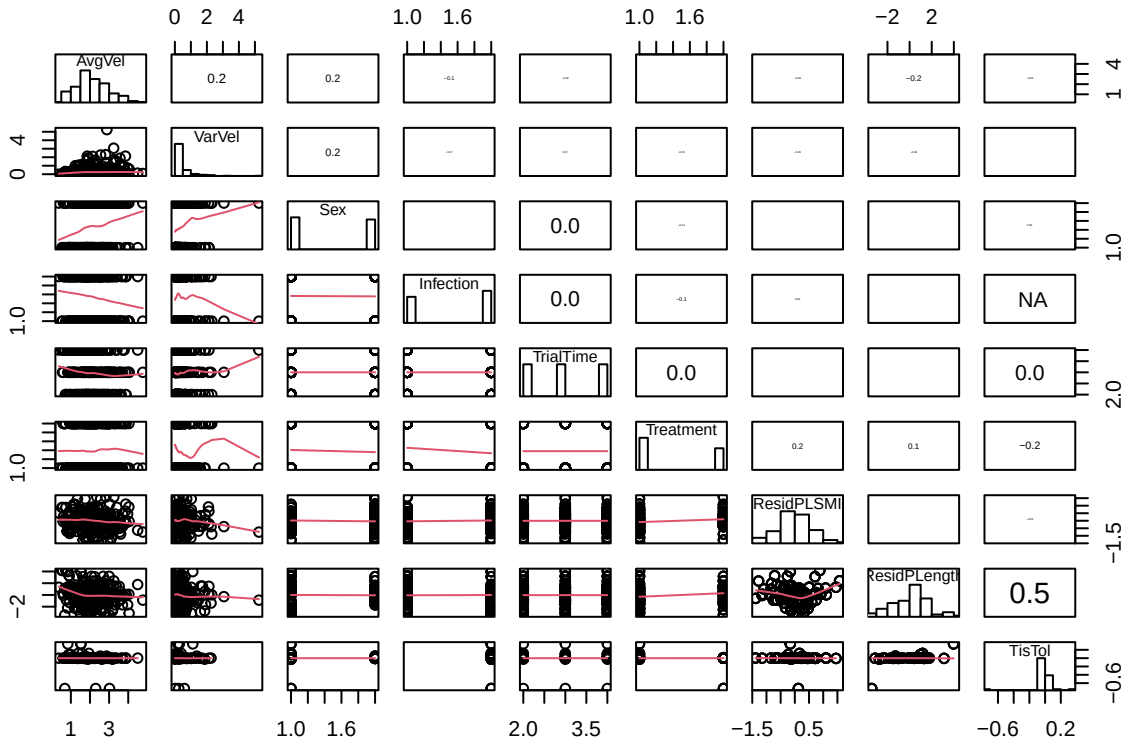
Is there sexual variation in individuals behavioral resilience?

polygon area as resilience and how this relates to other metrics

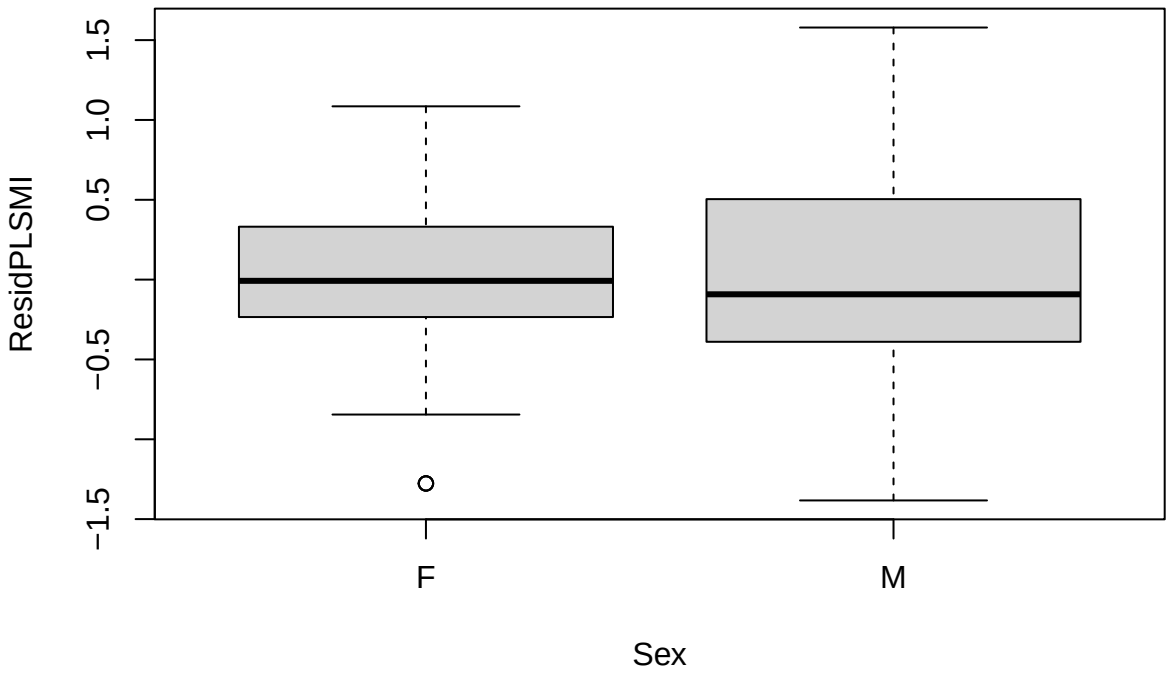
3 Do infected and uninfected individuals differ in their average and variation in activity?

3.1 Visually inspection of the explanatory variables that will be used in the analyses

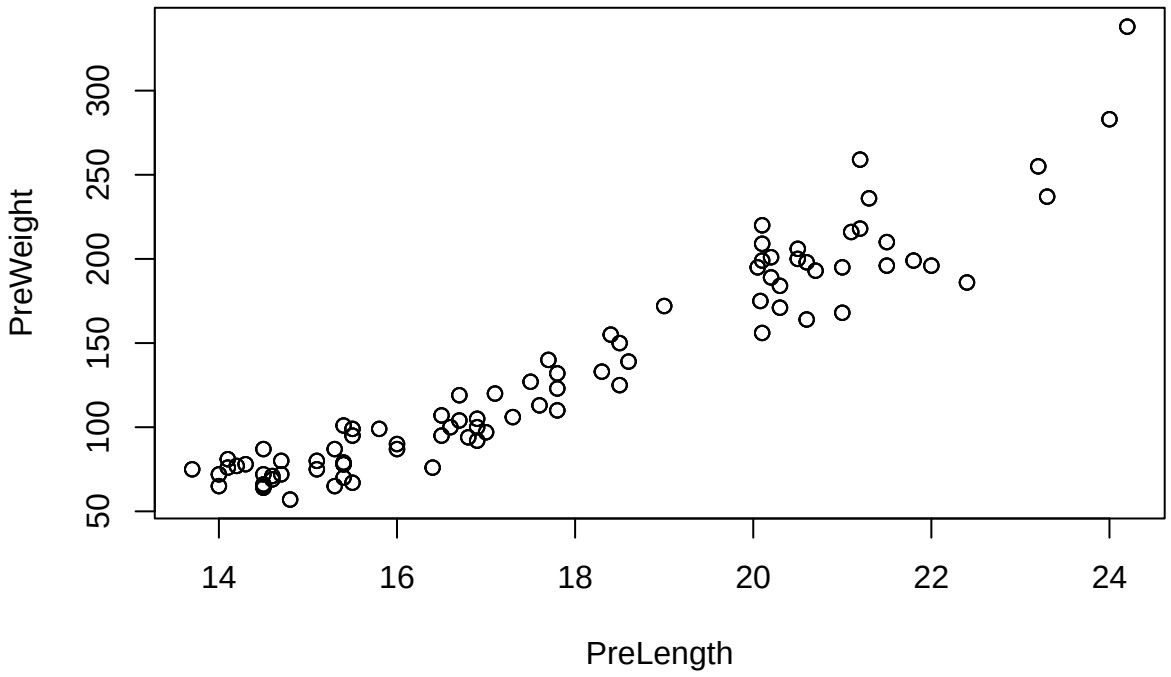
```
pairs(~AvgVel + VarVel + Sex + Infection + TrialTime + Treatment + ResidPLSMI + ResidPLength +
    TisTol, lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav8)
```




```
boxplot(ResidPLSMI ~ Sex, IndBehav8)
```




```
plot(PreWeight ~ PreLength, IndBehav8)
```



3.2 Does infection or sex variation impact the average activity of individauls?

Individuals had 3 behavioral trials per time period of infection (i.e. 3 behavioral trials before infection) and therefore using preliminary analysis we showed that there is no difference due to time of day of these recordings so we averaged and quantified the variance of the velocities for that day to get an average activity per trial time.

This analysis uses the average activity for each individual at each trial point.

3.2.1 Description, development, and fitting of linear model for the analysis

We will use a linear mixed model to analyze how average activity differs by infection status and sexual variation. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic
- $AvgVel_{det}$ = a + b_1TrialTime + b_2Infection * b_3Sex + b_4ScPSMI + b_5ScRPLength + b_6Treatment + a_i
- Stochastic
  - AvgVel ~ $N(AvgVel_{det}, \sigma^2)$
  - $a_i \sim N(0, \sigma^2_{fishID})$
- Fixed
  - TrialTime
  - Infection status
  - Sex
  - An interaction between Sex and Infection status
  - Scaled Pre-infection SMI

  - Scaled residuals from length and sex
  - VIE Treatment

- Random
  - fishID

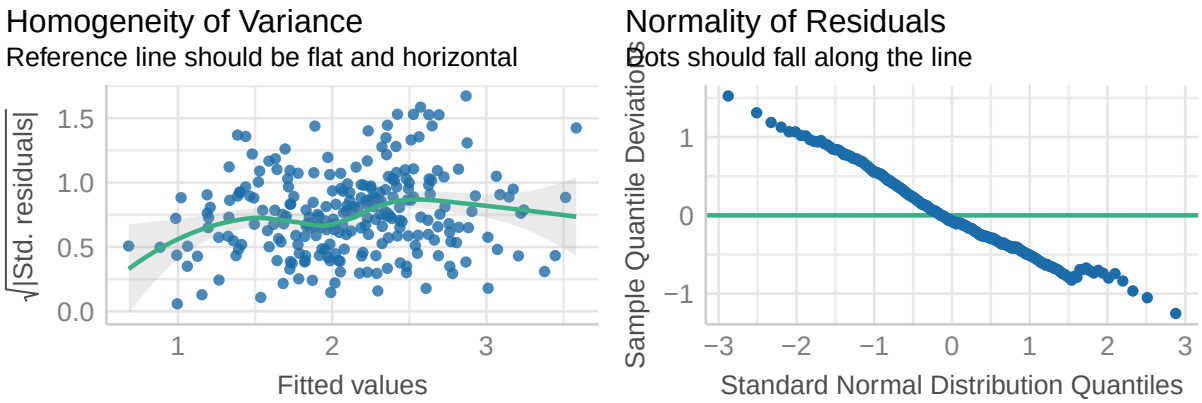
```
# Fit a linear model for checking what explanatory factors are important for
# Average activity Note this is a linear mixed model because we have multiple
# measures per fish and therefore, need to account for non-independence between
# measures.
AvgVelLM <- lmer(AvgVel ~ TrialTime + Infection * Sex + ScPSMI + ScRPLength + Treatment +
    (1 | fishID), IndBehav8)



# Summary to see the relationship of the variables.
summary(AvgVelLM)
```

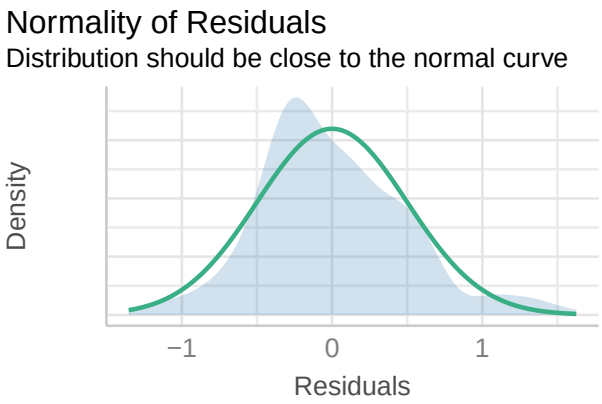
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: AvgVel ~ TrialTime + Infection * Sex + ScPSMI + ScRPLength +
##     Treatment + (1 | fishID)
##    Data: IndBehav8
##
## REML criterion at convergence: 561.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.3311 -0.5476 -0.1322  0.5362  2.7978
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  fishID   (Intercept) 0.3020   0.5496
##  Residual             0.3373   0.5808
## Number of obs: 250, groups:  fishID, 86
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)      2.01656    0.16934  11.908
## TrialTimeLate    0.13613    0.08857   1.537
## TrialTimeLater  -0.20354    0.09157  -2.223
## Infection1      -0.09947    0.19610  -0.507
## SexM             0.43701    0.20882   2.093
## ScPSMI          -0.11313    0.12255  -0.923
## ScRPLength      -0.04706    0.11945  -0.394
## TreatmentVIE     0.07066    0.14596   0.484
## Infection1:SexM -0.25897    0.28652  -0.904
##
## Correlation of Fixed Effects:
##             (Intr) TrilTmLt TrlTmLtr Infct1 SexM   ScPSMI ScRPLn TrtVIE
## TrialTimeLt -0.262
## TrialTimLtr -0.265  0.484
## Infection1  -0.675  0.000    0.005
## SexM        -0.611  0.000    0.005    0.528
```

```
## ScPSMI        0.063  0.000   -0.004   -0.009  0.046
## ScRPLength   -0.074  0.000    0.005    0.085  0.040 -0.803
## TreatmntVIE -0.419  0.000    0.001    0.088  0.028 -0.156  0.053
## Infctn1:SxM  0.444  0.000    0.013   -0.692 -0.742 -0.084 -0.037 -0.007
```

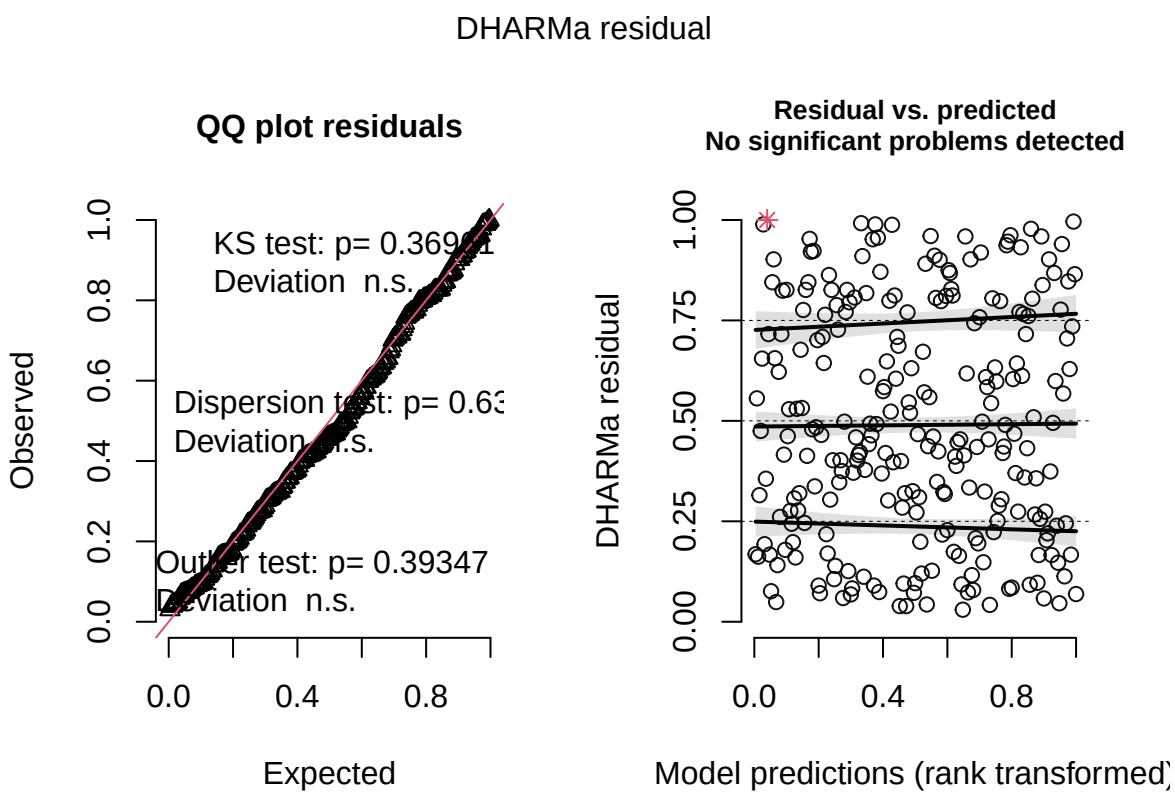
3.2.2 Validate that the model fits well and there are no problems

```
# Using the check_model function from the performmce package to check the
# model validation

check_model(AvgVelLM, check = c("qq", "normality", "homogeneity"))
```

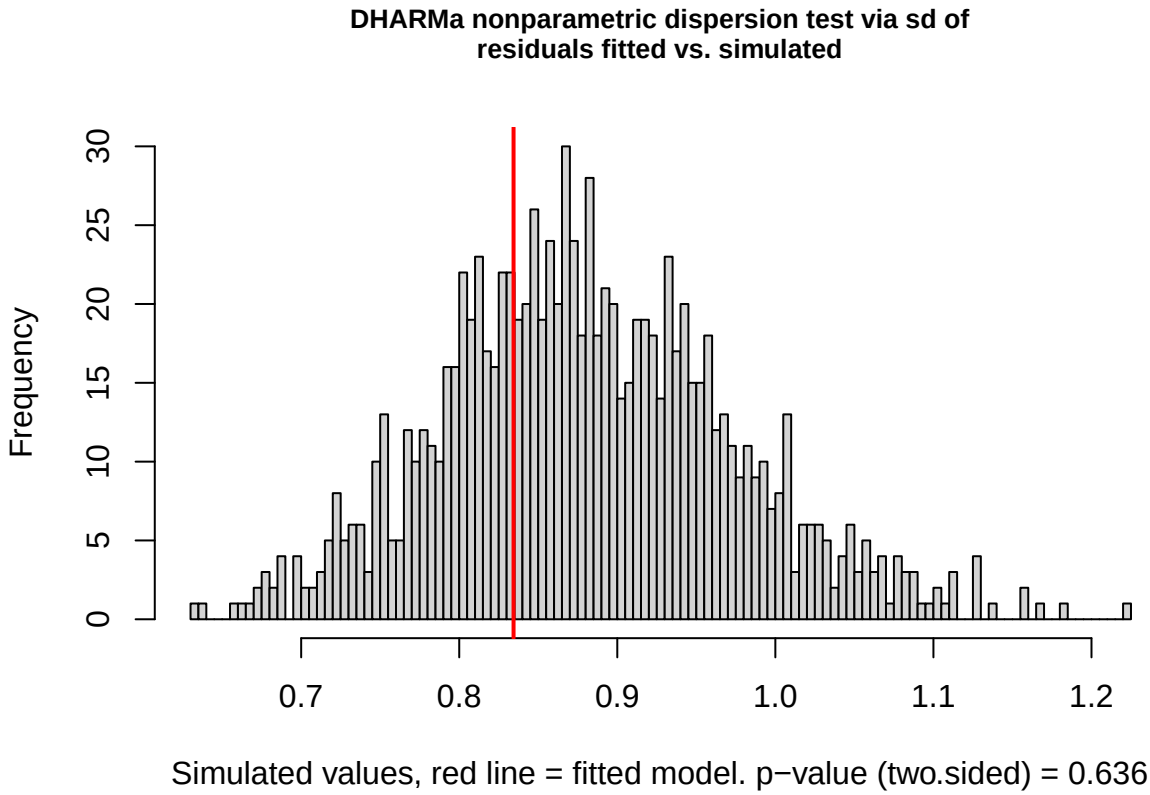





```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_AvgVelLM <- simulateResiduals(AvgVelLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_AvgVelLM)
```




```
# Testing for dispersion
testDispersion(sim_residuals_AvgVelLM)
```



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.94698, p-value = 0.636
## alternative hypothesis: two.sided
```

```
# All model validation looks good.
```

3.2.3 Testing the significance of factors in our model using a Kenward-Rodgers F test

```
# F test to test for signficance of slope of variables
Anova(AvgVelLM, test = "F", type = 3)
```

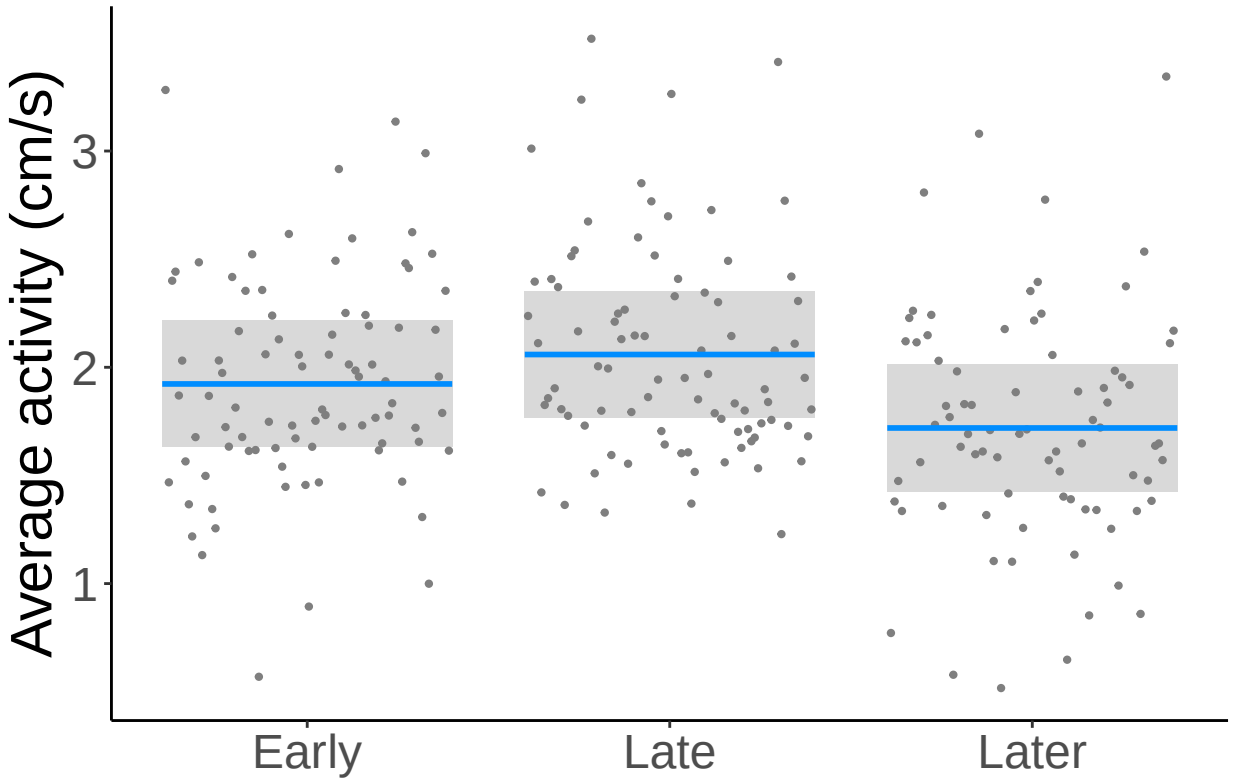
```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: AvgVel
##                       F Df  Df.res    Pr(>F)
## (Intercept)   141.8069  1  94.033 < 2.2e-16 ***
## TrialTime       6.9196  2 163.528  0.001304 **
## Infection       0.2573  1  77.719  0.613431
## Sex             4.3793  1  77.704  0.039642 *
## ScPSMI          0.8522  1  78.960  0.358746
## ScRPLength      0.1552  1  78.107  0.694694
## Treatment       0.2343  1  79.162  0.629676
## Infection:Sex   0.8169  1  78.805  0.368842
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.2.4 Visualize the important explanatory factors for average activity

```
# TrialTimeGraph
AvgVelbyTT = visreg(AvgVelLM, scale = "response", "TrialTime", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Average activity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22))
```

```
## Conditions used in construction of plot
## Infection: 1
## Sex: F
## ScPSMI: -0.07158284
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
## fishID: 1
```

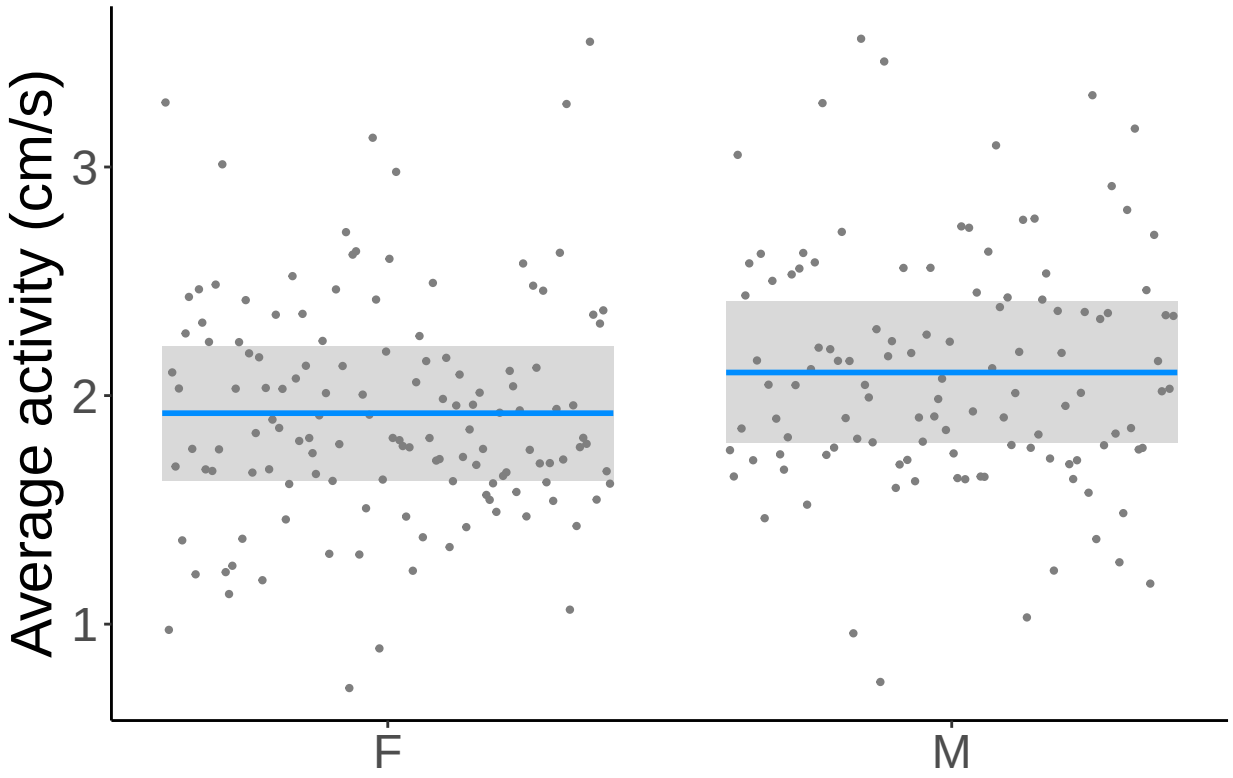
```
# Print the graph
print(AvgVelbyTT)
```




```
# TrialTimeGraph
AvgVelbySex = visreg(AvgVelLM, scale = "response", "Sex", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Average activity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22))
```

```
## Conditions used in construction of plot
## TrialTime: Early
## Infection: 1
## ScPSMI: -0.07158284
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
## fishID: 1
```

```
# Print the graph
print(AvgVelbySex)
```



3.3 Does infection or sex variation impact the variance in activity of individuals?

This analysis uses the variance of activity for each individual at each trial point.

3.3.1 Description, development, and fitting of linear model for the analysis

We will use a linear mixed model to analyze how variance in activity differs by infection status and sexual variation. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic
- $log(VarVel_{det}) = \text{a} + b_1\text{TrialTime} + b_2\text{Infection} + b_3\text{Sex} + b_4\text{ScPSMI} + b_5\text{ScRPLength} + b_6\text{Treatment} + a_i$
- Stochastic
  - $\log(\text{VarVel}) \sim N(log(VarVel_{det}), \sigma^2)$
  - $a_i \sim N(0, \sigma^2_{fishID})$
- Fixed
  - TrialTime
  - Infection
  - Sex

  - Scaled Pre-infection SMI
  - Scaled residuals from length and sex
  - VIE Treatment

- Random
  - fishID

3.3.2 Validate that the model fits well and there are no problems

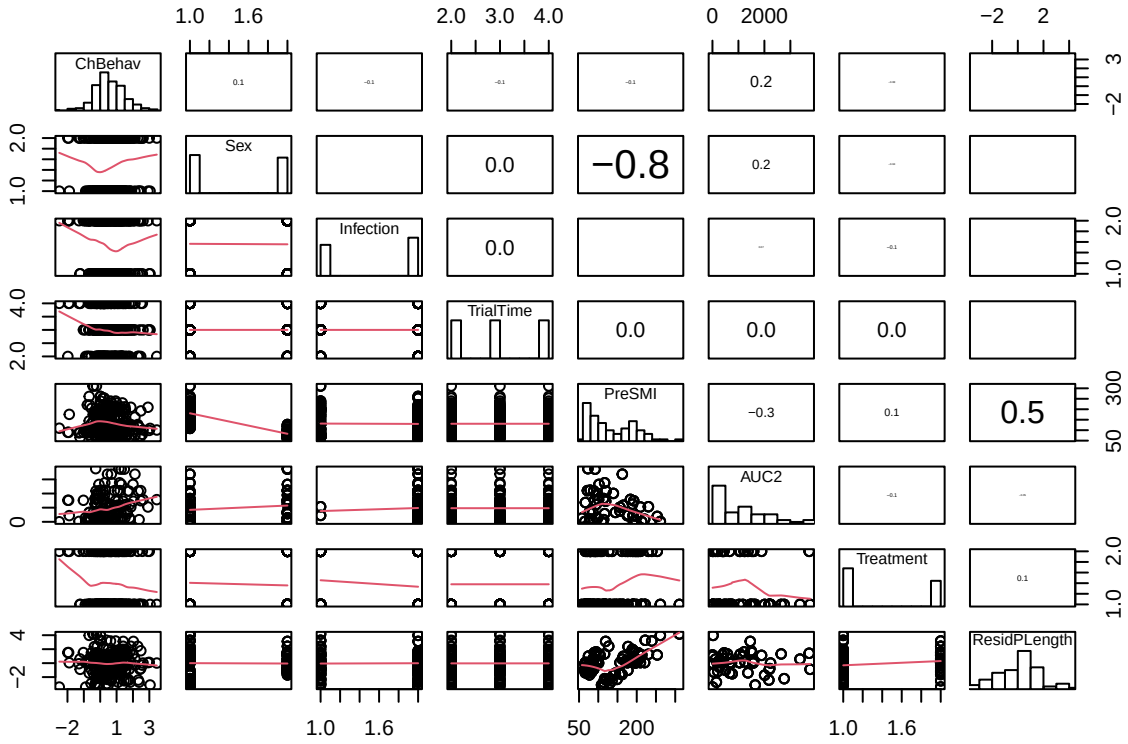
3.3.3 Testing the significance of factors in our model using a Kenward-Rodgers F test

3.3.4 Visualize the important explanatory factors for variance of activity

4 Do we see differences in change in behavior over time based on infection status and sexual variation?

4.1 Visually inspection of the explanatory variables that will be used in the analyses

```
pairs(~ChBehav + Sex + Infection + TrialTime + PreSMI + AUC2 + Treatment + ResidPLength,
    lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav8)
```



4.1.1 Description, development, and fitting of linear model for the analysis

We will use a linear mixed model to analyze how Change in activity differs by infection status and sexual variation. FishID is included as a random term to allow for non-independence of individuals due to multiple measurements per individual across time.

- Deterministic
- $ChVel_{det}$ = a + b_1TrialTime + b_2Infection * b_3Sex + b_4ScRPLength + b_5Treatment + a_i
- Stochastic
  - ChVel ~ $N(ChVel_{det}, \sigma^2)$
  - $a_i \sim N(0, \sigma^2_{fishID})$
- Fixed
  - TrialTime
  - Infection
  - Sex
  - Interaction between Sex and infection status
  - Scaled residuals from length and sex
  - VIE treatment of the fish
- Random
  - fishID

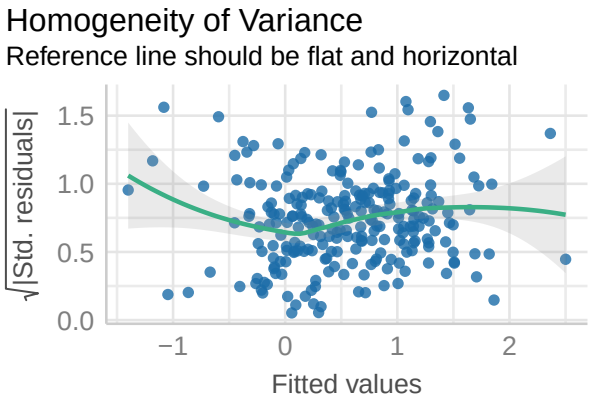
```
# Fit a linear model for checking what explanatory factors are important for
# Variance in activity Note this is a linear mixed model because we have
# multiple measures per fish and therefore, need to account for
# non-independence between measures.
ChVelLM <- lmer(ChBehav ~ TrialTime + Infection * Sex + ScChSMI + ScRPLength + Treatment +
    (1 | fishID), IndBehav8)

# Summary to see the relationship of the variables.
summary(ChVelLM)
```

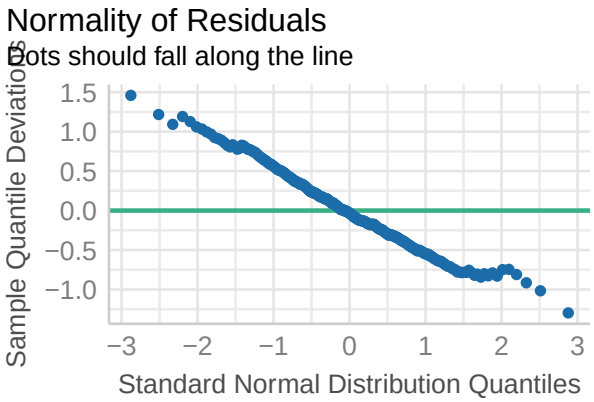
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: ChBehav ~ TrialTime + Infection * Sex + ScChSMI + ScRPLength +
##     Treatment + (1 | fishID)
##    Data: IndBehav8
##
## REML criterion at convergence: 589.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.43813 -0.54690 -0.06069  0.51319  2.71825
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  fishID   (Intercept) 0.4732   0.6879
##  Residual             0.3384   0.5817
## Number of obs: 250, groups:  fishID, 86
##
```

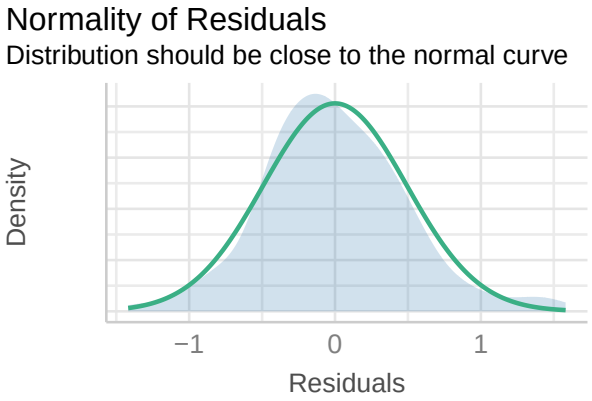
```
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)      0.74336    0.19952   3.726
## TrialTimeLate    0.13613    0.08871   1.535
## TrialTimeLater  -0.21812    0.09181  -2.376
## Infection1      -0.25721    0.23436  -1.097
## SexM             0.17036    0.24827   0.686
## ScChSMI         -0.12373    0.09156  -1.351
## ScRPLength       0.03524    0.09135   0.386
## TreatmentVIE    -0.21103    0.17270  -1.222
## Infection1:SexM  0.02620    0.33999   0.077
##
## Correlation of Fixed Effects:
##             (Intr) TrilTmLt TrlTmLtr Infct1 SexM   ScCSMI ScRPLn TrtVIE
## TrialTimeLt -0.222
## TrialTimLtr -0.225  0.483
## Infection1  -0.688  0.000    0.004
## SexM        -0.614  0.000    0.005    0.520
## ScChSMI     -0.117  0.000   -0.002    0.105 -0.051
## ScRPLength   0.008  0.000    0.003    0.080  0.138 -0.379
## TreatmntVIE -0.429  0.000    0.000    0.100  0.029  0.134 -0.164
## Infctn1:SxM  0.446  0.000    0.011   -0.682 -0.743  0.077 -0.192 -0.009
```

4.1.2 Validate that the model fits well and there are no problems

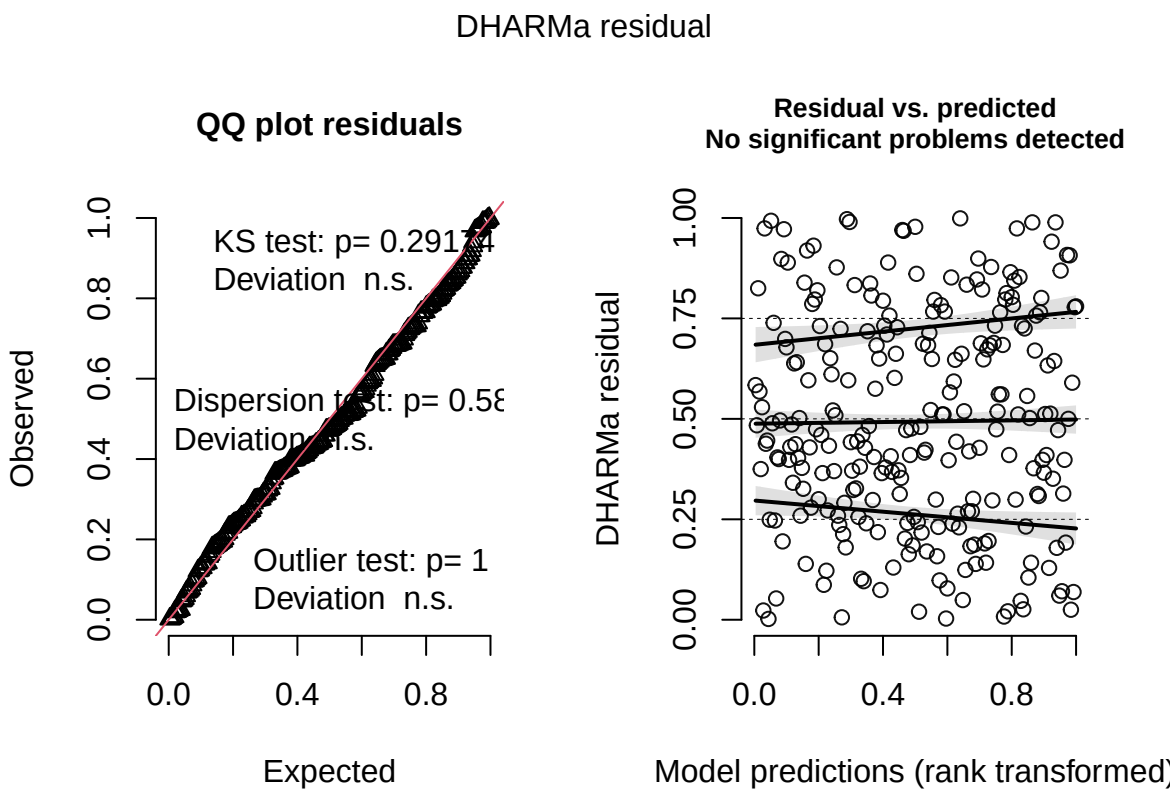
```
# Using the check_model function from the perforammce package to check the
# model validation

check_model(ChVelLM, check = c("qq", "normality", "homogeneity"))
```



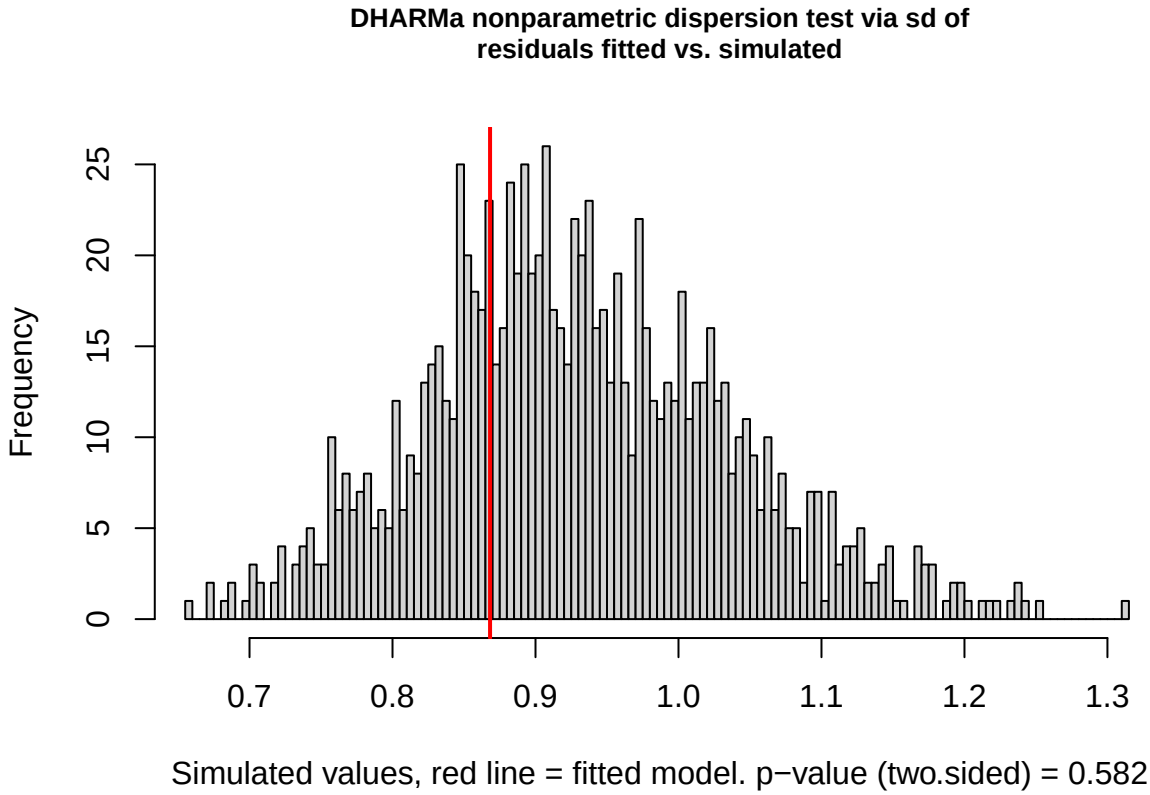





```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_ChVelLM <- simulateResiduals(ChVelLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_ChVelLM)
```




```
# Testing for dispersion
testDispersion(sim_residuals_ChVelLM)
```



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.93368, p-value = 0.582
## alternative hypothesis: two.sided
```

```
# There are some problems with this model validation. It doesnt look model
# breaking but definitely should look at other model error structures to
# resolve the issues.
```

4.1.3 Testing the significance of factors in our model using a Kenward-Rodgers F test

```
# F test to test for signficance of slope of variables
Anova(ChVelLM, test = "F", type = 3)
```

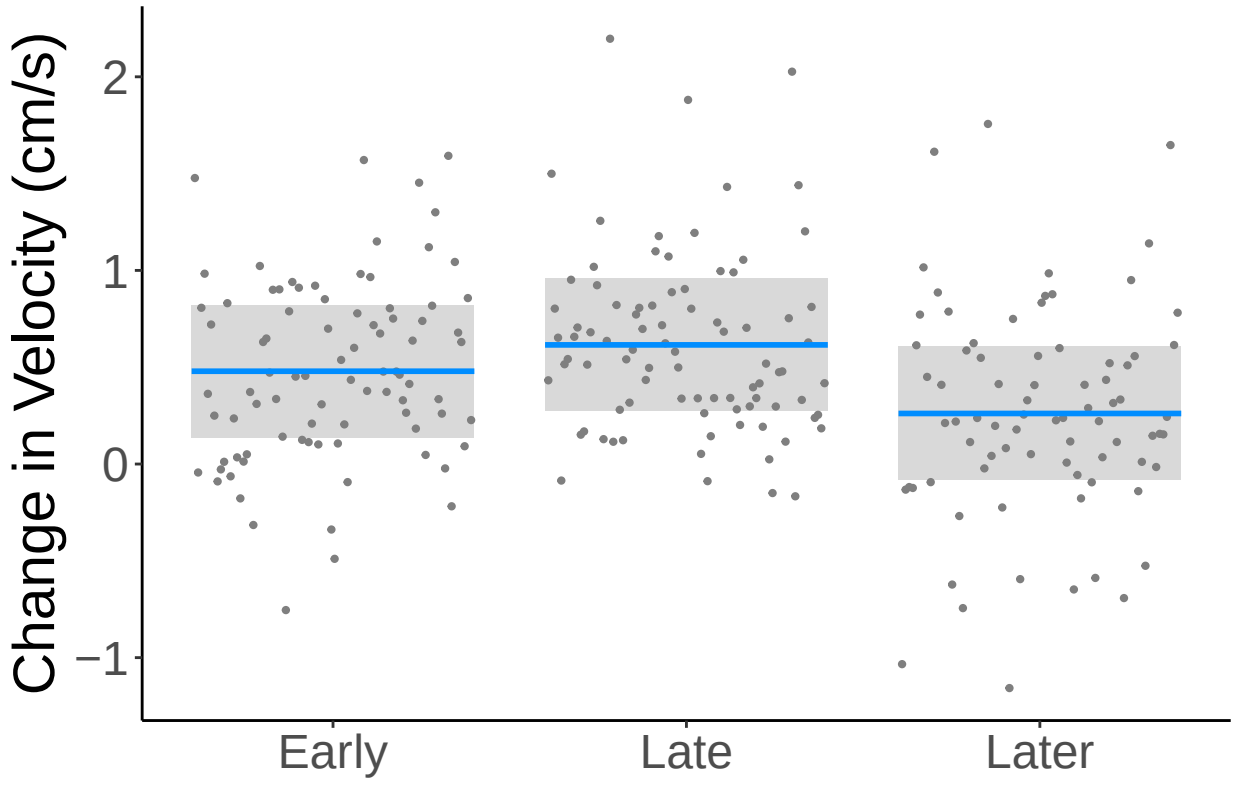
```
## Analysis of Deviance Table (Type III Wald F tests with Kenward-Roger df)
##
## Response: ChBehav
##                      F Df  Df.res    Pr(>F)
## (Intercept)   13.8809  1  89.549 0.0003404 ***
## TrialTime      7.5128  2 163.128 0.0007566 ***
## Infection      1.2045  1  78.084 0.2757989
## Sex            0.4708  1  78.095 0.4946356
## ScChSMI        1.8259  1  81.052 0.1803750
## ScRPLength     0.1488  1  80.227 0.7006941
## Treatment      1.4930  1  79.188 0.2253770
## Infection:Sex  0.0059  1  78.920 0.9387742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.1.4 Visualize the important explanatory factors for change in velocity

```
# TrialTimeGraph
ChVelbyTT = visreg(ChVelLM, scale = "response", "TrialTime", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Change in Velocity (cm/s)") +
    xlab(" ") + theme(text = element_text(size = 22))
```

```
## Conditions used in construction of plot
## Infection: 1
## Sex: F
## ScChSMI: 0.06939777
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
## fishID: 1
```

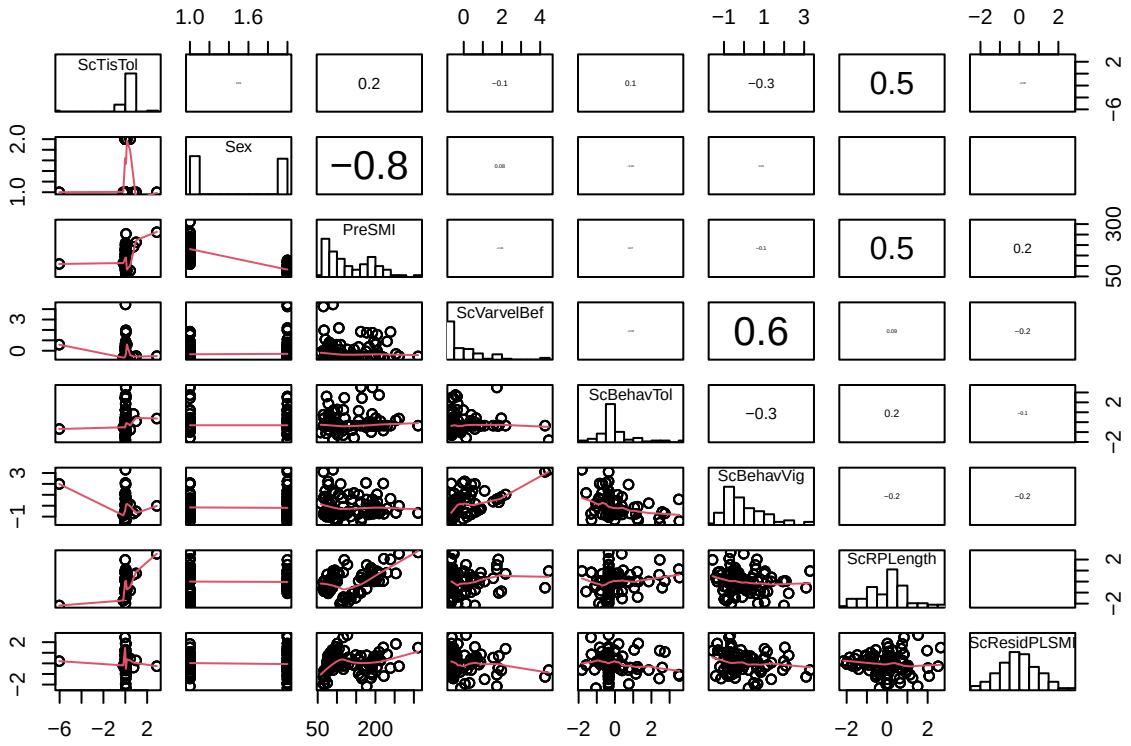
```
# Print the graph
print(ChVelbyTT)
```



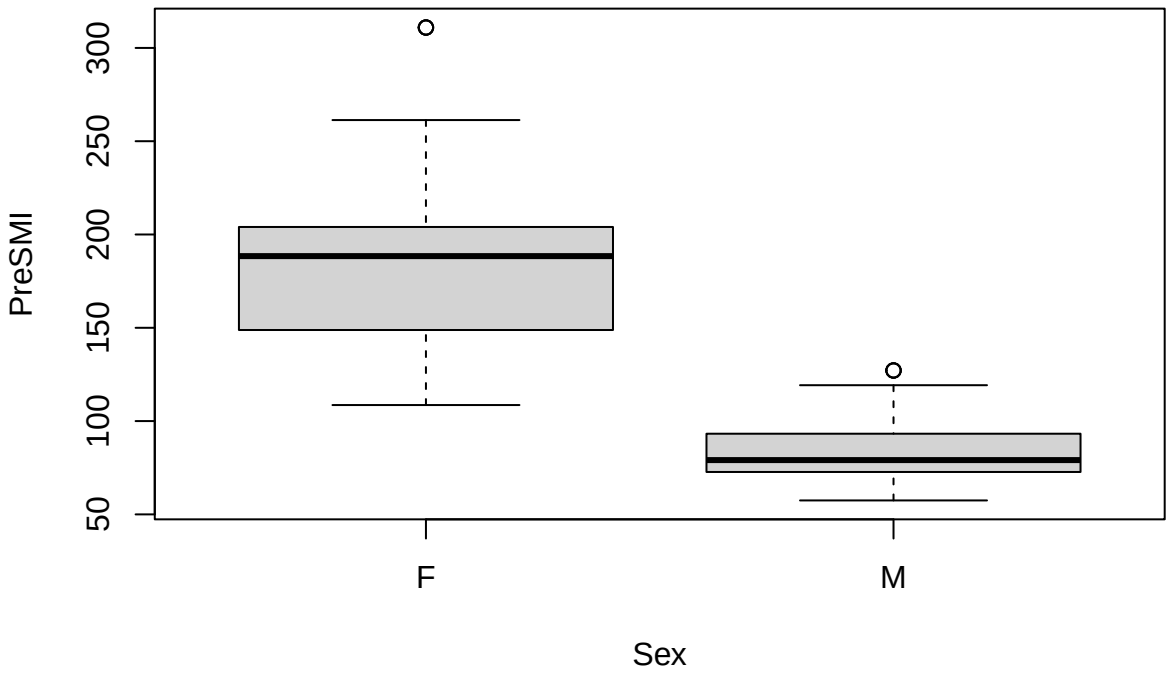
5 Is there sexual variation in host behavioral tolerance and behavioral vigor?

5.1 Visually inspection of the explanatory variables that will be used in the analyses

```
pairs(~ScTisTol + Sex + PreSMI + ScVarvelBef + ScBehavTol + ScBehavVig + ScRPLength +
    ScResidPLSMI, lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav8)
```




```
boxplot(PreSMI ~ Sex, IndBehav8)
```



5.2 What factors are important for host behavioral vigor and do is there any sexual variation in host behavioral vigor?

5.2.1 Description, development, and fitting of linear model for the analysis

We will use a linear model to analyze how behavioral vigor differs by sexual variation and other important host traits. Given each host only has one behavioral vigor measure, we do not need the fishID random effect used in previous models/

- Deterministic
- $BehavVig_{det}$ = a + b_1Sex + b_2ScResidPLSMI + b_3ScVarVelBef + b_4ScRPLength + b_5Treatment + b_6Sex:ScResidPLSMI + b_7Sex:ScRPLength + a_i
- Stochastic
  - BehavVig ~ $N(BehavVig_{det}, \sigma^2)$
  - $a_i \sim N(0, \sigma^2_{BehavGroup})$
- Fixed
  - Sex
  - Scaled residuals from Pre-infection SMI and length
  - Scaled variance in velocity before infection
  - Scaled residuals from length and sex
  - VIE treatment
  - An interaction between sex and Pre-infection SMI
  - An interaction between sex and Pre-infection Length
- Random
  - Behavior group of recording

```
# Fit a linear model for checking what explanatory factors are important for
# Variance in Velocity Note this is a linear mixed model because we have
# multiple measures per fish and therefore, need to account for
# non-independence between measures.
BehavVigLM <- glmmTMB(BehavVig ~ Sex * ScResidPLSMI + ScVarvelBef + ScRPLength +
    Treatment + Sex:ScRPLength + (1 | BehavGroup), family = Gamma("log"), IndBehav8)

# Summary to see the relationship of the variables.
summary(BehavVigLM)
```

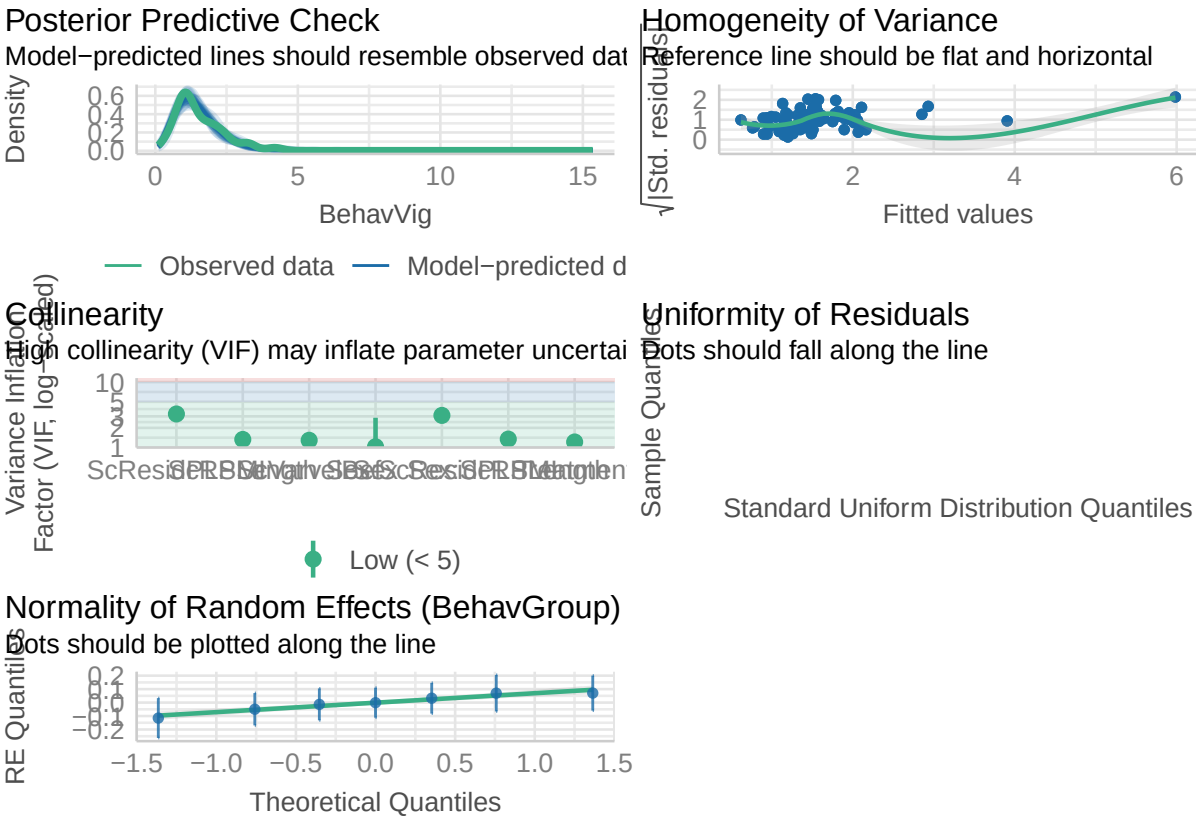
```
##  Family: Gamma  ( log )
## Formula:
## BehavVig ~ Sex * ScResidPLSMI + ScVarvelBef + ScRPLength + Treatment +
##     Sex:ScRPLength + (1 | BehavGroup)
## Data: IndBehav8
##
##      AIC      BIC   logLik deviance df.resid
##    458.1    493.6   -219.1    438.1      245
##
## Random effects:
##
```

```
## Conditional model:
##  Groups     Name        Variance Std.Dev.
##  BehavGroup (Intercept) 0.006553 0.08095
## Number of obs: 255, groups:  BehavGroup, 7
##
## Dispersion estimate for Gamma family (sigma^2): 0.168
##
## Conditional model:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)        0.43413    0.05416   8.016 1.09e-15 ***
## SexM              -0.07878    0.05367  -1.468  0.14212
## ScResidPLSMI       0.08813    0.04865   1.811  0.07009 .
## ScVarvelBef        0.27335    0.03065   8.918  < 2e-16 ***
## ScRPLength        -0.07718    0.03130  -2.466  0.01368 *
## TreatmentVIE      -0.03314    0.05862  -0.565  0.57186
## SexM:ScResidPLSMI -0.17948    0.05851  -3.067  0.00216 **
## SexM:ScRPLength   -0.19937    0.06927  -2.878  0.00400 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

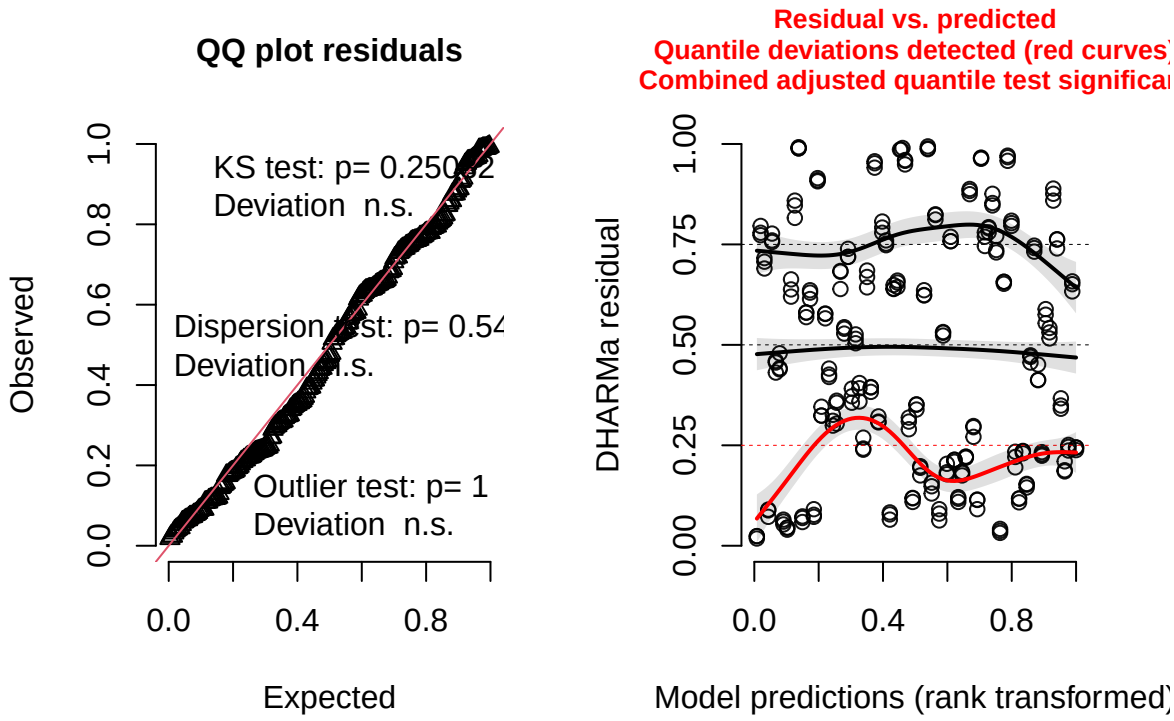
5.2.2 Validate that the model fits well and there are no problems

```
# Using the check_model function from the perforamnce package to check the
# model validation

check_model(BehavVigLM)
```

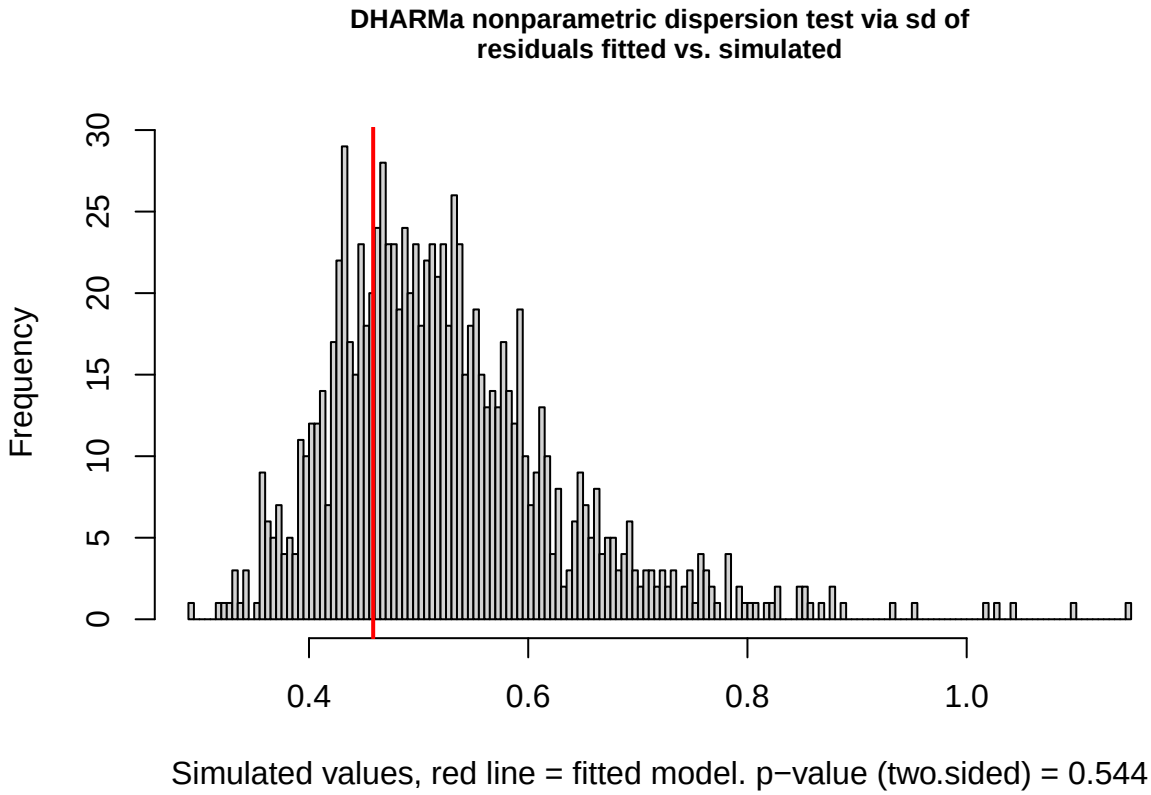
```
## `check_outliers()` does not yet support models of class `glmmTMB`.
```




```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_BehavVigLM <- simulateResiduals(BehavVigLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_BehavVigLM)
```




```
# Testing for dispersion
testDispersion(sim_residuals_BehavVigLM)
```



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.87322, p-value = 0.544
## alternative hypothesis: two.sided
```

5.2.3 Testing the significance of factors in our model

```
Anova(BehavVigLM, type = 3, test = "Chisq")
```

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: BehavVig
##                     Chisq Df Pr(>Chisq)
## (Intercept)      64.2599  1   1.09e-15 ***
## Sex               2.1549  1   0.142117
## ScResidPLSMI      3.2810  1   0.070088 .
## ScVarvelBef      79.5316  1  < 2.2e-16 ***
## ScRPLength        6.0791  1   0.013679 *
## Treatment         0.3196  1   0.571860
## Sex:ScResidPLSMI  9.4088  1   0.002159 **
## Sex:ScRPLength    8.2843  1   0.003999 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

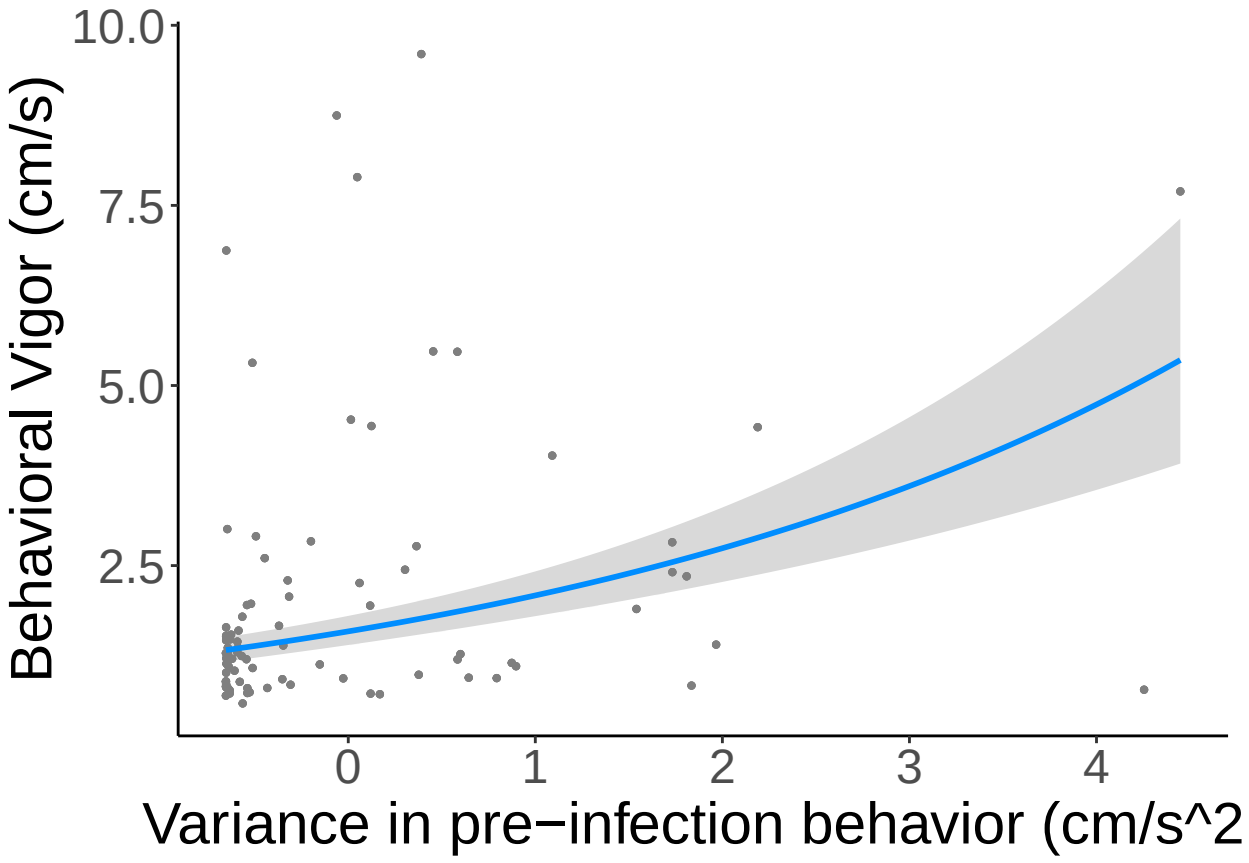
5.2.4 Visualize the important explanatory factors for behavioral vigor

```
# Variance in behaivor
BehavVigbyVar = visreg(BehavVigLM, scale = "response", "ScVarvelBef", partial = T,
    gg = TRUE) + theme_classic() + theme(legend.position = "none") + ylab("Behavioral Vigor (cm/s)") +
    xlab("Variance in pre-infection behavior (cm/s^2)") + theme(text = element_text(size = 22))
```

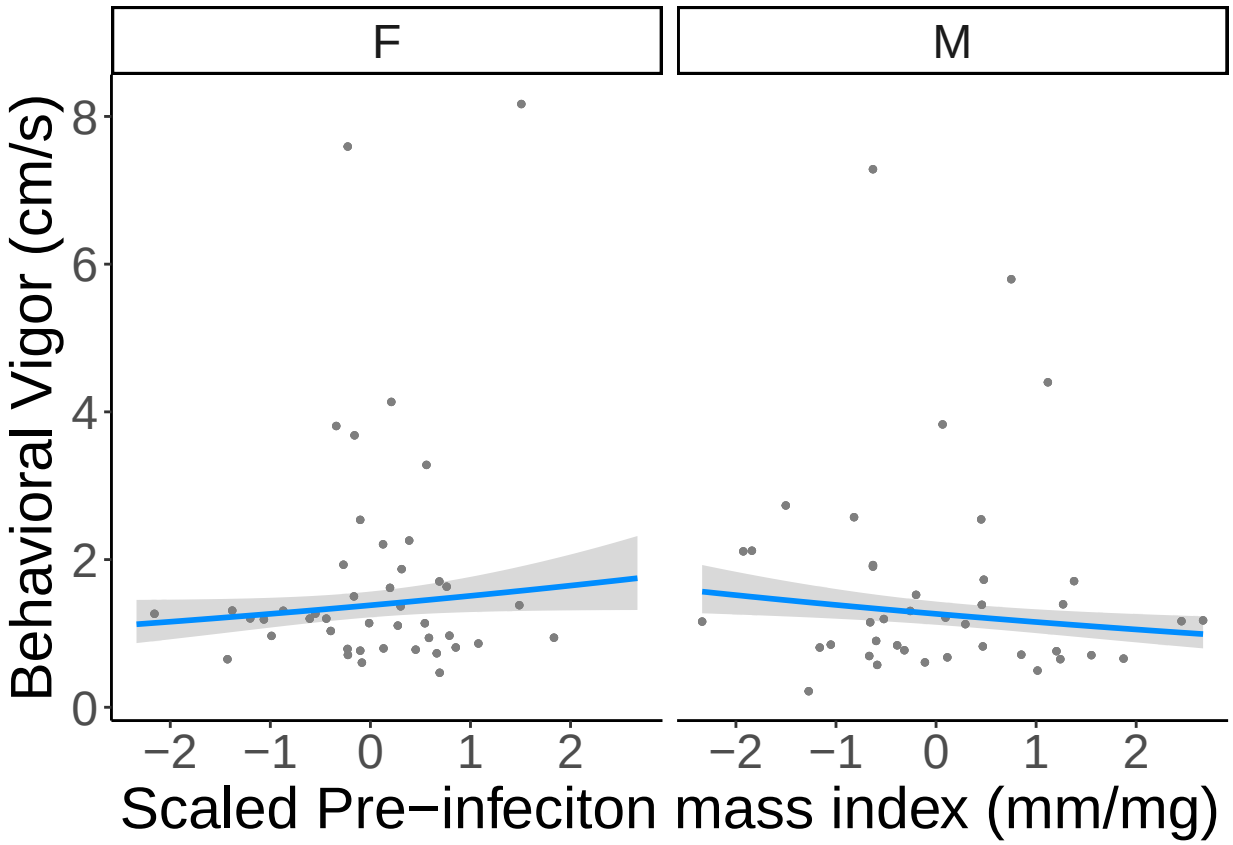
```
## Warning:   Note that you are attempting to plot a 'main effect' in a model that contains an
##   interaction.  This is potentially misleading; you may wish to consider using the 'by'
##   argument.
```

```
## Conditions used in construction of plot
## Sex: F
## ScResidPLSMI: -0.01293818
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
## BehavGroup: 4
```

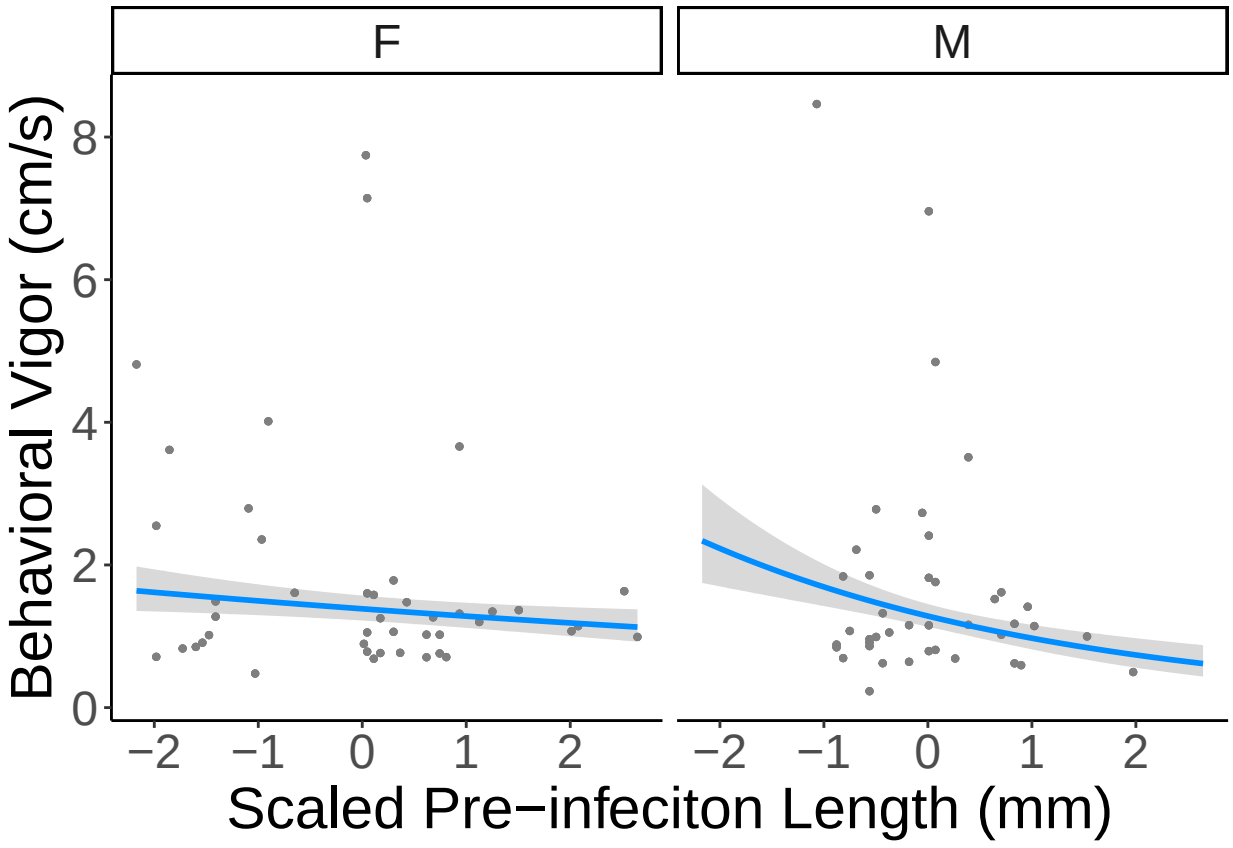
```
# Print the graph
print(BehavVigbyVar)
```




```
# TrialTimeGraph
BehavVigbySMI = visreg(BehavVigLM, scale = "response", "ScResidPLSMI", "Sex", partial = T,
    gg = TRUE) + theme_classic() + theme(legend.position = "none") + ylab("Behavioral Vigor (cm/s)") +
    xlab("Scaled Pre-infeciton mass index (mm/mg)") + theme(text = element_text(size = 22)) +
    facet_wrap(~Sex)
# Print the graph
print(BehavVigbySMI)
```




```
# TrialTimeGraph
BehavVigbyLength = visreg(BehavVigLM, scale = "response", "ScRPLength", "Sex", partial = T,
    gg = TRUE) + theme_classic() + theme(legend.position = "none") + ylab("Behavioral Vigor (cm/s)") +
    xlab("Scaled Pre-infeciton Length (mm)") + theme(text = element_text(size = 22)) +
    facet_wrap(~Sex)
# Print the graph
print(BehavVigbyLength)
```



5.2.5 Post-hoc analysis where we split by sex

```
IndBehav8M <- IndBehav8 %>%
    filter(Sex == "M")
```

```
## filter: removed 135 rows (52%), 126 rows remaining
```

```
IndBehav8F <- IndBehav8 %>%
    filter(Sex == "F")
```

```
## filter: removed 126 rows (48%), 135 rows remaining
```

```
# Fit a linear model for checking what explanatory factors are important for
# Variance in Velocity Note this is a linear mixed model because we have
# multiple measures per fish and therefore, need to account for
# non-independence between measures.  Females
BehavVigLMF <- glmmTMB(BehavVig ~ ScResidPLSMI + ScVarvelBef + Treatment + ScRPLength +
    (1 | BehavGroup), family = Gamma("log"), IndBehav8F)
# Males
BehavVigLMM <- glmmTMB(BehavVig ~ ScResidPLSMI + ScVarvelBef + ScRPLength * Treatment +
    (1 | BehavGroup), family = Gamma("log"), IndBehav8M)

# Summary to see the relationship of the varialbes for females
summary(BehavVigLMF)
```

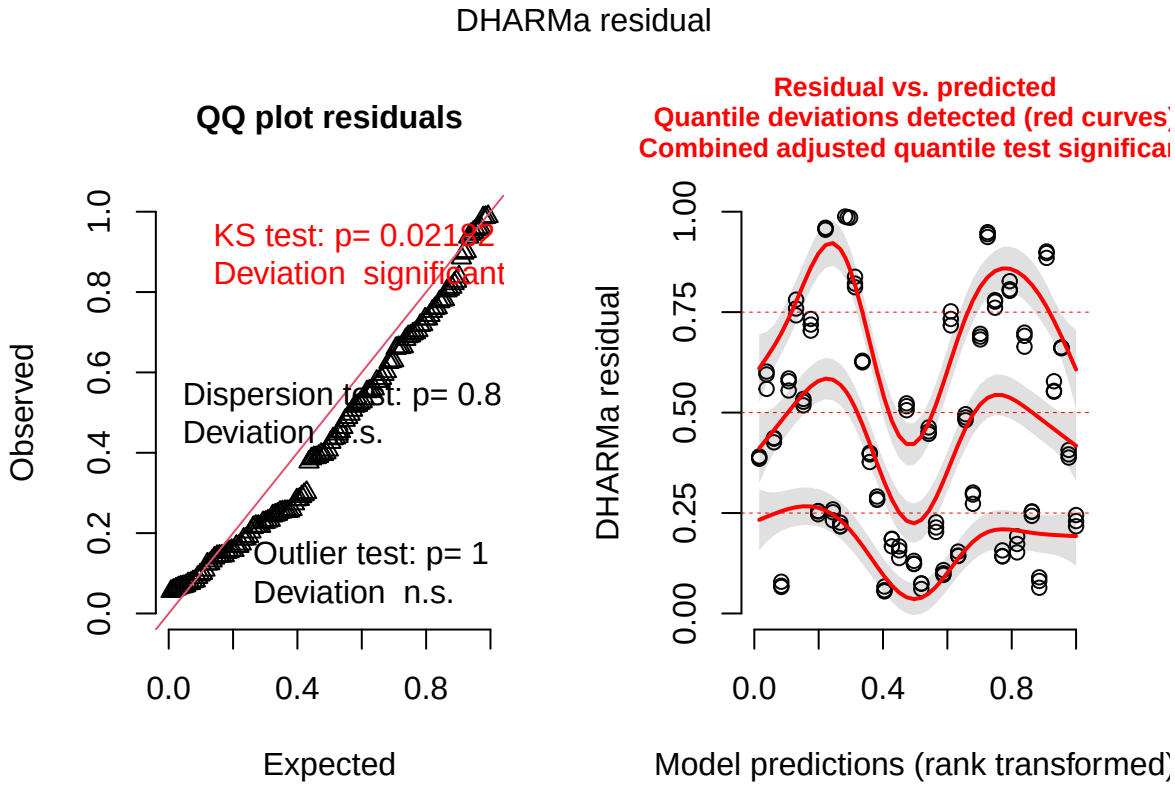
```
##  Family: Gamma  ( log )
## Formula:
## BehavVig ~ ScResidPLSMI + ScVarvelBef + Treatment + ScRPLength +
##     (1 | BehavGroup)
## Data: IndBehav8F
##
##      AIC      BIC   logLik deviance df.resid
##    205.3    225.5    -95.6    191.3      125
##
## Random effects:
##
## Conditional model:
##  Groups     Name        Variance Std.Dev.
##  BehavGroup (Intercept) 0.06519  0.2553
## Number of obs: 132, groups:  BehavGroup, 7
##
## Dispersion estimate for Gamma family (sigma^2): 0.112
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.51769    0.10826   4.782 1.73e-06 ***
## ScResidPLSMI  0.13492    0.04603   2.931  0.00338 **
## ScVarvelBef   0.22637    0.04584   4.938 7.87e-07 ***
## TreatmentVIE -0.15760    0.07191  -2.192  0.02841 *
## ScRPLength   -0.07310    0.02742  -2.666  0.00768 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Summary to see the relationship of the variables for males.
summary(BehavVigLMM)
```

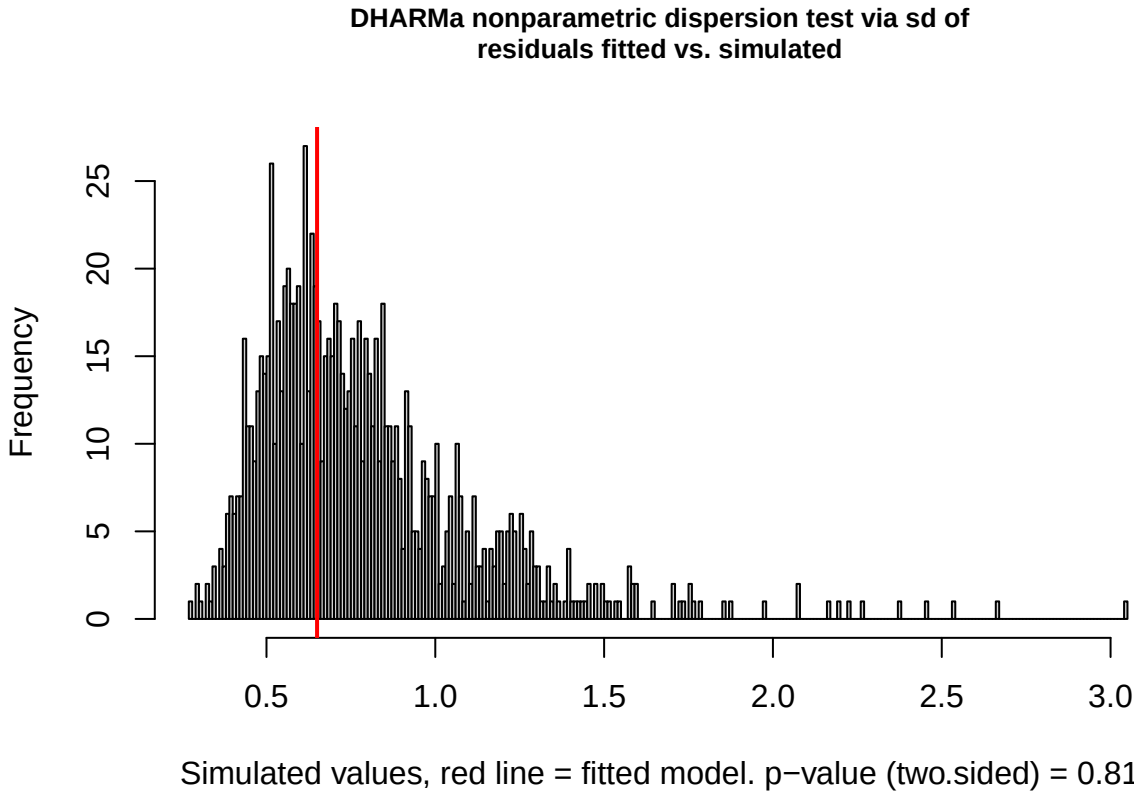
```
##  Family: Gamma  ( log )
## Formula:
## BehavVig ~ ScResidPLSMI + ScVarvelBef + ScRPLength * Treatment +
##     (1 | BehavGroup)
## Data: IndBehav8M
##
##      AIC      BIC   logLik deviance df.resid
##    234.1    256.6   -109.0    218.1      115
##
## Random effects:
##
## Conditional model:
##  Groups     Name        Variance Std.Dev.
##  BehavGroup (Intercept) 0.02584  0.1608
## Number of obs: 123, groups:  BehavGroup, 7
##
## Dispersion estimate for Gamma family (sigma^2): 0.179
##
## Conditional model:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)               0.27928    0.08383   3.332 0.000863 ***
## ScResidPLSMI             -0.07830    0.03948  -1.984 0.047309 *
## ScVarvelBef               0.26965    0.03978   6.778 1.21e-11 ***
```

```
## ScRPLength               -0.32815    0.09058  -3.623 0.000291 ***
## TreatmentVIE              0.14832    0.10852   1.367 0.171713
## ScRPLength:TreatmentVIE  0.07706    0.14541   0.530 0.596141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Using the Dharma package to check quantile residuals for female vigor model
# First simulating the quantile residuals
sim_residuals_BehavVigLMF <- simulateResiduals(BehavVigLMF, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_BehavVigLMF)
```

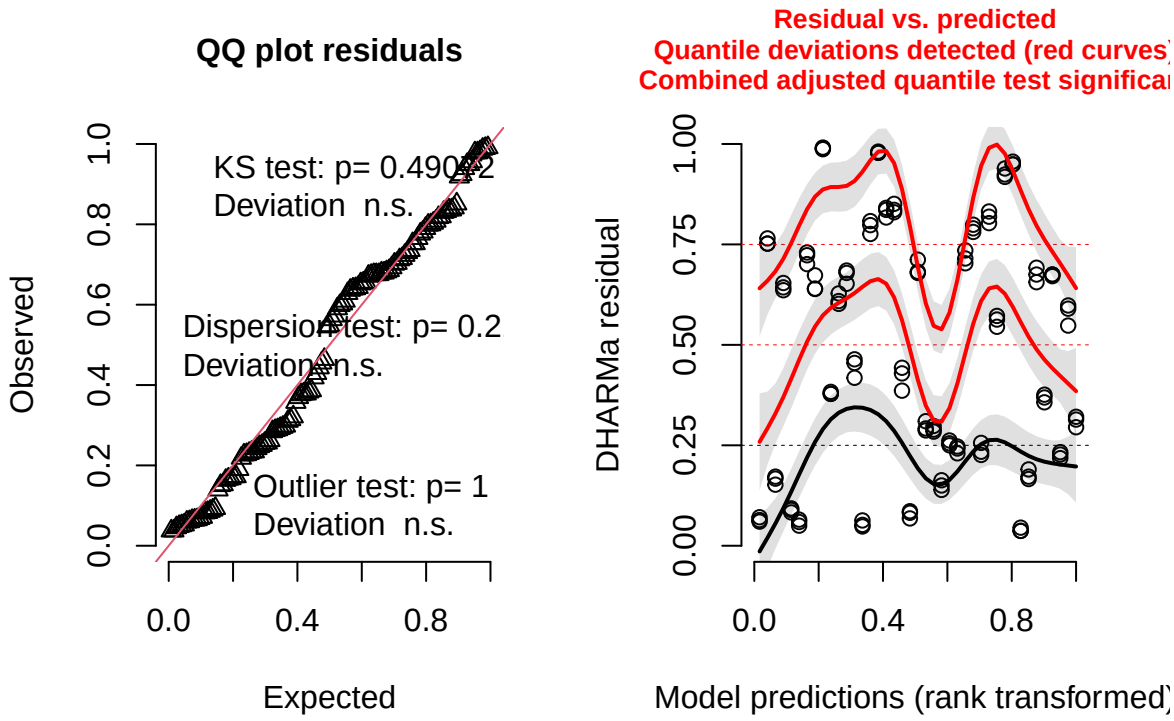



```
# Testing for dispersion
testDispersion(sim_residuals_BehavVigLMF)
```

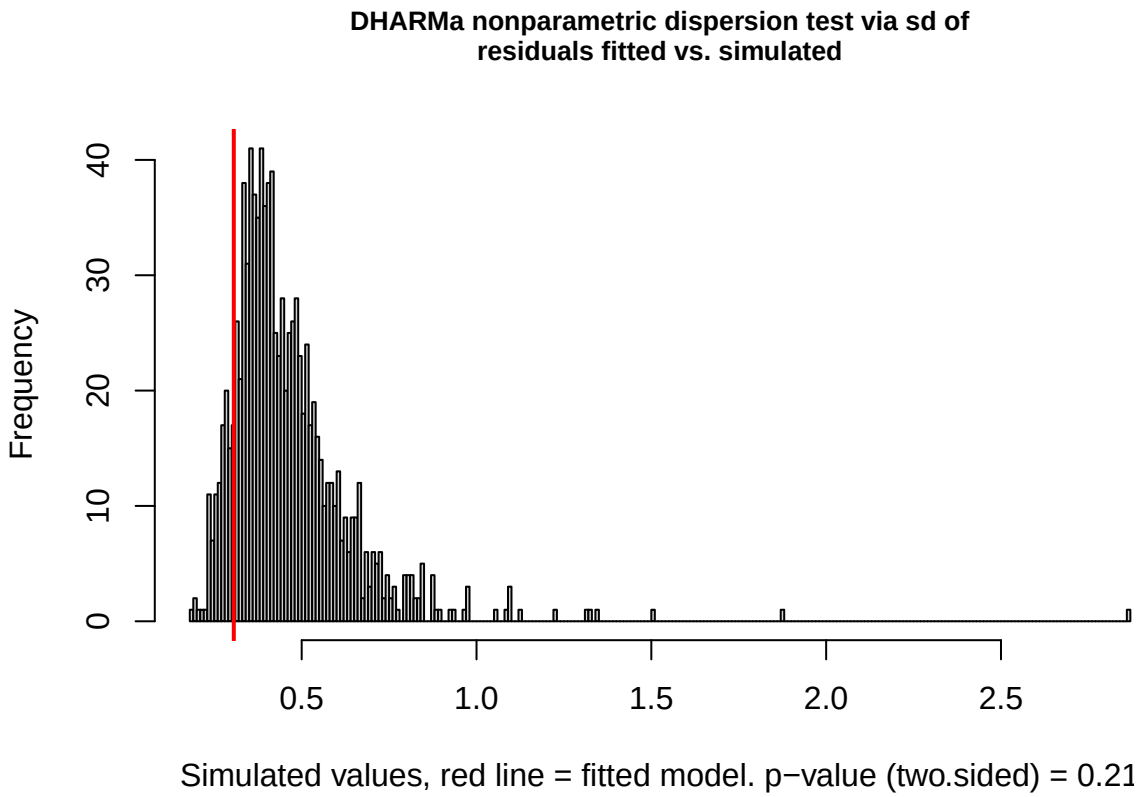



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.82615, p-value = 0.81
## alternative hypothesis: two.sided
```

```
# Using the Dharma package to check quantile residuals for female vigor model
# First simulating the quantile residuals
sim_residuals_BehavVigLMM <- simulateResiduals(BehavVigLMM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_BehavVigLMM)
```




```
# Testing for dispersion
testDispersion(sim_residuals_BehavVigLMM)
```



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.66116, p-value = 0.21
## alternative hypothesis: two.sided
```

5.2.6 Testing the significance of factors in our model

```
Anova(BehavVigLMF, type = 2, test = "Chisq")
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: BehavVig
##                Chisq Df Pr(>Chisq)
## ScResidPLSMI  8.5926  1   0.003375 **
## ScVarvelBef  24.3883  1  7.875e-07 ***
## Treatment     4.8032  1   0.028406 *
## ScRPLength    7.1076  1   0.007676 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova(BehavVigLMM, type = 2, test = "Chisq")
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: BehavVig
##                         Chisq Df Pr(>Chisq)
## ScResidPLSMI           3.9344  1    0.04731 *
## ScVarvelBef           45.9475  1  1.215e-11 ***
## ScRPLength            17.6976  1  2.590e-05 ***
## Treatment              2.4128  1    0.12035
## ScRPLength:Treatment   0.2809  1    0.59614
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.3 What factors are important for host behavioral tolerance and do is there any sexual variation in host behavioral tolerance?

5.3.1 Description, development, and fitting of linear model for the analysis

We will use a linear model to analyze how behavioral tolerance differs by sexual variation and other important host traits. Given each host only has one behavioral tolerance measure, we do not need the fishID random effect used in previous models/

- Deterministic
- $BehavTol_{det}$ = a + b_1Sex + b_2ScRPLength + b_3ScVarVelBef + b_4ScResidPLSMI + b_5ScBehavVig + b_5ScTisTol + b_6Sex:ScRPLength + a_i

- Stochastic
  - BehavTol ~ $N(BehavTol_{det}, \sigma^2)$
  - $a_i \sim N(0, \sigma^2_{BehavGroup})$
- Fixed
  - Sex
  - Scaled residual from length and sex
  - Scaled Pre-infection SMI
  - Scaled variance in velocity before infection
  - Scaled behavioral vigor
  - Scaled Tissue Tolerance
  - Sex by length residuals interaction

```
# We have some outliers that make interpreting the results for behavioral
# tolerance a pain so were removing them from the analysis

# Fit a linear model for behavioral tolerance Note this is a linear mixed model
# because we have multiple measures per fish and therefore, need to account for
# non-independence between measures.
BehavTolLM <- glmmTMB(BehavTol ~ Sex + ScResidPLSMI + ScChSMI + ScVarvelBef + ScRPLength +
    ScBehavVig + Treatment + (1 | BehavGroup), IndBehavI)



# Summary to see the relationship of the variables.
summary(BehavTolLM)
```

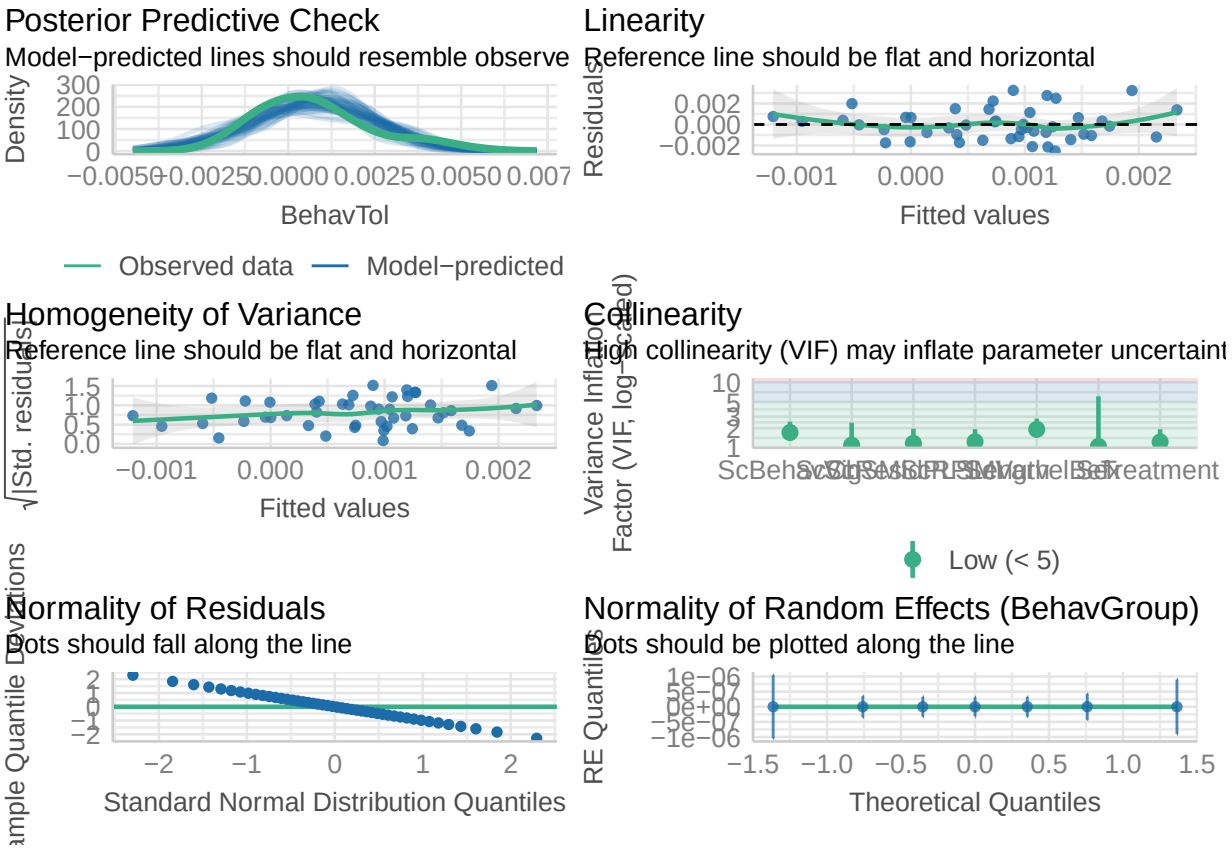
```
##  Family: gaussian  ( identity )
## Formula:
## BehavTol ~ Sex + ScResidPLSMI + ScChSMI + ScVarvelBef + ScRPLength +
##     ScBehavVig + Treatment + (1 | BehavGroup)
## Data: IndBehavI
##
##      AIC      BIC   logLik deviance df.resid
##   -453.1   -434.9    236.6   -473.1       36
##
## Random effects:
##
## Conditional model:
##  Groups     Name        Variance  Std.Dev.
##  BehavGroup (Intercept) 2.544e-14 1.595e-07
##  Residual               1.998e-06 1.413e-03
## Number of obs: 46, groups:  BehavGroup, 7
##
## Dispersion estimate for gaussian family (sigma^2): 2e-06
##
## Conditional model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.0011030  0.0003588   3.074  0.00211 **
## SexM         -0.0003081  0.0004281  -0.720  0.47167
## ScResidPLSMI -0.0002661  0.0002508  -1.061  0.28861
## ScChSMI       0.0001938  0.0002395   0.809  0.41851
```

```
## ScVarvelBef    0.0003069  0.0003149   0.975  0.32982
## ScRPLength     0.0002379  0.0002461   0.967  0.33367
## ScBehavVig    -0.0007011  0.0002607  -2.689  0.00717 **
## TreatmentVIE -0.0002562  0.0004786  -0.535  0.59240
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

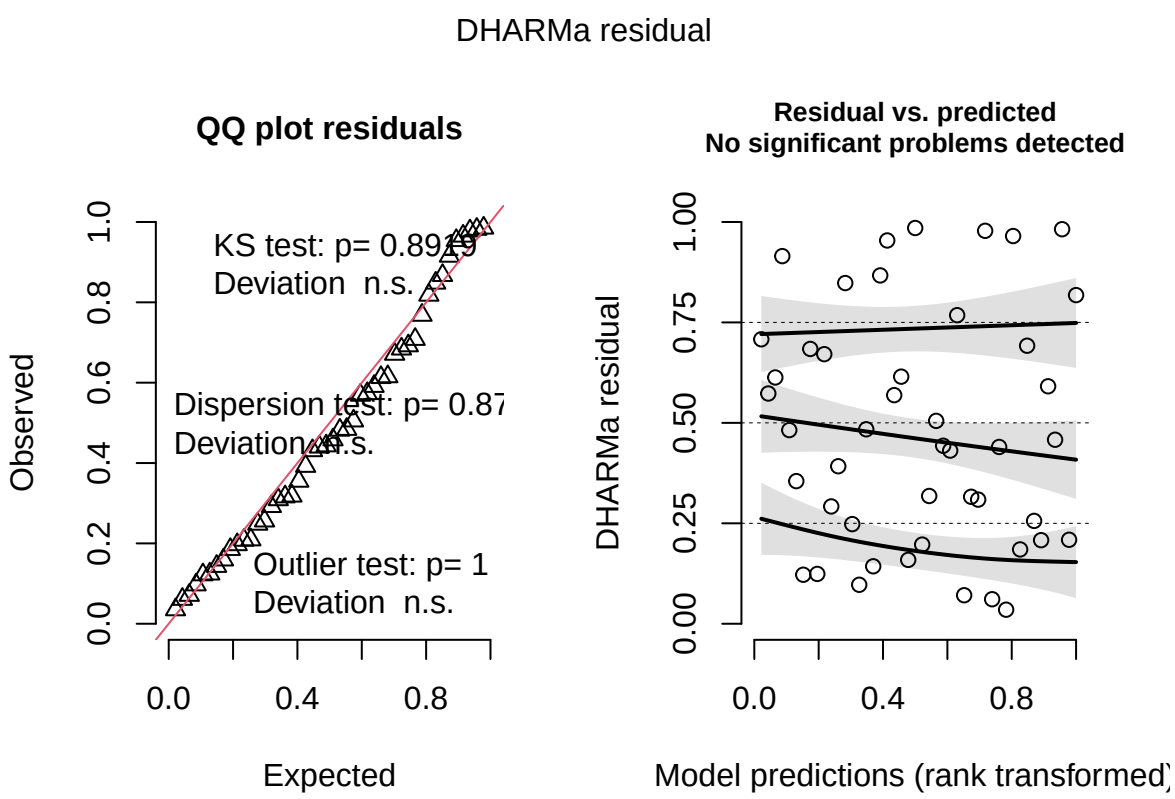
5.3.2 Validate that the model fits well and there are no problems

```
# Using the check_model function from the perforamnce package to check the
# model validation

check_model(BehavTolLM)
```

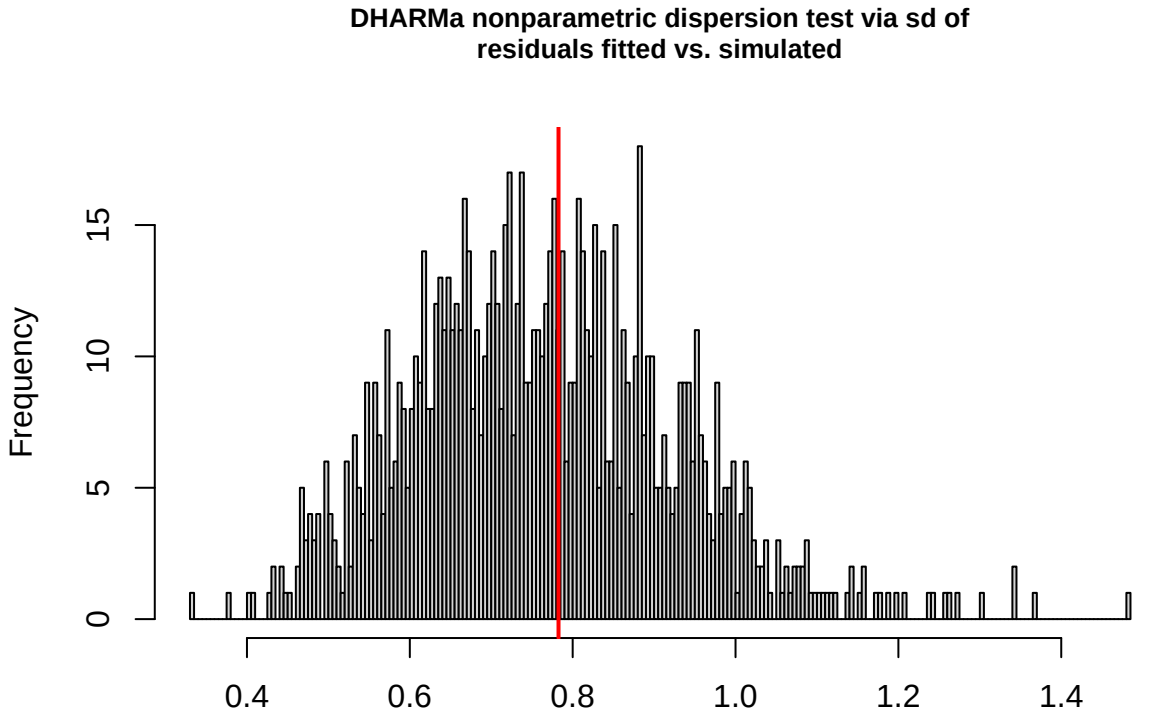
```
## `check_outliers()` does not yet support models of class `glmmTMB`.
```




```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_BehavTolLM <- simulateResiduals(BehavTolLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_BehavTolLM)
```




```
# Testing for dispersion
testDispersion(sim_residuals_BehavTolLM)
```



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 1.0218, p-value = 0.872
## alternative hypothesis: two.sided
```

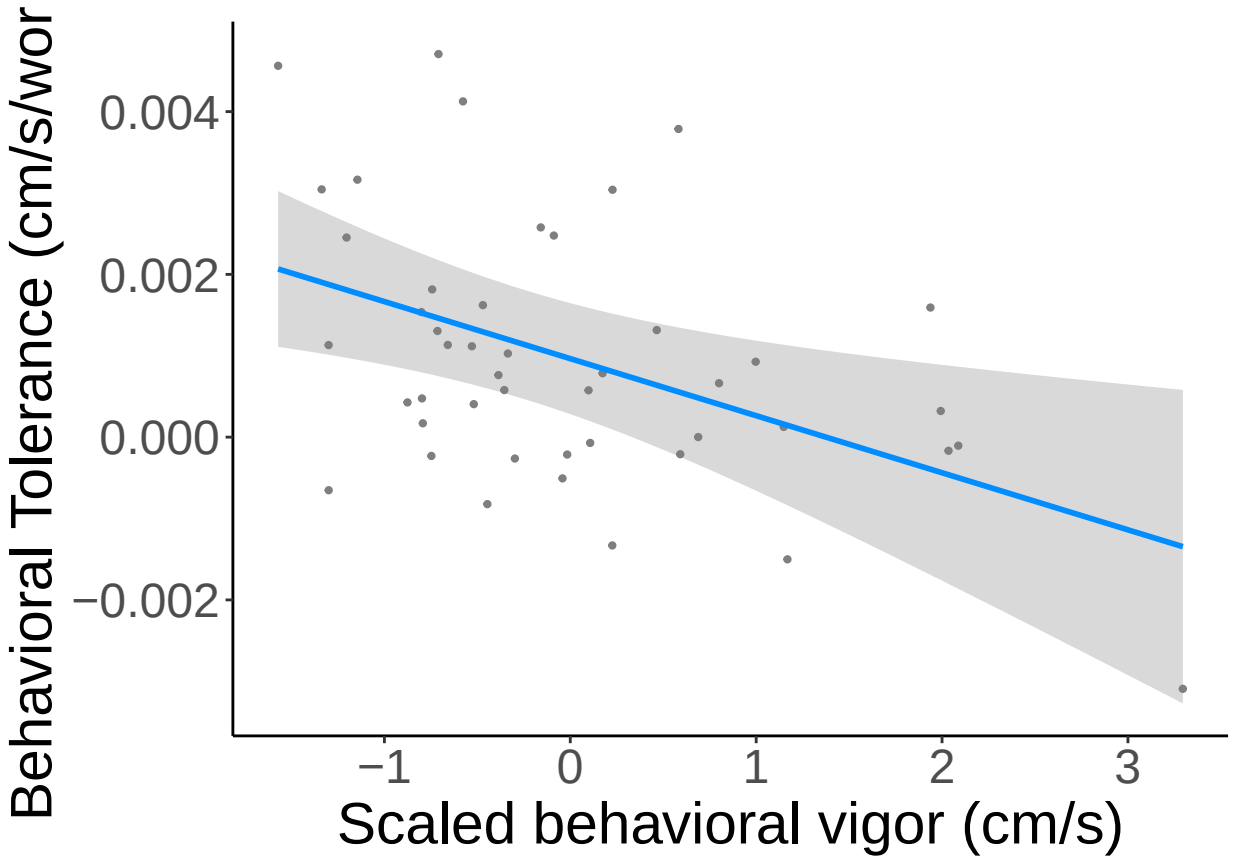
5.3.3 Testing the significance of factors in our model

```
Anova(BehavTolLM, Type = 3, test = "Chisq")
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: BehavTol
##                Chisq Df Pr(>Chisq)
## Sex           0.5181  1   0.471673
## ScResidPLSMI 1.1261  1   0.288606
## ScChSMI       0.6545  1   0.418508
## ScVarvelBef   0.9496  1   0.329817
## ScRPLength    0.9346  1   0.333672
## ScBehavVig    7.2301  1   0.007169 **
## Treatment     0.2866  1   0.592403
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.3.4 Visualize the important explanatory factors for behavioral tolerance

```
# TrialTimeGraph
BehavTolbyVig = visreg(BehavTolLM, scale = "response", "ScBehavVig", partial = T,
    gg = TRUE) + theme_classic() + theme(legend.position = "none") + ylab("Behavioral Tolerance (cm/s/w
    xlab("Scaled behavioral vigor (cm/s)") + theme(text = element_text(size = 22))
# Print the graph
print(BehavTolbyVig)
```

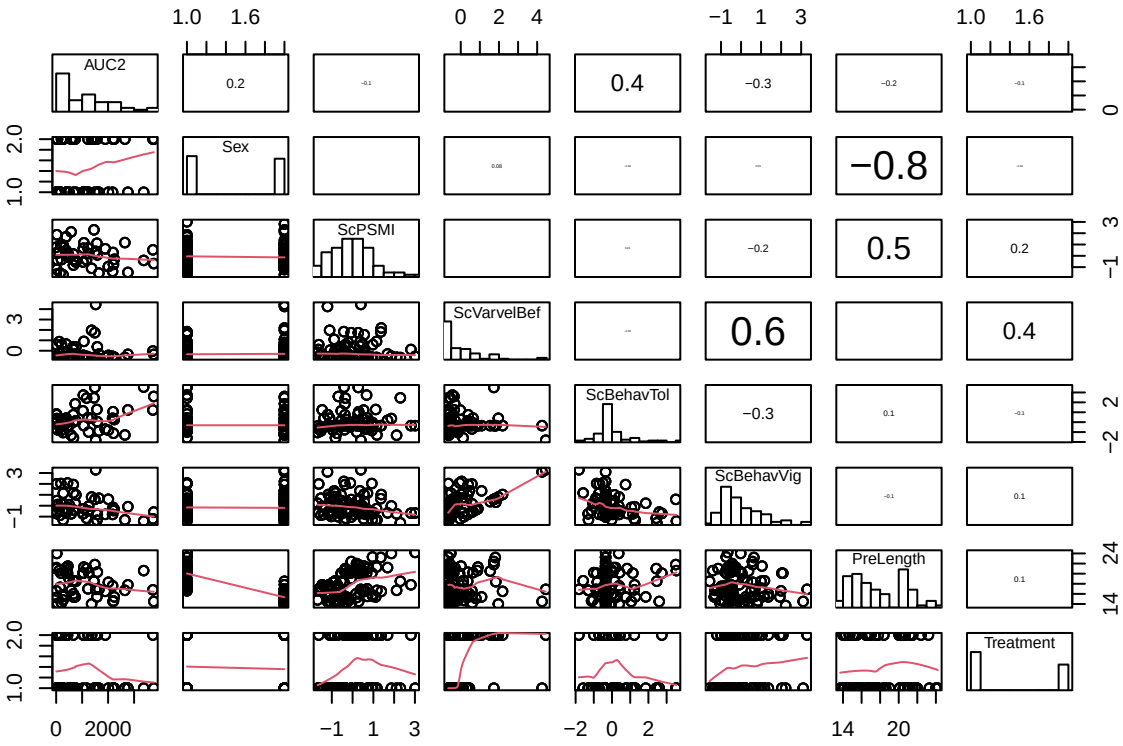


5.3.5 Post-hoc analysis for behavioral tolerance by sex

5.4 What factors are important for host infection intensity over the course of infection and do is there any sexual variation in host infection intensity?

5.5 Visually inspection of the explanatory variables that will be used in the analyses

```
pairs(~AUC2 + Sex + ScPSMI + ScVarvelBef + ScBehavTol + ScBehavVig + PreLength +
    Treatment, lower.panel = panel.smooth, diag.panel = panel.hist, upper.panel = panel.cor,
    data = IndBehav8)
```



5.5.1 Description, development, and fitting of linear model for the analysis

We will use a linear model to analyze how infection intensity differs by sexual variation and other important host traits. Given each host only has one infection intensity measure, we do not need the fishID random effect used in previous models.

- Deterministic
- AUC_{det} = a + b_1Sex + b_2ScBehavVig + b_3ScVarVelBef + b_4ScResidPLSMI + b_5ScRPLength + b_6Sex:ScBehavVig + b_7Sex:ScResidPLSMI
- Stochastic
  - AUC ~ $N(AUC_{det}, \sigma^2)$
- Fixed
  - Sex
  - Scaled behavioral vigor
  - Scaled Pre-infection SMI
  - Scaled variance in velocity before infection
  - Scaled Residuals of Length and Sex
  - VIE treatment
  - Interaction between Sex and behavioral vigor
  - Interaction between sex and body condition

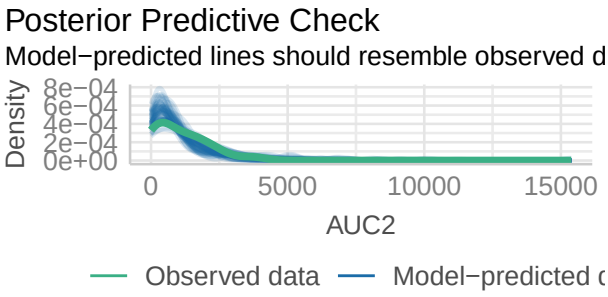
```
# Fit a linear model for behavioral tolerance Note this is a linear mixed model
# because we have multiple measures per fish and therefore, need to account for
# non-independence between measures.
AUCLM <- glm(AUC2 ~ Sex + ScBehavVig + ScBehavVig + ScVarvelBef + ScResidPLSMI +
```

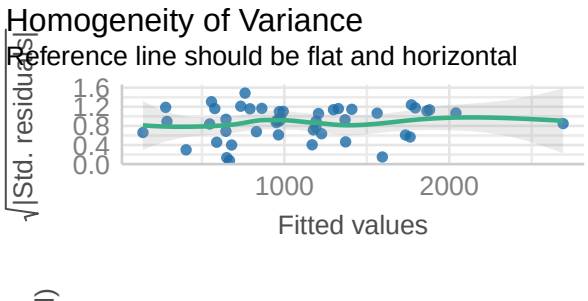
```
    ScRPLength + Treatment + Sex:ScResidPLSMI, family = Gamma(link = "log"), IndBehavI)

# Summary to see the relationship of the variables.
summary(AUCLM)
```

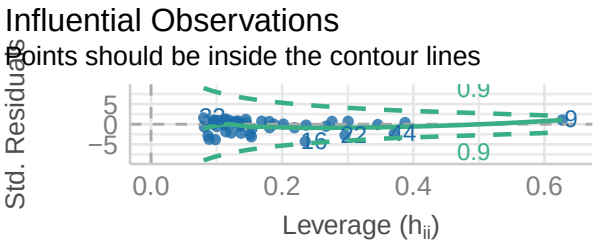
```
##
## Call:
## glm(formula = AUC2 ~ Sex + ScBehavVig + ScBehavVig + ScVarvelBef +
##     ScResidPLSMI + ScRPLength + Treatment + Sex:ScResidPLSMI,
##     family = Gamma(link = "log"), data = IndBehavI)
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        6.64270    0.19831  33.497  < 2e-16 ***
## SexM               0.56522    0.23528   2.402  0.02129 *
## ScBehavVig        -0.48120    0.14684  -3.277  0.00225 **
## ScVarvelBef        0.34499    0.17379   1.985  0.05439 .
## ScResidPLSMI      -0.30924    0.23284  -1.328  0.19205
## ScRPLength        -0.36544    0.13191  -2.770  0.00862 **
## TreatmentVIE      -0.01318    0.27719  -0.048  0.96233
## SexM:ScResidPLSMI  0.17665    0.29466   0.599  0.55241
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.6002634)
##
##     Null deviance: 70.454  on 45  degrees of freedom
## Residual deviance: 58.525  on 38  degrees of freedom
##   (2 observations deleted due to missingness)
## AIC: 741.17
##
## Number of Fisher Scoring iterations: 11
```

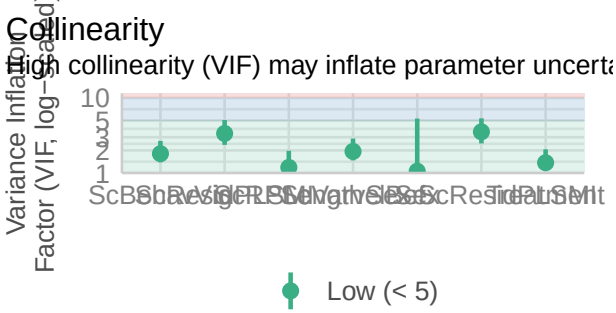
5.5.2 Validate that the model fits well and there are no problems

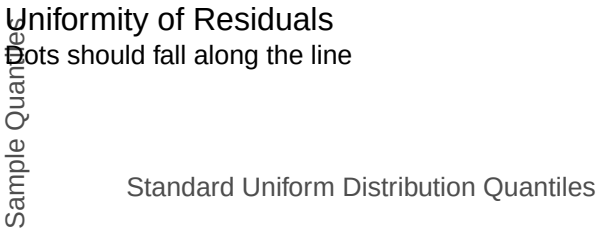
```
# Using the check_model function from the perforamnce package to check the
# model validation

check_model(AUCLM)
```



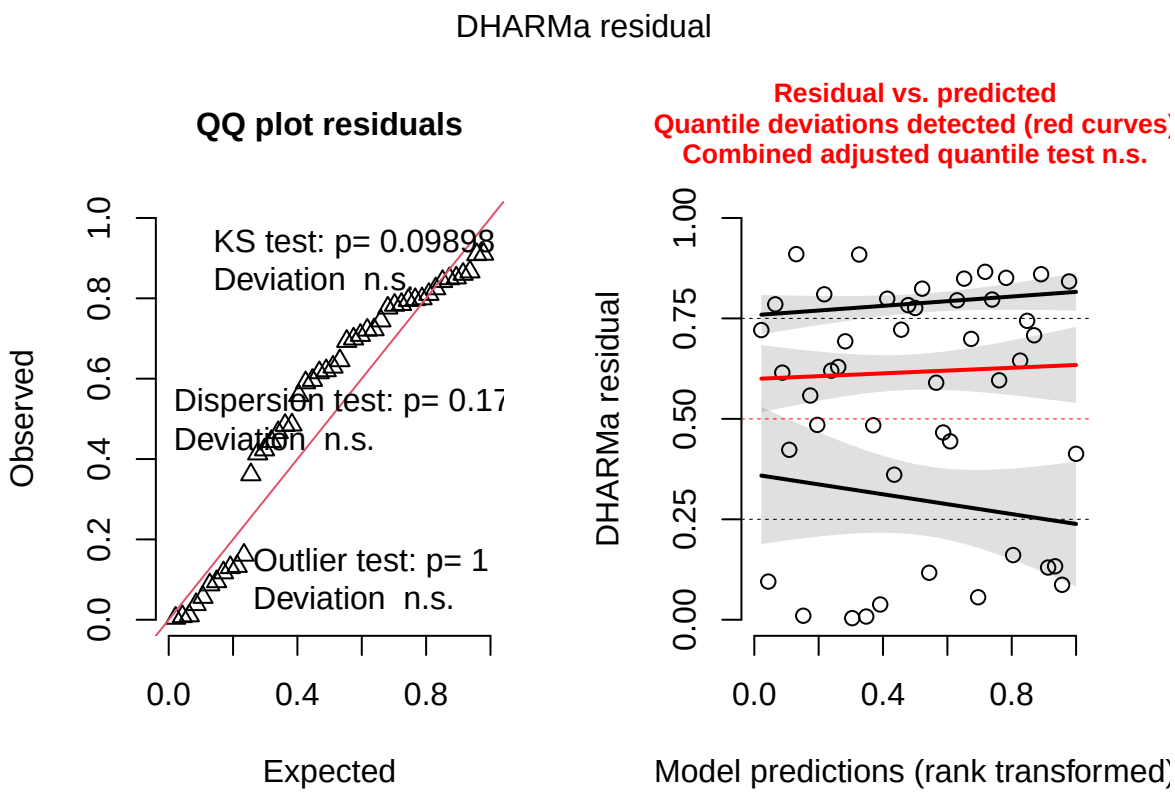




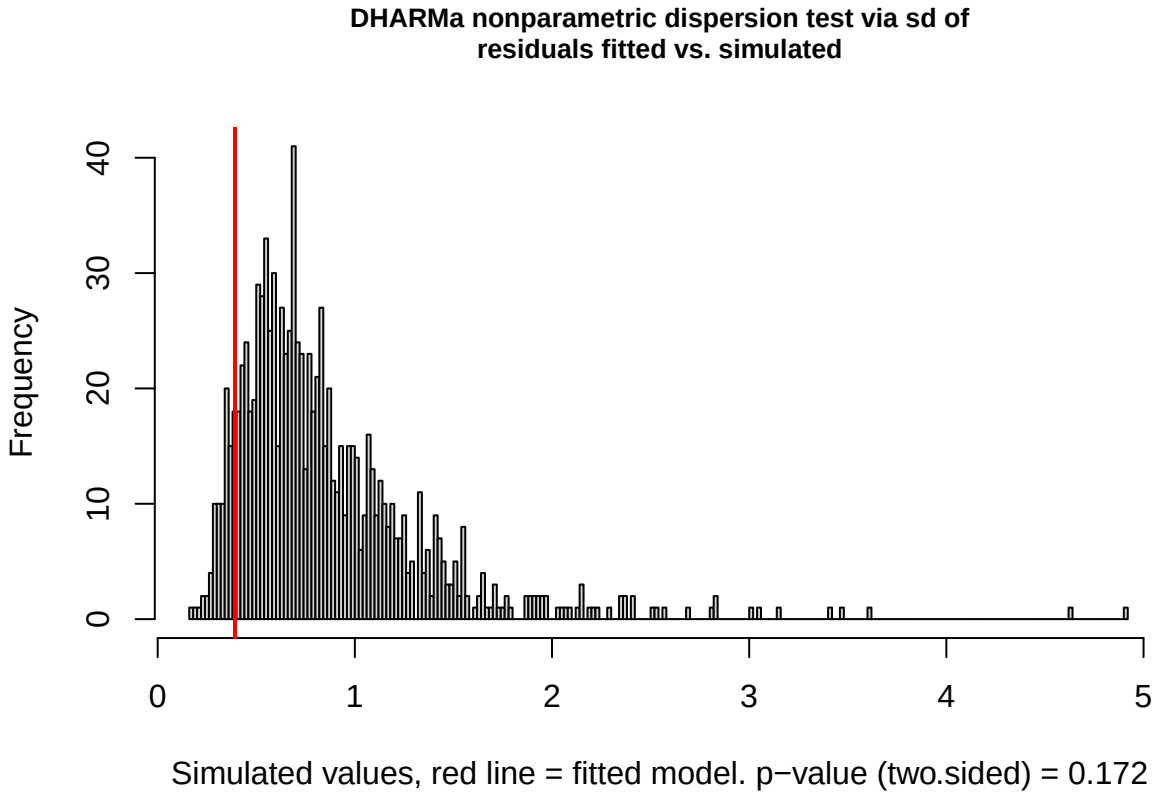





```
# Using the Dharma package to check quantile residuals First simulating the
# quantile residuals
sim_residuals_AUCLM <- simulateResiduals(AUCLM, 1000)
# Plotting the quantile residuals to test how quantile residuals look
plot(sim_residuals_AUCLM)
```




```
# Testing for dispersion
testDispersion(sim_residuals_AUCLM)
```



```
##
##  DHARMa nonparametric dispersion test via sd of residuals fitted vs.
##  simulated
##
## data:  simulationOutput
## dispersion = 0.45949, p-value = 0.172
## alternative hypothesis: two.sided
```

5.5.3 Testing the significance of factors in our model

```
summary(AUCLM)
```

```
##
## Call:
## glm(formula = AUC2 ~ Sex + ScBehavVig + ScBehavVig + ScVarvelBef +
##     ScResidPLSMI + ScRPLength + Treatment + Sex:ScResidPLSMI,
##     family = Gamma(link = "log"), data = IndBehavI)
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        6.64270    0.19831  33.497  < 2e-16 ***
## SexM               0.56522    0.23528   2.402  0.02129 *
## ScBehavVig        -0.48120    0.14684  -3.277  0.00225 **
## ScVarvelBef        0.34499    0.17379   1.985  0.05439 .
## ScResidPLSMI      -0.30924    0.23284  -1.328  0.19205
## ScRPLength        -0.36544    0.13191  -2.770  0.00862 **
## TreatmentVIE      -0.01318    0.27719  -0.048  0.96233
## SexM:ScResidPLSMI  0.17665    0.29466   0.599  0.55241
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.6002634)
##
##     Null deviance: 70.454  on 45  degrees of freedom
## Residual deviance: 58.525  on 38  degrees of freedom
##   (2 observations deleted due to missingness)
## AIC: 741.17
##
## Number of Fisher Scoring iterations: 11
```

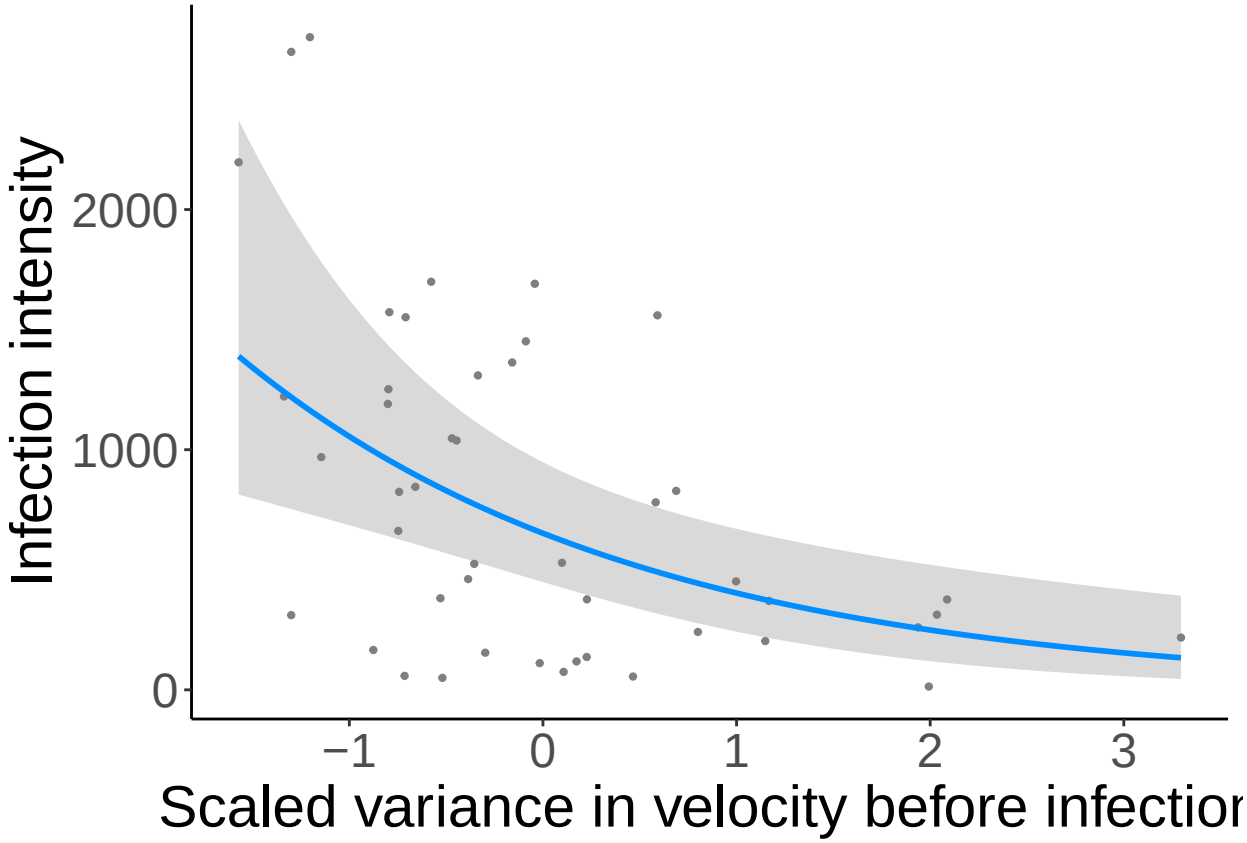
5.5.4 Visualize the important explanatory factors for infection intensity

```
# Behavioral Vigor graph
InfIntbyVig = visreg(AUCLM, scale = "response", "ScBehavVig", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Infection intensity") +
    xlab("Scaled variance in velocity before infection") + theme(text = element_text(size = 22))
```

```
## Conditions used in construction of plot
## Sex: F
## ScVarvelBef: -0.5119451
## ScResidPLSMI: -0.1015758
```

```
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
```

```
# Print the graph
print(InfIntbyVig)
```

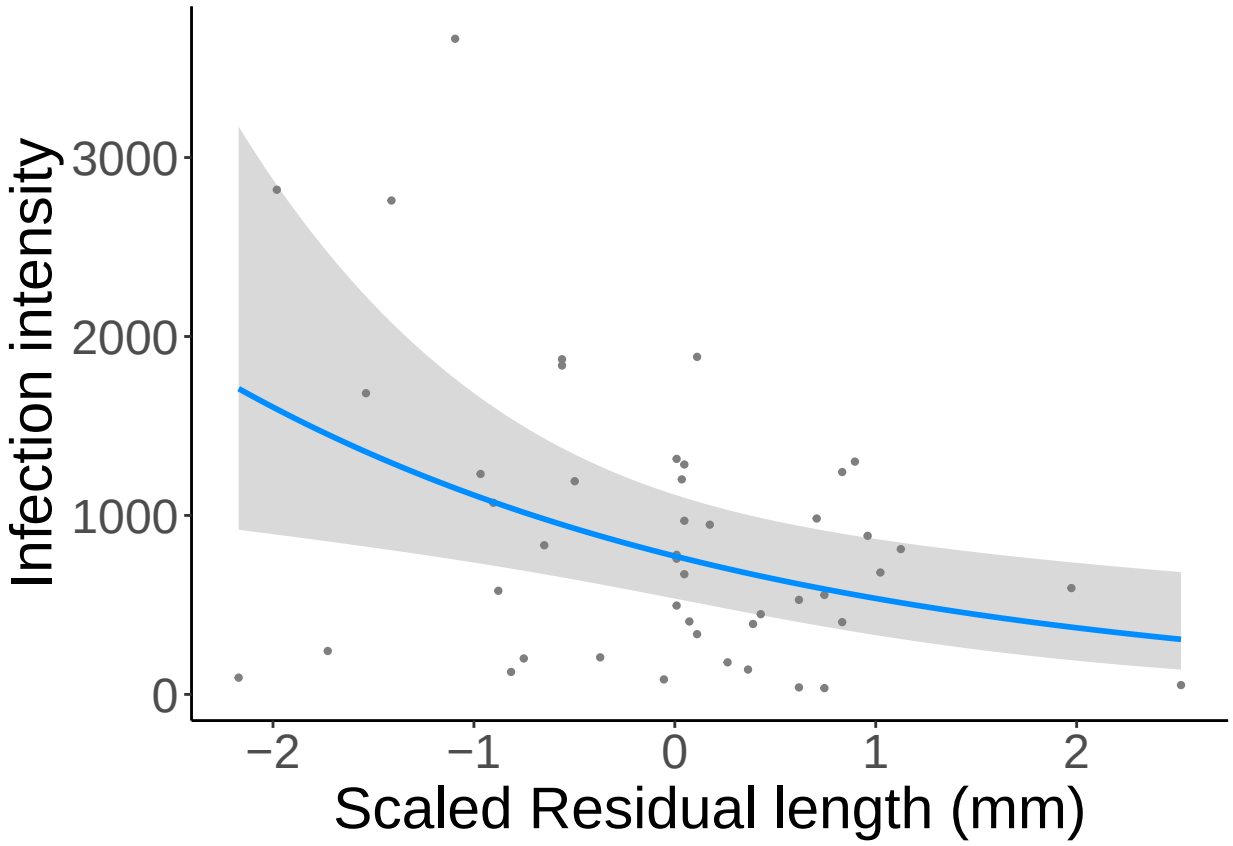



```
# Behavioral Vigor graph
InfIntbyLen = visreg(AUCLM, scale = "response", "ScRPLength", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Infection intensity") +
    xlab("Scaled Residual length (mm)") + theme(text = element_text(size = 22))
```

```
## Warning:   Note that you are attempting to plot a 'main effect' in a model that contains an
##   interaction.  This is potentially misleading; you may wish to consider using the 'by'
##   argument.
```

```
## Conditions used in construction of plot
## Sex: F
## ScBehavVig: -0.316863
## ScVarvelBef: -0.5119451
## ScResidPLSMI: -0.1015758
## Treatment: UNTOUCHED
```

```
# Print the graph
print(InfIntbyLen)
```

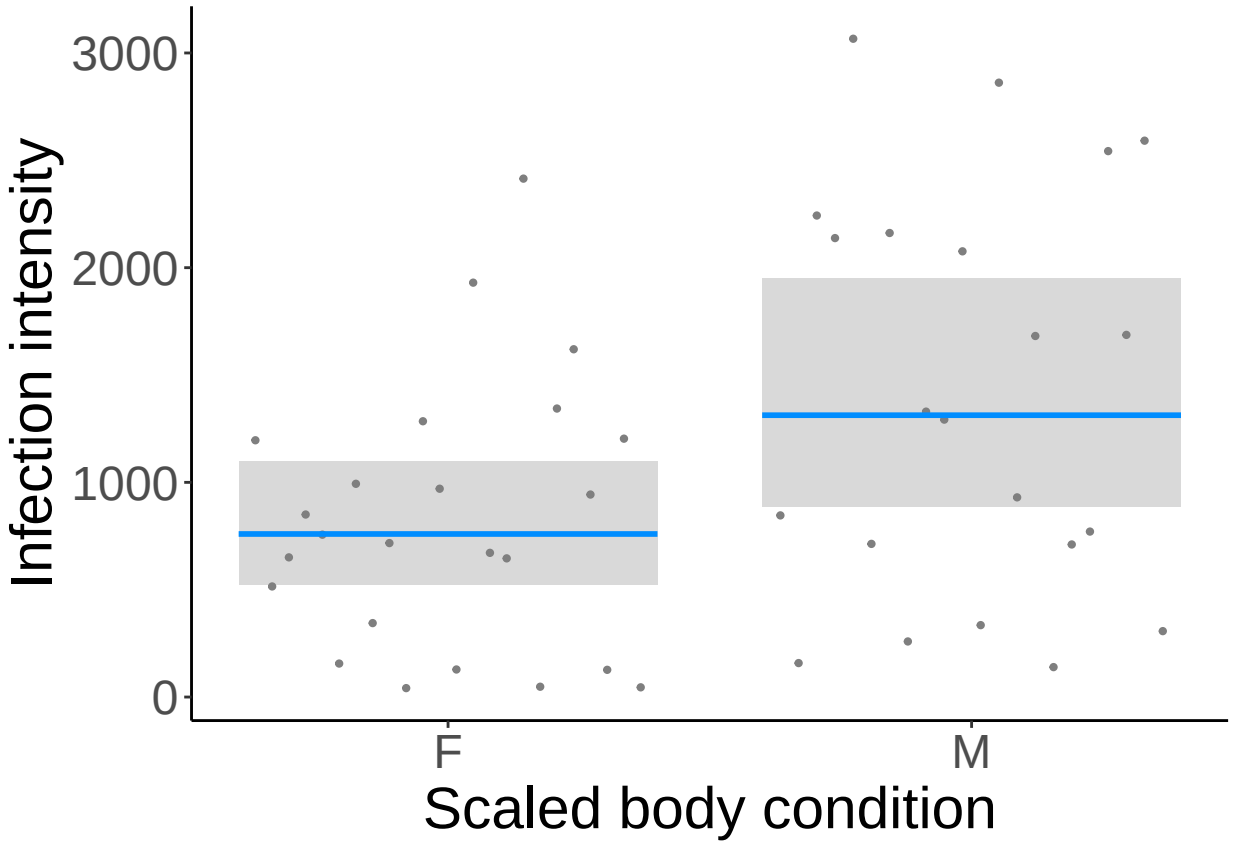



```
# Sex by SMI graph
InfIntbySex = visreg(AUCLM, scale = "response", "Sex", partial = T, gg = TRUE) +
    theme_classic() + theme(legend.position = "none") + ylab("Infection intensity") +
    xlab("Scaled body condition") + theme(text = element_text(size = 22))
```

```
## Warning:   Note that you are attempting to plot a 'main effect' in a model that contains an
##   interaction.  This is potentially misleading; you may wish to consider using the 'by'
##   argument.
```

```
## Conditions used in construction of plot
## ScBehavVig: -0.316863
## ScVarvelBef: -0.5119451
## ScResidPLSMI: -0.1015758
## ScRPLength: 0.04746859
## Treatment: UNTOUCHED
```

```
# Print the graph
print(InfIntbySex)
```


3000
2000
1000
0
Infection intensity
F
M
Scaled body condition