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

```r
# 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)
```

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())))
```

```
## New names:
## * 'fishID' -> 'fishID...2'
## * 'fishID' -> 'fishID...19'
## * 'AUC2' -> 'AUC2...21'
## * 'TrialTime' -> 'TrialTime...23'
## * 'Ratebf' -> 'Ratebf...24'
## * 'Recov' -> 'Recov...25'
## * 'TrialTime' -> 'TrialTime...31'
## * 'AUC2' -> 'AUC2...37'
## * 'Recov' -> 'Recov...40'
## * 'Ratebf' -> 'Ratebf...48'
```

```
# Filtering Data set so we can perform our analysis
IndBehav1 <- IndBehav %>%
    # Removing columns we dont need
select(-c(Blind.ID, fishID...19, TrialTime...31, AUC2...37, InfLength, Recov...40,
    Ratebf...48, LateLength, dayofinf, Countintial, Notes, InfWeight)) %>%
    # Renaming some of the columns with funky names for ease of coding and
    # naming
rename(fishID = fishID...2, AUC = AUC2...21, TrialTime = TrialTime...23, Ratebf = Ratebf...24,
    Recov = Recov...25)
```

```
## select: dropped 12 variables (Blind.ID, InfWeight, InfLength, LateLength, dayofinf, ...)
## rename: renamed 5 variables (fishID, AUC, TrialTime, Ratebf, Recov)
```

1.5 Calculate metrics needed for further analyses

1.5.1 Calculating all body condition metrics

```
IndBehav1 <- IndBehav %>%
    select(-c(Blind.ID, fishID...19, TrialTime...31, AUC2...37, InfLength, Recov...40,
        Ratebf...48, LateLength, dayofinf, Countintial, Notes, InfWeight)) %>%
    rename(fishID = fishID...2, AUC = AUC2...21, TrialTime = TrialTime...23, Ratebf = Ratebf...24,
        Recov = Recov...25)
```

```
## select: dropped 12 variables (Blind.ID, InfWeight, InfLength, LateLength, dayofinf, ...)
```

```
## rename: renamed 5 variables (fishID, AUC, TrialTime, Ratebf, Recov)
```

```
##### 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 541 rows (49%), 564 rows remaining
```

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

```
## filter: removed 565 rows (51%), 540 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 = 564   r = 0.9103118   r-square = 0.8286675
## Parametric P-values:   2-tailed = 1.890218e-217   1-tailed = 9.451089e-218
## Angle between the two OLS regression lines = 4.099607 degrees
##
## Regression results
##   Method Intercept     Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -1.114762 0.4206982         22.81640                NA
## 2     MA -1.155627 0.4341311         23.46717                NA
## 3    SMA -1.240856 0.4621474         24.80390                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS      -1.163000       -1.066525  0.4048487   0.4365478
## 2     MA      -1.205686       -1.106162  0.4178711   0.4505866
## 3    SMA      -1.289899       -1.193466  0.4465696   0.4782687
##
## Eigenvalues: 0.009625885 0.0002518475
##
## H statistic used for computing C.I. of MA: 0.0001893902
```

```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(PreSMI = PreWeight * ((20.2/PreLength)^0.4206982))
```

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

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

```r
# 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 = 540   r = 0.7784225   r-square = 0.6059415
## Parametric P-values:   2-tailed = 7.086423e-111    1-tailed = 3.543211e-111
## Angle between the two OLS regression lines = 6.76859 degrees
##
## Regression results
##   Method  Intercept      Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS -0.4620565 0.1938185         10.96898                NA
## 2     MA -0.4750867 0.1984879         11.22660                NA
## 3    SMA -0.6160126 0.2489888         13.98170                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS     -0.4990055      -0.4251076  0.1805813   0.2070557
## 2     MA     -0.5130172      -0.4373522  0.1849657   0.2120803
## 3    SMA     -0.6539329      -0.5800547  0.2361033   0.2625776
##
## Eigenvalues: 0.00431447 9.773707e-05
##
## H statistic used for computing C.I. of MA: 0.0001701002
```

```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(PreSMI = PreWeight * ((15.3/PreLength)^0.1938185))
```

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

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

Late Infection SMI####

```r
# Late infection SMI for females Take OLS 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 = 516   r = -0.127178   r-square = 0.01617424
## Parametric P-values:   2-tailed = 0.003807557    1-tailed = 0.001903778
## Angle between the two OLS regression lines = 57.58662 degrees
##
## Regression results
##   Method  Intercept       Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS   2.033718  -0.5976124       -30.86306                NA
## 2     MA 107.560934 -35.3034151       -88.37748                NA
## 3    SMA  14.504560  -4.6990243       -77.98610                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept  2.5%-Slope 97.5%-Slope
## 1    OLS      0.8050897        3.262347   -1.001497  -0.1937279
## 2     MA     64.2522935      331.105245 -108.822693 -21.0600644
## 3    SMA     13.3291795       15.785299   -5.120234  -4.3124649
##
## Eigenvalues: 0.1896132 0.008435402
##
## H statistic used for computing C.I. of MA: 0.0003658854
```

```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)^-0.5976124))
```

```
## 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 = 456   r = -0.06095677   r-square = 0.003715728
## Parametric P-values:   2-tailed = 0.1938354    1-tailed = 0.09691771
## Angle between the two OLS regression lines = 64.27059 degrees
##
## Regression results
##   Method  Intercept       Slope Angle (degrees) P-perm (1-tailed)
## 1    OLS   1.492739   -0.472241       -25.27860                NA
## 2     MA 348.385697 -124.982883       -89.54158                NA
## 3    SMA  21.760991   -7.747145       -82.64494                NA
##
## Confidence intervals
##   Method 2.5%-Intercept 97.5%-Intercept 2.5%-Slope 97.5%-Slope
## 1    OLS     -0.4947141        3.480193  -1.185443    0.2409615
## 2     MA    138.8775568     -682.296809 244.961036  -49.7839131
## 3    SMA     19.8652425       23.839279  -8.493107   -7.0667022
##
## Eigenvalues: 0.204658 0.003396832
##
## H statistic used for computing C.I. of MA: 0.0001459962
```

```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)^-0.472241))
```

```
## 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)
```

1.5.2 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 beahvioral vigor. The second way we could caluclate 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 teh 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: 1369.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8571 -0.6195 -0.0263  0.5647  4.1663
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  fishID   (Intercept) 6.101e-01 0.781095
##           Wormbf      1.231e-05 0.003509 -0.42
##  Residual             5.539e-01 0.744248
## Number of obs: 541, groups:  fishID, 48
##
## Fixed effects:
```

```
##              Estimate Std. Error t value
## (Intercept)    1.906      0.111   17.17
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 4.0577 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

```r
##### 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)
```

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

# Calculating the mean and variance in velocity 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 velocity 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 84 rows (8%), 1,020 rows remaining
```

```
## filter (grouped): removed 750 rows (74%), 270 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 91 unique values and 0% NA
```

```
## distinct (grouped): removed 179 rows (66%), 91 rows remaining
```

```r
##### creating a dataframe to calculate the mean and variance of velocity 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 84 rows (8%), 1,020 rows remaining

## filter (grouped): removed 760 rows (75%), 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 velocity 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 84 rows (8%), 1,020 rows remaining

## filter (grouped): removed 761 rows (75%), 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 velocity 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 84 rows (8%), 1,020 rows remaining

filter (grouped): removed 789 rows (77%), 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 velocity 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

```
##### subset down to infected individuals only and calculate point tolerance
##### metrics for each timeframe. #####
IndBehav4Inf <- IndBehav4 %>%
    group_by(fishID) %>%
    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 != 44)
```

group_by: one grouping variable (fishID)

filter (grouped): removed 38 rows (42%), 53 rows remaining

mutate (grouped): new variable 'EChBe' (double) with 50 unique values and 8% NA

new variable 'LChBe' (double) with 50 unique values and 8% NA

new variable 'LtrChBe' (double) with 43 unique values and 21% NA

new variable 'ERatebf' (double) with 23 unique values and 11% NA

new variable 'LRatebf' (double) with 46 unique values and 11% NA

new variable 'LtrRatebf' (double) with 36 unique values and 25% NA

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

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

filter (grouped): removed 4 rows (2%), 208 rows remaining

```r
##### Make two dataframes with pivot longer for wormbf and velocity 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 != 44)
```

filter (grouped): removed 38 rows (42%), 53 rows remaining

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

select: dropped 24 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

filter (grouped): removed 4 rows (2%), 208 rows remaining

```r
##### 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 38 rows (42%), 53 rows remaining

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

select: dropped 24 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 21 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 (TrialTin
```

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

```
## distinct (grouped): removed 624 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 19 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)
```

```
## distinct (grouped): removed 624 rows (75%), 208 rows remaining
```

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

##### 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)
```

```r
# 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, ) %>%
    group_by(fishID) %>%
    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 53 rows (58%), 38 rows remaining

## mutate (grouped): new variable 'EChBe' (double) with 38 unique values and 0% NA

##                   new variable 'LChBe' (double) with 38 unique values and 0% NA

##                   new variable 'LtrChBe' (double) with 38 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

## group_by: one grouping variable (fishID)

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

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

```r
##### 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 53 rows (58%), 38 rows remaining

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

## select: dropped 24 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 53 rows (58%), 38 rows remaining

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

select: dropped 24 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 20 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

distinct (grouped): removed 456 rows (75%), 152 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 19 variables (Sex, PreWeight, PreLength, Treatment, Infection, ...)

distinct (grouped): removed 456 rows (75%), 152 rows remaining

```
##### Merge the two 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)
```

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

# Remove some unwanted columns to clean up dataset in preparation for analysis
IndBehav5 <- IndBehav5 %>%
    select(-c(ERatebf, LRatebf, LtrRatebf, NRatebf_Early, NRatebf_Late, NRatebf_Later))
```

```
## select: dropped 6 variables (NRatebf_Early, NRatebf_Late, NRatebf_Later, ERatebf, LRatebf, ...)
```

```
View(IndBehav5)
```

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.

1.6 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)
Velocity: The average velocity of each fish during their behavior trial (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.

AUC: The area under the curve of infection over the total infection trajectory for each individual.

Recov: Whether or not the individual recovered from infection during the experimental trial. 1 - recovered, 0 - did not recover

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)

EChBe: The change in velocity between the early infection behavioral trial and the before trial. (cm/s)

LChBe: The change in velocity between the Late infection behavioral trial and the before trial. (cm/s)

LtrChBe: The change in velocity between the Later infection behavioral trial and the before trial. (cm/s)

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.

BehavTol:The behavioral tolerance of each individual calculated as the slope of a random effect model where fishID is the random effect term.

BehavVig:The behavioral Vigor of each individual calculated as the intercept of a random effect model where fishID is the random effect term