Test Report

Researcher A

2022-05-13

# Background

Insert information provided on the Session setup page

# Statistical Methods

## Mixed Effects Modeling

A mixed effects model is well suited for evaluating statistical differences between multiple experimental (treatment) groups across multiple times points. This can be accomplished by using a 2 factor model with interaction where the model terms are treatment, time, and treatment\*time (interaction). Mixed effects models can incorporate the within subject correlation between different time points. Some common correlation structures include auto-regressive (AR1) correlation structure which assumes that the correlation between time points decays at an exponential rate, compound symmetry (CS) correlation structure assumes that the correlation between time points is constant between any given time points, and unstructured is the most flexible and has no constraints.

Mixed effects models can model the variance/covariance structure of each group separately, i.e. the correlation matrix for each group is not required to be the same. While many of the typical regression assumptions are still applicable with mixed models especially that the model residuals are normally distributed.

## Checking Model Assumptions

### Normality and Transformation

A Shapiro-Wilk test is conducted to determine if the residuals of a linear model with a treatment, time, and treatment\*time interaction term. If the Shapiro-Wilk test is rejected, then a Box-Cox transformation is conducted to suggest an appropriate transformation. Then another Shapiro-Wilk test is conducted on a linear model with a treatment, time, and treatment\*time term using the transformed data to ensure that the transformation helped address the normality assumption. If the Shapiro-Wilk test is rejected, then we recommend further discussion with a statistician as a transformation did not make data follow the normality assumption.

No transformation was required or A [insert transformation] was applied to the data.

## Checking for Similar Variance between Groups

### Basic Model

We will denote the treatment groups (doses) and the non-wild type vehicle as the basic model. For this application, we require that the variance for each of the groups in the basic model are similar. To verify this assumption, first the variance is determined for each group and at each time point, and then averaged across the time points. A likelihood ratio test (LRT) is conducted between a model that estimates one common variance and a model that estimates an individual variance for each group. If the LRT is rejected and there is a 2 fold change between any group and the pooled variance then a statistician should be consulted. Alternatively, if we fail to reject the LRT and there is a 3 fold change between any group and the pooled variance then a statistician should be consulted. Otherwise, the application will move forward to the next modeling step.

### Controls and Wild Type

Once the variance for groups within basic model is determined to be similar, then a similar procedure is followed as above. A LRT is conducted for a model that has a common variance and a model that a common variance for the groups in the basic model and a different variance for the controls and wild type. If the LRT is rejected and there is a 2 fold change between the variance of controls or wild type and the basic model, then mixed effects model will estimate multiple variance components for the differing group. Alternatively, if we fail to reject the LRT and there is a 3 fold change between the variance of controls or wild type and the basic model, then mixed effects model will estimate the different variance components for the differing groups. Otherwise, the application will move forward to fit the final mixed model with one single common varaince component.

All groups have similar variance or [insert group] had a different variance from the basic model.

## Selection of Correlation Sturcture

The Akaike Information Criterion (AIC) is used to select the correlation structure that is most appropriate. The AIC strikes a balance between model complexity and quality of the model fit. The candidate correlation structures are AR1 (ARH1), CS (CSH), Toeplitz (TOEP), and unstructured (UN).

[insert correlation structure] had the smallest AIC.

## Comparison between Experimental Groups

There are 9 group comparisons that have relevant interpretation:

| Label | Group\_1 | Group\_2 | Purpose of Comparison |
| --- | --- | --- | --- |
| A | Wild Type | Vehicle | Verify Disease Model |
| B | Positive Control | Vehicle | Verify Positive Control |
| C | Wild Type | Dose | Which doses are similar to Wild Type? |
| D | Vehicle | Dose | Which treatment doses are effective? |
| E | Wild Type | Positive Control | Does Positive Control reverse disease? |
| F | Positive Control | Dose | Which doses are similar to Positive Control |
| G | Dose | Dose | Do doses differ from each other? |
| H | Negative Control | Vehicle | Rule out matrix effect |
| I | Negative Control | Dose | If H is rejected, then show Negative Control is not as good as treatment |

## Multiple Hypothesis Testing

Within the contrasts that involve treatment doses there are multiple comparisons that are being made and thus adjusting the p-values for these comparisons is important to maintain the desired family-wise Type I error rate. In many cases, the p-value adjust method may vary based on the comparison, but a simulation based methods (determines critical values and p-values based on sampling from the multivariate t distribution) are available that are well suited for most scenarios.

In addition to testing these hypotheses at a specific time point, testing of these hypothesis can be conducted across all time points. This will double the number of hypothesis tests conducted for each group comparison. Adjusting for multiple comparisons is necessary and will be accommodated using the simulation based method.

## Technical Replicates

In experiments that have technical replication, all of the data will be used to check the normality assumption and subsequently determine the most appropriate Box-Cox transformation (treating technical replicates as independent). After this step, the technical replicates will be averaged at each time point for each subject. The resulting dataset will be used for all of the remaining analysis steps.

## Change from Baseline Analysis

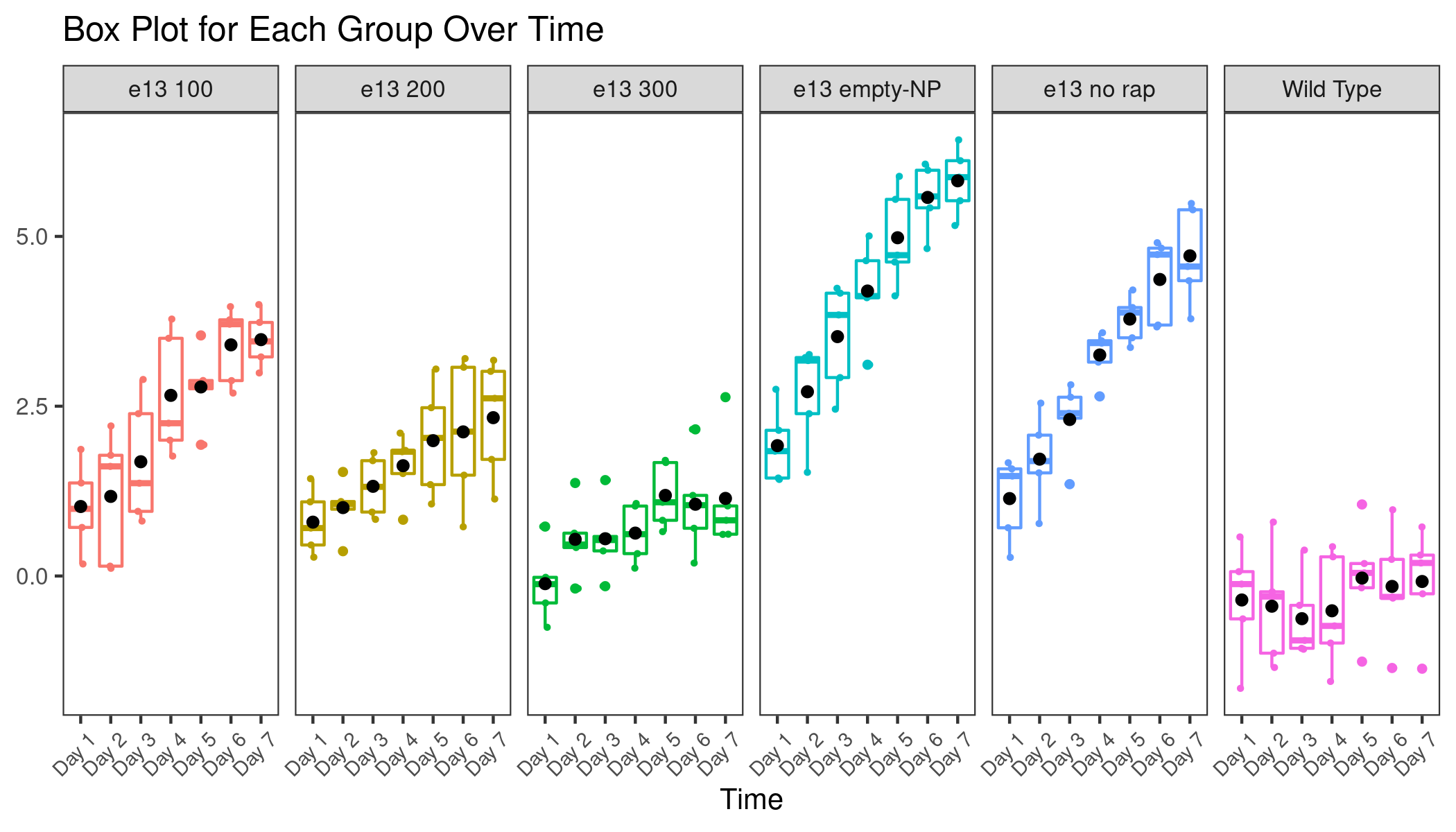
The user will be allowed to select whether the raw/transformed values are to be analysed or to analysis the change of each observation from a subject’s baseline measurement. The normality check and Box-Cox transformation will be conducted ignoring the baseline, then the change from baseline will be the response studied for the remainder of the analysis.

## Workflow Illustration

# Results

st <- storr\_rds("../storr")  
# # browser()  
# con <- connect\_table()  
# sessions <- tbl(con, 'sessions') %>%  
# arrange(desc(timestamp)) %>%  
# head(1)  
data <- st$get(  
 paste0(params$key)  
)

## Plots (One per page? or two per page?)



## Table 1

|  | | Original Scale | | |
| --- | --- | --- | --- | --- |
| Treatment | Times Included | Mean | Median | SE |
| e13 empty-NP | Average Over All Times | 4.10 | 4.16 | 0.092 |
| e13 empty-NP | Day 1 | 1.92 | 1.84 | 0.110 |
| Wild Type | Average Over All Times | -0.31 | -0.26 | 0.115 |
| Wild Type | Day 1 | -0.35 | -0.12 | 0.169 |

## Table 2

|  | | Original Scale | | |
| --- | --- | --- | --- | --- |
| Treatment | Times Included | Mean | Median | SE |
| e13 no rap | Average Over Time | 3.04 | 3.36 | 0.080 |
| e13 no rap | Day 1 | 1.14 | 1.47 | 0.123 |
| e13 100 | Average Over Time | 2.31 | 2.39 | 0.102 |
| e13 100 | Day 1 | 1.02 | 0.99 | 0.128 |
| e13 200 | Average Over Time | 1.60 | 1.48 | 0.093 |
| e13 200 | Day 1 | 0.79 | 0.71 | 0.094 |
| e13 300 | Average Over Time | 0.71 | 0.63 | 0.085 |
| e13 300 | Day 1 | -0.11 | -0.12 | 0.110 |

## Table 3