Test Report

Researcher A

2022-05-27

Table of Contents

[1 Session Summmary (Not final, maybe Freddy can help pass these as parameters into the markdown script) 2](#_Toc104567587)

[2 Background 2](#_Toc104567588)

[3 Statistical Methods 2](#_Toc104567589)

[3.1 Overview 2](#_Toc104567590)

[3.2 Mixed Effects Modeling 3](#_Toc104567591)

[3.3 Checking Model Assumptions 4](#_Toc104567592)

[3.3.1 Normality and Transformation 4](#_Toc104567593)

[3.4 Checking for Similar Variance between Groups 4](#_Toc104567594)

[3.4.1 Basic Model 4](#_Toc104567595)

[3.4.2 Controls and Wild Type 4](#_Toc104567596)

[3.5 Selection of Correlation Sturcture 5](#_Toc104567597)

[3.6 Comparison between Experimental Groups 5](#_Toc104567598)

[3.7 Multiple Hypothesis Testing 5](#_Toc104567599)

[3.8 Technical Replicates 6](#_Toc104567600)

[3.9 Change from Baseline Analysis 6](#_Toc104567601)

[4 Results 6](#_Toc104567602)

[4.1 Assumptions and Correlation Structure Results 6](#_Toc104567603)

[4.2 Plots 7](#_Toc104567604)

[4.3 Tables 12](#_Toc104567605)

[4.3.1 Table 1 12](#_Toc104567606)

[4.3.2 Table 2 13](#_Toc104567607)

[4.3.3 Table 3 14](#_Toc104567608)

[4.3.4 Table 4 15](#_Toc104567609)

[5 Appendix: Workflow Illustration 16](#_Toc104567610)

# 1 Session Summmary (Not final, maybe Freddy can help pass these as parameters into the markdown script)

|  |  |
| --- | --- |
| User Name |  |
| Consulting Statistician |  |
| Session ID | d24fad66-e6a3-4cf6-ba61-beae48eef1e1 |
| Results Location |  |

# 2 Background

Insert information provided on the Session setup page

# 3 Statistical Methods

## 3.1 Overview

This analysis is intended to compare groups from an experiment with multiple doses with data collected at multiple times. Pairwise comparisons will be conducted for relevant pairs and will be reported for a specific time point and over the entire course of the experiment. Mixed effects models will be used to model the differences between groups. The assumptions, for a linear mixed effects model,

• the explanatory variables are related linearly to the response,

• the errors have constant variance,

• the errors are independent, and

• the errors are normally distributed.

An explanation of how these are addressed is provided in the following sections.

Data is collected over time is referred to at longitudinal data. A common feature of repeated measurements on a subject is correlation which means that the value of the response on one occasion provides information about the likely value of the response on a future occasion. Also, data collected several measurements from a subject is referred to as repeated measures. A longitudinal analysis of within-individual change proceeds in 2 conceptually distinct stages. In the first stage, within-individual change is characterized in terms of some appropriate summary of the changes in the repeated measurements on each individual during the period of observation.

## 3.2 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.

## 3.3 Checking Model Assumptions

### 3.3.1 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.

## 3.4 Checking for Similar Variance between Groups

### 3.4.1 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 thedagroups 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.

### 3.4.2 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.

## 3.5 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).

## 3.6 Comparison between Experimental Groups

There are 9 group comparisons that have relevant interpretation:

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

## 3.7 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.

## 3.8 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.

## 3.9 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.

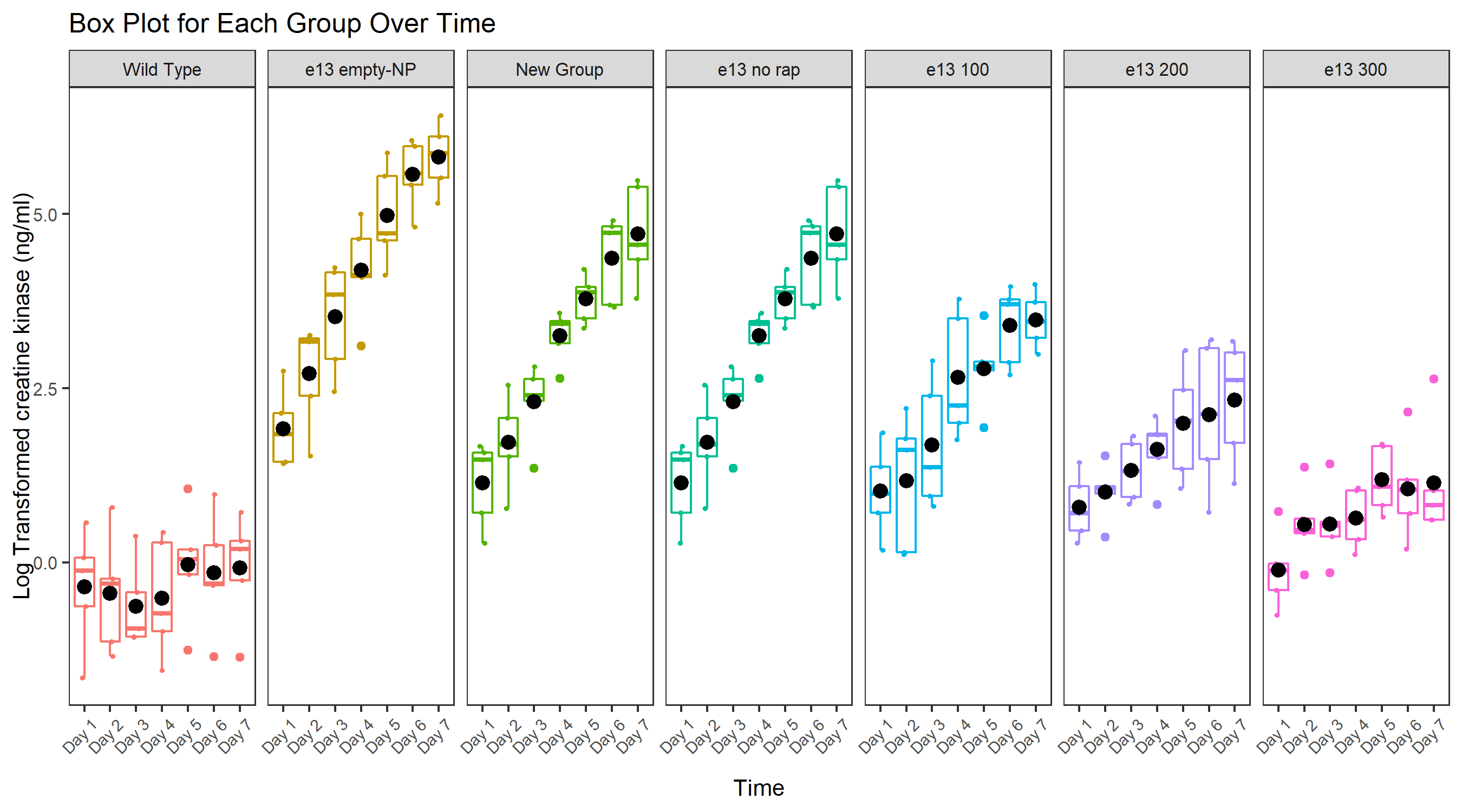
# 4 Results

## 4.1 Assumptions and Correlation Structure Results

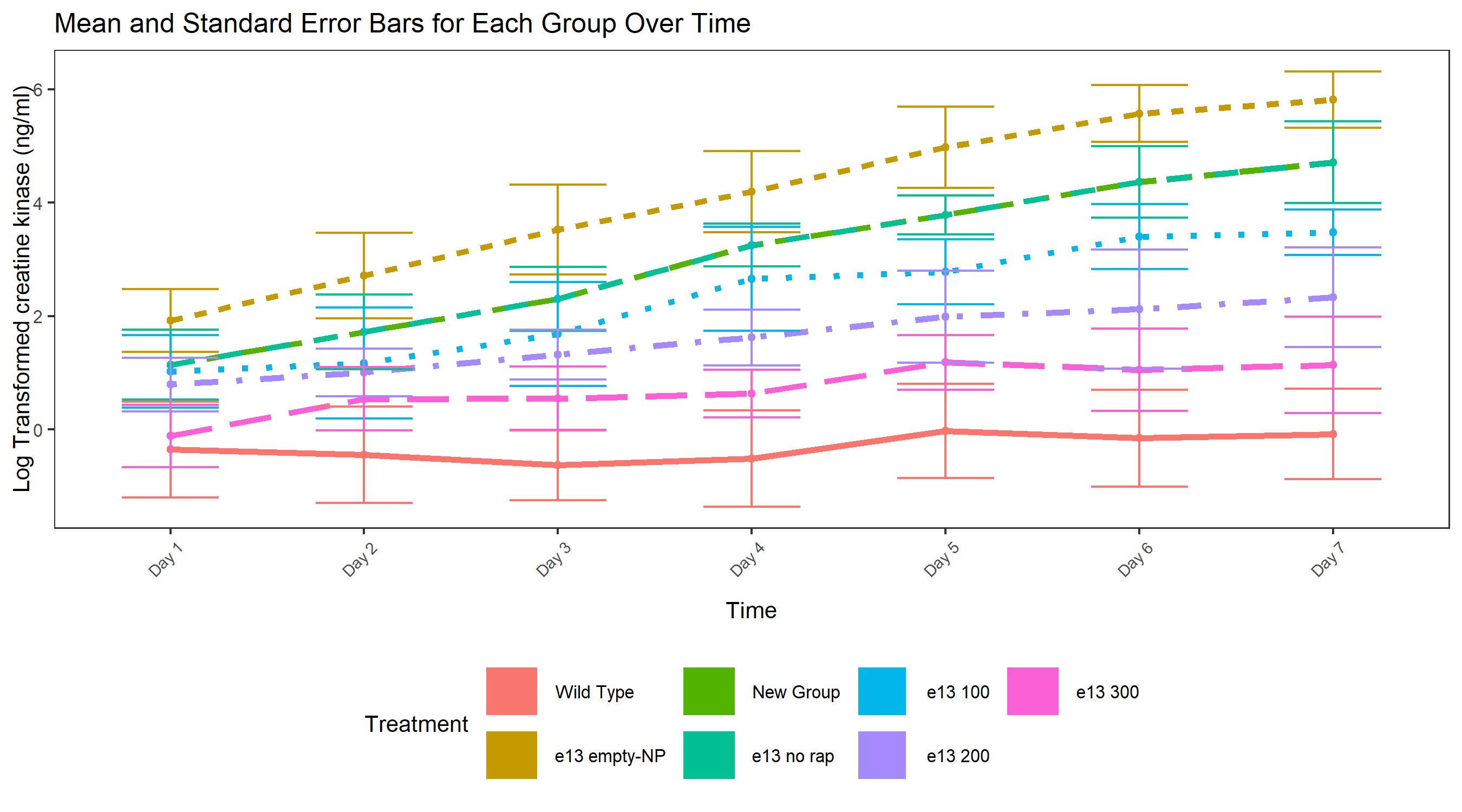
In order to meet the normality assumptions of the model, a log was applied to the data.. The correlation structure that lead to the lowest AIC was an AR1 correlation structure lead to the smallest AIC.

All groups have similar variance

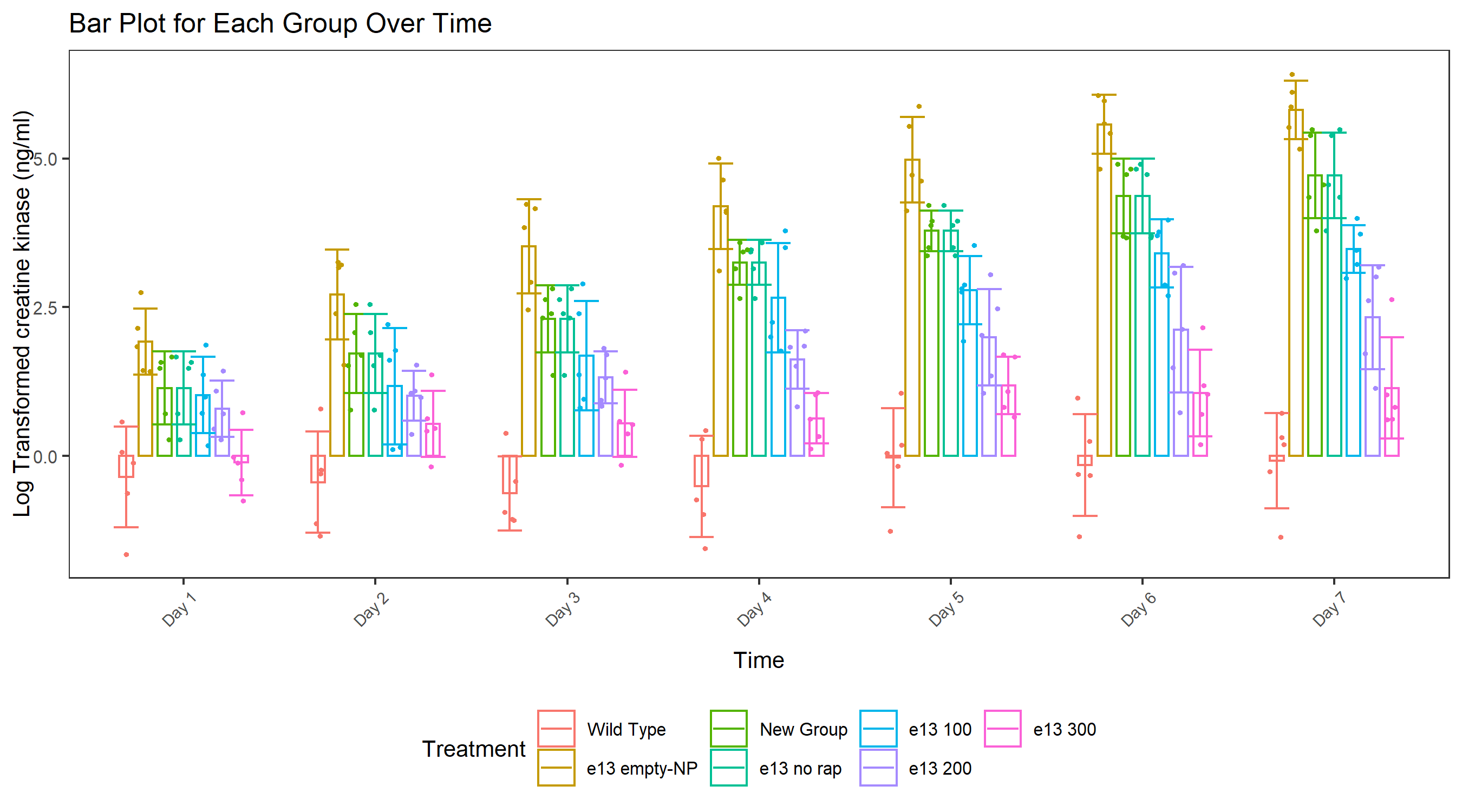
## 4.2 Plots



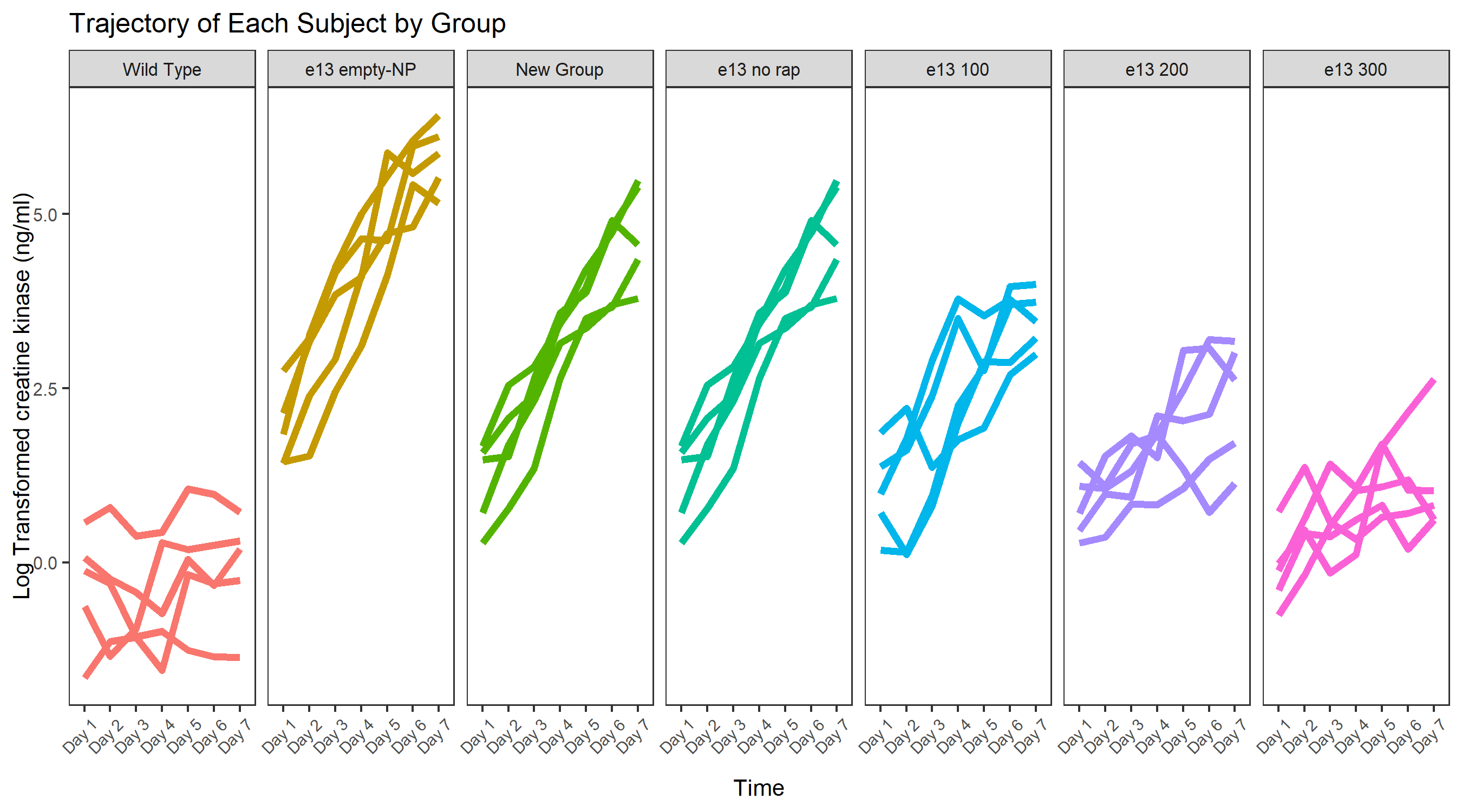
Box plot for each group and time.



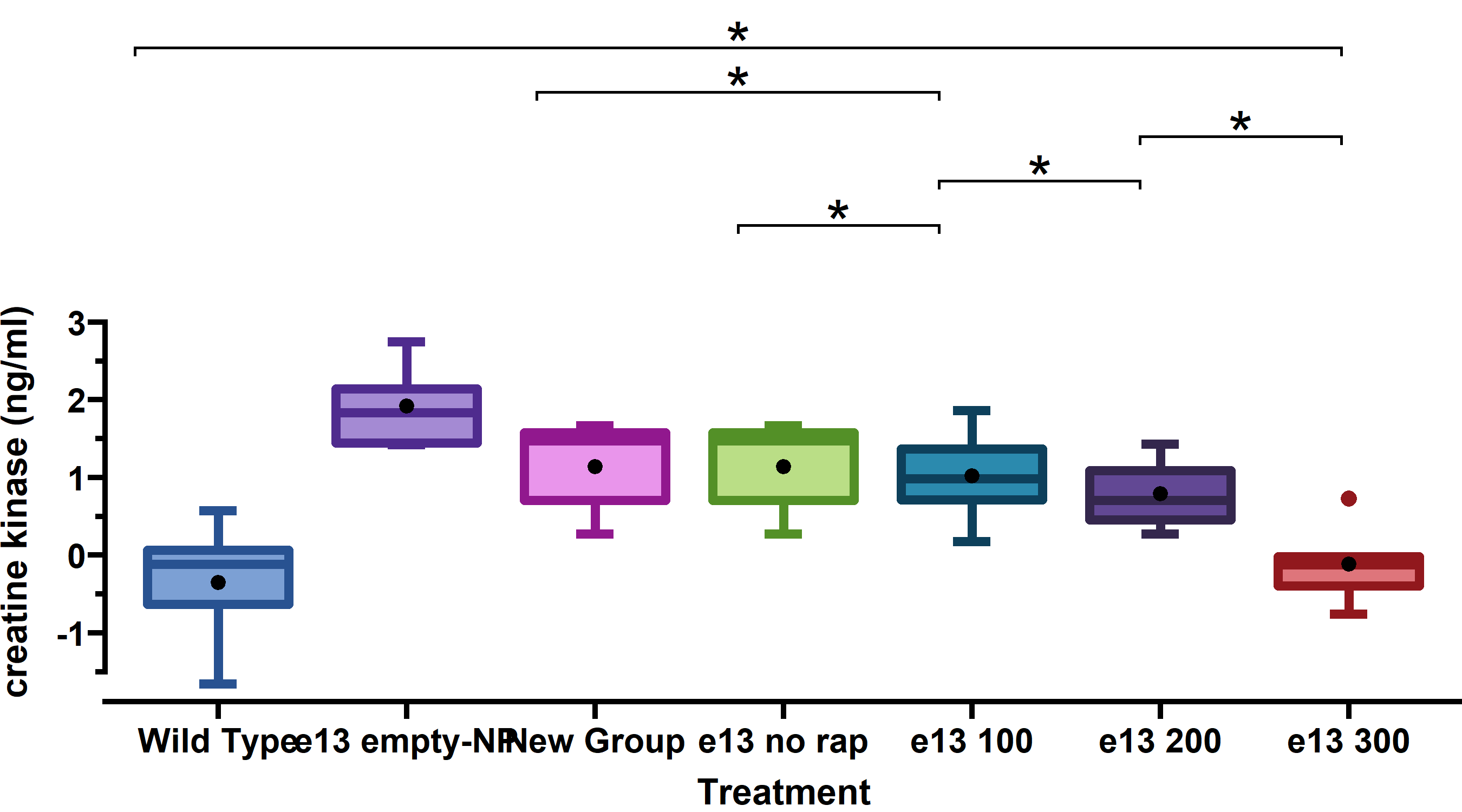
Group level trajectory across time where the line corresponds to the group average ofTransformed Scale. The vertical bars are standard error bars for each group at each time point.



Bar plot for each group and time.



Subject level trajectory across time where the line corresponds to the average of Transformed Scale response in the presence of techinical replicates.



Boxplot of Transformed Scale response values at Day 7. The stars above the horizontal bars correspond to the magnitude of the p-value comparing two groups where “\*”, “\*\*”, and “\*\*\*” represent 0.01 <= p-value < 0.05, 0.001 <= p-value < 0.01, and p-value < 0.001, respectively. A p-value summary will only be shown form p-value less than 0.05.

## 4.3 Tables

### 4.3.1 Table 1

|  | | Original Scale | | |
| --- | --- | --- | --- | --- |
| Treatment | Times Included | Mean | Median | SE |
| e13 empty-NP | Overall Average | 235.11 | 94.46 | 25.933 |
| e13 empty-NP | Day 7 | 546.40 | 420.56 | 70.626 |
| Wild Type | Overall Average | 1.12 | 1.03 | 0.110 |
| Wild Type | Day 7 | 1.44 | 1.34 | 0.163 |
| New Group | Overall Average | 61.58 | 30.52 | 5.961 |
| New Group | Day 7 | 181.00 | 95.30 | 29.765 |
| e13 no rap | Overall Average | 61.58 | 30.52 | 5.961 |
| e13 no rap | Day 7 | 181.00 | 95.30 | 29.765 |
| e13 100 | Overall Average | 25.63 | 16.69 | 1.968 |
| e13 100 | Day 7 | 56.15 | 55.76 | 4.532 |
| e13 200 | Overall Average | 8.75 | 5.12 | 0.906 |
| e13 200 | Day 7 | 17.17 | 14.05 | 2.684 |
| e13 300 | Overall Average | 2.88 | 2.23 | 0.302 |
| e13 300 | Day 7 | 4.86 | 2.79 | 1.064 |
| Group level summary on the original scale. | | | | |

### 4.3.2 Table 2

|  | | Transformed Scale | | Back Transform | | vs. e13 no rap | | vs. New Group | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Treatment | Times Included | Mean | SE | Transformed Mean | Transformed SE | LSMean Diff (95% CI) | p value | LSMean Diff (95% CI) | p value |
| e13 empty-NP | Overall Average | 4.10 | 0.228 | 60.55 | 13.83 | 1.06 (0.327, 1.8) | 0.004 |  |  |
| e13 empty-NP | Day 7 | 5.82 | 0.300 | 336.49 | 100.93 | 1.1 (0.159, 2.051) | 0.02 |  |  |
| Wild Type | Overall Average | -0.31 | 0.228 | 0.73 | 0.17 | -3.35 (-4.092, -2.618) | < 0.001 | -3.35 (-4.092, -2.618) | < 0.001 |
| Wild Type | Day 7 | -0.08 | 0.300 | 0.92 | 0.28 | -4.79 (-5.741, -3.849) | < 0.001 | -4.79 (-5.741, -3.849) | < 0.001 |
| New Group | Overall Average | 3.04 | 0.228 | 20.91 | 4.78 | 0 (-0.737, 0.737) | 1 |  |  |
| New Group | Day 7 | 4.71 | 0.300 | 111.46 | 33.43 | 0 (-0.946, 0.946) | 1 |  |  |
| A Log transformation was applied to the data. Difference and CI are estimated using model based LSmean | | | | | | | | | |

### 4.3.3 Table 3

|  | | Transformed Scale | | Back Transform | | vs. e13 100 | | vs. e13 200 | | vs. e13 300 | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Treatment | Times Included | Mean | SE | Transformed Mean | Transformed SE | LSMean Diff (95% CI) | p value | LSMean Diff (95% CI) | p value | LSMean Diff (95% CI) | p value |
| e13 no rap | Overall Average | 3.04 | 0.228 | 20.91 | 4.78 | 0.73 (-0.146, 1.597) | 0.133 | 1.44 (0.57, 2.313) | < 0.001 | 2.33 (1.455, 3.199) | < 0.001 |
| e13 no rap | Day 7 | 4.71 | 0.300 | 111.46 | 33.43 | 1.23 (0.119, 2.349) | 0.024 | 2.38 (1.269, 3.497) | < 0.001 | 3.57 (2.458, 4.684) | < 0.001 |
| e13 100 | Overall Average | 2.31 | 0.228 | 10.12 | 2.31 |  |  | 0.72 (-0.154, 1.586) | 0.137 | 1.6 (0.732, 2.471) | < 0.001 |
| e13 100 | Day 7 | 3.48 | 0.300 | 32.45 | 9.73 |  |  | 1.15 (0.038, 2.26) | 0.04 | 2.34 (1.226, 3.449) | < 0.001 |
| e13 200 | Overall Average | 1.60 | 0.228 | 4.95 | 1.13 |  |  |  |  | 0.89 (0.016, 1.755) | 0.045 |
| e13 200 | Day 7 | 2.33 | 0.300 | 10.29 | 3.09 |  |  |  |  | 1.19 (0.079, 2.299) | 0.031 |
| e13 300 | Overall Average | 0.71 | 0.228 | 2.04 | 0.47 |  |  |  |  |  |  |
| e13 300 | Day 7 | 1.14 | 0.300 | 3.13 | 0.94 |  |  |  |  |  |  |
| A Log transformation was applied to the data. Difference and CI are estimated using model based LSmean | | | | | | | | | | | |

### 4.3.4 Table 4

|  | | vs. Wild Type | | vs. New Group | | vs. e13 empty-NP | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Treatment | Times Included | LSMean Diff (95% CI) | p value | LSMean Diff (95% CI) | p value | LSMean Diff (95% CI) | p value |
| e13 100 | Overall Average | 2.63 (1.758, 3.501) | < 0.001 | -0.73 (-1.597, 0.146) | 0.133 | -1.79 (-2.662, -0.916) | < 0.001 |
| e13 100 | Day 7 | 3.56 (2.448, 4.674) | < 0.001 | -1.23 (-2.347, -0.121) | 0.024 | -2.34 (-3.453, -1.225) | < 0.001 |
| e13 200 | Overall Average | 1.91 (1.041, 2.785) | < 0.001 | -1.44 (-2.314, -0.57) | 0.001 | -2.51 (-3.377, -1.633) | < 0.001 |
| e13 200 | Day 7 | 2.41 (1.298, 3.526) | < 0.001 | -2.38 (-3.497, -1.269) | < 0.001 | -3.49 (-4.601, -2.374) | < 0.001 |
| e13 300 | Overall Average | 1.03 (0.156, 1.9) | 0.016 | -2.33 (-3.199, -1.455) | < 0.001 | -3.39 (-4.262, -2.518) | < 0.001 |
| e13 300 | Day 7 | 1.22 (0.11, 2.337) | 0.026 | -3.57 (-4.684, -2.458) | < 0.001 | -4.68 (-5.789, -3.563) | < 0.001 |
| A Log transformation was applied to the data. Difference and CI are estimated using model based LSmean | | | | | | | |

# 5 Appendix: Workflow Illustration

