

## Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature: \_\_\_\_\_  
Katherine Catevenis

\_\_\_\_\_  
Date

Analysis of the Effect of C3 Transferase  
on Behavioral Outcomes with Informative Right-Censoring

By

Katherine Catevenis  
Master of Science in Public Health  
Biostatistics

---

John J. Hanfelt, Ph.D.  
Thesis Advisor

Analysis of the Effect of C3 Transferase  
on Behavioral Outcomes with Informative Right-Censoring

By

Katherine Catevenis  
B.A.  
Washington and Jefferson College  
2010

Thesis Committee Chair: John J. Hanfelt, Ph.D.

An abstract of  
A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science in Public Health  
in Biostatistics  
2012

## Abstract

### Analysis of the Effect of C3 Transferase on Behavioral Outcomes with Informative Right-Censoring

By Katherine Catevenis

This study aimed to slow the degradation rates of the motor functions of mice with Amyotrophic Lateral Sclerosis (ALS). Two treatment groups, high titer AAV2 C3 and low titer AAV2 C3, were investigated to determine their effect on the degradation of the SOD1 mice, measured through the Basso-Beattie-Bresnahan (BBB) locomotor rating. In this experiment each mouse served as its own control through one side being treated and the other side used as a control. An analysis that adjusted for the correlation between the sides of each animal and the presence of informative right censoring was used. This method used an Empirical Bayes approach and was an extension of the method used by Pak [1999]. Twenty SOD1 gene mice were injected with their treatment on one side. Eight mice received the high titer treatment, eight mice received the low titer treatment, and four mice received a control GFP treatment. The BBB score was measured once before the injection and then on specific days after the injection. A test of informative right censoring was performed for each side of the mice for each treatment group. None of the treatment groups were found to have informative right censoring. Three types of slope estimates were compared for each animal, ordinary least squares, Empirical Bayes assuming no informative right censoring, and Empirical Bayes adjusted for informative right censoring. None of the treatment groups showed a significant difference between the decline of function in the treated versus the untreated side ( $p > 0.70$ ).

Analysis of the Effect of C3 Transferase  
on Behavioral Outcomes with Informative  
Right-Censoring

By  
Katherine Catevenis  
B.A.  
Washington and Jefferson College  
2010

Thesis Committee Chair: John J. Hanfelt, Ph.D.

A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science in Public Health  
in Biostatistics  
2012

## Acknowledgements

I am extremely grateful for my amazing thesis advisor, Dr. John J. Hanfelt, whose encouragement, guidance, and support made this thesis possible.

I am also thankful for Rebecca Zhang for taking time to read my thesis.

Finally, I would like to thank my parents, Dr. James and Gloria Catevenis, for their love and support throughout my education.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Methods</b>	<b>5</b>
2.1	Basic Model and Assumptions . . . . .	5
2.2	Implications of the Model . . . . .	7
2.2.1	Distribution of OLS Estimates of Slope . . . . .	7
2.2.2	Posterior Distribution of the True Slopes . . . . .	8
2.3	Inference . . . . .	9
<b>3</b>	<b>Results</b>	<b>14</b>
3.1	Checking Assumptions . . . . .	14
3.2	Model Based Results . . . . .	14
<b>4</b>	<b>Discussion</b>	<b>20</b>
4.1	Discussion of Results . . . . .	20
4.2	Further Work . . . . .	20
<b>5</b>	<b>Appendices</b>	<b>23</b>
<b>A</b>	<b>Figures</b>	<b>23</b>
<b>B</b>	<b>R Code</b>	<b>28</b>

## List of Tables

1	High Treatment Group Slope Estimates . . . . .	15
2	Low Treatment Group Slope Estimates . . . . .	16
3	GFP Treatment Group Slope Estimates . . . . .	17
4	Variance Component Estimates . . . . .	17
5	Testing Informative Right Censoring . . . . .	18
6	Comparison of Average Un-weighted Estimate and Average Empirical Bayes Estimate . . . . .	19
7	Testing Difference in Average Treated vs. Average Untreated Sides of the Mice in Each Treatment Group . . . . .	19

## List of Figures

1	BBB Score by Day After Injection . . . . .	23
2	Average BBB Score by Day By Side of Animal for All Treatment Groups . .	24
3	OLS Slope By Dropout Time By Side of the Animal for All Treatment Groups	25
4	Empirical Bayes Slope Estimate By Dropout Time By Side of the Animal for All Treatment Groups . . . . .	26
5	Boxplots of Residuals by Timeperiod by Side of the Animal for All Treatment Groups . . . . .	27

# 1 Introduction

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disorder that affects both the upper and lower motor neurons. The disease is extremely progressive with the mean survival time from diagnosis being only three to five years [Rowland and Shneider, 2001]. It is estimated that 20,000 - 30,000 people are living with ALS in the United States and approximately 5,000 are diagnosed each year. There are two types of ALS, sporadic and familial. Sporadic ALS (SALS) accounts for approximately 90% of ALS cases. These cases have no family history and no clear risk factors for the disease [ALS, 2010]. The remaining 10% of ALS cases are categorized as familial ALS (FALS). These cases are genetic and several genes and loci have been identified to cause the clinical phenotype [Kiernan et al., 2011].

One gene that has been identified to cause the clinical phenotype is the SOD1 gene. The SOD1 gene normally encodes for copper-zinc ion-binding superoxide dismutase. However, mutations in this gene cause FALS and accounts for approximately 20-25% of FALS cases. More than 100 mutations have been identified thus far [Matsumoto et al., 2006].

In ALS there is an abnormal activation of a cellular pathway, the Rho kinase (ROCK) pathway. This pathway inhibits regeneration of the motor neuron axons. It has been proven that C3 transferase inhibits the Rho kinase pathway. This inhibition suggests that inhibition of the Rho kinase pathway may be a way to promote axonal regeneration [Gopalakrishnan et al., 2008].

The primary purpose of this study was to determine if direct spinal chord injection of C3 transferase using a viral vector would promote regeneration or delay degeneration of motor neuron axons, therefore slowing the disease progress. The secondary purposes of the study were:

1. To estimate the rate of degradation for both the treated side and the untreated side of the mice;
2. To take into account the informative dropout in order to reduce the bias of the esti-

mated rate of degradation; and

3. To estimate the correlation in the degradation rate between the treated and untreated sides of the mice.

In a study performed by Hoberman et al. [2011] twenty mice with SOD1 G93A, a mutation of the SOD1 gene, were separated into three treatment groups. Each mouse received a 2 micro-liter unilateral spinal chord injection at the level of lumbar enlargement of one of the three treatments. A viral vector, AAV2, was used to deliver the C3 transferase directly into the spinal chord. The first treatment group, High, was given high titer AAV2 C3 ( $1.0 \times 10^{12}$  PFU/mL AAV2-C3-GFP). The second group, Low, was given low titer AAV2 C3 (1:10 dilution of AAV2-C3-GFP). The third group, GFP, was given control AAV2-GFP. GFP, green fluorescent protein, is used as a marker for expression. The mice were given the injection on only one side of their spinal chord, so each mouse served as its own control. There were several forms of behavioral testing used in the study. The first was the Basso-Beattie-Bresnahan (BBB) locomotor rating. The BBB rating can range from 0, which stands for complete paralysis, to 21, which stands for normal locomotion. The second form of behavioral testing used was called the ALS motor score, developed by Matsumoto et al. [2006]. This scoring system is used after disease onset and ranges from 0, extreme motor dysfunction, to 5, normal motor function. The third form of behavioral testing is hindgrip strength measurement. Also measured are the mice's weight and survival time. At disease endpoint, the mice were perfused with 4% paraformaldehyde. Their spinal chord, L4 ventral roots, sciatic nerves and gastrocnemius muscle were harvested and analyzed for counts of motor neurons and axons. The researchers conducting this study were blinded to the behavioral testing and the tissue analysis.

In order to analyze the data from this study a new method needs to be introduced. This method must take into account different aspects of the data.

The first aspect is that the data are longitudinal. Longitudinal data are one type of correlated data, where the response variable is measured repeatedly over time. These repeated

observations introduce the complicating feature: their correlation. In other words the data are no longer independent, so standard methods of analysis cannot be used. In order to obtain a precise estimate of the C3 transferase effect and a valid estimate of its standard error, taking into account their correlation is extremely important. A naïve analysis, or one that does not take into account their correlation, may lead to incorrect inferences.

The second aspect is that one side of the animal's brain is treated while the other side is left untreated. In this study, the behavioral testing data on each side of the mouse will generally be correlated due to the shared environment. Ignoring this correlation could result in inefficient estimators of the degradation rate and a biased standard error. Hence, this additional correlation needs to be taken into account in the statistical model.

Informative right censoring is the third aspect of the data that needs to be taken into account. In many longitudinal studies, some subjects may not be followed for the entire study time frame due to death or withdrawal. When this occurs, the outcome variable is right censored. Many of the statistical methods that exist assume that the right censoring is uninformative. In other words, that the right censoring is independent of the outcome variable. For this study, it would mean that the death of a mouse is not associated with its rate of degradation. However, the rate of degradation may be associated with the reason for the mouse's death or withdrawal. Due to the right censoring, the missing data cannot be ignored. Furthermore, since informative right censoring is present, standard likelihood based analyses and general multivariate repeated measures analyses are not appropriate. These analyses could result in biased and misleading estimations since they would place more weight on animals that completed the study rather than those who died early.

In the analysis, the data should be tested for whether the right censoring is informative or not. If right censoring and the rate of degradation are not independent then the analysis should adjust appropriately.

In order to account for the aspects of the data discussed, the statistical methodology of Pak [1999] is extended in order to allow different degradation rates for each side of the animal.

## 2 Methods

### 2.1 Basic Model and Assumptions

Measurements of the BBB locomotor rating were recorded at a baseline time, considered time 0, and also on separate days after the injection,

$t = (0, 4, 7, 11, 14, 18, 21, 26, 28, 32, 35, 39, 42, 46, 49, 53, 56, 60, 63, 67, 70, 74)$ .

Each animal had a different follow up time depending on when they were censored. There was one mouse, mouse 4, whose follow up was started on day 3 after the injection. In order to examine the relationship between the two doses of C3 transferase and the rate of degradation, an Empirical Bayes approach was selected. The basic model starts with the response variable, BBB locomotor rating, and the explanatory variable, number of days post injection. The BBB locomotor ratings observed in a specific treatment group are denoted by  $Y_{ijk}, i = 1, \dots, n; j = 1, \dots, J_i; k = 1, \dots, m_{ij}$  where subscript  $i$  refers to the given animal,  $j$  refers to the side of the animal, and  $k$  refers to the longitudinal measurement. Therefore, a specific treatment group has  $n$  animals,  $J_i$  sides to the animal, and  $m_{ij}$  measurements per side. In all cases, the animal will have 2 sides ( $J_i = 2$ ), a treated side ( $J_i = 1$ ) and an untreated side ( $J_i = 2$ ). Three distributional assumptions were made regarding the response variable, the BBB locomotor scale, as follows:

1. The expectation of the response variable is linear in time.
2. Each side of the animal has its own true degradation rate,  $\{\beta_{ij}\}$ ; and
3. The response variables, conditional on  $\{\beta_{ij}\}$ , are observed with errors that are mutually independent identically distributed normal random variables.

These assumptions imply that:

$$Y_{ijk} | \beta_{ij} = \alpha_{ij} + \beta_{ij} t_{ijk} + \epsilon_{ijk}, \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2) \quad (1)$$

In addition to equation 1, we model the animal's true, but unknown, degradation rates,  $\{\beta_{ij}\}$ , as normal random variables. The degradation rates have marginal variances,  $A_1$  and  $A_2$ , for the treated and untreated sides respectively and within a given animal, a common pairwise correlation,  $\rho\sqrt{A_1A_2}$ . The similarity between the degradation rates on each side of the animal is measured by the parameter  $\rho$ . It is not uncommon for some animals to die before the end of the study due to reasons related to their degradation rate. For example, animals with fast degradation rates may tend to die or be sacrificed for ethical reasons before the end of the study. The time of last observation, or dropout time, is denoted by  $t_{ij}^* = t_{ijk_{m_{ij}}}$ . Following Pak [1999], the assumption that an animal's degradation rate,  $\beta_{ij}$ , is linearly related to its dropout time,  $t_{ij}^*$ , was applied. Therefore the probability model for the animal's degradation rates is:

$$\begin{bmatrix} \beta_{i1} \\ \beta_{i2} \end{bmatrix} \sim N_2 \left( \mu_i = \begin{bmatrix} \gamma_0 + \gamma_1 t_{i1}^* \\ \gamma_2 + \gamma_3 t_{i2}^* \end{bmatrix}, W = \begin{bmatrix} A_1 & \rho\sqrt{A_1A_2} \\ \rho\sqrt{A_1A_2} & A_2 \end{bmatrix} \right) \quad (2)$$

Note that if  $\gamma_1$  and  $\gamma_3$  are not equal to 0 then dropout is informative and biased results may be obtained if naïve methods are used to estimate the animal's degradation rate. One example is an experiment where animals with fast degradation rates tend to die earlier than animals with slow degradation rates, in which case  $\gamma_1$  and  $\gamma_3$  are less than 0. By contrast, a standard random-effects analysis that adapts equations 1 and 2 but naïvely assumes that  $\gamma_1$  and  $\gamma_3$  are equal to 0 would lead to the underestimation of the true average degradation rates in the population. The goal of this analysis is to conduct inference on the true, individual animal slopes,  $\{\beta_{ij}\}$ , and the average slopes

$$\bar{\beta} = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n \beta_{i1} \\ \frac{1}{n} \sum_{i=1}^n \beta_{i2} \end{bmatrix} \quad (3)$$

based on the model specified by equations 1 and 2. Prior to estimating the true degradation rates, the previous model assumptions will be assessed using several plots. In order to

validate the linear relationship between the response variable, BBB locomotor scale, and the time variable several scatter plots will be examined. A crude assessment of informative right censoring can also be obtained from these plots. However, formal t-tests will be used to test for informative right censoring later in the analysis. To comply with model assumptions, the variance of the slope estimates needs to be constant over time. If animal specific slopes were not allowed in this model, then the variances would fan out because of the longitudinal aspect of the data. In other words, there would not be constant variance over time. Residuals take into account the individual slopes and therefore, should have constant variance if there is no dropout. A boxplot of the residuals grouped into 4 timepoints will be used to validate this assumption.

## 2.2 Implications of the Model

When informative right censoring is present, to obtain estimates of the true degradation rate using standard software such as PROC MIXED, PROC GLM, PROC GLIMMIXED, or PROC GENMOD in SAS is inappropriate. Informative right censoring needs to be taken into account in the analysis as well as the multiple sides per animal, ie each animal serving as its own control. Here, the R programming environment was used.

### 2.2.1 Distribution of OLS Estimates of Slope

The ordinary least-squares estimates of an animal's degradation rate is given by

$$b_{ij,OLS} = \frac{\sum_{k=1}^{m_{ij}} (t_{ijk} - \bar{t}_{ij})(y_{ijk} - \bar{y}_{ij\bullet})}{\sum_{k=1}^{m_{ij}} (t_{ijk} - \bar{t}_{ij\bullet})}, i = 1, \dots, n; j = 1, 2 \quad (4)$$

Following equation 1, the ordinary least-squares estimates,  $\{b_{ij,OLS}\}$ , given the true slopes,  $\{\beta_{ij}\}$ , are conditionally independent normally distributed random variables. In mathematical notation:

$$b_{ij,OLS}|\beta_{ij} \stackrel{\text{iid}}{\sim} N(\beta_i, V_{ij}) \quad \text{where} \quad V_{ij} = \frac{\sigma^2}{\sum (t_{ijk} - \bar{t}_{ij\bullet})^2} \quad (5)$$

As implied by equations 2 and 5, the marginal distribution of the ordinary least-squares estimates is

$$\hat{\beta}_i \sim N_2(\mu_i, V_i + W) \quad (6)$$

$$\text{where } \mu_i = \begin{bmatrix} \gamma_0 + \gamma_1 t_{i1}^* \\ \gamma_2 + \gamma_3 t_{i2}^* \end{bmatrix}$$

$$W = \begin{bmatrix} A_1 & \rho\sqrt{A_1 A_2} \\ \rho\sqrt{A_1 A_2} & A_2 \end{bmatrix}$$

$$V_i = \begin{bmatrix} V_{i1} & 0 \\ 0 & V_{i2} \end{bmatrix}$$

### 2.2.2 Posterior Distribution of the True Slopes

Also continuing from equations 2 and 5, the posterior distribution of the true slopes,  $\{\beta_{ij}\}$ , given the ordinary least squares estimates,  $\{b_{ij,OLS}\}$ , can be derived [Gelman et al., 1995].

$$\beta_i|\hat{\beta}_i \sim N_2(\mu_i^*, W_i^*) \quad (7)$$

$$\text{where } \mu_i^* = (W^{-1} + V_i^{-1})^{-1}(W^{-1}\mu_i + V_i^{-1}\hat{\beta}_i)$$

$$W_i^* = (W^{-1} + V_i^{-1})^{-1}$$

The posterior distribution takes into account the ordinary least squares estimates of the slopes as well as the dropout times for each animal.

## 2.3 Inference

By using the methods of Pak [1999], inference can be conducted on the model parameters  $\theta = (A_1, A_2, \rho, \gamma)$  where  $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)^T$ . First, we estimate  $\sigma_\epsilon^2$  using the standard ordinary least squares method based on the residual sum of squares.

$$\sigma_\epsilon^2 = \frac{\sum_{i=1}^n \sum_{j=1}^2 \sum_{k=1}^{m_{ij}} (y_{ijk} - \hat{\alpha}_{ij} - t_{ijk} b_{ij,OLS})^2}{\sum_{i=1}^n \sum_{j=1}^2 (m_{ij} - 2)} \quad (8)$$

Next, the marginal variance ( $F$ ),  $A_1$ ,  $A_2$ , and  $\rho$  are initialized assuming that  $\gamma_1$  and  $\gamma_3$  are equal to 0. The F matrix is calculated as follows:

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = \frac{1}{n-1} \sum_{i=1}^n \begin{bmatrix} V_{i1} & 0 \\ 0 & V_{i2} \end{bmatrix} (\hat{\beta}_i - \bar{\beta})(\hat{\beta}_i - \bar{\beta})^T \quad \text{where} \quad \bar{\beta} = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n \hat{\beta}_{i,1} \\ \frac{1}{n} \sum_{i=1}^n \hat{\beta}_{i,2} \end{bmatrix} \quad (9)$$

By using a method of moments approach that equates the observed F matrix with its expected value, initial estimates of  $A_1$ ,  $A_2$ , and  $\rho$  can be calculated.

$$\hat{A}_1 = \frac{F_{11} - 1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{V_{i1}}}$$

$$\hat{A}_2 = \frac{F_{22} - 1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{V_{i2}}}$$

$$\hat{\rho} = \frac{F_{12} + F_{21}}{\frac{\sqrt{A_1 A_2}}{n} \left( \sum_{i=1}^n \frac{1}{V_{i1}} + \frac{1}{V_{i2}} \right)}$$

Once the initial estimates of  $A_1$ ,  $A_2$ , and  $\rho$  are obtained, the informative dropout needs to be taken into account since previously it was assumed that there was no informative right censoring (ie  $\{\gamma_1, \gamma_3 = 0\}$ ). To do this, the gamma vector needs to be updated. Weighted least squares is used to update  $\gamma$  using our estimates of  $A_1$ ,  $A_2$ , and  $\rho$ .

$$\hat{\gamma} = (Z^T D^{-1} Z)^{-1} Z^T D^{-1} \hat{\beta} \quad (10)$$

$$\begin{aligned} \text{where } \hat{\beta} &= \begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_n \end{bmatrix}, \\ \hat{Z} &= \begin{bmatrix} \hat{Z}_1 \\ \vdots \\ \hat{Z}_n \end{bmatrix}, \quad \text{where } Z_i = \begin{bmatrix} 1 & t_{i1}^* & 0 & 0 \\ 0 & 0 & 1 & t_{i2}^* \end{bmatrix}, \\ \text{and } D &= \begin{bmatrix} D_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & D_n \end{bmatrix} \quad \text{where } D_i = \text{var}(\hat{\beta}_i) = V_i + W \end{aligned}$$

Subsequently, the initial estimates of  $A_1$ ,  $A_2$ , and  $\rho$  need to be updated using the new gamma vector.

$$\hat{\mu}_i = \begin{bmatrix} \hat{\gamma}_0 + \hat{\gamma}_1 t_{i1}^* \\ \hat{\gamma}_2 + \hat{\gamma}_3 t_{i2}^* \end{bmatrix} \quad (11)$$

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \frac{1}{n-1} \sum_{i=1}^n \begin{bmatrix} V_{i1} & 0 \\ 0 & V_{i2} \end{bmatrix} (\hat{\beta}_i - \hat{\mu}_i)(\hat{\beta}_i - \hat{\mu}_i)^T \quad (12)$$

By using a method of moments that equates the observed G matrix with its expected value, we are able to update  $A_1$ ,  $A_2$ , and  $\rho$  denoting these  $A_{1,new}$ ,  $A_{2,new}$  and  $\rho_{new}$ .

$$\hat{A}_{1,new} = \frac{G_{11} - 1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{V_{i1}}}$$

$$A_{2,\hat{new}} = \frac{G_{22} - 1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{V_{i2}}}$$

$$\rho_{\hat{new}} = \frac{G_{12} + G_{21}}{\frac{\sqrt{A_1 A_2}}{n} \left( \sum_{i=1}^n \frac{1}{V_{i1}} + \frac{1}{V_{i2}} \right)}$$

The alternation of updating  $A_1$ ,  $A_2$ , and  $\rho$ , and  $\gamma$  are continued until the estimates of  $A_1$ ,  $A_2$ ,  $\rho$ , and  $\gamma$  have converged.

$$\begin{aligned} |\hat{A}_1 - A_{1,\hat{new}}| &< 0.0001 \\ |\hat{A}_2 - A_{2,\hat{new}}| &< 0.0001 \\ |\hat{\rho} - \rho_{\hat{new}}| &< 0.0001 \end{aligned} \tag{13}$$

Once the final estimates of  $A_1$ ,  $A_2$ , and  $\rho$ , and  $\gamma$  have been obtained, the variance of  $\gamma$  can be calculated and used to perform a Wald test for informative dropout (ie  $\{\gamma_1, \gamma_3 = 0\}$ ).

$$var(\hat{\gamma}) = (Z^T D^{-1} Z)^{-1} \tag{14}$$

Finally, the Empirical Bayes estimates of the true individual slopes can be calculated [Gelman et al., 1995].

$$\begin{aligned} \hat{\beta}_{eb} &= (W^{-1} + V^{-1})^{-1} (W^{-1} \hat{\mu} + V^{-1} \hat{\beta}) \\ &= (W^{-1} + V^{-1})^{-1} (W^{-1} [Z(Z^T D^{-1} Z)^{-1} Z^T D^{-1} + V^{-1}] \hat{\beta}) \\ &= H \hat{\beta} \end{aligned} \tag{15}$$

$$\text{where } \hat{\mu} = \begin{bmatrix} \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_n \end{bmatrix} \quad (16)$$

$$\text{and } H = (W^{-1} + V^{-1})^{-1}(W^{-1} [Z(Z^T D^{-1} Z)^{-1} Z^T D^{-1} + V^{-1}])$$

The variance-covariance matrix of the Empirical Bayes estimates can be calculated using the H matrix also.

$$\begin{aligned} \text{var}(\hat{\beta}_{eb}) &= H \text{var}(\hat{\beta}) H^T \\ &= H D H^T \end{aligned} \quad (17)$$

The Empirical Bayes estimates of the average true slopes of the treated and untreated populations is given by

$$\bar{\beta}_{eb} = \frac{1}{n} Q^T \hat{\beta}_{eb} \quad (18)$$

$$\text{where } Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ \vdots & \end{bmatrix}$$

Its variance is given by

$$\text{Var}(\bar{\beta}_{eb}) = \frac{1}{n^2} Q^T D Q \quad (19)$$

The point estimates and standard errors which were obtained through Empirical Bayes

can be compared using Wald tests. For different contrasts the Q matrix will change however equations 18 and 19 are used. Contrasts between the treated and untreated sides for each treatment group were assessed to determine the treatment effect. Also, contrasts testing the treated slope equal to 0 and the untreated slope equal to 0 were performed.

$$Z = \frac{T}{\sqrt{Var(T)}} \sim N(0,1) \quad (20)$$

$$\text{where } T = \frac{1}{n} Q^{*T} \hat{\beta}_{eb} \quad (21)$$

$$\text{and } Var(T) = \frac{1}{n^2} Q^{*T} D Q^*, \quad (22)$$

$$\text{where } Q^* = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \vdots \end{bmatrix} \quad (23)$$

## 3 Results

### 3.1 Checking Assumptions

To begin the analysis, the assumptions were checked for the model. The first assumption was that the response variable, BBB Score, is linear in time. This was checked using a scatter-plot of the BBB Score by Time for each animal where each side was shown separately (Figure 1). This plot indicates that BBB Score is linear in time, satisfying the assumption. Another plot was used to assess this assumption (Figure 2). This plot shows the average BBB score pooled by day. Each combination of side of the mouse and treatment group was allotted its own line, indicated in the legend. The satisfaction of the linearity assumption is better shown through this figure (Figure 2). Figure 1 can also be used to assess whether the right censoring was informative. It seems, although not clear, as though the animals who had a faster decline in their BBB score also died faster, which would indicate informative right censoring. In the plot it is clear that after day 42, which every mouse survived to, the BBB scores are more scattered and rapidly decline. The second assumption that needed to be checked was that the variances of the slope estimates were constant over time. This was examined through a boxplot of the residuals for the treated and untreated sides of each treatment group split up into four equal time periods (Figure 5). This plot suggest that there is relatively constant variance over time for each combination of side and treatment group.

### 3.2 Model Based Results

After checking the assumptions of the model, the slope estimates for each of the animals were calculated. First, they were calculated with ordinary least squares. Then they were calculated using an empirical bayes approach assuming no informative right censoring and finally using an empirical bayes approach adjusting for informative right censoring. The estimates for each treatment group (High, Low, and GFP) can be seen in Tables 1, 2, and 3

respectively.

Table 1: High Treatment Group Slope Estimates

Mouse	Side	OLS Slope	Std. Error	Empirical Bayes Estimate without IRC	Std. Error	Empirical Bayes Estimate with IRC	Std. Error
1	Treated	-0.1255	0.0441	-0.2078	0.1639	-0.1303	0.1608
	Untreated	-0.1827	0.0371	-0.2339	0.1440	-0.1837	0.1381
3	Treated	0.0276	0.0718	-0.2584	0.1734	-0.0121	0.1603
	Untreated	-0.0945	0.0605	-0.2680	0.1517	-0.1046	0.1376
5	Treated	-0.3747	0.0580	-0.2388	0.1682	-0.3719	0.1586
	Untreated	-0.4412	0.0489	-0.2547	0.1474	-0.4224	0.1362
6	Treated	-0.1835	0.0407	-0.1994	0.1630	-0.1851	0.1618
	Untreated	-0.2326	0.0343	-0.2282	0.1433	-0.2316	0.1389
12	Treated	-0.2639	0.0441	-0.2078	0.1639	-0.2628	0.1608
	Untreated	-0.3012	0.0371	-0.2339	0.1440	-0.2973	0.1381
14	Treated	-0.4729	0.0642	-0.2472	0.1704	-0.4444	0.1587
	Untreated	-0.3750	0.0541	-0.2604	0.1492	-0.3729	0.1363
16	Treated	-0.2577	0.0718	-0.2584	0.1734	-0.2471	0.1603
	Untreated	-0.1863	0.0605	-0.2680	0.1517	-0.2036	0.1376
18	Treated	-0.1446	0.0350	-0.1797	0.1617	-0.1436	0.1641
	Untreated	-0.1424	0.0295	-0.2149	0.1422	-0.1459	0.1408

In all three treatment groups the slope estimates and their standard errors have the same relationship. The OLS standard errors are the smallest when compared to their slope estimates. Next, the Empirical Bayes with informative right censoring's standard errors are larger than the OLS standard errors when compared with their slope estimates but smaller than the Empirical Bayes without informative right censoring's standard errors.

Next, the variance components of each treatment group and side of the animal were calculated (Table 4). The variance components for the high treatment group and low treatment groups were very similar,  $\sigma_{\epsilon,1}^2 = 11.5479\sigma_{\epsilon,2}^2 = 8.1949$  versus  $\sigma_{\epsilon,1}^2 = 11.1712\sigma_{\epsilon,2}^2 = 8.2969$ . The variance components for the GFP treatment group were reversed from the high and low treatment groups,  $\sigma_{\epsilon,1}^2 = 8.7158\sigma_{\epsilon,2}^2 = 11.8807$ .

Table 2: Low Treatment Group Slope Estimates

Mouse	Side	OLS Slope	Std. Error	Empirical Bayes Estimate without IRC	Std. Error	Empirical Bayes Estimate with IRC	Std. Error
2	Treated	-0.2306	0.0370	-0.1477	0.1403	-0.2236	0.1075
	Untreated	-0.2260	0.0338	-0.1517	0.1337	-0.2211	0.1072
4	Treated	-0.1442	0.0410	-0.2014	0.1414	-0.1478	0.1033
	Untreated	-0.1055	0.0374	-0.2129	0.1347	-0.1163	0.1034
7	Treated	-0.3633	0.0631	-0.3088	0.1493	-0.3495	0.1028
	Untreated	-0.3548	0.0576	-0.3354	0.1416	-0.3532	0.1030
8	Treated	-0.3693	0.0631	-0.3088	0.1493	-0.3618	0.1028
	Untreated	-0.4667	0.0576	-0.3354	0.1416	-0.4387	0.1030
9	Treated	-0.1832	0.0434	-0.2014	0.1421	-0.1840	0.1029
	Untreated	-0.1690	0.0396	-0.2129	0.1353	-0.1741	0.1033
10	Treated	-0.1681	0.0436	-0.2014	0.1421	-0.1762	0.1029
	Untreated	-0.2796	0.0396	-0.2129	0.1353	-0.2696	0.1033
13	Treated	-0.3105	0.0570	-0.2857	0.1469	-0.3068	0.1009
	Untreated	-0.3448	0.0520	-0.3191	0.1395	-0.3387	0.1015
19	Treated	-0.1742	0.0570	-0.2857	0.1469	-0.1895	0.1009
	Untreated	-0.1369	0.0520	-0.3091	0.1395	-0.1663	0.1015

The initial variances of the treated and untreated sides were denoted under the without informative right censoring column as  $\hat{A}_1$  and  $\hat{A}_2$  respectively. Both the Low and GFP treatment groups marginal variances decreased when informative right censoring was adjusted for. However, in the High treatment group the variances increased when informative right censoring was adjusted for. All of the variance estimates were of approximately the same magnitude.

The correlation between the treated and untreated sides of the mice was given by  $\hat{\rho}$ . Under the without IRC column in Table 4 the initial correlation can be found. This was calculated before adjusting for informative right censoring. After adjusting for informative right censoring, the correlations for each treatment group decrease. This was due to adjusting for informative right censoring. The GFP treatment group's correlation decreased the most

Table 3: GFP Treatment Group Slope Estimates

Mouse	Side	OLS Slope	Std. Error	Empirical Bayes Estimate without IRC	Std. Error	Empirical Bayes Estimate with IRC	Std. Error
11	Treated	-0.1993	0.0283	-0.1519	0.1467	-0.1954	0.0952
	Untreated	-0.1882	0.0331	-0.1694	0.1521	-0.1875	0.0984
15	Treated	-0.1784	0.0354	-0.2468	0.1482	-0.1867	0.0908
	Untreated	-0.2116	0.0413	-0.2422	0.1541	-0.2144	0.0926
17	Treated	-0.2852	0.0504	-0.3676	0.1525	-0.3053	0.0925
	Untreated	-0.3560	0.0584	-0.3348	0.1597	-0.3450	0.0952
20	Treated	-0.5016	0.0557	-0.3934	0.1544	-0.4716	0.0956
	Untreated	-0.3411	0.0651	-0.3547	0.1621	-0.3504	0.0992

going from 0.8060 to 0.2258. The High treatment group had the highest final correlation, 0.6120.

Table 4: Variance Component Estimates

Treatment Group	OLS Estimates		Without IRC			With IRC		
	$\hat{\sigma}_{\epsilon,1}^2$	$\hat{\sigma}_{\epsilon,2}^2$	$\hat{A}_1$	$\hat{A}_2$	$\hat{\rho}$	$\hat{A}_1$	$\hat{A}_2$	$\hat{\rho}$
High	11.5479	8.1949	0.0249	0.0193	0.6244	0.0270	0.0199	0.6120
Low	11.1712	9.2969	0.0183	0.0167	0.3749	0.0115	0.0115	0.3308
GFP	8.7158	11.8807	0.0207	0.0220	0.8060	0.0086	0.0091	0.2258

After the variance components were estimated, the slopes for each treatment group and side of the animal combination were tested for informative right censoring (Table 5). The test of informative right censoring for the treated side of the High treatment group was not statistically significant ( $p = 0.6503$ ). The test for the untreated side of the High treatment group was also not statistically significant ( $p = 0.7206$ ). Both sides of the Low treatment group were not statistically significant either ( $p = 0.1740$ ,  $p = 0.1135$ ). Similarly, both sides of the GFP treatment group were not statistically significant ( $p = 0.0576$ ,  $p = 0.1779$ ). In order to further assess the informative right censoring, plots of the estimated slopes (OLS and Empirical Bayes) versus each animal's dropout time were examined (Figures 3 and 4,

respectively). These plots indicated that the animal's slopes increased with an increase in their dropout time, indicating that there was informative right censoring. There was a discrepancy between the test results and the plots, which suggested a lack of power.

Table 5: Testing Informative Right Censoring

Treatment Group	Side	Slope Estimates		Test of Informative Dropout	
		$\gamma$	SE( $\gamma$ )	Test Statistic	P-value
High	Treated( $\gamma_1$ )	0.0028	0.0062	0.4533	0.6503
	Untreated( $\gamma_3$ )	0.0019	0.0053	0.3577	0.7206
Low	Treated( $\gamma_1$ )	0.0076	0.0056	1.3593	0.1740
	Untreated( $\gamma_3$ )	0.0087	0.0055	1.5825	0.1135
GFP	Treated( $\gamma_1$ )	0.0086	0.0045	1.8985	0.0576
	Untreated( $\gamma_3$ )	0.0064	0.0047	1.3473	0.1779

The average slope estimates, both OLS and Empirical Bayes, were compared using Table 6. All of the slopes estimates using the OLS method were larger than those obtained through the Empirical Bayes method. The highest degradation rate was found in the treated side of the GFP treatment group using the OLS estimates. This was followed by the untreated side of the GFP treatment group. Subsequently, the untreated side of the Low treatment group had the next fastest degradation rate and then the untreated side of the High treatment group followed by the treated side of the low treatment group and finally the treated side of the high treatment group. This pattern was not the same for the average Empirical Bayes estimates. Of the slopes obtained through the Empirical Bayes method, the treated side of the High treatment group had the fastest decaying population. This was then followed by the untreated side of the GFP treatment group and the treated side of the Low treatment group. Next was the untreated side of the Low treatment group, the untreated side of the High treatment group, and finally was the treated side of the GFP treatment group.

Using the Empirical Bayes estimates of the slopes, the equality of the treated and untreated sides was tested for each treatment group (Table 7). The treated side of the High and Low treatment groups had a slower degradation rate on average than the untreated side on average in the same treatment group. Whereas, the treated side of the GFP treatment

Table 6: Comparison of Average Un-weighted Estimate and Average Empirical Bayes Estimate

Treatment Group	Total Mice	Side	Slope Estimates and Standard Errors			
			Average Unweighted Estimate	Std. Error	Average Empirical Bayes Estimate	Std. Error
High	8	Treated	-0.2244	0.1551	-0.2247	0.0038
		Untreated	-0.2445	0.1190	-0.2453	0.0028
Low	8	Treated	-0.2429	0.0917	-0.2424	0.0018
		Untreated	-0.2604	0.1241	-0.2598	0.0017
GFP	4	Treated	-0.2911	0.1477	-0.2898	0.0026
		Untreated	-0.2742	0.0866	-0.2743	0.0029

group had a faster degradation rate than the untreated side on average. However when tested, none of the treatment groups treated and untreated sides were significantly different ( $p_j > 0.70$ ).

Table 7: Testing Difference in Average Treated vs. Average Untreated Sides of the Mice in Each Treatment Group

Treatment Group	Difference of the Means	Std. Error	Test Statistic	P-Value
High	0.0206	0.0030	0.3778	0.7056
Low	0.0174	0.0025	0.3451	0.7300
GFP	-0.0154	0.0046	-0.2286	0.8192

## 4 Discussion

### 4.1 Discussion of Results

This study aimed to detect a difference between the treated and untreated sides of mice in three different treatment groups. A new methodology had to be created due to the potential for informative right censoring. In this study, there were no statistically significant results. No differences between the Empirical Bayes estimates for the treated and untreated sides of each treatment group were found, even though the methodology took into account the correlation between the two sides and the possibility of informative right censoring. However, the slopes estimates became much more precise when using the Empirical Bayes approach rather than the OLS approach, which was expected. Informative right censoring was not found to be present in this study. One treatment group, GFP, was close to finding informative dropout. Figures 3 and 4 indicate that there is informative dropout since the dropout time is linearly related to the slope estimates. When informative right censoring was not taken into account, the slope estimates were misleadingly high because the animals with faster degradation rates dropped out. Even though the test of informative dropout was not significant, it is still important to use the new method of analysis. Using a naïve analysis would ignore the fact that the  $\gamma$  estimates for each side ( $\hat{\gamma}_1$  and  $\hat{\gamma}_3$ ) were positive, indicating that the slope estimates increase with an increase in the dropout time. It is believed that with more power the informative dropout could be detected. The sample size was very small for the three treatment groups. Perhaps if more animals could be tested, the ability to detect informative right censoring would increase.

### 4.2 Further Work

Following from the suggestion of a lack of power in this study many opportunities for further work arise. The first is to develop a study design for researchers which would have adequate power to detect differences between the sides of the animals as well as to detect

informative right censoring. This would need to account for the number of treatment groups as well as the cost per animal. It would not be feasible to suggest a large sample size when the SOD1 gene mice are expensive. This could be done through simulation studies both creating new data as well as sampling from this data. Another opportunity for further work is to develop diagnostics for this methodology to test the assumptions. The diagnostics done here could be further developed into actual tests.

## References

Als (amyotrophic lateral sclerosis) fact sheet, 2010.

Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. 1st edition, 1995.

Sujatha M. Gopalakrishnan, Nicole Teusch, Christiane Imhof, Margot H. M. Bakker, Mark Schurdak, David J. Burns, and Usha Warrior. Role of rho kinase pathway in chondroitin sulfate proteoglycan-mediated inhibition of neurite outgrowth in pc12 cells. *Journal of Neuroscience Research*, 86(10):2214–2226, 2008.

Alex Hoberman, Eleanor Donnelly, Thais Federici, Griffin Baum, Cary Leung, Nicholas Boulis, and Robert Gross. Direct spinal chord gene therapy with c3 transferase in an als model. 2011.

Matthew C. Kiernan, Steve Vucic, Benjamin C. Cheah, Martin R. Turner, Andrew Eisen, Orla Hardiman, James R. Burrell, and Margaret C. Zoing. Amyotrophic lateral sclerosis. *The Lancet*, 377(9769):942–955, 2011.

Arifumi Matsumoto, Yohei Okada, Masanori Nakamichi, Masaya Nakamura, Yoshiaki Toyama, Gen Sobue, Makiko Nagai, Masashi Aoki, Yasuto Itoyama, and Hideyuki Okano. Disease progression of human sod1 (g93a) transgenic als model rats. *Journal of Neuroscience Research*, 83(1):119–133, 2006.

Hyejin Pak. *The Effect of bFGF Mutants on Tumor Growth Rate: An Investigation using Tumor Growth Curve Analysis with Correlated Tumors and Informative Right Censoring*. Master's of science in public health thesis, 1999.

Lewis P. Rowland and Neil A. Shneider. Amyotrophic lateral sclerosis. *New England Journal of Medicine*, 344(22):1688–1700, 2001.

## 5 Appendices

### A Figures

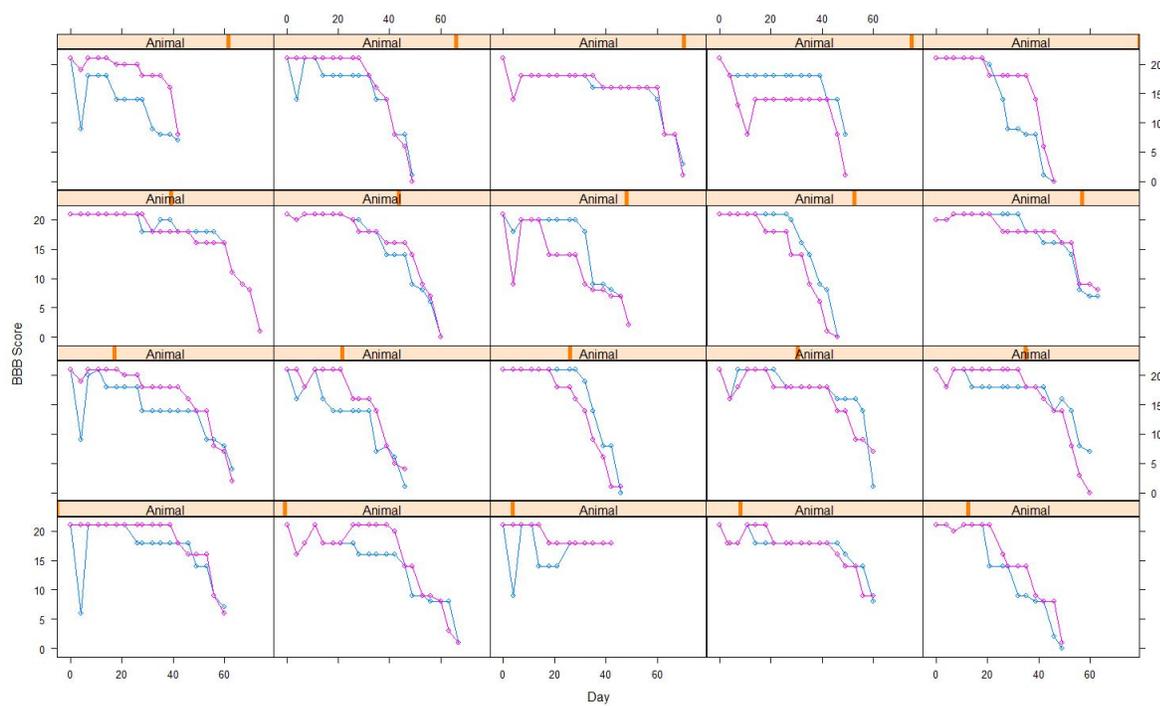


Figure 1: BBB Score by Day After Injection

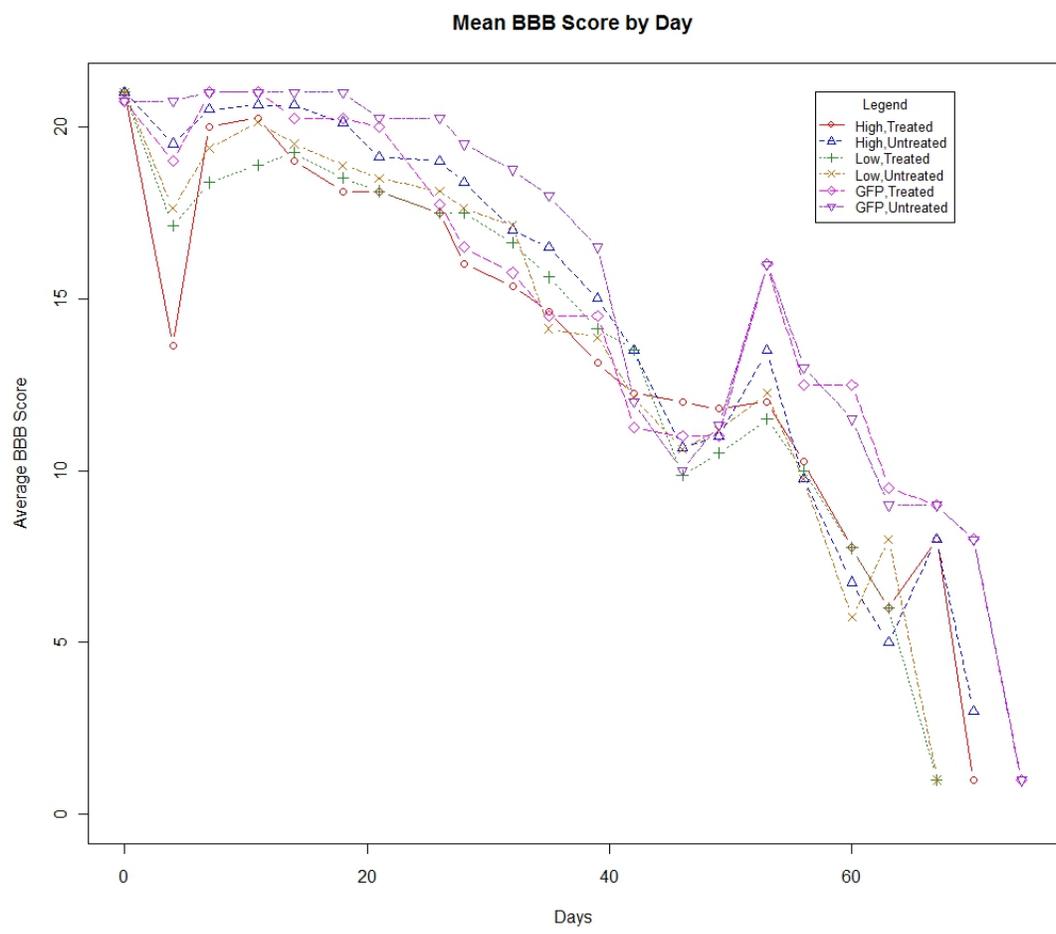


Figure 2: Average BBB Score by Day By Side of Animal for All Treatment Groups

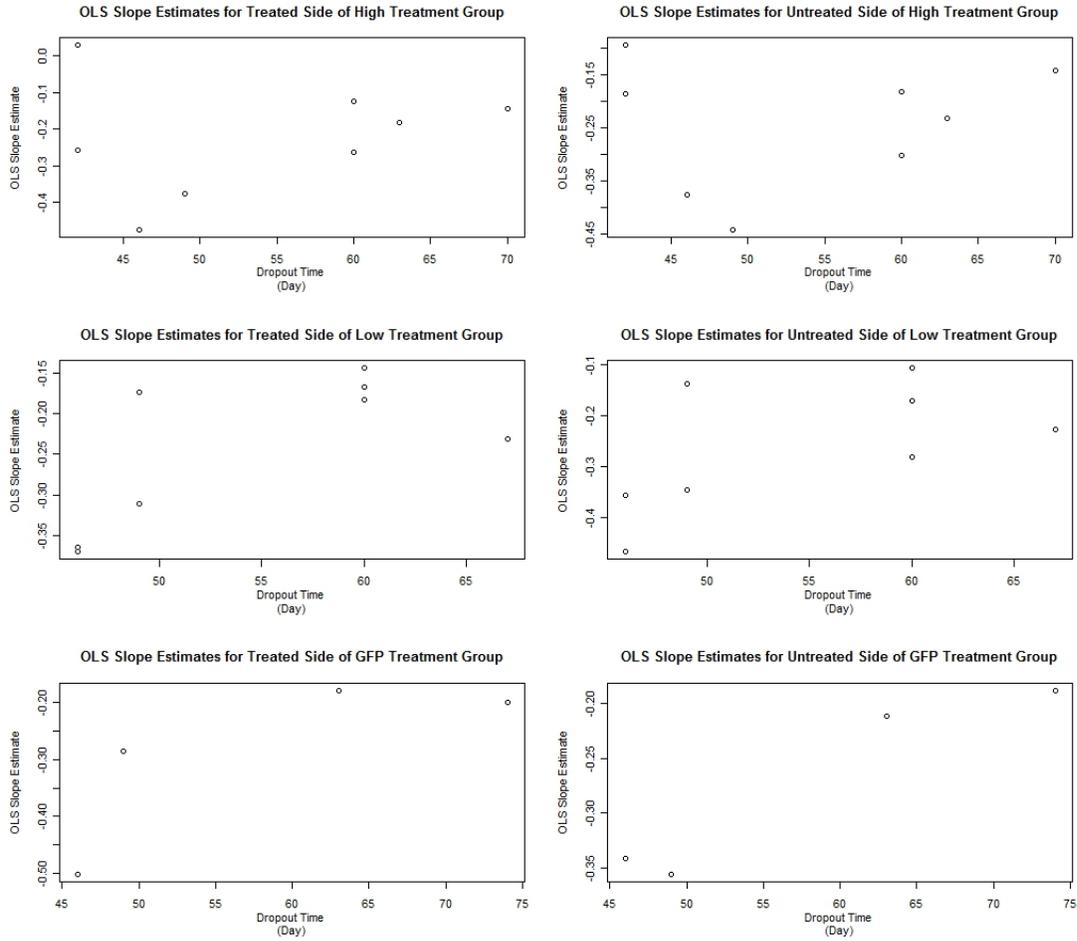


Figure 3: OLS Slope By Dropout Time By Side of the Animal for All Treatment Groups

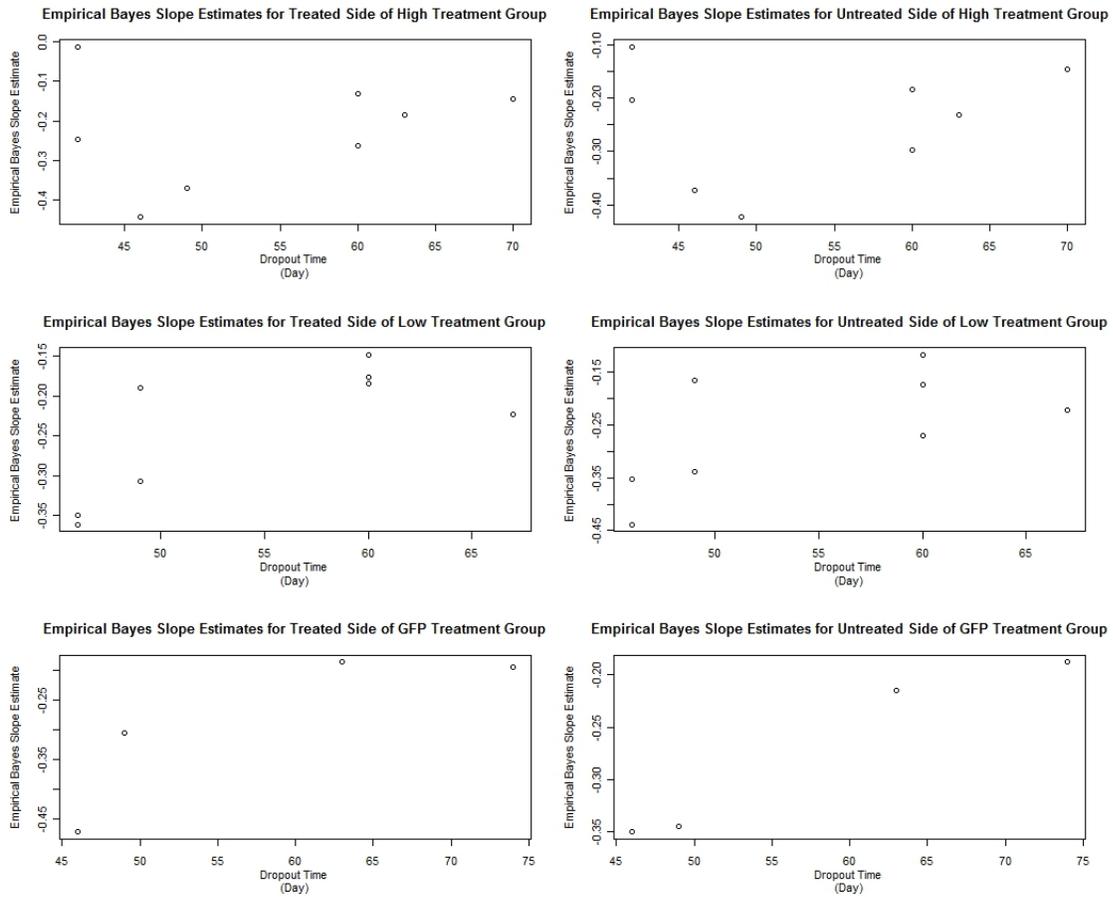


Figure 4: Empirical Bayes Slope Estimate By Dropout Time By Side of the Animal for All Treatment Groups

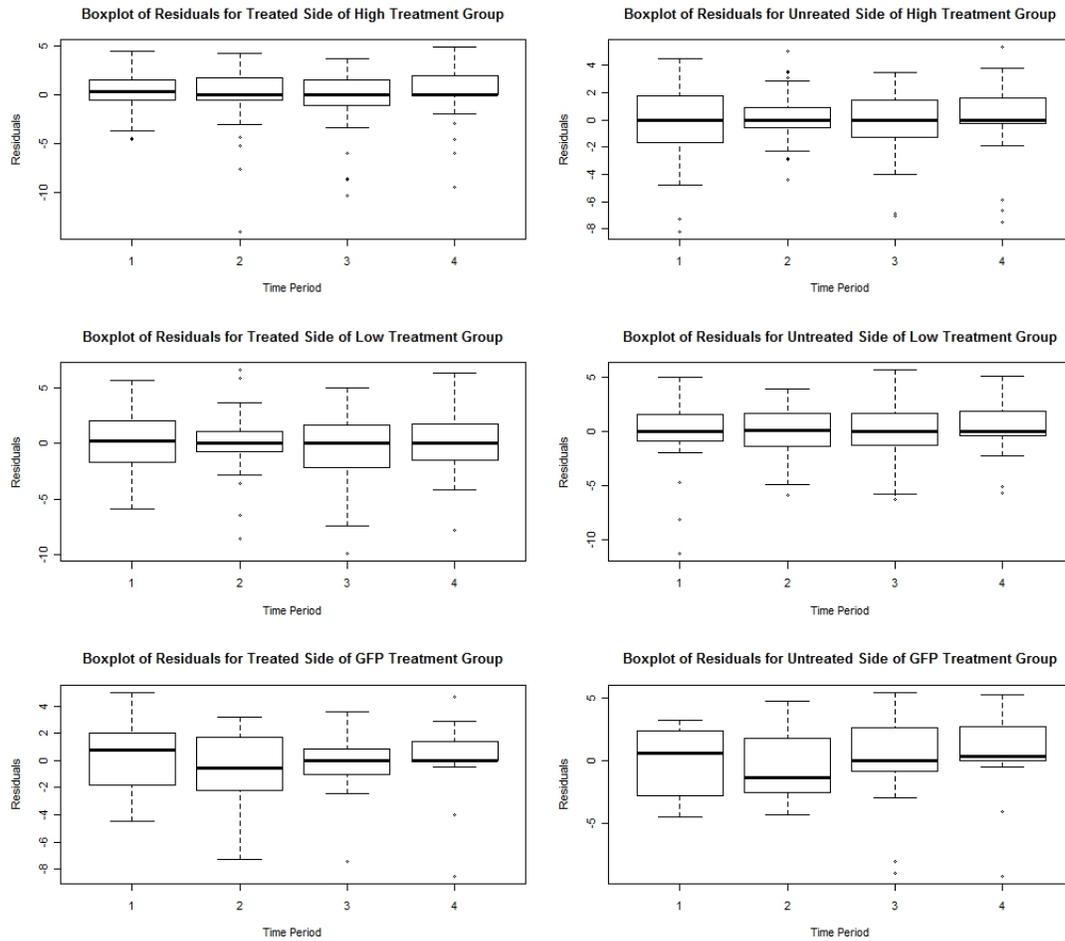


Figure 5: Boxplots of Residuals by Timeperiod by Side of the Animal for All Treatment Groups

## B R Code

Copyrighted by Katherine Catevenis, 2012

```

behdata=read.csv(file = "E : \\ Data \\ Behavioral_Data_Reorganized_CSV.csv")
siden=ifelse(behdata$Side==behdata$Treatment_Side,"treated","untreated")
behdata=data.frame(behdata,siden)

nanimal=20

nobstwo=matrix(0,nanimal,1)
for(i in 1:nanimal) {nobstwo[[i]] < -sum(behdata$Animal == i)}
nobs=nobstwo/2
maxobs=max(nobs)

# creating two matrices with all BBB data one for each side of the animals ie Y
bigytx=matrix(NA,nanimal,maxobs)
bigyuntx=matrix(NA,nanimal,maxobs)
for (i in 1:nanimal) {
bigytx[i,1:nobs[i,1]]=behdata$BBB[behdata$Animal==i & behdata$siden=="treated"]
bigyuntx[i,1:nobs[i,1]]=behdata$BBB[behdata$Animal==i & behdata$siden=="untreated"]
}

# creating a matrix with all day data, ie t
bigt=matrix(NA,nanimal,maxobs)
for (i in 1:nanimal) {
bigt[i,1:nobs[i,1]]=-behdata$Day[behdata$Animal==i & behdata$siden=="treated"]
}

bigytxhigh=rbind(bigytx[1,],bigytx[3,], bigytx[5,], bigytx[6,], bigytx[12,], bigytx[14,], bigytx[16,],
bigytx[18,])
bigytxlow=rbind(bigytx[2,],bigytx[4,], bigytx[7,], bigytx[8,], bigytx[9,], bigytx[10,], bigytx[13,],
bigytx[19,])
bigytxgfp=rbind(bigytx[11,],bigytx[15,],bigytx[17,],bigytx[20,])

```

```

bigyuntxhigh=rbind(bigyuntx[1,],bigyuntx[3,],bigyuntx[5,],bigyuntx[6,], bigyuntx[12,], bigyuntx[14,],
bigyuntx[16,], bigyuntx[18,])
bigyuntxlow=rbind(bigyuntx[2,],bigyuntx[4,],bigyuntx[7,],bigyuntx[8,], bigyuntx[9,], bigyuntx[10,],
bigyuntx[13,], bigyuntx[19,])
bigyuntxgfp=rbind(bigyuntx[11,],bigyuntx[15,],bigyuntx[17,],bigyuntx[20,])
bigthigh=rbind(bigt[1,],bigt[3,],bigt[5,],bigt[6,], bigt[12,], bigt[14,], bigt[16,], bigt[18,])
bigtlow=rbind(bigt[2,],bigt[4,],bigt[7,],bigt[8,], bigt[9,], bigt[10,], bigt[13,], bigt[19,])
bigtgfp=rbind(bigt[11,],bigt[15,],bigt[17,],bigt[20,])
nobshigh=rbind(nobs[1,1],nobs[3,1],nobs[5,1],nobs[6,1],nobs[12,1], nobs[14,1], nobs[16,1], nobs[18,1])
nobslow=rbind(nobs[2,1],nobs[4,1],nobs[7,1],nobs[8,1],nobs[9,1], nobs[10,1], nobs[13,1], nobs[19,1])
nobsgfp=rbind(nobs[11,1],nobs[15,1],nobs[17,1],nobs[20,1])
iest=function(bigytx,bigyuntx,bigt,nobs,nanimal) {
maxobs=max(nobs)
print(maxobs)
# creating a matrix with all model beta estimates for each side along with a matrix of
residuals to plot
# for diagnostics
slopes.individualtx=matrix(0,nanimal,2)
slopes.individualuntx=matrix(0,nanimal,2)
residualstx=matrix(0,nanimal,maxobs)
residualsuntx=matrix(0,nanimal,maxobs)
for ( i in 1:nanimal) {
mtx=lm(bigytx[i,]~bigt[i,])
muntx=lm(bigyuntx[i,]~bigt[i,])
slopes.individualtx[i,1]=mtx$coefficients[1]
slopes.individualtx[i,2]=mtx$coefficients[2]
slopes.individualuntx[i,1]=muntx$coefficients[1]

```

```

slopes.individualuntx[i,2]=muntx$coefficients[2]
residualstx[i,1:length(mtx$residuals)]=mtx$residuals
residualsuntx[i,1:length(muntx$residuals)]=muntx$residuals
}
# creating a matrix of dropout times for each animal
dropout=matrix(NA,nanimal,1)
for (i in 1:nanimal) {
dropout[i,1]=bigt[i,nobs[i,1]]
}
# plot slopes by dropout time
#plot(slopes.individualtx[,2]~dropout,xlab="Dropout Time",ylab="Treated Side Slope Estimate" )
#plot(slopes.individualuntx[,2]~dropout,xlab="Dropout Time",ylab="Untreated Side Slope Estimate" )
# sse
ssetx=diag(residualstx%*%t(residualstx))
sseuntx=diag(residualsuntx%*%t(residualsuntx))
# sst
t4=matrix(0,nanimal,maxobs)
for (i in 1:nanimal) {
t4[i,1:nobs[i,1]]=bigt[i,1:nobs[i,1]]-mean(bigt[i,1:nobs[i,1]])
}
sst=diag(t4%*%t(t4))
# sigma
sigmatx=sum(ssetx)/(sum(nobs)-2*nanimal)
sigmauntx=sum(sseuntx)/(sum(nobs)-2*nanimal)
print (sigmatx)

```

```

print (sigmauntx)
# variance of ols estimate of slope given animals true slope
v1tx=sigmatx/sst
v1untx=sigmauntx/sst
meanslopetx=mean(slopes.individualtx[,2])
meanslopeuntx=mean(slopes.individualuntx[,2])
seslopetx=sqrt(var(slopes.individualtx[,2]))
seslopeuntx=sqrt(var(slopes.individualuntx[,2]))
print (seslopetx)
print (seslopeuntx)
meanslope=matrix(rbind(meanslopetx,meanslopeuntx),nanimal*2,1,byrow=TRUE)
print(meanslope)
# slopes tx untx
slopes=matrix(NA,nanimal*2,1)
k=1
for (i in 1:nanimal) {
slopes[k,1]=slopes.individualtx[i,2]
k=k+1
slopes[k,1]=slopes.individualuntx[i,2]
k=k+1
}
bmm=slopes-meanslope
diagmat=function(v1mat) {
iseq=seq(1,((2*nanimal)-1),2)
jseq=seq(2,(2*nanimal),2)
diagv1=matrix(0,2*nanimal,2*nanimal)
for (i in iseq){

```

```

diagv1[i,(i:(i+1))]=v1mat[i,1:2]
}
for (j in jseq) {
diagv1[j,((j-1):j)]=v1mat[j,1:2]
}
return (diagv1)
}
v1=matrix(0,nanimal*2,2)
k=1
for (i in 1:nanimal) {
v1[k,1]=v1tx[i]
v1[k+1,2]=v1untx[i]
k=k+2
}
diagv1=diagmat(v1)
# calculate f
# v1 inv * bmm * t(bmm)
# 40 x 40 * 40 x 1 * 1 x 40
v1invdiag=1/diagv1
v1invdiag[v1invdiag=="Inf"]=0
bigf=v1invdiag%%*%%bmm%%*%%t(bmm)
sumdiag=function(mat)
{ sum22=matrix(0,2,2)
for (i in 1:(nrow(mat)-1))
{ mat22=matrix(c(mat[i,i],mat[i,i+1],mat[i+1,i],mat[i+1,i+1]), 2, 2, byrow=TRUE)
sum22=sum22+mat22
i=i+2
}
}

```

```

}
return(sum22)
}
v1txinv=1/v1tx
v1untxinv=1/v1untx
f=(1/(nanimal-1))*sumdiag(bigf)
a1i=(f[1,1]-1)/((1/nanimal)*sum(v1txinv))
a2i=(f[2,2]-1)/((1/nanimal)*sum(v1untxinv))
pi=(f[1,2]+f[2,1])/((sqrt(a1i*a2i)/nanimal)* ( sum(v1txinv) + sum(v1untxinv)))
z=matrix(0,nanimal*2,4)
j=1
for (i in 1:nanimal) {
z[j,1]=1
z[j,2]=dropout[i,1]
z[j,3]=0
z[j,4]=0
z[j+1,1]=0
z[j+1,2]=0 z[j+1,3]=1
z[j+1,4]=dropout[i,1]
j=j+2
} iestout=matrix(NA,(nanimal*2),(13+(2*maxobs)))
iestout[1,1]=a1i
iestout[1,2]=a2i
iestout[1,3]=pi
iestout[1:(nanimal*2),4:5]=v1
iestout[1:(nanimal*2),6]=slopes
iestout[1:(nanimal*2),7:10]=z

```

```

iestout[1:nanimal,11]=dropout
iestout[1:nanimal,12]=slopes.individualtx[,2]
iestout[1:nanimal,13]=slopes.individualuntx[,2]
iestout[1:nanimal,14:(13+maxobs)]=residualstx
iestout[1:nanimal,(14+maxobs):(13+(2*maxobs))]=residualsuntx
return (iestout)
}
#residualstx[i,1:length(mtx$residuals)]=mtx$residuals
#residualsuntx[i,1:length(muntx$residuals)]=muntx$residuals
#using for each treatment group
estihigh=iest(bigytxhigh,bigyuntxhigh,bighigh,nobshigh,8)
estilow=iest(bigytxlow,bigyuntxlow,bigtlow,nobslow,8)
estigfp=iest(bigytxgfp,bigyuntxgfp,bigtgfp,nobsgfp,4)
# function for updating estimates
update=function(a1i,a2i,pi,v1,b,z)
{ cycle=1
a1dif=1
a2dif=1
pdif=1
flag=0
while (flag==0 & cycle < 100 )
{
n=nrow(v1)/2
w=matrix(c(a1i,(pi*sqrt(a1i*a2i)),(pi*sqrt(a1i*a2i)),a2i), 2*n, 2 , byrow=TRUE)
varb=v1+w
print(varb)
diagmat=function(v1mat) {

```

```

iseq=seq(1,((2*n)-1),2)
jseq=seq(2,(2*n),2)
diagv1=matrix(0,2*n,2*n)
for (i in iseq){
diagv1[i,(i:(i+1))]=v1mat[i,1:2]
}
for (j in jseq) {
diagv1[j,((j-1):j)]=v1mat[j,1:2]
}
return (diagv1)
}
d=diagmat(varb)
vdiag=diagmat(v1)
gammah=((solve((t(z))%*%(solve(d))%*%z)) %* % (t(z)) % * %( solve(d)) %*% b)
muh=z%*%gammah
print(muh)
v1inv=1/v1
v1inv[v1inv=="Inf"]=0
v1invdiag=diagmat(v1inv)
gpart=v1invdiag%*%(b-muh)%*%(t(b-muh))
sumdiag=function(mat)
{ sum22=matrix(0,2,2)
for (i in 1:(nrow(mat)-1))
{ mat22=matrix(c(mat[i,i],mat[i,i+1],mat[i+1,i],mat[i+1,i+1]), 2, 2, byrow=TRUE)
sum22=sum22+mat22
i=i+2
}

```

```

return(sum22)
}
sumg=sumdiag(gpart)
g=(1/(n-2))*sumg
#make sure this is 2 x 2
# update a1 a2 and p
a1new=(g[1,1]-1)/((1/n)*sum(v1inv[,1]))
a2new=(g[2,2]-1)/((1/n)*sum(v1inv[,2]))
pnew=(g[1,2]+g[2,1])/((sqrt(a1new*a2new)/n)* ( sum(v1inv[,1])+ sum(v1inv[,2])))
wnew=matrix(c(a1new,(pnew*sqrt(a1new*a2new))),(pnew*sqrt(a1new*a2new)),a2new), 2*n,
2, byrow=TRUE)
wnewdiag=diagmat(wnew)
varbnew=v1+wnew
dnew=diagmat(varbnew)
gammahnew=((solve((t(z))%*%(solve(dnew))%*%z)) %* % (t(z)) %* % (solve(dnew)) %*%b)
row.names(gammahnew)=c("gamma_0", "gamma_1", "gamma_2", "gamma_3")
colnames(gammahnew)=c("Estimates")
vargammanew=solve((t(z))%*%(solve(dnew))%*%z)
a1dif=abs(a1new-a1i)
a2dif=abs(a2new-a2i)
pdif=abs(pnew-pi)
flag = ifelse(max(c(a2dif,a1dif,pdif))>0.0001, 1, 0)
cycle=cycle+1
if (cycle==100) { print("Did not Converge")}
a1i=a1new
a2i=a2new
pi=pnew

```

```

}
# bayes estimates
h=solve(solve(wnewdiag)+solve(vdiag))%*%((solve(wnewdiag)%*% z %* % solve(t(z) % *
% solve(dnew) % * % z) % * % t(z)% * % solve(dnew))+ solve(vdiag))
mustar=h%*%b
varmustar=h%*%dnew%*%t(h)
tmustar=t(mustar)
# bayes estimates using only gamma_0 and gamma_2 as estimates for mu
gam02=matrix(rbind(gammahnew[1],gammahnew[3]),2*n,1)
mustargam02=solve((solve(wnewdiag)+solve(vdiag))) %*% (( solve(wnewdiag)%*%gam02)+
(solve(vdiag)%*%b))
varmustargam02=solve(solve(wnewdiag)+solve(vdiag))
tmustargam02=t(mustargam02)
# testing bayes estimates
testbe=function(q,mu,n,d) {
t=(1/n)*(t(q)%*%mu)
var_t=(1/(n*n))*(t(q)%*%d%*%q)
z=t/sqrt(var_t)
p=2*(1-pnorm(abs(z)))
testing=cbind(t,var_t,z,p)
return(testing)
}
qtxuntx=matrix(rbind(1,-1),2*n,1)
qtx0=matrix(rbind(1,0),2*n,1)
quntx=matrix(rbind(0,1),2*n,1)
# testing treated vs untreated sides
txuntx=testbe(qtxuntx,mustar,n,dnew)

```

```

tx0=testbe(qtx0,mustar,n,dnew)
untx0=testbe(quntx,mustar,n,dnew)
# outputting estimates for everything
estimates=matrix(NA,((4*n)+13),2*n)
estimates[1,1]=a1new
estimates[2,1]=a2new
estimates[3,1]=pnew
estimates[4,1:4]=gammahnew
estimates[5:8,1:4]=vargammanew
estimates[9,1:(2*n)]=tmustar
estimates[10:((2*n)+9),1:(2*n)]=varmustar
estimates[((2*n)+10),1:(2*n)]=tmustargam02
estimates[((2*n)+11):((4*n)+10),1:(2*n)]=varmustargam02
estimates[((4*n)+11),1:4]=txuntx
estimates[((4*n)+12),1:4]=tx0
estimates[((4*n)+13),1:4]=untx0
return (estimates)
}
esthigh=update(estihigh[1,1],estihigh[1,2],estihigh[1,3],estihigh[1:16,4:5], estihigh[1:16,6], estihigh[1:16,7:10])
estlow=update(estilow[1,1],estilow[1,2],estilow[1,3],estilow[1:16,4:5], estilow[1:16,6], estilow[1:16,7:10])
estgfp=update(estigfp[1,1],estigfp[1,2],estigfp[1,3],estigfp[1:8,4:5], estigfp[1:8,6], estigfp[1:8,7:10])
# testing for informative right censoring (ie gamma 1 and gamma 3 = 0)
testirc=function(gamma,vargamma) {
segamma=sqrt(vargamma)
tgamma=gamma/segamma
pgamma=2*(1-pnorm(abs(tgamma)))

```

```
test=rbind(tgamma,pgamma)
return (test)
}
# high treatment group
vcgammahigh=esthigh[5:8,1:4]
gammahigh=esthigh[4,1:4]
gamma1high=gammahigh[2]
gamma3high=gammahigh[4]
vargamma1high=vcgammahigh[2,2]
vargamma3high=vcgammahigh[4,4]
pgam1high=testirc(gamma1high,vargamma1high)
pgam3high=testirc(gamma3high,vargamma3high)
pgam1high
pgam3high
#low treatment group
vcgammalow=estlow[5:8,1:4]
gammalow=estlow[4,1:4]
gamma1low=gammalow[2]
gamma3low=gammalow[4]
vargamma1low=vcgammalow[2,2]
vargamma3low=vcgammalow[4,4]
pgam1low=testirc(gamma1low,vargamma1low)
pgam3low=testirc(gamma3low,vargamma3low)
pgam1low
pgam3low
# gfp treatment group
vcgammagfp=estgfp[5:8,1:4]
```

```

gammagfp=estgfp[4,1:4]
gamma1gfp=gammagfp[2]
gamma3gfp=gammagfp[4]
vargamma1gfp=vcgammagfp[2,2]
vargamma3gfp=vcgammagfp[4,4]
pgam1gfp=testirc(gamma1gfp,vargamma1gfp)
pgam3gfp=testirc(gamma3gfp,vargamma3gfp)
pgam1gfp
pgam3gfp
#iestout[1:nanimal,11]=dropout
#iestout[1:nanimal,12]=slopes.individualtx[,2]
#iestout[1:nanimal,13]=slopes.individualuntx[,2]
# creating some plots
par(mfrow=c(3,2))
# treated OLS by dropout time for high tx
plot(estihigh[1:8,12]~estihigh[1:8,11], main="OLS Slope Estimates for Treated Side of High
Treatment Group", xlab="Dropout Time (Day)", ylab="OLS Slope Estimate")
# untreated OLS by dropout time for high tx
plot(estihigh[1:8,13]~estihigh[1:8,11], main="OLS Slope Estimates for Untreated Side of
High Treatment Group", xlab="Dropout Time (Day)", ylab="OLS Slope Estimate")
# treated OLS by dropout time for low tx
plot(estilow[1:8,12]~estilow[1:8,11], main="OLS Slope Estimates for Treated Side of Low
Treatment Group", xlab="Dropout Time (Day)", ylab="OLS Slope Estimate")
# untreated OLS by dropout time for low tx
plot(estilow[1:8,13]~estilow[1:8,11], main="OLS Slope Estimates for Untreated Side of Low
Treatment Group", xlab="Dropout Time (Day)", ylab="OLS Slope Estimate")
# treated OLS by dropout time for gfp tx plot(estigfp[1:4,12]~estigfp[1:4,11], main="OLS

```

```

Slope Estimates for Treated Side of GFP Treatment Group", xlab="Dropout Time (Day)",
ylab="OLS Slope Estimate")
# untreated OLS by dropout time for gfp tx
plot(estigfp[1:4,13]~estigfp[1:4,11], main="OLS Slope Estimates for Untreated Side of GFP
Treatment Group", xlab="Dropout Time (Day)", ylab="OLS Slope Estimate")
# plots for final bayes slope estimates
# estimates[9,1:(2*n)]=tmustar
#need to manipulate the slopes
bshightx=esthigh[9,c(1,3,5,7,9,11,13,15)]
bshighuntx=esthigh[9,c(2,4,6,8,10,12,14,16)]
bslowtx=estlow[9,c(1,3,5,7,9,11,13,15)]
bslowuntx=estlow[9,c(2,4,6,8,10,12,14,16)]
bsgfptx=estgfp[9,c(1,3,5,7)]
bsgfpuntx=estgfp[9,c(2,4,6,8)]
par(mfrow=c(3,2))
# treated Bayes by dropout time for high tx
plot(bshightx~estihigh[1:8,11],main="Empirical Bayes Slope Estimates for Treated Side of
High Treatment Group", xlab="Dropout Time (Day)", ylab="Empirical Bayes Slope Esti-
mate")
# untreated Bayes by dropout time for high tx
plot(bshighuntx~estihigh[1:8,11],main="Empirical Bayes Slope Estimates for Untreated Side
of High Treatment Group", xlab="Dropout Time (Day)", ylab="Empirical Bayes Slope Es-
timate")
# treated Bayes by dropout time for low tx
plot(bslowtx~estilow[1:8,11],main="Empirical Bayes Slope Estimates for Treated Side of
Low Treatment Group", xlab="Dropout Time (Day)", ylab="Empirical Bayes Slope Esti-
mate")

```

```

# untreated Bayes by dropout time for low tx
plot(bslowuntx~estilow[1:8,11],main="Empirical Bayes Slope Estimates for Untreated Side
of Low Treatment Group", xlab="Dropout Time (Day)", ylab="Empirical Bayes Slope Es-
timate")

# treated Bayes by dropout time for gfp tx
plot(bsgfp tx~estigfp[1:4,11],main="Empirical Bayes Slope Estimates for Treated Side of
GFP Treatment Group", xlab="Dropout Time (Day)", ylab="Empirical Bayes Slope Esti-
mate")

# untreated Bayes by dropout time for gfp tx
plot(bsgfpuntx~estigfp[1:4,11],main="Empirical Bayes Slope Estimates for Untreated Side
of GFP Treatment Group", xlab="Dropout Time (Day)", ylab="Empirical Bayes Slope Es-
timate")

# boxplots of residuals
#iestout[1:nanimal,14:(13+maxobs)]=residualstx
#iestout[1:nanimal,(14+maxobs):(13+(2*maxobs))]=residualsuntx
grp21=matrix(c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4,4),8,21)
grp20=matrix(c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4,4),8,20)
grp22=matrix(c(1,1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4,4),4,22)
hrt=estihigh[1:8,14:34]
hru=estihigh[1:8,35:55]
lrt=estilow[1:8,14:33]
lru=estilow[1:8,34:53]
grt=estigfp[1:4,14:35]
gru=estigfp[1:4,36:57]
par(mfrow=c(3,2))
boxplot(hrt~grp21,main="Boxplot of Residuals for Treated Side of High Treatment Group",
xlab="Time Period", ylab="Residuals")

```

```

boxplot(hru~grp21,main="Boxplot of Residuals for Untreated Side of High Treatment Group",
xlab="Time Period", ylab="Residuals")
boxplot(lrt~grp20,main="Boxplot of Residuals for Treated Side of Low Treatment Group",
xlab="Time Period", ylab="Residuals")
boxplot(lru~grp20,main="Boxplot of Residuals for Untreated Side of Low Treatment Group",
xlab="Time Period", ylab="Residuals")
boxplot(grt~grp22,main="Boxplot of Residuals for Treated Side of GFP Treatment Group",
xlab="Time Period", ylab="Residuals")
boxplot(gru~grp22,main="Boxplot of Residuals for Untreated Side of GFP Treatment Group",
xlab="Time Period", ylab="Residuals")
# line graph with average BBB by timepoint for each tx group
#bigytxhigh=rbind(bigytx[1,],bigytx[3,],bigytx[5,],bigytx[6,], bigytx[12,], bigytx[14,], bigytx[16,],
bigytx[18,])
#bigytxlow=rbind(bigytx[2,],bigytx[4,],bigytx[7,],bigytx[8,], bigytx[9,], bigytx[10,], bigytx[13,],
bigytx[19,])
#bigytxgfp=rbind(bigytx[11,],bigytx[15,],bigytx[17,],bigytx[20,])
#bigyuntxhigh=rbind(bigyuntx[1,],bigyuntx[3,],bigyuntx[5,],bigyuntx[6,],bigyuntx[12,], bigyuntx[14,],
bigyuntx[16,], bigyuntx[18,])
#bigyuntxlow=rbind(bigyuntx[2,],bigyuntx[4,],bigyuntx[7,],bigyuntx[8,],bigyuntx[9,], bigyuntx[10,],
bigyuntx[13,], bigyuntx[19,])
#bigyuntxgfp=rbind(bigyuntx[11,],bigyuntx[15,],bigyuntx[17,],bigyuntx[20,])
#bigthigh=rbind(bigt[1,],bigt[3,],bigt[5,],bigt[6,],bigt[12,], bigt[14,], bigt[16,], bigt[18,])
#bigtlow=rbind(bigt[2,],bigt[4,],bigt[7,],bigt[8,],bigt[9,], bigt[10,], bigt[13,], bigt[19,])
#bigtgfp=rbind(bigt[11,],bigt[15,],bigt[17,],bigt[20,])
avgyht=matrix(NA,1,maxobs)
avgyhu=matrix(NA,1,maxobs)
avgylt=matrix(NA,1,maxobs)

```

```

avgylu=matrix(NA,1,maxobs)
avgygt=matrix(NA,1,maxobs)
avgygu=matrix(NA,1,maxobs)
for (i in 1:maxobs) {
avgyht[1,i]=mean(bigytxhigh[,i],na.rm=TRUE)
avgyhu[1,i]=mean(bigyuntxhigh[,i],na.rm=TRUE)
avgylt[1,i]=mean(bigytxlow[,i],na.rm=TRUE)
avgylu[1,i]=mean(bigyuntxlow[,i],na.rm=TRUE)
avgygt[1,i]=mean(bigytxgfp[,i],na.rm=TRUE)
avgygu[1,i]=mean(bigyuntxgfp[,i],na.rm=TRUE)
}
t=c(0, 4, 7, 11, 14, 18, 21, 26, 28, 32, 35, 39, 42, 46, 49, 53, 56, 60, 63, 67, 70, 74)
par(mfrow=c(1,1))
plot(c(0,74), c(0,21), type="n", xlab="Days", ylab="Average BBB Score" )
lines(t,avgyht,type="b",lwd=1.5,lty=1,col="red",pch=1)
lines(t,avgyhu,type="b",lwd=1.5,lty=2,col="blue",pch=2)
lines(t,avgylt,type="b",lwd=1.5,lty=3,col="forestgreen",pch=3)
lines(t,avgylu,type="b",lwd=1.5,lty=4,col="darkorange3",pch=4)
lines(t,avgygt,type="b",lwd=1.5,lty=5,col="magenta",pch=5)
lines(t,avgygu,type="b",lwd=1.5,lty=6,col="purple",pch=6)
title("Mean BBB Score by Day")
legend(57,21, c("High,Treated","High,Untreated","Low,Treated","Low, Untreated", "GFP,Treated",
"GFP,Untreated"), cex=0.8, col=c("red","blue","forestgreen","darkorange3","magenta","purple"),
pch=1:6, lty=1:6, title="Legend")
sebshigh=sqrt(diag(esthigh[10:25,1:16]))
sebslow=sqrt(diag(estlow[10:25,1:16]))
sebsgfp=sqrt(diag(estgfp[10:17,1:8]))

```

```

## Histogram for ALS Score
histogram(data1$ALS,main="Histogram of ALS Score")
## Histogram for BBB Score
histogram(data1$BBB,main="Histogram of BBB Score")
## Histogram for BMS Score
histogram(data1$BMS,main="Histogram of BMS Score")
##Histogram for Rotarod
histogram(data1$Rotarod_1,main="Histogram of Rotarod_1 Score")
histogram(data1$Rotarod_2,main="Histogram of Rotarod_2 Score")
histogram(data1$Rotarod_3,main="Histogram of Rotarod_3 Score")
## Histogram for Gripstrength
histogram(data1$Gripstrength_1,main="Histogram of Gripstrength_1 Score")
histogram(data1$Gripstrength_2,main="Histogram of Gripstrength_2 Score")
histogram(data1$Gripstrength_3,main="Histogram of Gripstrength_3 Score")
## plotting ALS Score by Day for each Mouse
par(mfrow=c(4,5))
plot(data1$Day[Animal==1],data1$ALS[Animal==1], main="Scatterplot of ALS Score by
Day for Mouse 1", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==2],data1$ALS[Animal==2], main="Scatterplot of ALS Score by
Day for Mouse 2", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==3],data1$ALS[Animal==3], main="Scatterplot of ALS Score by
Day for Mouse 3", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==4],data1$ALS[Animal==4], main="Scatterplot of ALS Score by
Day for Mouse 4", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==5],data1$ALS[Animal==5], main="Scatterplot of ALS Score by
Day for Mouse 5", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==6],data1$ALS[Animal==6], main="Scatterplot of ALS Score by

```

```
Day for Mouse 6", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==7],data1$ALS[Animal==7], main="Scatterplot of ALS Score by
Day for Mouse 7", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==8],data1$ALS[Animal==8], main="Scatterplot of ALS Score by
Day for Mouse 8", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==9],data1$ALS[Animal==9], main="Scatterplot of ALS Score by
Day for Mouse 9", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==10],data1$ALS[Animal==10], main="Scatterplot of ALS Score
by Day for Mouse 10", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==11],data1$ALS[Animal==11], main="Scatterplot of ALS Score
by Day for Mouse 11", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==12],data1$ALS[Animal==12], main="Scatterplot of ALS Score
by Day for Mouse 12", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==13],data1$ALS[Animal==13], main="Scatterplot of ALS Score
by Day for Mouse 13", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==14],data1$ALS[Animal==14], main="Scatterplot of ALS Score
by Day for Mouse 14", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==15],data1$ALS[Animal==15], main="Scatterplot of ALS Score
by Day for Mouse 15", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==16],data1$ALS[Animal==16], main="Scatterplot of ALS Score
by Day for Mouse 16", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==17],data1$ALS[Animal==17], main="Scatterplot of ALS Score
by Day for Mouse 17", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==18],data1$ALS[Animal==18], main="Scatterplot of ALS Score
by Day for Mouse 18", xlab="Day", ylab="ALS Score")
plot(data1$Day[Animal==19],data1$ALS[Animal==19], main="Scatterplot of ALS Score
by Day for Mouse 19", xlab="Day", ylab="ALS Score")
```

```

plot(data1$Day[Animal==20],data1$ALS[Animal==20], main="Scatterplot of ALS Score
by Day for Mouse 20", xlab="Day", ylab="ALS Score")
##OR another way...
xyplot(data1$ALS ~ data1$Day — Animal, data=data1, groups=Side, xlab="Day", ylab="ALS
Score", type=c("p","l"))
## plotting BBB Score by Day by Side and Animal with both Sides in One panel colored
differently
xyplot(data1$BBB ~ data1$Day — Animal, data=data1, groups=Side)
## adding lines to connect the dots
xyplot(data1$BBB ~ data1$Day — Animal, data=data1, groups=Side, xlab="Day", ylab="BBB
Score", type=c("p","l"))
## plotting BMS score by day by side and animal with both sides in one panel
xyplot(data1$BMS ~ data1$Day — Animal, data=data1, groups=Side, xlab="Day", ylab="BMS
Score", type=c("p","l"))
## plotting Rotarod score same way (only using the 1st rotarod, maybe should use aver-
age?)
xyplot(data1$Rotarod_1 ~ data1$Day — Animal, data=data1, groups=Side, xlab="Day",
ylab="Rotarod Score", type=c("p","l"))
## same for gripstrength
xyplot(data1$Gripstrength_1 ~ data1$Day — Animal, data=data1, groups=Side, xlab="Day",
ylab="Gripstrength Score", type=c("p","l"))

```