A LONGITUDINAL ANALYSIS ON THE EFFECT OF DROPOUT ON INSTINCTUAL TRAUMA RESPONSE TREATMENT USING MAXIMUM LIKELIHOOD METHODS AND MULTIPLE IMPUTATION

by

GEORGIA SMITH

B.S., University of Virginia's College at Wise, 2021

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Masters of Science

Department of Biostatistics

2023

This thesis for the Master of Science degree by

Georgia Smith

has been approved for the

Department of Biostatistics

by

Matthew Strand, Chair

Diane Fairclough

Anne McFarland

Date: 18 August 2023

Georgia Smith (M.S., Biostatistics)

A Longitudinal Analysis on the Effect of Dropout on Instinctual Trauma Response Treatment Using Maximum Likelihood Methods and Multiple Imputation

Thesis directed by Prof. Matthew Strand

ABSTRACT

Instinctual Trauma Response therapy (ITR) is is an evidence-supported model for treatment of PTSD and trauma which was developed by Tinnin and Gantt (2009). ITR treatment is implemented during intensive treatment interventions which are generally performed during 3-5 consecutive 8 hour days of treatment. Despite being in practice for over 40 years, ITR has never been thoroughly quantitatively studied. This study aims to quantify the effect of ITR therapy utilizing a longitudinal dataset of five psychiatric measures which quantify symptomology collected by practitioners from 2010-2017. The collected data has minimal demographic variables and a lare amount of dropout with only 100 subjects out of 455 (21.98%) having completed the last follow up. To analyze this data we use Linear Mixed Models with random effects and non-simple correlation structures. We found that the dropout could not be assumed to be missing completely at random so we performed two analyses on each measure. First, we used mixed models which accounted for dropout using a random effect for last recorded timepoint. We then performed multiple imputation using chained equations and modelled the mixed models with the multiple imputed data. We compared the bias and precision of the estimates for these two methods for each of the five measurements. For all five measures we found there was a significant decrease in the score from the baseline measurement to all four of the post treatment measurements in both modelling strategies. We found that there was very little change after the post treatment measurements with subjects platauing after treatment. Overall, our findings support claims that ITR is a valid treatment method for trauma and that it results in significantly better treatment outcomes.

The form and content of this abstract are approved. I recommend its publication.

Approved: Matthew Strand

TABLE OF CONTENTS

CHAPTER

I INTRODUCTION **1**

Background 1

Primary Research Goals 2

Data 2

Data Cleaning 5

II METHODS **7**

Statistical Background of Modelling Methods 7

Missingness Mechanisms Background 7

Modelling Approaches to Missing Data 9

Methods Background 12

Modelling Choices 12

Exploratory Data Analysis 13

Missingness 20

Multiple Imputation 24

Model Selection 24

IES 28

SCL 29

TAS 30

log(DES) 30
log(DRS) 32

III RESULTS **38**

Measure Change Over Time 38
DES 38
DRS 40
IES 40
SCL 40
TAS 41
Impact of Dropout on Measures 41
IES 42
TAS 43
SCL 44
DES 45
DRS 45
Change Over Time 45
IES 47
TAS 48
SCL 49
DES 50
DRS 52
Multiple Imputation and Available Case Comparison 54
IES 54
TAS 55
SCL 56
DES 58

DRS 59

IV DISCUSSION **61**

Model-Based Results 61

IES 61

SCL 62

TAS 62

DES 62

DRS 63

Dropout 63

Change after Post-Treatment 64

Limitations 64

Conclusions 65

Future Reccomendations 65

REFERENCES **66**

APPENDIX

Code Appendix 68

SAS Code 68

R Code 111

LIST OF TABLES

TABLE

1	Demographic and Descriptive Measures By Dropout Timepoint	3
2	Description of All Available Variables	4
3	Subset of Non-Unique IDs Identified by Discrepancies in Age	5
4	Subset of Modified Non-Unique IDs Identified by Discrepancies in Age	5
5	Identification of Non-Unique Primary Diagnoses by ID	6
6	Regression Models Investigating the Effect of Dropout Timepoint on Baseline Scores, Post-Treatment Scores, and the Difference Between Baseline and Post-Treatment Scores	22
7	Correlation between Demographic Covariates and Dropout Timepoint	23
8	Initial Correlation Structure Selection Output	34
9	Fixed Effect Model Selection Output	35
10	Correlation and Random Effect Selection for IES	35
11	Correlation and Random Effect Selection for SCL	36
12	Correlation and Random Effect Selection for TAS	36
13	Correlation and Random Effect Selection for log(DES)	37
14	Correlation and Random Effect Selection for log(DRS)	37
15	Estimated change from Baseline to 6 Months For all Measures by Model	39
16	Contrasts For IES in the Available Case Model	42

17 Contrasts For TAS in the Available Case Model 44

18 Contrasts For SCL in the Available Case Model 45

19 Contrasts For DES in the Available Case Model 45

20 Contrasts For DRS in the Available Case Model 46

LIST OF FIGURES

FIGURE

1 Spaghetti Plot of Score Over Time By Dropout Timepoint and Measurement Type 14

2 Spaghetti Plot of Score Over Time By Dropout Timepoint and Measurement Type with LOCF Removed 15

3 Spaghetti Plot of Score Over Days Since Baseline By Dropout Timepoint and Measurement Type with LOCF Removed 16

4 Spaghetti Plot of Score Over Days Since Post Treatment By Dropout Timepoint and Measurement Type with LOCF Removed 17

5 Spaghetti Plot of Score over Days Since Post Treatment by Measurement Type with a LOESS average 18

6 Histogram of Outcomes to Test Normality 19

7 Upset Plot of the Top 10 Missing Variables 21

8 Difference in Baseline Estimates between Baseline Dropout and All other Dropout Timepoints 43

9 Total Change Across Dropout Timepoints for TAS 44

10 Estimated Values of Measures Across Timepoints 46

11 Estimated Change Between Timepoints for IES 48

12 Estimated Change Between Timepoints for TAS 49

13 Estimated Change Between Timepoints for SCL 50
14 Estimated Change Between Timepoints for DES 52
15 Estimated Change Between Timepoints for DRS 53
16 β Estimates for IES 55
17 β Estimates for TAS 56
18 β Estimates for SCL 57
19 β Estimates for log(DES) 59
20 β Estimates for log(DRS) 60

CHAPTER I

INTRODUCTION

Background

Instinctual Trauma Response (ITR) Treatment is an evidence-supported model for the treatment of PTSD and trauma which was developed by Gantt and Tinnin (2009). ITR is informed by art therapy, narrative therapy, and parts therapy, and based on studies of animal survival strategies, clinical observations, and brain science to combine verbal and nonverbal processing (Tinnin, Bills, and Gantt 2014). To date, ITR has not been extensively quantitatively studied despite being in practice for over 40 years. An extensive quantitative study is needed in order to quantify the efficacy of ITR for funding and for evidence for future patients.

Gantt and Tinnin (2007) did an outcome study on ITR and found that out of the 72 patients, 45% met the criteria for recovery, 44% were improved, 8% were unchanged, and two patients were worse after treatment. In patients with PTSD, 77% of the 22 PTSD patients recovered and the remaining 23% improved. And, in the 50 patients presenting dissociation 32%, 54% improved, 12% were unchanged, 4% were worse. While this study did show a decrease in symptomology post treatment, it did not account for things such as attrition, variation between subjects, and how a measure decreases over time. We analyzed a new

cohort of 455 subjects and will account for these things in our analysis. Based on Gantt and Tinnin (2007), though, we do expect to see a decrease in psychiatric measures indicating improved outcomes for patients. Note that all subject data was collected in a prior outcome analysis and was deidentified prior to our access to it.

Primary Research Goals

Our provided data included 455 subjects with five psychiatric measures that quantify symptomology. A decrease in scores indicates a decrease in symptomology and improved health outcomes for all five measures. With that in mind, our primary research question is whether or not we see a significant decrease in measures over the follow-up period.

This dataset included significant dropout meaning that the analyses had to be carefully done. We used multiple methods to account for this dropout and then compared to ensure our conclusions are correct. This dropout also gave us a secondary goal of quantifying the effect of dropout on measures over time. Our tertiary goal is to determine if measures change over time. ITR treatment includes psychoeducation which allows patients to treat themselves after their initial treatment with a practitioner, potentially leading to a continuing decrease in psychiatric measures over time post-treatment.

Data

This dataset was provided by Help For Trauma, a trauma therapy agency which provides ITR as their primary treatment under the supervision of creators Dr. Louis Tinnin and Dr. Linda Gantt. Help for Trauma continuously collected measures provided by practitioners, who followed up with their patients. The dataset spans individuals who received

ITR treatment from 2010-2017 from Help For Trauma practitioners. Practitioners collected measures pre and post-treatment and contacted patients for follow-up measures at 1-week post-treatment, 3 months post-treatment, and 6 months post-treatment. Although, many people did not respond to these contact attempts leading to a high dropout rate.

Basic demographics and descriptive measures, stratified by the last follow-up timepoint, are described in Table 1. Out of 455 total patients, 279 (61.3%) dropped out before the 3-month post-treatment follow-up timepoint. The data has skewed diagnoses with those with a primary diagnosis of adjustment disorder comprising 1.1% of individuals. The subjects are majority female with 75.6% female. This is not abnormal for trauma data with women experiencing PTSD two to three times as often as men do, on average (Olff 2017).

Table 1: Demographic and Descriptive Measures By Dropout Timepoint

	Baseline	Post Treatment	1 Week	3 Months	6 Months	Overall
	(N=10)	(N=119)	(N=150)	(N=76)	(N=100)	(N=455)
Treatment Days						
Mean (SD)	7.50 (2.64)	6.71 (3.07)	7.40 (2.89)	7.85 (3.42)	8.17 (4.01)	7.52 (3.34)
Median [Min, Max]	7.50 [5.00, 10.0]	5.00 [2.00, 15.0]	5.00 [1.00, 18.0]	5.00 [5.00, 20.0]	10.0 [2.00, 25.0]	5.00 [1.00, 25.0]
Missing	0 (0%)	29 (24.4%)	7 (4.7%)	1 (1.3%)	2 (2.0%)	39 (8.6%)
Age						
Mean (SD)	31.0 (16.9)	38.1 (11.8)	39.0 (13.0)	40.4 (14.2)	38.0 (12.3)	38.6 (12.9)
Median [Min, Max]	21.0 [17.0, 57.0]	39.0 [16.0, 82.0]	38.5 [15.0, 66.0]	43.0 [17.0, 65.0]	39.0 [13.0, 66.0]	39.0 [13.0, 82.0]
Missing	0 (0%)	4 (3.4%)	6 (4.0%)	3 (3.9%)	2 (2.0%)	15 (3.3%)
Primary Diagnosis						
Missing	0 (0%)	2 (1.7%)	0 (0%)	0 (0%)	0 (0%)	2 (0.4%)
PTSD	4 (40.0%)	60 (50.4%)	80 (53.3%)	37 (48.7%)	46 (46.0%)	227 (49.9%)
DID	1 (10.0%)	35 (29.4%)	54 (36.0%)	32 (42.1%)	45 (45.0%)	167 (36.7%)
DDNOS	0 (0%)	18 (15.1%)	15 (10.0%)	7 (9.2%)	8 (8.0%)	48 (10.5%)
Adjustment.Disorder	1 (10.0%)	2 (1.7%)	1 (0.7%)	0 (0%)	1 (1.0%)	5 (1.1%)
Missing.1	4 (40.0%)	2 (1.7%)	0 (0%)	0 (0%)	0 (0%)	6 (1.3%)
Gender						
Female	5 (50.0%)	89 (74.8%)	117 (78.0%)	55 (72.4%)	78 (78.0%)	344 (75.6%)
Male	5 (50.0%)	30 (25.2%)	33 (22.0%)	21 (27.6%)	22 (22.0%)	111 (24.4%)

Help For Trauma collects five measures from simple questionnaires to quantify the effect of traumatic event(s) on subjects' lives: DES, DRS, IES, SCL, and TAS. DES, or the Dissociative Experiences Scale, screens for dissociative symptoms. For DES, patients rate symptoms 0%-100% and all questions are averaged to receive a score. DRS, or the Dissociative Regression Scale, measures the impact of dissociation on functioning. For DRS,

patients rate symptoms 0%-100% and all questions are averaged to receive a score. TAS or the Toronto Alexithymia Scale measures patients' ability to verbalize feelings. For TAS, patients rate questions 1-5 and answers are summed to get a final score of 20-100. IES, or the Impact of Events Scale, measures the impact of a single event on intrusion and avoidance symptoms over the past 7 days. For IES, patients rate symptoms 0-5 and all questions are summed to receive a score. SCL, or the Symptom Checklist – 45, measures general psychiatric distress. For SCL, patients rate questions 0-4 and answers are summed to receive a score. A description of all variables can be found in Table 2.

Table 2: Description of All Available Variables

Name	Description
ID	Unique Subject IDs
Follow-Up Timepoint	Factor representing follow up timepoint which measure was collected
Dissociative Regression Scale	Average of DRS symptom rating
Dissociative Experience Scale	Average of DES symptom rating
Toronto Alexithymia Scale	Sum of TAS symptom rating
Symptom Checklist 45	Sum of SCL symptom rating
Impact of Events Scale	Sum of IES symptom rating
Treatment Days	Number of treatment days
Age	Numeric Age of Subject
Gender	Gender (Male or Female)
Primary Diagnosis	Factor representing Primary Diagnosis Endorsement
Hallucinations	Boolean representing hallucination endorsement based on responses to the DES and SCL
Type I S	Boolean representing Single Trauma type
Type I M	Boolean representing multiple different traumas
Type II	Boolean representing repeated traumas of the same type
Alexithymia Post Treatment	Boolean representing if alexithymia is present post-treatment
Alexithymia Pre Treatment	Boolean representing if alexithymia is present before treatment

Data Cleaning

The data were received in 11 Microsoft Excel sheets which had multiple repeated subject IDs with non-unique data. To find unique subject data the provided data was analyzed and duplicate IDs for non-unique subjects were identified by discrepancies in age as seen in Table 3. These IDs were then modified to reflect this age to create new unique IDs as seen in Table 4.

Table 3: Subset of Non-Unique IDs Identified by Discrepancies in Age

ID	Gender	Age	Treatment Days	Primary Diagnosis
2	1	24	10	DDNOS
2	2	58	5	PTSD
3	2	29	3	PTSD
3	1	40	10	DDNOS
4	2	32	10	PTSD
4	2	52	10	DID

Table 4: Subset of Modified Non-Unique IDs Identified by Discrepancies in Age

Old ID	New ID	Gender	Age	Treatment Days	Primary Diagnosis
2	2-24	1	24	10	DDNOS
2	2-58	2	58	5	PTSD
3	3-29	2	29	3	PTSD
3	3-40	1	40	10	DDNOS
4	4-32	2	32	10	PTSD
4	4-52	2	52	10	DID

Diagnoses were constant across repeated subjects except for two subjects identified in Table 5. These subjects' diagnoses were set as PTSD post-identification based on the attribution of the discrepancies to input error.

Overall, this data presented multiple complications on the front end due to poor collection methods. It also contained a large amount of dropout. We analyzed the data further

Table 5: Identification of Non-Unique Primary Diagnoses by ID

ID	Gender	Age	Primary Diagnosis
187	1	38	None
187	1	NA	PTSD
529	1	62	PTSD
529	1	62	s

to ensure assumptions for modelling methods can be met and that the missing data was handled in a responsible matter.

CHAPTER II

METHODS

Statistical Background of Modelling Methods

Longitudinal studies follow subjects over time with either continuous or repeated collection of measures (Coggon, Barker, and Rose 2009). These studies often have great size and complexity and must be analyzed using practices to account for the correlation within-subject measures over time. To analyze our data we will be using a linear mixed model approach (LMM) which will fit a regression-type model utilizing fixed effects with random effects and a non-simple error-covariance structure to account for between-subject variability.

Missingness Mechanisms Background

While linear mixed models are very robust to missing data, some missing data mechanisms are nonignorable. The basis of this classification is three mechanisms of missingness pioneered by Rubin (1976): missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). To investigate the properties of missingness mechanisms we will be using the notation outlined in Van Buuren (2018) in his book **Flexible Imputation of Missing Data**. Let R be some matrix that represents subjects' responses where $R = 0$ refers to missing entries and $R = 1$ refers to observed entries. Let Y_{obs} be the

observed data and Y_{mis} be the missing data. If $Y_{obs} = Y$ then there is no missing data. Let ψ contain the parameters of the missing data model and θ contain the parameters for the full data Y.

MCAR data is data in which the probability that an entry is missing is only dependent on some parameters ψ. That is, $P(R = 0|Y_{obs}, Y_{mis}, \psi) = P(R = 0|\psi)$. MAR data is missing data that has a pattern to the missingness but can predicted by parameters ψ and Y_{obs}, the observed data. In other words, $P(R = 0|Y_{obs}, Y_{mis}, \psi) = P(R = 0|Y_{obs}, \psi)$. MNAR data is that which the probability of missingness depends on some unobserved information where $P(R = 0|Y_{obs}, Y_{mis}, \psi)$ does not simplify.

To determine if the missingness pattern can be classified as MCAR, multiple methods can be used. One common robust method that is used is Little's test for missing completely at random (Little 1988a) and is easy to implement using packages distributed in languages such as `R` and `SAS`. Little's test evaluates if there is any significant difference between the means of different missing-value patterns using a χ^2 test. In this test, the null hypothesis is that there is no significant difference between the means of different missing-value patterns. While this method does provide good initial analysis, it has low statistical power and Type II errors can occur (i.e. failing to detect MAR). Additionally, Little's MCAR test can detect MCAR when missing observations are MNAR (Enders 2010; as cited in Nicholson, Deboeck, and Howard 2017).

There is no test commonly used and validated to differentiate between MAR and MNAR. In fact, it is impossible since the data needed to test is missing. To differentiate between the two patterns, a more qualitative approach must be employed and potential unmeasured dropout predictors should be considered. Instead of being classified by the mechanisms of missingness set forth by Rubin (1976), the missingness is often separated into two categories: ignorable and nonignorable missingness based on their impact on likelihood

inference.

Ignorable missingness was defined by Little and Rubin (2019) in their book **Statistical Analysis with Missing Data** as data that is MAR and parameter spaces ψ and θ are distinct. Assuming ignorability implies that the available data is sufficient to correct for the effects of missingness.

Modelling Approaches to Missing Data

In the case of MCAR, complete case analysis can be used, which removes any entries with missing data. Doing this erroneously can lead to bias and loss of precision. When using a complete case, it is also important to note it gives a loss of statistical power by decreasing the amount of subjects in the model. The most popular method and an alternative that works for all ignorable missingness (MAR and MCAR) is multiple imputation, which performs single imputation multiple times and then pools the results to get unbiased estimations (Nakai and Ke 2011).

Single imputation replaces missing values with an estimate of the true value of the missing value. There are multiple ways this estimate can be generated. One popular method in longitudinal analyses is the last observation carried forward (LOCF), in which all missing values attributed to dropout are replaced with the last available measure for the subject. This is often unrealistic since it assumes no change after dropout and can often result in bias. Mean imputation is also used, where a missing value is replaced with the mean of the non-missing values. This can alter the distribution of a variable and cause an underestimation of standard error. Hot deck imputation is a method commonly used in survey modeling and replaces missing values with values from subjects with similar data that are complete. Effective hot deck imputation often requires a large sample size so that a sufficiently similar responding subject can be found. Hot deck imputation can also cause distortions of correlations and covariance. All of the above single imputation methods are not best practices due to their

limitations with regards to standard error or biases of the predictors (Nakai and Ke 2011).

There are also prediction methods to estimate the true values of missing responses. These methods calculate a regression line and use the estimate from the regression line to estimate the true value of a dropout. To allow for variation between subjects, we can add random noise around the regression line along with parameter uncertainty. Parameter uncertainty can be calculated using Bayesian sampling or bootstrap resampling of the observed data. Doing this type of imputation using noise and parameter uncertainty is easily implementable, though can still result in underestimation of uncertainty or poor efficiency (Van Buuren 2018).

Predictive mean matching (PMM) is a nearest neighbors strategy that performs very well proposed by Rubin (1986) and Little (1988b). It is a very popular strategy because it is optimized for each target variable, has straightforward calculations, and is easy to include nominal and ordinal variables (Van Buuren 2018). Multiple studies have found that it outperforms other methods. Marshall et al. (2010) found that for Cox proportional hazards models predictive mean matching with regression switching outperformed data augmentation, regression switching, and flexible additive imputation models for incomplete skewed data but adds the caveat that it would not be recommended for data with more than 50% missingness. Shaw et al. (2023) extended the high performance of PMM to longitudinal models and confirmed that for longitudinal data PMM outperformed other imputation methods.

Predictive mean matching works by calculating the predicted value of each Y_{mis} by forming a subset of neighbors with complete cases that have predicted values closest to a predicted value of a missing entry. One value is selected from the subset of neighbors and their non-missing value replaces the missing value. Predictive mean matching is robust against misspecification and handles transformed data well (Van Buuren 2018).

Once a single imputation prediction method is selected, multiple imputation performs

the single imputation multiple times to get a variation of predicted values and performs model-building on each imputation. The results are then pooled, or combined, to get one set of estimates for a model. Multiple imputation outperforms single imputation because it allows for the variability of the predicted values and thus corrects the standard error to allow for accurate hypothesis testing (Nakai and Ke 2011).

Combining Multiple Imputed Data

Pooling models fit on multiple imputed data is a straightforward process. Given some set of quantities Q in the population Q is the estimate of those quantities that is unbiased and confidence valid. The posterior mean of $P(Q|Y_{obs}) = E(E[Q|Y_{obs}, Y_{mis}]|Y_{obs})$, or the average of the posterior means of Q over the multiple imputed data. This can be calculated as so,

$$\bar{Q} = \frac{1}{m}\sum_{l=1}^{m} Q_l,$$

where m is the number of imputation and Q_l is the estimate of l^{th} imputation. The variance $V(Q|Y_{obs})$ can be calculated as $E[V(Q|Y_{obs}, Y_{mis})|Y_{obs}] + V[E(Q|Y_{obs}, Y_{mis})|Y_{obs}]$. For $m = \infty$ imputation variance $T = \bar{U}_\infty + B_\infty$ where $\bar{U}_\infty$ are the between components $E[V(Q|Y_{obs}, Y_{mis})|Y_{obs}]$ and B_∞ are the within components $V[E(Q|Y_{obs}, Y_{mis})|Y_{obs}]$ of the variance. $\bar{U}$ and B for some finite m can be calculated $\bar{U} = \frac{1}{m}\sum_{l=1}^{m} \bar{U}_l$, and $B = \frac{1}{m-1}\sum_{l=1}^{m}(Q_l - \bar{Q})(Q_l - \bar{Q})'$. Since for a finite m $\bar{Q}$ only approximates $\bar{Q}_\infty$, more is needed. Rubin (1987) found that the variance between these estimates is approximately equal to $\frac{B_\infty}{m}$ and since B approximates B_∞, variance T of $\bar{U}$ is $\bar{U} + (1 + \frac{1}{m})B$ and is also the variance of $Q - \bar{Q}$ when $\bar{Q}$ is unbiased (Van Buuren 2018).

With aggregated estimates and variance $\bar{Q}$ and T statistical hypothesis tests become trivial except for degrees of freedom. Rubin (1987) developed a method for calculating

degrees of freedom that accounts for missing data and multiple imputations. Let $\lambda = \frac{B+B/m}{T}$ as the proportion of variation that is attributable to missing data and $r = \frac{B+B/m}{U} = \frac{\lambda}{1-\lambda}$ as the "relative increase in variance due to nonresponse" (Rubin 1987; as cited in Van Buuren 2018). Then, degrees of freedom $\nu = (m-1)(1+\frac{1}{r^2})$. This was adapted by Barnard and Rubin (1999) to account for errors when the sample size is small. Let $\nu_{obs} = \frac{\nu_{com}+1}{\nu_{com}+3}\nu_{com}(1-\lambda)$ where ν_{com} be the degrees of freedom for $\bar{Q}$ for complete data, with k parameters in a model with sample size n, $\nu_{com} = n-k$. ν_{adj} then, as the degrees of freedom adjusted for small sample size used for multiple imputation is then $\frac{\nu\nu_{obs}}{\nu+\nu_{obs}}$ where ν is the original degrees of freedom method defined in Rubin (1987). Note that this ν_{adj} converges to ν as ν_{com} approaches ∞ (Van Buuren 2018).

Methods Background

All analyses were performed on an OSX system with 16 GB of memory. SAS 9.4 was used for all mixed model analyses while R 4.2 was used for basic statistics, visualizations, and data pre-processing. For all hypothesis tests, a decision criteria of $\alpha = 0.05$ was used based on standards in the field. For multiple imputation `PROC MI` from SAS was used with estimates pooled using `PROC MIANALYZE`. For all mixed models `PROC MIXED` was utilized with `ESTIMATE` and `CONTRAST` matrices from SAS. Model selection was performed using the `ALLMIXED4` macro in SAS (Fernandez 2007). All R and SAS code is presented in the appendix.

Modelling Choices

To determine the optimal way of modeling this data we first performed exploratory data analysis. Once the basic data had been analyzed, we investigated the missingness

to determine patterns and mechanisms of the missing data. With that information, we continued to select modeling strategies and investigated parameter options.

Exploratory Data Analysis

Our first step in exploratory data analysis was to determine if the data was approximately normal and linear. We plotted histograms and spaghetti plots to assess this for each of our five outcomes. Figure 1 shows the spaghetti plot of scores over time faceted by dropout timepoint and measurement type. Those who had a last follow-up, or dropout, at the post-treatment time point have completely horizontal lines. That is, they show no change. We are unsure if this is due to merging error from multiple Excel spreadsheets or user input error as all data was collected and stored in multiple Excel spreadsheets which is known to result in errors. This pattern is also present in the Post-Treatment to 1 Week subject group among subjects whose last measurement was taken at 1-week post-treatment. We presume that this is due to erroneously performed baseline-carried-forward as a method for dealing with dropout. For subjects whose last observation is at post-treatment or 1-week post-treatment, we removed timepoints from subjects where there was a baseline carried forward entry, the updated values can be seen in Figure 2.

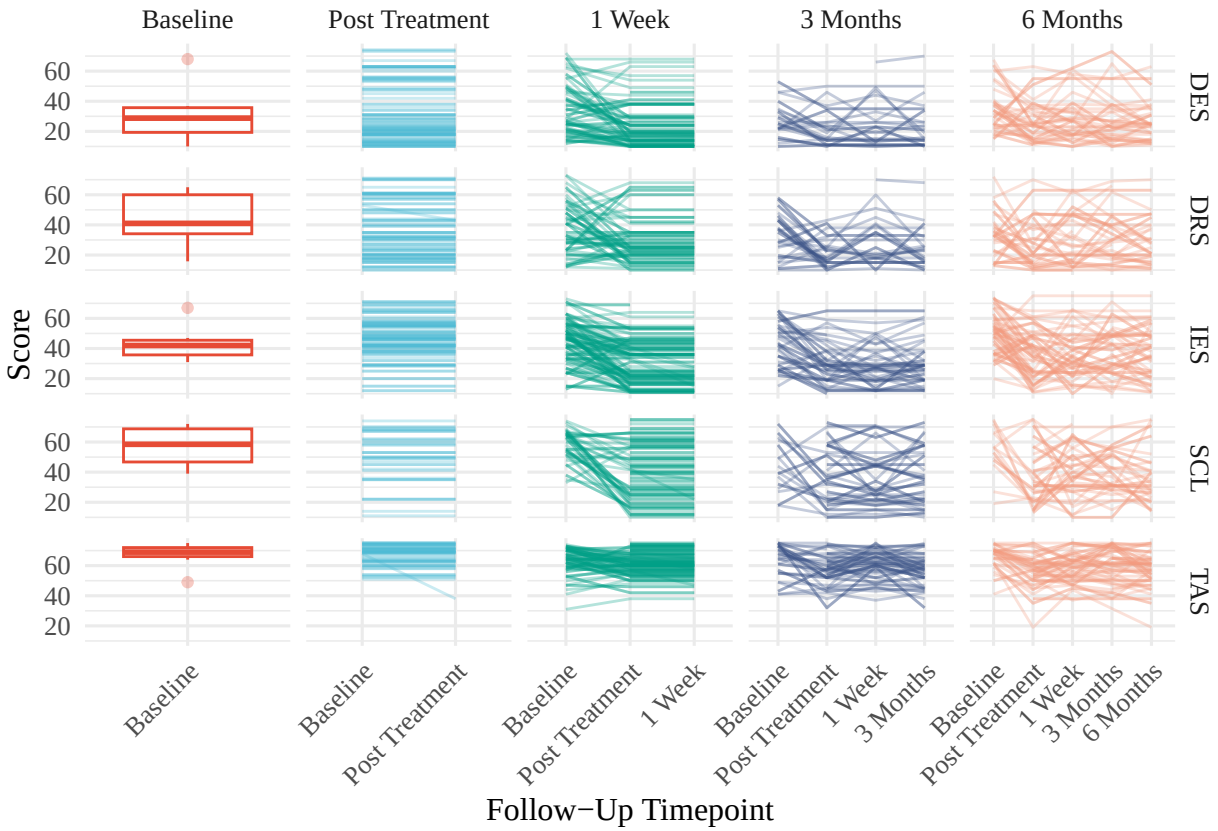

Figure 1: Spaghetti Plot of Score Over Time By Dropout Timepoint and Measurement Type

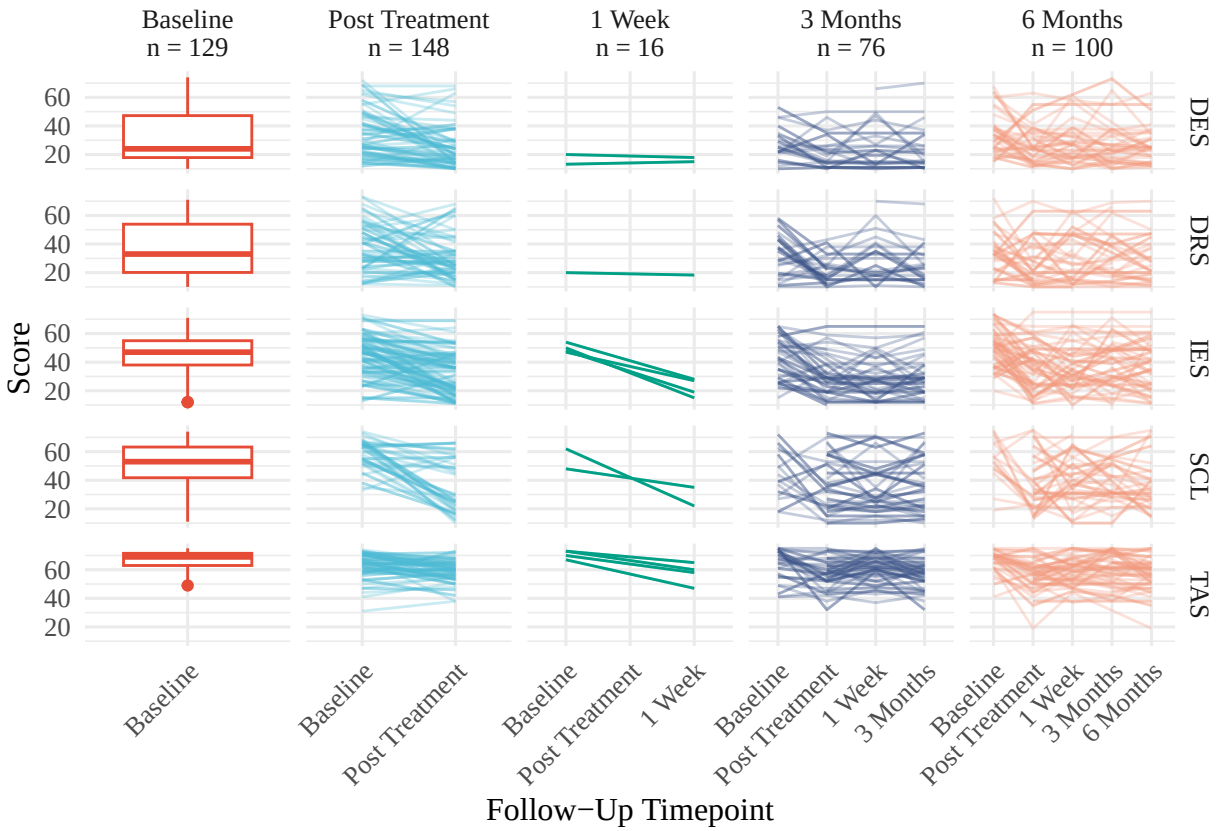

Figure 2: Spaghetti Plot of Score Over Time By Dropout Timepoint and Measurement Type with LOCF Removed

Based on Figure 2, a linear assumption could be upheld in an LMM with a random intercept to account for between-subject variance and a random slope to account for within-subject variance. Another view, which shows the patterns of the data clearer, uses days since baseline as the x variable. Days since baseline is calculated as treatment days + follow-up days. For example, instead of the one-week timepoint being a factor, it would instead be number of treatment days + 7. Post-treatment would be the number of treatment days. This view is seen in Figure 3.

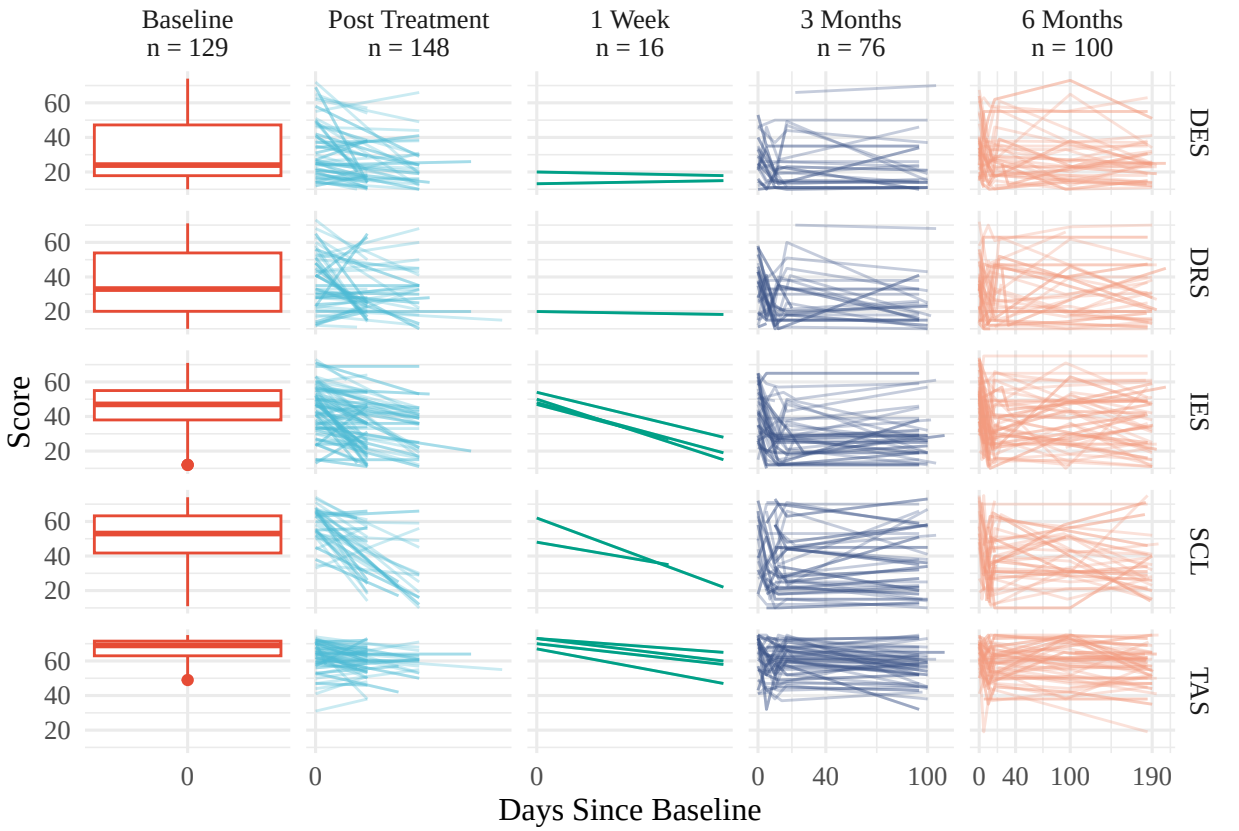

Figure 3: Spaghetti Plot of Score Over Days Since Baseline By Dropout Timepoint and Measurement Type with LOCF Removed

We performed one more transformation of the data to see patterns even clearer: we centered post-treatment at zero. This was done for two reasons. First, it gives clear increments to our post-treatment follow-up treatments at 7 days, 30 days, and 180 days. Second, there is a pattern of steep dropoff from baseline to post-treatment. By transforming the data to be centered at post-treatment, we can see this across subjects in Figure 4.

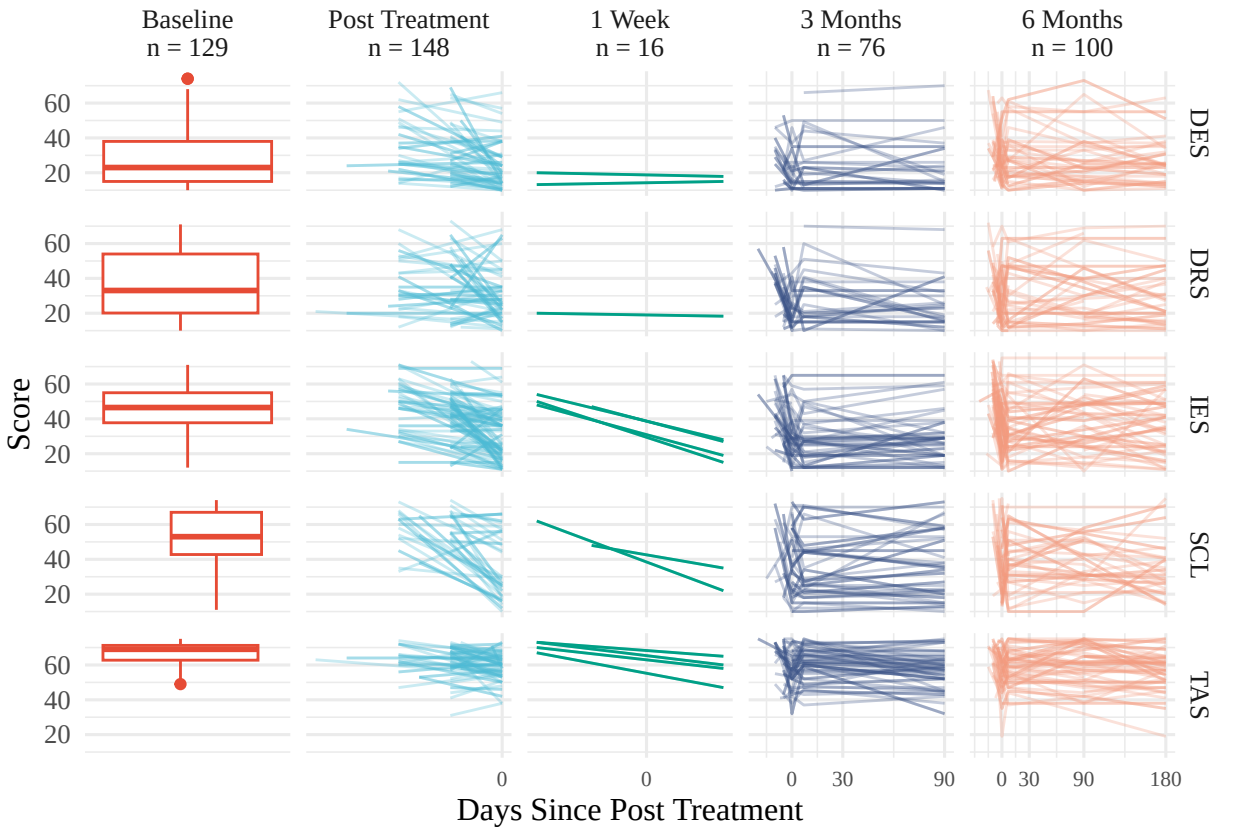

Figure 4: Spaghetti Plot of Score Over Days Since Post Treatment By Dropout Timepoint and Measurement Type with LOCF Removed

Based on the steep dropoff from Baseline to Post Treatment, we hypothesized that a knot at Post Treatment would work well in our model. Figure 5 supports this with the LOESS average across all subjects showing that the steep drop-off with leveling out is the pattern.

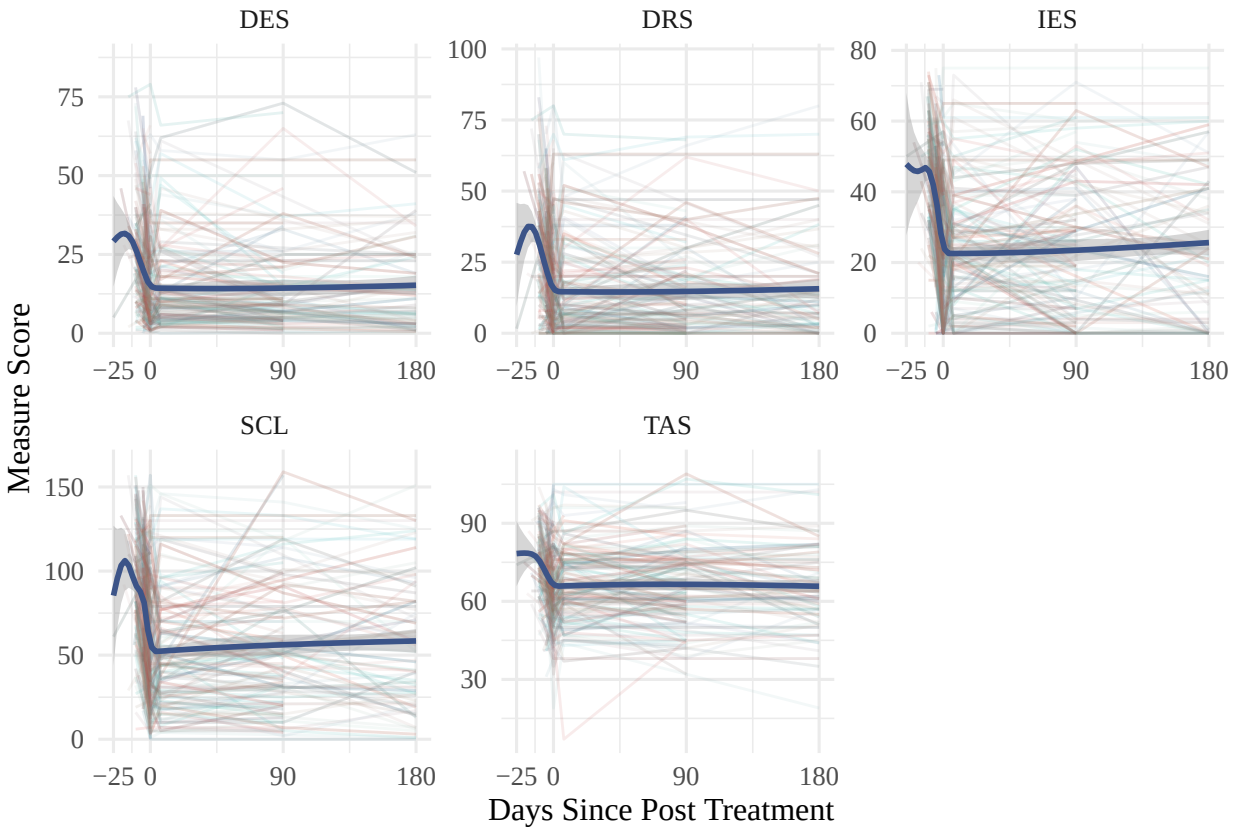

Figure 5: Spaghetti Plot of Score over Days Since Post Treatment by Measurement Type with a LOESS average

To investigate normality, we plotted a histogram of outcomes by measure, and also the log transformation of the measure to see if a simple transformation would work in the case of skewed data. This can be seen in Figure 6. From this, we found DES and DRS were skewed, but are approximately normal when log transformed. The rest of the measures are approximately normal.

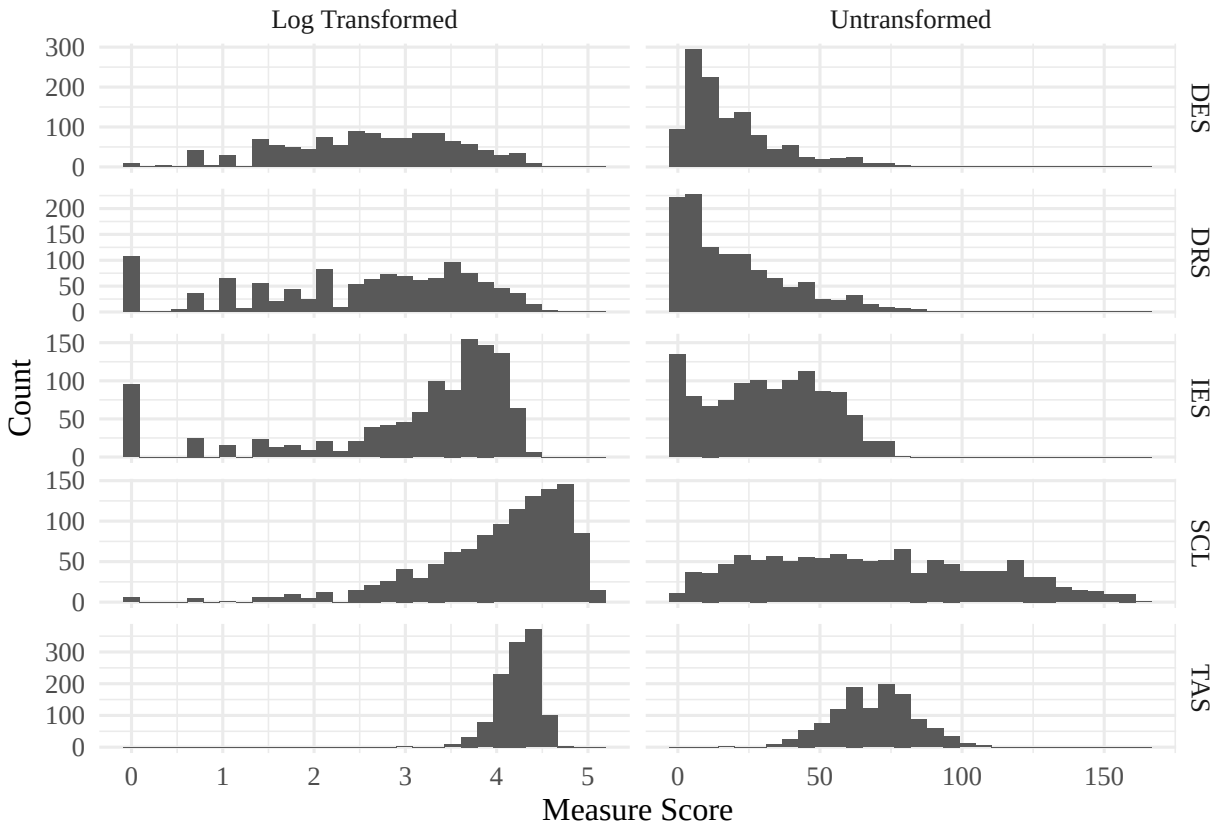

Figure 6: Histogram of Outcomes to Test Normality

Other Considerations

One additional assumption for linear mixed models is the assumption that responses between subjects are independent or accounted for by our random effects. One issue with this assumption is that we have subjects nested in clinicians, that is - there are multiple clinicians who each treat their subjects and report their measures. We have no information on which subjects were treated by which practitioners. This is an issue because practitioners could be an extraneous variable and affect the between-subject independence. We will assume between-subject independence with the caveat that our estimates may have a slight upward bias and increased variability based on the simulations by Schielzeth et al. (2020). Schielzeth et al. (2020) found that fixed effects remain unaffected by a missing random effect predictor

while “the variance arising from unmodelled lower-level predictors [were] largely absorbed by the residual variance”.

Missingness

We analyzed the missingness within the dataset to identify if the dropout was random or if it was correlated with any variables. Based on the literature discussed above, we used multiple methods to investigate the missingness. First, we applied Little’s MCAR test to the data. The data were transformed into wide format with columns for each of the five symptom quantifying measures and each of the five time points within those measures. This test, as noted above, performs a hypothesis test with the null being that there is no pattern to the missingness. The test found p-value $p = 1$ and we deemed it inconclusive since the test is known to have limitations. We continued investigating the mechanisms of missingness using graphic visualizations. Figure 7 shows the upset plot of the top ten most missing variables with their patterns. The most predominant patterns are those with dropout, while there are some subjects with missing datapoints at mid-way points (such as missing the 1-week follow-up) and returning at a later timepoint (3-months or 6-months follow-up). There is no clear pattern concerning other variables contributing to missingness based on this plot.

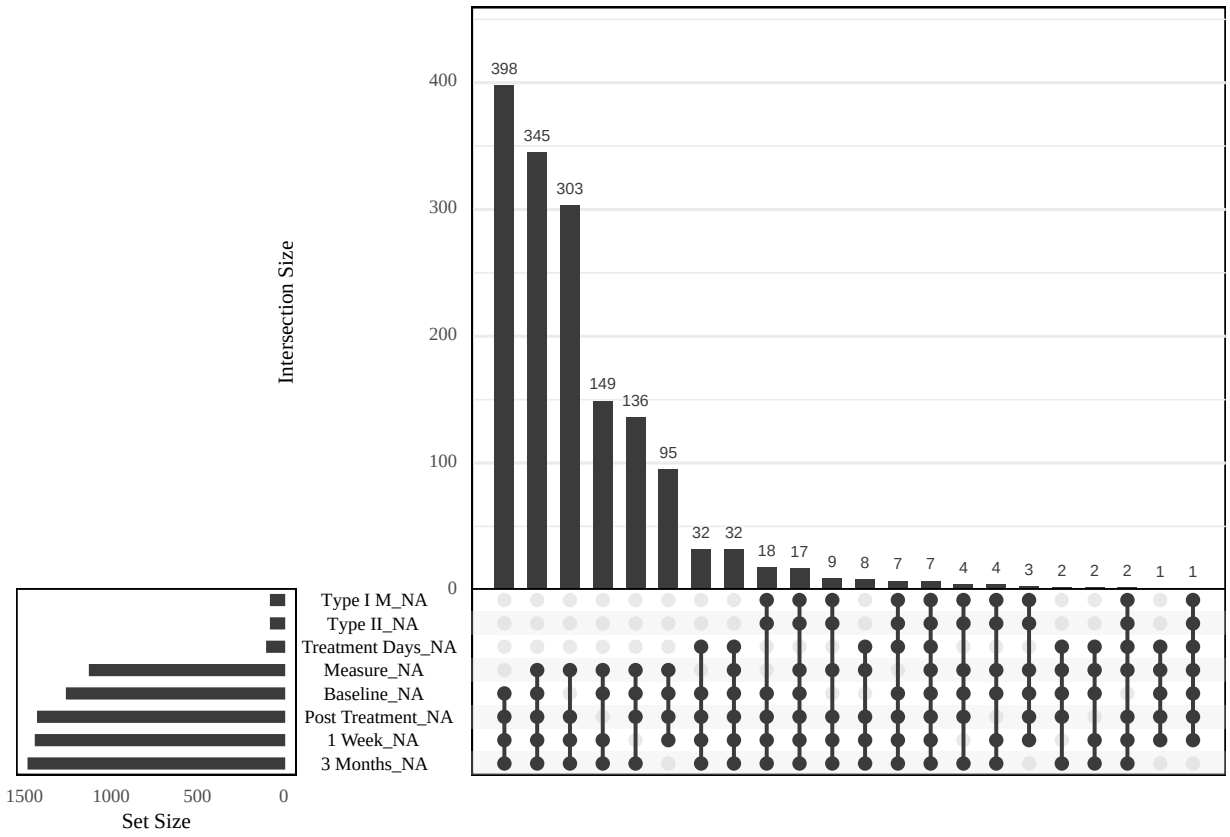

Figure 7: Upset Plot of the Top 10 Missing Variables

For each outcome, we then built three linear regression models to identify if the dropout timepoint as a category was a significant predictor of baseline score, post-treatment score, and the change from baseline to post-treatment to identify if people with worse scores were more likely to drop out. This was done due to a hypothesis that dropout was an indicator of a worse treatment outcome. The results from these models can be seen in Table 6. Those who dropped out at the 3-month and 6-month time points had significantly higher post-treatment TAS scores than those whose last follow-up timepoint was post-treatment. Those whose last follow-up time point was 6 months had a significantly higher baseline IES score than other time points.

We also looked at the demographic variables to identify if any were correlated with the

Term	Estimate	P-Value
Baseline Score		
Post Treatment	-3.93	0.08
1 Week	0.34	0.95
3 Months	-4.56	0.09
6 Months	-3.71	0.13
Change from Baseline to Post		
1 Week	8.36	0.25
3 Months	-3.24	0.11
6 Months	-0.57	0.76
Post-Treatment Score		
1 Week	5.37	0.47
3 Months	-2.28	0.27
6 Months	-0.62	0.75

(a) DES

Term	Estimate	P-Value
Baseline Score		
Post Treatment	-4.67	0.07
1 Week	6.21	0.37
3 Months	-4.26	0.17
6 Months	-4.95	0.08
Change from Baseline to Post		
1 Week	21.50	0.01
3 Months	-3.84	0.11
6 Months	-1.01	0.64
Post-Treatment Score		
1 Week	11.16	0.26
3 Months	-4.18	0.14
6 Months	-0.63	0.81

(b) DRS

Term	Estimate	P-Value
Baseline Score		
Post Treatment	-2.64	0.08
1 Week	5.10	0.20
3 Months	-1.14	0.52
6 Months	1.82	0.26
Change from Baseline to Post		
1 Week	11.55	0.11
3 Months	-2.23	0.27
6 Months	-0.52	0.78
Post-Treatment Score		
1 Week	-3.49	0.60
3 Months	-3.98	0.03
6 Months	-4.98	<2e-16

(c) TAS

Term	Estimate	P-Value
Baseline Score		
Post Treatment	-0.03	0.99
1 Week	6.12	0.28
3 Months	-1.75	0.50
6 Months	4.90	0.04
Change from Baseline to Post		
1 Week	12.58	0.19
3 Months	-4.43	0.10
6 Months	0.46	0.85
Post-Treatment Score		
1 Week	9.59	0.34
3 Months	-2.02	0.49
6 Months	-4.51	0.09

(d) IES

Term	Estimate	P-Value
Baseline Score		
Post Treatment	-1.56	0.71
1 Week	4.32	0.70
3 Months	-5.85	0.24
6 Months	1.88	0.68
Change from Baseline to Post		
1 Week	3.54	0.85
3 Months	-8.18	0.12
6 Months	-0.35	0.94
Post-Treatment Score		
1 Week	9.28	0.62
3 Months	-2.46	0.65
6 Months	-2.77	0.57

(e) SCL

Table 6: Regression Models Investigating the Effect of Dropout Timepoint on Baseline Scores, Post-Treatment Scores, and the Difference Between Baseline and Post-Treatment Scores

Table 7: Correlation between Demographic Covariates and Dropout Timepoint

Variable Name	Test	P-Value
Age	ANOVA	0.095
Hallucination	Chi-Squared	0.281
Treatment Days	Fisher's Exact	0.001
Type I M	ANOVA	<2e-16
Type I S	Chi-Squared	0.003
Alexithymia Pre.	Chi-Squared	0.379
Type II	Chi-Squared	0.016
Alexithymia Post	Chi-Squared	<2e-16
Primary Diagnosis	Chi-Squared	<2e-16
Gender	Chi-Squared	<2e-16

dropout timepoint. This is shown in Table 7. Diagnosis, number of treatment days, hallucination, Type II, alexithymia pre, and alexithymia post were all significantly correlated with dropout timepoint based on their respective hypothesis test. Age was borderline significant with a p-value of 0.095.

All of this indicates that the dropout cannot be classified as ignorable missingness and more work needed to be done to account for the dropout. Based on the literature above, we decided to do multiple analyses. First, we used linear mixed models and account for dropout with an interaction between the follow-up time and dropout time. This model accounts for the variability between dropout time points. Then, we compared these estimates with an analysis done using multiple imputations. This analysis used predictive mean matching to impute missing values in the data. We compared the bias and precision of the estimates from the complete case linear mixed model to the multiple imputed linear mixed model.

Multiple Imputation

For the imputation, we needed to choose the covariates used in the imputation and the number of imputations to run. We used 50 imputations based on a paper from White, Royston, and Wood (2011) which found that using the same number of imputations as the percentage of incomplete cases. We found that the average percentage of missingness across time points is 49.37/%. We chose to use predictive mean matching for all continuous missing variables because of its flexibility and performance. For binary variables, we used a logistic regression and for nominal data used a discriminant analysis. Each imputed variable needed covariates specified to correctly estimate the missing values. Including all possible covariates is normally recommended but with five time points across five measures this would include at minimum 25 variables. This can often lead to an overburdened model.

To avoid this we chose the covariates for the demographic variables based on a correlation test between variables and for the measures we used all demographic variables and a cross-lagged format for the relationship between the measures. The cross-lagged format uses all time points within a measure and the current time point for all other measures. This means that if we are imputing IES at the 6-month timepoint, the 6-month value for DES, DRS, SCL, and TAS was used, and all time points but the 6-month timepoint for IES under the assumption that the values not used are represented by the single timepoint (Van Buuren 2018).

For the covariates in the multiply imputed mixed model we used the same covariates selected for the available case mixed model, for ease of comparison between models.

Model Selection

For both the non-imputed models and the multiple imputed models, we utilized the SAS macro ALLMIXED4 written by Fernandez (2007). This macro follows a set of steps to find

the best linear mixed model for a given input. First, we investigated correlation structures for the data with the only fixed effects being follow-up timepoint, dropout timepoint, and an interaction between dropout timepoint and follow-up timepoint. Then, using the best correlation structure identified, we investigated every possible combination of fixed effects based on Akaike's information criterion (AIC) and minimal description length (MDL). We chose the model based on AIC, MDL, and model parsimony Then, with the selected fixed effects we reran the correlation structure investigation to confirm the steps above.

We tested five different correlation structures for this data: First order Auto-Regressive, Compound Symmetric, Spatial Power, Exponential Spatial, and Unstructured. For the spatial time structures we tested two spatial formats. One was the spatial correlation structure with time as days. Consider a subject who had 7 treatment days and no missing values. Their covariance matrix R would be

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi^7 & \phi^{14} & \phi^{37} & \phi^{187} \\ \phi^7 & 1 & \phi^7 & \phi^{30} & \phi^{180} \\ \phi^{14} & \phi^7 & 1 & \phi^{23} & \phi^{173} \\ \phi^{37} & \phi^{30} & \phi^{23} & 1 & \phi^{160} \\ \phi^{187} & \phi^{180} & \phi^{173} & \phi^{160} & 1 \end{bmatrix}.$$

If the subject was instead missing the 3 week follow up timepoint but had all of the other timepoints their covariance matrix R would be

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi^7 & \phi^{14} & \phi^{187} \\ \phi^7 & 1 & \phi^7 & \phi^{180} \\ \phi^{14} & \phi^7 & 1 & \phi^{173} \\ \phi^{187} & \phi^{180} & \phi^{173} & 1 \end{bmatrix}.$$

The AR(1) covariance structure would be

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi^1 & 1 & \phi^1 & \phi^2 \\ \phi^2 & \phi^1 & 1 & \phi^1 \\ \phi^3 & \phi^2 & \phi^1 & 1 \end{bmatrix},$$

since the data only contained timepoint records for subjects with scores recorded. Thereforew, we introduced what we call spatial with time as class which doesn't account for the space between timepoints, but does account for those with a missing timepoint. This yields a spatial with time as class/categorical covariance matrix for the subject with all timepoints but the 3-month follow up timepoint

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^4 \\ \phi^1 & 1 & \phi^1 & \phi^3 \\ \phi^2 & \phi^1 & 1 & \phi^2 \\ \phi^4 & \phi^3 & \phi^2 & 1 \end{bmatrix}.$$

For all five outcomes, the unstructured correlation structures failed to converge so they were not reported in this initial step. We then tested every possible covariate combination added to the base model. The possible covariates were primary diagnosis, gender, treatment days, age, and trauma type. Note that hallucination, Alexithymia Pre, and Alexithymia Post results are all based off of the measures so they were not included as possible covariates in this step. Outcomes for the correlation structure selection and the fixed effect selection can be seen in Tables 8 and 9, respectively. Note that for the fixed effect selection, we report those models which were within 5 points of the AIC of the lowest-scoring model.

Note that Follow $\vec{\text{Up}}$Timepoint$_{ij}$ represents the follow up timepoint for subject i at timepoint j and is a vector of all follow up timepoints such that

$$\text{Follow Up}\vec{\text{Timepoint}}_{ij} = \begin{bmatrix} \mathbf{1}_{\text{Post Treatment}}(\text{Follow Up Timepoint}_{ij}) \\ \mathbf{1}_{\text{1 Week}}(\text{Follow Up Timepoint}_{ij}) \\ \mathbf{1}_{\text{3 Months}}(\text{Follow Up Timepoint}_{ij}) \\ \mathbf{1}_{\text{6 Months}}(\text{Follow Up Timepoint}_{ij}) \end{bmatrix},$$

where $\mathbf{1}_{\text{1 Week}}(\text{Follow Up Timepoint}_{ij})$ is an indicator function for if Follow Up Timepoint for subject i at time j is the 1 week follow up timepoint. The $\vec{\beta}$ for this covariate is a 1×4 vector representing the estimated change in the outcome score for each follow-up timepoint when compared to baseline (baseline is the reference variable). Dropout $\vec{\text{T}}\text{imepoint}_i$ is a vector which represents the dropout group for subject i where

$$\text{Dropout }\vec{\text{T}}\text{imepoint}_i = \begin{bmatrix} \mathbf{1}_{\text{Post Treatment}}(\text{Dropout Timepoint}_i) \\ \mathbf{1}_{\text{1 Week}}(\text{Dropout Timepoint}_i) \\ \mathbf{1}_{\text{3 Months}}(\text{Dropout Timepoint}_i) \\ \mathbf{1}_{\text{6 Months}}(\text{Dropout Timepoint}_i) \end{bmatrix},$$

where $\mathbf{1}_{\text{1 Week}}(\text{Dropout Timepoint}_j)$ is an indicator function for if the dropout timepoint for subject i is the 1 week follow up timepoint. The $\vec{\beta}$ for this covariate is a 1×4 vector representing the estimated change in the outcome score for each dropout timepoint when compared to baseline dropout (baseline is the reference variable).

The operation $\text{Follow Up}\vec{\text{Timepoint}}_{ij}\text{Dropout }\vec{\text{T}}\text{imepoint}_{ij}^{'}$ is the outer product of the two vectors which creates a 4×4 matrix. $\text{vec}(\text{Follow Up}\vec{\text{Timepoint}}_{ij}\text{Dropout }\vec{\text{T}}\text{imepoint}_{ij}^{'})$ is the vectorization of this into a 16×1 vector and the associated $\vec{\beta}$ is a 1×16 vector representing the estimated change in outcome score for the follow up time point and the dropout timepoint.

IES

For IES, exponential spatial with time as days since post-treatment was the best-performing correlation structure by over 100 points. For the fixed effects, including primary diagnosis, treatment days, and age outperformed other models in both AIC and MDL. Including gender caused both an MDL and AIC improvement of approximately 0.5 points. We chose not to include gender to keep the model as simple as possible while still accounting for the demographic variables. The confirmatory correlation structure AICs and random effects AICs for IES are located in Table 10, showing that first-order autoregressive outperformed any spatial correlation structure and including a random slope for time in days did not improve the model. Our final model for IES for subject i at time j is

$$\begin{aligned} IES_{ij} =& \beta_0 + \vec{\beta}_1 \times \text{Follow Up } \vec{\text{T}}\text{imepoint}_{ij} + \vec{\beta}_2 \times \text{Dropout } \vec{\text{T}}\text{imepoint}_i \\ &+ \vec{\beta}_3 \times \text{vec}(\text{Follow Up Timepoint}_{ij}\text{Dropout Timepoint}_i') \\ &+ \vec{\beta}_4 \times \text{Primary } \vec{\text{D}}\text{iagnosis}_i + \beta_5 \times \text{Treatment Days}_i \\ &+ \beta_6 \times \text{Age}_i + b_{0i} + \epsilon_{ij}, \end{aligned}$$

where $b_{0i} \sim N(0, \sigma_b^2)$ is the random intercept for subject i and $\vec{\epsilon_i} = \left(\epsilon_{i1}, \ldots, \epsilon_{ir_i}\right)^t \sim N(0, R_i)$ where r_i is the last timepoint for subject i and

$$R_i = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \phi^4 \\ \phi & 1 & \phi & \phi^2 & \phi^3 \\ \phi^2 & \phi & 1 & \phi & \phi^2 \\ \phi^3 & \phi^2 & \phi & 1 & \phi \\ \phi^4 & \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

for some estimate of between-timepoint correlation ϕ.

SCL

For SCL, exponential spatial with time as days since post-treatment was the best-performing correlation structure by over 300 points. For the fixed effects, including primary diagnosis, treatment days, and age outperformed other models in MDL while including gender caused an AIC improvement of approximately 3 points. We chose not to include gender to keep the model as simple as possible while still accounting for the demographic variables. The confirmatory correlation structure AICs and random effects AICs for SCL are in Table 11 found that spatial power for time as class outperformed any other correlation structure and including a random slope for time in days did not improve the model. Our final model for SCL for subject i at time j is

$$\begin{aligned} SCL_{ij} =& \beta_0 + \vec{\beta}_1 \times \text{Follow Up } \vec{\text{T}}\text{imepoint}_{ij} + \vec{\beta}_2 \times \text{Dropout } \vec{\text{T}}\text{imepoint}_i \\ &+ \vec{\beta}_3 \times \text{vec}(\text{Follow Up Timepoint}_{ij}\text{Dropout Timepoint}_i') \\ &+ \vec{\beta}_4 \times \text{Primary } \vec{\text{D}}\text{iagnosis}_i + \beta_5 \times \text{Treatment Days}_i \\ &+ \beta_6 \times \text{Age}_i + b_{0i} + \epsilon_{ij}, \end{aligned}$$

where $b_{0i} \sim N(0, \sigma_b^2)$ is the random intercept for subject i and $\vec{\epsilon_i} = \left(\epsilon_{i1}, \ldots, \epsilon_{ir_i}\right)^t \sim N(0, R_i)$ where r_i is the last timepoint for subject i and

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi^{d_{1,2}} & \phi^{d_{1,3}} & \phi^{d_{1,4}} & \phi^{d_{1,5}} \\ \phi^{d_{2,1}} & 1 & \phi^{d_{2,3}} & \phi^{d_{2,4}} & \phi^{d_{2,5}} \\ \phi^{d_{3,1}} & \phi^{d_{3,2}} & 1 & \phi^{d_{3,4}} & \phi^{d_{3,5}} \\ \phi^{d_{4,1}} & \phi^{d_{4,2}} & \phi^{d_{4,3}} & 1 & \phi^{d_{4,5}} \\ \phi^{d_{5,1}} & \phi^{d_{5,2}} & \phi^{d_{5,3}} & \phi^{d_{5,4}} & 1 \end{bmatrix}$$

for some estimate of between-timepoint correlation ϕ and distance between timepoints k and j, d_{jk}.

TAS

For TAS, exponential spatial with time as days since post-treatment was the best-performing initial correlation structure by over 200 points. For the fixed effects, including treatment days and age outperformed other models in MDL and AIC. While including primary diagnosis was only a slight difference in AIC, we chose not to include it to keep the model as simple as possible while. The confirmatory correlation structure AICs and random effects AICs for TAS (Table 12) found that spatial power for time as class outperformed any other correlation structure and including a random slope for time in days did not improve the model. Our final model for TAS for subject i at time j is

$$\begin{aligned} TAS_{ij} =& \beta_0 + \vec{\beta}_1 \times \text{Follow Up } \vec{\text{T}}\text{imepoint}_{ij} + \vec{\beta}_2 \times \text{Dropout } \vec{\text{T}}\text{imepoint}_i \\ &+ \vec{\beta}_3 \times \text{vec}(\text{Follow Up Timepoint}_{ij}\text{Dropout Timepoint}_i') \\ &+ \beta_4 \times \text{Treatment Days}_i + \beta_5 \times \text{Age}_i + b_{0i} + \epsilon_{ij}, \end{aligned}$$

where $b_{0i} \sim N(0, \sigma_b^2)$ is the random intercept for subject i and $\vec{\epsilon_i} = \left(\epsilon_{i1}, \ldots, \epsilon_{ir_i}\right)^t \sim N(0, R_i)$ where r_i is the last timepoint for subject i and

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi^{d_{1,2}} & \phi^{d_{1,3}} & \phi^{d_{1,4}} & \phi^{d_{1,5}} \\ \phi^{d_{2,1}} & 1 & \phi^{d_{2,3}} & \phi^{d_{2,4}} & \phi^{d_{2,5}} \\ \phi^{d_{3,1}} & \phi^{d_{3,2}} & 1 & \phi^{d_{3,4}} & \phi^{d_{3,5}} \\ \phi^{d_{4,1}} & \phi^{d_{4,2}} & \phi^{d_{4,3}} & 1 & \phi^{d_{4,5}} \\ \phi^{d_{5,1}} & \phi^{d_{5,2}} & \phi^{d_{5,3}} & \phi^{d_{5,4}} & 1 \end{bmatrix}$$

for some estimate of between-timepoint correlation ϕ and distance between timepoints k and j, d_{jk}.

log(DES)

For DES, the score was naturally logged for all models to deal with the skewness as seen in Figure 6. Initial correlation tests found that exponential spatial with time as days

since post-treatment was the best-performing correlation structure by over 50 points. For the fixed effects, including primary diagnosis, treatment days, and age outperformed other models in MDL and AIC. While including gender caused created a comparable AIC, we chose not to include gender to keep the model as simple as possible. The confirmatory correlation structure AICs and random effects AICs for DES (Table 13) found that spatial power for time as class outperformed any other correlation structure and including a random slope for time in days created comparable fit statistics. We chose to include a random slope for days to account for more variance within subjects over time. Our final model for DES for subject i at time j is

$$\begin{aligned}\log DES_{ij} =&\beta_0 + \vec{\beta}_1 \times \text{Follow Up } \vec{\text{T}}\text{imepoint}_{ij} + \vec{\beta}_2 \times \text{Dropout } \vec{\text{T}}\text{imepoint}_i \\ &+ \vec{\beta}_3 \times \text{vec}(\text{Follow Up Timepoint}_{ij}\text{Dropout Timepoint}_i') \\ &+ \vec{\beta}_4 \times \text{Primary } \vec{\text{D}}\text{iagnosis}_i + \beta_5 \times \text{Treatment Days}_i \\ &+ \beta_6 \times \text{Age}_i + b_{0i} + b_{1i} \times \text{Follow Up Days}_{ij} + \epsilon_{i,j},\end{aligned}$$

where b_{0i} is the random intercept for subject i and b_{1i} is the random slope for days since post-treatment for subject i. $\vec{b_i} \sim N(0, G_i)$ is a 2×1 vector for the random intercept and random slope where G_i is the 2×2 variance-covariance matrix with 0-covariance for random intercept and random slope. $\vec{\epsilon_i} = \left(\epsilon_{i1}, \ldots, \epsilon_{ir_i}\right)^t \sim N(0, R_i)$ where r_i is the last timepoint for subject i and

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi^{d_{1,2}} & \phi^{d_{1,3}} & \phi^{d_{1,4}} & \phi^{d_{1,5}} \\ \phi^{d_{2,1}} & 1 & \phi^{d_{2,3}} & \phi^{d_{2,4}} & \phi^{d_{2,5}} \\ \phi^{d_{3,1}} & \phi^{d_{3,2}} & 1 & \phi^{d_{3,4}} & \phi^{d_{3,5}} \\ \phi^{d_{4,1}} & \phi^{d_{4,2}} & \phi^{d_{4,3}} & 1 & \phi^{d_{4,5}} \\ \phi^{d_{5,1}} & \phi^{d_{5,2}} & \phi^{d_{5,3}} & \phi^{d_{5,4}} & 1 \end{bmatrix}$$

for some estimate of between-timepoint correlation ϕ and distance between timepoints k and j, d_{jk}.

log(DRS)

For DRS, the score was natural logged for all models to deal with the skewness as seen in Figure 6. Initial correlation tests found that exponential spatial with time as days since post-treatment was the best-performing correlation structure by over 100 points. For the fixed effects, including primary diagnosis, treatment days, and age outperformed other models in MDL and AIC. While including gender caused created a comparable AIC, we chose not to include gender to keep the model as simple as possible. The confirmatory correlation structure AICs and random effects AICs for DRS (Table 14) found that spatial power for time as class outperformed any other correlation structure and including a random slope created comparable fit statistics. We chose to include a random slope for days to account for more variance within subjects over time. Our final model for DRS for subject i at time j is

$$\begin{aligned}\log DRS_{ij} =&\beta_0 + \vec{\beta}_1 \times \text{Follow Up } \vec{\text{T}}\text{imepoint}_{ij} + \vec{\beta}_2 \times \text{Dropout } \vec{\text{T}}\text{imepoint}_i \\ &+ \vec{\beta}_3 \times \text{vec}(\text{Follow Up Timepoint}_{ij}\text{Dropout Timepoint}_i') \\ &+ \vec{\beta}_4 \times \text{Primary } \vec{\text{D}}\text{iagnosis}_i + \beta_5 \times \text{Treatment Days}_i \\ &+ \beta_6 \times \text{Age}_i + b_{0i} + b_{1i} \times \text{Follow Up Days}_{ij} + \epsilon_{i,j},\end{aligned}$$

where b_{0i} is the random intercept for subject i and b_{1i} is the random slope for days since post-treatment for subject i. $\vec{b_i} \sim N(0, G_i)$ is a 2×1 vector for the random intercept and random slope where G_i is the 2×2 variance-covariance matrix with 0-covariance for random intercept and random slope. $\vec{\epsilon_i} = \left(\epsilon_{i1}, ..., \epsilon_{ir_i}\right)^t \sim N(0, R_i)$ where r_i is the last timepoint for subject i and

$$R = \sigma_\epsilon^2 \begin{bmatrix} 1 & \phi^{d_{1,2}} & \phi^{d_{1,3}} & \phi^{d_{1,4}} & \phi^{d_{1,5}} \\ \phi^{d_{2,1}} & 1 & \phi^{d_{2,3}} & \phi^{d_{2,4}} & \phi^{d_{2,5}} \\ \phi^{d_{3,1}} & \phi^{d_{3,2}} & 1 & \phi^{d_{3,4}} & \phi^{d_{3,5}} \\ \phi^{d_{4,1}} & \phi^{d_{4,2}} & \phi^{d_{4,3}} & 1 & \phi^{d_{4,5}} \\ \phi^{d_{5,1}} & \phi^{d_{5,2}} & \phi^{d_{5,3}} & \phi^{d_{5,4}} & 1 \end{bmatrix}$$

for some estimate of between-timepoint correlation ϕ and distance between timepoints k and j, d_{jk}.

Table 8: Initial Correlation Structure Selection Output

Correlation Structure	AICC
IES	
Exponential Spatial Time as Days	8959.35
AR(1)	9198.67
Spatial Power Follow-up Timepoint	9199.61
Exponential Spatial Follow-Up Timepoint	9218.81
Compound Symmetric	9220.82
log(DES)	
Exponential Spatial Time as Days	2376.34
AR(1)	2452.89
Spatial Power Follow-up Timepoint	2454.23
Exponential Spatial Follow-Up Timepoint	2466.99
Compound Symmetric	2469.00
log(DRS)	
Exponential Spatial Time as Days	3161.74
AR(1)	3272.85
Spatial Power Follow-up Timepoint	3274.71
Exponential Spatial Follow-Up Timepoint	3281.50
Compound Symmetric	3283.51
SCL	
Exponential Spatial Time as Days	10683.44
AR(1)	11065.77
Spatial Power Follow-up Timepoint	11066.90
Exponential Spatial Follow-Up Timepoint	11077.36
Compound Symmetric	11079.37
TAS	
Exponential Spatial Time as Days	8478.80
AR(1)	8768.34
Spatial Power Follow-up Timepoint	8768.40
Exponential Spatial Follow-Up Timepoint	8786.54
Compound Symmetric	8788.55

Table 9: Fixed Effect Model Selection Output

Fixed Effects	AIC_C	MDL
IES		
Primary Diagnosis + Treatment Days + Age	8496.595	4292.231
Treatment Days + Age	8513.462	4293.246
Primary Diagnosis + Gender + Treatment Days + Age	8498.405	4294.975
Gender + Treatment Days + Age	8514.702	4295.730
SCL		
Primary Diagnosis + Treatment Days + Age	9946.108	5015.142
Primary Diagnosis + Gender + Treatment Days + Age	9943.053	5015.460
Primary Diagnosis + Gender + Trauma Type + Treatment Days + Age	9944.595	5021.729
TAS		
Treatment Days + Age	7938.522	4005.775
Gender + Treatment Days + Age	7940.723	4008.740
Primary Diagnosis + Treatment Days + Age	7938.620	4013.243
Trauma Type + Treatment Days + Age	7942.896	4013.536
Primary Diagnosis + Gender + Treatment Days + Age	7940.834	4016.189
log(DES)		
Primary Diagnosis + Treatment Days + Age	2067.570	1075.873
Primary Diagnosis + Gender + Treatment Days + Age	2068.791	1078.329
Primary Diagnosis + Trauma Type + Treatment Days + Age	2069.122	1082.166
Primary Diagnosis + Gender + Trauma Type + Treatment Days + Age	2070.558	1084.710
log(DRS)		
Primary Diagnosis + Treatment Days + Age	2864.010	1474.093
Primary Diagnosis + Gender + Treatment Days + Age	2866.091	1476.979
Primary Diagnosis + Trauma Type + Treatment Days + Age	2866.942	1481.076

Table 10: Correlation and Random Effect Selection for IES

Correlation Structure	AIC	AICC	BIC
Random Intercept			
First-Order Autoregressive for Time as Class	7736.2	7737.4	7823.6
Spatial Power for Time as Class	7739.4	7740.6	7826.9
Spatial Power for Time as Days	7751.6	7752.9	7839.1
Exponential Spatial for Time as Days	7755.9	7757.0	7839.5
Exponential Spatial for Time as Class	7755.9	7757.0	7839.5
Random Intercept Random Slope for Time in Days			
First-Order Autoregressive for Time as Class	7738.2	7739.5	7829.4
Spatial Power for Time as Class	7741.4	7742.7	7832.6
Spatial Power for Time as Days	7751.6	7752.9	7839.1
Exponential Spatial for Time as Days	7755.9	7757.0	7839.5
Exponential Spatial for Time as Class	7755.9	7757.0	7839.5

Table 11: Correlation and Random Effect Selection for SCL

Correlation Structure	AIC	AICC	BIC
Random Intercept			
Spatial Power for Time as Class	8890.8	8892.2	8982.1
First-Order Autoregressive for Time as Class	8899.1	8900.4	8990.3
Exponential Spatial for Time as Class	8902.7	8903.9	8990.2
Spatial Power for Time as Days	8904.2	8905.6	8995.5
Exponential Spatial for Time as Days	8904.2	8905.6	8995.5
Random Intercept Random Slope for Time in Days			
Spatial Power for Time as Class	8892.3	8893.7	8987.4
First-Order Autoregressive for Time as Class	8900.7	8902.1	8995.7
Exponential Spatial for Time as Class	8904.7	8906.0	8996.0
Spatial Power for Time as Days	8906.2	8907.7	9001.3
Exponential Spatial for Time as Days	8906.2	8907.7	9001.3

Table 12: Correlation and Random Effect Selection for TAS

Correlation Structure	AIC	AICC	BIC
Random Intercept			
Spatial Power for Time as Class	7058.9	7059.8	7131.2
Spatial Power for Time as Days	7065.0	7065.9	7137.3
First-Order Autoregressive for Time as Class	7067.3	7068.1	7139.5
Exponential Spatial for Time as Class	7069.8	7070.6	7138.2
Exponential Spatial for Time as Days	7071.8	7072.6	7144.0
Random Intercept Random Slope for Time in Days			
Spatial Power for Time as Class	7060.7	7061.6	7136.8
Spatial Power for Time as Days	7065.0	7065.9	7137.3
First-Order Autoregressive for Time as Class	7069.2	7070.2	7145.3
Exponential Spatial for Time as Class	7069.8	7070.6	7138.2
Exponential Spatial for Time as Days	7071.8	7072.6	7144.0

Table 13: Correlation and Random Effect Selection for log(DES)

Correlation Structure	AIC	AICC	BIC
Random Intercept			
Spatial Power for Time as Class	1821.9	1823.3	1913.2
First-Order Autoregressive for Time as Class	1826.7	1828.0	1917.9
Exponential Spatial for Time as Class	1829.6	1830.8	1917.0
Spatial Power for Time as Days	1830.9	1832.2	1922.1
Exponential Spatial for Time as Days	1830.9	1832.2	1922.1
Random Intercept Random Slope for Time in Days			
Spatial Power for Time as Class	1821.9	1823.4	1917.0
First-Order Autoregressive for Time as Class	1827.0	1828.4	1922.0
Exponential Spatial for Time as Class	1830.6	1832.0	1921.9
Spatial Power for Time as Days	1831.9	1833.4	1927.0
Exponential Spatial for Time as Days	1831.9	1833.4	1927.0

Table 14: Correlation and Random Effect Selection for log(DRS)

Correlation Structure	AIC	AICC	BIC
Random Intercept			
Spatial Power for Time as Class	2532.0	2533.2	2619.4
First-Order Autoregressive for Time as Class	2535.5	2536.8	2623.0
Exponential Spatial for Time as Class	2537.5	2538.6	2621.2
Spatial Power for Time as Days	2539.5	2540.7	2626.9
Exponential Spatial for Time as Days	2539.5	2540.7	2626.9
Random Intercept Random Slope for Time in Days			
Spatial Power for Time as Class	2532.0	2533.2	2619.4
First-Order Autoregressive for Time as Class	2535.5	2536.8	2623.0
Exponential Spatial for Time as Class	2537.5	2538.6	2621.2
Spatial Power for Time as Days	2539.5	2540.7	2626.9
Exponential Spatial for Time as Days	2539.5	2540.7	2626.9

CHAPTER III

RESULTS

Our primary goal in this analysis was to identify if we see a significant decrease in measures over the follow-up period with a secondary and tertiary goal to quantify the effect of dropout and determine measures change over time. The below results present our findings.

Measure Change Over Time

For all measures, we found a significant decrease in score over time. DES's and DRS's changes are reported as a percentage change while the rest of the measures are presented as a point change. All of the measures' expected changes over time by model can be found in Table 15.

DES

For DES, we found an estimated decrease between 30-50%. The available case analysis found a significant estimated decrease in DES from Baseline to 6 Months of 40% (95% CI: 30%, 49%; $p < 0.001$). The multiple imputation found that that was an underestimate with a significant estimated decrease in DES from Baseline to 6 Months of 46% (95% CI: 41%, 50%; $p < 0.001$). This indicates that the multiple imputation found that the subjects that

Table 15: Estimated change from Baseline to 6 Months For all Measures by Model

Model	Estimate	95% Confidence Interval	P-Value
DES			
Available Case	-40%	(-49%, -30%)	p<.0001
Multiple Imputation	-46%	(-50%, -41%)	p<.0001
DRS			
Available Case	-48%	(-58%, -35%)	p<.0001
Multiple Imputation	-54%	(-58%, -48%)	p<.0001
IES			
Available Case	-17.33	(-20.91, -13.75)	p<.0001
Available Case	-0.65	(-0.87, -0.43)	p<.0001
Multiple Imputation	-17.97	(-19.8, -16.15)	p<.0001
SCL			
Available Case	-29.12	(-35.85, -22.38)	p<.0001
Multiple Imputation	-30.55	(-33.64, -27.46)	p<.0001
TAS			
Available Case	-7.3	(-9.97, -4.62)	p<.0001
Available Case	-7.3	(-9.97, -4.62)	p<.0001
Multiple Imputation	-7.14	(-8.24, -6.04)	p<.0001

dropped out were likely to have better DES outcomes.

DRS

For DRS, we found an estimated decrease between 35-58%. The available case analysis found a significant estimated decrease in DRS from Baseline to 6 Months of 48% (95% CI: 35%, 58%; $p < 0.001$). The multiple imputation found that that was an underestimate with a significant estimated decrease in DRS from Baseline to 6 Months of 54% (95% CI: 48%, 58%; $p < 0.001$). This indicates that the multiple imputation found that the subjects that dropped out were likely to have better DRS outcomes.

IES

For IES, we found an estimated decrease between 13-21 points. The available case analysis found a significant estimated decrease in IES from Baseline to 6 Months of 17.33 points (95% CI: 13.75, 20.91; $p < 0.001$). The multiple imputations found approximately the same with a decrease in IES from Baseline to 6 Months of 17.97 points (95% CI: 16.16, 19.8; $p < 0.001$). This indicates that the multiple imputation found that the subjects that dropped out have approximately the same IES change as subjects who did not drop out.

SCL

For SCL, we found an estimated decrease between 22-34 points. The available case analysis found a significant estimated decrease in SCL from Baseline to 6 Months of 29.12 points (95% CI: 22.38, 35.85; $p < 0.001$). The multiple imputation found approximately the same with a decrease in SCL from Baseline to 6 Months of 30.55 points (95% CI: 27.46, 33.64; $p < 0.001$). This indicates that the multiple imputation found that the subjects that dropped out have approximately the same SCL change as subjects who did not drop out.

TAS

For TAS, we found an estimated decrease between 4-10 points. The available case analysis found a significant estimated decrease in TAS from Baseline to 6 Months of 7.3 points (95% CI: 4.62, 9.97; $p < 0.001$). The multiple imputation found approximately the same with a decrease in TAS from Baseline to 6 Months of 7.14 points (95% CI: 6.04, 8.24; $p < 0.001$). This indicates that the multiple imputation found that the subjects that dropped out have approximately the same TAS change as subjects who did not drop out.

Impact of Dropout on Measures

Once we reached conclusions about how the measures performed over time, we performed hypothesis tests to investigate the effect of dropout timepoint on the available case models for each measure. For each model, we performed two F-Tests asking the questions '**Does baseline measure differ across dropout time points?**' and '**Does total change differ across time points?**'. Total change across dropout timepoints for some measure M at dropout time $d = \text{Dropout Timepoint}_i$ for subject i is defined as

$$\begin{aligned} M_{id} - M_{i1} = &\beta_{ad} \times \text{Follow Up Timepoint}_{id} - \beta_{a1} \times \text{Follow Up Timepoint}_{i1} \\ &+ \beta_{bd} \times \text{Follow Up Timepoint}_{id} \times \text{Dropout Timepoint}_i \\ &- \beta_{b1} \times \text{Follow Up Timepoint}_1 \times \text{Dropout Timepoint}_i \end{aligned}$$

where β_{ad} is the estimated Beta coefficient for the last follow-up timepoint, β_{a1} is the estimated Beta coefficient for the baseline follow-up timepoint, β_{bd} is the estimated Beta coefficient for the interaction between the last follow up timepoint and the dropout timepoint, and β_{b1} is the estimated Beta coefficient for the interaction between the baseline follow up timepoint and the dropout timepoint.

Table 16: Contrasts For IES in the Available Case Model

Hypothesis	P-Value
Does the Baseline Measure Differ Across Dropout Timepoints?	0.0280
Does the Total Change Differ Across Dropout Timepoints?	0.2141

IES

For IES, the contrasts are shown in table 16. Total change did not differ across dropout timepoint with the F-Test yielding a P-Value $p = 0.2141$. Baseline measure was significantly different across dropout time points with $p = 0.0280$. We investigated the baseline scores further as shown in Figure 8 with point estimates and error bars for the 95% confidence intervals. Subjects who dropped out at 3 Months and Post Treatment did not have a significant difference in baseline scores compared to those who dropped out at baseline (p=0.997 and p=0.1144, respectively). Those who dropped out at 1 week have borderline significantly lower baseline scores (p=0.0695). Note that the 95% confidence interval does contain zero. Those who drop out at Six Months have significantly lower baseline scores than those who drop out at baseline (p=0.0183). Overall there seem to be slight differences in baseline scores between dropout timepoint, but the total change of IES over time does not differ between dropout groups.

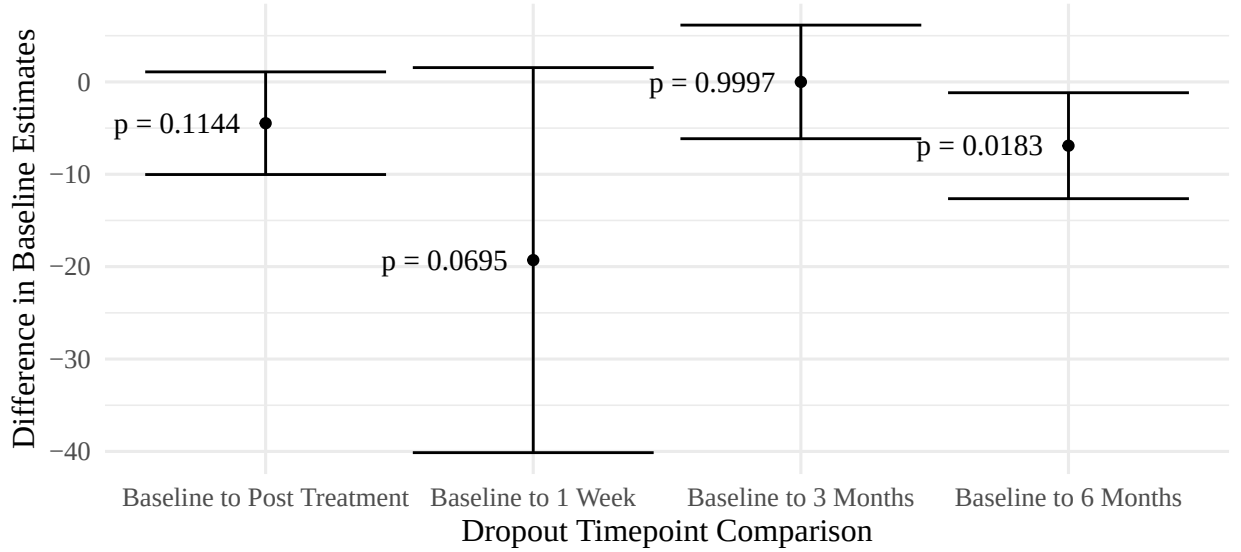

Figure 8: Difference in Baseline Estimates between Baseline Dropout and All other Dropout Timepoints

TAS

For TAS, the contrasts are shown in table 17. Baseline TAS did not differ between dropout time points with the F-Test yielding a P-Value of $p = 0.2679$ Total change was significantly different across dropout time points with $p = 0.0151$. We investigated the difference in total change amount between dropout time points (Figure 9) with point estimates and error bars for the 95% confidence intervals. Those who dropped out at 1 Week did not have a significantly different total change in TAS than those who dropped out at post-treatment with $p = 0.3722$. Subjects who dropped out at 3 months and 6 months have a total change that is significantly higher than those who drop out at Post-treatment with $p = 0.041$ and $p = 0.0021$, respectively. Overall, while baseline scores are approximately equal between dropout groups, those who drop out at 3 months and 6 months have a significantly greater decrease in their TAS scores than those who drop out at 1 week or Post-Treatment.

Table 17: Contrasts For TAS in the Available Case Model

Hypothesis	P-Value
Does the Baseline Measure Differ Across Dropout Timepoints?	0.2679
Does the Total Change Differ Across Dropout Timepoints?	0.0151

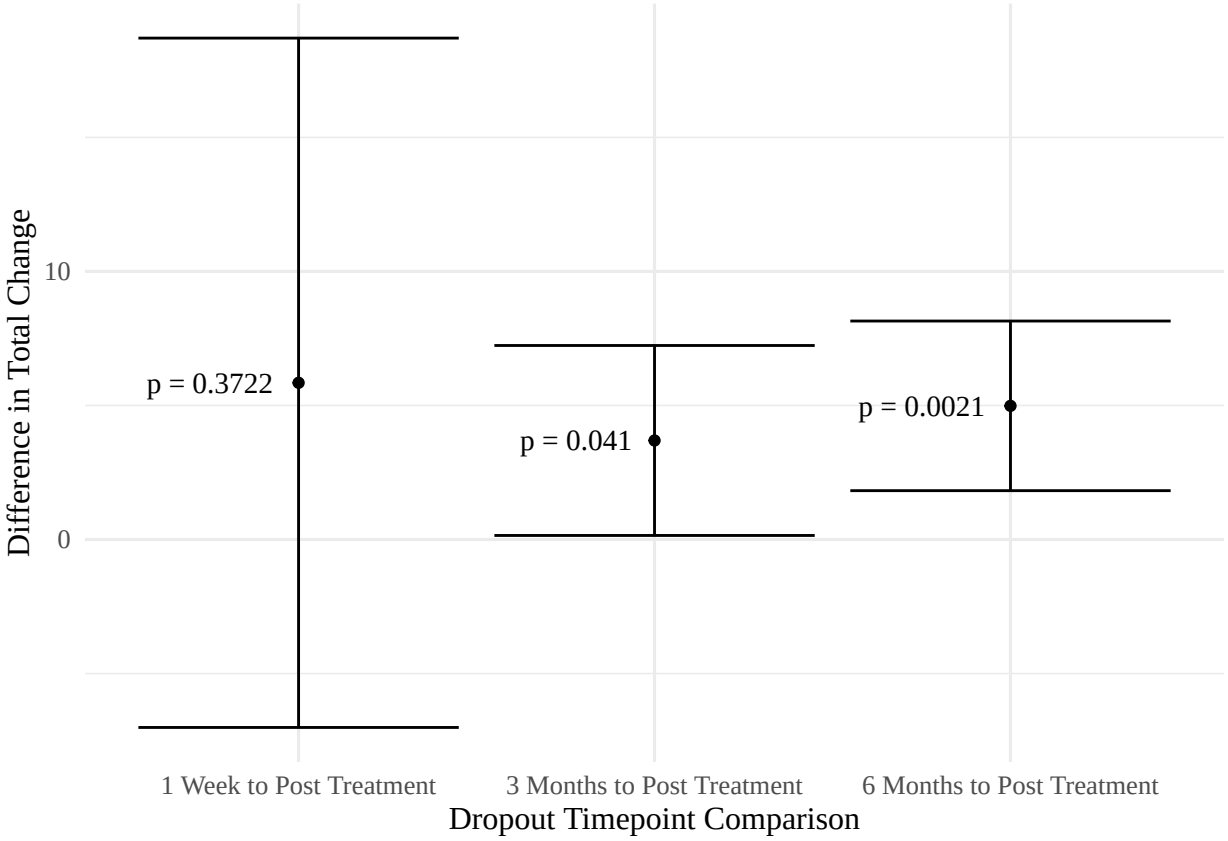

Figure 9: Total Change Across Dropout Timepoints for TAS

SCL

For SCL, the contrasts are shown in table 18. Baseline SCL did not differ between dropout timepoints with the F-Test yielding a P-Value of $p = 0.3652$. Total change also did not differ across dropout time points with $p = 0.5903$. Based on these F-Tests we did not investigate further. Overall, SCL does not appear to be affected by dropout timepoint.

Table 18: Contrasts For SCL in the Available Case Model

Hypothesis	P-Value
Does the Baseline Measure Differ Across Dropout Timepoints?	0.3652
Does the Total Change Differ Across Dropout Timepoints?	0.5903

DES

For DES, the contrasts are shown in table 19. Baseline DES did not differ between dropout timepoints with the F-Test yielding a P-Value of $p = 0.6317$. Total change also did not differ across dropout time points with $p = 0.1815$. Based on these F-Tests we did not investigate further. Overall, DES does not appear to be affected by dropout timepoint.

DRS

For DRS, the contrasts are shown in table 20. Baseline DRS did not differ between dropout timepoints with the F-Test yielding a P-Value of $p = 0.4371$. Total change also did not differ across dropout time points with $p = 0.2624$. Based on these F-Tests we did not investigate further. Overall, DES does not appear to be affected by dropout timepoint.

Change Over Time

After the dropout timepoint and overall change were analyzed, we continued to investigate how the measures changed over time. We utilized the SAS `LSMEANS` function to get

Table 19: Contrasts For DES in the Available Case Model

Hypothesis	P-Value
Does the Baseline Measure Differ Across Dropout Timepoints?	0.6371
Does the Total Change Differ Across Dropout Timepoints?	0.1815

Table 20: Contrasts For DRS in the Available Case Model

Hypothesis	P-Value
Does the Baseline Measure Differ Across Dropout Timepoints?	0.4371
Does the Total Change Differ Across Dropout Timepoints?	0.2624

averages of scores at each time point. We also used the SAS `MIANALYZE` to pool estimates from the multiple imputation model. Investigating the means plots shown in Figure 10, while we do see a large change from baseline to post-treatment, there does not appear to be a great change after post-treatment with subjects leveling out. We will investigate more thoroughly within each outcome.

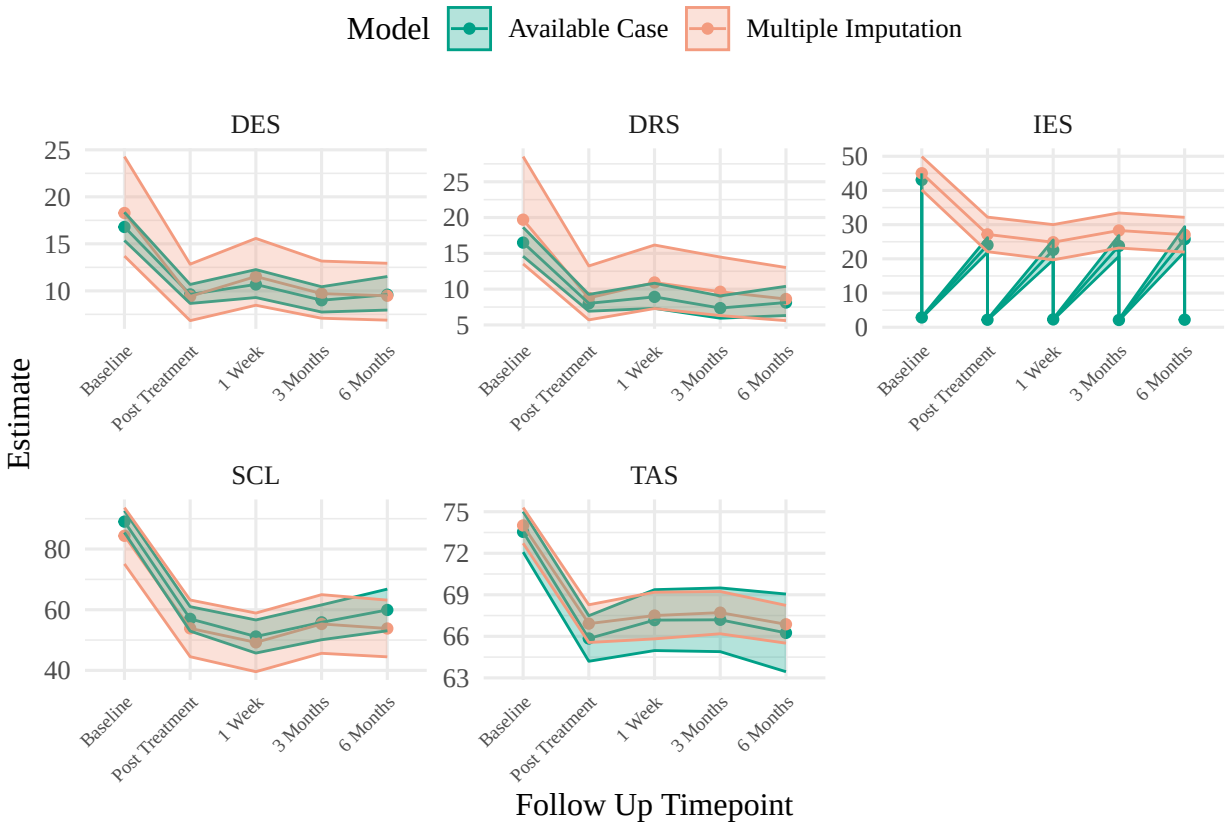

Figure 10: Estimated Values of Measures Across Timepoints

IES

The estimated change between all time points for IES is presented in Figure 11. As seen previously, there is a significant decrease from baseline to all four other follow-up time points in both the multiply imputed and available case models. There is no significant change in IES from Post-Treatment to any of the following time points. However, the multiply imputed model displayed borderline significance ($p = 0.0836$). From one week the multiply imputed model found a significant increase in IES to 3 months ($p = 0.0201$), and both the multiply imputed and available case found a borderline significant increase from 1 week to 6 months with $p = 0.0841$ and $p = 0.0887$, respectively. Neither the multiply imputed model nor the available case model found a significant change from 3 months to 6 months. Overall, we see that for IES there is a slight continuation in the decrease of IES from post-treatment to 1 week, and then a slight increase back to the post-treatment score from 1 week to 6 months. There is no overall significant change from post-treatment to 6 months.

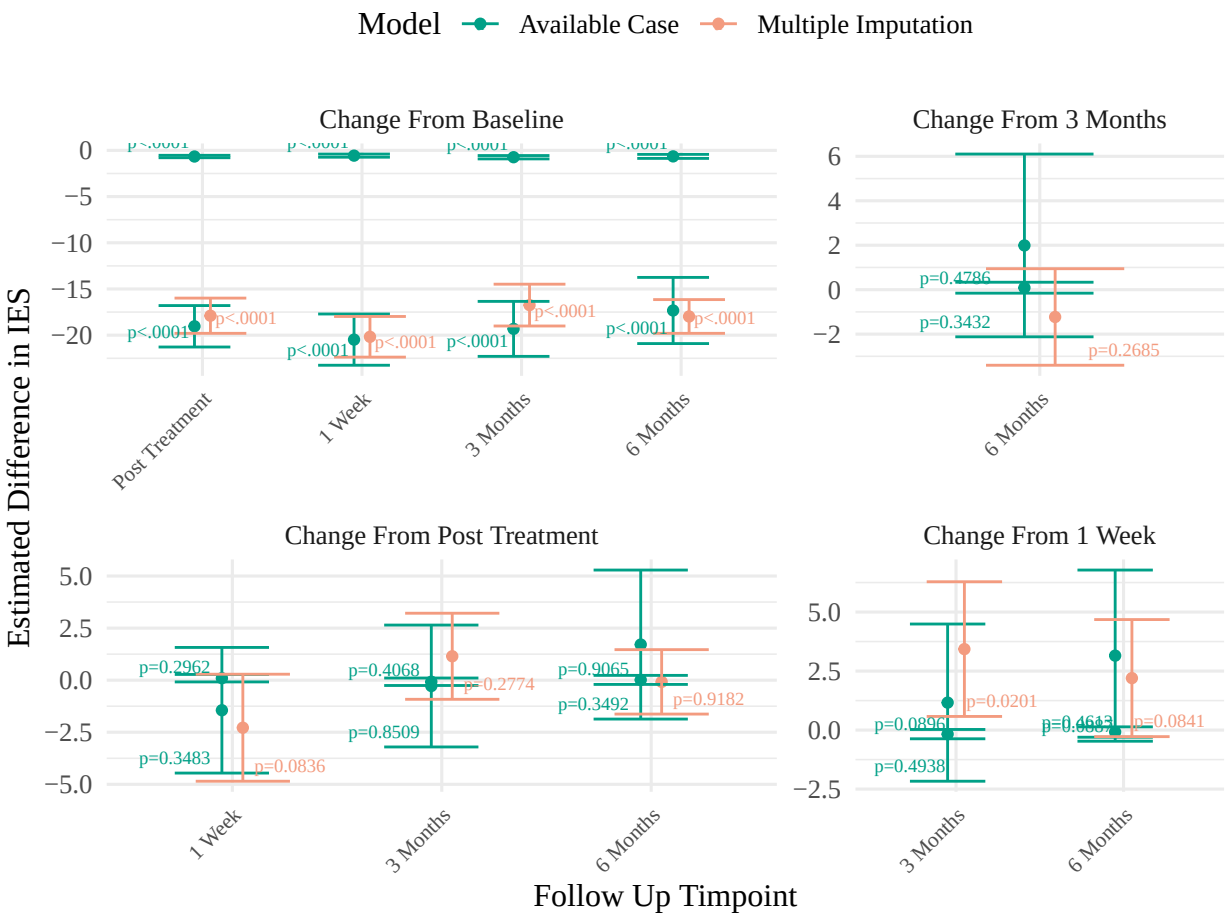

Figure 11: Estimated Change Between Timepoints for IES

TAS

The estimated change between all time points for TAS is presented in Figure 12. As seen previously, there is a significant decrease from baseline to all four other follow-up time points in both the multiply imputed and available case models. There is no significant or borderline significant change in TAS between time points after baseline.

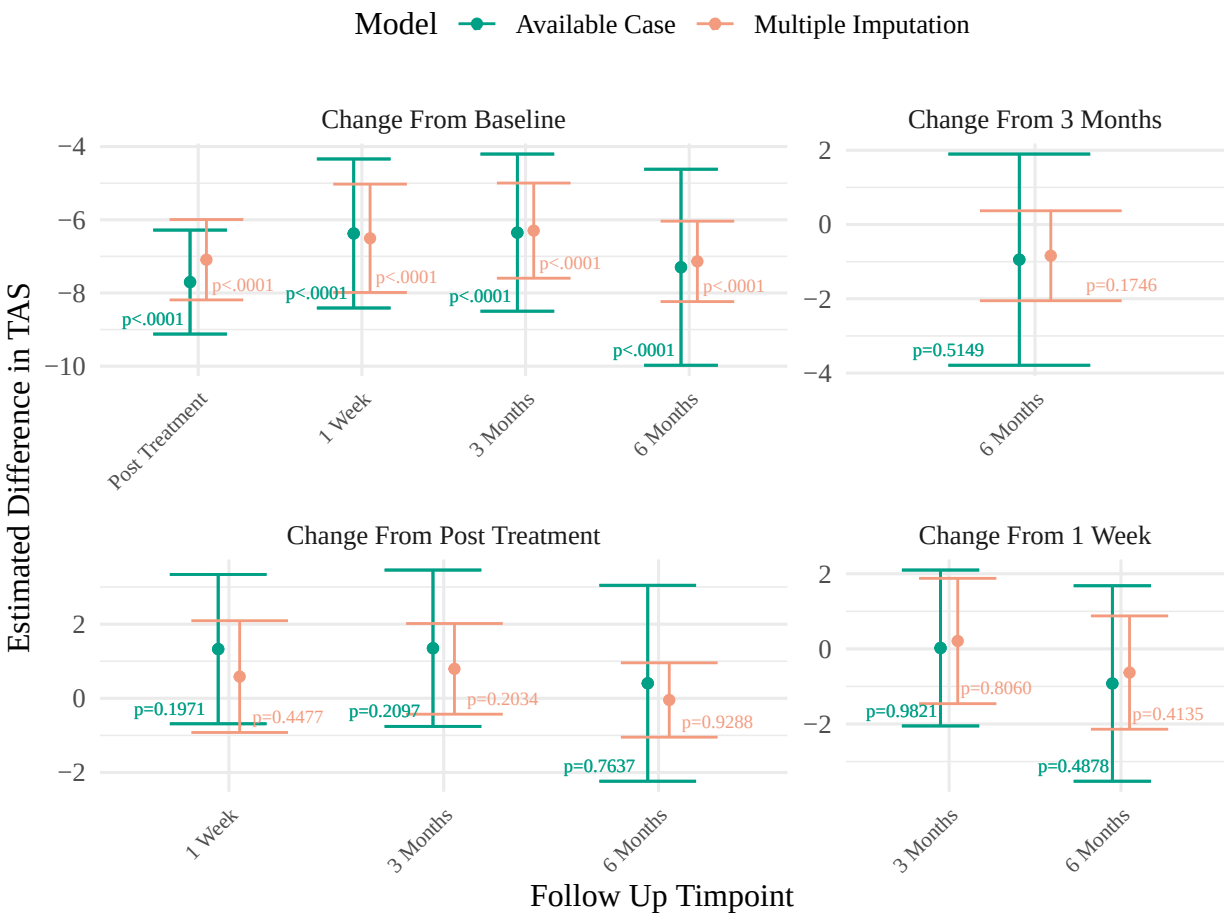

Figure 12: Estimated Change Between Timepoints for TAS

SCL

The estimated change between all time points for SCL is presented in Figure 13. As seen previously, there is a significant decrease from baseline to all four other follow-up time points in both the multiply imputed and available case models. SCL shows a significant decrease in score from Post Treatment to 1 Week in both the multiply imputed model and the available case model with $p = 0.0226$ and $p = 0.0416$, respectively. There is no significant change from Post Treatment to 3 months or 6 months. From one week the multiply imputed model found a significant increase in IES to 3 months ($p = 0.0039$), and both the multiply imputed and available case found a borderline significant increase from 1 week to 6 months

with $p = 0.0188$ and $p = 0.0108$, respectively. Neither the multiply imputed model nor the available case model found a significant change from 3 months to 6 months. Overall we see that for SCL there is a slight continuation in the decrease of SCL from post-treatment to 1 week, and then a slight increase back to the post-treatment score from 1 week to 6 months. There is no overall significant change from post-treatment to 6 months.

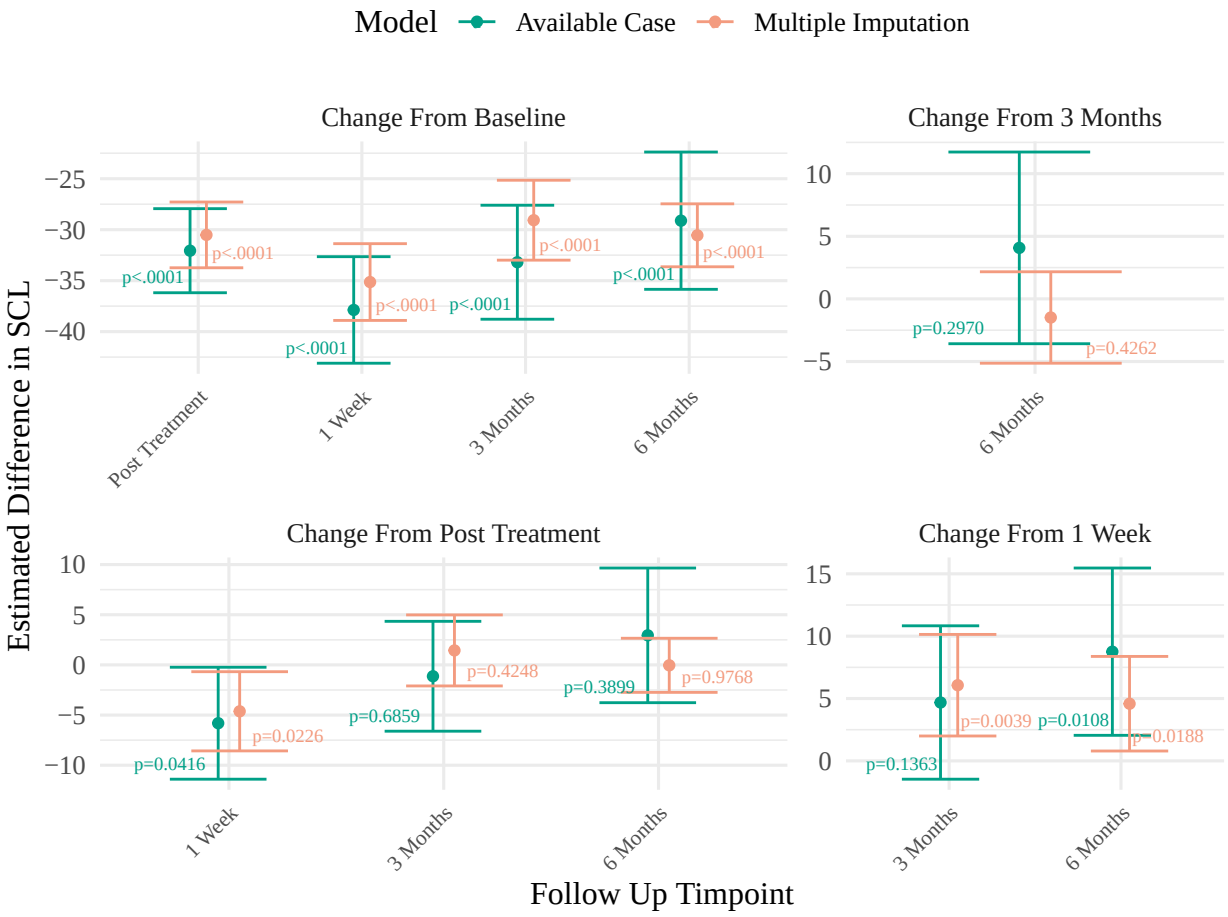

Figure 13: Estimated Change Between Timepoints for SCL

DES

The estimated change between all time points for DES is presented in Figure 14. As seen previously, there is a significant decrease from baseline to all four other follow-up time-

points in both the multiply imputed and available case models. There is no significant change in IES from Post-Treatment to any of the following time points for the available case model. The multiply imputed model has a significant increase in score from post-treatment to 1 week ($p = 0.0001$) with no significant change from post-treatment to 3 months or 6 months. Both the multiply imputed model and the available case model found a significant decrease in DES from 1 week to 3 months ($p = 0.0001$ and $p = 0.0181$, respectively). The available case model found no significant change from 1 week to 6 months while the multiply imputed model found a significant decrease in DES from 1 week to 6 months ($p < 0.001$). Neither the multiply imputed model nor the available case model found a significant change from 3 months to 6 months. Overall, we see that for DES there is a slight increase of DES from post-treatment to 1 week, and then a slight decrease back to the post-treatment score from 1 week to 6 months. There is no overall significant change from post-treatment to 6 months.

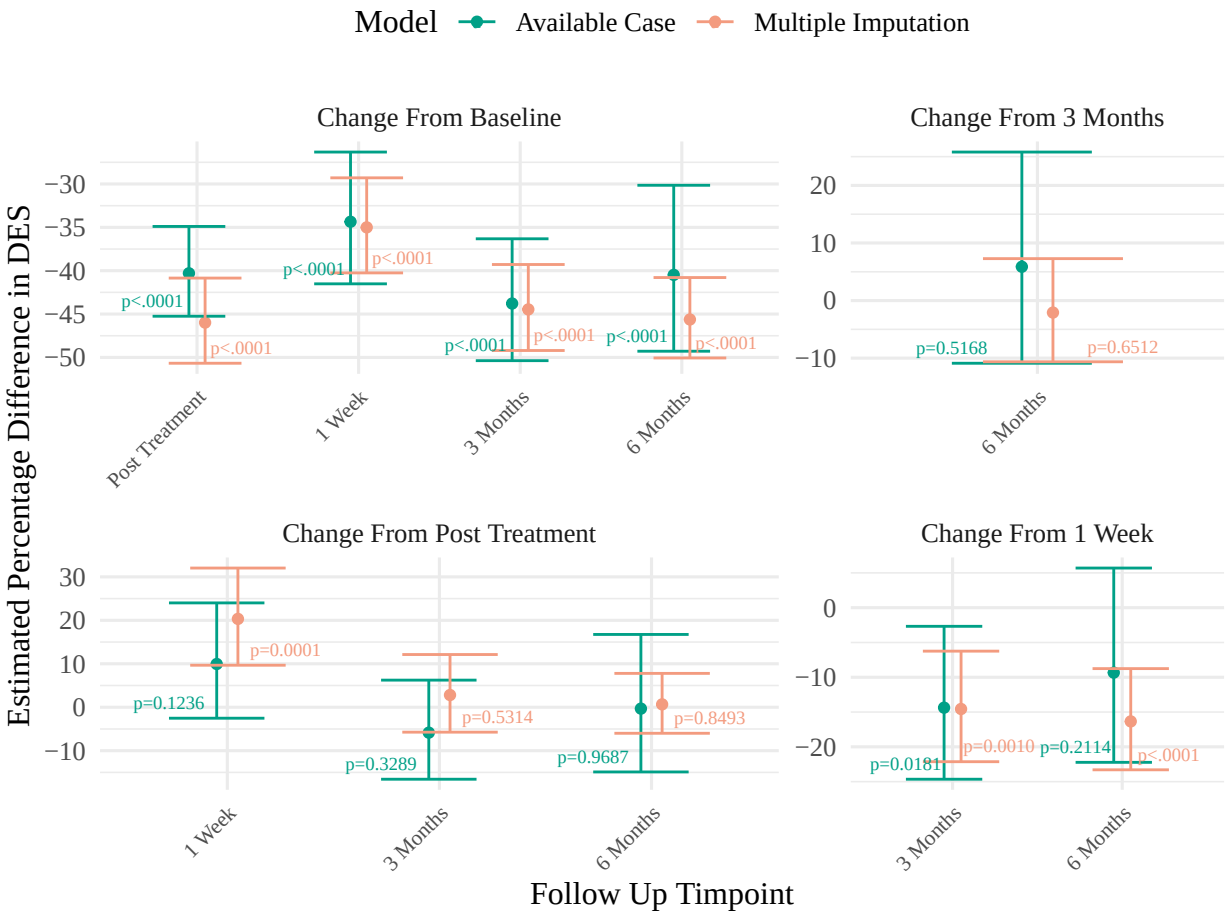

Figure 14: Estimated Change Between Timepoints for DES

DRS

The estimated change between all time points for DRS is presented in Figure 15. As seen previously, there is a significant decrease from baseline to all four other follow-up time points in both the multiply imputed and available case models. There is no significant change in IES from Post-Treatment to any of the following time points for the available case model. The multiply imputed model has a significant increase in score from post-treatment to 1 week ($p = 0.0066$) with no significant change from post-treatment to 3 months or 6 months. Both the multiply imputed model and the available case model found no significant change from 1 week to 3 months, but the multiply imputed model did find a significant decrease in DES

from 1 week to 6 months ($p = 0.0025$) while the available case model found no significant change from 1 week to 6 months. Neither the multiply imputed model nor the available case model found a significant change from 3 months to 6 months. Overall we see that for DRS there is a slight increase of DRS from post-treatment to 1 week, and then a slight decrease back to the post-treatment score from 1 week to 6 months. There is no overall significant change from post-treatment to 6 months.

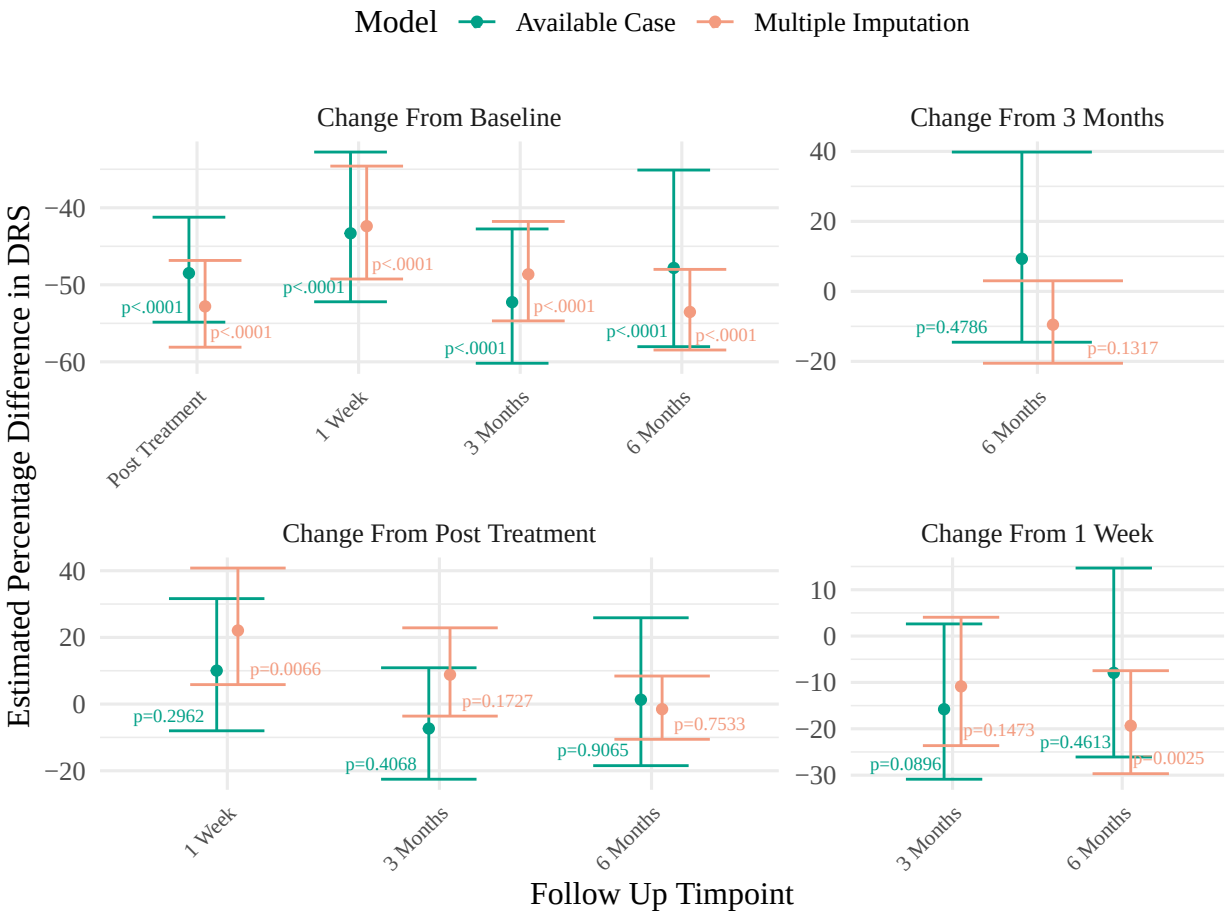

Figure 15: Estimated Change Between Timepoints for DRS

Multiple Imputation and Available Case Comparison

The results from the multiple imputation and the available case were approximately the same for all of our primary goals. Both modeling strategies found a steep decrease in scores from baseline to post-treatment with little to no change after. Below we investigate the differences between the models a bit further for each outcome.

IES

For the multiply imputed model and the available case model IES followed approximately the same trend with the multiply imputed model having slightly higher scores at all time points and a slight uptick at 3 months not seen in the available case model as seen in Figure 10. The average IES has a wider confidence interval in the multiply imputed model than in the available case model. In IES the change over time was approximately the same between the multiply imputed and available case models as discussed above. The estimates for other covariates in the model are presented in figure 16. The intercept is slightly lower for the multiply imputed model. Treatment days and age show little difference between the modeling methods. Primary diagnosis is also approximately the same, except those with no recorded diagnosis being slightly higher than those with a primary diagnosis. Overall, multiple imputation for IES did not significantly affect our results.

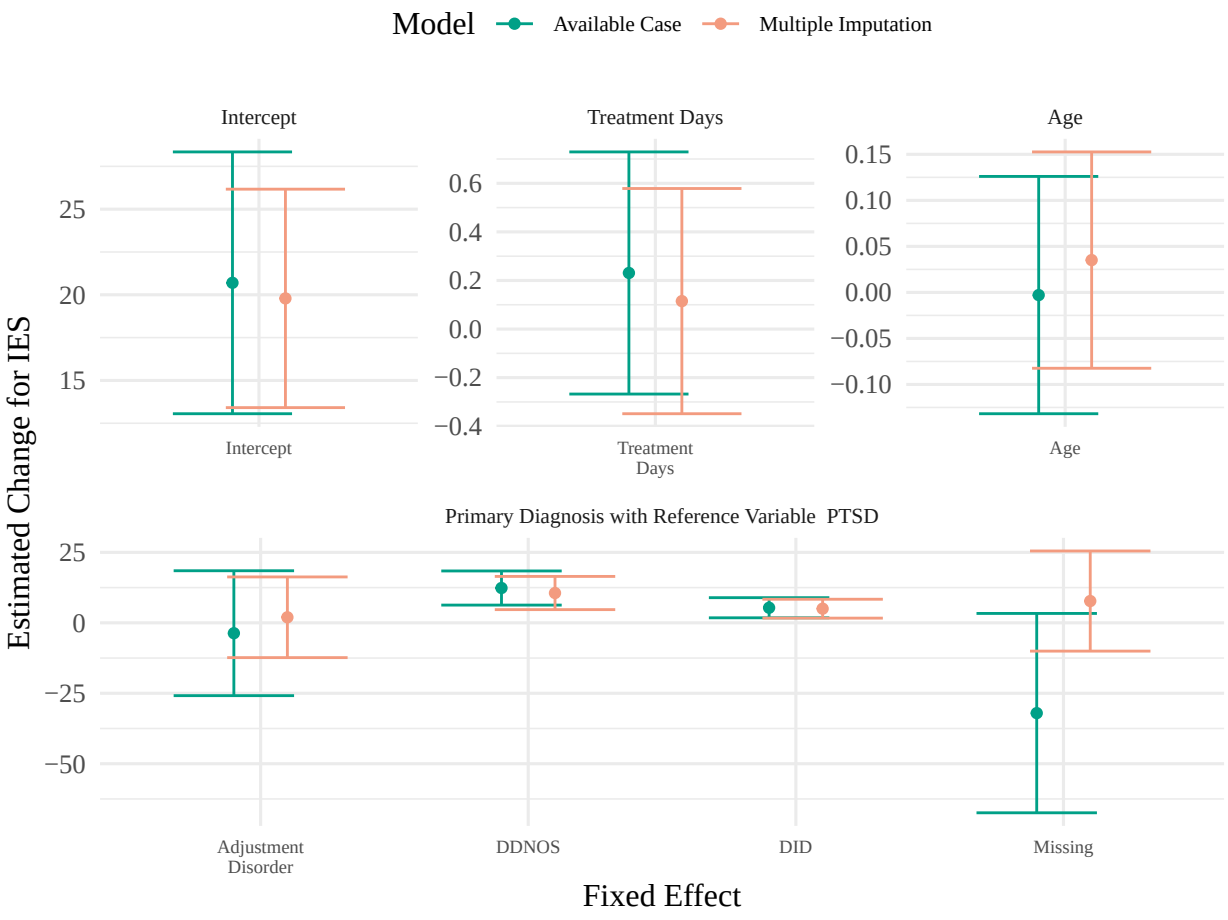

Figure 16: β Estimates for IES

TAS

For the multiply imputed model and the available case model TAS followed approximately the same trend with the multiply imputed model having slightly higher estimated scores at all time points as seen in Figure 10. The average TAS has a narrower confidence interval in the multiply imputed model than in the available case model. In TAS the change over time was approximately the same between the multiply imputed and available case models as discussed above. The estimates for other covariates in the model are presented in figure 17. The intercept is slightly lower for the multiply imputed model. Treatment days and age show little difference between the modeling methods. Overall, multiple imputation

for TAS did not significantly affect our results.

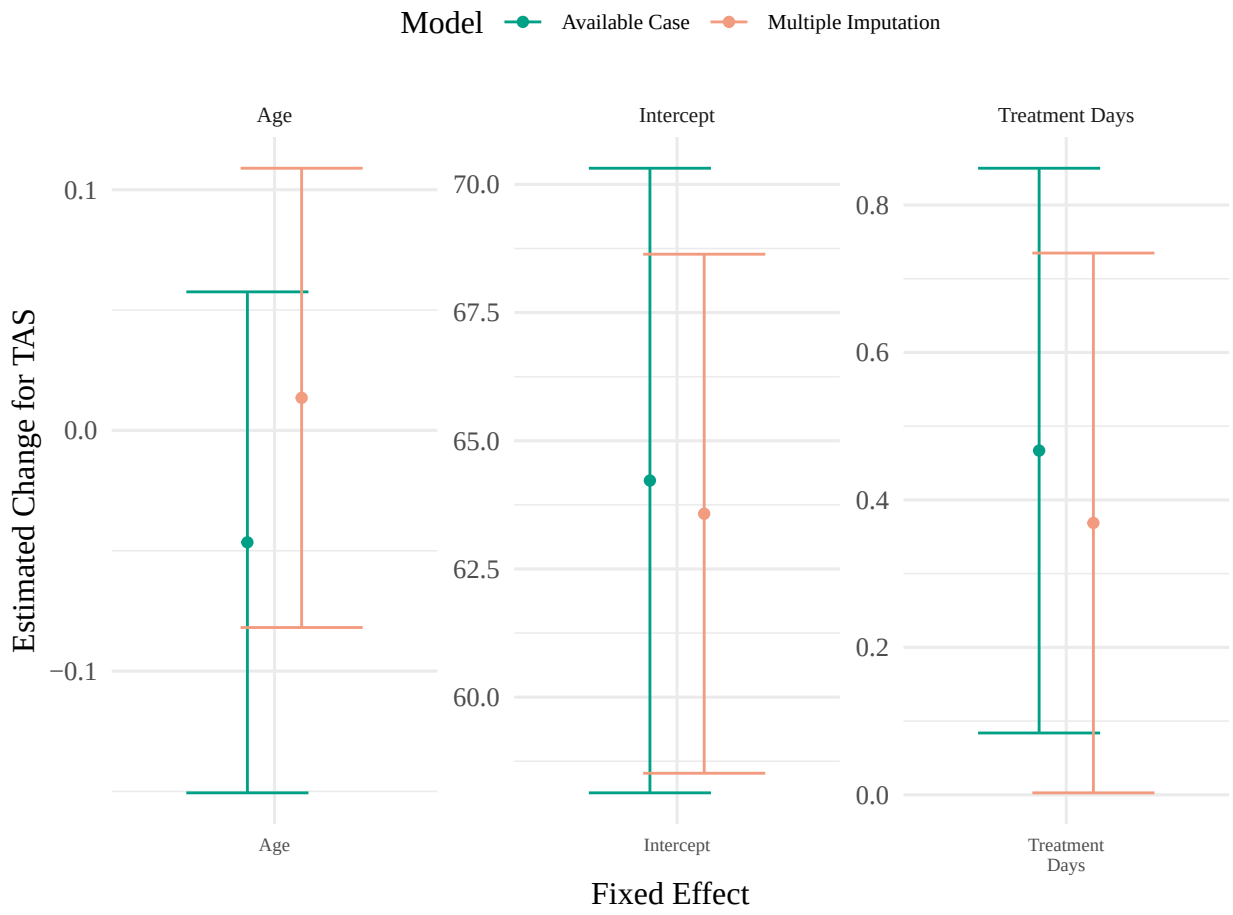

Figure 17: β Estimates for TAS

SCL

For the multiply imputed model and the available case model SCL followed approximately the same trend with the multiply imputed model having slightly higher scores at all time points and a slight downtick at 6 months not seen in the available case model as seen in Figure 10. The average SCL has a wider confidence interval in the multiply imputed model than in the available case model. In SCL the change over time was approximately the same between the multiply imputed and available case models as discussed above. The

estimates for other covariates in the model are presented in figure 18. The effect for women is slightly lower for the multiply imputed model. The intercept and treatment days show little difference between the modeling methods. Age is slightly lower in the available case model compared to the multiply imputed model. Primary diagnosis is also approximately the same, except those with no recorded diagnosis being slightly higher than those with a primary diagnosis. Overall, multiple imputation for SCL did not significantly affect our results.

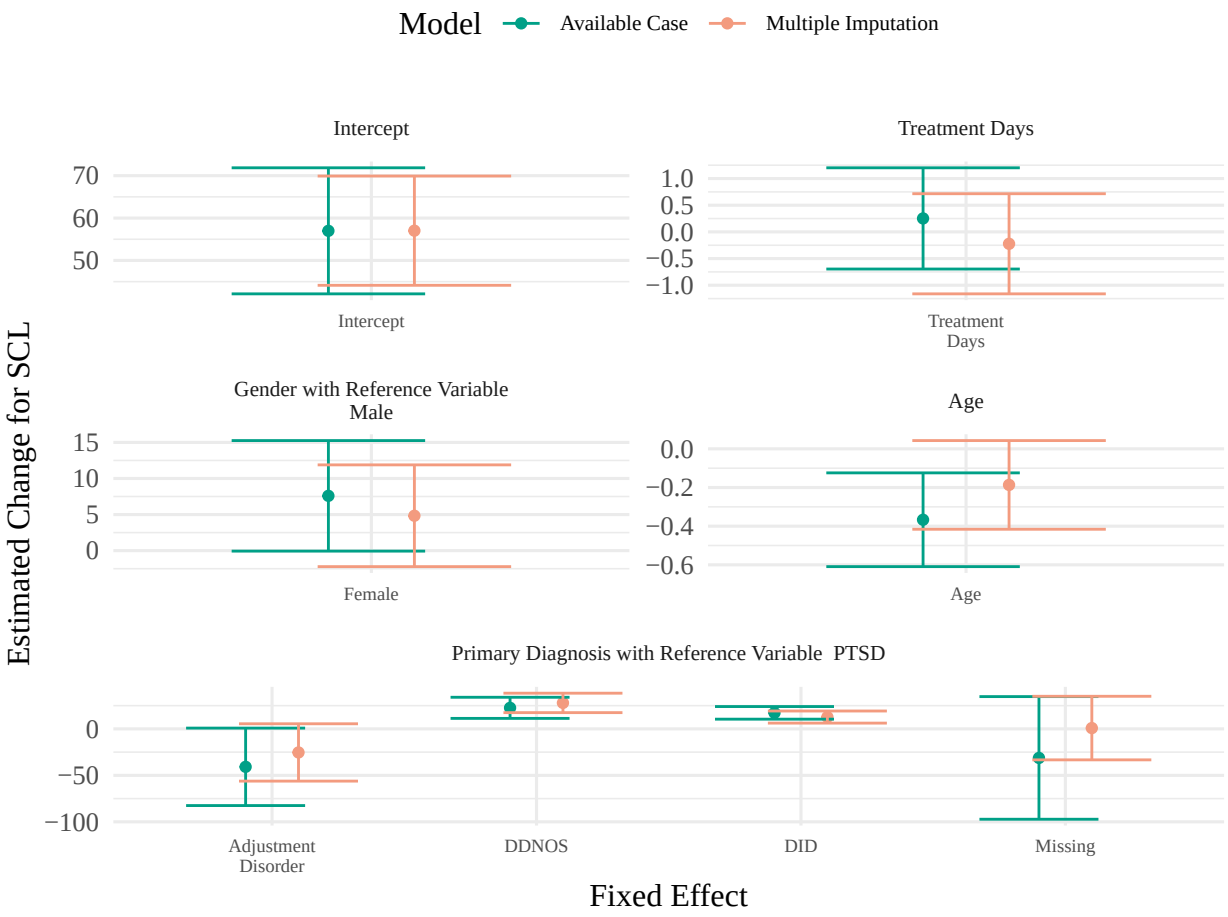

Figure 18: β Estimates for SCL

DES

For the multiply imputed model and the available case model DES followed approximately the same trend with the multiply imputed model having slightly higher scores at all time points as seen in Figure 10. The average DES has a wider confidence interval in the multiply imputed model than in the available case model. In DES the change over time was approximately the same between the multiply imputed and available case models as discussed above. The estimates for other covariates in the model are presented in figure 19. The intercept is slightly lower for the multiply imputed model. Treatment days and age show little difference between the modeling methods. Primary diagnosis is also approximately the same, except those with no recorded diagnosis being slightly higher than those with a primary diagnosis and those with adjustment disorder having a lower estimated β in the multiply imputed model than the available case model. Overall, multiple imputation for DES did not significantly affect our results.

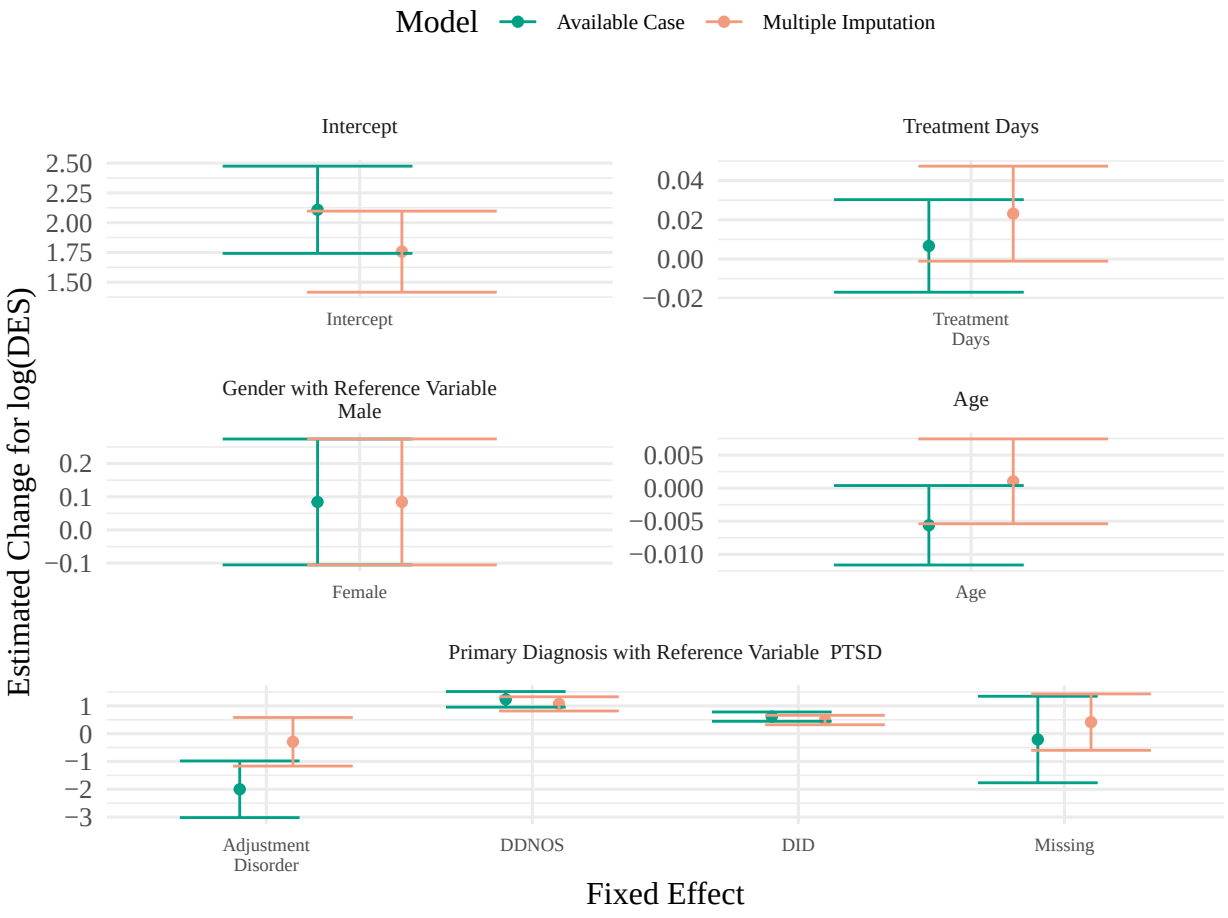

Figure 19: β Estimates for log(DES)

DRS

For the multiply imputed model and the available case model DRS followed approximately the same trend as seen in Figure 10. The average DRS has a wider confidence interval in the multiply imputed model than in the available case model. In DRS the change over time was approximately the same between the multiply imputed and available case models as discussed above. The estimates for other covariates in the model are presented in figure 20. The intercept is slightly lower for the multiply imputed model. Treatment days and age showed little difference between the modeling methods. Primary diagnosis is also approximately the same, except those with adjustment disorder have a lower estimated β in

the multiply imputed model than the available case model. Overall multiple imputation for DRS did not significantly affect our results.

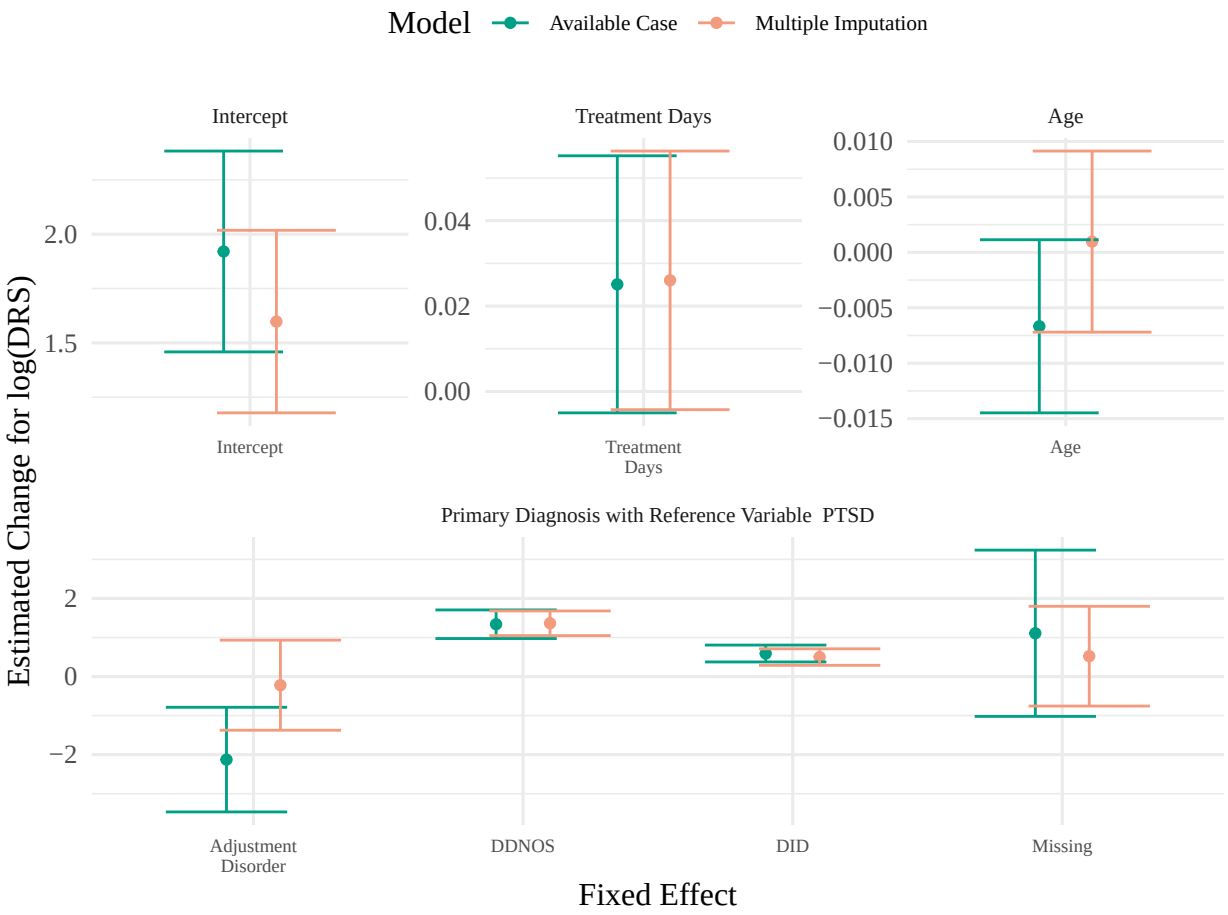

Figure 20: β Estimates for log(DRS)

CHAPTER IV

DISCUSSION

Model-Based Results

Overall, we found that the Intensive Trauma Response Treatment causes a steep initial decline in all five of the reported measures that are maintained after treatment has ended. This indicates that the ITR treatment causes a decrease in symptomology related to trauma that does not reoccur once treatment has ended. Note though, that due to the limitations of the data collected this cannot be generalized to anyone with trauma, but instead can be applied to those who have engaged in ITR from 2010-2017.

IES

As discussed above, for IES, or the Impact of Events Scale, we found an average decrease between -13 points to -20 points across the multiply imputed and the available case models. The average score for IES at baseline in this cohort is 43.02 (Quartiles: [33, 55]). A decrease of 20 points in IES from the mean would remove the clinical concern of PTSD, while a decrease of 13 points would take the subject below the cutoff for PTSD diagnosis (McCabe 2019). This shows that the estimated change in IES from baseline to 6 months is not only significant but clinically significant and truly represents a better outcome for subjects.

SCL

For the symptom checklists scores over 100 reflect severe distress, scores between 80 and 100 reflect moderate distress, and scores under 80 reflect mild distress. For SCL scores under 100 SCL and DES and DRS interact to find clinical diagnoses (Trauma Professionals 2012). The average score for SCL at baseline in this cohort is 89.82 (Quartiles: [64, 116]). We found an average change in SCL of approximately 30 points in both the multiply imputed model (95% CI: 27, 34) and the available case model (95% CI: 22, 36) from baseline to 6 months post-treatment. This would decrease the subjects with symptomology around average at baseline from moderate-severe distress to moderate or mild distress. Based on this the findings for SCL are clinically significant and represent a significant decrease in symptomology from baseline to 6 months post-treatment.

TAS

All measures under 62 in the Toronto alexithymia scale indicate no alexithymia. Scores above 62 represent the presence of alexithymia with scores over 70 showing chronic or severe alexithymia and scores from 62-70 showing mild or acute alexithymia (Trauma Professionals 2012). The average score for TAS at baseline in this cohort is 74.1 (Quartiles: [67, 81]). We found an average decrease in TAS from baseline to 6 months of 7 points in both the multiply imputed and available case models with 95% confidence intervals spanning 4 points to 10 points. A decrease in TAS of 4 from the mean of our cohort would not be a great change, while a 7-point decrease would show a significant decrease in the symptomology. Overall, the results found are clinically significant, but not to the same degree as our other measures.

DES

In DES, or the dissociative experience scale, scores under 20 do not indicate any diagnosis, while DES scores from 20-30 indicate PTSD, scores from 30-50 indicate a dissociative

disorder, and scores over 50 indicate a dissociative identity disorder (Trauma Professionals 2012). We found an estimated decrease in DES scores of 40% in the available case model (95% CI: 30, 49) and 46% in the multiply imputed model (95% CI: 41, 50). The average score for DES at baseline in this cohort is 24.46 (Quartiles: [11, 33]). For all subjects within the first three quartiles, this decrease would take them below the threshold for any diagnosis identified using DES. This is a clinically significant finding and represents a great decrease in dissociative symptomology from baseline to 6 months post-treatment.

DRS

DRS, or the dissociative regression scale, categorizes subjects into categories of regression present (DRS over 50) or no significant regression (DRS 50 or under) (Trauma Professionals 2012). We found an estimated decrease in DES scores of 48% in the available case model (95% CI: 35, 58) and 54% in the multiply imputed model (95% CI: 48, 58). The highest score in our cohort for DRS at baseline was 97. A decrease of 48% would give that subject a score of 50.44 and a decrease of 54% would give them a score of 44.62 putting them on the edge or below the threshold for the presence of regression indicating a clinically significant decrease in DRS score and representing a large decrease in dissociative regression from baseline to 6 months post-treatment.

Dropout

While initial analyses showed that dropout timepoint could affect a subject's outcome, we found very little evidence of this in our models. IES and TAS were the only measures that found any significant difference between dropout groups, with TAS showing a slightly more extreme decrease in TAS for those who dropped out at 6 months or 3 months than those who dropped out before 3 months. IES showed a slightly lower baseline score than those who dropped out at other time points. Overall, these differences were very minor and we concluded that dropout did not have a great impact on the outcomes.

Change after Post-Treatment

ITR is a time-bound intervention lasting a set amount of consecutive treatment days. Patients do receive emotional grounding tools which allow them to continue to self-treat after the intervention is over, which we hypothesized may lead to a continuing decrease in scores after treatment had finished. Our results, presented above, do not support this conclusion. While some measures had slight estimated increases or decreases at certain timepoint, no measure had an estimated significant change from post-treatment to 6 months. While this does not indicate that the self-treatment allows for a continuing decline in scores, those who receive the ITR intervention maintain the decrease in symptomology found at post-treatment up to the 6-month timepoint. This means that ITR provides lasting relief from the symptoms associated with trauma.

Limitations

While our dataset provided a substantial subject pool, the amount of data that could impact the independence between subjects was lacking. Because this data was collected by individual practitioners referred by the Help For Trauma company, the counselor could impact the between-subject independence, we do not have this data. We are also missing measures such as the year a subject went through ITR which could potentially impact the data, instead we have a general time range. Additionally, this was not a designed study, i.e. there was no study design to aid in data collection from the start. Instead, this data collected general clinic outcomes, so there is no randomization to account for potential confounders.

With any missing data analysis it is impossible to say if the data is MAR or MNAR with complete confidence. This is a large limitation of our analysis. If there are unobserved

predictors of missingness we cannot account or adjust for them. This highlights the importants of collecting a wide array of information for the subjects to account for everything possible in the case of dropout or missingness. Following up with subject who drop out to get a subset of measures or investigate reason of dropout could also aid in the missingness evaluation.

Conclusions

Overall, in the subjects contained within this dataset, we found a clinically significant decrease in every measure selected from baseline to post-treatment that was maintained for six months after treatment ended. This indicates that ITR is a valid model for the treatment of trauma and significantly decreases the symptomology associated with trauma. While there was one outcome where total change in outcome was significantly effected by dropout timepoint, this was not a clinically significant value. Therefore we concluded that dropout did not appear to significantly change the results.

Future Reccomendations

Future work concerning ITR should investigate the effect of intervention more thoroughly using a case-control study or a cohort study with thorough data collection protocols and randomization to account for confounders. Measures collected should be chosen to compare to work done on predominant trauma treatment protocols such as Eye Movement Desensitization and Reprocessing Therapy (EMDR) or Cognitive Behavioral Therapy (CBT) to more thoroughly be able to compare the interventions.

REFERENCES

Barnard, J. and Rubin, D. B. (1999). Miscellanea. small-sample degrees of freedom with multiple imputation. Biometrika, 86(4):948–955.

Carlson, E. B. and Putnam, F. W. (1993). Dissociation: Vol. 6, no. 1, p. 016-027: An update on the dissociative experience scale.

Coggon, D., Barker, D., and Rose, G. (2009). Epidemiology for the Uninitiated. John Wiley & Sons.

Enders, C. K. (2010). Applied missing data analysis. Guilford Publications.

Fernandez, G. (2007). Model selection in proc mixed—a user—friendly sas® macro application. In SAS Global Forum, volume 191, pages 1–20.

Gantt, L. and Tinnin, L. (2009). Support for a neurobiological view of trauma with implications for art therapy. The Arts in Psychotherapy, 36(3):148–153.

Gantt, L. and Tinnin, L. W. (2007). Intensive trauma therapy of ptsd and dissociation: An outcome study. The Arts in Psychotherapy, 34(1):69–80.

Li, P., Stuart, E. A., and Allison, D. B. (2015). Multiple imputation: a flexible tool for handling missing data. Jama, 314(18):1966–1967.

Little, R. J. (1988a). Missing-data adjustments in large surveys. Journal of Business & Economic Statistics, 6(3):287–296.

Little, R. J. (1988b). A test of missing completely at random for multivariate data with missing values. Journal of the American statistical Association, 83(404):1198–1202.

Little, R. J. and Rubin, D. B. (2019). Statistical analysis with missing data, volume 793. John Wiley & Sons.

Marshall, A., Altman, D. G., Royston, P., and Holder, R. L. (2010). Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. BMC medical research methodology, 10(1):1–16.

McCabe, D. (2019). The impact of event scale-revised (ies-r). New York University Rory Meyers College of Nursing, 19.

Molenberghs, G. and Verbeke, G. (2001). A review on linear mixed models for longitudinal data, possibly subject to dropout. Statistical Modelling, 1(4):235–269.

Nakai, M. and Ke, W. (2011). Review of the methods for handling missing data in longitudinal data analysis. International Journal of Mathematical Analysis, 5(1):1–13.

Nicholson, J. S., Deboeck, P. R., and Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. International Journal of Behavioral Development, 41(1):143–153.

of Trauma Professionals, I. A. (2012). Trauma therapist toolkit.

Olff, M. (2017). Sex and gender differences in post-traumatic stress disorder: an update. European journal of psychotraumatology, 8(sup4):1351204.

Rubin, D. B. (1976). Inference and missing data. Biometrika, 63(3):581–592.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. Journal of Business & Economic Statistics, 4(1):87–94.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys, volume 81. John Wiley & Sons.

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., and Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. Methods in ecology and evolution, 11(9):1141–1152.

Shaw, C., Wu, Y., Zimmerman, S. C., Hayes-Larson, E., Belin, T. R., Power, M. C., Glymour, M. M., and Mayeda, E. R. (2023). Comparison of imputation strategies for incomplete longitudinal data in lifecourse epidemiology. American Journal of Epidemiology, page kwad139.

Tinnin, L., Bills, L., and Gantt, L. (2014). Short-term treatment of simple and complex ptsd. pages 99–118.

Van Buuren, S. (2018). Flexible imputation of missing data. CRC press.

White, I. R., Royston, P., and Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine, 30(4):377–399.

Code Appendix

SAS Code

```
proc import datafile="super_wide_data.csv"
    out=data_wide
     dbms=csv replace;
     getnames=yes;
run;

DATA data_wide(drop=VAR1);
    SET data_wide;
RUN;


PROC SORT data=data_wide;
    by ID last;
    RUN;


PROC TRANSPOSE data=data_wide out=dw_temp(rename=(col1=value));
    by _all_;
    var Baseline_IES--Six_Months_TAS;
RUN;
```

```
data dw_temp(drop=Baseline_ies--six_months_tas);
set dw_temp;
run;


DATA dw_imp;
    set dw_temp;
    length measure $20 time $20;
    measure = scan(_name_, 3, '_');
    if measure = '' then
        measure = scan(_name_, 2, '_');
    time = substr(_name_, 1,
      length(_name_) - length(measure) - 1);
    format time timefmt;
RUN;


data dw_imp(drop=_name_);
set dw_imp;
run;

proc sort data=dw_imp;
    by ID  time ;
    run;
```

```
proc transpose data=dw_imp out=dwide_imp;
    by  ID time TX_DAYS last
        Gender Age Axis_I_Diagnosis
        HAL ALX_PRS_ ALX__POST
        Trauma_Type;
    id measure;
    var value;
    run;




DATA final_reg;
    SET dwide_imp;
    IF time='Baseline' THEN time_num = 1;
    ELSE IF time='Post_Treatment' THEN time_num = 2;
    ELSE IF time='One_Week' THEN time_num = 3;
    ELSE IF time='Three_Months' THEN time_num = 4;
    ELSE IF time='Six_Months' THEN time_num = 5;
    IF last='Baseline' THEN last_num = 1;
    ELSE IF last='Post Treatment' THEN last_num = 2;
    ELSE IF last='1 Week' THEN last_num = 3;
    ELSE IF last='3 Months' THEN last_num = 4;
    ELSE IF last='6 Months' THEN last_num = 5;
    logDRS = log(drs+1);
    logDES = log(des+1);
```

```
    IF time='Baseline' THEN numtime = 0-TX_DAYS;
    ELSE IF time='Post_Treatment' THEN numtime=0;
    ELSE IF time='One_Week' THEN numtime=7;
    ELSE IF time='Three_Months' THEN numtime=90;
    ELSE IF time='Six_Months' THEN numtime=180;
RUN;

proc sort data=final_reg;
    by id time_num last_num;
    run;



libname allmix4 base "\\Mac\Home\Documents\allmixed";
options sasmstore=allmix4 mstored;

data allmix4.final_reg;
set final_reg;
run;



%allmixed
( data_  = sas_final_reg
, respi = IES
,GLMSELECT =
, class    = time_num last  ID
```

```
, z =   _IES_cov_select1
, MODOPT=
, must = time_num last time_num*last
, fixed1 =
, fixed2 =
, Random =  random intercept  / subject=ID
, Repeat = repeated  / subject=ID type=
, sub = ID
, covari= ar(1) un cs
  sp(pow)(time_num)
  sp(pow)(numtime)
  sp(exp)(numtime)
  sp(exp)(time_num)
, explor = , graph = pdf, output = \\Mac\Home\Documents\allmixed\, lsmeans   =
, dir2 =    \\Mac\Home\Documents\allmixed\IES\
, start = , Stop =
)

/*sp(exp)(numtime) did best*/

%allmixed
( data_  = sas_final_reg
, respi = IES
,GLMSELECT =
, class    = time_num last
    Axis_I_Diagnosis Gender
    Trauma_type ID
```

```
, z =   _IES_fixed_select1
, MODOPT=
, must = time_num last time_num*last
, fixed1 = Axis_i_Diagnosis Gender Trauma_type
, fixed2 =TX_Days Age
, Random =  random intercept / subject=ID
, Repeat = repeated / subject=ID
       type=sp(exp)(numtime)
, sub = ID
, covari=
, explor = , graph = pdf, output = \\Mac\Home\Documents\allmixed\, lsmeans   =
, dir2 =    \\Mac\Home\Documents\allmixed\IES\
, start = , Stop =
)

/* we go with mdl for a more parsimonious
model here with Axis_i_Diagnosis TX_Days Age*/

%allmixed
( data_  = sas_final_reg
, respi = IES
,GLMSELECT =
, class    = time_num last
    Axis_I_Diagnosis Gender
    Trauma_type ID
, z =   _IES_cov_confirm
, MODOPT=
```

```
, must = time_num last time_num*last
    Axis_i_Diagnosis TX_Days Age
, fixed1 =
, fixed2 =
, Random =  random intercept  / subject=ID
, Repeat = repeated  / subject=ID type=
, sub = ID
, covari= ar(1) un cs sp(pow)
    (time_num)
    sp(pow)(numtime)
    sp(exp)(numtime)
    sp(exp)(time_num)
, explor = , graph = pdf, output = \\Mac\Home\Documents\allmixed\, lsmeans   =
, dir2 =    \\Mac\Home\Documents\allmixed\IES\
, start = , Stop =
)

/*ar(1) and sp(pow)(time_num) do about the same*/




proc import datafile="super_wide_data.csv"
     out=data_wide
     dbms=csv replace;
     getnames=yes;
```

```
run;

DATA data_wide(drop=VAR1);
    SET data_wide;
RUN;


PROC SORT data=data_wide;
    by ID last;
    RUN;


PROC TRANSPOSE data=data_wide out=dw_temp(rename=(col1=value));
    by _all_;
    var Baseline_IES--Six_Months_TAS;
RUN;


data dw_temp(drop=Baseline_ies--six_months_tas);
set dw_temp;
run;


DATA dw_imp;
    set dw_temp;
    length measure $20 time $20;
    measure = scan(_name_, 3, '_');
```

```
    if measure = '' then
      measure = scan(_name_, 2, '_');
    time = substr(_name_, 1,
      length(_name_) - length(measure) - 1);
    format time timefmt;
RUN;


data dw_imp(drop=_name_);
set dw_imp;
run;

proc sort data=dw_imp;
    by ID  time ;
    run;



proc transpose data=dw_imp out=dwide_imp;
    by  ID time TX_DAYS
        last Gender Age
        Axis_I_Diagnosis HAL ALX_PRS_
        ALX__POST Trauma_Type;
    id measure;
    var value;
    run;
```

```
DATA final_reg;
    SET dwide_imp;
    IF time='Baseline' THEN time_num = 1;
    ELSE IF time='Post_Treatment' THEN time_num = 2;
    ELSE IF time='One_Week' THEN time_num = 3;
    ELSE IF time='Three_Months' THEN time_num = 4;
    ELSE IF time='Six_Months' THEN time_num = 5;
    IF last='Baseline' THEN last_num = 1;
    ELSE IF last='Post Treatment' THEN last_num = 2;
    ELSE IF last='1 Week' THEN last_num = 3;
    ELSE IF last='3 Months' THEN last_num = 4;
    ELSE IF last='6 Months' THEN last_num = 5;
    logDRS = log(drs+1);
    logDES = log(des+1);
    IF time='Baseline' THEN numtime = 0-TX_DAYS;
    ELSE IF time='Post_Treatment' THEN numtime=0;
    ELSE IF time='One_Week' THEN numtime=7;
    ELSE IF time='Three_Months' THEN numtime=90;
    ELSE IF time='Six_Months' THEN numtime=180;
RUN;
```

```
proc sort data=final_reg;
    by id time_num last_num;
    run;

data final_reg_freq;
 set final_reg;
 if cmiss(of _all_) then delete;
run;
proc freq data=final_reg_freq;
  tables time_num*last_num / out=group_counts noprint;
run;

proc sql;
   create table weights as
   select
     last_num,
     count(*) as n
   from final_reg_freq
   group by last_num;
quit;

proc sql;
   create table final_reg_freq_weighted as
   select
     a.*,
     b.n / (select count(*) from final_reg_freq) as weight
   from final_reg_freq a
```

```
   join weights b
     on a.last_num = b.last_num;
quit;

%include "basic_models.sas";

ods exclude all;



/*----------------------*/
/*IES*/
/*----------------------*/
proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
  Axis_i_Diagnosis TX_Days Age/
      solution CL residual outp=Pred;
random intercept numtime / subject=ID  g;
repeated / subject=ID type=sp(pow)(numtime);
ods output FitStatistics=fits; run;

data fits;
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
```

```
      model_descriptions.csv,
      sp(pow)(numtime),
      IES, , Random Intercept
      Random Slope for Time in Days ,
     final_reg_freq, ML );


data fits;
set fits;
Value = .;
run;

     proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num
    time_num last_num
    Axis_i_Diagnosis TX_Days
    Age/ solution CL residual outp=Pred;
random intercept numtime / subject=ID  g;
repeated / subject=ID type=sp(exp)(numtime);
ods output FitStatistics=fits; run;

data fits;
set fits;
run;

 %saveall(fits  , IES,
```

```
      time-as-fact, corr_compare,
      model_descriptions.csv,
      sp(exp)(numtime),
      IES, ,
      Random Intercept Random Slope for Time in Days ,
     final_reg_freq, ML );
data fits;
set fits;
Value = .;
run;


proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
      Axis_i_Diagnosis TX_Days
      Age/ solution CL residual outp=Pred;
random intercept numtime / subject=ID  g;
repeated / subject=ID type=sp(pow)(time_num);
ods output FitStatistics=fits; run;
data fits;
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
```

```
      sp(pow)(time_num),
      IES, ,
      Random Intercept Random Slope for Time in Days ,
     final_reg_freq, ML );


data fits;
set fits;
Value = .;
run;




proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
  Axis_i_Diagnosis TX_Days
  Age/ solution CL residual outp=Pred;
random intercept numtime / subject=ID  g;
repeated / subject=ID type=sp(exp)(time_num);
ods output FitStatistics=fits; run;
data fits;
set fits;
run;

 %saveall(fits , IES,
```

```
      time-as-fact, corr_compare,
      model_descriptions.csv,
      sp(exp)(time_num),
      IES, ,
      Random Intercept Random Slope for Time in Days ,
     final_reg_freq, ML );



data fits;
set fits;
Value = .;
run;




proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
      Axis_i_Diagnosis TX_Days
      Age/ solution CL residual outp=Pred;
random intercept numtime / subject=ID  g;
repeated / subject=ID type=ar(1);
ods output FitStatistics=fits; run;


data fits;
```

```
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
      ar(1),
      IES, ,
      Random Intercept Random Slope for Time in Days ,
     final_reg_freq, ML );

data fits;
set fits;
Value = .;
run;




proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
    Axis_i_Diagnosis TX_Days
    Age/ solution CL residual outp=Pred;
random intercept  / subject=ID  g;
repeated / subject=ID type=sp(pow)(numtime);
```

```
ods output FitStatistics=fits; run;



data fits;
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
      sp(pow)(numtime),
      IES, , Random Intercept ,
     final_reg_freq, ML );
data fits;
set fits;
Value = .;
run;

     proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
    Axis_i_Diagnosis TX_Days
    Age/ solution CL residual outp=Pred;
random intercept / subject=ID  g;
repeated / subject=ID type=sp(exp)(numtime);
ods output FitStatistics=fits; run;
```

```
data fits;
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
      sp(exp)(numtime),
      IES, , Random Intercept ,
     final_reg_freq, ML );
data fits;
set fits;
Value = .;
run;


proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
    Axis_i_Diagnosis TX_Days
    Age/ solution CL residual outp=Pred;
random intercept / subject=ID  g;
repeated / subject=ID type=sp(pow)(time_num);
ods output FitStatistics=fits; run;

data fits;
```

```
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
      sp(pow)(time_num),
      IES, , Random Intercept,
     final_reg_freq, ML );
data fits;
set fits;
Value = .;
run;




proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
    Axis_i_Diagnosis TX_Days
    Age/ solution CL residual outp=Pred;
random intercept / subject=ID  g;
repeated / subject=ID type=sp(exp)(time_num);
ods output FitStatistics=fits; run;
```

```
data fits;
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
      sp(exp)(time_num),
      IES, , Random Intercept ,
     final_reg_freq, ML );
data fits;
set fits;
Value = .;
run;
```

```
proc mixed data=final_reg_freq method=ML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num last_num
    Axis_i_Diagnosis TX_Days
    Age/ solution CL residual outp=Pred;
random intercept / subject=ID  g;
```

```
repeated / subject=ID type=ar(1);
ods output FitStatistics=fits; run;

data fits;
set fits;
run;

 %saveall(fits , IES,
      time-as-fact, corr_compare,
      model_descriptions.csv,
      ar(1),
      IES, , Random Intercept  ,
     final_reg_freq, ML );
data fits;
set fits;
Value = .;
run;

proc mixed data=final_reg_freq method=REML;
class time_num gender last_num axis_i_diagnosis;
model IES = time_num*last_num time_num
    last_num Axis_i_Diagnosis TX_Days Age/ solution CL outp=Pred;
random intercept  / subject=ID  g;
repeated / subject=ID type=ar(1);
estimate 'total change comp last_num = 3-2'
    time_num*last_num 0 -1  1 0 0    1 0 0
                      -1 0 0    0 0   0
```

```
                    time_num 0 1 -1 0 0 / cl ;
estimate 'total change comp last_num = 4-2'
    time_num*last_num 0 -1  0 1 0    1 0 0
                   0 0 0    -1 0   0
                   time_num 0 1 0 -1 0 / cl ;
estimate 'total change comp last_num = 5-2'
    time_num*last_num 0 -1  0 0 1    1 0 0
                     0 0 0    0 0   -1
                     time_num 0 1 0 0 -1 / cl ;
contrast 'total_change last_num: 2 = 3 = 4 = 5 ?'
    time_num*last_num 0 -1 1 0 0    1 0 0
                     -1 0 0    0 0   0
                     time_num 0 1 -1 0 0,
    time_num*last_num 0 -1 0 1 0    1 0 0
                     0 0 0    -1 0   0
                     time_num 0 1 0 -1 0,
    time_num*last_num 0 -1 0 0 1    1 0 0
                     0 0 0    0 0   -1
                     time_num 0 1 0 0 -1;
contrast 'baseline last_num 1=2=3=4=5 or 1-2 = 1-3 = 1-4 = 1-5 = 0?'
    time_num*last_num 1 -1 0 0 0    0 0 0
                     0 0 0    0 0   0
                     last_num 1 -1 0 0 0,
    time_num*last_num 1 0 -1 0 0    0 0 0
                     0 0 0    0 0   0
                     last_num 1 0 -1 0 0 ,
    time_num*last_num 1 0 0 -1 0    0 0 0
```

```
                    0 0 0    0 0   0
                    last_num 1 0 0 -1 0 ,
    time_num*last_num 1 0 0 0 -1    0 0 0
                    0 0 0    0 0   0
                    last_num 1 0 0 0 -1;
estimate 'baseline last_num 1-2 '
    time_num*last_num 1 -1 0 0 0    0 0 0
                    0 0 0    0 0   0
                    last_num 1 -1 0 0 0/cl;
estimate 'baseline last_num 1-3 '
    time_num*last_num 1 0 -1 0 0    0 0 0
                    0 0 0    0 0   0
                    last_num 1 0 -1 0 0/cl;
estimate 'baseline last_num 1-4 '
    time_num*last_num 1 0 0 -1 0    0 0 0
                    0 0 0    0 0   0
                    last_num 1 0 0 -1 0/cl;
estimate 'baseline last_num 1-5 '
    time_num*last_num 1 0 0 0 -1    0 0 0
                    0 0 0    0 0   0
                    last_num 1 0 0 0 -1/cl;
lsmeans time_num / OM bylevel;
lsmeans time_num / OM DIFF bylevel;
ODS OUTPUT  FitStatistics=fits SolutionF=fixed CovParms=CP convergenceStatus=conv
                modelinfo=mod tests3=tests  Dimensions=dims estimates=ests lsmeans=lsmea
                contrasts=contrasts;
run;
```

```
 %saveall(contrasts fits fixed CP conv mod
      tests Pred dims ests
      lsmeans diffs, IES,
      time-as-fact, sas_output,
      model_descriptions.csv,
      Basic LMM for IES,
      IES,
Follow-Up Timepoint as Factor +
Last Timepoint as Factor +
Interaction Between Follow-Up and Last Timepoint +
Primary Diagnosis +
Treatment Days +
Age,
Random Intercept for Subject - AR(1)  ,
     final_reg_freq, REML );



proc import datafile="super_wide_data.csv"
    out=data_wide
     dbms=csv replace;
     getnames=yes;

    GUESSINGROWS=MAX;
run;
```

```
data data_wide(drop=last);
set data_wide;run;

ods exclude none;

PROC MI data=data_wide out=mice_out
    nimpute=52 seed=12345;
      class Trauma_Type gender
            axis_I_diagnosis alx_prs_
            alx__post hal;
      fcs plots=TRACE nbiter = 50
        regpmm(TX_DAYS=Age Gender
                Axis_I_Diagnosis ALX_PRS_
                ALX__POST / details);
      fcs plots=TRACE nbiter = 50
        regpmm(Age=TX_DAYS / details);
      fcs plots=TRACE nbiter = 50
        discrim(Axis_I_Diagnosis=TX_DAYS HAL
                Trauma_Type / details
                      classeffects=include);
      fcs plots=TRACE nbiter = 50
        logistic(HAL=Axis_I_Diagnosis ALX_PRS_
              ALX__POST Trauma_Type / details);
      fcs plots=TRACE nbiter = 50
          logistic(ALX_PRS_=TX_DAYS Gender
```

```
        HAL ALX__POST / details);
fcs plots=TRACE nbiter = 50
  logistic(ALX__POST=Age Gender HAL
      ALX_PRS_ / details);
fcs plots=TRACE nbiter = 50
  discrim(Trauma_Type=Axis_I_Diagnosis
      ALX_PRS_ /
      details classeffects=include);
fcs plots=TRACE nbiter = 50
    regpmm(Baseline_IES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_
        ALX__POST Trauma_Type Baseline_DES
        Baseline_SCL Baseline_DRS Baseline_TAS
        Post_Treatment_IES One_Week_IES
        Three_Months_IES Six_Months_IES / details);
fcs plots=TRACE nbiter = 50
    regpmm(Baseline_DES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Baseline_SCL
        Baseline_DRS Baseline_TAS
        Post_Treatment_DES One_Week_DES
        Three_Months_DES Six_Months_DES / details);
fcs plots=TRACE nbiter = 50
    regpmm(Baseline_SCL=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Baseline_DES
        Baseline_DRS Baseline_TAS
```

```
        Post_Treatment_SCL One_Week_SCL
        Three_Months_SCL Six_Months_SCL / details);
fcs plots=TRACE nbiter = 50
    regpmm(Baseline_DRS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Baseline_DES
        Baseline_SCL Baseline_TAS
        Post_Treatment_DRS One_Week_DRS
        Three_Months_DRS Six_Months_DRS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Baseline_TAS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Baseline_DES
        Baseline_SCL Baseline_DRS
        Post_Treatment_TAS One_Week_TAS
        Three_Months_TAS Six_Months_TAS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Post_Treatment_IES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Post_Treatment_DES
        Post_Treatment_SCL Post_Treatment_DRS
        Post_Treatment_TAS One_Week_IES
        Three_Months_IES Six_Months_IES / details);
fcs plots=TRACE nbiter = 50
    regpmm(Post_Treatment_DES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DES Post_Treatment_IES
```

```
        Post_Treatment_SCL Post_Treatment_DRS
        Post_Treatment_TAS One_Week_DES
        Three_Months_DES Six_Months_DES / details);
fcs plots=TRACE nbiter = 50
    regpmm(Post_Treatment_SCL=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_SCL Post_Treatment_IES
        Post_Treatment_DES Post_Treatment_DRS
        Post_Treatment_TAS One_Week_SCL
        Three_Months_SCL Six_Months_SCL / details);
fcs plots=TRACE nbiter = 50
    regpmm(Post_Treatment_DRS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DRS Post_Treatment_IES
        Post_Treatment_DES Post_Treatment_SCL
        Post_Treatment_TAS One_Week_DRS
        Three_Months_DRS Six_Months_DRS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Post_Treatment_TAS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_TAS Post_Treatment_IES
        Post_Treatment_DES Post_Treatment_SCL
        Post_Treatment_DRS One_Week_TAS
        Three_Months_TAS Six_Months_TAS / details);
fcs plots=TRACE nbiter = 50
    regpmm(One_Week_IES=TX_DAYS Age Gender
    Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
```

```
    Trauma_Type Baseline_IES Post_Treatment_IES
    One_Week_DES One_Week_SCL One_Week_DRS
    One_Week_TAS Three_Months_DES Three_Months_SCL
    Three_Months_DRS Three_Months_TAS
    Six_Months_IES / details);
fcs plots=TRACE nbiter = 50
    regpmm(One_Week_DES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DES Post_Treatment_DES
        One_Week_IES One_Week_SCL One_Week_DRS
        One_Week_TAS Three_Months_IES
        Three_Months_SCL Three_Months_DRS
        Three_Months_TAS Six_Months_DES / details);
fcs plots=TRACE nbiter = 50
    regpmm(One_Week_SCL=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_SCL Post_Treatment_SCL
        One_Week_IES One_Week_DES One_Week_DRS
        One_Week_TAS Three_Months_IES
        Three_Months_DES Three_Months_DRS
        Three_Months_TAS Six_Months_SCL / details);
fcs plots=TRACE nbiter = 50
    regpmm(One_Week_DRS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DRS Post_Treatment_DRS
        One_Week_IES One_Week_DES One_Week_SCL
        One_Week_TAS Three_Months_IES
```

```
        Three_Months_DES Three_Months_SCL
        Three_Months_TAS Six_Months_DRS / details);
fcs plots=TRACE nbiter = 50
    regpmm(One_Week_TAS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_TAS Post_Treatment_TAS
        One_Week_IES One_Week_DES One_Week_SCL
        One_Week_DRS Three_Months_IES
        Three_Months_DES Three_Months_SCL
        Three_Months_DRS Six_Months_TAS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Three_Months_IES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Post_Treatment_IES
        One_Week_IES Three_Months_DES
        Three_Months_SCL Three_Months_DRS
        Three_Months_TAS Six_Months_IES / details);
fcs plots=TRACE nbiter = 50
    regpmm(Three_Months_DES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DES Post_Treatment_DES
        One_Week_DES Three_Months_IES
        Three_Months_SCL Three_Months_DRS
        Three_Months_TAS Six_Months_DES / details);
fcs plots=TRACE nbiter = 50
    regpmm(Three_Months_SCL=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
```

```
        Trauma_Type Baseline_SCL Post_Treatment_SCL
        One_Week_SCL Three_Months_IES
        Three_Months_DES Three_Months_DRS
        Three_Months_TAS Six_Months_SCL / details);
fcs plots=TRACE nbiter = 50
    regpmm(Three_Months_DRS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DRS Post_Treatment_DRS
        One_Week_DRS Three_Months_IES
        Three_Months_DES Three_Months_SCL
        Three_Months_TAS Six_Months_DRS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Three_Months_TAS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_TAS Post_Treatment_TAS
        One_Week_TAS Three_Months_IES
        Three_Months_DES Three_Months_SCL
        Three_Months_DRS Six_Months_TAS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Six_Months_IES=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_IES Post_Treatment_IES
        One_Week_IES Three_Months_IES
        Six_Months_DES Six_Months_SCL
        Six_Months_DRS Six_Months_TAS / details);
fcs plots=TRACE nbiter = 50
    regpmm(Six_Months_DES=TX_DAYS Age Gender
```

```
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_DES Post_Treatment_DES
        One_Week_DES Three_Months_DES
        Six_Months_IES Six_Months_SCL
        Six_Months_DRS Six_Months_TAS / details);
 fcs plots=TRACE nbiter = 50
    regpmm(Six_Months_SCL=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_SCL Post_Treatment_SCL
        One_Week_SCL Three_Months_SCL
        Six_Months_IES Six_Months_DES
        Six_Months_DRS Six_Months_TAS / details);
 fcs plots=TRACE nbiter = 50
    regpmm(Six_Months_DRS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX__POST Trauma_Type
        Baseline_DRS Post_Treatment_DRS
        One_Week_DRS Three_Months_DRS
        Six_Months_IES Six_Months_DES
        Six_Months_SCL Six_Months_TAS / details);
 fcs plots=TRACE nbiter = 50
    regpmm(Six_Months_TAS=TX_DAYS Age Gender
        Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
        Trauma_Type Baseline_TAS Post_Treatment_TAS
        One_Week_TAS Three_Months_TAS
        Six_Months_IES Six_Months_DES
        Six_Months_SCL Six_Months_DRS / details)
plots=trace(mean(Six_Months_TAS)
```

```
     mean(Three_Months_TAS));
     var TX_DAYS Gender Age
          Axis_I_Diagnosis HAL ALX_PRS_ ALX__POST
          Trauma_Type Baseline_IES Baseline_DES
          Baseline_SCL Baseline_DRS Baseline_TAS
          Post_Treatment_IES Post_Treatment_DES
          Post_Treatment_SCL Post_Treatment_DRS
          Post_Treatment_TAS One_Week_IES One_Week_DES
          One_Week_SCL One_Week_DRS One_Week_TAS
          Three_Months_IES Three_Months_DES
          Three_Months_SCL Three_Months_DRS
          Three_Months_TAS Six_Months_IES Six_Months_DES
          Six_Months_SCL Six_Months_DRS Six_Months_TAS;
 ods output corr=micecor
    fcsmodel=fcs_output
    modelinfo=miceinfo
    paramerterestimats=micesparams
    varianceinfo=varinf;
run;
```

```
DATA mice_out(drop=VAR1 alx__post alx_prs_);
    SET mice_out;
RUN;



PROC SORT data=mice_out;
    by _imputation_ id
      tx_days age gender
      axis_i_diagnosis hal
      trauma_type;
RUN;

proc print data=mice_out(obs=10);run;

PROC TRANSPOSE data=mice_out out=long_temp(rename=(col1=value));
    by _imputation_ id tx_days
        age gender axis_i_diagnosis
        hal  trauma_type;
    var Baseline_IES--Six_Months_TAS;
RUN;



data long_temp(drop=col2);
set long_temp;
run;
```

```
proc print data=long_temp(obs=10);run;
DATA long_imp;
    set long_temp;
    length measure $20 time $20;
    measure = scan(_name_, 3, '_');
    if measure = '' then measure = scan(_name_, 2, '_');
    time = substr(_name_, 1,
    length(_name_) - length(measure) - 1);
RUN;

data long_imp(drop=_name_);
set long_imp;
run;



proc sort data=long_imp;
    by _imputation_ ID  time ;
    run;



proc transpose data=long_imp out=wide_imp;
    by  _imputation_ ID time
        TX_DAYS Gender Age
        Axis_I_Diagnosis HAL
        Trauma_Type;
    id measure;
```

```
    var value;
    run;


DATA final_mice;
    SET wide_imp;
    IF time='Baseline' THEN time_num = 1;
    ELSE IF time='Post_Treatment' THEN time_num = 2;
    ELSE IF time='One_Week' THEN time_num = 3;
    ELSE IF time='Three_Months' THEN time_num = 4;
    ELSE IF time='Six_Months' THEN time_num = 5;
    logDRS = log(drs+1);
    logDES = log(des+1);
    IF time='Baseline' THEN numtime = 0-TX_DAYS;
    ELSE IF time='Post_Treatment' THEN numtime=0;
    ELSE IF time='One_Week' THEN numtime=7;
    ELSE IF time='Three_Months' THEN numtime=90;
    ELSE IF time='Six_Months' THEN numtime=180;
RUN;

proc sort data=final_mice;
    by _imputation_ id time_num;
    run;



 %saveall(micecor fcs_output
```

```
          miceinfo micesparams
          varinf final_mice,MI,
      time-as-fact, sas_output_mi, model_descriptions.csv,
      PROC MI ,
      MI, Multiple Imputation Output,
      data_wide, REML );



ods exclude all;
proc mixed data = final_mice method = reml;
by _imputation_ ;
class time_num axis_i_diagnosis;
model IES = time_num Axis_i_Diagnosis
        TX_Days Age/ solution CL outp=Pred;
random intercept  / subject=ID  g;
repeated / subject=ID type=ar(1);
estimate "CHANGE -- Baseline-Post Treatment"
    time_num -1 1 0 0 0 / cl;
estimate "CHANGE -- Post Treatment-1 Week"
    time_num 0 -1 1 0 0 / cl;
estimate "CHANGE -- Post Treatment-3 Months"
    time_num 0 -1 0 1 0 / cl;
estimate "CHANGE -- Post Treatment-6 Months"
    time_num 0 -1 0 0 1 / cl;
estimate "CHANGE -- Baseline-1 Week"
    time_num -1 0 1 0 0 / cl;
```

```
estimate "CHANGE -- 1 Week-3 Months"
    time_num 0 0 -1 1 0 / cl;
estimate "CHANGE -- 1 Week-6 Months"
    time_num 0 0 -1 0 1 / cl;
estimate "CHANGE -- Baseline-3 Months"
    time_num -1 0 0 1 0 /cl;
estimate "CHANGE -- 3 Months-6 Months"
    time_num 0 0 0 -1 1 / cl;
estimate "CHANGE -- Baseline-6 Months"
    time_num -1 0 0 0 1 / cl;
lsmeans time_num;
lsmeans time_num / DIFF;
ods output solutionf=beta_ies estimates=esty_ies
  FitStatistics=fits SolutionF=fixed CovParms=CP
  convergenceStatus=conv modelinfo=mod tests3=tests  D
  imensions=dims lsmeans=lsmeans diffs=diffs; run;



data diffs;
set diffs;
comparison = time_num||' vs '||left(_time_num);
run;
proc print data=diffs; run;
proc sort data=diffs;by comparison _imputation_; run;
ods output parameterestimates=diffs_pool;
proc mianalyze data=diffs;
```

```
    by comparison ;
    modeleffects estimate;
    stderr stderr;run;


proc sort data=lsmeans;by time_num _imputation_; run;
ods output parameterestimates=lsmeans_pool;
proc mianalyze data=lsmeans;
    by time_num;
    modeleffects estimate;
    stderr stderr;run;



proc sort data=esty_ies; by label _imputation_; run;
proc print data=esty_ies;run;



data beta_ies;
   set beta_ies;
   by _imputation_;
   retain group_number;
   if first._imputation_ then group_number + 1;
   _Imputation_ = group_number;
   drop group_number;
run;
```

```
proc sort data=beta_ies; by _imputation_;

proc mianalyze parms(classvar=full)=beta_ies;
class time_num axis_i_diagnosis;
 modeleffects Intercept time_num
    Axis_i_Diagnosis TX_Days Age;
 ods output  parameterestimates=mianalyze_params varianceinfo=varinfo;
run;




/* For the lsmeans output  */
data esty_2_ies; set esty_ies;
    var = stderr**2; run;
proc sort data=esty_2_ies; by  label;
/* The between-MI variance */
proc means data = esty_2_ies mean var; by label;
var estimate; output out=esty_mean
    mean=est_mean var=bvar; run;
/* The within-MI variance  */
proc means data = esty_2_ies mean; by label;
var var; output out=esty_var mean=wvar; run;

proc sort data = esty_mean; by label; run;
```

```
proc sort data = esty_var; by  label; run;
data esty_means_ies;
format model $5.;
merge esty_mean esty_var;
by  label;
model = "IES";
/* Total Variance */
esty_var = wvar + (1 + (1/5))*bvar;
esty_stderr = sqrt(esty_var);
/* Calculate degrees of freedom */
df = (5-1)*((1 + wvar/((1+(1/5))*bvar))**2);
if df > 10000000 then df = 10000000;
else df = df;
t = tinv(0.975, df);
lcl = est_mean - t*esty_stderr;
ucl = est_mean + t*esty_stderr;
run;



 %saveall(fits fixed CP conv mod
        tests Pred dims beta_ies
        estsy_ies mianalyze_params
        varinfo esty_means_ies lmeans
        lsmeans_pool  diffs diffs_pool
        , IES,
      time-as-fact, sas_output_mi,
      model_descriptions.csv,
```

```
      MI LMM for IES,
      IES,
Follow-Up Timepoint as Factor +
Last Timepoint as Factor +
Interaction Between Follow-Up and Last Timepoint +
Primary Diagnosis +
Treatment Days + Age,
Random Intercept for Subject - AR(1)  ,
     final_mice, REML );
```

R Code

```
library(tidyverse)
options(kableExtra.latex.load_packages = FALSE)
library("kableExtra")
library(table1)
library(ggthemes)
library(pander)
library(extrafont)
font_import(pattern = "Times New", prompt = F)  # Import Times New Roman font
theme_set(theme_minimal(base_size = 12,
              base_family = "Times New Roman"))

long_data_timeeint = tibble(read.csv("../../long_data.csv")) %>%
mutate(time = factor(
time,
levels = c(
"Baseline",
"PostTreatment",
"1weekPost",
"3mosPost",
"6mosPost"
),
labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
))  %>%
select(-X) %>%
group_by(ID) %>%
```

```
mutate(last = factor(levels(time)[max(as.integer(time))])) %>%
ungroup() %>%
mutate(Axis.I.Diagnosis = factor(Axis.I.Diagnosis)) %>%
mutate(Axis.I.Diagnosis = factor(
Axis.I.Diagnosis,
levels(Axis.I.Diagnosis),
labels = c(
"Missing",
"PTSD",
"DID",
"DDNOS",
"Adjustment Disorder",
"Adjustment Disorder",
"Adjustment Disorder"
)
)) %>%
droplevels() %>%
mutate(Gender = factor(
Gender,
levels = c(1, 2),
labels = c("Female", "Male")
)) %>%
mutate(HAL = as.factor(HAL)) %>%
mutate(Type.II = as.factor(Type.II)) %>%
mutate(Type.I.S = as.factor(Type.I.S)) %>%
mutate(Type.I.M = as.factor(Type.I.M)) %>%
mutate(ALX.PRS. = as.factor(ALX.PRS.)) %>%
```

```
mutate(ALX..POST = as.factor(ALX..POST)) %>%
mutate(last = factor(
last,
levels =  c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months"),
labels =  c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
))

long_data_bcf <- long_data_timeeint %>%
select(-numtime, -TRS) %>%
pivot_longer(
cols = c(IES, DES, SCL, DRS, TAS),
names_to = "Measure",
values_to = "score"
) %>%
pivot_wider(names_from = time, values_from = score) %>%
rowwise() %>%
mutate(keep_row = if_else(
(
as.character(last) == "Post Treatment" &
`Post Treatment` == Baseline
) |
(as.character(last) == "1 Week" &
`1 Week` == `Post Treatment`) |
(as.character(last) == "1 Week" &
Baseline == `Post Treatment`) |
(as.character(last) == "1 Week" & Baseline == `1 Week`),
TRUE,
```

```
FALSE
)) %>%
mutate(keep_row = ifelse(is.na(keep_row), FALSE, keep_row)) %>%
mutate(keep_row = as.numeric(keep_row)) %>%
mutate(last  = as.numeric(last) - keep_row)  %>%
mutate(last = factor(
last,
levels = 1:5,
labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
)) %>%
pivot_longer(cols = Baseline:`6 Months`,
     values_to = 'score',
     names_to = 'time') %>%
mutate(time = factor(
time,
levels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months"),
labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
)) %>%
drop_na(score) %>%
filter(as.numeric(factor(time)) <= as.numeric(last)) %>%
pivot_wider(names_from = Measure, values_from = score)

tx_shifted = long_data_bcf %>%
select(-keep_row) %>%
mutate(id_num = as.numeric(ID),
is_num = !is.na(id_num))
```

```
# Find the maximum numeric ID
max_id <- max(tx_shifted$id_num, na.rm = TRUE)

# Create a replacement mapping for non-numeric IDs
non_num_ids <- tx_shifted %>% filter(!is_num) %>% distinct(ID)
non_num_ids$new_id <- max_id + seq_along(non_num_ids$ID)

tx_shifted = tx_shifted %>%
left_join(non_num_ids, by = "ID") %>%
mutate(ID = ifelse(is_num, id_num, new_id),
ID = as.integer(ID)  # Convert to integer) %>%
dplyr::select(-is_num, -id_num, -new_id)

tx_shifted_wide =  tx_shifted %>%
 select(
   ID,
   DRS,
   DES,
   TAS,
   SCL,
   IES,
   time,
   TX.DAYS,
   Age,
   Gender,
   Axis.I.Diagnosis,
   HAL,
```

```
    Type.I.S,
    Type.I.M,
    Type.II,
    ALX..POST,
    ALX.PRS.,
    last
  ) %>%
  group_by(ID, time) %>%
  fill(DRS:IES) %>%
  mutate(num_miss = is.na(DRS) + is.na(DES) + is.na(TAS) + is.na(SCL) +
         is.na(IES)) %>%
  slice_min(num_miss) %>%
  ungroup() %>%
  select(-num_miss) %>%
  pivot_wider(names_from = time,
              values_from = c(DRS, DES, IES, SCL, TAS)) %>%
  pivot_longer(DRS_Baseline:`TAS_6 Months`,
               names_to = "measure",
               values_to = "vals")  %>%
  mutate(time = sapply(measure, function(x)
    str_split(x, "_")[[1]][2]))

tx_shifted_vismiss = tx_shifted_wide %>%
  mutate(vals = sapply(vals, function(x)
    ifelse(is.na(x), NA, ""))) %>%
  pivot_wider(names_from = time, values_from = vals) %>%
  select(-measure, -last) %>% unique
```

```
tx_shifted_wide = tx_shifted_wide %>%
 mutate(measure = sapply(measure, function(x)
   str_split(x, "_")[[1]][1])) %>%
 pivot_wider(names_from = time, values_from = vals)

tx_shifted_long = tx_shifted %>%
 pivot_longer(cols = IES:TAS,
              names_to = "measure",
              values_to = "score")

data_wide = tx_shifted %>%
 select(
   ID,
   DRS,
   DES,
   TAS,
   SCL,
   IES,
   time,
   TX.DAYS,
   Age,
   Gender,
   Axis.I.Diagnosis,
   HAL,
   Type.I.S,
```

```
  Type.I.M,
  Type.II,
  ALX..POST,
  ALX.PRS.,
  -last
) %>%
 mutate(time = gsub(" ", "_", time)) %>%
 group_by(ID, time) %>%
 fill(DRS:IES) %>%
 mutate(num_miss = is.na(DRS) + is.na(DES) + is.na(TAS) + is.na(SCL) +
                   is.na(IES)) %>%
 slice_min(num_miss) %>%
 ungroup() %>%
 select(-num_miss) %>%
 pivot_wider(names_from = time,
             values_from = c(DRS, DES, IES, SCL, TAS))
```

```
label(long_data_timeeint$TX.DAYS) = "Treatment Days"
label(long_data_timeeint$Axis.I.Diagnosis) = "Primary Diagnosis"
label(long_data_timeeint$last) = "Dropout Timepoint"

tab_data = long_data_timeeint %>%
 select(ID, TX.DAYS, Age, Axis.I.Diagnosis, last, Gender) %>%
 distinct(ID, .keep_all = T)
```

```
t1kable(
 table1(
   ~ TX.DAYS + Age + Axis.I.Diagnosis + Gender |
     last,
   data = tab_data,
   label = "t1",
   caption = "Demographic and Descriptive Measures By Dropout Timepoint"
 ),
 booktabs = T,
 label = "table1"
) %>%
 kable_styling(latex_options = c("hold_position", "scale_down"))

vars = long_data_bcf %>% select(
 ID,
 time,
 DRS,
 DES,
 TAS,
 SCL,
 IES,
 TX.DAYS,
 Age,
 Gender,
 Axis.I.Diagnosis,
 HAL,
```

```
 Type.I.S,
 Type.I.M,
 Type.II,
 ALX..POST,
 ALX.PRS.
) %>%
 colnames


var_tibble = tibble(
 vname = vars,
 name = c(
   "ID",
   "Follow-Up Timepoint",
   "Dissociative Regression Scale",
   "Dissociative Experience Scale",
   "Toronto Alexithymia Scale",
   "Symptom Checklist 45",
   "Impact of Events Scale",
   "Treatment Days",
   "Age",
   "Gender",
   "Primary Diagnosis",
   "Hallucinations",
   "Type I S",
   "Type I M",
   "Type II",
   "Alexithymia Post Treatment",
```

```
   "Alexithymia Pre Treatment"
 ),
 desc = c(
"Unique Subject IDs",
"Factor representing follow up timepoint which measure was collected",
"Average of DRS symptom rating",
"Average of DES symptom rating",
"Sum of TAS symptom rating",
"Sum of SCL symptom rating",
"Sum of IES symptom rating",
"Number of treatment days",
"Numeric Age of Subject",
"Gender (Male or Female)",
"Factor representing Primary Diagnosis Endorsement",
"Boolean representing hallucination endorsement based on responses to the DES and SCL",
"Boolean representing Single Trauma type",
"Boolean representing multiple different traumas",
"Boolean representing repeated traumas of the same type",
"Boolean representing if alexithymia is present post-treatment",
"Boolean representing if alexithymia is present before treatment"
 )
)

var_tibble %>%
 select(name, desc) %>%
 knitr::kable(
   caption = "Description of All Available Variables",
```

```
   col.names = c("Name", "Description"),
   booktabs = TRUE,
   label = "tabdesc"
 ) %>%
 kable_styling(latex_options = c("hold_position")) %>%
 column_spec(c(1), width = "2in") %>%
 column_spec(2, width = "4.5in")



library(readxl)
library(broom)
library(tidyverse)
library(formatR)
library(parameters)
library(stringr)
library(gridExtra)
## load and clean data
make_data <- function(datapath) {
 temp = list.files(pattern = "*.csv", path = datapath)
 main_df = tibble()
 main_flag = TRUE
 for (file in temp) {
   temp_df = tibble(read.csv(paste(datapath, file, sep = "/")))
   if (!('Year' %in% colnames(temp_df))) {
     yr = str_extract(file, "[0-9]{4}")
     temp_df = temp_df %>%
       mutate(Year = rep(as.integer(yr), nrow(temp_df)))
```

```
  }
  temp_df = temp_df  %>%
    mutate(across(everything(), as.character))

  main_df = bind_rows(main_df, temp_df)
}
main_df = main_df %>%
  distinct()
for (col in colnames(main_df)) {
  tryCatch({
    main_df[col] = as.numeric(unlist(main_df[col]))
  }
  , warning = function(w) {

  })
}
main_df = main_df %>%
  drop_na(ID) %>%
  mutate(
    Gender_Update = case_when(
      is.na(Gender) ~ NA_real_,
      Gender == "F" ~ 1,
      Gender == "M" ~ 2,
      Gender == "1" ~ 1,
      Gender == "2" ~ 2,
      TRUE ~ as.numeric(Gender)
    )
```

```
)  %>%
mutate(Gender = Gender_Update) %>%
select(-Gender_Update) %>%
mutate(Improved = case_when(
  Improved == '' ~ "Missing",
  is.na(Improved) ~ "Missing",
  TRUE ~ Improved
)) %>%
mutate(Improved = as.factor(Improved)) %>%
mutate(
  Axis.I.Diagnosis = case_when(
    Axis.I.Diagnosis == "1" ~ "PTSD",
    Axis.I.Diagnosis == "2" ~ "DDNOS",
    Axis.I.Diagnosis == "3" ~ "DID",
    Axis.I.Diagnosis == "" ~ "None",
    TRUE ~ Axis.I.Diagnosis
  )
)   %>%
mutate(Gender = as.factor(Gender)) %>%
mutate(Treatment = X) %>%
select(-X) %>%
select(-year, -Diag) %>%
relocate(where(is.factor)) %>%
relocate(Treatment) %>%
relocate(Gender) %>%
relocate(ID) %>%
relocate(where(is.logical), .after = Improved) %>%
```

```
   relocate(c(Year, Age), .after = Gender)



 return(main_df)
}

#indentify and rename duplicates
identify_duplicates <- function(df) {
 df %>%
   arrange(ID, Age, Year) %>%
   group_by(ID) %>%
   mutate(diff_age = max(Age) - min(Age)) %>%
   mutate(diff_yr = max(Year) - min(Year)) %>%
   ungroup() %>%
   mutate(
     IDnew = case_when(
       diff_age == 0 ~ as.character(ID),
       is.na(diff_age) ~ as.character(ID),
       diff_age >= diff_yr + 1 ~ paste(ID, Age, sep = ("-")),
       diff_age <= diff_yr - 1 ~ paste(ID, Age, sep = ("-")),
       TRUE ~ as.character(ID)
     )
   ) %>%
   select(-diff_age, -diff_yr) %>%
   mutate(Axis.I.Diagnosis = as.factor(Axis.I.Diagnosis))  %>%
   return()
}
```

```
# load all data
itrdata_dups = make_data("../data_sheets-copy")
itrdata_dups_id = identify_duplicates(itrdata_dups)

itrdata_dupsrm = itrdata_dups_id %>%
 mutate(ID = IDnew) %>%
 select(-IDnew)

a1_diag = levels(itrdata_dupsrm$Axis.I.Diagnosis)
itrdata = itrdata_dupsrm %>%
 mutate(Axis.I.Diagnosis = ifelse(ID == '529' |
                                    ID == '187', a1_diag[9],
                                  levels(Axis.I.Diagnosis)[Axis.I.Diagnosis]))
itrdata = itrdata %>% mutate(Axis.I.Diagnosis = as.factor(Axis.I.Diagnosis))
a1_diag = levels(itrdata$Axis.I.Diagnosis)


itrdata  = itrdata %>%
 distinct(across(-Year), .keep_all = TRUE)


itrdata_dups %>%
 arrange(ID, Age, Year) %>%
 group_by(ID) %>%
 mutate(diff_age = max(Age) - min(Age)) %>%
```

```
 mutate(diff_yr = max(Year) - min(Year)) %>%
 ungroup() %>%
 relocate(diff_age, diff_yr, .after = ID) %>%
 filter(diff_age > 1) %>%
 select(-Year) %>%
 distinct() %>%
 head() %>%
 select(ID, Gender,  Age , TX.DAYS , Axis.I.Diagnosis) %>%
 knitr::kable(
   caption = "Subset of Non-Unique IDs Identified by Discrepancies in Age",
   col.names = c('ID', 'Gender',  'Age' , 'Treatment Days' , 'Primary Diagnosis'),
   booktabs = TRUE,
   label = "tab2"
 ) %>%
 kable_styling(latex_options = c("hold_position"))



itrdata_dups_id %>%
 arrange(ID, Age, Year) %>%
 group_by(ID) %>%
 mutate(diff_age = max(Age) - min(Age)) %>%
 mutate(diff_yr = max(Year) - min(Year)) %>%
 ungroup() %>%
 relocate(diff_age, diff_yr, .after = ID) %>%
 relocate(IDnew, .after = ID) %>%
 filter(diff_age > 1) %>%
```

```
 filter(diff_age - diff_yr > 0) %>%
 select(-Year) %>%
 distinct() %>%
 head() %>%
 select(ID, IDnew, Gender, Age , TX.DAYS , Axis.I.Diagnosis) %>%
 knitr::kable(
 caption =
 "Subset of Modified Non-Unique IDs Identified by Discrepancies in Age",
 col.names = c(
   'Old ID',
   'New ID',
   'Gender',
   'Age' ,
   'Treatment Days' ,
   'Primary Diagnosis'
 ),
 booktabs = TRUE,
 label = "tab3"
 ) %>%
 kable_styling(latex_options = c("hold_position"))



itrdata_dupsrm %>%
 distinct(ID, Axis.I.Diagnosis, .keep_all = T) %>%
 group_by(ID) %>%
 filter(n() > 1)  %>%
 ungroup() %>%
```

```
 select(ID, Gender, Age , Axis.I.Diagnosis) %>%
 knitr::kable(
   caption = "Identification of Non-Unique Primary Diagnoses by ID",
   col.names = c('ID', 'Gender', 'Age', 'Primary Diagnosis'),
   booktabs = TRUE,
   label = "tab4"
 ) %>%
 kable_styling(latex_options = c("hold_position"))
```

```
ggplot_timeint = long_data_timeeint %>%
 pivot_longer(
   cols = c(DRS, DES, TAS, SCL, IES),
   names_to = "Assessment",
   values_to = "Score"
 )

ggplot_timeint %>%
 ggplot(aes(x = time, y = Score)) +
 geom_line(aes(group = ID, color = last), alpha = .3) +
 geom_boxplot(
   data = filter(ggplot_timeint, as.character(last) == "Baseline"),
   aes(color = last),
   alpha = .3
 ) +
```

```
 facet_grid(Assessment ~ last, scales = "free_x") +
 ylim(c(10, 75)) +
 ylab("Score") +
 xlab("Follow-Up Timepoint") +
 theme(legend.position = "none",
       axis.text.x = element_text(angle = 45, hjust = 1)) +
 scale_color_manual(values = ggpubr::get_palette('npg', 5))



ggplot_long_data_bcf = long_data_bcf %>%
 pivot_longer(
   cols = c(DRS, DES, TAS, SCL, IES),
   names_to = "Assessment",
   values_to = "Score"
 )

ggplot_long_data_bcf = ggplot_long_data_bcf %>%
 left_join(
   ggplot_long_data_bcf %>% select(ID, last) %>% unique %>%
     group_by(last) %>% summarise(N = n()) %>%
     mutate(Label = paste0(as.character(last), '\nn = ', N)),
   by = join_by(last)
 ) %>%
 arrange(last) %>%
 mutate(Label = factor(
   Label,
   levels = c(
```

```
    unique(Label)[grepl('Baseline', unique(Label))],
    unique(Label)[grepl('Post', unique(Label))],
    unique(Label)[grepl('1 Week', unique(Label))],
    unique(Label)[grepl('3 Months', unique(Label))],
    unique(Label)[grepl('6 Months', unique(Label))]
  )
))

ggplot() +
 geom_line(data = ggplot_long_data_bcf,
           aes(
             group = ID,
             color = last,
             x = time,
             y = Score
           ),
           alpha = .3) +
 facet_grid(Assessment ~ Label, scales = "free_x") +
 geom_boxplot(
   data = subset(ggplot_long_data_bcf,
                 as.character(last) == "Baseline"),
   aes(color = last, x = time, y = Score)
 ) +
 geom_line(
   data = subset(ggplot_long_data_bcf,
                 as.character(last) == "1 Week"),
   aes(
```

```
    group = ID,
    color = last,
    x = time,
    y = Score
  )
) +
ylim(c(10, 75)) +
ylab("Score") +
xlab("Follow-Up Timepoint") +
theme(legend.position = 'none',
      axis.text.x = element_text(angle = 45, hjust = 1)) +
scale_color_manual(values = ggpubr::get_palette('npg', 5))


ggplot_long_data_bcf = long_data_bcf %>%
 pivot_longer(
   cols = c(DRS, DES, TAS, SCL, IES),
   names_to = "Assessment",
   values_to = "Score"
 ) %>%
 mutate(time = factor(
   time,
   labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
 ))
ggplot_long_data_bcf = ggplot_long_data_bcf %>%
 left_join(
   ggplot_long_data_bcf %>% select(ID, last) %>% unique %>%
```

```
    group_by(last) %>% summarise(N =n()) %>%
    mutate(Label = paste0(as.character(last), '\nn = ', N)),
  by = join_by(last)
) %>%
arrange(last) %>%
mutate(Label = factor(
  Label,
  levels = c(
    unique(Label)[grepl('Baseline', unique(Label))],
    unique(Label)[grepl('Post', unique(Label))],
    unique(Label)[grepl('1 Week', unique(Label))],
    unique(Label)[grepl('3 Months', unique(Label))],
    unique(Label)[grepl('6 Months', unique(Label))]
  )
)) %>%
mutate(numtime = ifelse(
  as.character(time) == 'Baseline',
  0,
  ifelse(
    as.character(time) == 'Post Treatment',
    TX.DAYS,
    ifelse(
      as.character(time) == '1 Week',
      TX.DAYS + 7,
      ifelse(as.character(time) ==
               '3 Months',
             TX.DAYS + 90, TX.DAYS +
```

```
                180)
    )
  )
))

ggplot() +
 geom_line(data = ggplot_long_data_bcf,
           aes(
             group = ID,
             color = last,
             x = numtime,
             y = Score
           ),
           alpha = .3) +
 facet_grid(Assessment ~ Label, scales = "free_x") +
 geom_boxplot(
   data = subset(ggplot_long_data_bcf,
                 as.character(last) == "Baseline"),
   aes(color = last, x = numtime, y = Score)
 )  +
 geom_line(
   data = subset(ggplot_long_data_bcf,
                 as.character(last) == "1 Week"),
   aes(
     group = ID,
     color = last,
     x = numtime,
```

```
    y = Score
  )
 ) +
 ylim(c(10, 75)) +
 ylab("Score") +
 xlab("Days Since Baseline") +
 theme(legend.position = 'none') +
 scale_color_manual(values = ggpubr::get_palette('npg', 5)) +
 scale_x_continuous(breaks = c(0, 40, 100, 190))



ggplot_long_data_bcf = long_data_bcf %>%
 pivot_longer(
   cols = c(DRS, DES, TAS, SCL, IES),
   names_to = "Assessment",
   values_to = "Score"
 ) %>%
 mutate(time = factor(
   time,
   labels = c("Baseline", "Post Treatment", "1 Week",
              "3 Months", "6 Months")
 ))
ggplot_long_data_bcf = ggplot_long_data_bcf %>%
 left_join(
   ggplot_long_data_bcf %>% select(ID, last) %>% unique %>%
     group_by(last) %>% summarise(N =n()) %>%
     mutate(Label = paste0(as.character(last), '\nn = ', N)),
```

```
  by = join_by(last)
) %>%
arrange(last) %>%
mutate(Label = factor(
  Label,
  levels = c(
    unique(Label)[grepl('Baseline', unique(Label))],
    unique(Label)[grepl('Post', unique(Label))],
    unique(Label)[grepl('1 Week', unique(Label))],
    unique(Label)[grepl('3 Months', unique(Label))],
    unique(Label)[grepl('6 Months', unique(Label))]
  )
)) %>%
mutate(numtime = ifelse(
  as.character(time) == 'Baseline',
  0,
  ifelse(
    as.character(time) == 'Post Treatment',
    TX.DAYS,
    ifelse(
      as.character(time) == '1 Week',
      TX.DAYS + 7,
      ifelse(as.character(time) ==
               '3 Months',
             TX.DAYS + 90, TX.DAYS +
               180)
    )
```

```
  )
)) %>%
 mutate(numtime = numtime - TX.DAYS)


ggplot() +
 geom_line(data = ggplot_long_data_bcf,
           aes(
             group = ID,
             color = last,
             x = numtime,
             y = Score
           ),
           alpha = .3) +
 facet_grid(Assessment ~ Label, scales = "free_x") +
 geom_boxplot(
   data = subset(ggplot_long_data_bcf,
                 as.character(last) == "Baseline"),
   aes(color = last, x = numtime, y = Score)
 )  +
 geom_line(
   data = filter(ggplot_long_data_bcf,
                 as.character(last) == "1 Week"),
   aes(
     group = ID,
     color = last,
     x = numtime,
     y = Score
```

```
  )
 ) +
 ylim(c(10, 75)) +
 ylab("Score") +
 xlab("Days Since Post Treatment") +
 theme(legend.position = 'none', axis.text.x = element_text(size = 8)) +
 scale_color_manual(values = ggpubr::get_palette('npg', 5)) +
 scale_x_continuous(breaks = c(0, 30, 90, 180))



ggplot_long_data_bcf = long_data_bcf %>%
 pivot_longer(
   cols = c(DRS, DES, TAS, SCL, IES),
   names_to = "Assessment",
   values_to = "Score"
 ) %>%
 mutate(time = factor(
   time,
   labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
 ))
ggplot_long_data_bcf = ggplot_long_data_bcf %>%
 left_join(
   ggplot_long_data_bcf %>% select(ID, last) %>% unique %>%
     group_by(last) %>% summarise(N =n()) %>%
     mutate(Label = paste0(as.character(last), '\nn = ', N)),
   by = join_by(last)
 ) %>%
```

```
arrange(last) %>%
mutate(Label = factor(
  Label,
  levels = c(
    unique(Label)[grepl('Baseline', unique(Label))],
    unique(Label)[grepl('Post', unique(Label))],
    unique(Label)[grepl('1 Week', unique(Label))],
    unique(Label)[grepl('3 Months', unique(Label))],
    unique(Label)[grepl('6 Months', unique(Label))]
  )
)) %>%
mutate(numtime = ifelse(
  as.character(time) == 'Baseline',
  0,
  ifelse(
    as.character(time) == 'Post Treatment',
    TX.DAYS,
    ifelse(
      as.character(time) == '1 Week',
      TX.DAYS + 7,
      ifelse(as.character(time) ==
               '3 Months',
             TX.DAYS + 90, TX.DAYS +
               180)
    )
  )
)) %>%
```

```
 mutate(numtime = numtime - TX.DAYS)



ggplot_long_data_bcf  %>%
 ggplot(aes(x = numtime, y = Score)) +
 geom_point(alpha = 0) +
 geom_line(alpha = .1, aes(color = as.factor(ID), group = as.factor(ID))) +
 geom_smooth(color = "#3C5488FF") +
 facet_wrap( ~ Assessment, scales = "free") +
 theme(legend.position = "none") +
 xlab("Days Since Post Treatment") +
 ylab("Measure Score") +
 scale_color_manual(values = ggpubr::get_palette('npg', length(unique(
   ggplot_long_data_bcf$ID
 )))) +
 scale_x_continuous(breaks = c(-25, 0, 90, 180))



ggplot_long_data_bcf = long_data_bcf %>%
 pivot_longer(
   cols = c(DRS, DES, TAS, SCL, IES),
   names_to = "Assessment",
   values_to = "Score"
 ) %>%
 mutate(time = factor(
```

```
    time,
    labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
  ))
ggplot_long_data_bcf = ggplot_long_data_bcf %>%
  left_join(
    ggplot_long_data_bcf %>% select(ID, last) %>% unique %>%
      group_by(last) %>% summarise(N = n()) %>%
      mutate(Label = paste0(as.character(last), '\nn = ', N)),
    by = join_by(last)
  ) %>%
  arrange(last) %>%
  mutate(Label = factor(
    Label,
    levels = c(
      unique(Label)[grepl('Baseline', unique(Label))],
      unique(Label)[grepl('Post', unique(Label))],
      unique(Label)[grepl('1 Week', unique(Label))],
      unique(Label)[grepl('3 Months', unique(Label))],
      unique(Label)[grepl('6 Months', unique(Label))]
    )
  )) %>%
  mutate(numtime = ifelse(
    as.character(time) == 'Baseline',
    0,
    ifelse(
      as.character(time) == 'Post Treatment',
      TX.DAYS,
```

```
    ifelse(
      as.character(time) == '1 Week',
      TX.DAYS + 7,
      ifelse(as.character(time) ==
               '3 Months',
             TX.DAYS + 90, TX.DAYS +
               180)
    )
  )
)) %>%
 mutate(numtime = numtime - TX.DAYS)



ggplot_long_data_bcf %>%
 mutate(logscore = log(Score + 1)) %>%
 pivot_longer(cols = c("Score", "logscore"), names_to = "score_val") %>%
 mutate(score_val = ifelse(score_val == "Score", "Untransformed",
    "Log Transformed")) %>%
 ggplot(aes(x = value)) +
 geom_histogram() +
 facet_grid(Assessment ~ score_val , scales = "free") +
 xlab("Measure Score") +
 ylab("Count")
```

```
colnames(tx_shifted_vismiss) =  c(
 'ID',
 'Treatment Days',
 'Age',
 'Gender',
 'Primary Diagnosis',
 'Hallucination',
 'Type I S',
 'Type I M',
 'Type II',
 'Alexithymia Pre',
 'Alexithymia Post',
 'Measure',
 'Baseline',
 'Post Treatment',
 '1 Week',
 '3 Months',
 '6 Months'
)



naniar::gg_miss_upset(tx_shifted_vismiss, nsets = 8)



tlist = list()
```

```
for (m in c("DES", "DRS", "TAS", "IES", "SCL")) {
 tx_shifted_wide_DRS = tx_shifted_wide %>%
   filter(measure == m) %>%
   mutate(post_diff = `Post Treatment` - Baseline)
 mod1 = lm(Baseline ~ last , data = tx_shifted_wide_DRS)
 mod2 = lm(`Post Treatment` ~ last, data = tx_shifted_wide_DRS)
 mod3 = lm(post_diff ~ last, data = tx_shifted_wide_DRS)

 baseline_aov = aov(Baseline ~ last, tx_shifted_wide_DRS)

 post_aov = aov(`Post Treatment` ~ last, tx_shifted_wide_DRS)

 m_mod = broom::tidy(mod1) %>%
   mutate(mod = 'lmb') %>%
   bind_rows(mutate(broom::tidy(mod2), mod = "lmp")) %>%
   bind_rows(mutate(broom::tidy(mod3), mod = "lmd")) %>%
   filter(term != "(Intercept)") %>%
   mutate(term = ifelse(term == "`Post Treatment`", "Post Treatment", term)) %>%
   mutate(term = ifelse(term == "last", "Dropout Timepoint", term)) %>%
   mutate(mod = ifelse(
     mod == "lmb",
     "Baseline Score",
     ifelse(
       mod == "lmp",
       "Post-Treatment Score",
       ifelse(mod == "lmd", "Change from Baseline to Post",
              "")
```

```
    )
  )) %>%
  mutate(term = gsub("last", "", term)) %>%
  relocate(mod) %>%
  mutate(measure = m)
 tlist[[m]] = m_mod
}

plist = tlist
cat("\\begin{table}[!htbp]\n")
cat("\\small\n")
for (i in names(tlist)) {
 # Begin subtable environment
 cat("\\begin{subtable}{.5\\textwidth}\n")
 cat("\\centering\n")
 plist[[i]] = tlist[[i]] %>%
  select(term, estimate, p.value) %>%
  mutate(estimate = round(estimate, 3),
         p.value = round(p.value, 3)) %>%
  mutate(p.value = format.pval(p.value)) %>%
  knitr::kable(booktabs = T,
               col.names = c("Term", "Estimate", "P-Value")) %>%
  pack_rows(index = table((tlist[[i]]$mod)))
 plist[[i]] %>% print
 cat("\\caption{", i, "}\n")
 cat(paste("\\label{subtab:", i, "}", sep = ""))
 cat("\\end{subtable}\n")
```

```
 if (match(i, names(tlist)) %% 2 == 0) {
   cat("\\\\")
 }
}
cat("\\caption{Regression Models}\n")
cat("\\label{tab:regmods}")

# End the table environment
cat("\\end{table}\n")



fixed_columns = c(
 "Age",
 "Gender",
 "Axis.I.Diagnosis",
 "TX.DAYS",
 "HAL",
 "Type.I.S",
 "Type.I.M",
 "Type.II",
 "ALX.PRS.",
 "ALX..POST"
)
outcomes = c("DRS", "TAS", "SCL", "IES")
predictors = fixed_columns
mydata = tx_shifted
```

```
predictors_test = tibble(vname = character(),
                         test = character(),
                         pval = numeric())

# loop through predictors and run appropriate test
for (i in predictors) {
 if (is.numeric(mydata[[i]]) | is.integer(mydata[[i]])) {
   # run ANOVA
   model <-
     lm(as.formula(paste(i, "last", sep = " ~ ")), data = mydata[c(i, "last")])
   anova_res <- anova(model)
   # output results
   predictors_test = predictors_test %>% add_row(vname = i,
                                                 test = "ANOVA",
                                                 pval = anova_res$`Pr(>F)`[1])
 } else  {
   # run chi-squared test
   test = "Chi-Squared"
   result <- tryCatch(
     chisq.test(table(mydata[[i]], mydata$last)),
     warning = function(w)
       NULL,
     error = function(e)
       NULL
   )
   # if chi-squared test throws a warning, use Fisher's exact test instead
   if (is.null(result)) {
```

```
    result <-
      fisher.test(table(mydata[[i]], mydata$last), simulate.p.value = T)
    test = "Fisher's Exact"
  }
  # output results
  predictors_test = predictors_test %>% add_row(vname = i,
                                                test = test,
                                                pval = result$p.value)
 }
}

predictors_test %>%
 mutate(pval = round(pval, 3)) %>%
 mutate(vname = factor(
   vname,
   labels = c(
     'Age',
     'Gender',
     'Primary Diagnosis',
     'Treatment Days',
     'Hallucination',
     'Type I S',
     'Type I M',
     'Type II',
     'Alexithymia Pre.',
     'Alexithymia Post'
   )
```

```
 )) %>%
 mutate(pval = base::format.pval(pval)) %>%
 knitr::kable(
   col.names = c("Variable Name", "Test", "P-Value"),
   booktabs = T,
   caption = "Correlation between Demographic Covariates and Dropout Timepoint",
   label = "cordemo"
 )
```

```
n = 0
total = c()
for (col in colnames(data_wide)[12:36]) {
 n = n + 1
 total = c(total, sum(is.na(data_wide[[col]])))
}

# proportion of missingness across all time points and measures
sum(total) / n / nrow(data_wide) # = 0.4937143

naniar::mcar_test(data_wide) # p = 1, statistic = 1177.471 df = 1706
```

```
outcomes = c("IES", "TAS", "SCL", "logDES", "logDRS")
```

```
allmixed_files = tibble()
for (oc in outcomes) {
 am_files = list.files(paste("../../../allmixed/",
      oc, sep = ""))
 am_files = am_files[grepl("xlsx$", am_files)]
 am_files = am_files[!grepl("^\\~\\$", am_files)]

 add_files = tibble(files = am_files) %>%
   mutate(outcome = oc)
 allmixed_files = bind_rows(allmixed_files, add_files)
}



cov_select = allmixed_files %>%
 filter(grepl("cov_select", files))
cov_selection_op = tibble()
for (file in cov_select$files) {
 row = cov_select[cov_select$files == file, ]
 add_cs = readxl::read_excel(paste("../../../allmixed",
      row$outcome, file, sep = "/")) %>%
   mutate(outcome = row$outcome)
 cov_selection_op = bind_rows(cov_selection_op, add_cs)
}

cov_selection_op = cov_selection_op %>%
```

```
 mutate(outcome = sapply(outcome, function(oc)
   ifelse(
     grepl("log", oc), paste("log(",
        str_split(oc, "log")[[1]][2], ")", sep = ""),
     oc
   ))) %>%
 mutate(COVARIANCE = factor(
   COVARIANCE,
   levels = c("ar(1)", "cs", "sp(exp)(nu", "sp(pow)(ti", "sp(exp)(ti"),
   labels = c(
     "AR(1)",
     "Compound Symmetric",
     "Exponential Spatial Time as Days",
     "Spatial Power Follow-up Timepoint",
     "Exponential Spatial Follow-Up Timepoint"
   )
 )) %>%
 mutate(AICC = round(AICC, 2)) %>%
 select(COVARIANCE, AICC, outcome)  %>%
 rename(`Correlation Structure` = COVARIANCE,
        Outcome = outcome) %>%
 mutate(Outcome = factor(Outcome)) %>%
 arrange(Outcome)

cov_selection_op %>%
 select(-Outcome) %>%
 knitr::kable(booktabs = TRUE,
```

```
              caption = "Initial Correlation Structure Selection Output",
              label = "corrinit") %>%
 pack_rows(index = table(cov_selection_op$Outcome))



fixed = allmixed_files %>%
 filter(grepl("fixed", files))
fixed_select = tibble()
for (file in fixed$files) {
 row = fixed[fixed$files == file, ]
 fixed_add = readxl::read_excel(paste("../../../allmixed",
        row$outcome, file, sep = "/")) %>%
   mutate(outcome = row$outcome)
 fixed_select = bind_rows(fixed_select, fixed_add)
}



fixed_report = fixed_select %>%
 mutate(outcome = sapply(outcome, function(oc)
   ifelse(
     grepl("log", oc), paste("log(",
        str_split(oc, "log")[[1]][2], ")", sep = ""),
     oc
   ))) %>%
```

```
group_by(outcome) %>%
mutate(mod_num = row_number()) %>%
ungroup() %>%
relocate(mod_num) %>%
arrange(outcome) %>%
mutate(FIXEDT = sapply(FIXEDT, function(x)
  gsub("[0-9]+[ ]*", "", x))) %>%
mutate(FIXEDT = str_replace_all(FIXEDT, "TX_D$", "TX_DAYS Age")) %>%
mutate(FIXEDT =
         sapply(FIXEDT, function(fix)
           ifelse(
             grepl("REF", fix),
             list(c("Base Model")),
             str_split(fix, " ")
           ))) %>%
unnest(FIXEDT) %>%
mutate(
  FIXEDT = str_replace_all(FIXEDT, "TX.*", "Treatment Days"),
  FIXEDT = str_replace_all(FIXEDT, "Trauma.*", "Trauma Type"),
  FIXEDT = str_replace_all(FIXEDT, "Axis.*", "Primary Diagnosis")
) %>%
group_by(mod_num, outcome) %>%
mutate(`Fixed Effects` = paste0(FIXEDT, collapse = " + ")) %>%
select(-FIXEDT) %>%
ungroup() %>%
ungroup() %>%
unique() %>%
```

```
 relocate(`Fixed Effects`) %>%
 filter(DELTAAIC_C < 5 | DELTAMDL < 5)  %>%
 mutate(outcome = factor(outcome, levels = c(
   "IES", "SCL", "TAS", "log(DES)", "log(DRS)"
 ))) %>%
 arrange(outcome, MDL, AIC_C) %>%
 relocate(outcome, MDL, AIC_C)
fixed_report %>%
 select(`Fixed Effects`, AIC_C, `MDL`) %>%
 knitr::kable(booktabs = TRUE,
              caption = "Fixed Effect Model Selection Output",
              label = "femodselect") %>%
 pack_rows(index = table(fixed_report$outcome))




outcomes = c("IES", "TAS", "SCL", "logDES", "logDRS")

corr_compare = read.csv("../corr_compare/model_descriptions.csv") %>%
 mutate(df = purrr::map(filename, function(x)
   read.csv(paste(
     "../corr_compare/", x, sep = ""
   )))) %>%
 mutate(desc = purrr::map(df, function(x)
   x[[1]])) %>%
 mutate(score = purrr::map(df, function(x)
```

```
  x[[2]])) %>%
unnest(desc, score) %>%
select(-df, -fixed_effects) %>%
mutate(outcome = sapply(outcome, function(oc)
  ifelse(
    grepl("log", oc), paste("log(",
       str_split(oc, "log")[[1]][2], ")", sep = ""),
    oc
  ))) %>%
pivot_wider(names_from = desc, values_from = score) %>%
mutate(`Correlation Structure` = sapply(model_description, function(corr)
  ifelse(
    grepl("pow", corr),
    "Spatial Power",
    ifelse(
      grepl("exp", corr),
      "Exponential Spatial",
      "First-Order Autoregressive"
    )
  ))) %>%
mutate(CS_var = sapply(model_description, function(corr)
  ifelse(grepl("numtime", corr),
         "Time as Days",
         "Time as Class"))) %>%
mutate(`Correlation Structure` = paste(`Correlation Structure`,
                                       CS_var, sep = " for ")) %>%
select(-notes,
```

```
        -model_description,
        -CS_var,
        -filename,
        -dataframe,
        -date_added,
        -time_added) %>%
 relocate(outcome, `Correlation Structure`) %>%
 arrange(outcome, random_effects, `AIC (Smaller is Better)`) %>%
 unique()


for (oc in unique(corr_compare$outcome)) {
 corr_compare %>%
   filter(outcome == oc) %>%
   arrange(outcome, random_effects, `AIC (Smaller is Better)`) %>%
   rename(AIC = `AIC (Smaller is Better)`,
          AICC = `AICC (Smaller is Better)`,
          BIC = `BIC (Smaller is Better)`) %>%
   select(-outcome, -random_effects, -`-2 Log Likelihood`) %>%
   knitr::kable(
     booktabs = TRUE,
     caption = paste("Correlation and Random Effect Selection for", oc),
     label = paste(oc, "corrconfirm2", sep = "")
   ) %>%
   pack_rows(index = table(
      filter(corr_compare, outcome == oc)$random_effects)
      ) %>%
```

```
   print

}
```

```
folders = c("sas_output", "sas_output_mi")
files = tibble()
for (foldername in folders) {
 files = rbind(files, mutate(read.csv(
   paste0("../", foldername, "/model_descriptions.csv", sep = "")
 ), folder = foldername))
}
files = files %>%
 mutate(model_description = as.factor(model_description)) %>%
 mutate(output = sapply(filename, function(x)
   unlist(str_split(x, "_")[[1]][1]))) %>%
 mutate(outcome = as.factor(outcome))
```

```
diffs =
 files %>%
 filter(grepl("diffs.*reg|diffs_pool.*", filename))

diffs_dat = diffs %>%
```

```
 mutate(path = paste('..', folder, filename, sep = "/")) %>%
 mutate(df = purrr::map(path, function(x) {
   read.csv(as.character(x))
 })) %>%
 unnest(df)


diffs_entire =  diffs_dat %>%
 select(
   -filename,
   -model_description,
   -fixed_effects,
   -random_effects,
   -dataframe,
   -notes,
   -date_added,
   -time_added,
   -output,
   -path
 ) %>%
 mutate(time_num = purrr::map2(time_num, comparison, function(time, comp)
   ifelse(is.na(time),
          as.numeric(str_split(comp, "vs")[[1]][1]), time))) %>%
 mutate(X_time_num = purrr::map2(X_time_num, comparison, function(time, comp)
   ifelse(is.na(time),
          as.numeric(str_split(comp, "vs")[[1]][2]), time))) %>%
 unnest(time_num, X_time_num) %>%
 mutate(time_num = factor(
```

```
  time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(X_time_num = factor(
  X_time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
                      "Available Case")) %>%
mutate(Estimate = Estimate * (-1)) %>%
mutate(Lower = ifelse(
  is.na(LCLMean),
  Estimate - (1.96 * StdErr),
  Estimate - (1.96 * StdErr)
)) %>%
mutate(Upper = ifelse(
  is.na(UCLMean),
  Estimate + (1.96 * StdErr),
  Estimate + (1.96 * StdErr)
)) %>%
select(outcome,
       Model,
       time_num,
```

```
        X_time_num,
        Estimate,
        Lower,
        Upper,
        Probt) %>%
mutate(
  Estimate = ifelse(
    grepl("log", outcome),
    paste(as.character(round((
      exp(Estimate) - 1
    ) * 100, 0)), "%",
    sep = ""),
    as.character(round(Estimate, 2))
  ),
  Upper = ifelse(
    grepl("log", outcome),
    paste(as.character(round((
      exp(Upper) - 1
    ) * 100, 0)), "%",
    sep = ""),
    as.character(round(Upper, 2))
  ),
  Lower = ifelse(
    grepl("log", outcome),
    paste(as.character(round((
      exp(Lower) - 1
    ) * 100, 0)), "%",
```

```
    sep = ""),
    as.character(round(Lower, 2))
  )
) %>%
mutate(Probt = ifelse(
  grepl("<", Probt),
  paste("p", Probt, sep = ""),
  paste("p=", Probt, sep = "")
)) %>%
mutate(time_num = paste("Change From", as.character(time_num))) %>%
mutate(time_num = factor(
  time_num,
  levels = c(
    "Change From Baseline",
    "Change From Post Treatment",
    "Change From 1 Week",
    "Change From 3 Months"
  )
)) %>%
filter(grepl("Baseline", as.character(time_num))) %>%
filter(grepl("6", X_time_num)) %>%
mutate(outcome = sapply(outcome, function(oc)
  ifelse(
    grepl("log", oc), str_split(oc, "log")[[1]][2],
    as.character(oc)
  )))  %>% arrange(outcome, Model)
```

```
diffs_entire %>%
 select(-outcome, -time_num, -X_time_num) %>%
 mutate(`95% Confidence Interval` = paste("(", Lower,
                                          ", ", Upper,
                                          ")", sep = "")) %>%
 select(-Lower, -Upper) %>%
 relocate(Model, Estimate, `95% Confidence Interval`, Probt) %>%
 rename(`P-Value` = Probt) %>%
 knitr::kable(booktabs = TRUE,
  caption =
    "Estimated change from Baseline to 6 Months For all Measures by Model",
  label = "changewhole") %>%
 pack_rows(index = table(diffs_entire$outcome))
## IES

ests =
 files %>%
 filter(grepl("^ests.*reg", filename)) %>%
 filter(grepl("13", date_added))

conts = files %>%
 filter(grepl("contrasts", filename)) %>%
 filter(grepl("13", date_added))

conts_dat = conts %>%
 mutate(path = paste('..', folder, filename, sep = "/")) %>%
 mutate(df = purrr::map(path, function(x) {
```

```
    read.csv(as.character(x))
 })) %>%
 unnest(df)
ests_dat = ests %>%
 mutate(path = paste('..', folder, filename, sep = "/")) %>%
 mutate(df = purrr::map(path, function(x) {
    read.csv(as.character(x))
 })) %>%
 unnest(df)



change_ests = ests_dat %>%
 select(
    -filename,
    -model_description,
    -fixed_effects,
    -random_effects,
    -dataframe,
    -notes,
    -date_added,
    -time_added,
    -output,
    -path,
    -folder
 ) %>%
 filter(grepl("total change", Label)) %>% relocate(Label) %>%
 mutate(split_label  = sapply(Label, function(x)
```

```
    str_split(str_split(x, " last_num = ")[[1]][2], "-"))) %>%
  relocate(split_label) %>%
  unnest(split_label) %>%
  mutate(split_label = factor(
    split_label,
    levels = as.character(1:5),
    labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
  )) %>%
  mutate(lab_level = rep(c("change_from", "change_to"), n() / 2)) %>%
  pivot_wider(values_from = split_label, names_from = lab_level) %>%
  mutate(change = paste(
    as.character(change_from),
    " to ",
    as.character(change_to),
    sep = ""
  )) %>%
  select(-Label) %>%
  relocate(outcome, change) %>%
  arrange(outcome, change)

conts_main = conts_dat %>%
  select(
    -filename,
    -model_description,
    -fixed_effects,
    -random_effects,
    -dataframe,
```

```
    -notes,
    -date_added,
    -time_added,
    -output,
    -path,
    -folder
  ) %>%
  mutate(
    Label = ifelse(
      grepl('total_change', Label),
      "Does the Total Change Differ Across Dropout Timepoints?",
      "Does the Baseline Measure Differ Across Dropout Timepoints?"
    )
  ) %>%
  arrange(outcome, Label)



baseline_ests = ests_dat %>%
  select(
    -filename,
    -model_description,
    -fixed_effects,
    -random_effects,
    -dataframe,
    -notes,
    -date_added,
    -time_added,
```

```
  -output,
  -path,
  -folder
) %>%
filter(!grepl("total change", Label)) %>%
mutate(change  = sapply(Label, function(x)
  str_split(x, " last_num ")[[1]][2])) %>%
relocate(change) %>%
mutate(change = sapply(change, function(x)
  str_split(x, "-")))  %>%
unnest(change) %>%
unique() %>%
mutate(change_lab = rep(c("change_from", "change_to"), n() / 2)) %>%
mutate(change = factor(
  change,
  levels = as.character(1:5),
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
pivot_wider(names_from = change_lab, values_from = change) %>%
unique() %>%
mutate(change = paste(
  as.character(change_from),
  " to ",
  as.character(change_to),
  sep = ""
)) %>%
```

```
 select(-Label) %>%
 relocate(outcome, change) %>%
 arrange(outcome, change)



oc = "IES"
conts_main  %>%
 filter(outcome == oc) %>%
 select(Label, ProbF) %>%
 mutate(ProbF = format.pval(ProbF)) %>%
 rename(`Hypothesis` = Label) %>%
 rename(`P-Value` = ProbF)  %>%
 knitr::kable(caption = "Contrasts For IES in  the Available Case Model",
              label = "iescontrasts",
              booktabs = TRUE)


baseline_ests %>%
 filter(outcome == oc) %>%
 mutate(change = factor(
   change,
   levels = c(
     'Baseline to Post Treatment',
     'Baseline to 1 Week',
     'Baseline to 3 Months',
     'Baseline to 6 Months'
   )
 )) %>%
```

```
 ggplot(aes(y = Estimate, x = change)) +
 geom_point() +
 geom_errorbar(aes(ymin = Lower, ymax = Upper)) +
 xlab("Dropout Timepoint Comparison") +
 ylab("Difference in Baseline Estimates") +
 geom_text(aes(label = paste("p =", format.pval(as.numeric(
   Probt
 )))), hjust = 1.2, family = "Times New Roman")



oc = "TAS"
conts_main  %>%
 filter(outcome == oc) %>%
 select(Label, ProbF) %>%
 mutate(ProbF = format.pval(ProbF)) %>%
 rename(`Hypothesis` = Label) %>%
 rename(`P-Value` = ProbF)  %>%
 knitr::kable(caption = "Contrasts For TAS in  the Available Case Model",
              label = "tascontrasts",
              booktabs = TRUE)


change_ests %>%
 filter(outcome == oc) %>%
 ggplot(aes(y = Estimate, x = change)) +
 geom_point() +
 geom_errorbar(aes(ymin = Lower, ymax = Upper)) +
 xlab("Dropout Timepoint Comparison") +
```

```
 ylab("Difference in Total Change") +
 geom_text(aes(label = paste("p =", Probt)), hjust = 1.2,
           family = "Times New Roman")
```

```
oc = "SCL"
conts_main  %>%
 filter(outcome == oc) %>%
 select(Label, ProbF) %>%
 mutate(ProbF = format.pval(ProbF)) %>%
 rename(`Hypothesis` = Label) %>%
 rename(`P-Value` = ProbF)  %>%
 knitr::kable(caption = "Contrasts For SCL in  the Available Case Model",
              label = "sclcontrasts",
              booktabs = TRUE)
```

```
oc = "logDES"
conts_main  %>%
 filter(outcome == oc) %>%
 select(Label, ProbF) %>%
 mutate(ProbF = format.pval(ProbF)) %>%
 rename(`Hypothesis` = Label) %>%
 rename(`P-Value` = ProbF)  %>%
 knitr::kable(caption = "Contrasts For DES in  the Available Case Model",
```

```
              label = "descontrasts",
              booktabs = TRUE)


oc = "logDRS"
conts_main  %>%
 filter(outcome == oc) %>%
 select(Label, ProbF) %>%
 mutate(ProbF = format.pval(ProbF)) %>%
 rename(`Hypothesis` = Label) %>%
 rename(`P-Value` = ProbF)  %>%
 knitr::kable(caption = "Contrasts For DRS in  the Available Case Model",
              label = "drscontrasts",
              booktabs = TRUE)


lsmeans =
 files %>%
 filter(grepl("lsmeans.*reg|lsmeans_pool.*", filename))

lsmeans_wdat = lsmeans  %>%
 mutate(path = paste('..', folder, filename, sep = "/")) %>%
 mutate(df = purrr::map(path, function(x) {
   read.csv(as.character(x)) %>% mutate(time_num = factor(time_num))
 })) %>%
 unnest(df)
```

```
lsmeans_wdat %>%
 select(
   -filename,
   -model_description,
   -fixed_effects,
   -random_effects,
   -dataframe,
   -notes,
   -date_added,
   -time_added,
   -output,
   -path
 ) %>%
 mutate(Effect = ifelse(grepl("mi", folder), "time_num", Effect)) %>%
 unique() %>%
 filter(Effect == "time_num") %>%
 mutate(time_num = factor(
   time_num,
   levels = 1:5,
   labels = c("Baseline", "Post Treatment",
              "1 Week", "3 Months", "6 Months")
 )) %>%
 mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
        "Available Case")) %>%
 mutate(Lower = ifelse(is.na(LCLMean),
                       Estimate - (1.96 * StdErr),
                       LCLMean)) %>%
```

```
mutate(Upper = ifelse(is.na(UCLMean),
                      Estimate + (1.96 * StdErr),
                      UCLMean)) %>%
select(outcome, Model, time_num, Estimate, Lower, Upper) %>%
mutate(
  Estimate = ifelse(grepl("log", outcome),
                    exp(Estimate) - 1,
                    Estimate),
  Upper = ifelse(grepl("log", outcome),
                 exp(Upper) - 1,
                 Upper),
  Lower = ifelse(grepl("log", outcome),
                 exp(Lower) - 1,
                 Lower)
) %>%
mutate(outcome = sapply(as.character(outcome), function(x)
  ifelse(grepl('log', x),
         str_split(x, "log")[[1]][2],
         x))) %>%
ggplot(aes(x = time_num, y = Estimate, color = Model)) +
geom_point() +
geom_line(aes(x = as.numeric(time_num))) +
geom_ribbon(aes(
  x = as.numeric(time_num),
  ymin = Lower,
  ymax = Upper,
  fill = Model
```

```
 ), alpha = .3) +
 facet_wrap( ~ outcome, scales = "free")  +
 scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
 scale_fill_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
 theme(legend.position = "top",
       axis.text.x = element_text(
         angle = 45,
         hjust = 1,
         size = 7
       )) + xlab("Follow Up Timepoint")



diffs =
 files %>%
 filter(grepl("diffs.*reg|diffs_pool.*", filename))


diffs_dat = diffs %>%
 mutate(path = paste('..', folder, filename, sep = "/")) %>%
 mutate(df = purrr::map(path, function(x) {
   read.csv(as.character(x))
 })) %>%
 unnest(df)



oc = "IES"
design <- matrix(c(1, 2, 4, 3), 2, 2)
layout(design)
```

```
plt =  diffs_dat %>%
 select(
   -filename,
   -model_description,
   -fixed_effects,
   -random_effects,
   -dataframe,
   -notes,
   -date_added,
   -time_added,
   -output,
   -path
 ) %>%
 mutate(time_num = purrr::map2(time_num, comparison, function(time, comp)
   ifelse(is.na(time),
          as.numeric(str_split(comp, "vs")[[1]][1]), time))) %>%
 mutate(X_time_num = purrr::map2(X_time_num, comparison, function(time, comp)
   ifelse(is.na(time),
          as.numeric(str_split(comp, "vs")[[1]][2]), time))) %>%
 unnest(time_num, X_time_num) %>%
 mutate(time_num = factor(
   time_num,
   levels = 1:5,
   labels = c("Baseline", "Post Treatment",
              "1 Week", "3 Months", "6 Months")
 )) %>%
 mutate(X_time_num = factor(
```

```
  X_time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
                      "Available Case")) %>%
mutate(Estimate = Estimate * (-1)) %>%
mutate(Lower = ifelse(
  is.na(LCLMean),
  Estimate - (1.96 * StdErr),
  Estimate - (1.96 * StdErr)
)) %>%
mutate(Upper = ifelse(
  is.na(UCLMean),
  Estimate + (1.96 * StdErr),
  Estimate + (1.96 * StdErr)
)) %>%
select(outcome,
       Model,
       time_num,
       X_time_num,
       Estimate,
       Lower,
       Upper,
       Probt) %>%
mutate(
```

```
  Estimate = ifelse(grepl("log", outcome),
                    exp(Estimate) - 1,
                    Estimate),
  Upper = ifelse(grepl("log", outcome),
                 exp(Upper) - 1,
                 Upper),
  Lower = ifelse(grepl("log", outcome),
                 exp(Lower) - 1,
                 Lower)
) %>%
mutate(Probt = ifelse(
  grepl("<", Probt),
  paste("p", Probt, sep = ""),
  paste("p=", Probt, sep = "")
)) %>%
mutate(time_num = paste("Change From", as.character(time_num))) %>%
mutate(time_num = factor(
  time_num,
  levels = c(
    "Change From Baseline",
    "Change From Post Treatment",
    "Change From 1 Week",
    "Change From 3 Months"
  )
)) %>%
filter(outcome == oc) %>%
ggplot(aes(x = X_time_num, y = Estimate, color = Model)) +
```

```
geom_point(position = position_dodge(0.2)) +
geom_errorbar(
  aes(ymin = Lower, ymax = Upper, fill = Model),
  alpha = 1,
  position = position_dodge(0.2)
) +
#facet_wrap(~time_num, scales = "free", shrink = F, ncol = 3)  +
geom_text(
  aes(
    label = Probt,
    color = Model,
    x = X_time_num,
    y = Lower
  ),
  position = position_dodge(1.1),
  size = 2.5,
  vjust = -.75,
  family = "Times New Roman"
) +
scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
scale_fill_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
theme(legend.position = "top",
      axis.text.x = element_text(
        angle = 45,
        hjust = 1,
        size = 8
      )) +
```

```
 xlab("Follow Up Timpoint") +
 ylab(paste("Estimated Difference in", oc)) +
 ggh4x::facet_manual(
   vars(time_num),
   design = design,
   widths = c(4, 2.2),
   scales = "free"
 )


print(plt)


oc = "TAS"
design <- matrix(c(1, 2, 4, 3), 2, 2)
layout(design)
plt =  diffs_dat %>%
 select(
   -filename,
   -model_description,
   -fixed_effects,
   -random_effects,
   -dataframe,
   -notes,
   -date_added,
   -time_added,
   -output,
```

```
  -path
) %>%
mutate(time_num = purrr::map2(time_num, comparison, function(time, comp)
  ifelse(is.na(time),
         as.numeric(str_split(comp, "vs")[[1]][1]), time))) %>%
mutate(X_time_num = purrr::map2(X_time_num, comparison, function(time, comp)
  ifelse(is.na(time),
         as.numeric(str_split(comp, "vs")[[1]][2]), time))) %>%
unnest(time_num, X_time_num) %>%
mutate(time_num = factor(
  time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(X_time_num = factor(
  X_time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
       "Available Case")) %>%
mutate(Estimate = Estimate * (-1)) %>%
mutate(Lower = ifelse(
  is.na(LCLMean),
  Estimate - (1.96 * StdErr),
```

```
  Estimate - (1.96 * StdErr)
)) %>%
mutate(Upper = ifelse(
  is.na(UCLMean),
  Estimate + (1.96 * StdErr),
  Estimate + (1.96 * StdErr)
)) %>%
select(outcome,
       Model,
       time_num,
       X_time_num,
       Estimate,
       Lower,
       Upper,
       Probt) %>%
mutate(
  Estimate = ifelse(grepl("log", outcome),
                    exp(Estimate) - 1,
                    Estimate),
  Upper = ifelse(grepl("log", outcome),
                 exp(Upper) - 1,
                 Upper),
  Lower = ifelse(grepl("log", outcome),
                 exp(Lower) - 1,
                 Lower)
) %>%
mutate(Probt = ifelse(
```

```
  grepl("<", Probt),
  paste("p", Probt, sep = ""),
  paste("p=", Probt, sep = "")
)) %>%
mutate(time_num = paste("Change From", as.character(time_num))) %>%
mutate(time_num = factor(
  time_num,
  levels = c(
    "Change From Baseline",
    "Change From Post Treatment",
    "Change From 1 Week",
    "Change From 3 Months"
  )
)) %>%
filter(outcome == oc) %>%
ggplot(aes(x = X_time_num, y = Estimate, color = Model)) +
geom_point(position = position_dodge(0.2)) +
geom_errorbar(
  aes(ymin = Lower, ymax = Upper, fill = Model),
  alpha = 1,
  position = position_dodge(0.2)
) +
#facet_wrap(~time_num, scales = "free", shrink = F, ncol = 3)  +
geom_text(
  aes(
    label = Probt,
    color = Model,
```

```
      x = X_time_num,
      y = Lower
    ),
    position = position_dodge(1.1),
    size = 2.5,
    vjust = -.75,
    family = "Times New Roman"
  ) +
  scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
  scale_fill_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
  theme(legend.position = "top",
        axis.text.x = element_text(
          angle = 45,
          hjust = 1,
          size = 8
        )) +
  xlab("Follow Up Timpoint") +
  ylab(paste("Estimated Difference in", oc)) +
  ggh4x::facet_manual(
    vars(time_num),
    design = design,
    widths = c(4, 2.2),
    scales = "free"
  )


print(plt)
```

```
oc = "SCL"
design <- matrix(c(1, 2, 4, 3), 2, 2)
layout(design)
plt =  diffs_dat %>%
 select(
   -filename,
   -model_description,
   -fixed_effects,
   -random_effects,
   -dataframe,
   -notes,
   -date_added,
   -time_added,
   -output,
   -path
 ) %>%
 mutate(time_num = purrr::map2(time_num, comparison, function(time, comp)
   ifelse(is.na(time),
          as.numeric(str_split(comp, "vs")[[1]][1]), time))) %>%
 mutate(X_time_num = purrr::map2(X_time_num, comparison, function(time, comp)
   ifelse(is.na(time),
          as.numeric(str_split(comp, "vs")[[1]][2]), time))) %>%
 unnest(time_num, X_time_num) %>%
 mutate(time_num = factor(
```

```
  time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(X_time_num = factor(
  X_time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
       "Available Case")) %>%
mutate(Estimate = Estimate * (-1)) %>%
mutate(Lower = ifelse(
  is.na(LCLMean),
  Estimate - (1.96 * StdErr),
  Estimate - (1.96 * StdErr)
)) %>%
mutate(Upper = ifelse(
  is.na(UCLMean),
  Estimate + (1.96 * StdErr),
  Estimate + (1.96 * StdErr)
)) %>%
select(outcome,
       Model,
       time_num,
```

```
        X_time_num,
        Estimate,
        Lower,
        Upper,
        Probt) %>%
mutate(
  Estimate = ifelse(grepl("log", outcome),
                    exp(Estimate) - 1,
                    Estimate),
  Upper = ifelse(grepl("log", outcome),
                 exp(Upper) - 1,
                 Upper),
  Lower = ifelse(grepl("log", outcome),
                 exp(Lower) - 1,
                 Lower)
) %>%
mutate(Probt = ifelse(
  grepl("<", Probt),
  paste("p", Probt, sep = ""),
  paste("p=", Probt, sep = "")
)) %>%
mutate(time_num = paste("Change From", as.character(time_num))) %>%
mutate(time_num = factor(
  time_num,
  levels = c(
    "Change From Baseline",
    "Change From Post Treatment",
```

```
    "Change From 1 Week",
    "Change From 3 Months"
  )
)) %>%
filter(outcome == oc) %>%
ggplot(aes(x = X_time_num, y = Estimate, color = Model)) +
geom_point(position = position_dodge(0.2)) +
geom_errorbar(
  aes(ymin = Lower, ymax = Upper, fill = Model),
  alpha = 1,
  position = position_dodge(0.2)
) +
#facet_wrap(~time_num, scales = "free", shrink = F, ncol = 3)  +
geom_text(
  aes(
    label = Probt,
    color = Model,
    x = X_time_num,
    y = Lower
  ),
  position = position_dodge(1.1),
  size = 2.5,
  vjust = -.75,
  family = "Times New Roman"
) +
scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
scale_fill_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
```

```
 theme(legend.position = "top",
       axis.text.x = element_text(
         angle = 45,
         hjust = 1,
         size = 8
       )) +
 xlab("Follow Up Timpoint") +
 ylab(paste("Estimated Difference in", oc)) +
 ggh4x::facet_manual(
   vars(time_num),
   design = design,
   widths = c(4, 2.2),
   scales = "free"
 )


print(plt)



oc = "logDES"
design <- matrix(c(1, 2, 4, 3), 2, 2)
layout(design)
plt =  diffs_dat %>%
 select(
   -filename,
   -model_description,
```

```
  -fixed_effects,
  -random_effects,
  -dataframe,
  -notes,
  -date_added,
  -time_added,
  -output,
  -path
) %>%
mutate(time_num = purrr::map2(time_num, comparison, function(time, comp)
  ifelse(is.na(time),
         as.numeric(str_split(comp, "vs")[[1]][1]), time))) %>%
mutate(X_time_num = purrr::map2(X_time_num, comparison, function(time, comp)
  ifelse(is.na(time),
         as.numeric(str_split(comp, "vs")[[1]][2]), time))) %>%
unnest(time_num, X_time_num) %>%
mutate(time_num = factor(
  time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(X_time_num = factor(
  X_time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
```

```
)) %>%
mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
       "Available Case")) %>%
mutate(Estimate = Estimate * (-1)) %>%
mutate(Lower = ifelse(
  is.na(LCLMean),
  Estimate - (1.96 * StdErr),
  Estimate - (1.96 * StdErr)
)) %>%
mutate(Upper = ifelse(
  is.na(UCLMean),
  Estimate + (1.96 * StdErr),
  Estimate + (1.96 * StdErr)
)) %>%
select(outcome,
       Model,
       time_num,
       X_time_num,
       Estimate,
       Lower,
       Upper,
       Probt) %>%
mutate(
  Estimate = ifelse(grepl("log", outcome),
                    100 * (exp(Estimate) - 1),
                    Estimate),
  Upper = ifelse(grepl("log", outcome),
```

```
                100 * (exp(Upper) - 1),
                Upper),
  Lower = ifelse(grepl("log", outcome),
                100 * (exp(Lower) - 1),
                Lower)
) %>%
mutate(Probt = ifelse(
  grepl("<", Probt),
  paste("p", Probt, sep = ""),
  paste("p=", Probt, sep = "")
)) %>%
mutate(time_num = paste("Change From", as.character(time_num))) %>%
mutate(time_num = factor(
  time_num,
  levels = c(
    "Change From Baseline",
    "Change From Post Treatment",
    "Change From 1 Week",
    "Change From 3 Months"
  )
)) %>%
filter(outcome == oc) %>%
ggplot(aes(x = X_time_num, y = Estimate, color = Model)) +
geom_point(position = position_dodge(0.2)) +
geom_errorbar(
  aes(ymin = Lower, ymax = Upper, fill = Model),
  alpha = 1,
```

```
  position = position_dodge(0.2)
) +
#facet_wrap(~time_num, scales = "free", shrink = F, ncol = 3)  +
geom_text(
  aes(
    label = Probt,
    color = Model,
    x = X_time_num,
    y = Lower
  ),
  position = position_dodge(1.1),
  size = 2.5,
  vjust = -.75,
  family = "Times New Roman"
) +
scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
scale_fill_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
theme(legend.position = "top",
      axis.text.x = element_text(
        angle = 45,
        hjust = 1,
        size = 8
      )) +
xlab("Follow Up Timpoint") +
ylab(paste("Estimated Percentage Difference in DES")) +
ggh4x::facet_manual(
  vars(time_num),
```

```
   design = design,
   widths = c(4, 2.2),
   scales = "free"
 )



print(plt)




oc = "logDRS"
design <- matrix(c(1, 2, 4, 3), 2, 2)
layout(design)
plt =  diffs_dat %>%
 select(
   -filename,
   -model_description,
   -fixed_effects,
   -random_effects,
   -dataframe,
   -notes,
   -date_added,
   -time_added,
   -output,
   -path
 ) %>%
 mutate(time_num = purrr::map2(time_num, comparison, function(time, comp)
```

```
  ifelse(is.na(time),
         as.numeric(str_split(comp, "vs")[[1]][1]), time))) %>%
mutate(X_time_num = purrr::map2(X_time_num, comparison, function(time, comp)
  ifelse(is.na(time),
         as.numeric(str_split(comp, "vs")[[1]][2]), time))) %>%
unnest(time_num, X_time_num) %>%
mutate(time_num = factor(
  time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(X_time_num = factor(
  X_time_num,
  levels = 1:5,
  labels = c("Baseline", "Post Treatment",
             "1 Week", "3 Months", "6 Months")
)) %>%
mutate(Model = ifelse(grepl("mi", folder), "Multiple Imputation",
       "Available Case")) %>%
mutate(Estimate = Estimate * (-1)) %>%
mutate(Lower = ifelse(
  is.na(LCLMean),
  Estimate - (1.96 * StdErr),
  Estimate - (1.96 * StdErr)
)) %>%
mutate(Upper = ifelse(
```

```
  is.na(UCLMean),
  Estimate + (1.96 * StdErr),
  Estimate + (1.96 * StdErr)
)) %>%
select(outcome,
       Model,
       time_num,
       X_time_num,
       Estimate,
       Lower,
       Upper,
       Probt) %>%
mutate(
  Estimate = ifelse(grepl("log", outcome),
                    100 * (exp(Estimate) - 1),
                    Estimate),
  Upper = ifelse(grepl("log", outcome),
                 100 * (exp(Upper) - 1),
                 Upper),
  Lower = ifelse(grepl("log", outcome),
                 100 * (exp(Lower) - 1),
                 Lower)
) %>%
mutate(Probt = ifelse(
  grepl("<", Probt),
  paste("p", Probt, sep = ""),
  paste("p=", Probt, sep = "")
```

```
)) %>%
mutate(time_num = paste("Change From", as.character(time_num))) %>%
mutate(time_num = factor(
  time_num,
  levels = c(
    "Change From Baseline",
    "Change From Post Treatment",
    "Change From 1 Week",
    "Change From 3 Months"
  )
)) %>%
filter(outcome == oc) %>%
ggplot(aes(x = X_time_num, y = Estimate, color = Model)) +
geom_point(position = position_dodge(0.2)) +
geom_errorbar(
  aes(ymin = Lower, ymax = Upper, fill = Model),
  alpha = 1,
  position = position_dodge(0.2)
) +
#facet_wrap(~time_num, scales = "free", shrink = F, ncol = 3)  +
geom_text(
  aes(
    label = Probt,
    color = Model,
    x = X_time_num,
    y = Lower
  ),
```

```
    position = position_dodge(1.1),
    size = 2.5,
    vjust = -.75,
    family = "Times New Roman"
  ) +
  scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
  scale_fill_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)]) +
  theme(legend.position = "top",
        axis.text.x = element_text(
          angle = 45,
          hjust = 1,
          size = 8
        )) +
  xlab("Follow Up Timpoint") +
  ylab(paste("Estimated Percentage Difference in DRS")) +
  ggh4x::facet_manual(
    vars(time_num),
    design = design,
    widths = c(4, 2.2),
    scales = "free"
  )


print(plt)
```

```
fixed_files =
 files %>%
 filter(grepl("^fixed.*reg|mianalyze", filename)) %>%
 filter(grepl("_1.csv", filename))



all_fixed = tibble()
for (fixed in unique(fixed_files$filename)) {
 cur_mod = filter(fixed_files, filename == fixed)
 outcome_name = cur_mod$outcome

 fn = cur_mod$filename
 folder = cur_mod$folder
 df = cur_mod$output
 path = paste("../", folder, "/", fn, sep = "")
 cur_fixed = read.csv(path)
 if (grepl('mi', folder)) {
   cur_fixed = cur_fixed %>%
     rename(Effect = Parm,
            Lower = LCLMean,
            Upper = UCLMean)  %>%
     mutate(Effect = tolower(Effect))
 } else {
   cur_fixed = cur_fixed %>%
     mutate(last_num = factor(
       last_num,
       levels = c(1:5),
```

```
      labels = c(
        "Dropout at Baseline",
        "Dropout at Post Treatment",
        "Dropout at 1 Week",
        "Dropout at 3 Months",
        "Dropout at 6 Months"
      )
    ))
}
cur_fixed = cur_fixed %>%
  mutate(time_num = factor(
    time_num,
    levels = c(1:5),
    labels = c("Baseline", "Post Treatment", "1 Week", "3 Months", "6 Months")
  )) %>%
  mutate(rn = row_number()) %>%
  mutate(Variable = sapply(rn, function(x)
    ifelse(
      Effect[x] == "Intercept",
      "Intercept",
      ifelse(
        Effect[x] == "intercept",
        "Intercept",
        ifelse(
          grepl("\\*", Effect[x]),
          paste(
            as.character(time_num[x]),
```

```
            " - ",
            as.character(last_num)[x],
            " Interaction",
            sep = ""
          ),
          ifelse(!(
            Effect[x] %in% c("Age", "TX_DAYS", "age", "tx_days")
          ),
          as.character(get(
            as.character(unique(Effect[x]))
          )[x]),
          Effect[x])
        )
      )
    ))) %>%
  mutate(var = Effect) %>%
  mutate(Effect = paste(Effect, ": ", Variable, sep = ""))
cur_fixed =
  cur_fixed %>%
  mutate(outcome = outcome_name, folder = folder)
cur_fixed = cur_fixed %>%
  select(var,
         Effect,
         Estimate,
         StdErr,
         Lower,
         Upper,
```

```
          DF,
          tValue,
          Probt,
          outcome,
          folder)
 all_fixed = bind_rows(all_fixed, cur_fixed)
}
```

```
oc = "IES"
plt_df =
 all_fixed %>%
 filter(grepl(oc, outcome)) %>%
 filter(!grepl("_num", Effect)) %>%
 mutate(
   Effect = str_replace_all(Effect, "Adjustm$", "Adjustment Disorder"),
   Effect = str_replace_all(Effect, "[aA]+xis_[iI]+.*:", "Primary Diagnosis:"),
   Effect = str_replace_all(Effect, "[tT]+[xX]+.*$", "Treatment Days"),
   Effect = str_replace_all(Effect, "[aA]+ge.*$", "Age"),
   Effect = str_replace_all(Effect, "[iI]+ntercept.*$", "Intercept"),
   var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
   var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
 ) %>%
 mutate(
   var = str_replace_all(var, "Adjustm$", "Adjustment Disorder"),
```

```
  var = str_replace_all(var, "[aA]+xis_[iI]+.*", "Primary Diagnosis"),
  var = str_replace_all(var, "[tT]+[xX]+.*$", "Treatment Days"),
  var = str_replace_all(var, "[aA]+ge.*$", "Age"),
  var = str_replace_all(var, "[iI]+ntercept.*$", "Intercept"),
  var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
  var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
) %>%
mutate(Effect = sapply(Effect, function(x)
  ifelse(grepl(":", x),
         str_split(x, ":")[[1]][2],
         x))) %>%
group_by(folder) %>%
mutate(folder = ifelse(
  grepl("mi", folder),
  "Multiple Imputation",
  ifelse(
    grepl("weight", folder),
    "Complete Case with Weighted Dropout Timepoint",
    "Available Case"
  )
)) %>%
mutate(outcome =  str_split(outcome, "_")[[1]][1]) %>%
group_by(var) %>%
mutate(ref = ifelse(n() > 1,
                    Effect[Estimate == 0])) %>%
relocate(ref) %>%
ungroup() %>%
```

```
 mutate(var = ifelse(
   !is.na(ref),
   paste(var, " with Reference Variable ", ref, sep = ""),
   var
 )) %>%
 filter(!is.na(StdErr)) %>%
 select(-ref) %>%
 mutate(Effect = str_wrap(Effect, width = 8)) %>%
 group_by(var) %>%
 mutate(sublevs = length(unique(Effect))) %>%
 mutate(var = ifelse(sublevs == 1, str_wrap(var, width = 30), var)) %>%
 mutate(var = factor(var)) %>%
 ungroup()

if (select(plt_df, var, sublevs) %>% unique %>% select(sublevs) %>% sum > 3) {
 long_vars = select(plt_df, var, sublevs) %>% unique %>%
    filter(sublevs > 1) %>% select(var) %>% unlist()
 long_vars = sapply(long_vars, as.numeric)
 short_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs == 1) %>% select(var) %>% unlist()
 short_vars = sapply(short_vars, as.numeric)

 if (length(short_vars) %% 2 == 0) {
   short_vars = matrix(short_vars, ncol = 2)
 } else{
   short_vars = matrix(short_vars, ncol = 3)
 }
```

```
 display = short_vars
 for (v in long_vars) {
   add_v = matrix(v, ncol = ncol(short_vars))
   display = rbind(short_vars, add_v)
 }
 plt = plt_df %>% ggplot(aes(x = Effect, y = Estimate, color = folder)) +
   geom_point(position = position_dodge(0.4)) +
   geom_errorbar(aes(ymin = Lower, ymax = Upper),
        position = position_dodge(0.4)) +
   facet_wrap( ~ var, scales = "free") +
   ggh4x::facet_manual(vars(var), design = display,  scales = "free") +
   labs(color = "Model") +
   ylab(paste("Estimated Change for", oc)) +
   xlab("Fixed Effect") +
   theme(
     legend.position = "top",
     axis.text.x = element_text(size = 7),
     legend.text = element_text(size = 8),
     strip.text = element_text(size = 8)
   ) +
   scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])
} else{
 plt =  plt_df %>%
   ggplot(aes(x = Effect, y = Estimate, color = folder)) +
   geom_point(position = position_dodge(0.4)) +
   geom_errorbar(aes(ymin = Lower, ymax = Upper),
          position = position_dodge(0.4)) +
```

```
   facet_wrap( ~ var, scales = "free") +
   labs(color = "Model") +
   ylab(paste("Estimated Change for", oc)) +
   xlab("Fixed Effect") +
   theme(
     legend.position = "top",
     axis.text.x = element_text(size = 7),
     legend.text = element_text(size = 8),
     strip.text = element_text(size = 8)
   ) +
   scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])

}

print(plt)



oc = "TAS"
plt_df =
 all_fixed %>%
 filter(grepl(oc, outcome)) %>%
 filter(!grepl("_num", Effect)) %>%
 mutate(
   Effect = str_replace_all(Effect, "Adjustm$", "Adjustment Disorder"),
   Effect = str_replace_all(Effect, "[aA]+xis_[iI]+.*:", "Primary Diagnosis:"),
   Effect = str_replace_all(Effect, "[tT]+[xX]+.*$", "Treatment Days"),
```

```
  Effect = str_replace_all(Effect, "[aA]+ge.*$", "Age"),
  Effect = str_replace_all(Effect, "[iI]+ntercept.*$", "Intercept"),
  var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
  var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
) %>%
mutate(
  var = str_replace_all(var, "Adjustm$", "Adjustment Disorder"),
  var = str_replace_all(var, "[aA]+xis_[iI]+.*", "Primary Diagnosis"),
  var = str_replace_all(var, "[tT]+[xX]+.*$", "Treatment Days"),
  var = str_replace_all(var, "[aA]+ge.*$", "Age"),
  var = str_replace_all(var, "[iI]+ntercept.*$", "Intercept"),
  var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
  var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
) %>%
mutate(Effect = sapply(Effect, function(x)
  ifelse(grepl(":", x),
         str_split(x, ":")[[1]][2],
         x))) %>%
group_by(folder) %>%
mutate(folder = ifelse(
  grepl("mi", folder),
  "Multiple Imputation",
  ifelse(
    grepl("weight", folder),
    "Complete Case with Weighted Dropout Timepoint",
    "Available Case"
  )
```

```
 )) %>%
 mutate(outcome =  str_split(outcome, "_")[[1]][1]) %>%
 group_by(var) %>%
 mutate(ref = ifelse(n() > 1,
                     Effect[Estimate == 0])) %>%
 relocate(ref) %>%
 ungroup() %>%
 mutate(var = ifelse(
   !is.na(ref),
   paste(var, " with Reference Variable ", ref, sep = ""),
   var
 )) %>%
 filter(!is.na(StdErr)) %>%
 select(-ref) %>%
 mutate(Effect = str_wrap(Effect, width = 8)) %>%
 group_by(var) %>%
 mutate(sublevs = length(unique(Effect))) %>%
 mutate(var = ifelse(sublevs == 1, str_wrap(var, width = 30), var)) %>%
 mutate(var = factor(var)) %>%
 ungroup()

if (select(plt_df, var, sublevs) %>% unique %>% select(sublevs) %>% sum > 3) {
 long_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs > 1) %>% select(var) %>% unlist()
 long_vars = sapply(long_vars, as.numeric)
 short_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs == 1) %>% select(var) %>% unlist()
```

```
short_vars = sapply(short_vars, as.numeric)


if (length(short_vars) %% 2 == 0) {
  short_vars = matrix(short_vars, ncol = 2)
} else{
  short_vars = matrix(short_vars, ncol = 3)
}
display = short_vars
for (v in long_vars) {
  add_v = matrix(v, ncol = ncol(short_vars))
  display = rbind(short_vars, add_v)
}
plt = plt_df %>% ggplot(aes(x = Effect, y = Estimate, color = folder)) +
  geom_point(position = position_dodge(0.4)) +
  geom_errorbar(aes(ymin = Lower, ymax = Upper),
       position = position_dodge(0.4)) +
  facet_wrap( ~ var, scales = "free") +
  ggh4x::facet_manual(vars(var), design = display,  scales = "free") +
  labs(color = "Model") +
  ylab(paste("Estimated Change for", oc)) +
  xlab("Fixed Effect") +
  theme(
    legend.position = "top",
    axis.text.x = element_text(size = 7),
    legend.text = element_text(size = 8),
    strip.text = element_text(size = 8)
  ) +
```

```
   scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])
} else{
 plt =  plt_df %>%
   ggplot(aes(x = Effect, y = Estimate, color = folder)) +
   geom_point(position = position_dodge(0.4)) +
   geom_errorbar(aes(ymin = Lower, ymax = Upper),
        position = position_dodge(0.4)) +
   facet_wrap( ~ var, scales = "free") +
   labs(color = "Model") +
   ylab(paste("Estimated Change for", oc)) +
   xlab("Fixed Effect") +
   theme(
     legend.position = "top",
     axis.text.x = element_text(size = 7),
     legend.text = element_text(size = 8),
     strip.text = element_text(size = 8)
   ) +
   scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])

}

print(plt)



oc = "SCL"
plt_df =
```

```
all_fixed %>%
filter(grepl(oc, outcome)) %>%
filter(!grepl("_num", Effect)) %>%
mutate(
  Effect = str_replace_all(Effect, "Adjustm$", "Adjustment Disorder"),
  Effect = str_replace_all(Effect, "[aA]+xis_[iI]+.*:", "Primary Diagnosis:"),
  Effect = str_replace_all(Effect, "[tT]+[xX]+.*$", "Treatment Days"),
  Effect = str_replace_all(Effect, "[aA]+ge.*$", "Age"),
  Effect = str_replace_all(Effect, "[iI]+ntercept.*$", "Intercept"),
  var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
  var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
) %>%
mutate(
  var = str_replace_all(var, "Adjustm$", "Adjustment Disorder"),
  var = str_replace_all(var, "[aA]+xis_[iI]+.*", "Primary Diagnosis"),
  var = str_replace_all(var, "[tT]+[xX]+.*$", "Treatment Days"),
  var = str_replace_all(var, "[aA]+ge.*$", "Age"),
  var = str_replace_all(var, "[iI]+ntercept.*$", "Intercept"),
  var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
  var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
) %>%
mutate(Effect = sapply(Effect, function(x)
  ifelse(grepl(":", x),
         str_split(x, ":")[[1]][2],
         x))) %>%
group_by(folder) %>%
mutate(folder = ifelse(
```

```
  grepl("mi", folder),
  "Multiple Imputation",
  ifelse(
    grepl("weight", folder),
    "Complete Case with Weighted Dropout Timepoint",
    "Available Case"
  )
)) %>%
mutate(outcome =  str_split(outcome, "_")[[1]][1]) %>%
group_by(var) %>%
mutate(ref = ifelse(n() > 1,
                    Effect[Estimate == 0])) %>%
relocate(ref) %>%
ungroup() %>%
mutate(var = ifelse(
  !is.na(ref),
  paste(var, " with Reference Variable ", ref, sep = ""),
  var
)) %>%
filter(!is.na(StdErr)) %>%
select(-ref) %>%
mutate(Effect = str_wrap(Effect, width = 8)) %>%
group_by(var) %>%
mutate(sublevs = length(unique(Effect))) %>%
mutate(var = ifelse(sublevs == 1, str_wrap(var, width = 30), var)) %>%
mutate(var = factor(var)) %>%
ungroup()
```

```
if (select(plt_df, var, sublevs) %>% unique %>% select(sublevs) %>% sum > 3) {
 long_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs > 1) %>% select(var) %>% unlist()
 long_vars = sapply(long_vars, as.numeric)
 short_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs == 1) %>% select(var) %>% unlist()
 short_vars = sapply(short_vars, as.numeric)

 if (length(short_vars) %% 2 == 0) {
   short_vars = matrix(short_vars, ncol = 2)
 } else{
   short_vars = matrix(short_vars, ncol = 3)
 }
 display = short_vars
 for (v in long_vars) {
   add_v = matrix(v, ncol = ncol(short_vars))
   display = rbind(short_vars, add_v)
 }
 plt = plt_df %>% ggplot(aes(x = Effect, y = Estimate, color = folder)) +
   geom_point(position = position_dodge(0.4)) +
   geom_errorbar(aes(ymin = Lower, ymax = Upper),
   position = position_dodge(0.4)) +
   facet_wrap( ~ var, scales = "free") +
   ggh4x::facet_manual(vars(var), design = display,  scales = "free") +
   labs(color = "Model") +
   ylab(paste("Estimated Change for", oc)) +
```

```
   xlab("Fixed Effect") +
   theme(
     legend.position = "top",
     axis.text.x = element_text(size = 7),
     legend.text = element_text(size = 8),
     strip.text = element_text(size = 8)
   ) +
   scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])
} else{
 plt =  plt_df %>%
   ggplot(aes(x = Effect, y = Estimate, color = folder)) +
   geom_point(position = position_dodge(0.4)) +
   geom_errorbar(aes(ymin = Lower, ymax = Upper),
   position = position_dodge(0.4)) +
   facet_wrap( ~ var, scales = "free") +
   labs(color = "Model") +
   ylab(paste("Estimated Change for", oc)) +
   xlab("Fixed Effect") +
   theme(
     legend.position = "top",
     axis.text.x = element_text(size = 7),
     legend.text = element_text(size = 8),
     strip.text = element_text(size = 8)
   ) +
   scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])

}
```

```
print(plt)


oc = "logDES"
plt_df =
 all_fixed %>%
 filter(grepl(oc, outcome)) %>%
 filter(!grepl("_num", Effect)) %>%
 mutate(
   Effect = str_replace_all(Effect, "Adjustm$", "Adjustment Disorder"),
   Effect = str_replace_all(Effect, "[aA]+xis_[iI]+.*:", "Primary Diagnosis:"),
   Effect = str_replace_all(Effect, "[tT]+[xX]+.*$", "Treatment Days"),
   Effect = str_replace_all(Effect, "[aA]+ge.*$", "Age"),
   Effect = str_replace_all(Effect, "[iI]+ntercept.*$", "Intercept"),
   var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
   var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
 ) %>%
 mutate(
   var = str_replace_all(var, "Adjustm$", "Adjustment Disorder"),
   var = str_replace_all(var, "[aA]+xis_[iI]+.*", "Primary Diagnosis"),
   var = str_replace_all(var, "[tT]+[xX]+.*$", "Treatment Days"),
   var = str_replace_all(var, "[aA]+ge.*$", "Age"),
   var = str_replace_all(var, "[iI]+ntercept.*$", "Intercept"),
   var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
   var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
 ) %>%
 mutate(Effect = sapply(Effect, function(x)
```

```
  ifelse(grepl(":", x),
         str_split(x, ":")[[1]][2],
         x))) %>%
group_by(folder) %>%
mutate(folder = ifelse(
  grepl("mi", folder),
  "Multiple Imputation",
  ifelse(
    grepl("weight", folder),
    "Complete Case with Weighted Dropout Timepoint",
    "Available Case"
  )
)) %>%
mutate(outcome =  str_split(outcome, "_")[[1]][1]) %>%
group_by(var) %>%
mutate(ref = ifelse(n() > 1,
                    Effect[Estimate == 0])) %>%
relocate(ref) %>%
ungroup() %>%
mutate(var = ifelse(
  !is.na(ref),
  paste(var, " with Reference Variable ", ref, sep = ""),
  var
)) %>%
filter(!is.na(StdErr)) %>%
select(-ref) %>%
mutate(Effect = str_wrap(Effect, width = 8)) %>%
```

```
 group_by(var) %>%
 mutate(sublevs = length(unique(Effect))) %>%
 mutate(var = ifelse(sublevs == 1, str_wrap(var, width = 30), var)) %>%
 mutate(var = factor(var)) %>%
 ungroup()

if (select(plt_df, var, sublevs) %>% unique %>% select(sublevs) %>% sum > 3) {
 long_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs > 1) %>% select(var) %>% unlist()
 long_vars = sapply(long_vars, as.numeric)
 short_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs == 1) %>% select(var) %>% unlist()
 short_vars = sapply(short_vars, as.numeric)

 if (length(short_vars) %% 2 == 0) {
   short_vars = matrix(short_vars, ncol = 2)
 } else{
   short_vars = matrix(short_vars, ncol = 3)
 }
 display = short_vars
 for (v in long_vars) {
   add_v = matrix(v, ncol = ncol(short_vars))
   display = rbind(short_vars, add_v)
 }
 plt = plt_df %>% ggplot(aes(x = Effect, y = Estimate, color = folder)) +
   geom_point(position = position_dodge(0.4)) +
   geom_errorbar(aes(ymin = Lower, ymax = Upper),
```

```
    position = position_dodge(0.4)) +
    facet_wrap( ~ var, scales = "free") +
    ggh4x::facet_manual(vars(var), design = display,  scales = "free") +
    labs(color = "Model") +
    ylab(paste("Estimated Change for log(DES)")) +
    xlab("Fixed Effect") +
    theme(
      legend.position = "top",
      axis.text.x = element_text(size = 7),
      legend.text = element_text(size = 8),
      strip.text = element_text(size = 8)
    ) +
    scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])
} else{
 plt =  plt_df %>%
    ggplot(aes(x = Effect, y = Estimate, color = folder)) +
    geom_point(position = position_dodge(0.4)) +
    geom_errorbar(aes(ymin = Lower, ymax = Upper), position = position_dodge(0.4)) +
    facet_wrap( ~ var, scales = "free") +
    labs(color = "Model") +
    ylab(paste("Estimated Change for", oc)) +
    xlab("Fixed Effect") +
    theme(
      legend.position = "top",
      axis.text.x = element_text(size = 7),
      legend.text = element_text(size = 8),
      strip.text = element_text(size = 8)
```

```
  ) +
  scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])

}

print(plt)



oc = "logDRS"
plt_df =
 all_fixed %>%
 filter(grepl(oc, outcome)) %>%
 filter(!grepl("_num", Effect)) %>%
 mutate(
   Effect = str_replace_all(Effect, "Adjustm$", "Adjustment Disorder"),
   Effect = str_replace_all(Effect, "[aA]+xis_[iI]+.*:", "Primary Diagnosis:"),
   Effect = str_replace_all(Effect, "[tT]+[xX]+.*$", "Treatment Days"),
   Effect = str_replace_all(Effect, "[aA]+ge.*$", "Age"),
   Effect = str_replace_all(Effect, "[iI]+ntercept.*$", "Intercept"),
   var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
   var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
 ) %>%
 mutate(
   var = str_replace_all(var, "Adjustm$", "Adjustment Disorder"),
   var = str_replace_all(var, "[aA]+xis_[iI]+.*", "Primary Diagnosis"),
   var = str_replace_all(var, "[tT]+[xX]+.*$", "Treatment Days"),
   var = str_replace_all(var, "[aA]+ge.*$", "Age"),
```

```
  var = str_replace_all(var, "[iI]+ntercept.*$", "Intercept"),
  var = str_replace_all(var, "[gG]+ender.*$", "Gender"),
  var = str_replace_all(var, "[tT]+rauma_.*$", "Trauma Type")
) %>%
mutate(Effect = sapply(Effect, function(x)
  ifelse(grepl(":", x),
         str_split(x, ":")[[1]][2],
         x))) %>%
group_by(folder) %>%
mutate(folder = ifelse(
  grepl("mi", folder),
  "Multiple Imputation",
  ifelse(
    grepl("weight", folder),
    "Complete Case with Weighted Dropout Timepoint",
    "Available Case"
  )
)) %>%
mutate(outcome =  str_split(outcome, "_")[[1]][1]) %>%
group_by(var) %>%
mutate(ref = ifelse(n() > 1,
                    Effect[Estimate == 0])) %>%
relocate(ref) %>%
ungroup() %>%
mutate(var = ifelse(
  !is.na(ref),
  paste(var, " with Reference Variable ", ref, sep = ""),
```

```
  var
)) %>%
 filter(!is.na(StdErr)) %>%
 select(-ref) %>%
 mutate(Effect = str_wrap(Effect, width = 8)) %>%
 group_by(var) %>%
 mutate(sublevs = length(unique(Effect))) %>%
 mutate(var = ifelse(sublevs == 1, str_wrap(var, width = 30), var)) %>%
 mutate(var = factor(var)) %>%
 ungroup()

if (select(plt_df, var, sublevs) %>% unique %>% select(sublevs) %>% sum > 3) {
 long_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs >  1) %>% select(var) %>% unlist()
 long_vars = sapply(long_vars, as.numeric)
 short_vars = select(plt_df, var, sublevs) %>% unique %>%
 filter(sublevs == 1) %>% select(var) %>% unlist()
 short_vars = sapply(short_vars, as.numeric)

 if (length(short_vars) %% 2 == 0) {
   short_vars = matrix(short_vars, ncol = 2)
 } else{
   short_vars = matrix(short_vars, ncol = 3)
 }
 display = short_vars
 for (v in long_vars) {
   add_v = matrix(v, ncol = ncol(short_vars))
```

```
    display = rbind(short_vars, add_v)
  }
  plt = plt_df %>% ggplot(aes(x = Effect, y = Estimate, color = folder)) +
    geom_point(position = position_dodge(0.4)) +
    geom_errorbar(aes(ymin = Lower, ymax = Upper),
    position = position_dodge(0.4)) +
    facet_wrap( ~ var, scales = "free") +
    ggh4x::facet_manual(vars(var), design = display,  scales = "free") +
    labs(color = "Model") +
    ylab(paste("Estimated Change for log(DRS)")) +
    xlab("Fixed Effect") +
    theme(
      legend.position = "top",
      axis.text.x = element_text(size = 7),
      legend.text = element_text(size = 8),
      strip.text = element_text(size = 8)
    ) +
    scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])
} else{
  plt =  plt_df %>%
    ggplot(aes(x = Effect, y = Estimate, color = folder)) +
    geom_point(position = position_dodge(0.4)) +
    geom_errorbar(aes(ymin = Lower, ymax = Upper),
    position = position_dodge(0.4)) +
    facet_wrap( ~ var, scales = "free") +
    labs(color = "Model") +
    ylab(paste("Estimated Change for", oc)) +
```

```
  xlab("Fixed Effect") +
  theme(
    legend.position = "top",
    axis.text.x = element_text(size = 7),
    legend.text = element_text(size = 8),
    strip.text = element_text(size = 8)
  ) +
  scale_color_manual(values = ggpubr::get_palette('npg', 7)[c(3, 5)])

}

print(plt)
```