Interventions that Reduce Mental Distress Across Months

Consulting Report prepared for Dr. Lang Wu

Trevor Kwan

April 22, 2021

Department of Statistics

University of British Columbia

Abstract

This study seeks to determine the effect of a treatment on reducing mental distress over time. Subjects provided information on their mental distress, gender, and education through a self-report questionnaire. Mixed-effects models found a significant decrease in mental distress over time across all subjects. No significant difference in mental distress was found between treatment and control groups. Results were limited by missing data and violations of model assumptions.

1. Introduction

Background The study investigates changes in subjects' mental distress over time in an intervention group and a control group.

Methods Subjects answered questionnaires at baseline (month = 0), 3, 6, 18, and 60 months. These questionnaires measured a number of variables, including mental distress (GSI), gender, and education.

Measurement (Variable Descriptions) The independent variable was group status. All subjects were randomly assigned into one of two groups: a treatment/intervention group and a control group. Other predictor variables, including gender and education, were measured in the questionnaires.

The dependent variable was the Global Severity Index (GSI). The GSI is used as a measure of mental distress in subjects. A higher GSI score implies a higher level of mental distress.

Objectives The main objectives of the study are to compare the treatment effect with the control effect on reducing mental distress, to study how mental distress changes over time in all groups, and to investigate whether changes in mental distress can be explained by the treatment variable or other predictor variables.

Statistical Questions of Interest There was interest in investigating whether subjects' mental distress in both treatment groups decrease significantly over time. There was also interest in testing the effectiveness of the treatment. If a significant difference in subjects' mental distress between the treatment groups is found, the treatment is considered effective.

Why is the problem important? If a treatment is found to be effective in reducing mental distress in subjects, this intervention can be recommended to people facing overwhelming levels of mental distress.

2. Exploratory Data Analysis (EDA)

A large number of observations seem to be missing (see Table 1). The rate of missing observations per month was found to be greater for later months in comparison to earlier months (see Table 2). Exploratory data analysis was done using only data without missing observations. The data were examined for trends corresponding to the research questions: trajectory of mental distress over time and differences in mental distress between treatment and control groups.

Table 1: Descriptive Summary Statistics

Statistic	Value
Number of Observations (Rows)	1355.0
Number of Observations with Missing Data (Rows)	317.0
Number of Complete Observations	1038.0
Number of Unique Subjects	271.0
Number of Unique Education Levels	14.0
Mean GSI	0.8
Standard Deviation GSI	0.4

Table 2: Rate of Missing Observations By Month

Month	Missing Rate
0	0.03
3	0.14
6	0.17
18	0.34
60	0.32

Large variations in GSI per month were found between individuals, indicating mixed-effects models in formal analysis to be useful (see Figure 1). Overall trends seemed to indicate that GSI scores were decreasing over time in both treatment and control groups (see Figure 1 and Figure 2). Grouping individuals by treatment status showed the median GSI of the treatment group to be slightly higher than the median GSI of the control group (see Figure 3). A two-sample t-test was conducted to explore the possibility of a significant difference in GSI scores across treatment groups, but the difference was not found to be significant (see Figure 3). A Wilcoxon test was also conducted since t-tests are sensitive to outliers, and this difference was found to be significant (see Figure 3). The figures in this section reflect only the data for which there were no missing observations.

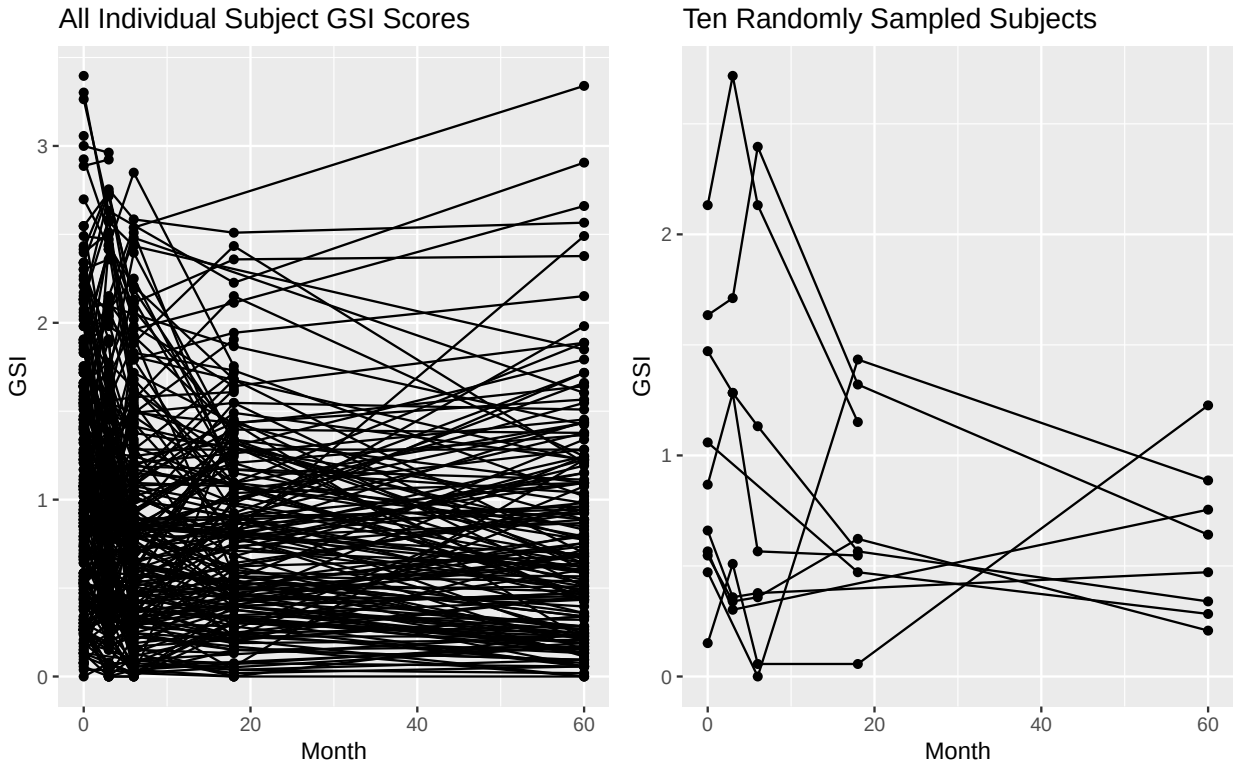


Figure 1: Individual GSI Scores Across Months

3. Formal Data Analysis

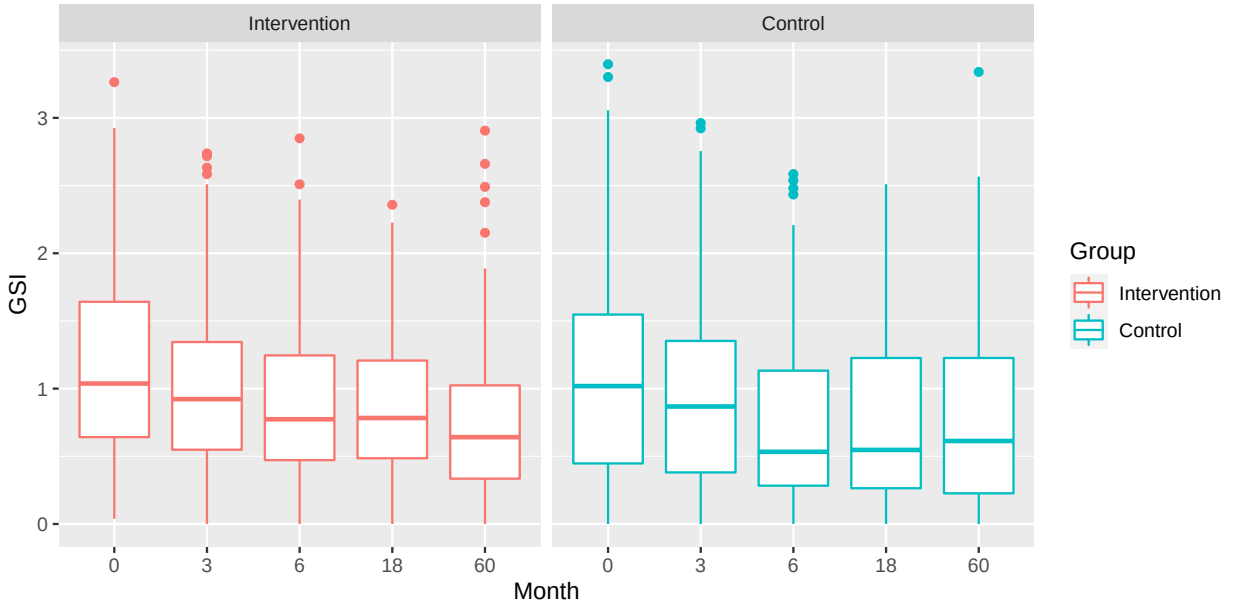


Figure 2: GSI Scores Across Months Grouped By Treatment Group

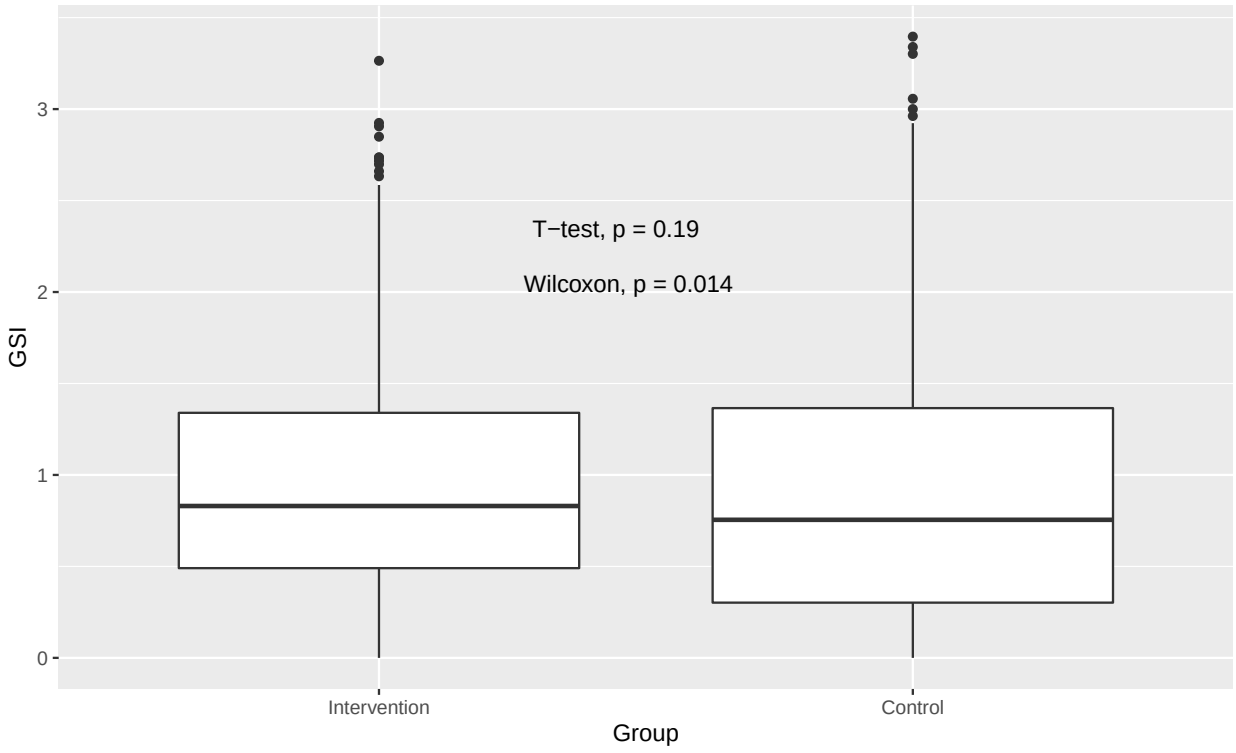


Figure 3: GSI Scores Across Treatment Groups

Missing Data

Of the 1355 observations, 317 were missing so the first step of the formal analysis was to determine how to handle the missing data. A complete cases method, where all missing observations were removed from the dataset was first attempted, but this required a strong assumption for the data to be missing completely at random. Because there was no information provided on how the scores were missing, this assumption made the complete cases method undesirable. A multiple imputation method using predictive means matching to estimate missing values was then conducted. This was done through a multiple imputation by chained equations (MICE) implementation in R (R Core Team 2021; Buuren and Groothuis-Oudshoorn 2021). The resulting dataset from multiple imputations was used for the formal analysis.

Means Comparison

Preliminary exploratory analyses suggests GSI scores of the treatment group to be higher than GSI scores of the control group. Using the full imputed dataset, a two-sample t-test found this difference to be non-significant. Examination of normality assumptions using the Shapiro-Wilk test found both the distribution of the treatment group and distribution of the control group to violate the normality assumption. Thus, a two-sample Wilcoxon test was required, and the test found the GSI difference in treatment groups to be significant (Derrick et al. 2017).

Table 3: Results of t-tests and wilcoxon tests for treatment vs. control group comparison

Estimated Difference	Standard Error	p-value (t-test)	p-value (Wilcoxon test)
-0.041	0.037	0.262	0.064

Mixed Effects Models

Mean comparison tests found non-significant differences between the two treatment groups, but a mixed-effects models approach was taken to account for the large variations between individuals and to appropriately model the clustered nature of the data, allowing for individual specific inference as well as population average inference. This allowed the correlation between within-subject measurements to be modeled appropriately. Mixed-effects models also automatically incorporate missing data, meaning the longitudinal data need not be balanced and each individual may have different repeated measurements (Pinheiro and Bates 2000).

Model selections were conducted using ANOVA, first by shortlisting models that produced significant results with the likelihood ratio test, then selecting the final model with the lowest AIC and BIC value (see appendix).

To investigate the difference in GSI scores between treatment groups, a mixed-effects model was constructed, taking the form:

$$GSI_{ijkl} = \beta_0 + \beta_1 treatment_{ijkl} + subject_i + month_j + gender_k + education_l + \epsilon_{ijkl} \quad (1)$$

where GSI is the mental distress score, $treatment$ is an indicator of the treatment group, $subject$ is the random effect for the i^{th} subject, $month$ is the random effect for the j^{th} month, $gender$ is the random effect for the k^{th} gender, $education$ is the random effect for the l^{th} education level, and ϵ_{ijkl} is the random error. This model showed non-significant decreases in GSI when switching from the treatment group to the control group (Table 4, see appendix for full summary).

Table 4: Estimated mean scores and p-values associated with parameters (intercept, and parameter estimates for treatment vs. control groups) for the fitted mixed model of the form shown in Equation 1

Group	Estimated Mean Score	Parameter Estimate	Standard Error	p-value
Treatment	0.924	0.924	0.199	0.011
Control	0.892	-0.033	0.059	0.657

Mixed-Effects Regression

The previous model only provided information on whether the treatment groups were significantly different in GSI scores, but did not specifically investigate whether or not subjects' mental distress in both of the treatment groups decreased significantly over time. To examine the effect of month on GSI scores, a mixed-effects regression approach was taken, in the form:

$$GSI_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})month_{ij} + b_{2i}gender_{ij} + \epsilon_{ij} \quad (2)$$

but this time the month variable was the fixed effect and the variance between month and gender were shared and varied by subject. This model showed that there was a significant decrease in GSI score of approximately 0.005 for each additional 1 month increase (Table 5, see appendix for full summary). Examining the residuals from the regression showed slight evidence of heteroscedasticity, but showed little evidence of violations of the normality assumption for this model (see appendix).

Table 5: Parameter estimates and p-values for the fitted mixed-effects regression model of the form shown in Equation 2

Parameter	Estimated Value	Standard Error	p-value
Mean GSI at Month = 0	1.003	0.035	< 0.001
Estimated Per- Month Change in GSI	-0.005	0.001	< 0.001

4. Discussion

The purpose of this study was to test whether subjects' mental distress would change over time, and whether this change in distress would be different between a treatment/intervention group and a control group. To do this, exploratory data analysis was first performed. Using only complete cases, a non-significant difference in GSI was found between the two treatment groups using the t-test, but a Wilcoxon test found that difference to be significant. Specifically, the median GSI of the treatment group was found to be slightly higher than the median GSI of the control group. Additional preliminary analyses were also done by plotting individual subjects' GSI trend over time. Large variations were found between individuals, suggesting a mixed-effects model may be useful. All formal analyses were performed using data imputed with the multiple imputation method, and this imputation was based on the assumption that the data were missing at random.

The first model was a comparison of means between the treatment and control group. Two different tests: the two-sample t-test and Wilcoxon-test were used for testing. These tests confirmed our preliminary results: the t-test was non-significant and the Wilcoxon-test was significant. The Wilcoxon-test was more robust because the Shapiro-Wilk test found the normality assumptions of the t-test to be violated. The second model was a mixed-effects means comparison model. This model was built upon the previous model, testing the effects of the treatment groups on the GSI score but also incorporating the random effects of subject, month, gender, and education. The results were consistent with the first t-test model, finding no significant difference between the treatment groups. The last model was a mixed-effects regression model testing the

effect of the months variable on GSI score while incorporating the random effect of shared variance between month and gender varying by subject. Results of the final model found a small but significant decrease in GSI score per month. All mixed-effects models were selected based on AIC values, BIC values, and likelihood ratio test significance.

Limitations

Missing data was a major issue in this study. Missing observations were treated with the multiple imputation method relying on the assumption that the data were missing at random. This was preferable to the complete cases method which required an even stronger assumption of the data to be missing completely at random.

Another concern was the violation of the constant variance assumption in the mixed-effects regression model. Examination of residuals in the regression model found evidence of heteroscedasticity, which could potentially bias the decision to reject the null hypothesis.

A third issue was the violation of the normality assumption in the means comparison t-test model. This reduces the power of the t-test and decreases the ability to detect a true difference.

A final limitation was an incongruency between the data available to be used and the research objective of the study. When assessing the relationship between GSI, treatment groups, and months, it would be preferable to account for the effects of more variables in addition to the random effects of gender and education.

Future Recommendations

Recommendations for study design and data collection are offered to lead to more robust studies in the future. First, it would be preferable to have a more even spread of monthly time points where data are collected. Second, more samples at both the individual measurement level and the subject level would allow for more accurate results and a greater number of complete observations.

5. Conclusion

In conclusion, results found a significant decrease in GSI scores for each additional month across all subjects. However, there was not enough evidence to conclude that this change in GSI scores was different between treatment and control subjects. Subsequent studies may have broader results if current limitations on missing data, violation of assumptions, and data collection are resolved. A further study with a larger, more comprehensive sample would be recommended.

References

Buuren, Stef van, and Karin Groothuis-Oudshoorn. 2021. *Mice: Multivariate Imputation by Chained Equations*.

Derrick, Ben, Bethan Russ, Deirdre Taylor, and Paul White. 2017. "Test Statistics for the Comparison of Means for Two Samples That Include Both Paired and Independent Observations." *Journal of Modern Applied Statisitical Methods* 16 (1). https://doi.org/10.22237/jmasm/1493597280.

Pinheiro, José C, and Douglas M Bates. 2000. *Mixed-Effects Models in S and S-Plus*. New York, NY [u.a.]: Springer. http://www.worldcat.org/search?qt=worldcat_org_all&q=1441903178.

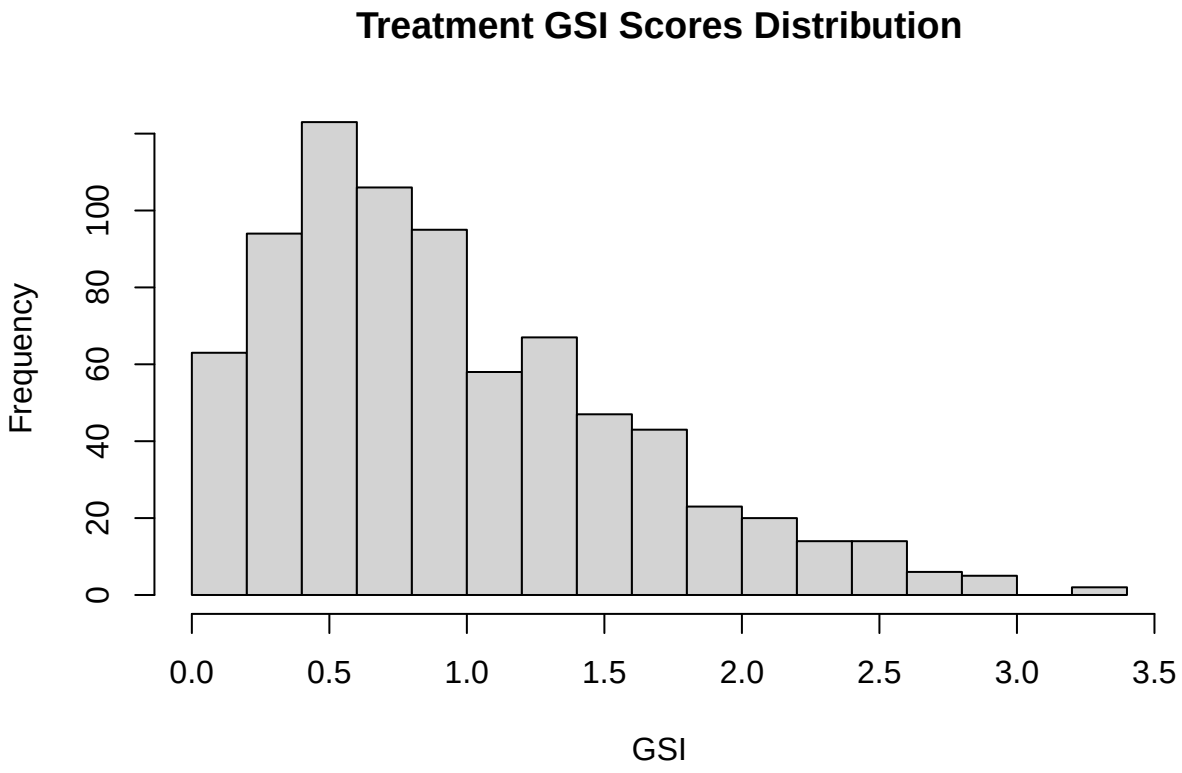
R Core Team. 2021. *R: A Language and Enviroment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org/.

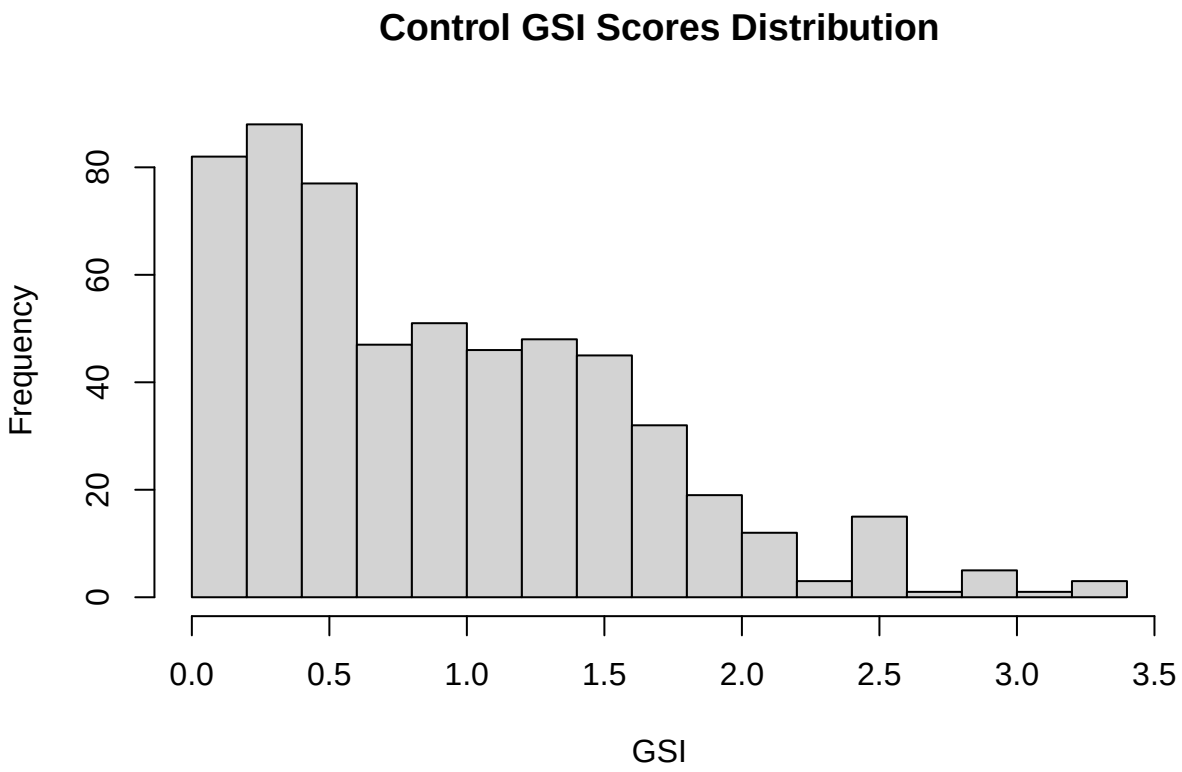
Appendix

Examination of Normality Assumptions

```
##
##  Shapiro-Wilk normality test
##
## data:  dat.treat$GSI
## W = 0.93904, p-value < 2.2e-16
```

```
##
##  Shapiro-Wilk normality test
##
## data:  dat.control$GSI
## W = 0.9301, p-value = 9.871e-16
```





Mixed-Effects Model Selection

Table 6: Mixed-Effects Model Selection with only treatment variable as fixed variable

	AIC	BIC	p-value
me.f	2335.253	2356.100	NA
me.d	2290.152	2321.422	0
me.c	2216.039	2252.520	0
me.b	2336.669	2383.573	1
me.e	2336.669	2383.573	NA
me.a	2288.773	2356.523	0

Table 7: Mixed-Effects Model Selection with only month variable as fixed variable

	AIC	BIC	p-value
me.1	2235.399	2266.669	NA
me.5	2231.644	2268.125	0.0164397
me.3	2239.772	2286.676	1.0000000
me.4	2239.656	2286.560	NA
me.2	2239.855	2307.605	0.0991448
me.6	2249.026	2327.199	1.0000000

Fit statistics for Mixed-Effects Treatment Comparison Model (Equation 1):

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: GSI ~ NEW_GRP + (1 | month) + (1 | DDS1) + (1 | DDS4) + (1 |      SN)
##    Data: dat.mi
##
## REML criterion at convergence: 2207.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8935 -0.5908 -0.1015  0.5120  4.7305
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  SN       (Intercept) 0.17575  0.4192
##  DDS4     (Intercept) 0.05472  0.2339
##  month    (Intercept) 0.02104  0.1450
##  DDS1     (Intercept) 0.04441  0.2107
##  Residual             0.20794  0.4560
## Number of obs: 1355, groups:  SN, 271; DDS4, 14; month, 5; DDS1, 2
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)   0.92450    0.19860   2.87620   4.655   0.0205 *
## NEW_GRP      -0.03260    0.05882 266.39571  -0.554   0.5798
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr)
## NEW_GRP -0.429
```

Fit statistics for Mixed-Effects GSI Across Months Regression (Equation 2):

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: GSI ~ month + (1 + month + DDS1 | SN)
##    Data: dat.mi
##
## REML criterion at convergence: 2240.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2012 -0.5207 -0.1204  0.4735  4.5831
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  SN       (Intercept) 2.644e-01 0.514209
##           month       3.325e-05 0.005766 -0.55
##           DDS1        3.934e-02 0.198346 -0.32 -0.13
##  Residual             1.964e-01 0.443145
## Number of obs: 1355, groups:  SN, 271
```

```
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  1.003e+00  3.508e-02  2.664e+02  28.601  < 2e-16 ***
## month       -4.984e-03  6.464e-04  2.674e+02  -7.711 2.47e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##       (Intr)
## month -0.534
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.136568 (tol = 0.002, component 1)
## Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

Model Diagnostics for Mixed-Effects Regression Model (Equation 2)

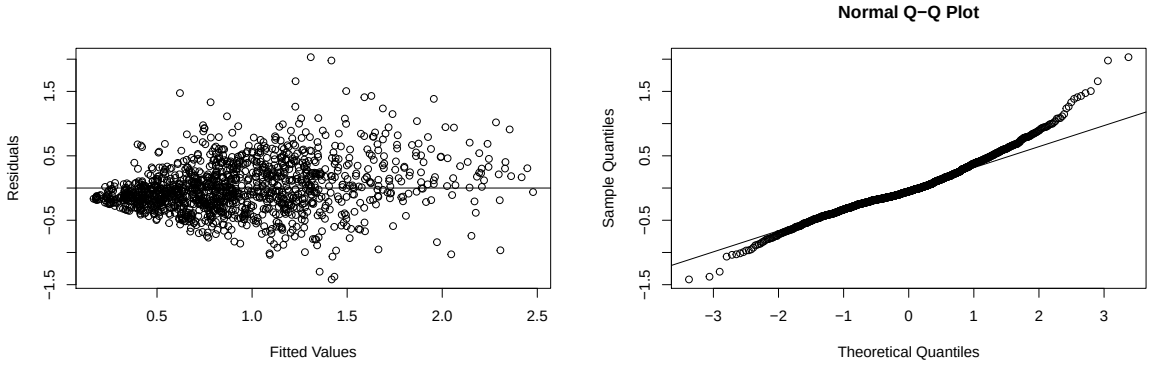


Figure 4: Residual and Normal-Quantile plots for the mixed effect regression model of Equation 2