LONGITUDINAL ANALYSIS OF MENTAL DISTRESS

SHUYI TAN

Abstract. The potential negative impact of mental distress on health and life quality is gaining increasing recognition as an important public health issue. Effective intervention for mental distress is expected to reduce subjects' GSI over time. Based on a longitudinal study that measure the GSI scores of subjects in a intervention group and treatments respectively at five time points, data were analysed using descriptive statistics, visualizations, linear mixed-effect (LME) modelling and generalized linear modeling (GEE). Results from both exploratory and confirmatory analyses reveals that the effectiveness of the intervention is not significant. Subjects from the intervention group show significant decrease over time in the GSI with 0.44 to 0.77 unit of decrement per month, while subjects from from the control group present non-significant reduction in GSI over time with around 0.22-unit decrement per month. In addition to time and intervention, the change in GSI can be partially explained by gender and education. Overall, male subjects with fewer years of education tend to have lower level of mental distress.

1. Introduction

The persistent COVID-19 pandemic has a massive impact on global mental health. [1] It is crucial to raise awareness of mental health because untreated mental illness will contribute to higher medical expenses, poorer performance, fewer employment opportunities and increased risk of suicide.[2] One of the most sensitive indicator of mental distress is the Global Severity Index (GSI), which can quantify a patient's severity-of-illness and provides a single composite score for measuring the outcome of a treatment program based on reducing symptom severity. Namely, a higher GSI score revealing a severer mental health condition. [4]

In this longitudinal study, 271 subjects are randomly assigned into an intervention group and a control group, in which their changes in mental distress over time are investigated based on their responses to questionnaires. The main objective of the study is to investigate whether there exist significant reduction in the subjects' mental distresses over time in each of the group, and to explore the effectiveness of the intervention by checking if the difference in subjects' mental distress between the treatment and the control group is significant.

This report intends to provide statistical analyses on the given longitudinal dataset. The report is organized as follows. The exploratory data analysis presented in Section 2 will shed light on some initial empirical results. Section 3 will provide confirmatory analyses to validate the results from Section 2. Two-sample t-test and Wilcoxon test will be performed to compare the difference between two groups, and the linear mix effect (LME) model and generalized estimating equation (GEE) model will be employed to model and evaluate the changes in GSI over time.

Date: April 17, 2021.

2. **Exploratory Data Analysis**

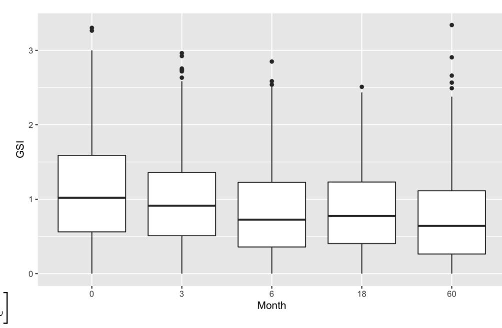
An overview of the dataset is provided to facilitate the reading of following analyses. Furthermore, univariate analyses are performed to seek potential associations between the response and explanatory variables.

2.1. **Dataset Overview.** In this study, 271 subjects were recruited and randomly assigned into an intervention group ($N = 156$) and a control group ($N = 115$). Mental distress in subjects were measured at the month= 0 (baseline), 3, 6, 18, and 60 months, based on their answers to questionnaires. Description of variables are shown in Table 1.

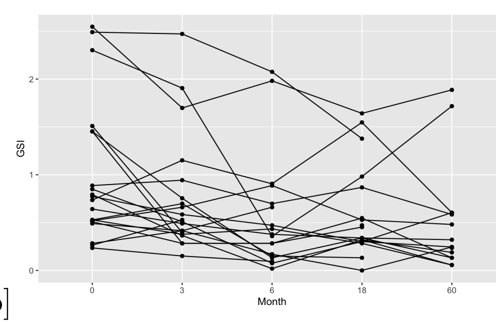
Variable	**Type**	**Range/Level**
Study Group	Factor	Intervention, Control
Month	Factor	0, 3, 6, 18, 60 (in months)
Gender	Factor	Female, Male
Education	Numeric	8 - 21
GSI	Numeric	0.00 - 3.39

Table 1. Description of Variables

The main variable of interest in this study is the value of GSI at each time point. Plot[a] of Figure 1 demonstrates the integral longitudinal trend of the GSI over time, which reveals an overall decreasing trend except the slight increase at the time point of 18 months. Checking each box in plot[a] of Figure 1, the distributions of GSI at each time points are similar and approximately symmetric. Plot[b] of Figure 2 shows the individual trajectory of GSI among 20 randomly-selected subjects, which suggests that there may exist correlation as well as large variations between subjects. The phenomenon of large variations between individuals indicates that a linear mixed effect model may be appropriate for this data.



[a]



[b]

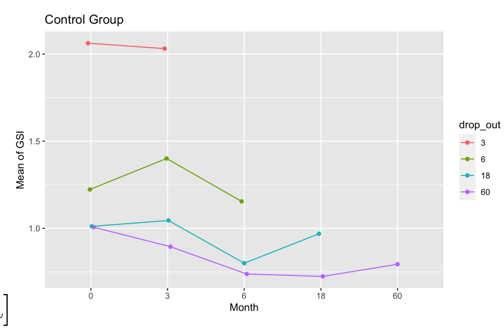
Figure 1. Left: Longitudinal Trend of GIS for All Subjects; Right: Individual Trajectory of GSI of 20 Randomly-selected Subjects

2.2. **Data Missingness.** There are 1355 observations in total for 271 subjects. The proportions of data missingness at each time point are reported in Table 2. As shown in Table 2, the proportions of data missingness increase over time. Unfortunately, the knowledge about the missing mechanisms of this study is not clear. Intuitively, subjects whose GSI are missing at all time points or available only at one time points will be removed. The histogram of GSI scores at each time point demonstrates similar shapes, which indicates similar distribution of GSI scores among different periods.

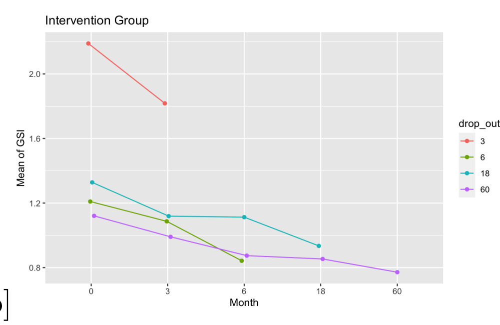
Time Point (in months)	0	3	6	18	60
Count	10	38	52	105	98
Rate	3.7%	14.0%	19.2%	38.7%	36.2%

Table 2. Data Missingness at Each Time Point

To further alleviate the selection bias, subjects are stratified based on the time point of drop-out, and the following figure (Figure 2) plots the mean response (GSI) over time for the control group and intervention group respectively. The plots suggest that subjects who complete only the measurement at month 3 have a much higher mean baseline GSI compared to all other subjects. Furthermore, subjects who complete the whole study at 60 months have lower mean GSI than that of subjects that do not complete the entire measurements. In summary, it suggests that the "completers" and the "drop-out" subjects differ with respect to their GSI score. In other words, it would be inappropriate to only retain subjects with available records at all time points to do a "complete-case" analysis. At this point, the data will be assumed to be approximately missing at random (MAR), which means that the probability of a missing response is independent of its current and future responses conditional on the observed past responses and the covariates. [7] Whether to ignore or impute these missing values will depend on specific models that are used in the confirmatory modeling section.



[a]



[b]

Figure 2. Patterns of Mean GSI by Dropout Time

2.3. **Outliers.** By checking the plot [a] of Figure 1 again, it can be clearly observed that there are a few extremely values in the dataset (black dots at the top of each box), which may lead to misleading results in both empirical and modeling results in later analyses. One of the commonly-used methods to detect and tackle outliers is the Interquartile Range (IQR) method. The idea is that any observations that are more than 1.5 IQR below Q1 or more than 1.5 IQR above Q3 are considered outliers, with IQR describing the middle 50% of values when ordered from lowest to highest.[5] The following analysis will be based on the dataset with outliers removed.

2.4. **Univariate Analysis.** The effectiveness of the intervention is of particular interest in this study. The boxplot from Figure 3 along with the descriptive statistics in Table 3 present the cross-sectional effect of intervention in comparison with the control group. It can be observed that in both groups the GSI decrease over time. Overall, the GSI scores of subjects in the intervention groups are slightly higher than those in the control group.

However, the mean response of two groups are very close at the thrid month and $60th$ month.

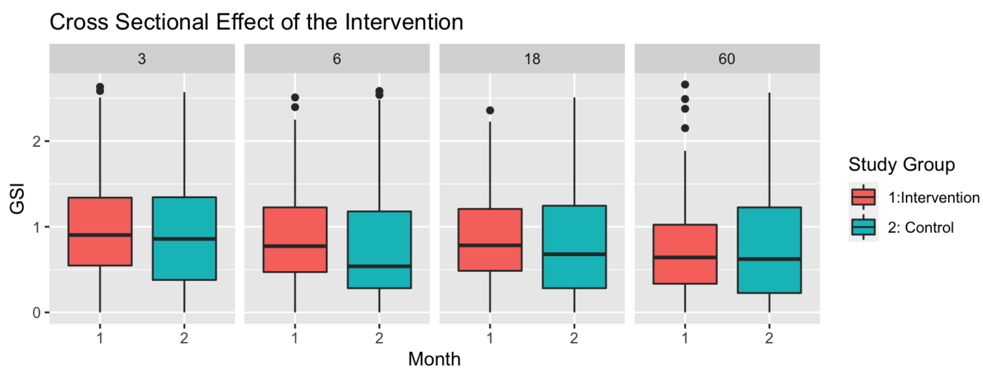


FIGURE 3. Cross-sectional Effect of the Intervention at Each Time Point

Month	3		6		18		60	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Control	0.941	0.687	0.800	0.691	0.794	0.607	0.756	0.607
Internvention	0.999	0.598	0.874	0.578	0.865	0.526	0.751	0.554

TABLE 3. Cross-Sectional Descriptive Statistics

In order to confirm if there is a significant difference in subjects' mental distress between the treatment and the control group, two-sample t-test and Wilcoxon test will be conducted at each time point. The normality of data at each time point is assured by Shapiro–Wilk test, and the assumption of homogeneous variance is also confirmed by a F test at each time point. The results of both tests are summarized in Table 4. As shown in Table 4, the p-values from both tests are large, which is a straightforward signal indicating that the difference between two groups is not statistically significant.

	Two-sample T-test		**Wixcoxon Test**	
Month	Test Statistic	P-value	Test Statistic	P-value
3	0.642	0.522	6585.0	0.2435
6	0.818	0.414	6448.5	0.1003
18	0.784	0.434	3711.5	0.1874
60	- 0.050	0.960	3572	0.7819

TABLE 4. Two-sample T-test and Wilcoxon Test between the Intervention group and the Control Group

3. **Confirmatory Data Analysis**

In order to validate the results from exploratory data analysis, linear mixed effect (LME) models and GEE models will be applied and compared in this section. We will first analyze the change of GSI over time in each group separately (Section 3.1 and Section 3.2), and then the intervention will be treated as a predictor along with other covariates in the models over the whole dataset (Section 3.3).

Linear mixed effect (LME) models can be thought of as a trade-off between noisy individual regressions and less-noisy aggregate that may lose important differences by averaging samples, which are an extension of simple linear models to allow both fixed and random effects. [6] LME models make specific assumptions about the variation in observations attributable to variation within a subject and to variation among subjects. From plot[b] of Figure 2, we have observed large variations between individuals and there may exist correlations among subjects, which suggests that a LME model may be appropriate and give useful insights in this study. Since a LME model use all data simultaneously, the missing values can be ignored here because the data is assumed to be MAR.

The generalized estimating equation (GEE) model is an alternative model to analyze the longitudinal data, which does not require distributional assumptions. Compared to the LME model, a GEE model does not allow for individual-specific inference in absence of random effects. Instead, it models the average response as an extension of generalized linear models (GLM) to longitudinal data. [3] It is worth awareness that the standard GEE method is valid only if the data are missing completely at random (MCAR), but it can lead to biased results if the data are missing at random (MAR).[7] Therefore, before fitting a GEE model to the data, missing data will be imputed using the method of multiple imputation with the R package 'MICE', which generates imputed datasets based on observed values. Since the GEE estimates are consistent even if the variance-covariance matrix is misspecified, the GEE models in the following modeling analysis will always adopt an exchangeable covariance structure. For the purpose of sensitivity analysis, we will also look at the model fit of a weighted GEE model, which is complementary to the standard GEE model when the data is missing at random (MAR) rather than missing completely at random (MCAR) before imputation.

3.1. **Intervention Group.** To investigate if the subjects' mental distresses in the intervention group decrease significantly over time, an LME model and two GEE models will be applied to the data of the intervention group. To determine the choice of the random effect for the LME model, a simple linear model is built for each individual to check how the estimates vary across individuals based on a 95% confidence interval (Figure 4 in the appendix). It reveals that random effect may be required for both intercept and slope in a single linear mixed effect model because it is shown in the plot that many confidence intervals do not overlap. To reach a more rigorous decision, both a random intercept and a random intercept and slope model are fitted to the data. By conducting an ANOVA on these two models, the AIC of the random intercept and slope model is higher than that of random intercept model, while the BIC values of both models are very close. Moreover, the likelihood ratio test showed significant result (p = 0.0043) if we use the commonly-used significance level ($\alpha = 0.05$). Hereby a random intercept and slope model will be adopted. In terms of model assumptions, besides the assumption of large between-individual variations, a LME model should also satisfy the assumptions including the normality of residuals and homogeneity of variances of residuals, which have been verified and the results are shown in Figure 5 and Figure 6. Additionally, a GEE model and a weighted GEE model is fitted to the imputed data and unimputed data respectively. Detailed outputs from LME and GEE are summarized in Table 5.

Overall, the results from three models are mostly consistent. The regression estimates of the time variable *month* from three models confirms our observations from EDA (Section

Model		Estimate	SE	P-value
LME	Intercept	1.0293	0.0464	0
	Month	-0.0047	0.0008	0
GEE	Intercept	1.0276	0.0417	0
	Month	-0.0044	0.0008	0
Weighted GEE	Intercept	1.0276	0.0460	$<2 \times 10^{-16}$
	Month	-0.0044	-5.452	5×10^{-8}

Table 5. Results of the LME and GEE models for the Intervention Group

2.4) that the mental distress in the intervention group decrease over time, which is 0.47-unit decrement in the LME and 0.44-unit decrement in GEE per month. Moreover, the p-values of *month* indicates its statistical significance.

3.2. **Control Group.** Following the same procedure as the analysis of the intervention interval, the changes in the mental distress of subjects from the control group are also investigated. An LME model with random intercept is selected and the diagnostic plots for model selection and assumptions (Figures 7, 8, and 9) are provided in the appendix. The modeling outputs are displayed in Table 6 below.

Model		Estimate	SE	P-value
LME	Intercept	0.9343	0.0615	0
	Month	-0.0023	0.0008	0.0051
GEE	Intercept	0.9340	0.0596	0
	Month	-0.0023	0.0009	0.0078
Weighted GEE	Intercept	0.9340	0.0627	0
	Month	-0.0023	0.0008	0.0052

Table 6. Results of the LME and GEE models for the Control Group

It can be seen from Table 6 that the GSI score in the control group is decreasing over time, which echoes with the results from EDA (Section 2.4). To further quantify the magnitude of the decrease, on average there is approximately 0.0022 to 0.0023-unit decrement in GSI per month in the control group. However, the p-values are all larger than 0.05, which is the commonly-used significance level. Compared to the results from the intervention group, the decrease of GSI over time in the control group seems to be less statistically significant or even not statistically significant.

3.3. **The Whole Dataset.** In addition to analyzing two groups separately, an attempt to adjust the treatment as a binary predictor along with other explanatory variables to investigate whether the changes in mental distress over time can be partially explained by the treatment and other variables (e.g., gender and education) is launched. The modeling procedure was repeated as in previous subsections. For LME modeling, a random intercept model is selected based on the values of AIC, BIC, and a likelihood ratio test. The model outputs are summarized in Table 7. The assumptions of normality of homogeneity of variance of residuals are verified and presented in Figures 10 and 11.

Model		**Estimate**	**SE**	**P-value**
LME	Intercept	1.5126	0.2159	0
	Intervention	0.0609	0.0673	0.366
	Month	-0.0036	0.0005	0
	Female	0.2805	0.0712	0
	Education	-0.0546	0.0144	0
GEE	Intercept	1.5100	0.2083	0
	Intervention	0.0608	0.0650	0.349
	Month	0.0036	0.0006	0
	Female	0.2802	0.0687	$4.49 \cdot 10^{-5}$
	Education	-0.0544	0.0139	$8.712 \cdot 10^{-5}$
Weighted GEE	Intercept	1.5100	0.2086	0
	Intervention	0.0608	0.0679	0.371
	Month	-0.0036	0.0006	0
	Female	0.2802	0.0667	$2.62 \cdot 10^{-5}$
	Education	-0.0544	0.0136	$6.34 \cdot 10^{-5}$

Table 7. Results of the LME and GEE models over the Whole Dataset

It can be observed from Table 7 that the estimates from the three models are very close. Each predictor will be inspected respectively. For the intervention, with not going through intervention as the baseline factor, we can see that there is a positive association between the intervention and GSI. However, the effect of the intervention is not significant with large p-values, which echoes with the finding in the EDA (section 2.4) that the difference in GSI between the intervention group and control group is not significant. For the time variable *month*, its estimate confirms the overall decreasing trend of the GSI over time as a significant predictor. Meanwhile, it is of interest whether the changes in mental distress over time are partially affected by gender and education. As shown in Table 7, both gender and education are significant covariates to GSI. Specifically, female yields approximately 0.2802-unit higher GSI than the overall mean, and there is 0.005-unit decrement in GSI per year of education.

4. **Conclusion and Limitations**

Based on the results from exploratory analysis and confirmatory analysis, subjects' mental distress in each of the groups both demonstrate a decreasing trend, with the decrease in the intervention being significant while that in the control being less or even not significant. For the effectiveness of the intervention, the current data does not support a significant difference in GSI score between two groups of subjects, which is further confirmed by the LME and GEE models in Section 3.3. In addition, it is found that there exist statistically significant associations between our response GSI and gender and education. Specifically, male subjects with fewer years of education are more likely to have lower GSI.

There exists some limitations regarding these analyses. The dataset contains a relatively large percentage of missing data, which imposes one of the biggest uncertainty on the analysis because of the lack of knowledge in the data missing mechanism of this study. More information on the collection procedure and subject backgrounds will be required to support the validity of the data missing assumption.

References

[1] Kola, Lola. *Global mental health and COVID-19.* Lancet Psychiatry. 2020;7(8):655-657. doi:10.1016/S2215-0366(20)30235-2

[2] Holthaus, jeans. *The Importance of Mental Health Awareness.* Christian Mental Health Service. Received from https://www.pinerest.org/mental-health-awareness-blog/

[3] Columbia University Mailman School of Public Health. *Repeated Measures Analysis* Columbia University. Received from https://www.publichealth.columbia.edu/research/population-health-methods/repeated-measures-analysis

[4] Derogatis, Leonard R *Quickly measure psychological symptoms.* Brief Symptom Inventory (BSI). Pearson.

[5] PennState. *3.2 - Identifying Outliers: IQR Method.* Penn State Eberly College of Science. Penn State University. Received from https://online.stat.psu.edu/stat200/lesson/3/3.2

[6] UCLA: Statistical Consulting Group. *INTRODUCTION TO LINEAR MIXED MODELS*. Institute for Digital Research & Education. University of California Los Angeles. Receivde from https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/

[7] Lin, Guixin. Rodriguez,Robert N. *Weighted Methods for Analyzing Missing Data with the GEE Procedure* SAS Institute Inc.Paper SAS166-2015

Appendix A.
Intervention Group

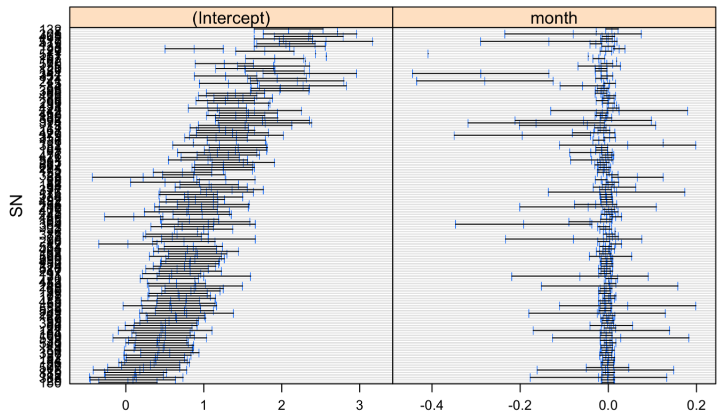
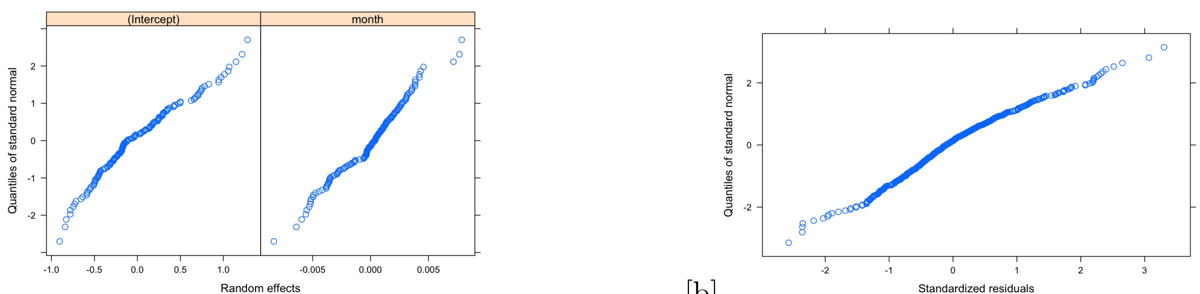


Figure 4. Individual Linear Regression Fits for the Intervention Group



[a] [b]

Figure 5. Left: QQ Plot for Random Effect for the Intervention Group; Right: QQ Plot for Residual for the Intervention Group

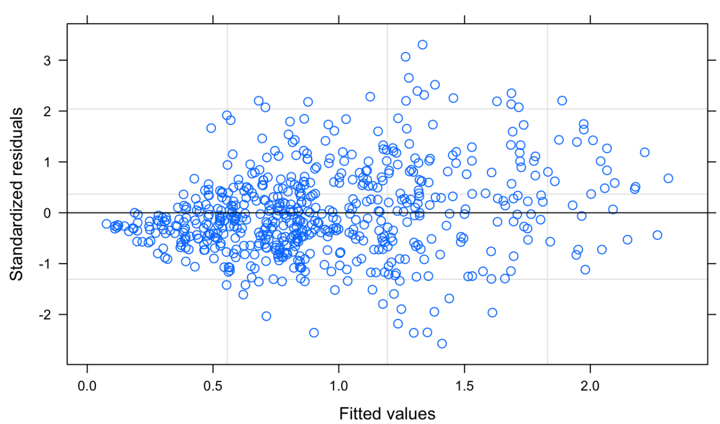


Figure 6. Residual vs. Fitted Value of the Random Intercept and Slope Model for the Intervention Group

Appendix B.
Control Group

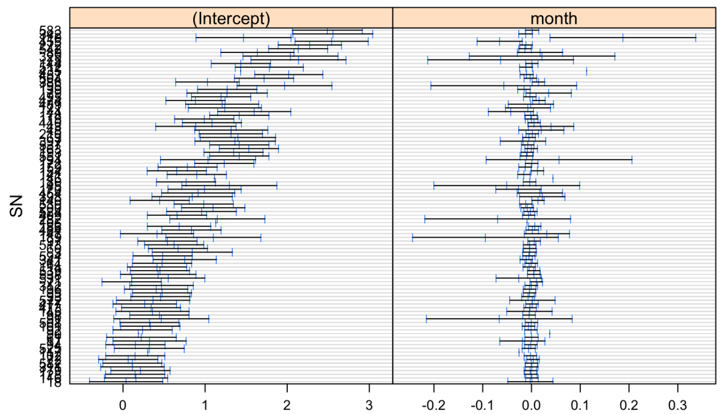
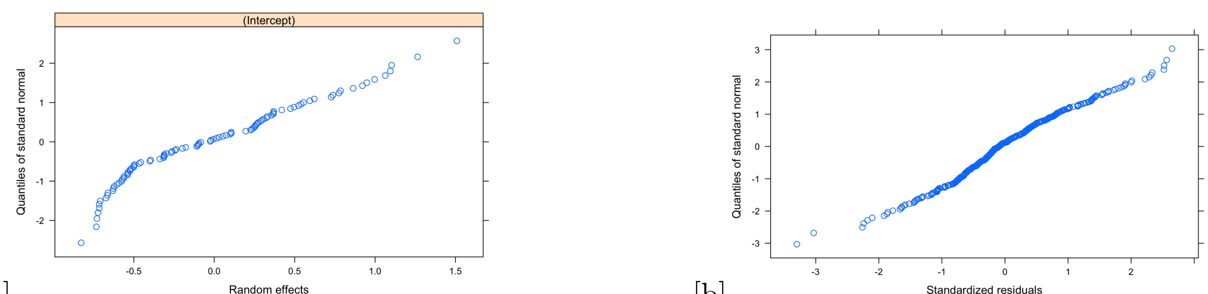


Figure 7. Individual Linear Regression Fits for the Control Group



[a] [b]

Figure 8. Left: QQ Plot for Random Effect for the Control Group; Right: QQ Plot for Residual for the Control Group

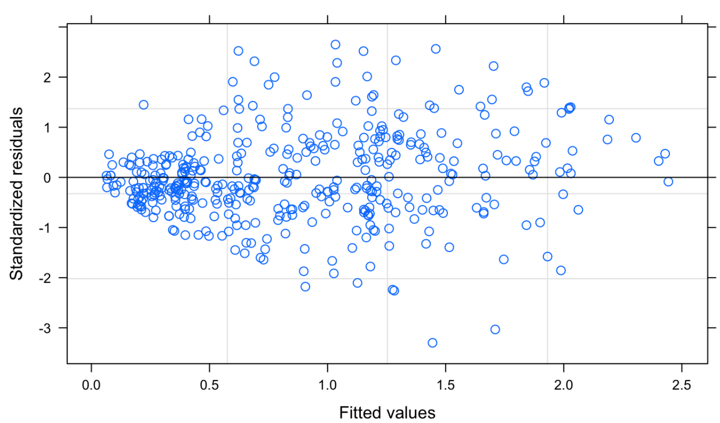


Figure 9. Residual vs. Fitted Value of the Random Intercept Model for the Control Group

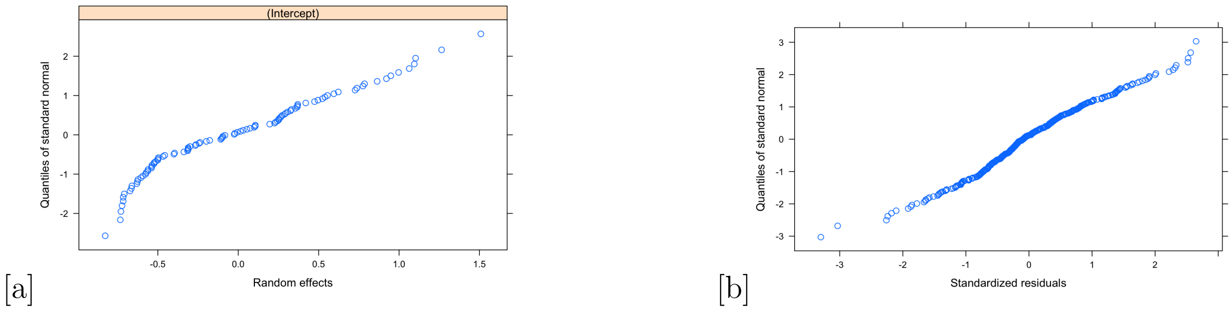


Figure 10. Left: QQ Plot for Random Effect for the Whole Dataset; Right: QQ Plot for Residual for the for the Whole Dataset

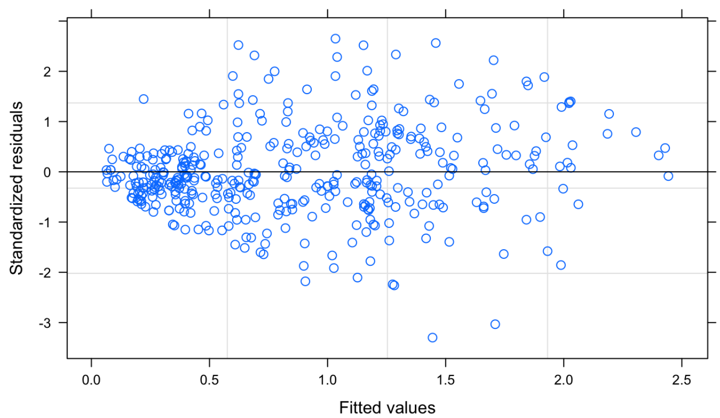


Figure 11. Residual vs. Fitted Value of the Random Intercept Model for the for the Whole Dataset