

AN ANALYTICAL APPROACH COMPARING REPEATED-MEASURES ANALYSIS OF
VARIANCE (ANOVA) AND MIXED MODELS IN A DOUBLE-BLIND PLACEBO-
CONTROLLED CLINICAL TRIAL

by

Amie E. Sagady

BA, University of Michigan, 2003

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Amie Sagady

It was defended on

May 23, 2005

and approved by

Benjamin Handen, Ph.D., Associate Professor, Departments of Psychiatry and Pediatrics, School of Medicine, University of Pittsburgh

Sheryl F. Kelsey, Ph.D., Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Thesis Advisor: Ada O. Youk, Ph.D., Research Assistant Professor, Graduate School of Public Health, University of Pittsburgh

Ada O. Youk, Ph.D.

AN ANALYTICAL APPROACH COMPARING REPEATED-MEASURES ANALYSIS OF
VARIANCE (ANOVA) AND MIXED MODELS IN A DOUBLE-BLIND PLACEBO-
CONTROLLED CLINICAL TRIAL

Amie E. Sagady, M.S.

University of Pittsburgh, 2005

ABSTRACT

Longitudinal studies are common in many areas of public health. A usual method to analyze longitudinal data is by repeated-measures analysis of variance (ANOVA). A newer method, the mixed models approach, is gaining more acceptance due to the available use of computer programs. It is of public health importance to review the advantages of the recent mixed models approach to analyzing longitudinal data.

The main characteristic of longitudinal studies is that the outcome of interest is measured on the same individual at several points in time. The standard approach to analyzing this type of data is the repeated-measures ANOVA, but this type of design assumes equal correlation between individuals and either includes data from individuals with complete observations only or imputes missing data, both of which suffer from the ineffective use of available data. Alternatively, the mixed model approach has the ability to model the data more accurately because it can take into account the correlation between repeated observations, as well as uses data from all individuals regardless of whether their data are complete.

This thesis first reviews the literature on the repeated-measures ANOVA and mixed models techniques. Data from a placebo-controlled clinical trial of the drug methylphenidate

(MPH) looking at the social/play behavior of children with attention deficit hyperactivity disorder (ADHD) and mental retardation (MR) are analyzed using repeated-measures ANOVA, repeated-measures ANOVA with the last observation carried forward (LOCF) and mixed models techniques. P-values and parameter estimates for the three methods are compared.

MPH had a significant effect on the variables Withdrawn and Intensity in both of the repeated-measures analyses. With the repeated-measures with LOCF, MPH had a significant effect on the variables Activity Intensity Level and Sociability. The mixed models analysis found MPH to have a significant effect on the variables Intensity and Activity Intensity Level. The parameter estimates for the two repeated-measures ANOVA analyses were almost identical, but the mixed model parameter estimates were different. Mixed models should be used to analyze these data as assumptions of the repeated-measures ANOVA are violated. Mixed models also take into account the missing data and correlated outcomes.

TABLE OF CONTENTS

1.	CHAPTER ONE	1
1.1.	INTRODUCTION	1
1.2.	STATEMENT OF THE PROBLEM	3
2.	CHAPTER TWO	4
2.1.	REVIEW OF THE LITERATURE	4
2.1.1.	Repeated-Measures ANOVA.....	4
2.1.2.	Mixed Model.....	6
3.	CHAPTER THREE	8
3.1.	METHODS	8
3.1.1.	Subjects.....	8
3.1.2.	Setting.....	8
3.1.3.	Procedure	8
3.1.4.	Dependent Measures.....	9
4.	CHAPTER FOUR.....	12
4.1.	ANALYSIS.....	12
4.1.1.	Assessing Normality	12
4.1.2.	Repeated-Measures ANOVA Analysis.....	12
4.1.3.	Mixed Model Analysis.....	14
5.	CHAPTER FIVE	15

5.1. RESULTS	15
5.1.1. Descriptive Analysis	15
5.1.2. Missing Data	18
5.1.3. Normality Assumption.....	18
5.1.4. Model Results	21
5.1.5. Parameter Estimate and Confidence Interval Results.....	23
6. CHAPTER SIX.....	30
6.1. DISCUSSION.....	30
APPENDIX.....	33
PROGRAMS FOR ANALYSIS.....	33
BIBLIOGRAPHY.....	47

LIST OF TABLES

Table 1. ADHD/MR Subject Demographic Information.....	15
Table 2. Means and Standard Deviations for Drug Conditions.....	17
Table 3. P-values for each model.....	22

LIST OF FIGURES

Figure 1. Q-Q Plots of Transformed Percentile Variables.....	19
Figure 2. Q-Q Plots of Transformed Numeric Variables.....	20
Figure 3. Withdrawn Parameter Estimates.	23
Figure 4. Solitary Parameter Estimates.....	24
Figure 5. Interactive Parameter Estimates.	25
Figure 6. Rough and Tumble Parameter Estimates.	26
Figure 7. Rule Violation Parameter Estimates.....	27
Figure 8. Intensity Parameter Estimates.	28
Figure 9. Aggression Parameter Estimates.	29

1. CHAPTER ONE

1.1. INTRODUCTION

Longitudinal studies are gaining in popularity for their relevance in areas of public health, medicine, and the social sciences. With these studies, the outcome of interest is measured on the same individual at several different points in time. Statistical techniques that take into account the repeated observations on each individual are required. These repeated observations are correlated because the observations of one individual over time are not independent of each other (Twisk, 2003). Standard approaches, such as analysis of variance (ANOVA) and regression techniques are commonly used to analyze longitudinal data, mostly because of simplicity; however, they have limitations. Recent approaches such as Mixed Models (which is also regression model) are becoming more common because of available computer programs and the ability to model longitudinal data more accurately. Two major issues in longitudinal studies include correlation among repeated outcome measurements and missing data.

The standard repeated-measures analysis assumes equal correlation between individuals. Assuming equal correlation when it is not can result in a higher Type I error (the probability of concluding the treatments differ, when in reality they do not differ), lower statistical power (the probability of detecting a difference between the treatments if the treatments do in fact differ), and incorrect confidence intervals (Edwards, 2000).

Alternatively, more recently developed statistical approaches can take into account different correlations between observations in longitudinal data.

Another issue in longitudinal data analysis is missing data. Missing data occurs when a measurement is not available at a certain time or a subject refuses to complete some of the tests. There are three types of missing data: (1) missing completely at random (MCAR: not dependent on either observed or unobserved outcomes); (2) missing at random (MAR: dependent on observed outcomes but not the unobserved outcomes); (3) missing not at random (MNAR: dependent on the unobserved outcomes). Frequently, missing data are related to the outcome, and thus the data are not MCAR. Certain types of analysis assume that missing data are MCAR, even when they are not, and may therefore bias estimates of treatment effects and the associated standard errors (Little and Rubin, 1987).

Standard statistical approaches to this problem have included using data only from individuals with complete observations or imputing missing data, one common method being carrying the last observation forward (LOCF). Both of these approaches suffer from the ineffective use of the available data (Petkova et al., 2002). Omitting individuals with incomplete data can lead to sample bias, as the group of people who complete data may not be representative of the whole population, and imputing missing data by LOCF can lead to biased treatment effects (Gueorguieva and Krystal, 2004). Alternatively, more recently developed statistical approaches allow all available information on each subject, weighted appropriately, to be incorporated into the analysis.

1.2. STATEMENT OF THE PROBLEM

Dr. Benjamin Handen and his colleagues from the University of Pittsburgh have conducted double-blind, placebo-controlled clinical trials with the drug methylphenidate (MPH) among children with attention deficit hyperactivity disorder (ADHD) and mental retardation (MR). Outcome measures studied have included behavioral ratings, work output, measures of learning, attention and impulsivity, and direct observation of peer social interactions (Handen, et al., 1992). Other outcome measures studied included measures of independent play and restricted academic task (Handen et al., 1995). Both of these studies were analyzed using repeated-measures ANOVA, and contained missing data. For the 1992 study, a second analysis was conducted using mean replacement for the incomplete or missing data. For the 1995 study, missing data were imputed using a maximum likelihood technique. Using a mixed models analysis instead of repeated-measures ANOVA could result in a more accurate estimate of treatment effects because all of the data for each subject are incorporated into the analysis and the correlation between repeated outcomes can be modeled realistically.

Dr. Handen conducted another double-blind, cross-over design clinical trial from 1992-1994 which looked at the use of MPH and the social and play behavior in 23 children with ADHD and MR. Because these data were never analyzed, this thesis will address the aim of the study by comparing results using the repeated-measures ANOVA approach with complete data only, the repeated-measures ANOVA approach using the LOCF imputation method and a mixed models approach. Results of the three methods are compared to determine if the three approaches give the same results regarding the effects of the drug condition on the social and play behaviors using p-values, parameter estimates and their corresponding confidence intervals.

2. CHAPTER TWO

2.1. REVIEW OF THE LITERATURE

2.1.1. Repeated-Measures ANOVA

The repeated-measures ANOVA model compares group means of a dependent variable across repeated measurements of time. There is a within-subjects factor, as well a between-subjects factor. Time is often the within-subjects factor because different measurements on the same subject are at different times, and treatment is often the between-subjects factor because levels of treatment can change only between subjects (Littell et al., 1996). This method of analysis has strong assumptions about the data, many of which are not typically satisfied in medical data.

The model can be expressed in matrix notation as

$$y_{ij} = \mu + \pi_i + \tau_j + e_{ij},$$

where

y_{ij} = response from subject i at time j ,

μ = overall mean,

π_i = random effect for subject i which is constant over all occasions,

τ_j = fixed effect of time j ,

e_{ij} = random error component specific to subject i at time j (Davis, 2002).

The first assumption is equal variance of the outcome across treatment groups. Often in placebo-controlled clinical trials, the treatment group has a larger variance at the end of the study than does the placebo group. Although the two groups start out at

similar severity levels because of the inclusion/exclusion criteria, the group that gets the active medication usually responds to the treatment, which results in an increased variance of the outcome. However, the variance of the placebo group remains almost as it was at baseline.

The second assumption is there should be balanced data; that is, after repeated observations, all individuals should have the same number of measurements. Unfortunately, studies do have missing values due to missed visits, dropouts, or subjects not completing all questionnaires, etc. Repeated-measures ANOVA requires the use of data from subjects with complete measurements only or imputation of missing data. LOCF is the most common imputation technique. In this method, the last observed value of the variable is substituted. The problem is that this technique could introduce bias.

The third assumption is equal time intervals between repeated observations on the same unit. The repeated-measures ANOVA model treats time as a nominal variable, which can be a limitation. For example, suppose that measurements for two subjects are taken on the same day. If the second measurement of the first subject is taken seven days after the initial measurement, and the second measurement of the second subject is taken nine days after the initial measurement, one cannot assume that these two measurements were taken at the same level of “time” (Petkova, et al., 2002).

One last assumption is equal correlations between repeated observations on a subject, meaning that the correlation between measurements at time 1 and time 2 are the same as the correlation between measurements at time 1 and time 3, etc. However, consecutive observations on the same subject tend to be more highly correlated than observations on the same subject taken farther apart in time. When this happens, the

Type I error is inflated, and there is an overestimation of the statistical significance of the treatment effect (Gueorguieva and Krystal, 2004).

2.1.2. Mixed Model

The mixed model is a regression model that analyzes correlated continuous outcomes as a function of fixed effects, while simultaneously modeling individual subject parameters as random effects (Edwards, 2000). A factor is a fixed effect when the levels in the study represent all the possible levels of a factor, or at least all the levels about which an inference is to be made. A random effect is when the levels in the study represent a random sample of the target population (Littell et al., 1996).

The model can be expressed in matrix notation as

$$y = X\alpha + Z\beta + e,$$

where

$$y = (y_1, y_2, y_3, \dots, y_n)' = \text{observed values,}$$

$$\alpha = (\mu, \alpha_1, \alpha_2, \dots, \alpha_p)' = \text{fixed effects parameters,}$$

$$\beta = (\beta, \beta_1, \beta_2, \dots, \beta_q)' = \text{random effect/coefficient parameters,}$$

$$e = (e_1, e_2, e_3, \dots, e_n)' = \text{residuals (Brown and Prescott, 2003).}$$

Unlike the standard repeated-measures ANOVA model, the mixed model is based on less restrictive assumptions. The correlation between observations is one assumption that may be relaxed with mixed models. In mixed models, the covariance structure of the data can adequately model data in which observations are not independent. Mixed models can also be applied to unbalanced data and to repeated measurements taken at unequal time intervals.

The mixed model has many important advantages over the standard repeated-measures ANOVA. It uses all available data on subjects, is unaffected by missing data as long as data are missing at random, can flexibly model time effects, and allows the use of realistic variance and correlation patterns, which results in a more accurate treatment effect and standard error estimates and helps control Type I error (Gueorguieva and Krystal, 2004).

There have been several examples of mixed models used in medicine that have appeared in literature, but the use of mixed models has not yet become routine (Brown, 2003). This can be attributed to the complexity of the models. Recent introduction of reliable software to the general public is likely to change this situation.

3. CHAPTER THREE

3.1. METHODS

This double-blind, cross-over design study looked at the use of MPH and the social and play behavior in children with ADHD and MR.

3.1.1. Subjects

Twenty-three children with moderate MR to borderline intellectual functioning served as subjects. They were also diagnosed with ADHD, based on a score of 15 or more points on both the Parent and Teacher Conners Hyperactivity Index. Exclusionary criteria included autism, pervasive developmental disorder, or significant motor impairment.

3.1.2. Setting

Subjects participated in a 6-week laboratory school program on Saturdays at Children's Hospital of Pittsburgh. The program ran from 9:00am until 3:00pm for six consecutive weeks. During the week the subjects remained in their regular special education classrooms.

3.1.3. Procedure

The study involved a double-blind, crossover design with 0.3mg/kg and 0.6 mg/kg doses of MPH and a placebo. Subjects were randomized into one of six drug-placebo orders (1) placebo, .3 mg/kg, .6 mg/kg, (2) placebo, .6 mg/kg, .3 mg/kg, (3) .3 mg/kg, placebo, .6 mg/kg (4) .3 mg/kg, .6 mg/kg, placebo (5) .6 mg/kg, placebo, .3 mg/kg or (6) .6 mg/kg, .3 mg/kg, placebo). The first two Saturdays of the study were used to take baseline

measurements during the sessions. The next three Saturdays were used to take measurements with the three different drug conditions. The last Saturday was used as a make-up day in case a subject missed one of the previous Saturday sessions. A dose of MPH was given twice daily during weeks 3 to 5, the first with breakfast and the second with lunch.

Parents gave subjects the Saturday morning MPH dose at 8:15am as confirmed by project staff. Program personnel gave the Saturday lunch MPH dose and dependent measures were taken between 9:15am and 11:15am and between 12:45pm and 2:45pm. Each Monday, program staff contacted school nurses to ensure that medication had been received. Parents were asked to return all prescription bottles, including any unused pills. In order to answer questions and to ensure compliance with the protocol, both teachers and parents were contacted weekly (Handen, et al., 1992).

3.1.4. Dependent Measures

A twenty-two minute play session, involving three to four subjects placed in a 17 x 22 foot playroom, was conducted during each drug condition with live coding by three trained research assistants from behind a one-way mirror. In addition, sessions were videotaped in the event questions arose regarding the coding. Prior to each child entering the room, 12 toys were arranged in a circle in the center of the room. Five toys were chosen to be particularly appealing to boys (e.g., He-Men characters, GI Joe characters, Rock'em & Sock'em Robots, a bo-bo doll, and a pinball game), three toys were chosen to be particularly appealing to girls (e.g., Barbie Dolls, a make-up vanity, and pompoms), and four toys were chosen to be appealing to both genders (e.g., Playdoh, an electronic flute, crayons with coloring books, and a Waterfuls toy) (Handen, et al., 1995). A teacher

sat the group down and briefly reviewed the playroom rules (e.g., no hitting) and then remained seated in the room by the door (reading a magazine) during the play session.

The coding involved a 10-second observe/5-second record system with coders shifting their observation of subjects every 2.5 minutes, so that a sample of behavior was taken from the beginning, middle, and end of the session for each subject. Coders were unaware of dose level. The following behaviors were recorded:

Withdrawn. This is defined as sitting or standing alone and not engaging with either play materials or peers. The child may be staring or watching others play, but does not respond to teacher or peer initiatives. The child remains this way for the entire interval.

Solitary. This is defined as playing with toys alone. The child is not engaging in any verbal or physical activity with others, but the child may be moving from one activity to another.

Interactive. This is defined as exhibiting prosocial behavior toward one or more peers (not adults). Examples include asking for help or assistance from a peer, laughing with another peer (but not at a peer), or offering a greeting to another peer.

Withdrawn, Solitary and Interactive are mutually exclusive, meaning a child can only be coded as one out of the three.

Rough and Tumble. This is defined as engaging in vigorous physical play with other peers. Examples include running, skipping, chasing other peers, and wrestling with peers.

Rule Violation. This is defined when a child displays negative affect, aggressive behavior, or is non-compliant to prescribed rules. Examples include leaving the room,

climbing on tables, turning out the lights, and refusing to comply with adult requests either verbally or by physical action.

Intensity. This is defined when a child exhibits a high level of physical energy or affect (positive or negative) while interacting with peers or objects during play group. Examples include gross motor movements such as running, jumping, skipping, having tantrums, pounding, slapping and hitting. Intensity is always coded with rough and tumble.

At the end of all playgroup sessions, the coder completes a global rating for each child that is intended to record the child's overall behavior during play. Global ratings are scaled on a Likert scale from 0 to 4. Each child was rated for the following:

Activity Intensity Level. This is the overall level of physical activity or intensity during play. The range is from 0 (not at all active) to 4 (extremely active).

Sociability. This is the overall level of social interaction with peers during play. The range is from 0 (readily interacts or initiates interactions) to 4 (unresponsive to social overtures).

Aggression. This is the overall level of antisocial or inappropriate interaction with peers or teacher during play. The range is from 0 (interactions are prosocial) to 4 (most interactions are antisocial towards peers) (Handen, 1992).

4. CHAPTER FOUR

4.1. ANALYSIS

4.1.1. Assessing Normality

Normality was assessed using Quantile-Quantile (Q-Q) Plots. This is a plot of the actual versus the predicted values, which should be close to a straight line. If the points roughly follow the straight line, the data were considered normal. If the data were not approximately normal, then data transformations were performed to normalize the data, specifically arcsine square root transformations for the percentile data (Social/Play Behaviors) and square root transformations for the numeric data (Global Ratings).

4.1.2. Repeated-Measures ANOVA Analysis

The aim of the study was to examine the changes in social and play behavior with the use of MPH among those children with ADHD and MR. This study was first analyzed using Univariate repeated-measures ANOVA in SPSS 11.5. Specifically, the general linear model, repeated-measures command was used. Any subjects with missing data were deleted from the analysis. Each of the nine dependent variables (Withdrawn, Solitary, Interactive, Rough and Tumble, Rule Violation, Intensity, Activity Intensity Level, Sociability and Aggression) were analyzed separately to determine the effect of MPH on the variables. The independent variable or within-subjects factor was drug condition (placebo, .3 mg/kg, .6 mg/kg dose) and the between-subjects variable was drug-placebo order (1-6 as described previously). The main effects were compared for the drug

condition as well as drug-placebo order using the Bonferroni alpha correction (alpha divided by the number of groups). Because MPH is cleared from one's system in four hours, there were no carryover effects from week to week as in many crossover designs.

The repeated-measures ANOVA has a sphericity assumption (measure of homogeneity of variances of differences between levels), and the test is known as Mauchly's W. When the test is significant, the critical value needs to be corrected. For this analysis, when Mauchly's W was > 0.9 , the sphericity assumption was satisfied and no correction was necessary. When Mauchly's W was between 0.7 and 0.9, the sphericity assumption was not satisfied and the Geisser-Greenhouse correction was used. This is an adjustment to the Univariate test-statistic degrees of freedom and the correction factor can be found by multiplying the error degrees of freedom and the degrees of freedom for any given effect. The new degrees of freedom are decreased by this correction factor, which results in a p-value that is larger than the uncorrected p-value. When Mauchly's W was < 0.7 , the sphericity assumption was not satisfied and was violated so severely that the correction was not possible and the multivariate ANOVA test was used.

To perform this analysis, the data needed to be in wide format, where each row provided all of a single subject's data. The row tells SPSS to which subject the data corresponds, and the column tells SPSS to which condition each score belongs. The dataset provided was already in SPSS in wide format, so no data conversion needed to be done.

The study was analyzed a second time using repeated-measures ANOVA analysis, using the LOCF technique to impute missing data. For each subject, missing values were replaced by the last observed value of that variable.

4.1.3. Mixed Model Analysis

The study was analyzed a third time using Mixed Models in SPSS 11.5. Specifically, the mixed models linear command was used. Again, each of the nine dependent variables above were analyzed separately to determine the effect of MPH on the dependent variables. The random effect was the subject because the sample is a random sample of the target population. The fixed effects were the drug condition (placebo, .3 mg/kg, .6 mg/kg dose) and the drug-placebo order (1-6 as described previously). The repeated covariance type was “Compound Symmetry.” This is the simplest covariance structure and assumes that the covariances between all time points are constant. This was used because if the objective is obtaining a reliable treatment estimate and standard error, a compound symmetry pattern is likely to be robust. Estimates of overall treatment effects will differ little between models using different using covariance patterns in datasets with few repeated measurements (Brown and Prescott, 2003). The main effects were compared for the fixed effects as well as drug-placebo order using the Bonferroni alpha correction (alpha divided by the number of groups).

For the Mixed Model analysis, the data needed to be converted to long format, where there was one score per observation (yielding 3 observations per subject for each drug condition). The data conversion was done using SPSS using the restructure command.

5. CHAPTER FIVE

5.1. RESULTS

5.1.1. Descriptive Analysis

Table 1 displays the demographic characteristics of the 23 children with ADHD and MR. The mean age was 8.6 years, with the range from 6.4 to 13.3 years of age. The mean IQ was 64.1, with the range from 44 to 77. Twelve of the children with ADHD were males (52.2%), while eleven were female (47.8%). Sixteen of the children were Caucasian (69.6%).

Table 1. ADHD/MR Subject Demographic Information

Variable	Mean	Standard Deviation	Range
Age (months)	102.8	22.9	77-160
IQ	64.1	9.4	44-77

Variable	Frequency	Percentage
Sex		
Males	12	52.2
Females	11	47.8
Race		
Caucasian	16	69.6
Other	7	30.4
Total ADHD Subjects	23	100.0

Table 2 shows the means and standard deviations for each of the nine outcome variables for each of the drug conditions (placebo, .3 mg/kg, and .6 mg/kg dose). The means and standard deviations for the social/play behaviors are percentile data, meaning that a mean of 0.28 is 0.28% and a standard deviation of 0.8 is 0.8%. The means and standard deviations for the Global Ratings data are regular numerical data. The means and standard deviations for the variables Withdrawn, Rough and Tumble, and Rule Violation are much lower than the other social/play behaviors because they had mostly zero values.

Table 2. Means and Standard Deviations for Drug Conditions

	Mean (%)	Standard Deviation (%)
Social/Play		
Withdrawn Placebo	0.28	0.8
Withdrawn .3 mg/kg	2.40	3.9
Withdrawn .6 mg/kg	3.67	13.1
Solitary Placebo	44.59	25.3
Solitary .3 mg/kg	46.02	29.3
Solitary .6 mg/kg	48.64	26.5
Interactive Placebo	45.13	21.9
Interactive .3 mg/kg	43.95	29.3
Interactive .6 mg/kg	42.93	25.0
Rough & Tumble Placebo	4.91	9.4
Rough & Tumble .3 mg/kg	4.29	1.4
Rough & Tumble .6 mg/kg	1.31	5.5
Rule Violation Placebo	5.10	10.8
Rule Violation .3 mg/kg	3.60	6.8
Rule Violation .6 mg/kg	3.46	5.5
Intensity Placebo	34.48	29.5
Intensity .3 mg/kg	27.72	30.4
Intensity .6 mg/kg	22.04	24.1
Global Ratings	Mean	SD
Activity Intensity Placebo	2.52	0.9
Activity Intensity .3 mg/kg	2.09	1.0
Activity Intensity .6 mg/kg	1.96	1.0
Sociability Placebo	3.44	0.8
Sociability .3 mg/kg	3.21	0.8
Sociability .6 mg/kg	3.24	0.8
Aggression Placebo	1.06	1.0
Aggression .3 mg/kg	1.07	1.0
Aggression .6 mg/kg	1.00	0.7

5.1.2. Missing Data

Some subjects had adverse side effects, resulting in a need to discontinue a particular dose of MPH, leading to missing values in the dataset. Each of the nine dependent variables described above were assessed with placebo, .3 mg/kg of MPH and .6 mg/kg of MPH. This means that each subject should have twenty-seven data observations. Three of the twenty-three subjects are missing at least one observation. Subject 80 had eighteen missing data points. Subjects 98 and 108 had nine missing data points.

5.1.3. Normality Assumption

The Q-Q plots for the original data did not show normality. Therefore, the percentile data variables of Withdrawn, Solitary, Interactive, Rough and Tumble, Rule Violation and Intensity were transformed using an arcsine-square root transformation and the numeric global ratings of Activity Intensity Level, Sociability and Aggression were transformed using a square root transformation. Figure 1 shows the Q-Q plots for the percentile transformed variables and Figure 2 shows the Q-Q plots for the numeric transformed variables. The plots show that the data were approximately normal because they all roughly follow the straight line. The variables of Withdrawn, Rough and Tumble and Rule Violation had mostly zero values, which is why those transformed values still did not follow the straight line closely.

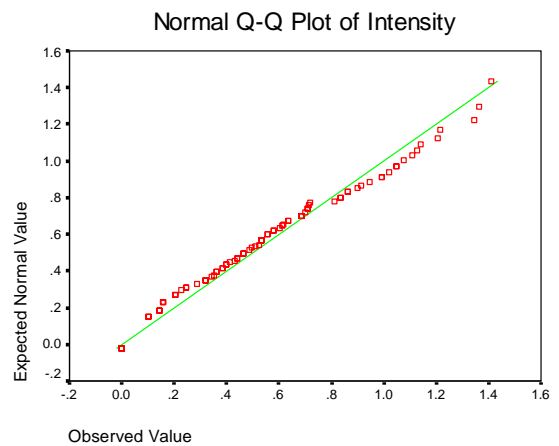
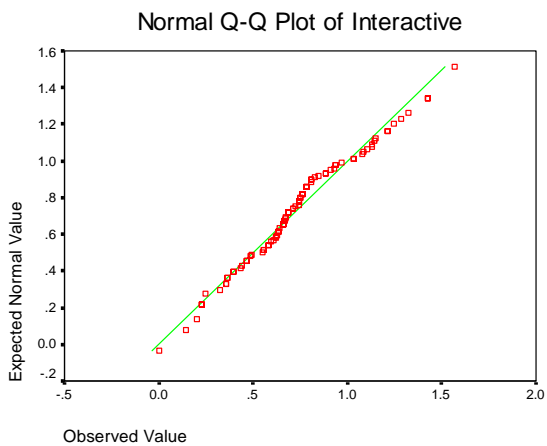
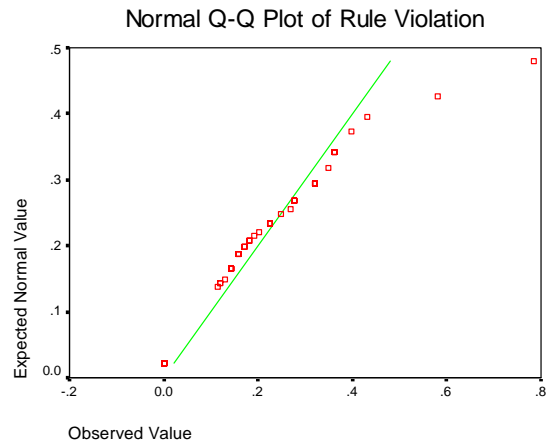
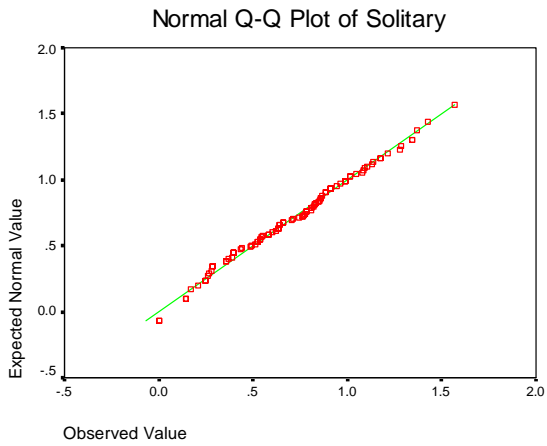
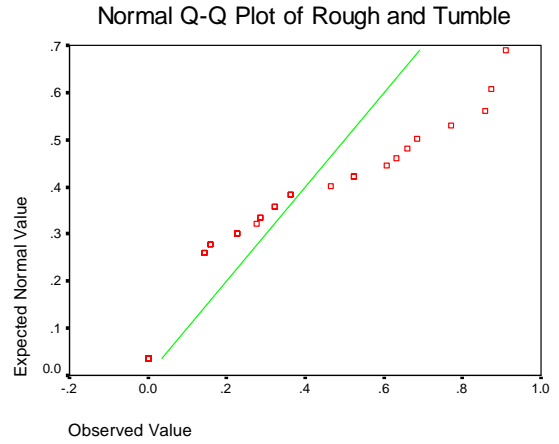
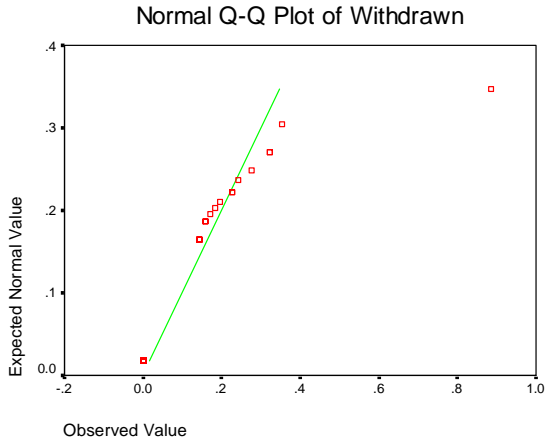


Figure 1. Q-Q Plots of Transformed Percentile Variables.

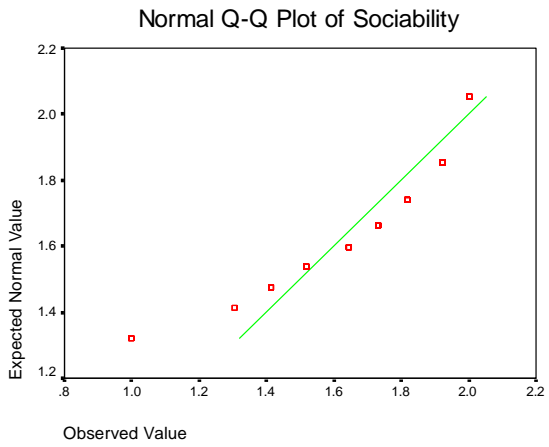
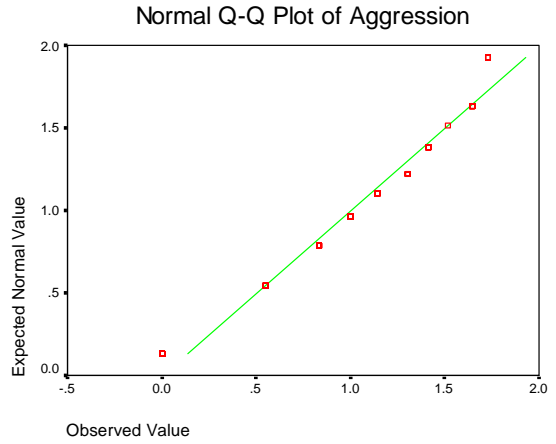
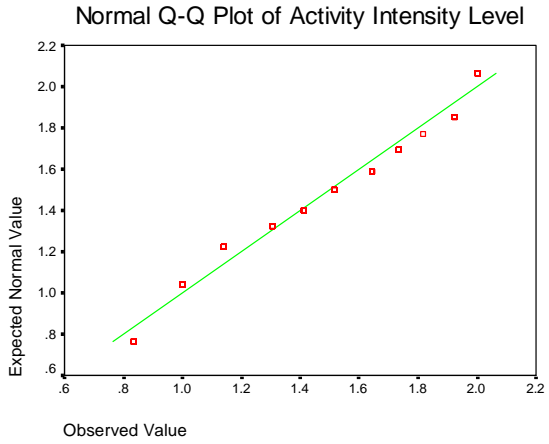


Figure 2. Q-Q Plots of Transformed Numeric Variables.

5.1.4. Model Results

Table 3 shows the p-values of the three analyses for each of the nine outcome variables. P-values are considered statistically significant when they are less than 0.05. Drug-placebo order was taken out of the model when it was not significant at the 0.05 level.

The repeated-measures ANOVA and repeated-measures ANOVA with LOCF analyses found MPH to have a significant effect on the social/play behavior of Withdrawn. All three analyses found MPH to have a significant effect on the social/play behavior of Intensity. The repeated-measures ANOVA with LOCF and Mixed Model analyses found the global rating of Activity Intensity Level to be significant and the repeated-measures ANOVA with LOCF analysis found the global measure of Sociability significant.

In the repeated-measures ANOVA analysis the variable Withdrawn had a significant difference between the placebo and .3 mg/kg of MPH. A significant difference was found between placebo and .6 mg/kg of MPH for Intensity in all three analyses. For Activity Intensity Level, the repeated-measures ANOVA with LOCF and mixed model analyses found a significant difference between placebo and .6 mg/kg of MPH and the repeated-measures ANOVA with LOCF found that drug-placebo order 2 (placebo, .6 mg/kg, .3 mg/kg) was significantly different from drug-placebo order 3 (.3 mg/kg, placebo, .6 mg/kg). Lastly, for Sociability, the repeated-measures ANOVA with LOCF analysis found a significant difference between placebo and .6 mg/kg of MPH.

Table 3. P-values for each model.

Outcome	Repeated-Measures ANOVA	Repeated-Measures ANOVA with LOCF	Mixed Models
Social/Play			
Withdrawn	0.013 ²	0.028	0.212
Solitary	0.614	0.790	0.753
Interactive	0.712	0.947	0.896
Rough & Tumble	0.107	0.112	0.100
Rule Violation	0.667	0.619	0.692
Intensity	0.001 ³	0.001 ³	0.001 ³
Global Ratings			
Activity Intensity Level	0.065 ¹	0.039 ^{1,3,4}	0.017 ^{1,3}
Sociability	0.057 ¹	0.027 ^{1,3}	0.565 ¹
Aggression	0.696	0.637	0.752

¹ In the models where drug-placebo order was significant and left in the model, the p-values are in bold.

² A significant difference was found between placebo and .3 mg/kg of MPH.

³ A significant difference was found between placebo and .6 mg/kg of MPH.

⁴ Drug-placebo order 2 (placebo, .6 mg/kg, .3 mg/kg) was significantly different from drug-placebo order 3 (.3 mg/kg, placebo, .6 mg/kg).

5.1.5. Parameter Estimate and Confidence Interval Results

Figure 3 shows the parameter estimates and corresponding confidence intervals for the outcome variable Withdrawn. For the placebo and .3 mg/kg drug conditions, the repeated-measures ANOVA and repeated-measures ANOVA with LOCF models gave almost identical results. For the .6 mg/kg drug condition, all three models gave almost the same parameter estimates and confidence intervals.

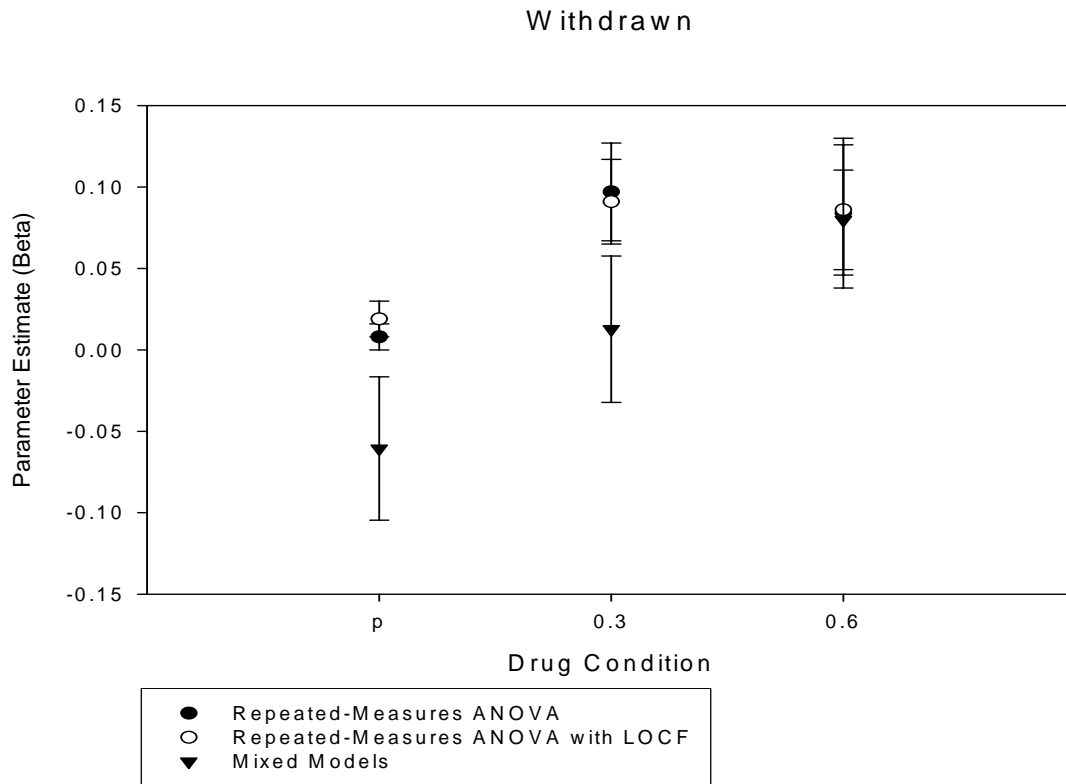


Figure 3. Withdrawn Parameter Estimates.

Figure 4 shows the parameter estimates and corresponding confidence intervals for the outcome variable Solitary. For the placebo and .3 mg/kg drug conditions, the repeated-measures ANOVA and repeated-measures ANOVA with LOCF models gave almost identical results. For the .6 mg/kg drug condition, all three models gave almost identical parameter estimates and confidence intervals.

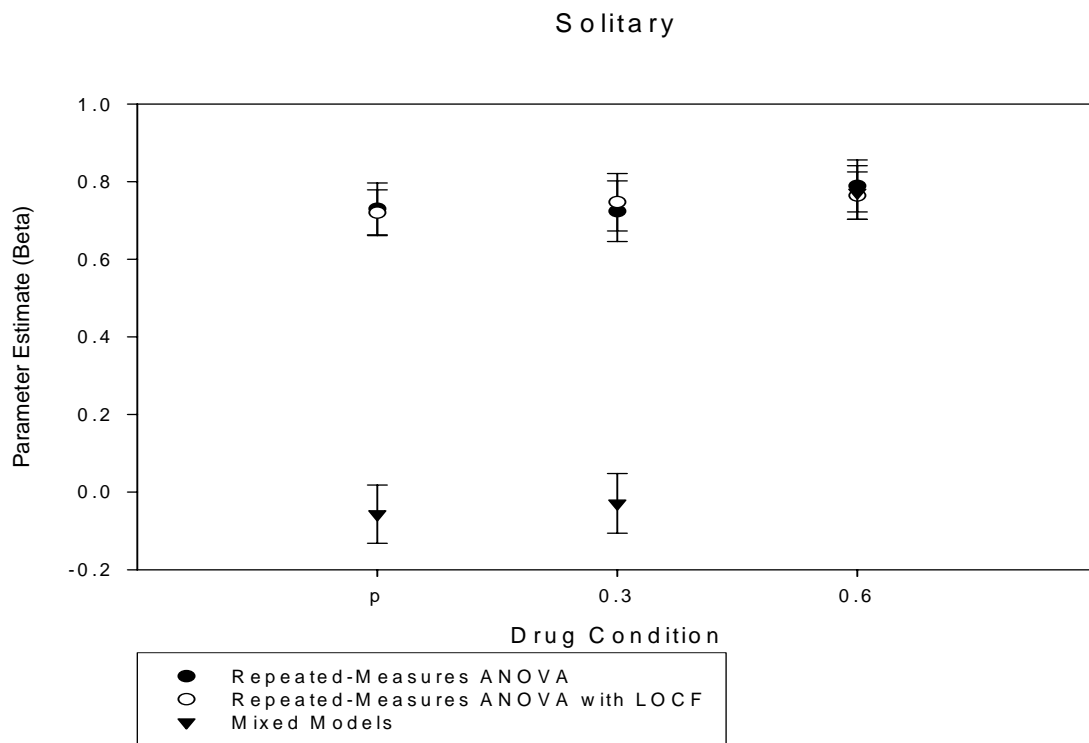


Figure 4. Solitary Parameter Estimates.

Figure 5 shows the parameter estimates and corresponding confidence intervals for the outcome variable Interactive. For the placebo and .3 mg/kg drug conditions, the repeated-measures ANOVA and repeated-measures ANOVA with LOCF models gave similar results. For the .6 mg/kg drug condition, all three models gave almost the exact same parameter estimates and confidence intervals.

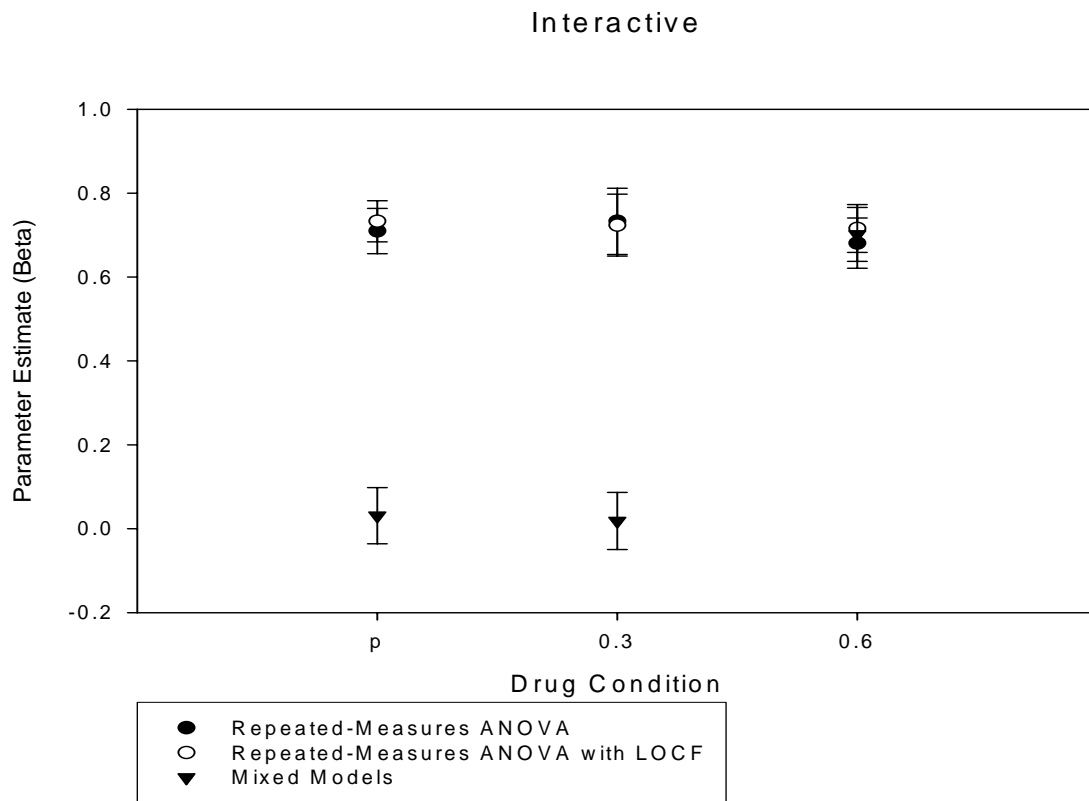


Figure 5. Interactive Parameter Estimates.

Figure 6 shows the parameter estimates and corresponding confidence intervals for the outcome variable Rough and Tumble. For the placebo and .3 mg/kg drug conditions, the repeated-measures ANOVA and repeated-measures ANOVA with LOCF models gave similar results. For the .6 mg/kg drug condition, all three models gave almost the same parameter estimates and confidence intervals.

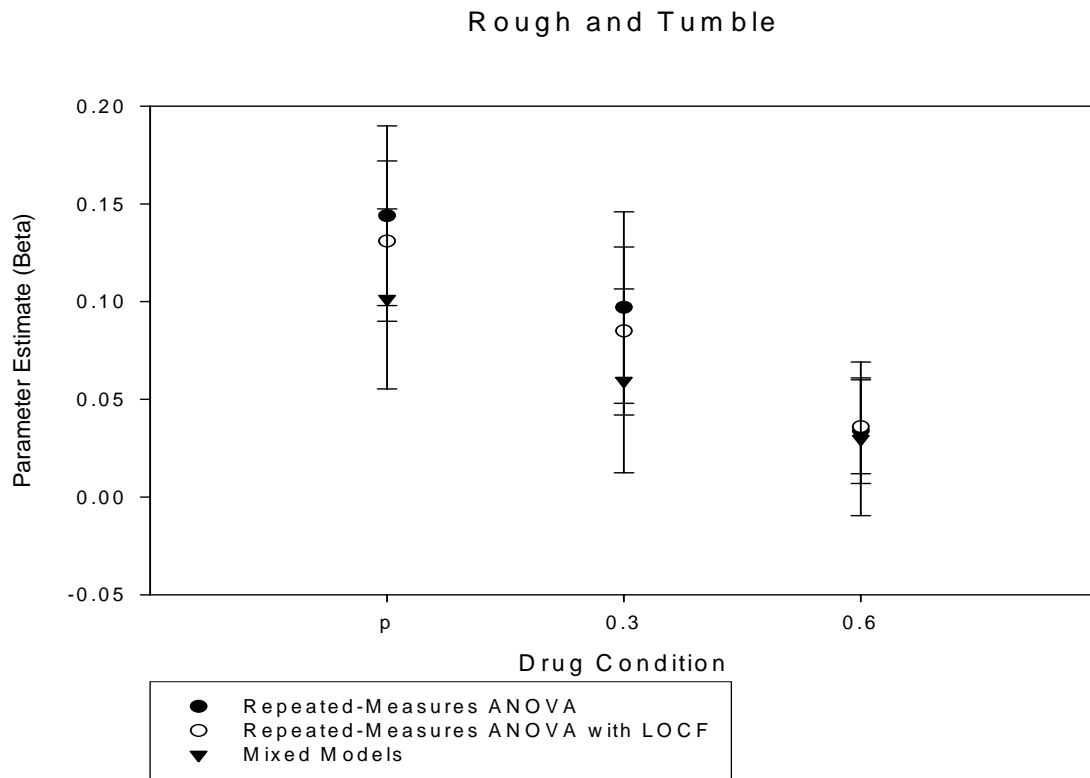


Figure 6. Rough and Tumble Parameter Estimates.

Figure 7 shows the parameter estimates and corresponding confidence intervals for the outcome variable Rule Violation. For the placebo and .3 mg/kg drug conditions, the repeated-measures ANOVA and repeated-measures ANOVA with LOCF models gave similar results. For the .6 mg/kg drug condition, all three models gave similar parameter estimates and confidence intervals.

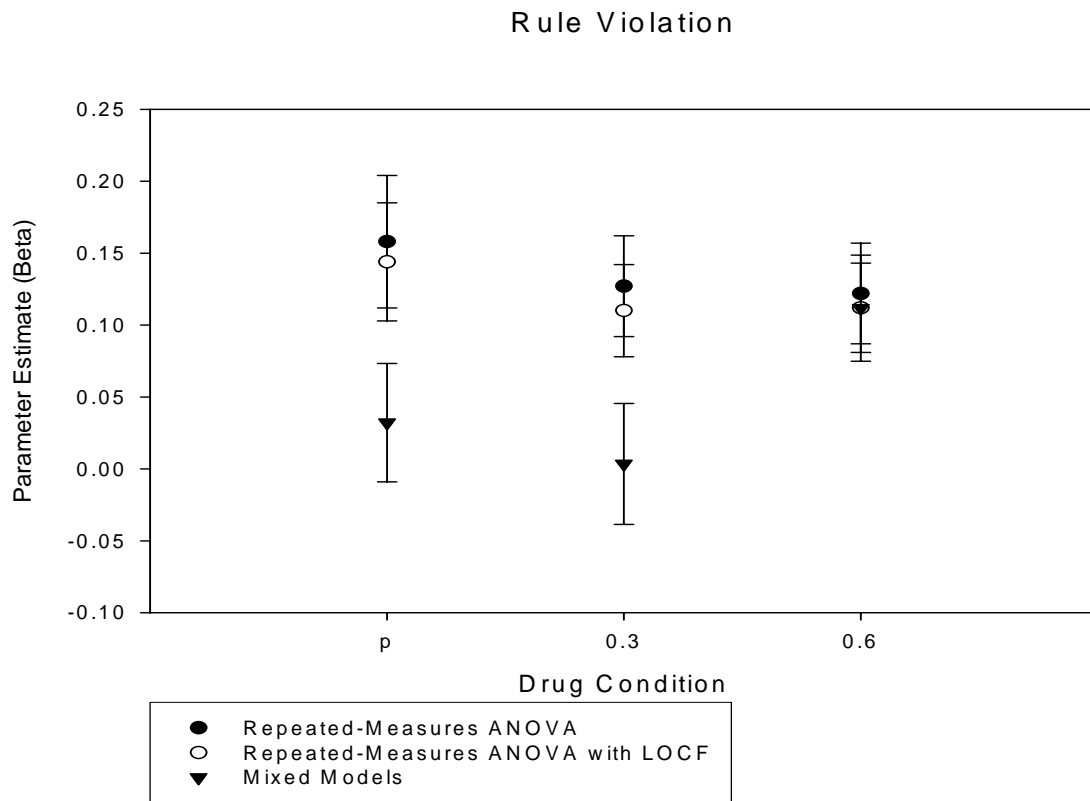


Figure 7. Rule Violation Parameter Estimates.

Twisk, J. (2003) *Applied Longitudinal Data Analysis for Epidemiology*, Cambridge University Press, Cambridge.