Paper 1798-2014

Comparison of Five Analytic Techniques for Two-Group, Pre-Post Repeated Measures Designs Using SAS ®

J. Madison Hyer, Georgia Regents University & Jennifer L. Waller, Georgia Regents University

ABSTRACT

There has been debate regarding which method to use to analyze repeated measures continuous data when the design includes only two measurement times. Five different techniques can be applied and give similar results when there is little to no correlation between pre- and post-test measurements and when data at each time point are complete: 1) analysis of variance on the difference between pre- and post-test, 2) analysis of covariance on the differences between pre- and post-test controlling for pre-test, 3) analysis of covariance on post-test controlling for pre-test, 4) multiple analysis of variance on post-test and pre-test, and 5) repeated measures analysis of variance. However, when there is missing data or if a moderate to high correlation between pre- and post-test measures exists under an intent-to-treat analysis framework, bias is introduced in the tests for the ANOVA, ANCOVA, and MANOVA techniques. A comparison of Type III sum of squares (where estimated), F-tests and p-values for a complete case and an intent-to-treat case are presented. The analysis using a complete case data set shows that all five methods produce similar results except for the repeated measures ANOVA due to a moderate correlation between pre- and post-test measures. However, significant bias is introduced for the tests when using the intent-to-treat data set.

INTRODUCTION

Many clinical trial designs use a pre-post repeated-measures, two-group design to assess the effect of a treatment over time compared to a control condition on a continuous variable. To analyze the data and determine the effect of the intervention, several analytic techniques exist, but the question becomes which method to use in the presence of a moderate to high correlation of the outcome measure between the two time points or in the presence of missing data at post-test. For non-statisticians, performing an analysis of variance (ANOVA), an analysis of covariance (ANCOVA) on the difference between pre- and post-test measures, or a multiple ANOVA (MANOVA) on both pre- and post-test is easier than performing a repeated measures mixed model. However, if a moderate to high correlation exists between the continuous measures at the two measurement times, the results of the ANOVA, ANCOVA and MANOVA are biased as the correlation between time points is not taken into account in the estimation. Similarly, if missing data exists at post-test the ANOVA, ANCOVA and MANOVA methods do not utilize all subjects with pre-test data, rather the analysis is done on subjects with complete data rather than an analysis on all available data. To overcome these biases, a repeated measures ANOVA using PROC MIXED can be performed to model the correlation structure between the measurement times and utilize all available data.

METHODS

STUDY DESIGN

This study examined school attendance in children between the ages of 8 to 11, who were overweight or obese (BMI $\geq 85^{th}$ percentile), healthy enough to participate in an exercise intervention, any race or ethnicity, and recruited from a public school system in the southeastern US. The study design was a randomized control clinical trial with two groups, exercise and control, from 4 cohorts measured at three time points, baseline, post-initial, and one-year follow-up. The study enrolled 175 study participants of which, 124 contained data for baseline school attendance.

Children randomized to the exercise intervention spent 40 minutes per day in two 20-miniute sessions in activities eliciting vigorous exercise. The activities focused on intensity, enjoyment, and safety of the children and consisted of different games and activities that were easy to comprehend, not competitive, and brought about sporadic intervals of vigorous movement. Children allocated to the control group spent 60 minutes participating in instructor-led sedentary activities, including board games, watching videos, or similar sedentary activities.

OUTCOME OF INTEREST

The main outcome variable for this analysis was school absences, measured at two time points- baseline and post intervention. The data collected for school absences before the study consists of the student's attendance record for the academic year immediately prior to the academic year of the study. The data collected after the completion of the study reflects the attendance record for the students during the academic year of the study.

MAIN INDEPENDENT VARIABLES

The main independent variables of interest were group (exercise versus control) and measurement time (baseline and post-test).

DATA SETS FOR ANALYSIS

COMPLETE CASE DATA SET

The complete case data set contained all individuals with complete data for all outcome and independent variables at baseline and post-test. Of the 175 study participants, 78 had baseline and post-test data. All other design variables (cohort, race and sex) had complete data on the 175 participants. The complete case data set was only limited by the missing data for school absences.

INTENT TO TREAT DATA SET

The intent-to-treat data set contained all individuals with available data at baseline and posttest. Only participants who had attendance data at baseline were included in the analysis. Of the 175 participants, 124 had school attendance data at baseline.

STATISTICAL ANALYSIS

All statistical analysis was performed using SAS® 9.3. Five different analytical approaches were used to examine similarity or differences in Type-III sums of squares (where estimated), F-tests and statistical significance of the group or group by time effect using the complete case and intent-to-treat data sets: 1) analysis of variance (ANOVA) on the difference between post-test and pre-test school absences using PROC GLM, 2) analysis of covariance (ANCOVA) on the difference between post-test and pre-test school absences controlling for pre-test school absences using PROC GLM, 3) analysis of covariance (ANCOVA) on post-test school absences controlling for pre-test school absences using PROC GLM, 4) multiple analysis of variance (MANOVA) on post-test and pre-test school absences using PROC GLM, and 5) repeated measures analysis of variance on school absences using PROC MIXED. Two different sets of analyses were performed: one set disregarding the design variables in the analysis and one controlling for the design variables in the analysis. When controlling for the design variables, each model contarined variables of cohort, race and sex. Both the complete case and intent-to-treat data sets contained black and white/other race groups, male and female gender, and control or exercise intervention groups. The complete data set contains 3 cohorts and the intent-to-treat data set contains 4 cohorts, as little attendance data has been collected for one cohort at this time.

ANOVA ON DIFFERENCES (POST-TEST - PRE-TEST) USING PROC GLM

For the analysis of ANOVA on the difference between pre- and post-test, the outcome variable was the difference between post-test attendance and pre-test attendance data. The independent effects considered were cohort, race, sex, and group. SAS code used to perform the analysis is

```
      PROC GLM data=x;
        class cohort race sex group;
        model post_pre_absence=cohort race sex group;
        means group;
        lsmeans group;
      title 'ANOVA of Exercise vs Control on';
      title2 'Difference of Post and Pre-test School Absences';
  run;
```

ANCOVA ON DIFFERENCES (POST-TEST - PRE-TEST) CONTROLLING FOR PRE-TEST USING PROC GLM

For this analysis, the outcome variable was the difference between post-test attendance and pre-test attendance. The independent effects were cohort, race, sex, and group controlling for pre-test absence data - the absence data for the academic year prior to the study. The SAS code to perform the ANCOVA analysis is

```
      PROC GLM data=x;
        class cohort race sex group;
        model post_pre=cohort race sex absence_pre group;
        means group;
        lsmeans group;
      title 'ANCOVA of Exercise vs Control';
```

```
    title2 'Difference of Post and Pre-test School Absences';
    title3 'Controlling for Pre-test School Absences';
run;
```

ANCOVA ON POST-TEST CONTROLLING FOR PRE-TEST USING PROC GLM

The outcome variable was post-test absence data, absences for the academic year of the study. The independent effects were cohort, race, sex, and group controlling for pre-test absence data - the absence data for the academic year prior to the study. The SAS code to perform the ANCOVA controlling for pre-test is

```
    PROC GLM data=x;
      class cohort race sex group;
      model absence_post=cohort race sex absence_pre group;
      means group;
      lsmeans group;
    title 'ANCOVA of Exercise vs Control on Post-test School Absences';
    title2 'Controlling for Pre-test School Absences';
run;
```

MANOVA ON POST-TEST AND PRE-TEST USING PROC GLM

For the MANOVA analysis, the outcome variables were absence data for the academic year of the study (post-test) and absence data for the academic year immediately prior to the study (pre-test). The independent effects ere cohort, race, sex, and group. The SAS code to perform the MANOVA is

```
    PROC GLM data=x;
      class cohort race sex group;
      model absence_pre absence_post =cohort race sex group;
      repeated time / print printh;
      means group;
      lsmeans group;
    title 'MANOVA of Exercise vs Control on Post-test  and Pre-test School
Absences';
    title2 'Controlling for Pre-test School Absences';
run;
```

REPEATED MEASURES ANOVA USING PROC MIXED

For repeated measures ANOVA, the outcome variable was absence data with each individual having two rows of data, one row for their data at pre-test and one row for their data at post-test. The design effects were cohort, race, sex, and main effects of group, time and the group by time interaction were used. Because there are only two time points, an unstructured correlation structure is used. The Ken-Ward-Roger option is used to obtain the correct denominator degrees of freedom for the F-tests. The SAS code used to perform the repeated measures ANOVA in PROC MIXED is

```
    PROC MIXED data=x_long;
      class subjectid cohort race sex group time;
      model absence=cohort race sex group|time / ddfm=kenwardroger;
      repeated / subject=subjectid type=un r rcorr;
      lsmeans group|time / pdiff;
    title 'Repeated Measures ANOVA of Group and Time on School Absences';
run;
```

As an additional analysis when controlling for the study design, the race by time, sex by time, and cohort by time interactions were included to compare the MIXED results to the MANOVA results.

RESULTS

INTENT-TO-TREAT PRE-TEST DESCRIPTIVE STATISTICS

Recall, 124 of the 175 study participants contained data for baseline attendance. Briefly, 59 (47.6%) and 65 (52.4%) were allocated to control and exercise, respectively. One hundred and five (84.7%) were black and 19 (15.3%) were white/other. Fifty-one (41.1%) were male and 73 (58.9%) were female. There were no significant differences at pretest between race, gender or age between those allocated to the control or exercise group ($p = 0.1395$ for race, $p = 0.5263$ for gender, and $p = 0.8861$ for age).

COMPLETE CASE PRE-TEST DESCRIPTIVE STATISTICS

Of the 124 children enrolled in the study containing pre-test attendance data, 78 contained data absence at the post-test. Of those 78 with post-test data, 38 (48.7%) were in the control group and 40 (51.3%) were in the case group. Sixty-four (82.1%) were black and 14 (17.9%) were white/other. Thirty-five (44.9%) were male and 43 (53.1%) were female. There were no significant differences at pretest between race, gender or age between those allocated for control or exercise group (p =0.1983 for race, p =0.1793 for gender, and p =0.7471 for age) among those with both pre- and post-test data.

ANALYTIC TECHNIQUE COMPARISON

For the following four tables, acronyms used for the the analyses presented are: analysis of variance on the difference between pre- and post-test (ANOVA diff); analysis of covariance on the differences between pre- and post-test con- trolling for pre-test (ANCOVA diff); analysis of covariance on post-test controlling for pre-test (ANCOVA post); multiple analysis of variance on post-test and pre-test (MANOVA); repeated measures ANOVA using PROC MIXED (MIXED).

ANALYSIS NOT CONTROLLING FOR DESIGN VARIABLES

Tables 1 and 2 show the results of the five analysis techniques regarding Type III sum of squares (where estimated), F-tests and p-values for the complete case and intention to treat analysis, respectively, with group as the independent effect without design variables. Presented for the ANOVA models are the estimates and tests for the group effect, for the ANCOVA model are the estimates and tests for the group effect, and for the MANOVA model are the estimates and tests for the group by effect, and for the repeated measures mixed model are the estimates and tests for the group by time interaction.

	n	Type III SS	df_{num}	df_{dem}	F	p-value
ANOVA $diff$ (group)	78	100.92	1	76	2.18	.1436
ANCOVA $diff$ (group)	78	54.02	1	75	1.90	.1720
ANCOVA $post$ (group)	78	54.02	1	75	1.90	.1720
MANOVA (group×time)	78	100.92	1	76	2.18	.1436
MIXED (group×time)	202	·	1	76	2.18	.1436

Table 1: Results of the five analysis techniques for the complete case analysis with group as the independent effect without design variables

	n	Type III SS	df_{num}	df_{dem}	F	p-value
ANOVA $diff$ (group)	78	100.92	1	76	2.18	.1436
ANCOVA $diff$ (group)	78	54.02	1	75	1.90	.1720
ANCOVA $post$ (group)	78	54.02	1	75	1.90	.1720
MANOVA (group×time)	78	100.92	1	76	2.18	.1436
MIXED (group×time)	202	·	1	96	1.51	.2226

Table 2: Results of the five analysis techniques for the intention to treat case analysis with group as the independent effect without design variables

For the complete case (Table 1), the results of Type III SS, F-tests and p-values are identical for the ANOVA, MANOVA, and MIXED analyses. Similarly, results of Type III SS, F-tests and p-values are identical for ANCOVA on differences and ANCOVA on post-test.

For the intent-to-treat case (Table 2), the results of Type III SS, F-tests and p-values are identical for the ANOVA and MANOVA analyses. Similarly, results of Type III SS, F-tests and p-values are identical for ANCOVA on differences and ANCOVA on post-test. However, the MIXED analysis produced different, less significant, results from the other four analyses.

Comparing the complete case (Table 1) with the intent to treat case (Table 2) for the group effect, the first four analyses produced identical results. However, the fifth, the MIXED analysis, produced different results between the complete and intent-to-treat cases due to inclusion of subjects with missing data at post-test and the moderate-to high correlation between pre- and post-test in the intent to treat case.

ANALYSIS CONTROLLING FOR DESIGN VARIABLES

Tables 3 and 4 show the results of the five analysis techniques regarding Type III sum of squares, F-tests and p-values for the complete case and intent-to-treat analysis, respectively, with group and the study design variables as the independent effects. Presented for the ANOVA model are the estimates and tests for the group effect, for the ANCOVA models are the estimates and tests for the group effect, and for the MANOVA model are the estimates and tests for the group effect, and for the repeated measures mixed model are the estimates and tests for the group by time interaction.

	n	Type III SS	df_{num}	df_{dem}	F	p-value
ANOVA $diff$ (group)	78	78.40	1	72	1.75	.1895
ANCOVA $diff$ (group)	78	74.82	1	71	2.70	.1048
ANCOVA $post$ (group)	78	74.82	1	71	2.70	.1048
MANOVA (group×time)	78	78.40	1	72	1.75	.1895
MIXED (group×time) (cohort×time) (race×time) (sex×time)	202	·	1	72	1.75	.1895
MIXED (group×time) controlled for design	202	·	1	76	2.18	.1436

Table 3: Results of the six analysis techniques for the complete case analysis with group and the study design variables as the independent effects

	n	Type III SS	df_{num}	df_{dem}	F	p-value
ANOVA $diff$ (group)	78	78.40	1	72	1.75	.1895
ANCOVA $diff$ (group)	78	74.82	1	71	2.70	.1048
ANCOVA $post$ (group)	78	74.82	1	71	2.70	.1048
MANOVA (group×time)	78	78.40	1	72	1.75	.1895
MIXED (group×time) (cohort×time) (race×time) (sex×time)	202	·	1	101	1.42	.2365
MIXED (group×time) controlled for design	202	·	1	106	1.76	.1880

Table 4: Results of the six analysis techniques for the intention to treat case analysis with group and the study design variables as the independent effects

For the complete data set (Table 3), the results of Type III SS (where calculated), F-tests and p-values are identical for the ANOVA, MANOVA, and MIXED (controlling for design variables by time interactions) analyses. Similarly, results of Type III SS, F-tests and p-values are identical for ANCOVA on differences and ANCOVA on post-test. The last analysis, MIXED (controlled for design), produced different results - less significant than the ANCOVA analyses but more significant than ANOVA and MANOVA analyses as well as the MIXED with design variables by time interactions.

For the intent-to-treat data set (Table 4), the results of Type III SS (where calculated), F-tests and p-values are identical for the ANOVA and MANOVA analyses. Similarly, results of Type III SS (where calculated), F-tests and p-values are identical for ANCOVA on differences and ANCOVA on post-test. However, with the intent-to-treat data set, MIXED with design variables by time interactions and MIXED controlled for design analyses produced different results. The MIXED controlled for design analysis produced results less significant than the ANCOVA analyses but slightly more significant than ANOVA and MANOVA analyses as well as the MIXED with design variables by time interactions analysis. The MIXED with design variables by time interactions produced the least significant results.

Comparing the complete case data set (Table 3) with the intent-to-treat data set (Table 4), with group as the independent effect with design variables, the first four analyses produced identical results. Additionally, while the first four are identical, the MIXED analyses produced different results between the complete and intent-to-treat cases. The results for both MIXED analyses are less significant for the intent-to-treat case.

CONCLUSIONS

Examination of the results indicate that using a complete data set when there is a moderate to large correlation between outcome measurements introduces a bias towards statistical significance. Additionally, when there exists missing data, the intent-to-treat analysis using PROC MIXED, produced tests that were less biased toward

significance than other analysis techniques- due to making use of all available data and incorporating the correlation between time points.

Thus, in the presence of a moderate to high correlation between outcome measurements and/or the presence of missing data, using all available data and utilizing the estimated correlation between measurement times produced less biased towards statistical significance results. It is recommended that a MIXED model analysis be used in this situation.

REFERENCES

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Name: J. Madison Hyer
Organization: Georgia Regents University
Address: Department of Biostatistics and Epidemiology
City, State ZIP: Augusta, GA 30912
Email: madison.hyer@gmail.com

Name: Jennifer L. Waller
Organization: Georgia Regents University
Address: Department of Biostatistics and Epidemiology, AE-1012
City, State ZIP: Augusta, GA 30912
Email: jwaller@gru.edu