

Mixed Effects Models

Yan Wang, Bristol-Myers Squibb, Wallingford, CT

ABSTRACT

The MIXED procedure has been commonly used at the Bristol-Myers Squibb Company for quality of life and pharmacokinetic/pharmacodynamics modeling. This paper provides two examples to explore the use of the SAS PROC MIXED procedure. One example is a phase 3 neuroscience study, where we use this example to demonstrate the longitudinal data analysis. The other example is a phase 2, PK, HIV, cross-over study. The paper describes the programs that have been used to carry out these analyses, and the interpretation of the outputs.

INTRODUCTION

Why do we use the PROC MIXED procedure?

Before we deep dive in to the examples, let's first try to understand why we use the MIXED procedure. As we know, SAS supports several statistical methods for the analysis of longitudinal or repeated data, such as PROC GLM. What's the strength of using the MIXED procedure?

The MIXED procedure is flexible. The name of mixed means that the model can contain both fixed effect parameters and random effect parameters. And GLM procedure can only handle fixed effects. PROC MIXED uses all the available data in the analysis, which means no observation is dropped from the analysis, even when some of their values at certain visits or time points are missing. PROC MIXED offers a wide variety of covariance structures, which enables us to directly address the within-subject correlation structure and incorporate it into a statistical model, especially for the analysis that includes between groups effects as well as within subject effects. In terms of syntax, PROC MIXED has the same grammar style as PROC GLM. Both have the similar statements, such as MODEL & LSMEANS, etc.

CASE STUDY 1: LONGITUDINAL DATA ANALYSIS

The first case study is a phase 3 neuroscience trial that is designed as a randomized, double-blind, and placebo-controlled study of the safety & efficacy of Aripiprazole in the treatment of patients with major depressive disorder. The research hypothesis or the interest of the study is to test whether aripiprazole is significantly better than placebo in reducing depressive symptoms in patients with major depressive disorder, measured by Montgomery Asberg Depression Rating Scale (MADRS). The MADRS is a 10-item psychological questionnaire. It is commonly used by clinicians to differentiate moderate and severe depression, and it is also often used as the efficacy assessment of a patient's depression level. In this practical study, the data are collected over time at a weekly basis. This study includes two treatment groups: Aripiprazole (Ari) and Placebo. The primary efficacy outcome measure is the change from the end of phase B, which is the baseline, to the end of Phase C, which is the end treatment period in the MADRS total score.

In the statistical analysis plan (SAP), the endpoint is stated as "The Change from end of phase B in MADRS individual item scores to every study week in phase C." The SAP also details what SAS procedure and options were used. MIXED is the model of choice for this analysis, and uses the restricted maximum likelihood (REML) method for estimation. REML is the default method in the MIXED procedure, so there is no need to specify this explicitly in the code. The time variable is included in both CLASS and MODEL statements.

```
CODE: PROC MIXED DATA = Bms_eff ;
        CLASS Treat Pid Week ;
        MODEL Change = Base_tot Base_tot*Week Treat Week Treat*Week
                / SOLUTION OUTF = Outp DDFM = KENWARDROGER ;
        REPEATED Week / TYPE = UN SUBJECT = Pid ;
        LSMEANS Treat*Week / AT MEANS CL DIFF SLICE = Week ;
                                MAKE "LSMEANS" OUT = Gtmeans ;
                                MAKE "DIFFS" OUT = Diffs ;
```

The CLASS statement is the same as in PROC GLM. The model considers the Treatment group, Subject ID, and Week as categorical variables. Those variables can be either character or numeric.

The MODEL statement first specifies the response (dependent) variable, then lists the independent variables after the equal (=) sign. The variable “Change” is the dependent variable, whose value is the differences between the total MADRS score per week during treatment minus the baseline total MADRS score by subject. The procedure allows only for one dependent variable in the model. In this study, the independent variables are “baseline total score”, “treatment”, and “week”. They cover majority of the main effects of the design. The model also contains 2 interaction terms: “baseline total score” by “week” and “treatment” by “week”, which model the interactions between the main effects. By specifying the SOLUTION option on the model statement, we request t-tests and standard errors for each fixed effect. The OUTP option creates an output dataset named “Outp” as well. The DDFM (degrees of freedom) calculation is based on Kenward & Roger method.

The REPEATED statement specifies the repeated measures variable “week”.

The option TYPE specifies the covariance structures – UN means unstructured covariance.

The SUBJECT option identifies the subjects in the mixed model.

The LSMEANS statement computes least squares means for the interaction effect between the treatment and week. The AT option sets the effect equal to the product of the individual means rather than the mean of the product (as with standard LS-means calculations). The AT MEANS option leaves covariates equal to their mean values (as with standard LS-means) and incorporates this adjustment to cross products of covariates. The CL option which produces the confidence limits. The DIFF option computes differences of the least squares means. The SLICE option considers only the differences of LSEAMS within the same week by different treatment group.

Due to the fact that the LSMEANS statement doesn’t have an OUT option, so the MAKE statement converts any table produced by the PROC MIXED into a SAS data set for further process.

LOG:

Ensure “NOTE: Convergence Criteria Met” in the log file, otherwise re-examine the model.

OUTPUT:

The program generate a lot of tables, we shall focus on the key tables for the purpose of showing the main analysis results.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
BASE_TOT	1	348	6.90	0.0090
BASE_TOT*WEEK	5	314	0.92	0.4670
TREAT	1	344	27.87	<.0001
WEEK	5	312	1.09	0.3633
TREAT*WEEK	5	314	6.10	<.0001

The “Type 3 Tests of Fixed Effects” table contains the hypothesis tests for the significance of each of the fixed effects. The TYPE3 is the default test, which enables the procedure to produce the exact F tests. (Please note that the F- and p-values are identical to those from PROC GLM.)

The table shows the tests for a significant regression relationship between the response variable and the predictor variables. By checking the P value at the 0.05 level, the baseline total score, the treatment, and the treatment by week interaction are significant, and they all have a statistically significant association with the response variable. A useful way to explain the significant interactions is to graphically display them. How do we graph the interaction between the Treat & the Week? We use the data set created from the LSMEANS statement for the graph.

Least Squares Means						
Effect	TREAT	Study Week	Estimate	Standard Error	Lower	Upper
TREAT*WEEK	Ari	9	-3.0922	0.3792	-3.8383	-2.3462
TREAT*WEEK	Ari	10	-6.4347	0.4704	-7.3600	-5.5095
TREAT*WEEK	Ari	11	-8.2243	0.5319	-9.2704	-7.1783
...						
Partial Output						

The “Least Squares Means” table is important. The graph is constructed from the estimated least squares means and their confidence intervals for the interaction effect between the treatment and the week from this table. They estimate the marginal means over a population.

Tests of Effect Slices						
Effect	Study Week	BASE_TOT	Num DF	Den DF	F Value	Pr > F
TREAT*WEEK	9	25.84	1	344	1.95	0.1636
TREAT*WEEK	10	25.84	1	341	19.67	<.0001
TREAT*WEEK	11	25.84	1	342	20.73	<.0001
TREAT*WEEK	12	25.84	1	337	31.29	<.0001
TREAT*WEEK	13	25.84	1	332	28.36	<.0001
TREAT*WEEK	14	25.84	1	334	12.68	0.0004

The “Tests of Effect Slices” table displays the significance tests on the differences of the least squares means for the two treatments at each week. For example, at week 9, there is no significant treatment difference between placebo and Ari. However, for the rest of the weeks, they are significantly different. These significant test results for the interaction effects are consistent with the graph, which shown below.

FIGURE:

Below is the figure produced from the data set “Gtmeans”, which is the result from the LSMEANS statement.

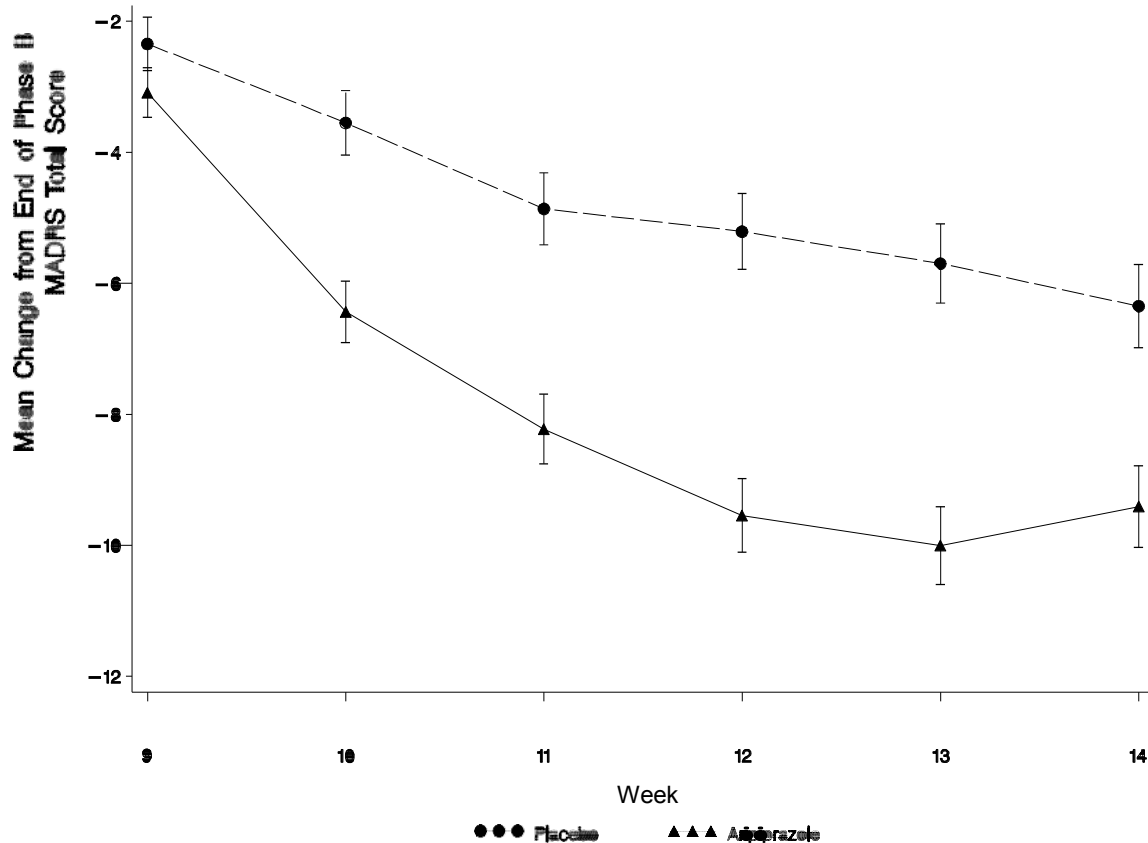


Figure 1. Average Profiles of Mean Change from End of Phase B to Phase C in MADRS Total Score, OC Data Set, Efficacy Sample

In this figure, the top dashed line indicates the placebo group, and the bottom solid line displays the ARI group. The x axis shows the time points by week, the y axis is the response/dependent variable which means “change” in total score from baseline to end of treatment.

Clearly, it shows that the two treatment groups start off at the same level of depression. The lines for the two treatment groups start to separate from week 9 and the difference continues to grow. This is consistent with the between treatment test, which indicates that the treatment variable is significant.

In the graph we see that the two groups have non-parallel lines that decrease over time, which indicates that the interaction term of “week” and “treat” is significant.

ANALYSIS FINDING:

The interpretation from the table “Type 3 Tests of Fixed Effects” concludes:

- There is a statistical difference for mean of change from baseline to the end of treatment score between two treatment groups (P<0.0001).
- There is a statistically significant interaction between treatment groups and week (P<0.0001).
- The result shows a strong evidence that the drug is significantly better than placebo over time.

CASE STUDY 2: CROSSOVER DATA ANALYSIS

This is an open-label, randomized, 3-period, 3-treatment, crossover study. Table 1 below shows what the crossover data look like for this study: at each period, subjects are randomized to one of the three treatments. The variable SEQUENCE captures the sequence order of the randomized drugs per subject. The variable CMAX is the log-transformed as a dependent variable in the model example.

USUBJID	PERIOD	TRT	SEQUENCE	CMAX
AI438010-1-1	1	B	BAC	3200
AI438010-1-1	2	A	BAC	4230
AI438010-1-1	3	C	BAC	5310
AI438010-1-10	1	A	ACB	2880
AI438010-1-10	2	C	ACB	3340
AI438010-1-10	3	B	ACB	2360

Table 1. Data layout

CODE:

```
PROC MIXED DATA = Indat ;
    CLASS Usubjid Sequence Period Trt ;
    MODEL LCMAX = Sequence Period Trt / SOLUTION DDFM = KR ;
    REPEATED TRT / SUBJECT = Usubjid TYPE = UN ;
    ESTIMATE "B VS A" Trt -1 1 0 / CL ALPHA = 0.10 ;
    ESTIMATE "C VS A" Trt -1 0 1 / CL ALPHA = 0.10 ;
    ESTIMATE "C VS B" Trt 0 -1 1 / CL ALPHA = 0.10 ;
    LSMEANS Trt / CL ALPHA = 0.10 ;
    ODS OUTPUT TESTS3 = Tests3 ESTIMATES = Estimates LSMEANS = Lsmeans ;
```

In this model, the dependent variable, LCMAX, is the logarithm of Cmax. Variables including sequence, period, and trt could potentially impact the value of Cmax. The ESTIMATE statement provides a mechanism for obtaining custom hypothesis tests among LS-means. For examples, in the above code, the first estimate statement is to test whether the mean of drug B is the same as that of drug A. We set the significance level at .1, which gives the two-sided 90% confidence interval.

OUTPUT:

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
SEQUENCE	5	12	0.83	0.5513
PERIOD	2	29.4	5.38	0.0102
TRT	2	15.2	34.28	<.0001

Table 2. Type 3 Tests of Fixed Effects

The F test from table 2 shows that the sequence order of taking the drugs is not significant, which means that the variable SEQUENCE is not a main effect. The variables period and trt are significant at .1 significance level. Since treatment is significant, then we can explore the pairwise comparison of the treatments.

Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
B VS A	-0.2190	0.04832	16.6	-4.53	0.0003	0.1	-0.3031	-0.1348
C VS A	0.1311	0.03824	15.8	3.43	0.0035	0.1	0.06431	0.1979
C VS B	0.3501	0.04139	16.1	8.46	<.0001	0.1	0.2778	0.4223

Table 3. Output from the ESTIMATE statement

Table 3 shows all the pairwise comparisons among the 3 TRT groups. The P values indicate that the pairwise differences in means of the treatments are highly significantly different from each other at the .01 significance level.

Least Squares Means									
Effect	Treatment	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
TRT	A	8.0936	0.07374	11.9	109.76	<.0001	0.1	7.9621	8.2251
TRT	B	7.8747	0.07995	11.8	98.50	<.0001	0.1	7.7320	8.0174
TRT	C	8.2247	0.07618	12.3	107.96	<.0001	0.1	8.0892	8.3602

Table 4. Output from the LSMEANS statement

It is straight forward to interpret the above results, that is, LCMAX is statistically significant different between the different treatments

CONCLUSION

- MIXED model extends the GLM by the addition of random effect parameters.
- MIXED model can easily be fitted to longitudinal data.
- PROC MIXED implements maximum likelihood to estimate the covariance parameters.
- Kenward-Roger degrees of freedom adjustment is used as the standard operating procedure for longitudinal models.

ACKNOWLEDGMENTS

Thanks to Richard Perry for the sponsorship. Thanks to my manager David Jurek for the support and the review, also thanks to Ming Zhou for the review.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact me at:

Yan Wang
Bristol-Myers Squibb
5 Research Parkway
Wallingford, CT 06443
Work Phone: 203-677-6273
yan.wang@bms.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.