PharmaSUG China 2024 - Paper SA-10041

LSMEANS vs ESTIMATE in proc mixed

Gloria Zhong, Merck Serono R&D Hub; Sukie Gao, Merck Serono R&D Hub;
Wayne Yang, Merck Serono R&D Hub

ABSTRACT

In a recent project involving the assessment of bioequivalence between two formulations using the Proc mixed procedure in SAS, the LSMEANS and ESTIMATE statements yielded identical least square mean differences. Although the output names generated by ODS for these statements varied, the estimated differences in treatment effects were consistent across output listings. This included identical results for the absolute values of t-statistics, p-values, and the confidence interval bounds, all of which could be corroborated through ODS traceability. This paper aims to elucidate the reasons behind this consistency and to offer guidance on choosing between these statements when constructing models.

INTRODUCTION

The evaluation of bioequivalence is a statistical process integral to the development and approval of generic pharmaceuticals. It involves comparing the bioavailability of a test drug to that of a reference drug to ensure they are within an acceptable range of variability. The Proc mixed procedure in SAS software stands as a cornerstone analytical tool for this purpose, allowing researchers to model and analyze data where responses are continuous and may be correlated. LSMEANS and ESTIMATE statements are pivotal in deriving least square mean differences-a critical measure in bioequivalence studies. Despite their distinct functionalities and output designations within the Output Delivery System (ODS), they often yield identical estimates for treatment differences. This convergence raises questions about the underlying mechanisms of these statements and the circumstances that might lead to the selection of one over the other. This paper delves into the principles of LSMEANS and ESTIMATE, exploring their similarities and differences, and provides a framework for practitioners to determine the most appropriate statement for their models.

Understanding LSMEANS and ESTIMATE

The LSMEANS statement computes least squares means (LS-means) or adjusted means, which are predictions of the marginal means. By using the "obsmargins" option, these means are based on the actual proportion of covariates/strata in the sample, providing a more accurate representation of the population of interest. This approach is particularly useful in the presence of an unbalanced design or missing data, as it does not affect the treatment effect and its inference. The LS-means offer a way to understand the effect of each level of a treatment while accounting for the real distribution of other factors.

In contrast, the ESTIMATE statement is designed to provide a direct estimate of specified linear combinations of the fixed-effects parameters in the model. This can include the difference between treatment levels or a comparison to a predefined constant. The ESTIMATE statement gives researchers the flexibility to test specific hypotheses about the fixed effects, which is why it is often employed in conjunction with LSMEANS to validate the treatment effect differences.

Take a bioequivalence project as an example. The primary objective of project is to build up model on fixed effect of sequence, treatment and period, and a random effect of subject within sequence. Ratios of geometric means and corresponding 90% CIs for Test/Reference is estimated, respectively:

```
proc mixed data=&dat;
class subject period sequence treatment;
model &var=sequence period treatment/cl residual;
random intercept / subject = subject(sequence);
lsmeans treatment / cl diff alpha = 0.1;
estimate 'Test vs Reference' treatment -1 1 0 / cl alpha = 0.1;
ods output lsmeans=&var.3;
ods output estimates=&var.5;
ods output diffs=&var.2;
ods output covparms=&var.1;
run;
```

Code above illustrates two statements in the same model. Without specifying sorting order of treatment then by default the treatment groups would be sorted by alphabetical order, e.g., point estimate of treatment group 'Test' and 'Reference' log value difference will end up in ratio of 'Ref/Test' when back transformed to original scale. (Edupganti 2011)

PROC MIXED selects the degrees of freedom to match those displayed in the "Tests of Fixed Effects" table for the final effect listed in the ESTIMATE statement. (SAS 2015)

Solution for Fixed Effects

Effect	sequence_	treatment_	Period	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept				9.4972	0.07318	48	129.79	<.0001	0.05	9.3501	9.6443
sequence_	Ref-Test			-0.01912	0.09523	37	-0.20	0.8420	0.05	-0.2121	0.1738
sequence_	Test-Ref			0							
PERIOD			1	-0.02337	0.03802	37	-0.61	0.5426	0.05	-0.1004	0.05367
PERIOD			2	0							
treatment_		Ref		0.007744	0.03802	37	0.20	0.8397	0.05	-0.06929	0.08478
treatment_		Test		0							

ESTIMATE 'label' < fixed-effect values . . . >
< | random-effect values . . . > < / options > **;**

The benefit of applying ESTIMATE statement is that ratio of 'Test/Ref' or 'Ref/Test' can be determined in this statement by setting up effect order. i.e. coefficient -1 matches the treatment group which should be placed as the denominator in treatment ratio.

LSMEANS fixed-effects < / options > **;**

DIFF< =difftype **>**

The LSMEANS statement computes least squares means (LS-means) of fixed effects. DIFF option requests that differences of the LS-means be displayed. (SAS 2015)

Both options provide output table containing the same result. Table estimate is produced by ESTIMATE statement. Table diff is produced by LSMEANS/diff.

Consistency in Estimates and Statistical Inference

Despite the fundamental differences in their application, LSMEANS and ESTIMATE have shown to produce consistent estimates for treatment differences across various types of models. This consistency extends to the absolute values of t-statistics, p-values, and the bounds of confidence intervals. The t-value provides the magnitude of difference relative to the variability observed in the data, the p-value offers the probability of observing such a difference by chance, and the confidence intervals provide a range in which the true treatment difference is expected to lie with a certain level of confidence.

The LSMEANS statement computes least squares means of fixed effects, which are equivalent to arithmetic means when the design is balanced (Jerry W. Davis 2017). In unbalanced designs or designs with missing data, where the simple arithmetic mean would be biased, this statement would be very useful. Least squares means are the only option for calculating treatment level means within the mixed model procedures. (Jerry W. Davis 2017)

The ESTIMATE statement specifies a hypothesis about the linear combination of the fixed effects. It is typically used to estimate the effect size of contrasts or comparisons that are not directly provided by the model.

CL option in each statement provides confidence intervals for these means and their differences. When you specify the alpha = 0.1 option in both statements, you are requesting the confidence intervals for the mean differences or estimates at the 90% confidence level.

For both the LSMEANS and ESTIMATE statements, the treatment difference is calculated as a linear combination of the fixed effects parameters. The confidence interval for this difference is then constructed using the standard error of this difference (derived from the variance-covariance matrix) and the appropriate quantile of the t-distribution (which depends on the alpha level and the degrees of freedom from the model). In summary, the reason why LSMEANS and ESTIMATE produce the same result for treatment difference is that they are both based on the same model and statistical theory. For ESTIMATE statement, a proper coefficients configuration will produce exactly the same result of LSMEANS, and for LSMEANS statement, the use of option E can obtain the linear combination coefficient matrix, which can be used in ESTIMATE statement to obtain exactly the same result of LSMEANS. An example is provided in Discussion section (Example 2).

Discussion

When conducting Mixed-effects models for repeated measures (MMRM) analysis in SAS, although the ESTIMATE and LSMEANS statements can provide identical information, there are some scenarios in which they are not interchangeable. Understanding their functions and appropriate use cases is also crucial. Here are the main differences between the ESTIMATE and LSMEANS statements and when they cannot be used interchangeably:

LSMEANS:

- Calculates and compares the least squares means (LSMEANS) of treatment groups, which are the means adjusted for the entire model.
- Primarily used for direct comparison of the average effects between different treatment groups.
- Automatically adjusts means in unbalanced designs to make them comparable. (SAS 2015)

An unbalanced study design occurs when the sample sizes across different groups (such as treatment groups) are not equal. This can happen due to various reasons such as differential dropout rates, recruitment challenges, missing visits, or other logistical issues during the study. In clinical studies, unbalanced designs are common, and it is crucial to use statistical methods that can appropriately handle this unbalance. Simple arithmetic means (simple averages) can be misleading because they might be more heavily influenced by the groups with more data points. As they are based on data only. LS-means, based on model, adjust for this by considering the model's structure and the overall data layout, providing a more accurate representation of each factor's effect. (Idaho n.d.)In unbalanced factorial experiments, LS means for each factor mimic the main effects means but are adjusted for imbalance. (Lenth January 2016)

LSMEANS can be seen as a special form of ESTIMATE, consider an example where you have a model with a single categorical treatment factor with three levels (A, B, C). Here's how you might use both LSMEANS and ESTIMATE in PROC MIXED:

```
proc mixed data=test;
   class trt;
   model response = trt;
   lsmeans trt; /* Compute LS-means for each level of treatment */
   estimate 'LS-mean A' trt 1 0 0; /* Manually specifying the LS-mean for
level A */
   estimate 'LS-mean B' trt 0 1 0; /* Manually specifying the LS-mean for
level B */
   estimate 'LS-mean C' trt 0 0 1; /* Manually specifying the LS-mean for
level C */
run;
```

Each ESTIMATE statement for a specific treatment level with coefficients (1 for the level of interest and 0 for others) mimics the computation of LS-means. It essentially estimates the mean of the response variable for that level, controlling for other factors in the model, which is what LSMEANS does.

ESTIMATE:

- Used to estimate custom linear combinations of fixed effects in the model.
- Can be used to construct more complex hypothesis tests, such as tests for specific linear combinations or interaction effects. (SAS 2015)
- Provides greater flexibility in specifying complex model effects and interactions. (Lemus 2016)

An example in SAS help documents outlined how ESTIMATE is used for a complex hypothesis as SLICE and LSMEANS statements cannot be used for this more complex contrast. (SAS 2015)

Scenarios Where They Are Not Interchangeable

When Testing Complex Parameter Combinations: If you need to test specific linear combinations of parameters in the model, such as the interaction effects between multiple variables at a specific time point, the ESTIMATE statement is necessary. LSMEANS only provides the average effects of treatment groups and is not directly suitable for complex linear hypothesis testing. For example, in studies involving multiple dose levels or multiple time points, the analysis may need to consider dose-response relationships or temporal dynamic effects. This may require layered analysis of data at different doses or points in time, or the inclusion in the model of the interaction of dose and time as continuous variables.

When Dealing with Non-standard Effects: When dealing with non-standard effects (such as non-linear effects or polynomial effects), ESTIMATE offers the flexibility to customize these effects. LSMEANS is generally not applicable for such complex model settings. For example, in multi-center or large-scale

studies, the data structure may be more complex, involving multiple levels of data (e.g., data from different experimental centers). In this case, it may be necessary to use a multilevel model to properly analyze the data to account for the hierarchy of the data and potential group effects.

When Specific Comparisons Rather Than Average Effects Are Needed: If the focus of the analysis is to compare specific treatment contrasts (such as Treatment A at Time Point 1 versus Treatment B at Time Point 2), and these contrasts are more than just comparisons of average effects, the ESTIMATE statement can precisely define these contrasts. Although LSMEANS can compare effects between different treatment groups, it does not provide the option to construct these specific contrasts. For example, in some studies, it may be necessary to adjust covariates that affect pharmacokinetic parameters, such as age, weight, sex, etc. The introduction of these covariates can be accounted for by including their interaction with the primary therapeutic effect in the model.

Example 1

Example code for 2X2 classic BE study design:

```
proc mixed data=&dat;
     class subject period sequence treatment(ref="Ref");
     model &var=sequence period treatment/cl residual;
     random intercept / subject = subject(sequence);
     lsmeans treatment / cl diff alpha = 0.1;
     estimate 'Test vs Reference' treatment 1 -1 / cl alpha = 0.1;
     ods output diffs=&var.1 estimates=&var.2;
run;
```

Below Display1 and Display 2 show that the LSMEAN statement gives the same Least Square Mean Difference result as the ESTIMATE statement.

Display 1. Least Square Mean Differences - LSMEANS statement

Differences of Least Squares Means										
Effect	Actual Treatment	Actual Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
TRTA	Test	Ref	-0.04448	0.03471	18	-1.28	0.2163	0.1	-0.1047	0.01571

Display 2. Least Square Mean Differences - ESTIMATE statement

Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Test vs Ref	-0.04448	0.03471	18	-1.28	0.2163	0.1	-0.1047	0.01571

Example 2

(1) Use ESTIMATE to custom linear combinations of fixed effects in the model:

Consider a simple study with two treatments and five repeated measures for the primary efficacy endpoint. Then the model statement in PROC MIXED for the longitudinal analysis can be: MODEL AVAL = TRT VIS TRT*VIS; where AVAL is the primary efficacy measure, TRT and VIS are the treatment and visit variables, respectively.

Dummy datasets as below:

```
data test;
    seed = 6342454;
    retain usubjid 0;
    retain last_trtn .;
    do rep = 1 to ceil(ranuni(seed) * 3) + 3;
       do trtn = 1 to 2;
           if last_trtn ne trtn then do;
               usubjid + 1;
               last_trtn = trtn;
           end;
           do vis = 1 to 5;
               aval = 3 + trtn + vis + trtn * vis + rannor(seed);
               output;
           end;
       end;
    end;
    keep usubjid trtn vis aval;
run;
```

A partial display of the dataset:

usubjid	trtn	vis	aval
1	1	1	5.706282
1	1	2	8.212366
1	1	3	7.008195
1	1	4	11.32771
1	1	5	13.11545
2	2	1	8.464592
2	2	2	11.59902
2	2	3	13.72082
2	2	4	17.08439
2	2	5	19.39241

Get the coefficients needed for the ESTIMATE statements through the LSMEANS statement: The E option in LSMEANS statement (Table 1). Zeros in this table are shown as blanks for clarity. Note in the CLASS statement, the order of variables is important. The ordering of Row1 - Row10 in Table 1 typically depends on the order in which the variables are specified in the CLASS statement and the setting of the ORDER= option in the PROC or CLASS statement.

```
proc mixed data = test;
     class trtn vis;
     model aval=trtn vis trtn*vis/solution;
     lsmean trtn*vis/e diff;
run;
```

Table1: Coefficients for trtn*vis Least Squares Means

t1v1	t1v2	t1v3	t1v4	t1v5	t2v1	t2v2	t2v3	t2v4	t2v5

Coefficients for trtn*vis Least Squares Means												
Effect	**trtn**	**vis**	**Row1**	**Row2**	**Row3**	**Row4**	**Row5**	**Row6**	**Row7**	**Row8**	**Row9**	**Row10**
Intercept			1	1	1	1	1	1	1	1	1	1
trtn	1		1	1	1	1	1					
trtn	2							1	1	1	1	1
vis		1	1					1				
vis		2		1					1			
vis		3			1					1		
vis		4				1					1	
vis		5					1					1
trtn*vis	1	1	1									
trtn*vis	1	2		1								
trtn*vis	1	3			1							
trtn*vis	1	4				1						
trtn*vis	1	5					1					
trtn*vis	2	1						1				
trtn*vis	2	2							1			
trtn*vis	2	3								1		
trtn*vis	2	4									1	
trtn*vis	2	5										1

Table 2: Transpose the coefficients table for easy viewing against code

The highly parts are the coefficients of T1V1, T2V1, and T1V1 versus T2V1 respectively.

	Effect	Intercept	trtn	trtn	vis	vis	vis	vis	vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis
	trtn		1	2						1	1	1	1	1	2	2	2	2	2
	vis				1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
t1v1	Row1	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
t1v2	Row2	1	1			1					1								
t1v3	Row3	1	1				1					1							
t1v4	Row4	1	1					1					1						
t1v5	Row5	1	1						1					1					
t2v1	Row6	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
t2v2	Row7	1		1		1										1			
t2v3	Row8	1		1			1										1		
t2v4	Row9	1		1				1										1	
t2v5	Row10	1		1					1										1
t1v1- t2v1		0	1	-1	0	0	0	0	0	1	0	0	0	0	-1	0	0	0	0

For ESTIMATE statements, note that trailing zero coefficients can be omitted.

```
proc mixed data=test;
     class trtn vis;
     model aval=trtn vis trtn*vis/solution;

     estimate "T1V1: LSM in vis1 & trt1"
               intercept 1
               trtn 1 0
               vis 1 0 0 0 0
               trtn*vis 1 0 0 0 0 0 0 0 0 0;

     estimate "T2V1: LSM in vis1 & trt2"
               intercept 1
               trtn 0 1
               vis 1 0 0 0 0
               trtn*vis 0 0 0 0 0 1 0 0 0 0;

     estimate "T1V1-T2V1: LSM diff between trt1 & trt2 at vis1"
               trtn 1 -1
               trtn*vis 1 0 0 0 0 -1 0 0 0 0;

     estimate "T1V5: LSM in vis5 & trt1"
               intercept 1
               trtn 1 0
               vis 0 0 0 0 1
               trtn*vis 0 0 0 0 1 0 0 0 0 0;

     estimate "T2V5: LSM in vis5 & trt2"
```

```
                intercept 1
                trtn 0 1
                vis 0 0 0 0 1
                trtn*vis 0 0 0 0 0 0 0 0 0 1;

      estimate "T1V5-T2V5: LSM diff between trt1 & trt2 at vis5"
                trtn 1 -1
                trtn*vis 0 0 0 0 1 0 0 0 0 -1;
run;
```

Result 1. Least Square Mean & Mean Differences - LSMEANS statement

Least Squares Means

Effect	trtn	vis	Estimate	Standard Error	DF	t Value	Pr > |t|
trtn*vis	1	1	6.1333	0.4208	40	14.58	<.0001
trtn*vis	1	2	8.3951	0.4208	40	19.95	<.0001
trtn*vis	1	3	10.1476	0.4208	40	24.12	<.0001
trtn*vis	1	4	11.5402	0.4208	40	27.43	<.0001
trtn*vis	1	5	14.0028	0.4208	40	33.28	<.0001
trtn*vis	2	1	7.4162	0.4208	40	17.62	<.0001
trtn*vis	2	2	10.5711	0.4208	40	25.12	<.0001
trtn*vis	2	3	13.7351	0.4208	40	32.64	<.0001
trtn*vis	2	4	16.9652	0.4208	40	40.32	<.0001
trtn*vis	2	5	20.2921	0.4208	40	48.22	<.0001

Differences of Least Squares Means

Effect	trtn	vis	_trtn	_vis	Estimate	Standard Error	DF	t Value	Pr > |t|
trtn*vis	1	1	1	2	-2.2619	0.5951	40	-3.8	0.0005
trtn*vis	1	1	1	3	-4.0143	0.5951	40	-6.75	<.0001
trtn*vis	1	1	1	4	-5.407	0.5951	40	-9.09	<.0001
trtn*vis	1	1	1	5	-7.8695	0.5951	40	-13.22	<.0001
trtn*vis	1	1	2	1	-1.2829	0.5951	40	-2.16	0.0372
....									
trtn*vis	1	5	2	5	-6.2893	0.5951	40	-10.57	<.0001

Result 2. Least Square Mean & Mean Differences - ESTIMATE statement

Estimates					
Label	Estimate	Standard Error	DF	t Value	Pr > |t|
T1V1: LSM in vis1 & trt1	6.1333	0.4208	40	14.58	<.0001
T2V1: LSM in vis1 & trt2	7.4162	0.4208	40	17.62	<.0001
T1V1-T2V1: LSM diff between trt1 & trt2 at vis1	-1.2829	0.5951	40	-2.16	0.0372
T1V5: LSM in vis5 & trt1	14.0028	0.4208	40	33.28	<.0001
T2V5: LSM in vis5 & trt2	20.2921	0.4208	40	48.22	<.0001
T1V5-T2V5: LSM diff between trt1 & trt2 at vis5	-6.2893	0.5951	40	-10.57	<.0001

(2) Use ESTIMATE to construct more complex hypothesis tests

Suppose you want to test the average of T1V1 and T2V1 versus the average of T1V5 and T2V5, the LSMEANS statements cannot be used for this more complex contrast, where the ESTIMATE statements provide greater flexibility. In below table (Table 3):

Row No. 15 is the coefficients for average of T1V1 and T2V1, which is calculate from row No. 4 and 9.

Row No.16 is the coefficients for average of T1V5 and T2V5, which is calculate from row No.8 and 13.

Row No.17 is the coefficients for average of T1V1 and T2V1 versus the average of T1V5 and T2V5, which is calculate from row No.15 and 16.

Table 3: Coefficients for complex hypothesis tests

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1		Effect	Intercept		trtn	trtn		vis	vis	vis	vis	vis		trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis	trtn *vis
2		trtn			1	2								1	1	1	1	1	2	2	2	2	2
3		vis						1	2	3	4	5		1	2	3	4	5	1	2	3	4	5
4	t1v1	Row1	1		1	0		1	0	0	0	0		1	0	0	0	0	0	0	0	0	0
5	t1v2	Row2	1		1				1						1								
6	t1v3	Row3	1		1					1						1							
7	t1v4	Row4	1		1						1						1						
8	t1v5	Row5	1		1	0		0	0	0	0	1		0	0	0	0	1	0	0	0	0	0
9	t2v1	Row6	1		0	1		1	0	0	0	0		0	0	0	0	0	1	0	0	0	0
10	t2v2	Row7	1			1			1											1			
11	t2v3	Row8	1			1				1											1		
12	t2v4	Row9	1			1					1											1	
13	t2v5	Row10	1		0	1		0	0	0	0	1		0	0	0	0	0	0	0	0	0	1
14																							
15	1/2 (t1v1 + t2v1)		1		0.5	0.5		1	0	0	0	0		0.5	0	0	0	0	0.5	0	0	0	0
16	1/2 (t1v5 + t2v5)		1		0.5	0.5		0	0	0	0	1		0	0	0	0	0.5	0	0	0	0	0.5
17	1/2 (t1v1 + t2v1) - 1/2 (t1v5 + t2v5)		0		0	0		1	0	0	0	-1		0.5	0	0	0	-0.5	0.5	0	0	0	-0.5

```
proc mixed data = test;
     class trtn vis;
     model aval=trtn vis trtn*vis/solution;

     estimate "avg T1V1+T2V1: Average LSM of trt1 & trt2 at vis1"
               intercept 1
               trtn 0.5 0.5
               vis 1 0 0 0 0
               trtn*vis 0.5 0 0 0 0 0.5 0 0 0 0;

     estimate "avg T1V5+T2V5: Average LSM of trt1 & trt2 at vis5"
               intercept 1
               trtn 0.5 0.5
               vis 0 0 0 0 1
               trtn*vis 0 0 0 0 0.5 0 0 0 0 0.5;

     estimate "avg T1V1+T2V1 - avg T1V5+T2V5:
               average of T1V1 and T2V1 versus the average of T1V5 and
T2V5"
               vis 1 0 0 0 -1
               trtn*vis 0.5 0 0 0 -0.5 0.5 0 0 0 -0.5;
run;
```

Result 3: Output from ESTIMATE statement for complex hypothesis tests

Estimates					
Label	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
avg T1V1+T2V1: Average LSM of trt1 & trt2 at vis1	6.7747	0.2975	40	22.77	<.0001
avg T1V5+T2V5: Average LSM of trt1 & trt2 at vis5	17.1475	0.2975	40	57.63	<.0001
avg T1V1+T2V1 - avg T1V5+T2V5: average of T1V1 and T2V1 versus the average of T1V5 and T2V5	-10.3727	0.4208	40	-24.65	<.0001

CONCLUSION

For the comparison of average effects between different treatment groups in MMRM analysis, the LSMEANS and ESTIMATE in PROC MIXED can yield identical least square mean differences.

LSMEAN primarily used for direct comparison of the average effects between different treatment groups. The advantage of using LSMEANS is that it directly calculates the adjusted mean of the treatment group, which is very intuitive and easy to understand and interpret. And when there is an unbalanced study

design, LSMEANS automatically adjusts for this, ensuring that the resulting mean of the treatment groups is comparable, which is very common in actual clinical trials. And the steps for making intergroup comparisons and generating confidence intervals using LSMEANS are relatively simple and easy to implement, especially in standard treatment group comparisons.

The advantage of using ESTIMATE is that it allows users to customize complex linear combinations, which is useful when you need to test the impact of a particular combination of parameters in a model. Besides, when the analysis involves complex interactions or non-standard effects, the ESTIMATE provides the possibility to conduct these analyses, and it can precisely define and test these complex effects. In addition, ESTIMATE is very useful for exploratory data analysis, helping researchers to discover potential patterns and relationships in the data.

The choice of which to use should be based on the specific needs of the analysis and the predefined statistical hypotheses. However, the construction of ESTIMATE statements can get very complicated when study designs become complex, programming can be tedious and challenging. Admittedly, it's one of the tasks that many analysts find daunting. (David J., 2010).

REFERENCES

Edupganti, Shilpa. 2011. "Proc Mixed - Right Options to get Right Output." *NESUG.*

Idaho, Collage of agricultural and life sciences University of. n.d. *LSMEANS.* Staitsitcal Programs SAS workshop - Proc Mixed.

Jerry W. Davis, University of Georgia, Griffin Campus. 2017. "Linear Mixed Models with Random Effects."

Lenth, Russell V. January 2016. "Least-square Means: the R package lsmeans." *Journal of Statistical software* (Journal of Statistical Software) Volumn 69 Issue 1.

SAS. 2015. "SAS/STAT® 14.1 User's Guide."

David J. Pasta. 2010." Practicalities of Using ESTIMATE and CONTRAST Statements"

Lemus, Hector, 2016. San Diego State University. "Demystifying the CONTRAST and ESTIMATE Statement." SAS WUSS

CONTACT INFORMATION

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

Gloria Zhong
Merck Serono
Gloria.zhong@merckgroup.com

Sukie Gao
Merck Serono
Sukie.Gao@merckgroup.com

Wayne Yang
Merck Serono
Wayne.Yang@merckgroup.com