

July 23, 2021

To: Lesley Inker, MD, MS
Tom Greene, PhD
Hocine Tighiouart, MS
Hiddo L. Heerspink, PhD

From: Edward F. Vonesh, PhD

Topic: Amendment to technical report summarizing different methods for determining and characterizing the potential timing of acute effects based on select CKD clinical trials.

1. Introduction

This amendment to the technical report issued on December 10, 2020 addresses the issue of taking a pseudo-random sampling approach to selecting a subset of subjects from the larger trials (ALTITUDE and EMPA-REG) versus taking a purely random sample using PROC SURVEYSELECT. In this amended report, we present results from: 1) the ALTITUDE study for random samples of 800, 400, and 200 subjects per treatment group, 2) the EMPA-REG study for random samples of 600, 300, and 100 subjects per treatment group and 3) the Zuchelli study with all 121 subjects included. The random samples for the ALTITUDE and EMPA-REG studies were chosen so as to represent a relatively large (600-800), moderately large (300-400) and relatively small Phase 2 trial of chronic kidney disease (CKD). Table 1 lists the three studies along with the total number of subjects per treatment group.

Table 1: List of Studies and Frequency of Subjects per Treatment Group

<i>Study ID</i>	<i>Study</i>	<i>Treatment</i>	<i>Frequency</i>
1	ALTITUDE	Placebo/Control	4090
		Treatment	4060
2	EMPA-REG	Placebo/Control	2321
		Treatment	4615
9	Zuchelli	Placebo/Control	61
		Treatment	60

For each study, the CKD-EPI analytical team developed an algorithm for determining a balanced set of designated visit times (in months) based on the frequency of measurements occurring within defined windows of time (in months). This was done so as to avoid the nuances associated with unbalanced irregularly spaced visits. Follow-up was limited to two years so as to more closely mimic how a Phase 2 clinical trial might be conducted.

Determination of the timing of an acute effect was determined using one of three modeling approaches:

- 1) A semi-parametric repeated measures ANOVA (RM-ANOVA) approach,
- 2) A semi-parametric repeated measures change-from-baseline ANCOVA (RM-ANCOVA) approach using centered baseline eGFR values as a covariate,
- 3) A parametric two-stage linear spline mixed-effects model.

A more thorough description of these approaches was provided in the original technical report. Section 2 briefly describes the methodology used for random sampling of subjects while Section 3 presents a summary of results.

2. Methods

Rather than use a pseudo random sample as described in the original technical report (i.e., taking the first N subjects from each treatment arm with an arbitrarily selected starting subject), the SAS program now uses the SAS procedure SURVEYSELECT to select a truly random sample from each treatment arm. Below is the SAS code used to create a SAS dataset with a random sample of N subjects per treatment arm.

```
proc SurveySelect data=TempStudy Out=Study_Sample
                method=PPS seed=&seed N=&Sample;
  Strata Treatment;
  Cluster SampleID;
  Size eGFR0;
run;
```

The option, method=PPS, requests that the selection of subjects be done using a probability proportional to size and without replacement. The `Strata Treatment;` statement requests that a stratified random sample be taken according to the two treatment groups while the `Cluster SampleID;` statement requests that all observations clustered within a subject be collected where `SampleID` is simply a created numeric ID number for each unique subject. The `Size eGFR0;` statement requests subjects be selected using a probability that is proportional to their starting eGFR values, `eGFR0`. The macro variable `&seed` specifies a fixed starting seed for the random selection process so that the same sample would be selected at a later date if needed while the macro variable `&Sample` simply specifies what sample size one wishes to use for each treatment group.

Some additional modifications to the SAS program were also made so as to accommodate the random sampling of subjects. Below is an example of SAS code utilizing five SAS macro programs that 1) generates a random sample of 800 subjects per treatment arm from the ALTITUDE study (`%Study`); 2) fits a RM-ANOVA model and a RM-ANCOVA model (`%RM_ANOVA`) to the randomly selected subjects; 3) fits a parametric two-stage linear spline mixed-effects model (`%GetKnot(Listing=close, MaxKnots=12)`) to the randomly selected subjects; and 4) summarizes results for each of the models in a series of plots (`%SGplots`). The RM-ANOVA and RM-ANCOVA models can be used to determine the timing and magnitude of an acute treatment effect based solely on the observed times (months) at which eGFR measurements are obtained. In contrast, the fully parametric linear spline mixed-effects model can be used to determine, via interpolation, the optimal timing of an acute treatment effect based on monthly increments rather than only months where eGFR is measured.

Example: The ALTITUDE study with a random sample of 800 subjects per treatment from the overall cohort.

```
%ClearTempDirectories;
%Study(StudyID=1, Sample=800, Random=YES, Seed=3638569, Stratum=Overall);
%RM_ANOVA;
%GetKnot(Listing=close, MaxKnots=12);
%SGplots;
```

3. Results

Table 2 summarizes the estimated timing and magnitude of an acute treatment effect for each of the 3 studies considered in this amendment. The “optimal” knot for the ANOVA and ANCOVA models were determined based solely on the months where eGFR was measured while the “optimal” knot for the linear spline mixed-effects model is based on the month that provides the best fit (lowest AIC) whether it be an observed month or an interpolated month. As shown in Appendices 1-3, the optimal knot for the ANOVA and ANCOVA models coincide with each other for all three studies. For the ALTITUDE study, there were some moderate differences in the timing of the acute effects based on the ANCOVA approach versus that of the linear spline mixed-effects modeling approach. This may be the result of having different estimates of the intercept for the linear spline model. If one assumes a common intercept (as would be expected for RCT’s) and one computes the acute treatment effect based on this

assumption together with the estimated acute and chronic slopes, the acute treatment effect from the linear spline mixed-effects model is fairly concordant with results obtained under the change-from-baseline RM-ANCOVA approach for all three studies.

Table 2. Summary of the optimal timing and magnitude of acute treatment effects (Treated-Control) by study with the optimal timing (knot) for the ANCOVA model determined on the basis of those months where eGFR measurements were taken while the optimal knot for the linear spline mixed-effects model is based on the observed or interpolated month that provides the best fit.

Study N per Treatment	ANCOVA Model		Linear Spline Mixed-Effects Model		
	Optimal* Timing (Knot)	Acute Treatment Effect (95% CL) (mL/min/1.73m ²)	Optimal Timing (Knot)	Acute Treatment Effect (95% CL) Estimated Intercepts (mL/min/1.73m ²)	Acute Treatment Effect (95% CL) Common Intercepts (mL/min/1.73m ²)
ALTITUDE	Month		Month		
N=(800,800)	3	-1.53 (-2.48, -0.57)	5	-0.40 (-2.62, 1.82)	-1.09 (-2.03, -0.15)
N=(400,400)	3	-2.24 (-3.63, -0.85)	7	-4.83 (-8.18, -1.48)	-2.21 (-3.66,-0.75)
N=(200,200)	3	-0.53 (-2.57, 1.50)	4	-0.87 (-5.42, 3.67)	-1.53 (-3.43, 0.37)
EMPA-REG	Month		Month		
N=(600,600)	1	-3.24 (-4.19, -2.29)	1	-3.82 (-6.05, -1.60)	-2.36 (-3.21, -1.51)
N=(300,300)	1	-3.47 (-4.80, -2.14)	1	-3.17 (-6.23, -0.10)	-2.65 (-3.84, -1.47)
N=(100,100)	1	-1.79 (-4.40, 0.82)	1	-4.02 (-9.35, 1.31)	-1.88 (-4.03, 0.27)
Zuchelli					
N=(61,60)	6	0.49 (-2.23, 3.21)	7	0.23 (-4.10, 4.55)	0.24 (-1.96, 2.44)

Detailed graphical results for the three (3) studies are shown in Appendices 1-3 according to the order in which the studies are listed in Table 1. An initial call to the macro %*Study* is done to obtain a random sample per treatment group of 800, 400, and 200 for the ALTITUDE study, and 600, 300 and 100 for the EMPA-REG study. This was done so as to reflect possible sample sizes one might target for a Phase 2 trial. For each study, a total of 7 graphs depicting the potential timing and size of an acute treatment effect are shown for each considered sample size. Compared with the original report, using a purely random sample of subjects per treatment group yields much more consistent results for the ALTITUDE study with regards to timing and magnitude of the acute treatment effects.

It should be noted that model convergence with a positive-definite Hessian matrix and full-rank positive-definite random-effects covariance matrix was achieved for both the ALTITUDE and Zuchelli studies. For the EMPA-REG trial, model convergence was achieved for each knot considered in the algorithms along with a full-rank positive-definite Hessian matrix. However, at the selected optimal knot of 1 month, the covariance matrix of the random effects under the linear spline mixed-effects model was not positive definite as the variance of the random acute slope effect was 0. This is not unexpected as it merely reflects the fact that there is only a single post-baseline time point with which to estimate a random acute slope effect. The linear spline mixed-effects model was run a second time using a knot at month 1 but without an acute random slope effect. The estimates were identical to two decimal places with minor differences in the confidence limits indicating that a value of 0 for the acute slope variance gives essentially the same results as dropping the acute slope random effect all-together.

4. Conclusions

The analyses presented here illustrate the importance of using a purely random sample of subjects when estimating the timing and magnitude of acute treatment effects based on sample sizes that may be more realistic for phase 2 trials. The two large studies and one small study selected here will be used in a manuscript for publication describing an overall strategy for estimating the timing and magnitude of acute treatment effects in

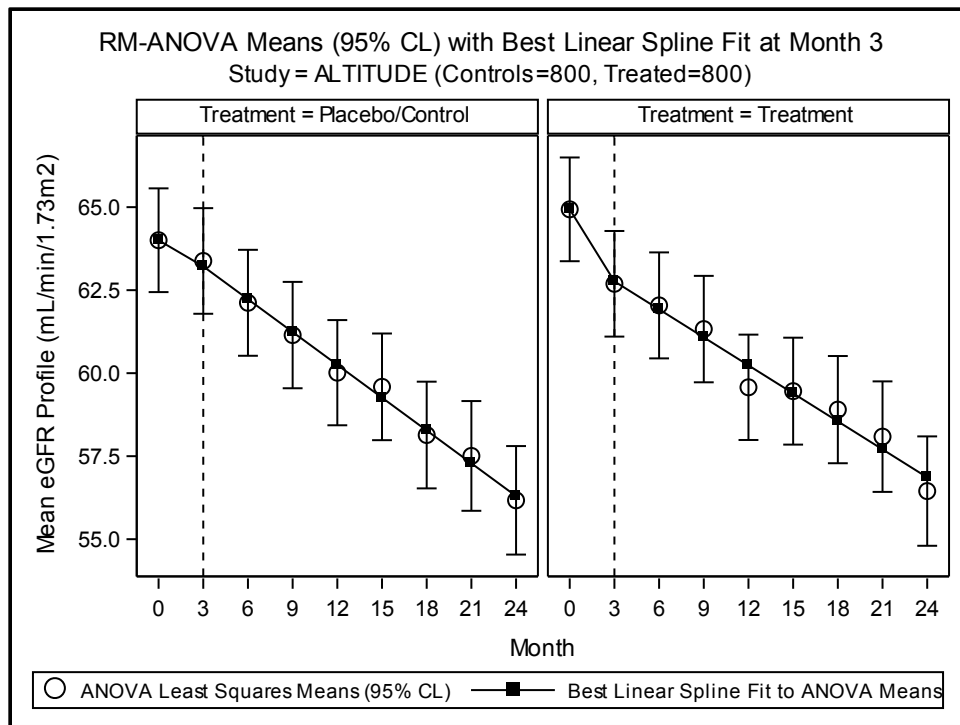
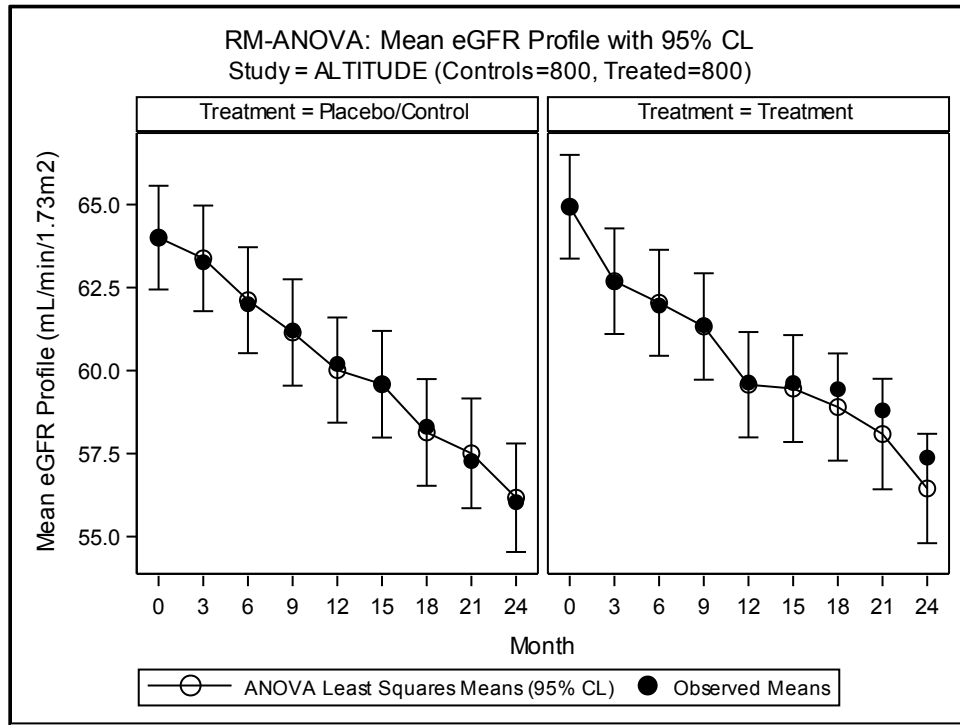
CKD trials – both large and small. There are advantages and disadvantages to using either the RM-ANCOVA modeling approach or the linear spline-mixed-effects modeling approach. By comparing how concordant or discordant results from both methods are and how well those results agree with clinical experience and knowledge of the underlying intervention, one can perhaps better assess what a reasonable choice would be when planning a phase 3 trial.

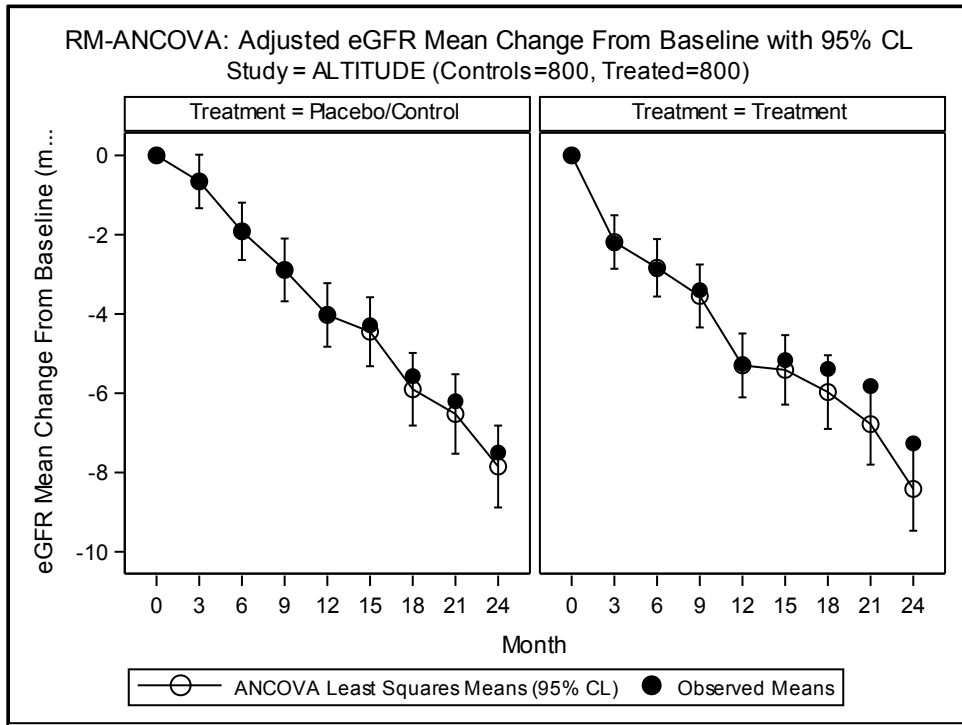
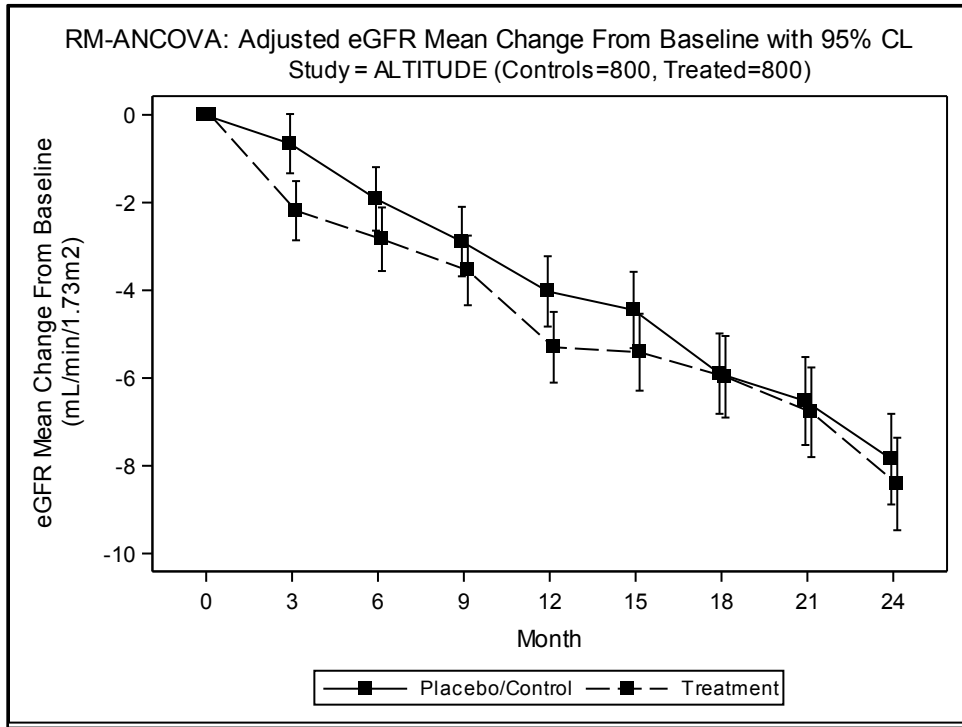
References

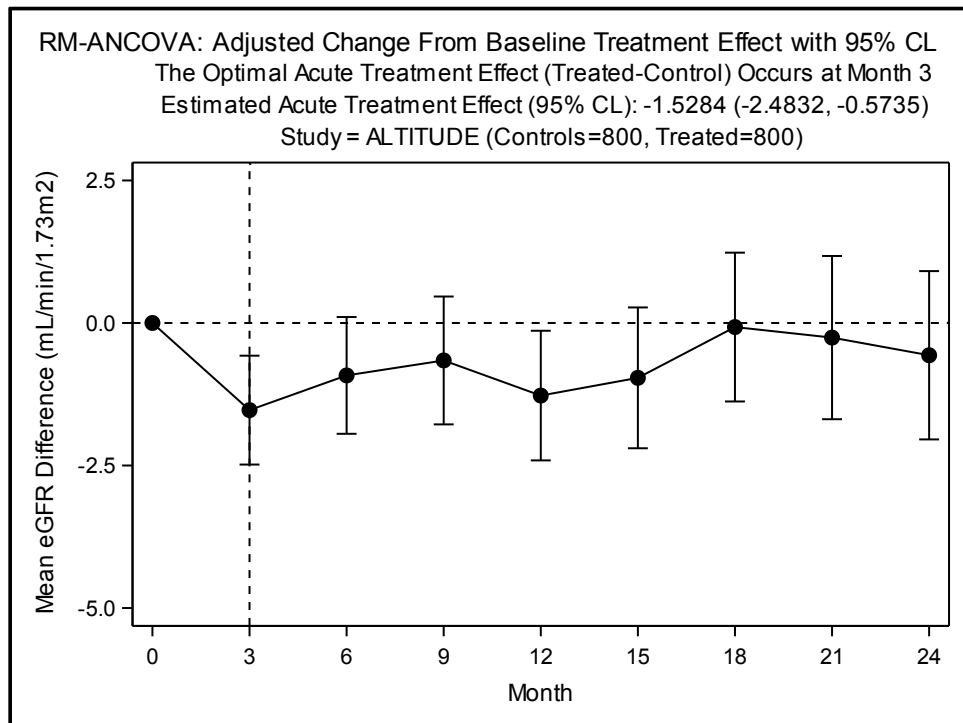
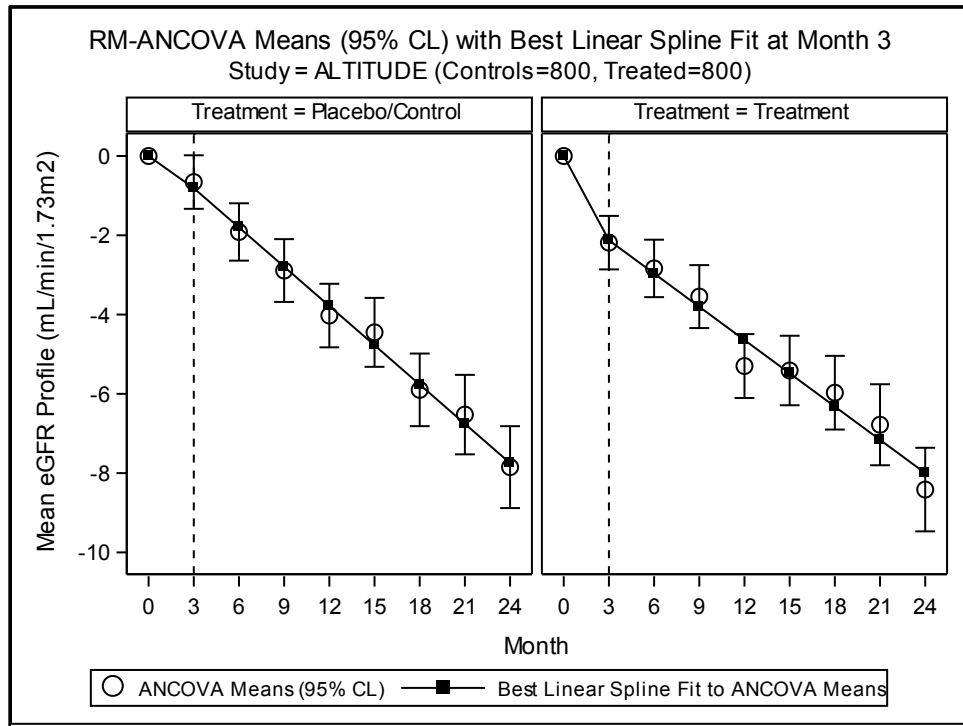
1. Vonesh, E., Tighiouart, H., Ying, J., Heerspink, H. L., Lewis, J., Staplin, N., et al. (2019). Mixed-effects models for slope-based endpoints in clinical trials of chronic kidney disease. *Statistics in Medicine* 38, 4218-4239.
2. Vonesh, E. Technical report summarizing different methods for determining and characterizing the potential timing of acute effects based on select CKD clinical trials. December 10, 2020.

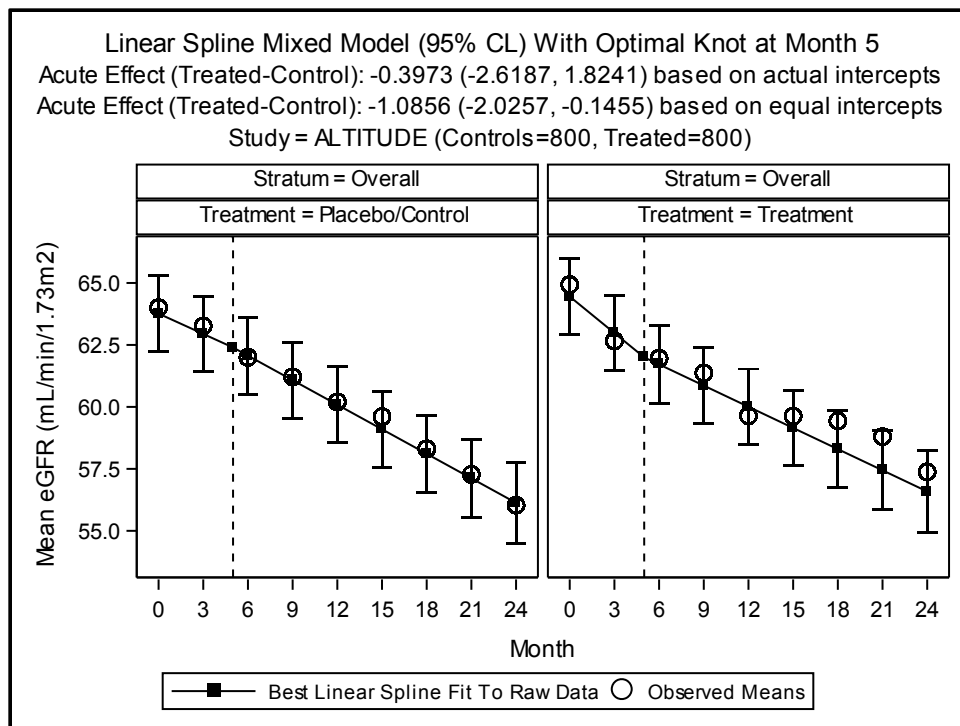
Appendix 1: ALTITUDE results

Case 1: Sample Size=800 per Treatment group

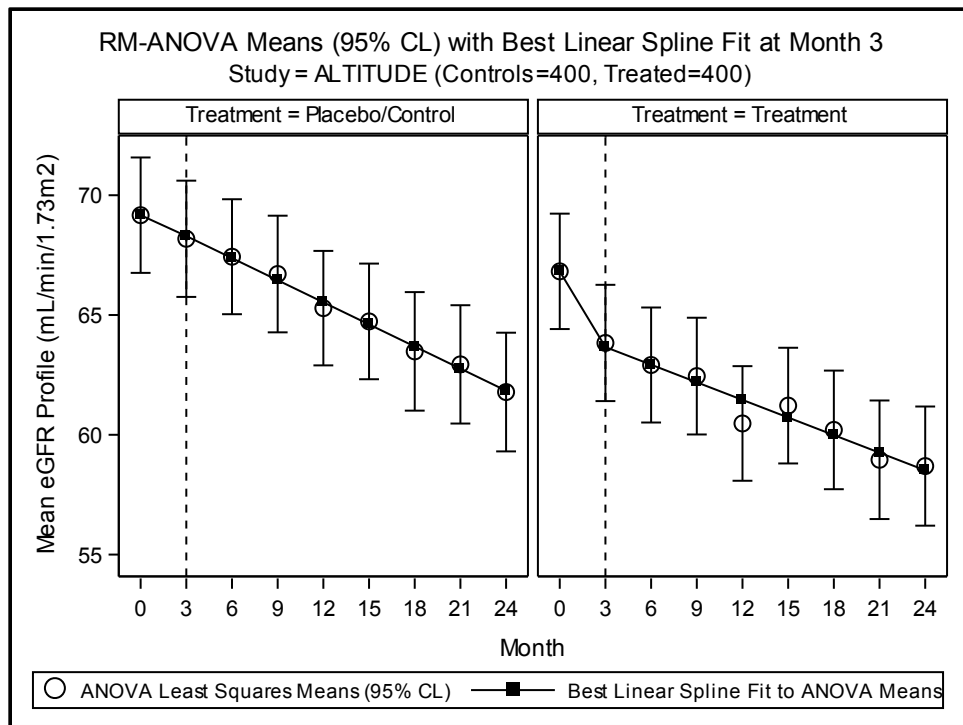
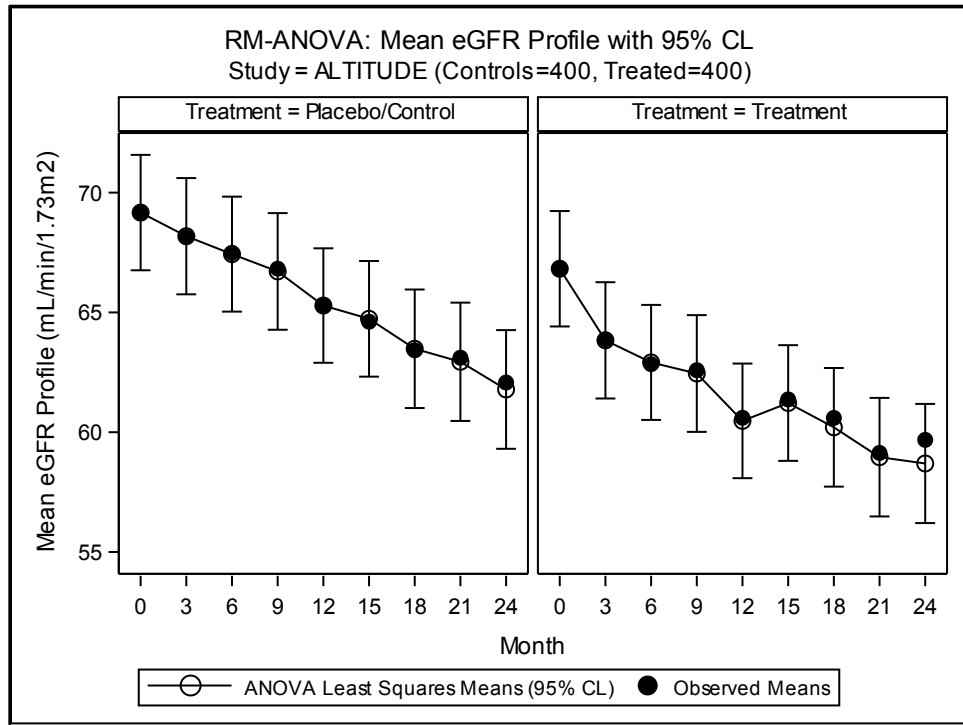


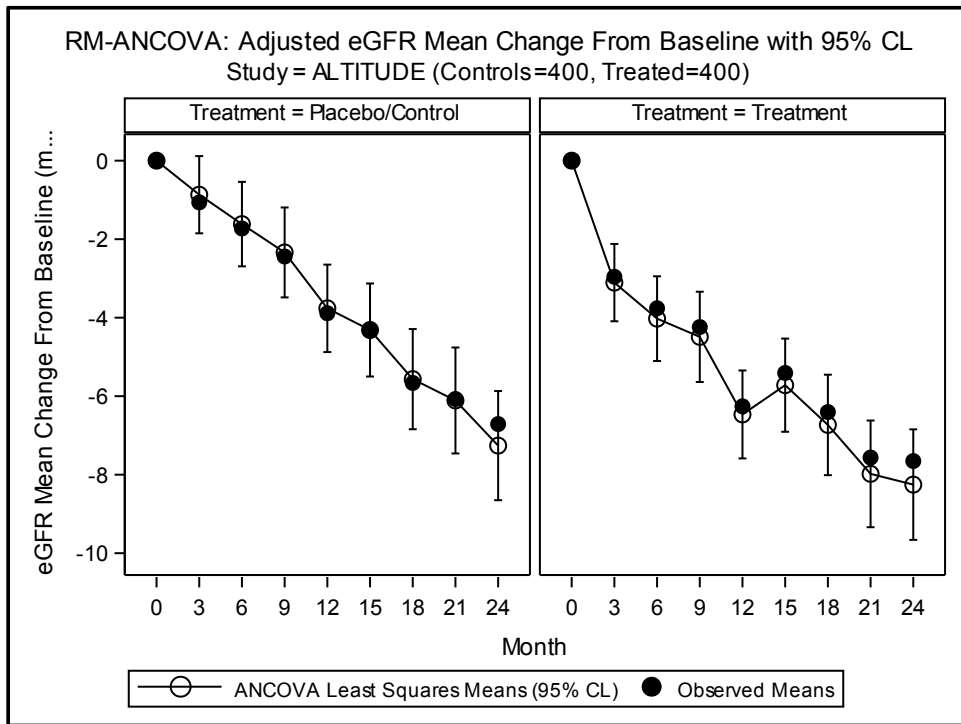
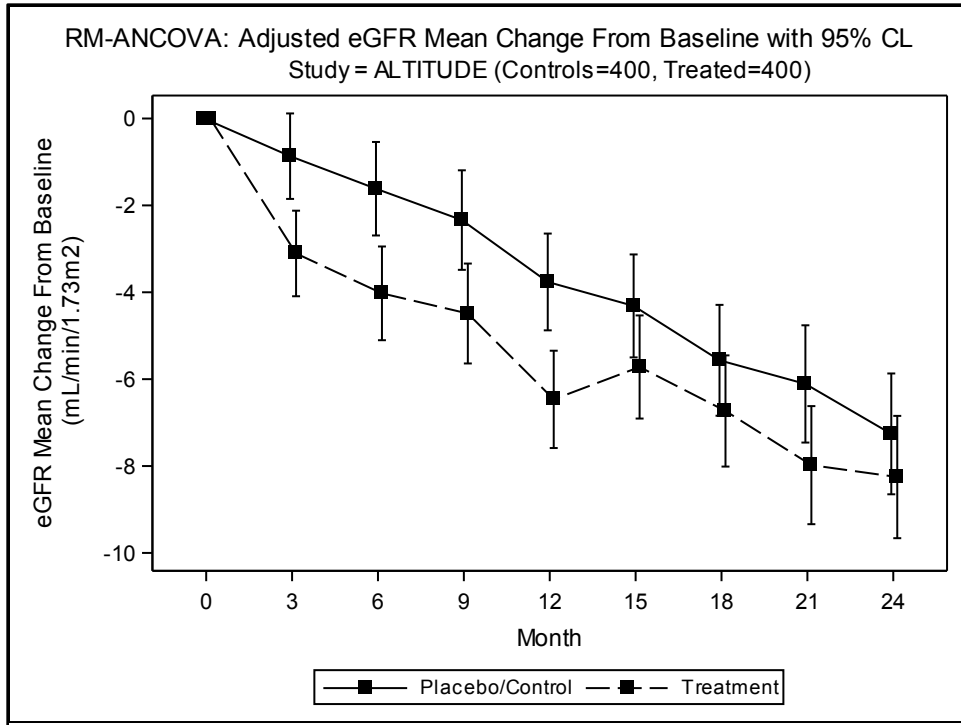


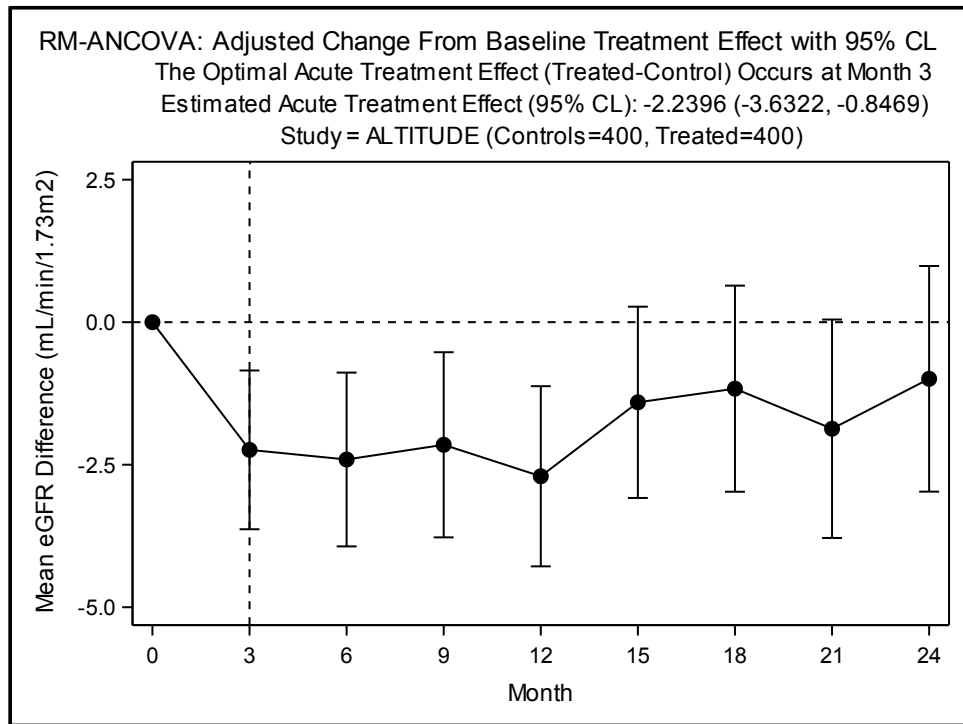
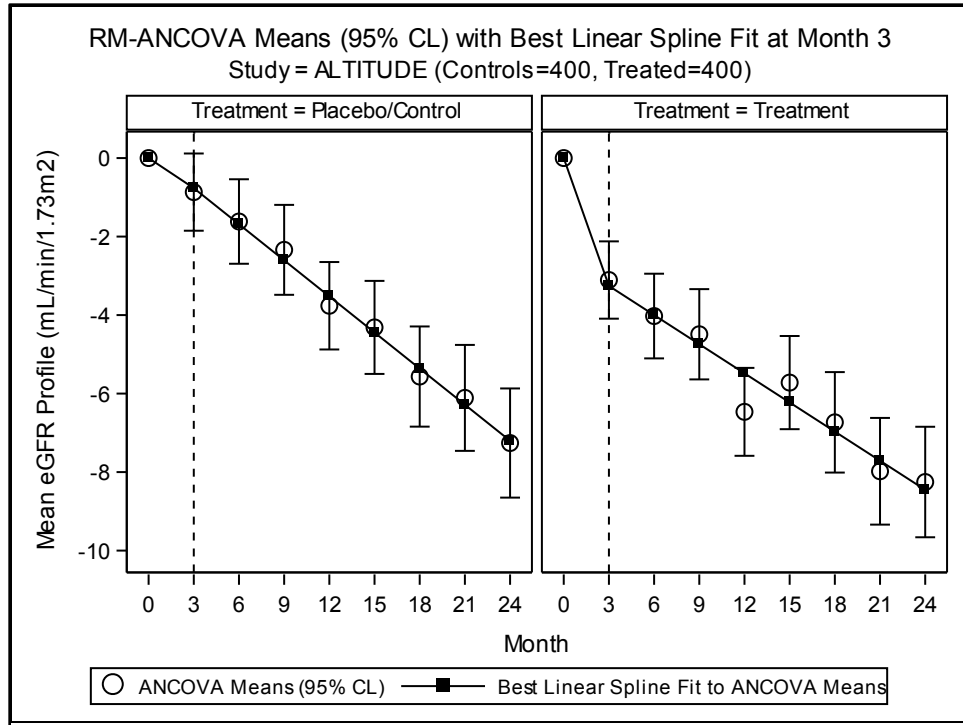


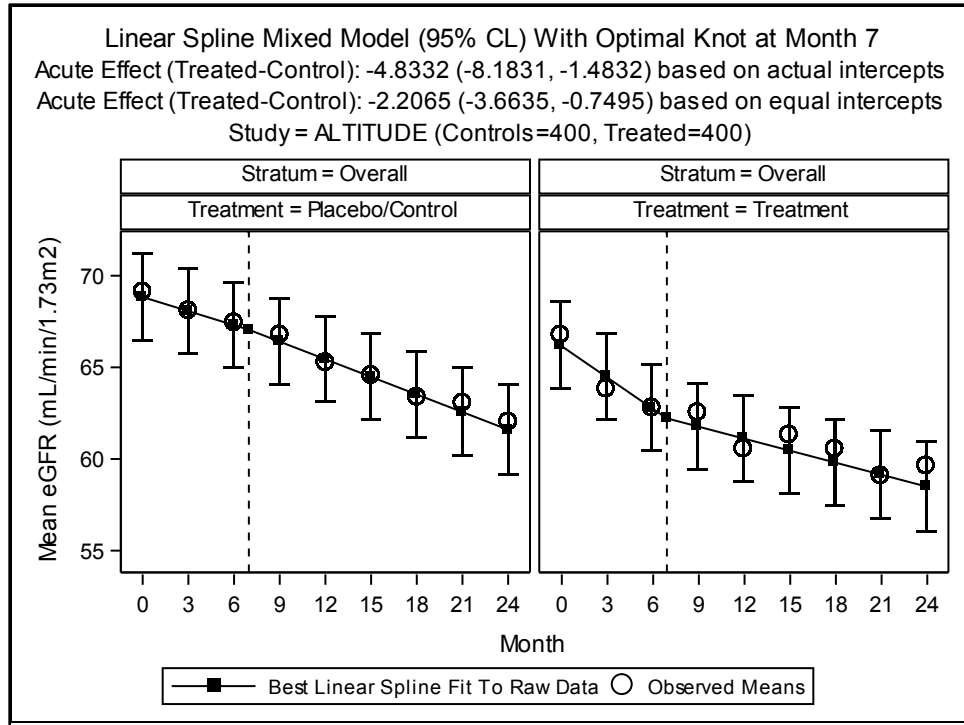


Appendix 1: ALTITUDE results
 Case 2: Sample Size=400 per Treatment group

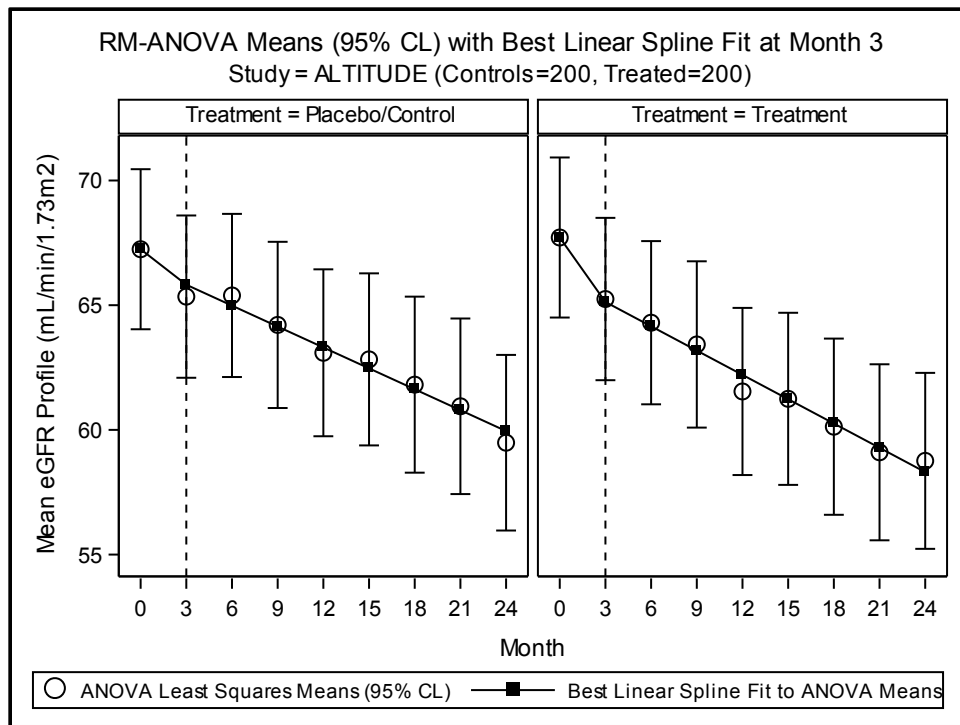
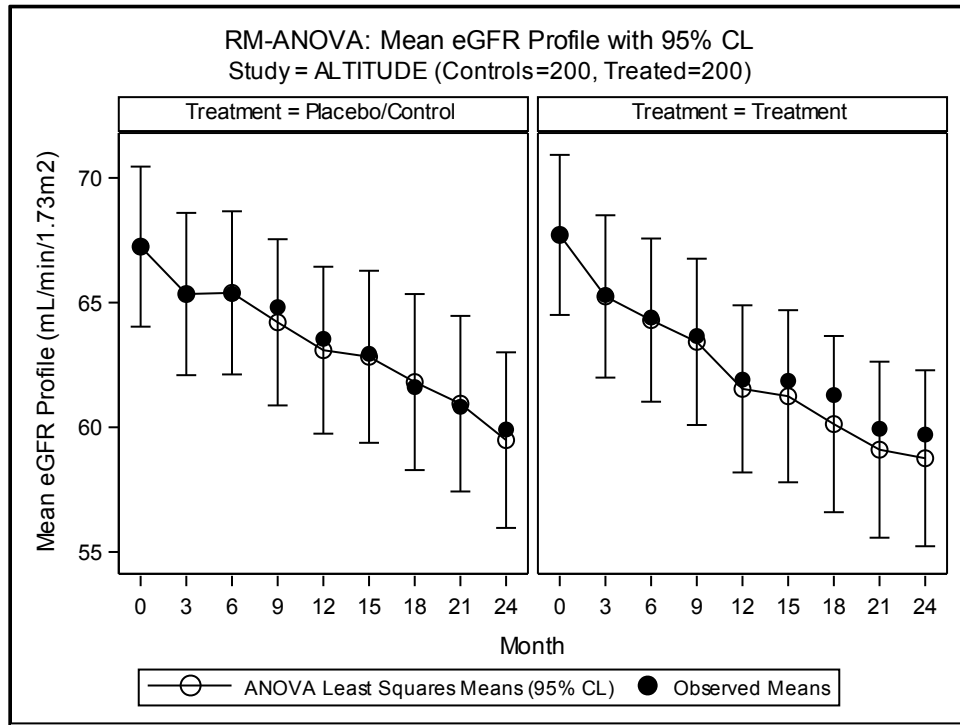


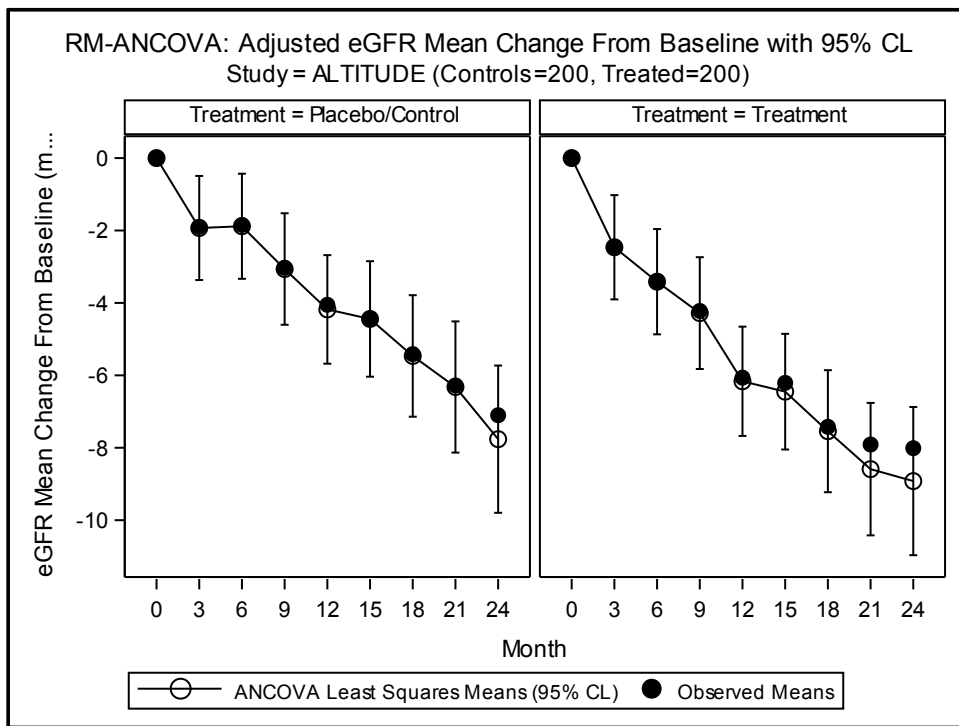
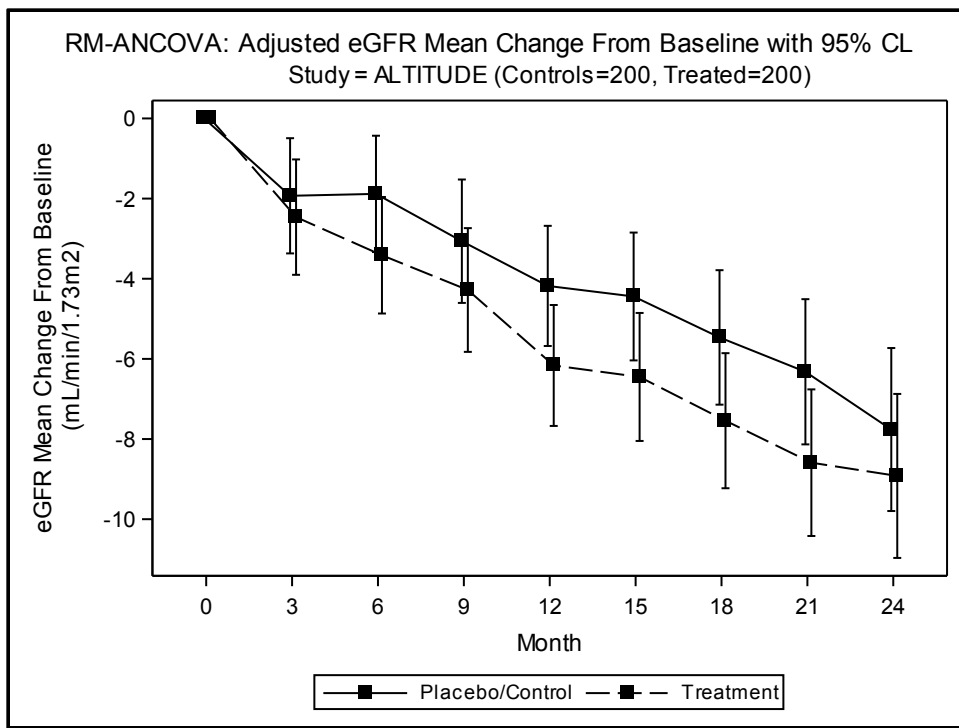


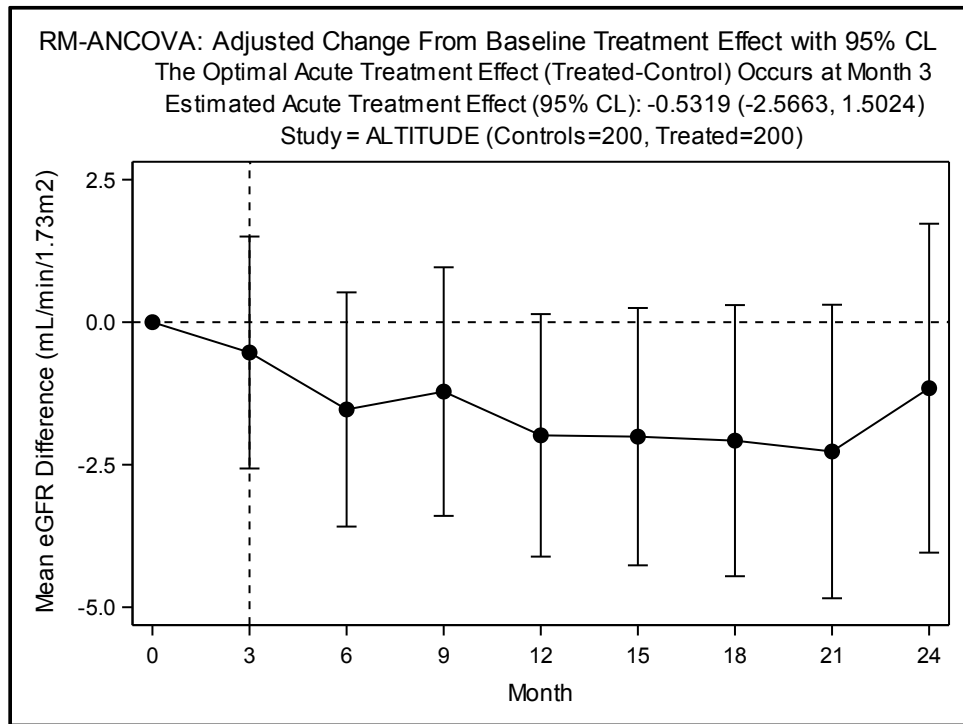
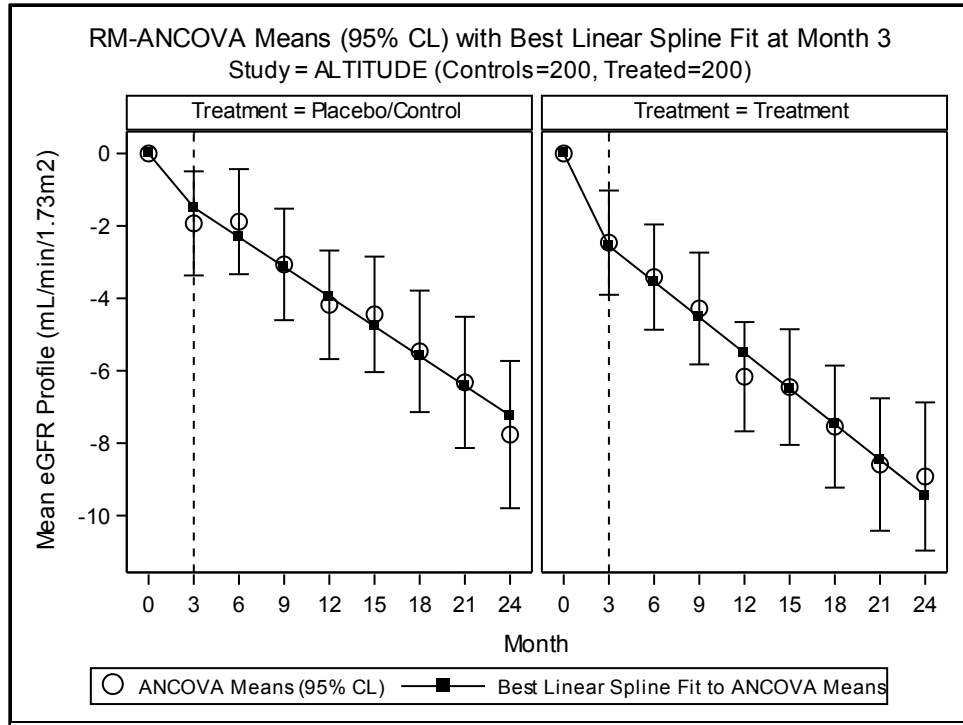


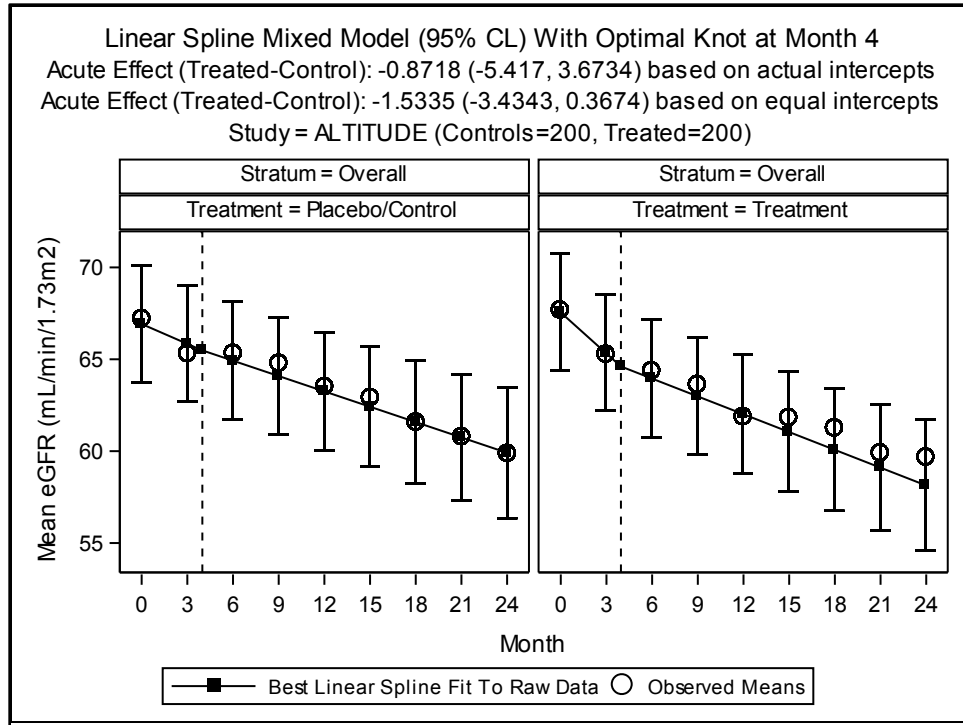


Appendix 1: ALTITUDE results
 Case 3: Sample Size=200 per Treatment group



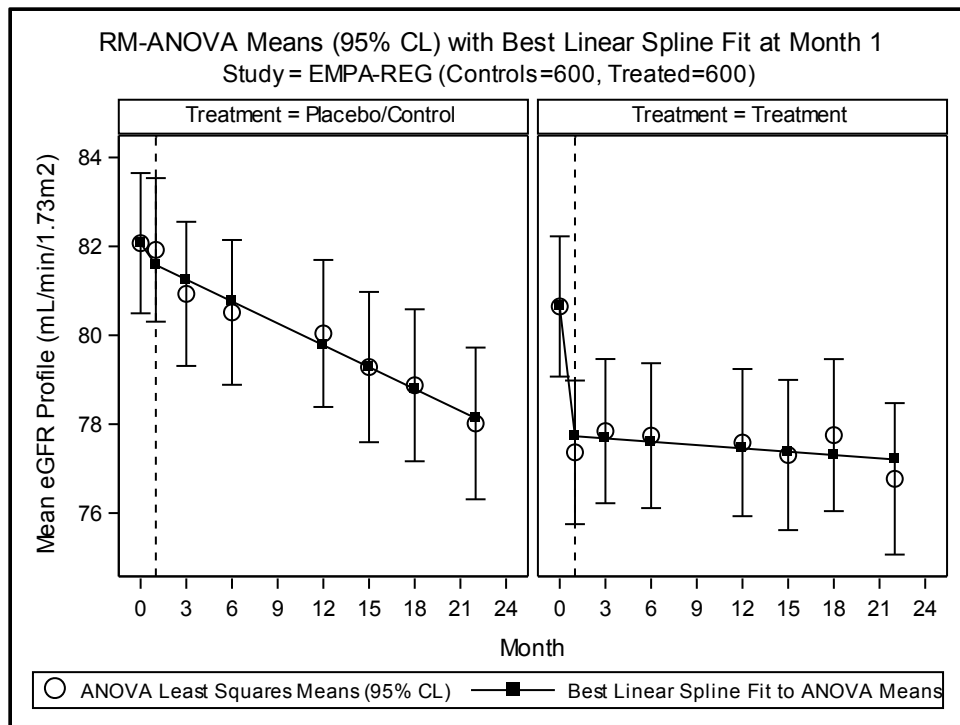
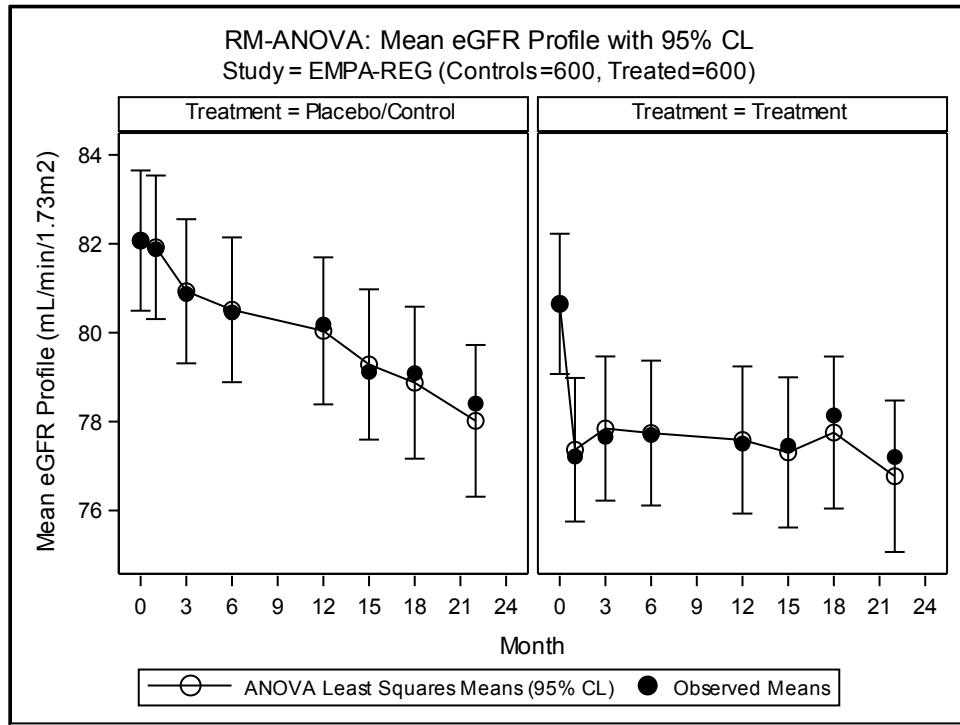


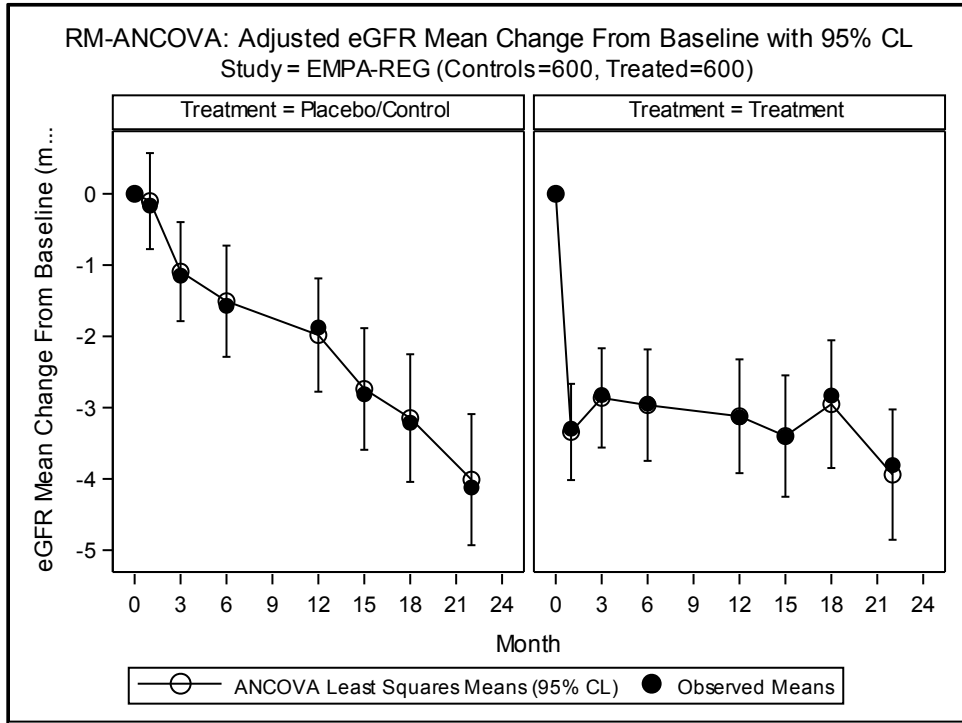
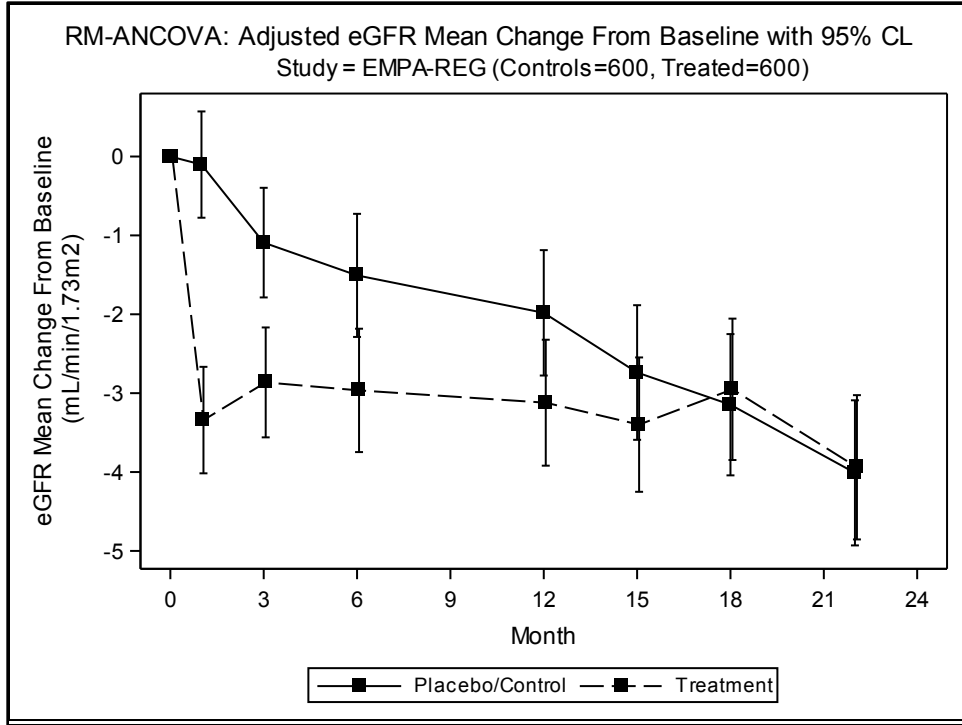


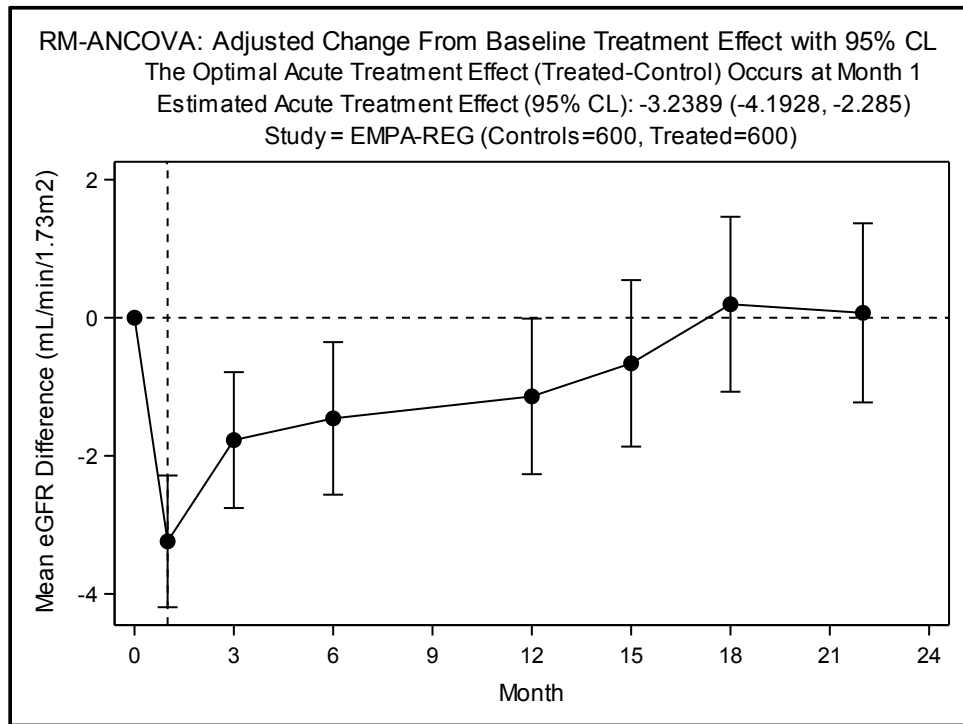
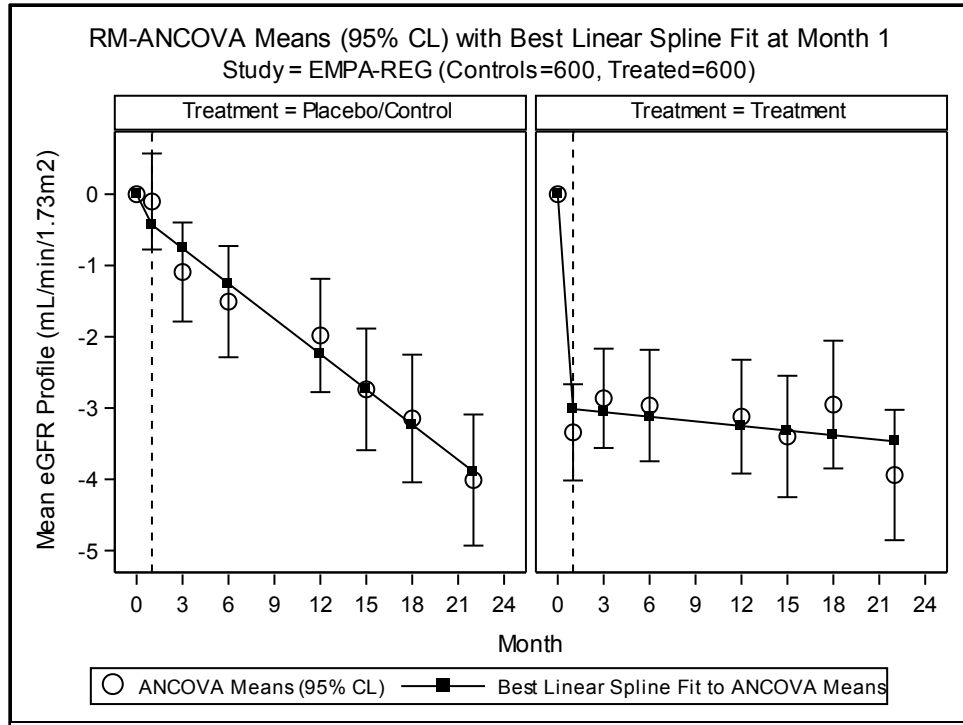


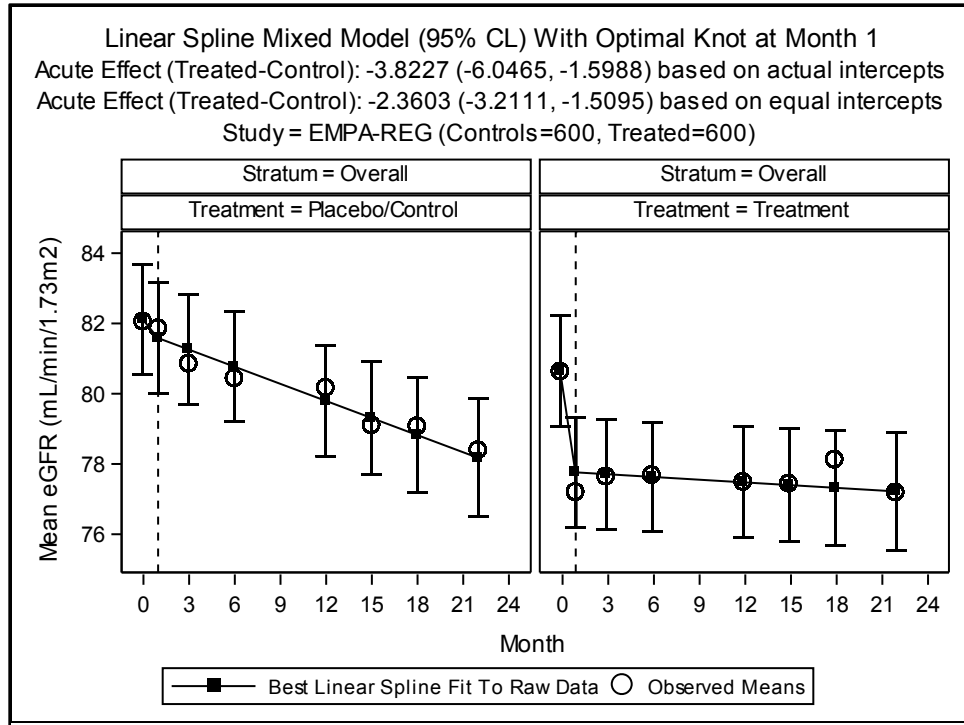
Appendix 2: EMPA-REG results

Case 1: Sample Size=600 per Treatment group



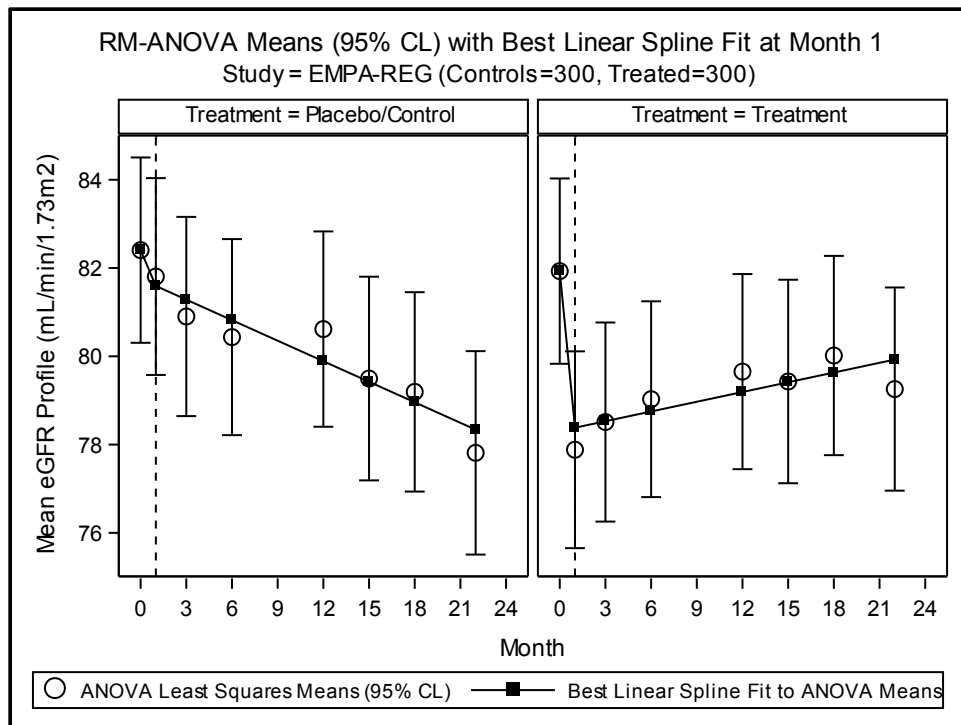
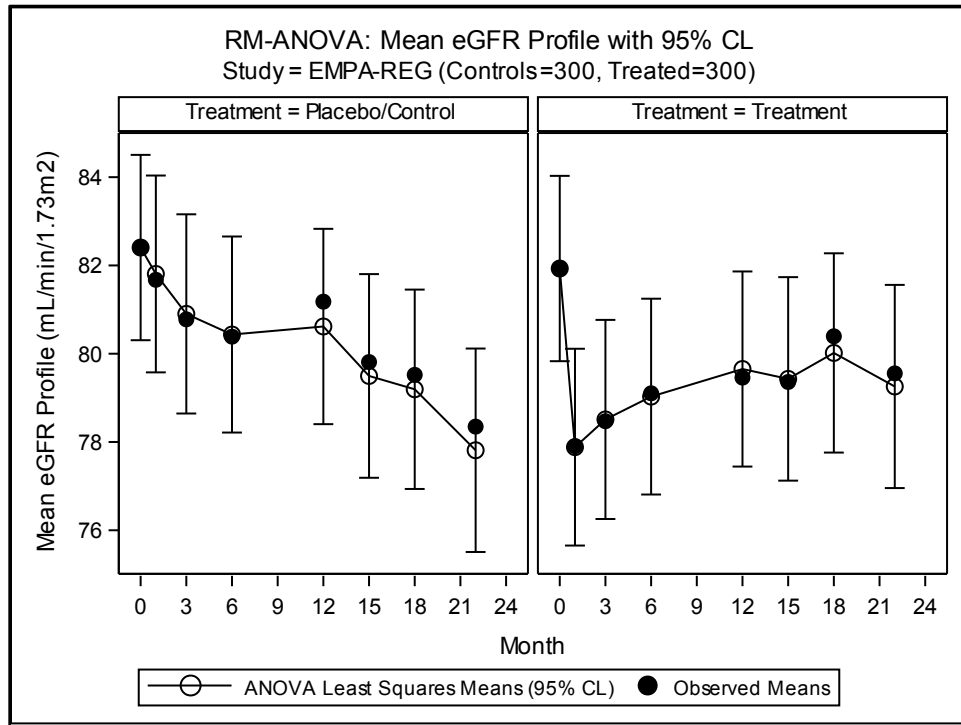


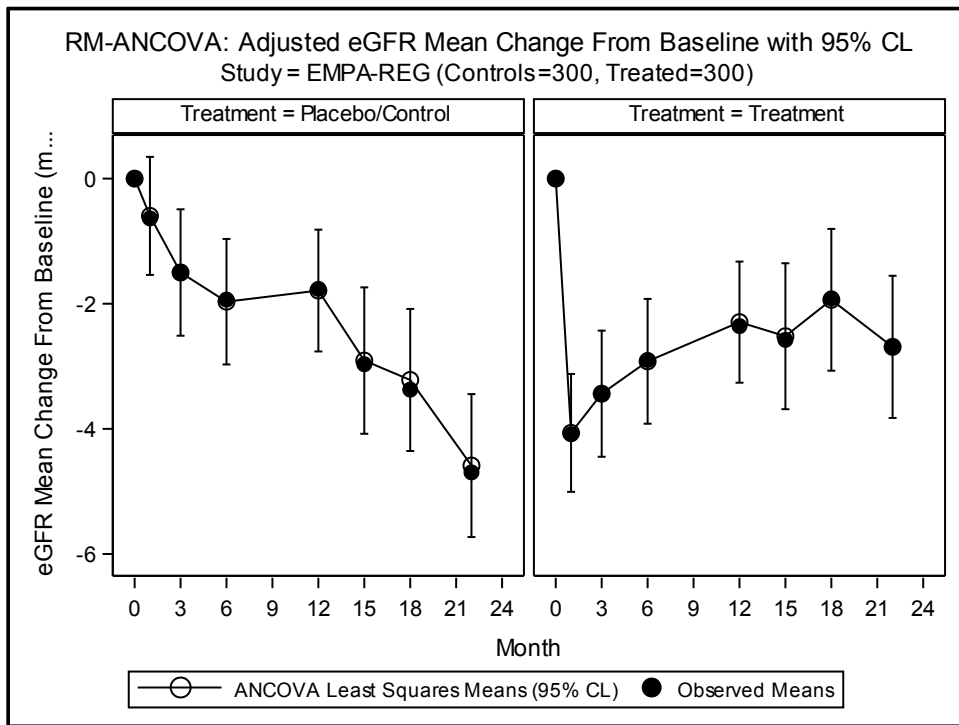
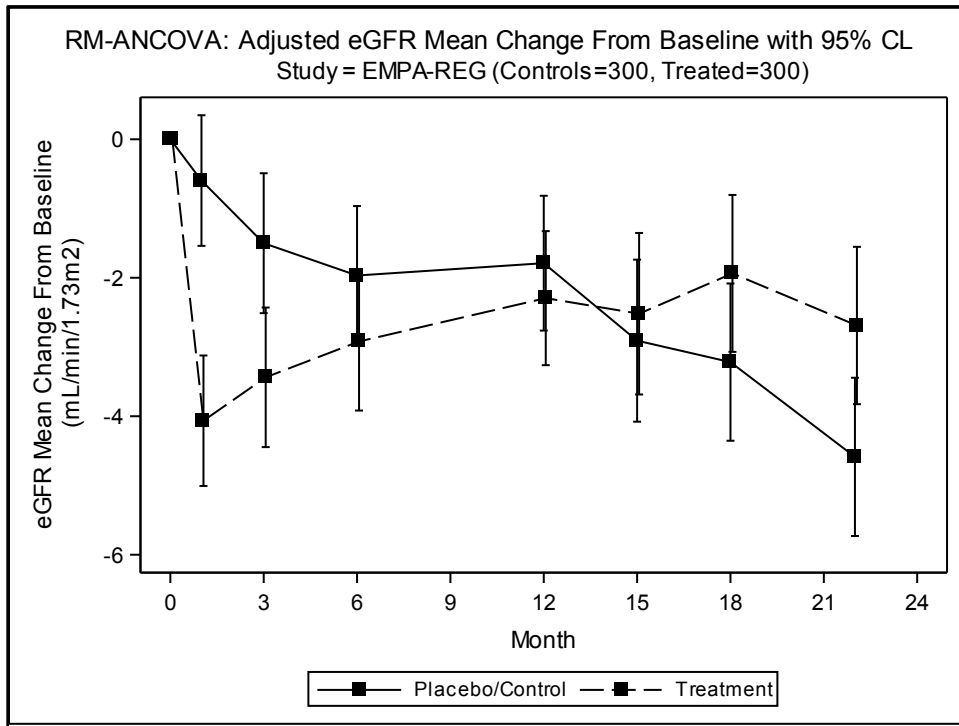


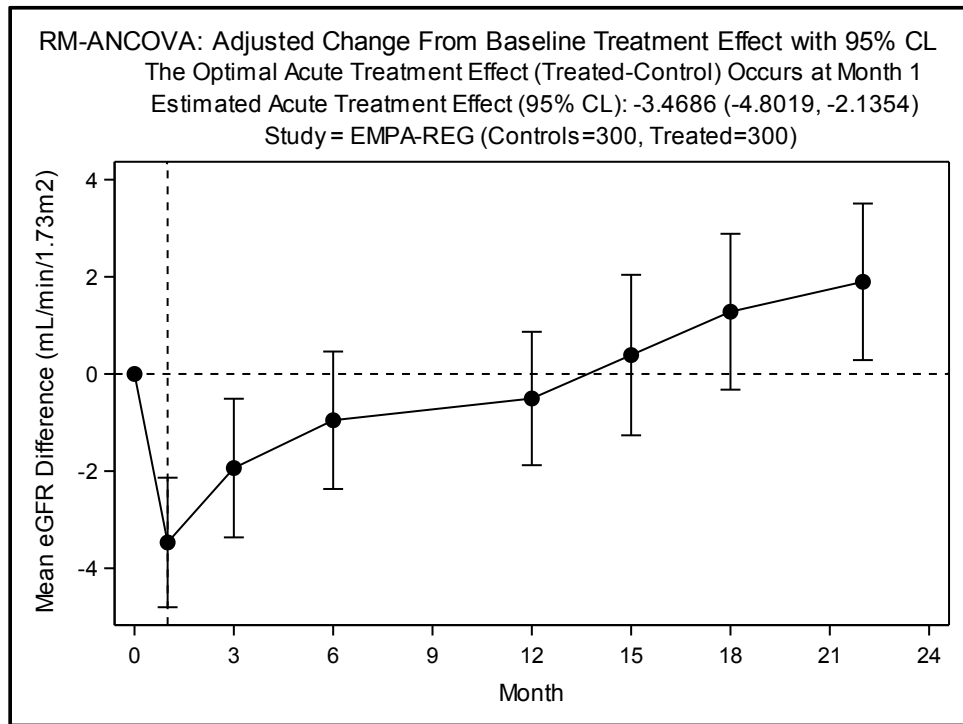
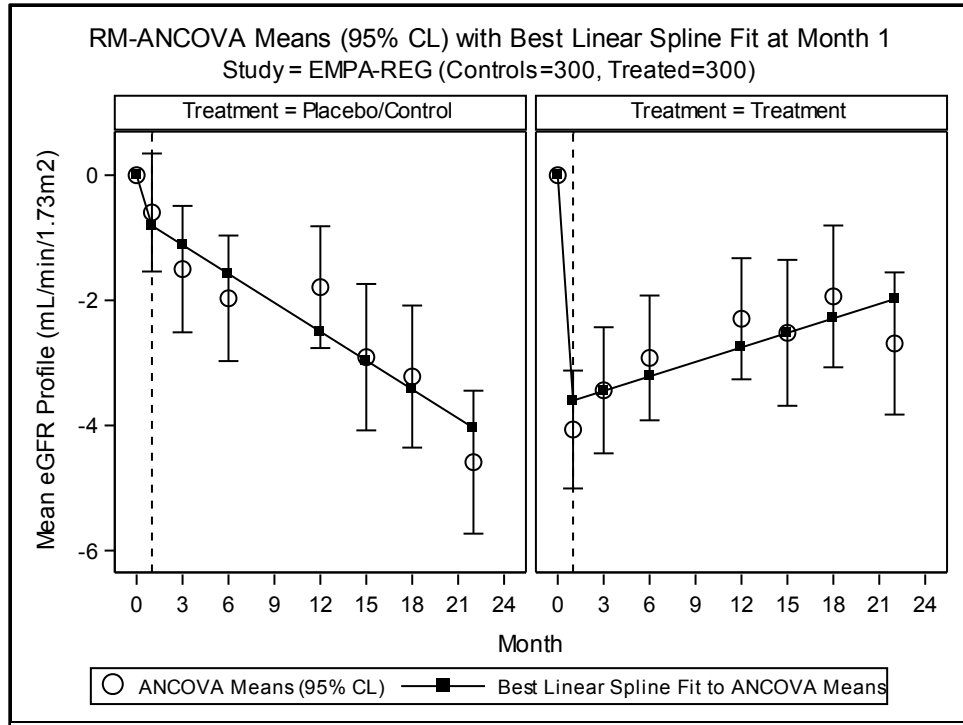


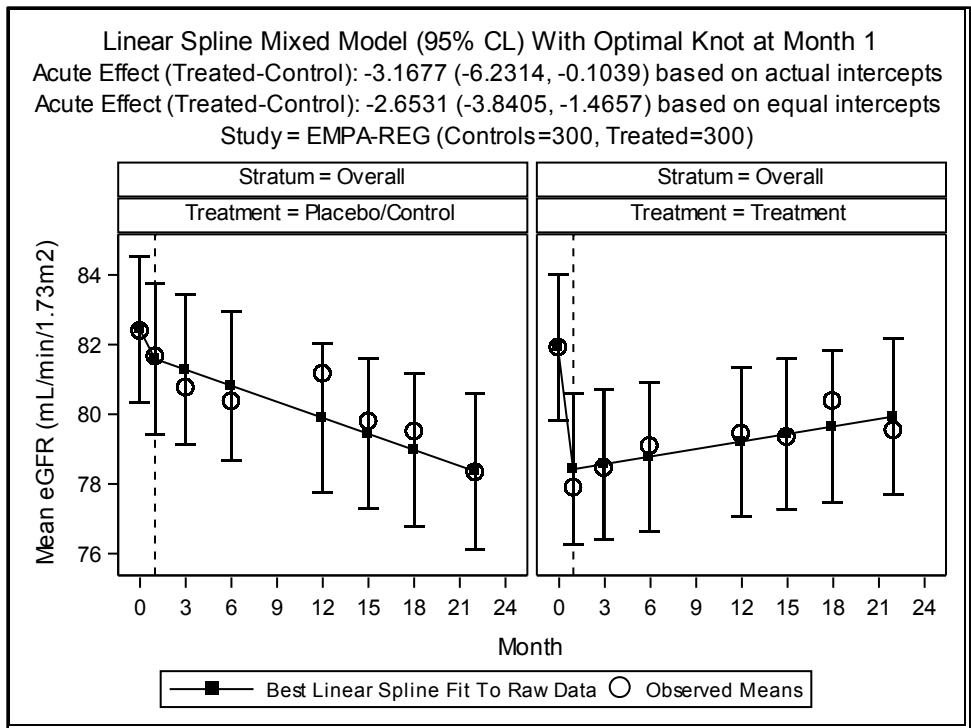
Appendix 2: EMPA-REG results

Case 2: Sample Size=300 per Treatment group

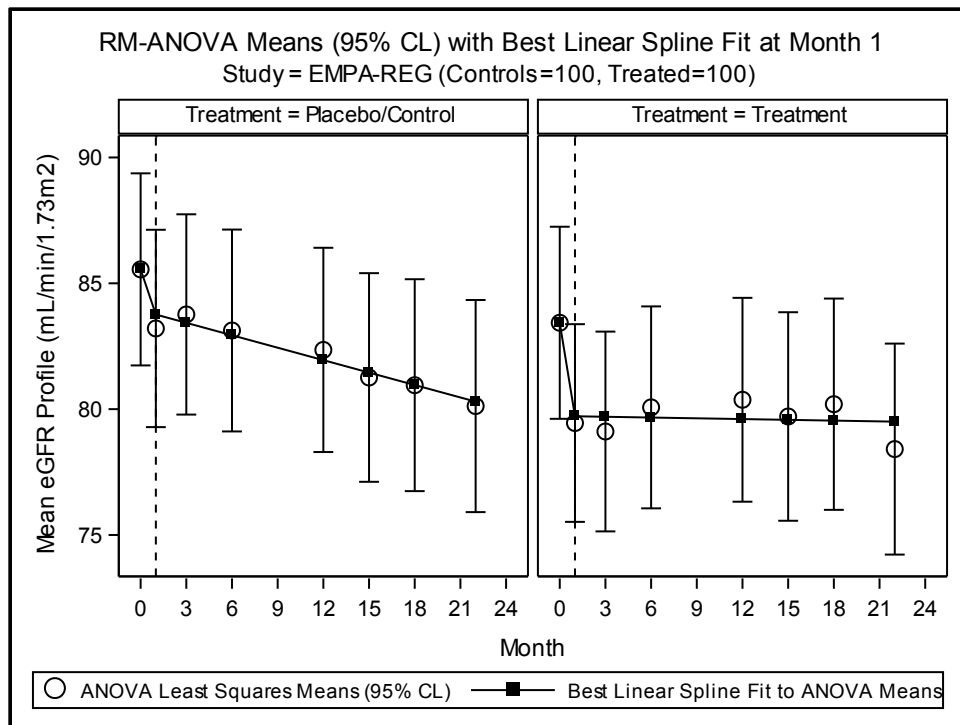
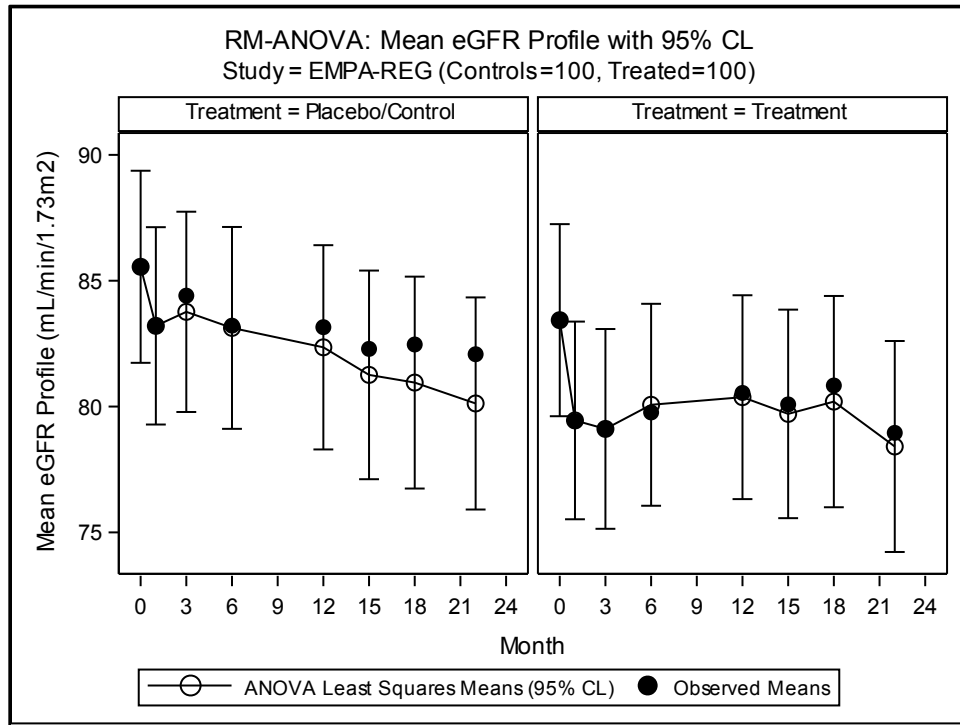


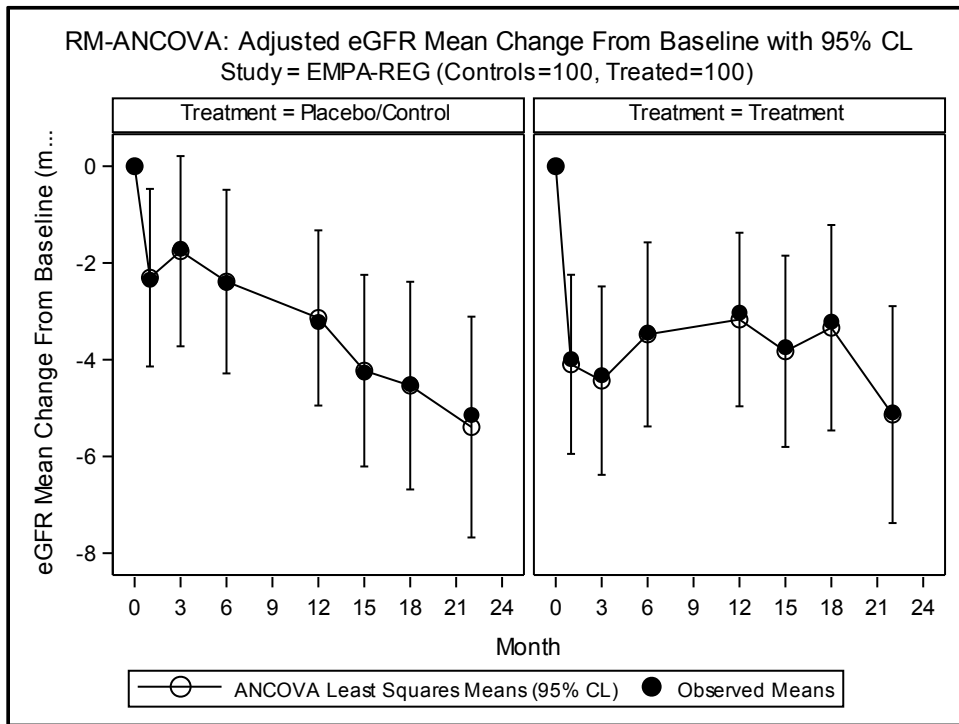
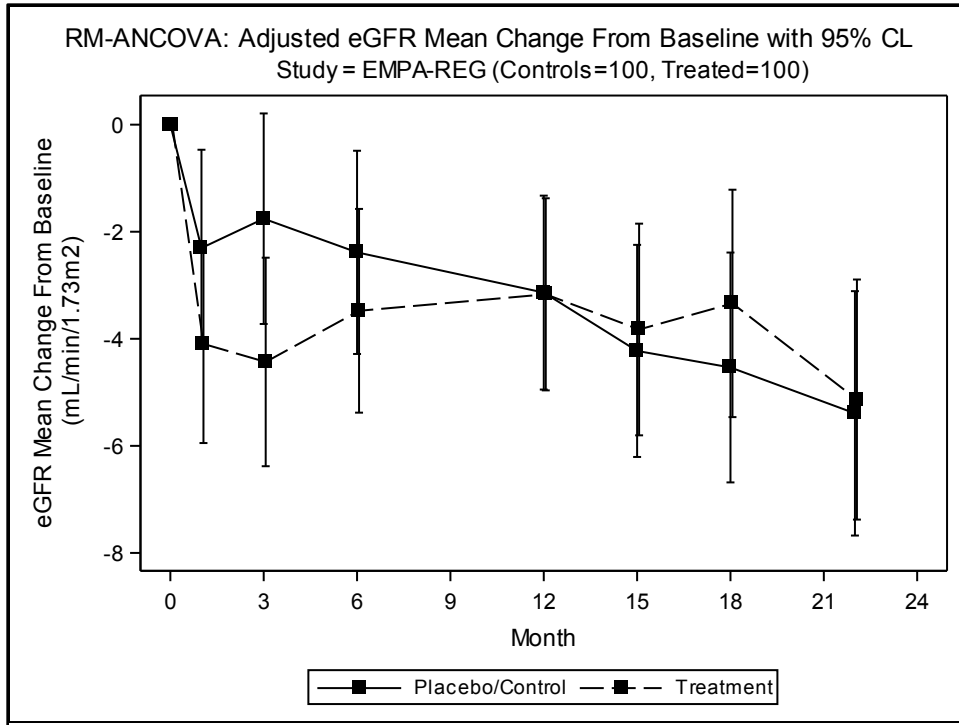


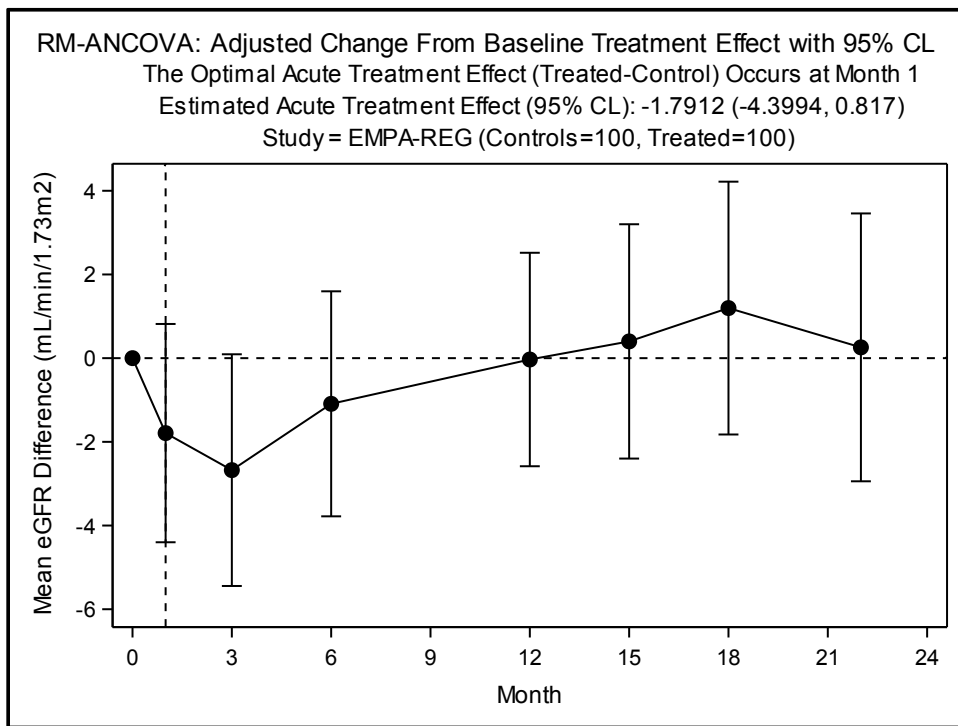
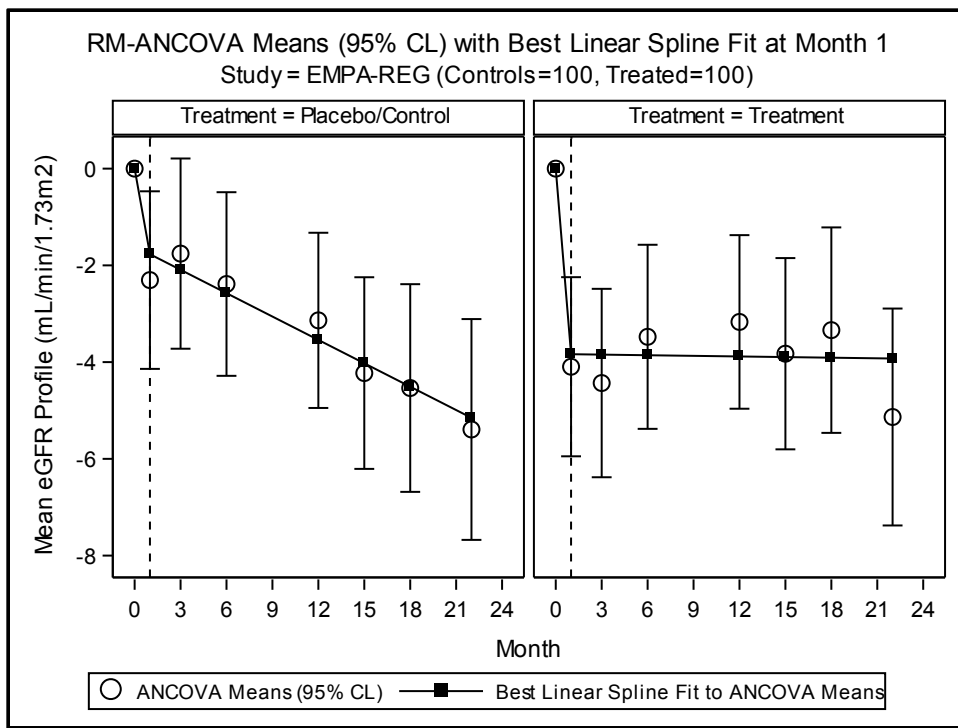


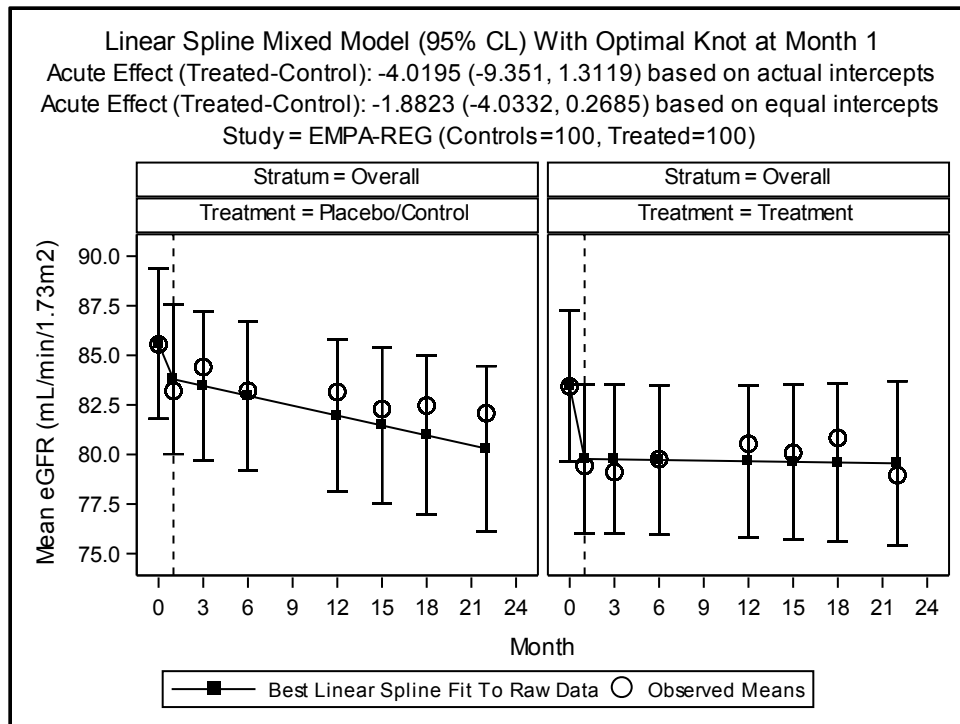


Appendix 2: EMPA-REG results
 Case 3: Sample Size=100 per Treatment group









Appendix 3: Zuchelli results

Sample Size=61 Control, 60 Treatment

