

Vonesh Statistical Consulting, LLC

Vonesh Statistical Consulting, LLC
1928 Forest Creek Lane
Libertyville, IL 60048

Phone: 847-367-1678,
Fax: 847-367-1802
Email: Edward@Vonesh-Stats.com

December 10, 2020

To: Lesley Inker, MD, MS
Tom Greene, PhD
Hocine Tighiouart, MS

From: Edward F. Vonesh, PhD

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

1. Introduction

This report is intended to summarize several methods one can use to determine and characterize the potential timing and magnitude of acute treatment effects for Phase 2 or Phase 3 clinical trials of chronic kidney disease (CKD). A total of nine (9) studies were selected to illustrate how one might determine if and when an acute treatment effect might occur when comparing a control group with an intervention treatment group. Table 1 lists the nine studies along with their corresponding 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
3	IDNT(CCB)	Placebo/Control	556
		Treatment	572
4	IDNT(CNTRL)	Placebo/Control	563
		Treatment	572
5	MDRD-A(BP)	Placebo/Control	285
		Treatment	299
6	MDRD-A(DIET)	Placebo/Control	293
		Treatment	291
7	MDRD-B(BP)	Placebo/Control	123
		Treatment	132
8	MDRD-B(DIET)	Placebo/Control	129
		Treatment	126
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 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 is provided Section 2 while Section 3 presents a summary of results using graphical representations.

2. Methods

Determination of the timing of an acute treatment effect can be based on one of two basic methods. One method would be to use a semi-parametric mean profile analysis in which predicted mean eGFR values over time are obtained using a repeated measures analysis of variance (RM-ANOVA) model or predicted mean change-from-baseline eGFR values over time are obtained using a repeated measures analysis of covariance (RM-ANCOVA) model with centered baseline eGFR values serving as the covariate. The timing of an acute treatment effect can be ascertained by comparing these mean profiles over time. Of the two modeling approaches, the RM-ANCOVA approach is generally preferred as it is statistically more efficient and powerful than the RM-ANOVA approach.

A second approach would be to fit a smooth parametric linear spline mixed-effects model with a knot established at an optimally determined point in time at which there is a transition from an acute eGFR trajectory phase (acute slope) to a chronic eGFR trajectory phase (chronic slope). Here we consider the two-stage linear spline mixed-effects model used by Vonesh et al. (2019) to obtain good starting values for a full likelihood-based linear spline mixed-effects model.

2.1. Repeated Measures Profile Analysis – A Two-Step Algorithm

In this section we describe a two-step approach for determining the likely timing of an acute treatment effect based on a repeated measures profile analysis using either a RM-ANOVA model or a RM-ANCOVA model. In step 1, one fits either a RM-ANOVA model or a RM-ANCOVA model from which predicted least squares means are obtained at the time points where eGFR measurements are taken. In step 2, one fits a simple linear spline model with a single fixed knot or change point to the least squares means from the RM-ANOVA or RM-ANCOVA model. By varying the knot at each of the designated visit times, one can determine an “optimal” knot by selecting the linear spline model whose fixed-point knot provides the best fit to the predicted least square means as determined by Akaike’s information criterion (AIC). This is done using a common knot for both the control group and treatment group assuming the mean response for the treatment group (or, in some cases, the control group) will be the predominant factor in ascertaining the timing and magnitude of an acute treatment effect if such an effect occurs.

Once an “optimal” timing (knot) of an acute effect has been determined, the acute treatment effect under the RM-ANCOVA model is computed simply as the difference in the least squares mean change from baseline values for the treatment group vs control group at that “optimal” knot.

2.2. Two-Stage Linear Spline Mixed-Effects Analysis

A parametric two-stage linear spline mixed-effects model with a single fixed knot is used to fit a smoothed linear response curve having an acute linear response phase followed by a chronic linear response phase. In the first stage, a standard linear spline mixed-effects model is fit to the observed data assuming homogeneous within-subject and between-subject variability. Subject-specific predicted means at each time point are computed based

on this initial model from which a power-of-the-mean variance function of these means can then be used as weights in a second-stage linear spline mixed-effects model in order to account for within-subject heteroscedasticity. As was done with the profile analysis, one can determine an optimal knot that best describes the timing of an acute effect by varying the knot at designated time points, fitting the linear spline model at each designated knot, and choosing that knot which provides the best fit among the linear spline mixed-effects models considered. This can be accomplished using one of two approaches. The first approach would be to vary the knot at each of the observed visit times and the second approach would be to vary the knot using 1 month increments even if eGFR was not measured at a given month. For each of these two approaches, one would compare the fit of the models and select the model whose fixed-point knot provides the best fit as determined by Akaike's information criterion (AIC). A common knot for both the control and treatment groups is assumed in order to ascertain the timing and magnitude of an acute treatment effect at a given point in time.

Once a fixed "optimal" knot for the two treatment groups has been determined, an estimate of the acute treatment effect would simply be the predicted mean difference between the treated versus control group at that fixed point in time. Such an estimate could take into account both the estimated intercepts and acute slopes for the two treatment groups. In this case, the resulting estimate would be an unbiased estimate of the acute treatment effect whenever the true but unknown population intercepts are unequal between the two groups. However, for randomized controlled trials, a better estimate might be one that assumes the population intercepts are in fact equal for the two groups. In that case, the two assumed equal population intercepts would cancel when predicting the expected acute treatment effect. As such, a better estimate under the assumption of equal population intercepts would be one that ignores the estimated intercepts in the construction of the predicted mean difference. Both of these approaches to estimating an acute treatment effect are presented here.

2.3. SAS program

The methods described in Sections 2.1-2.2 were implemented in a SAS program (SAS 9.4) using a series of SAS macros with options as follows:

- 1) **%ClearTempDirectories**; This macro simply clears all temporary datasets that were created from previous calls to the macros listed below. This should always be the first macro called when performing a specific set of analyses for a given study.
- 2) **%Study(StudyID= , Sample= , N1= , Stratum=)**; This macro identifies the study to be analyzed and creates a SAS dataset, **Study**, containing the required data to be analyzed for the given study. The macro contains options for a) restricting the sample size of each treatment group within the given study, b) using a pseudo-random sample of subjects from the given study and c) selecting whether the analysis is for the overall cohort of subjects from the given study or for cohorts of subjects stratified by what stage of CKD they are in at baseline. Below is a description of the options.
 - a) **StudyID=** This refers to the study ID number as listed in Table 1. For example, StudyID=1 refers to the ALTITUDE study
 - b) **Sample=** This refers to a select sample size for each treatment group. For example, Sample=600 would designate using 600 subjects per treatment group. This is useful for selecting sample sizes more in line with a Phase 2 trial. For example, the ALTITUDE study had over 4,000 subjects per treatment group whereas a Phase 2 trial would likely have only 200 or 300 subjects per treatment group.
 - c) **N1=** This defines which subject to start sampling from if one wishes to take a pseudo-random sample from within each treatment group. The value of N1 one chooses is based on a created variable, SampleID, which takes values from 1 to N within a given treatment group of size N as determined by the sorted value of new_id which is the assigned subject ID variable. This option is intended for use only for selecting a pseudo random starting point. Its default value of N1=1 should suffice in which case the macro will select the first N subjects. Setting N1=51, for example would take subject 51 through subject 51+N.

- d) **Stratum=** This macro variable defines whether subsequent analyses are to be carried out either on an overall cohort of subjects regardless of what stage of CKD subjects are in at baseline or if analyses are to be carried out on subjects stratified into one of four categories of CKD based on what stage of CKD they are in at baseline. The options are **Stratum=Overall** (this is the default value) and **Stratum=CKD**.

When **Stratum=Overall**, subsequent calls to the macros **%RM_ANOVA**, **%GetKnot** and **%SGplots** will produce results from the RM-ANOVA and RM-ANCOVA models (from **%RM_ANOVA**) as well as the linear spline-mixed-effects model (from **%GetKnot**) which are then plotted in a series of graphs with a call to the macro **%SGplots**. When **Stratum=CKD** is specified, the macro **%RM_ANOVA** is ignored (even if one calls the macro) while the macros **%GetKnot** and **%SGplots** will produce results from a linear spline-mixed-effects model that incorporates a stratified variable, CKD, that takes the values of 'CKD 1-2', 'CKD 3a', 'CKD 3b' or 'CKD 4-5' depending on what stage of CKD a subject is in based on the subject's baseline eGFR value. In this case, one gets separate regression parameter estimates under the stratified linear spline mixed-effects model according to CKD Stages 1-2, 3a, 3b and 4-5. This stratified analysis **ONLY** applies to the linear spline mixed-effects analysis and not to the RM-ANOVA or RM-ANCOVA analyses.

- 3) **%RM_ANOVA**; This macro runs both the RM-ANOVA model and RM-ANCOVA model for a given study. It then uses the RM-ANOVA and RM-ANCOVA predicted least squares means as observations and fits a sequence of linear spline models to these means using, in sequence, each non-zero observation time as a single fixed knot. The macro then identifies the knot that provides the best fit to the least squares means. This macro will run only when the macro variable **Stratum=Overall** is specified within the call to macro **%Study**.
- 4) **%GetKnot(Listing=, MaxKnots=, ObsKnots=)**; This macro fits the two-stage linear spline mixed-effects model to the observed eGFR values with options for a) displaying or not displaying the results of each call to PROC MIXED, b) setting the maximum time point to be considered as a potential knot and c) whether knots are selected based on observed visit times or 1-month incremental times. Below is a more thorough description of these options.
 - a) **Listing=** This instructs the macro program to either list (**Listing=**) or not list (**Listing=CLOSE**) all of the PROC MIXED calls. By default the option is **Listing=CLOSE** as otherwise a very large amount of SAS output is generated. It can be useful to run **Listing=** (i.e., blank) so as to see all the output but only when one wishes to look at how things work or as a check on why possible errors occur if and when they occur.
 - b) **MaxKnots=** This refers to the maximum knot among the possible knots one wishes to evaluate. For example **MaxKnots=12** means only values of knot ≤ 12 are considered. The default value is taken to be 12 assuming that if a significant or nearly significant acute effect occurs, it will happen within the first year of follow-up.
 - c) **ObsKnots=** This refers to whether one uses values of knot that only occur at the observed visit times or values of knot that range from 1 to **MaxKnots** in increments of 1 month. The two options are **ObsKnots=YES** or **ObsKnots=NO**. For example, if **MaxKnots=12** and one specifies **ObsKnots=YES** then only knots at the observed visit times occurring up to and including month 12 (if observed) are used to identify an optimal knot. If one specifies **ObsKnots=NO** then knots at visit times ranging from 1 month up to and including 12 months in 1 month increments are used to identify an optimal knot. The default is **ObsKnots=YES**.
- 5) **%SGplots**; This macro takes fitted values from the RM-ANOVA and RM-ANCOVA models and the linear spline mixed-effects model and provides a graphical display of how well the models fit the observed means and, in the case of the RM-ANOVA and RM-ANCOVA models, the least squares means. A reference vertical line identifying the month where the "best" fit knot (i.e., acute treatment effect) occurs is displayed for each model. In each case, the best fitting model is the one having the lowest AIC value among the models considered as described in Sections 2.1 and 2.2.

Below are two examples illustrating the sequence of how to run these SAS macros. The first example is for the ALTITUDE study for the overall cohort while the second example is for the IDNT(CNTRL) study both for the overall cohort and for subjects stratified by what stage of CKD they are in at baseline.

Example 1: The ALTITUDE study with a maximum of 600 subjects per treatment from the overall cohort.

```
%ClearTempDirectories;
%Study(StudyID=1, Sample=600, N1=1, Stratum=Overall);
%RM_ANOVA;
%GetKnot(Listing=close, MaxKnots=12, ObsKnots=YES);
%GetKnot(Listing=close, MaxKnots=12, ObsKnots=NO);
%SGplots;
```

Example 2-a: The IDNT(CNTRL) study with a maximum of 600 subjects per treatment from the overall cohort.

```
%ClearTempDirectories;
%Study(StudyID=4, Sample=600, N1=1, Stratum=Overall);
%RM_ANOVA;
%GetKnot(Listing=close, MaxKnots=12, ObsKnots=YES);
%GetKnot(Listing=close, MaxKnots=12, ObsKnots=NO);
%SGplots;
```

Example 2-b: The IDNT(CNTRL) study with a maximum of 600 subjects per treatment with subjects stratified according to what stage of CKD they are in at the start of follow-up.

```
%ClearTempDirectories;
%Study(StudyID=4, Sample=600, N1=1, Stratum=CKD);
%GetKnot(Listing=close, MaxKnots=12, ObsKnots=YES);
%SGplots;
```

3. Results

Detailed graphical results for the nine (9) studies are shown in Appendices 1-9 according to the Study ID's listed in Table 1. An initial call to the macros for each study was done restricting the maximum sample size per treatment group to 600 (a macro option). This was done so as to reflect the maximum sample size one might encounter in a Phase 2 trial. Of the nine studies, only two studies exceeded this threshold, the ALTITUDE study (Study ID 1) which had 8,150 subjects in total and the EMPA-REG study (Study ID 2) which had 6,936 subjects in total. The remaining studies had less than 600 subjects per treatment group with sample sizes ranging from just under 600 subjects per treatment group (the two IDNT studies) to 60 and 61 subjects per treatment (the Zuchelli study). For each study, a total of 8 graphs depicting the potential timing and size of an acute treatment effect are shown based on the overall cohort of subjects. To illustrate what impact sample size has on the ability to detect an acute treatment effect, analyses for the ALTITUDE study were repeated with sample sizes restricted to 100, 200 and 300 subjects per treatment group (see Appendix 1 and Table 2). Also, for the two IDNT studies (Study ID's 3 and 4), two additional graphs are presented depicting the timing of an acute treatment effect across four groups of subjects stratified according to what stage of CKD they are in at baseline (see Appendices 3 and 4).

3.1 General Results Across Studies

Estimates of the “optimal” timing of an acute treatment effect (i.e., “optimal” knot) based on the two-step profile analysis applied to both the RM-ANOVA and RM-ANCOVA models were the same within each study but did vary across studies. Likewise, estimates of the “optimal” knot determined from a linear spline mixed-effects model based only on months where eGFR was observed (up to 12 months) versus any month (up to 12 months) were similar if not the same within each study. In all but one case, that being the ALTITUDE study with sample size=300 per group, the difference between the two “optimal” knots from the linear spline mixed-effects model were within 1 to 2 months of each other. While similar within studies, these “optimal” knots also varied across studies.

3.2 Specific Results by Study

Table 2 summarizes the estimated timing and magnitude of an acute treatment effect for each of the 9 studies with “optimal” knots determined based solely on the months where eGFR was measured. For most studies there was a notable difference between the “optimal” knot selected on the basis of the two-step RM-ANCOVA approach versus that based on the linear spline-mixed-effects approach with the former resulting in an equal or earlier timing of an acute treatment effect in all but two cases. For the three studies where the timing of an acute treatment effect was the same (the EMPA-REG and two IDNT studies), there was fair agreement between the two approaches. The fact that the timing was the same and magnitude of the acute effects was similar for these three studies may simply reflect the greater power these studies had over those studies with smaller sample sizes. In general, the RM-ANCOVA model and the linear spline-mixed-effects model with assumed equal population intercepts resulted in more precise estimates (narrower confidence intervals) of the acute treatment effects compared with the linear spline-mixed-effects model utilizing actual estimated intercepts.

Table 2. Summary of the optimal timing and magnitude of acute treatment effects (Treated-Control) by study with the optimal timing (knot) determined on the basis of those months where eGFR measurements were taken.

	ANCOVA Model		Linear Spline Mixed-Effects Model		
Study N per Treatment	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) Equal Intercepts (mL/min/1.73m ²)
ALTITUDE N=(600,600) N=(300,300) N=(200,200) N=(100,100)	Month 3 6 6 21	-1.48 (-2.51, -0.44) -2.14 (-3.76, -0.53) -2.62 (-4.79, -0.45) -3.35 (-7.66, 0.96)	Month 6 15 12 12	-0.83 (-3.11, 1.45) 0.93 (-2.16, 4.02) -0.14 (-4.15, 3.88) -5.83 (-11.65, -0.02)	-1.50 (-2.53, -0.47) -0.80 (-2.41, 0.82) -1.54 (-3.62, 0.53) -2.97 (-5.99, 0.06)
EMPA-REG N=(600,600)	1	-3.29 (-4.27, -2.31)	1	-2.20 (-4.44, 0.04)	-2.86 (-3.73, -2.00)
IDNT(CCB) N=(556,572)	3	-1.29 (-2.51, -0.06)	3	-1.49 (-3.70, 0.73)	-0.53 (-1.49, 0.43)
IDNT(CNTRL) N=(563,572)	3	-1.27 (-2.50, -0.05)	3	-1.50 (-3.77, 0.77)	-0.65 (-1.62, 0.31)
MDRD-A(BP) N=(285,299)	8	0.08 (-1.00, 1.15)	12	0.81 (-1.26, 2.88)	0.49 (-0.65, 1.62)
MDRD-A(DIET) N=(293,291)	8	1.72 (0.65, 2.79)	12	3.81 (1.76, 5.86)	1.33 (0.20, 2.46)
MDRD-B(BP) N=(123,132)	20	0.47 (-1.28, 2.21)	8	0.62 (-1.23, 2.46)	0.76 (-0.43, 1.94)
MDRD-B(DIET) N=(129,126)	4	2.03 (1.05, 3.02)	8	2.62 (0.80, 4.43)	1.46 (0.29, 2.63)
Zuchelli N=(61,60)	6	0.49 (-2.23, 3.21)	8	0.25 (-4.12, 4.61)	0.24 (-1.96, 2.44)

* The ANCOVA “optimal” knots are the same “optimal” knots obtained when using the ANOVA model.

When examining the individual studies, both graphically and on the basis of Table 2, the presence of a significant negative acute treatment effect is most evident in the ALTITUDE, EMPA-REG, and the two IDNT studies while a significant positive acute treatment effect was present in both the MDRD-A(DIET) and MDRD-B(DIET) studies for the treated group versus the placebo/control group. There is little evidence of any acute treatment effect for either the MDRD-A(BP) or MDRD-B(BP) studies. Interestingly, the Zuchelli study suggests there is a neutral “acute” phase for both the control group and treated groups with little or no change in the mean eGFR profile over a period of approximately 6-8 months after which both groups show a rate of decline in eGFR that is similar for the two treatment groups. The net result for the Zuchelli study was that there was no evidence of any differential acute or chronic treatment effect between the two groups over time.

4. Discussion

There are notable differences between “optimal” knots selected on the basis of the RM-ANOVA/RM-ANCOVA approach versus the “optimal” knots selected on the basis of the linear spline-mixed-effects approach. With the RM-ANOVA/RM-ANCOVA approach, the “optimal” knot is based on fitting a linear spline model to the predicted least squares means from either the RM-ANOVA or RM-ANCOVA model and selecting the knot that provides the best fit. The issue one must consider with this approach is that the least squares means are the predicted (imputed) means that one would expect if the data were in fact balanced and complete which, of course, is not the case. This is nicely illustrated with the two IDNT studies as shown in Appendices 3 and 4. Here we see that observed means for the overall group show good agreement with the least squares means through the first 12 months of follow-up after which there is a clear and significant departure between the observed means and the predicted least squares means. This suggests that the separation between observed and predicted means may be due in part to patient dropout in the later stages of follow-up. As evidence supporting this, the stratified analyses according to what stage of CKD patients are in at baseline show that the observed means for subjects in Stages 1-2 and Stage 3a track much more closely with the least squares means as compared with subjects in Stage 3b and Stages 4-5 where dropout due to death or ESKD is much more likely to occur.

In contrast, the linear spline mixed-effects approach has the advantage of selecting the “optimal” knot based directly on the observed data taking into account variation within- and between-subjects and by directly modeling acute and chronic trajectories over time. If dropout is ignorable, then this approach should be optimal assuming the linear spline model is the correct parametric model over a two-year period. In this case, the predicted mean profiles are unbiased while the observed means merely reflect the degree to which random dropout, related possibly to lower observed eGFR values, invokes departure from the model-based predicted means. However, when dropout is informative or non-ignorable, then this approach will introduce some degree of bias in the estimated slopes which in turn would manifest itself in some degree of bias in the mean profile over time.

Finally results from the ALTITUDE study suggests that sample size has a significant impact on the ability to accurately determine the timing and magnitude of an acute treatment effect as seen in Table 2 and Appendix 1. Under the null hypothesis that treatment group intercepts are equal, power calculations for detecting an acute treatment effect between two treatment groups will be equivalent to power calculations for detecting differences in the acute slopes between two treatment groups. This was illustrated with power calculations based on results from the IDNT(CNTRL) group and published in Web Appendix A of the supplemental material of Vonesh et al. (2019). In that analysis, the power to detect an acute treatment effect at 4 months of 0.20 mL/min/1.73m² (or, equivalently, a difference in the acute slopes of -2.42 mL/min/1.73m²/year) ranged from 0.205 to 0.362 for sample sizes ranging from 300 to 600 per treatment group (Table A.3 of Web Appendix A). It should be noted that although the acute slopes may not differ significantly, they can nonetheless exert a significant impact on the chronic and total slopes depending on the magnitude of change that occurs following the acute phase. On that basis, the timing of a suspected acute effect, even one that is marginal in magnitude, may be more important in a Phase 2 trial than attempting to power the Phase 2 trial to detect an acute treatment effect.

5. Conclusions

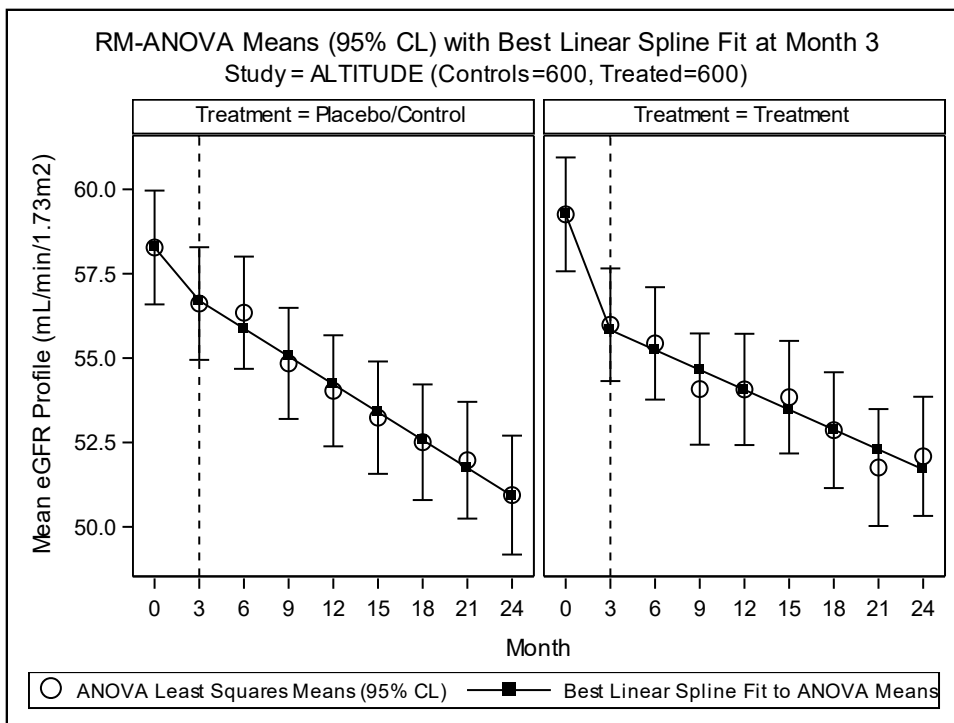
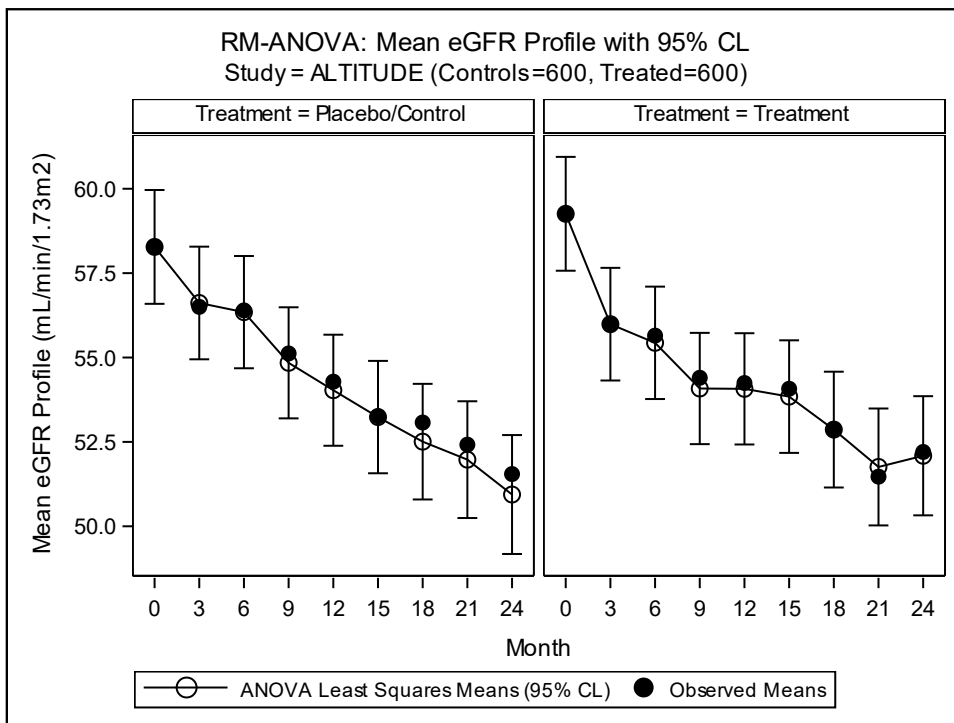
The analyses presented here illustrate just how difficult it can be to identify the timing and magnitude of an acute treatment effect for CKD trials. In weighing the advantages and disadvantages of the repeated measures profile analysis approach versus the linear spline mixed-effects modeling approach one must consider a number of other factors which this exercise fails to do. Most notable among such factors would be clinical input into how the

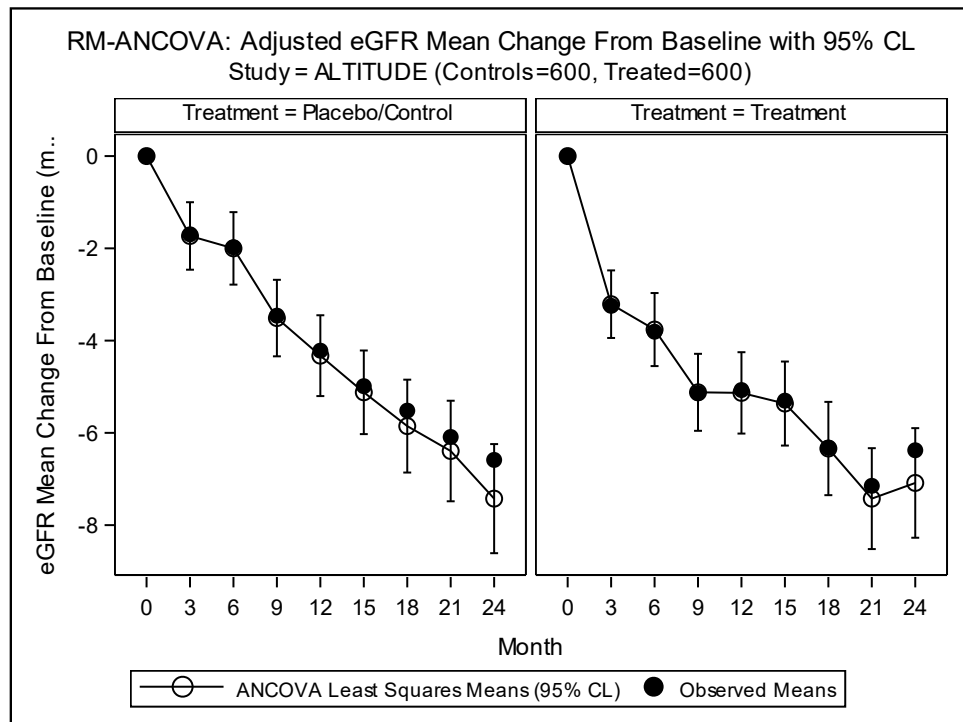
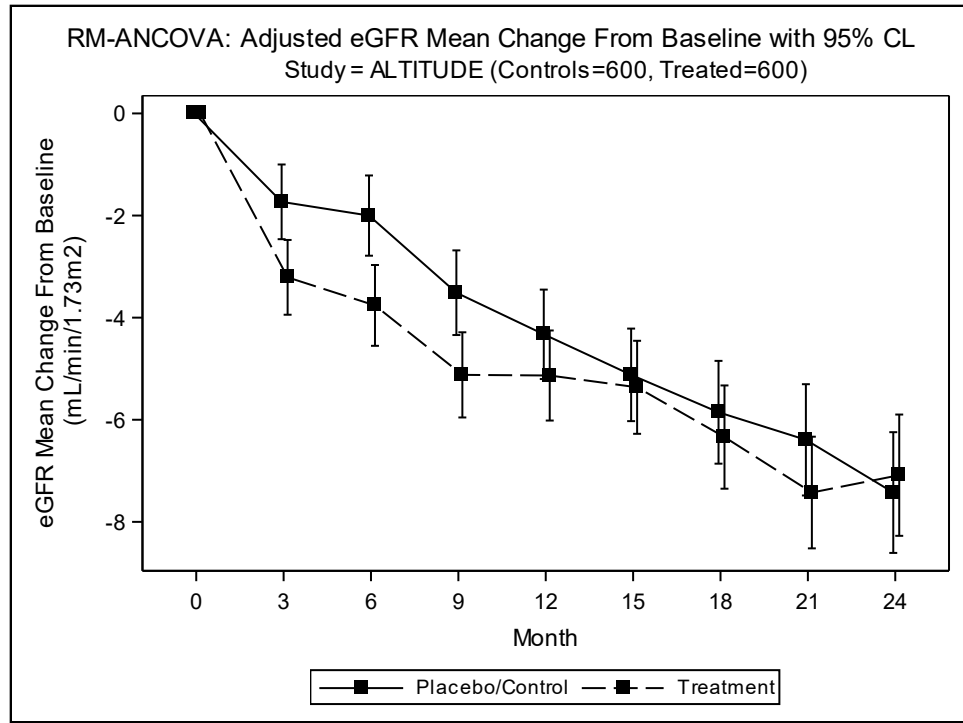
treatment intervention (i.e., biological mechanism) is expected to manifest itself over time. In the absence of such knowledge, if we assume an acute effect is most likely to occur within say the first 12 months, then the simplicity and semi-parametric approach of the RM-ANOVA or, preferably, RM-ANCOVA approach may prove to be less prone to error when estimating the timing of an acute effect.

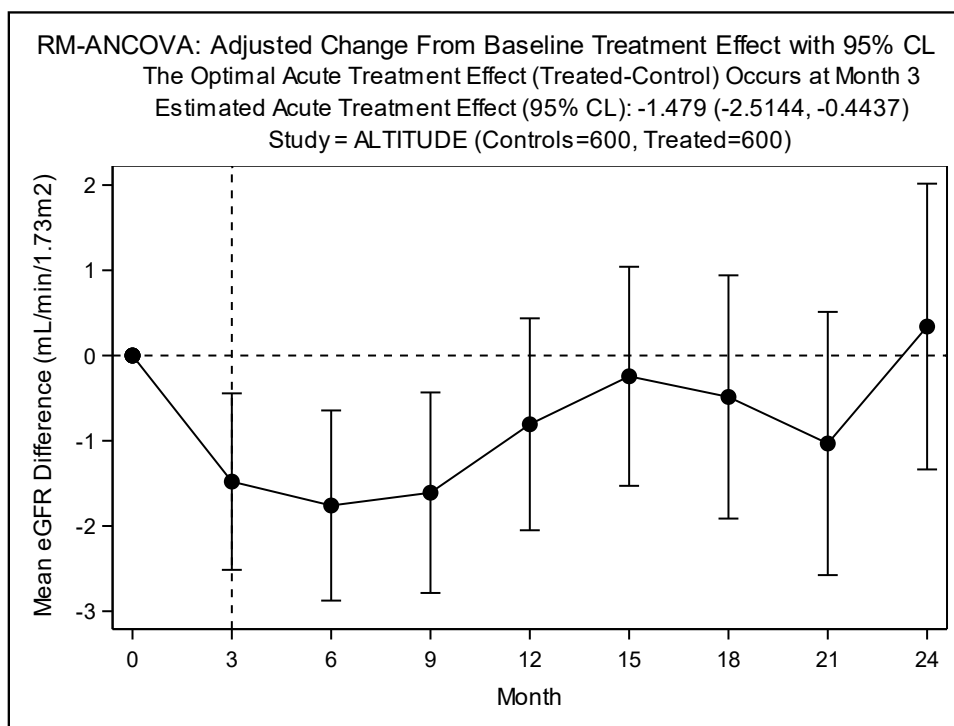
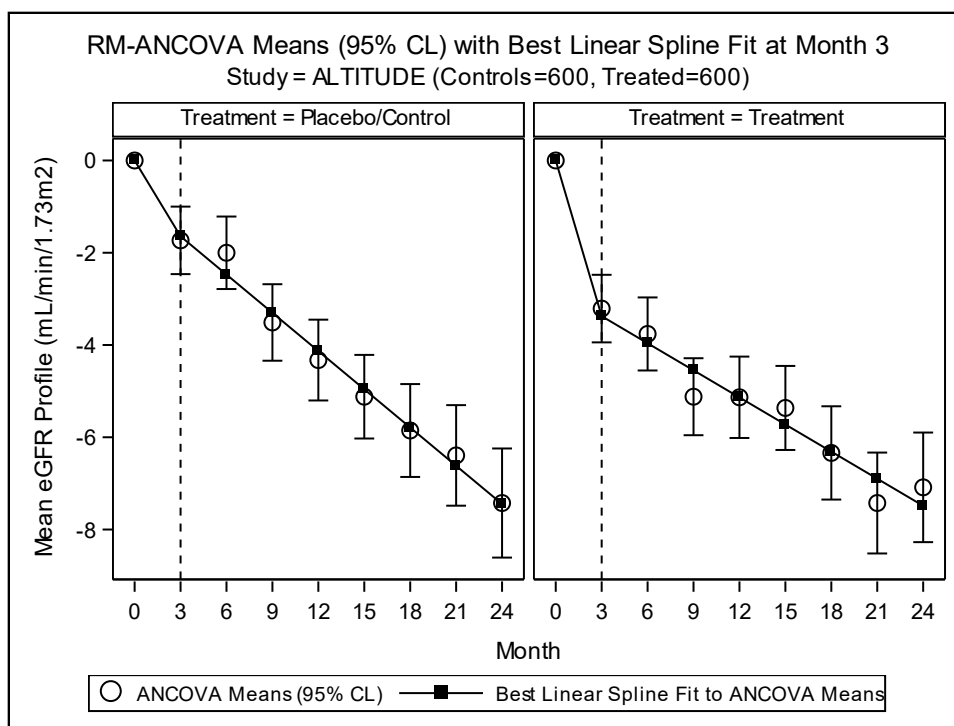
References

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.

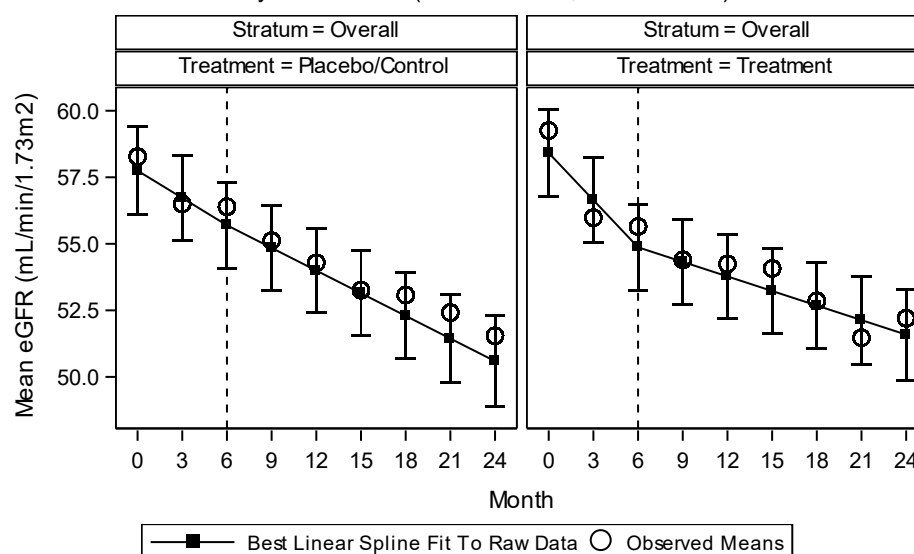
Appendix 1: ALTITUDE results
Case 1: Sample Size=600 per Treatment group



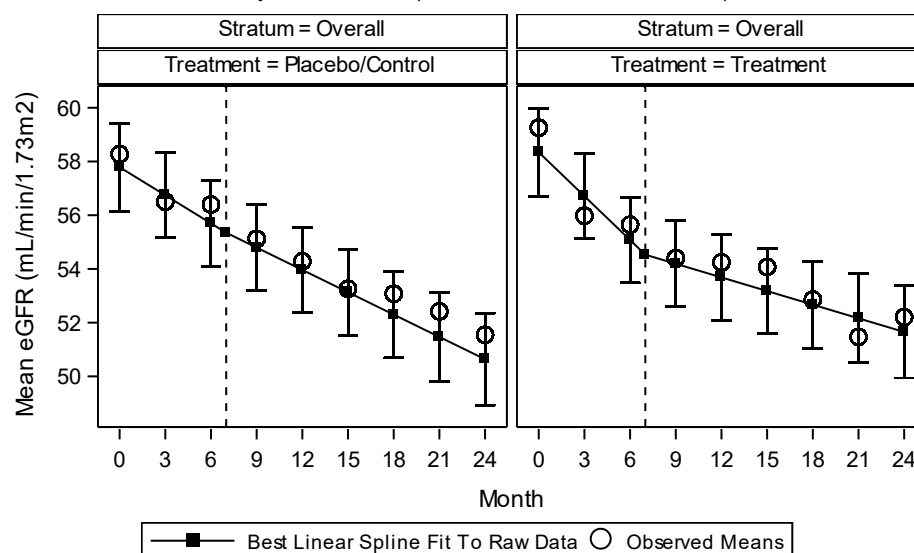




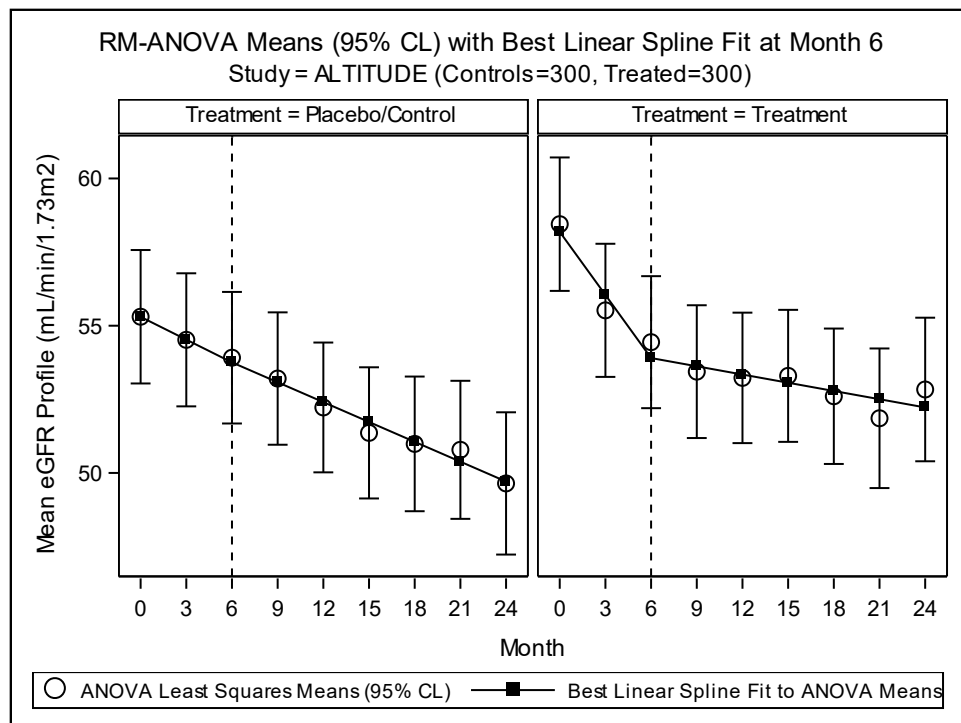
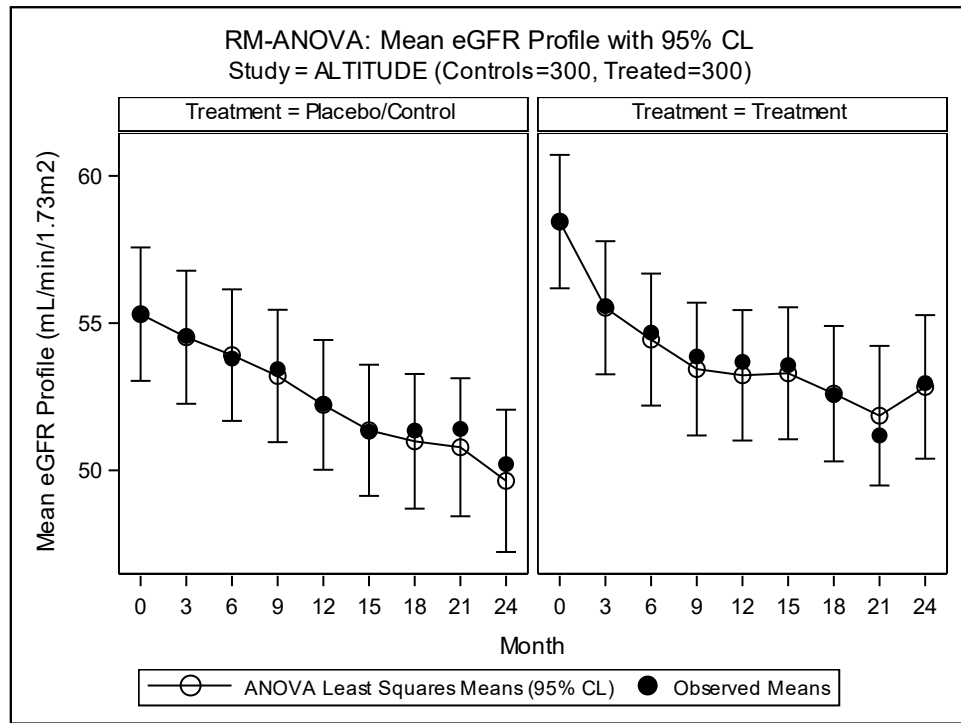
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 6
 Acute Effect (Treated-Control): -0.829 (-3.1073, 1.4492) based on actual intercepts
 Acute Effect (Treated-Control): -1.4966 (-2.5251, -0.4681) based on equal intercepts
 Study = ALTITUDE (Controls=600, Treated=600)

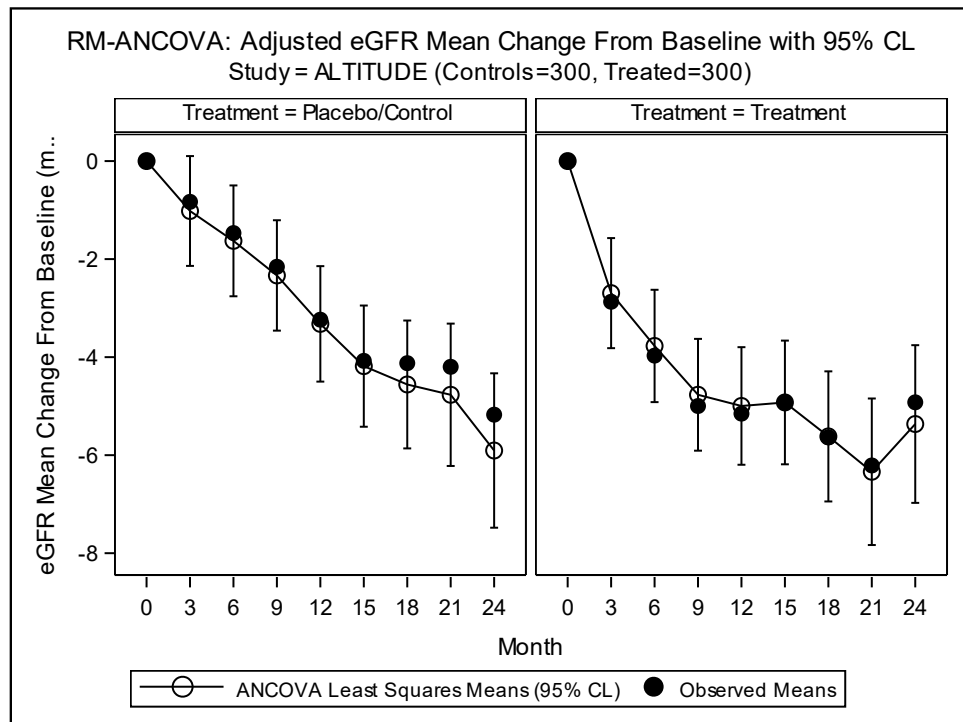
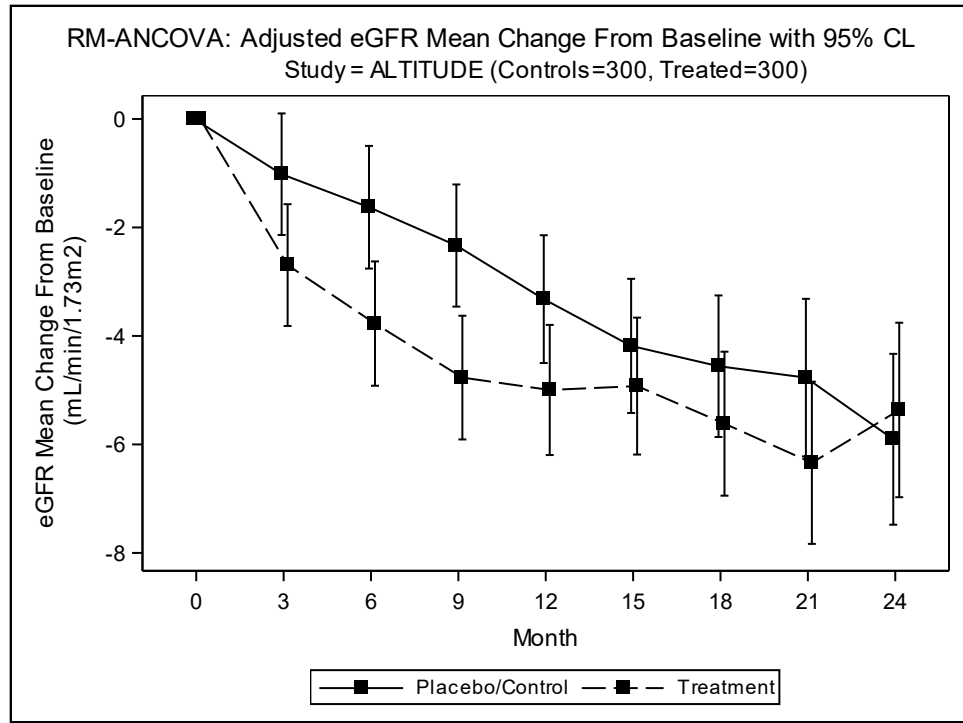


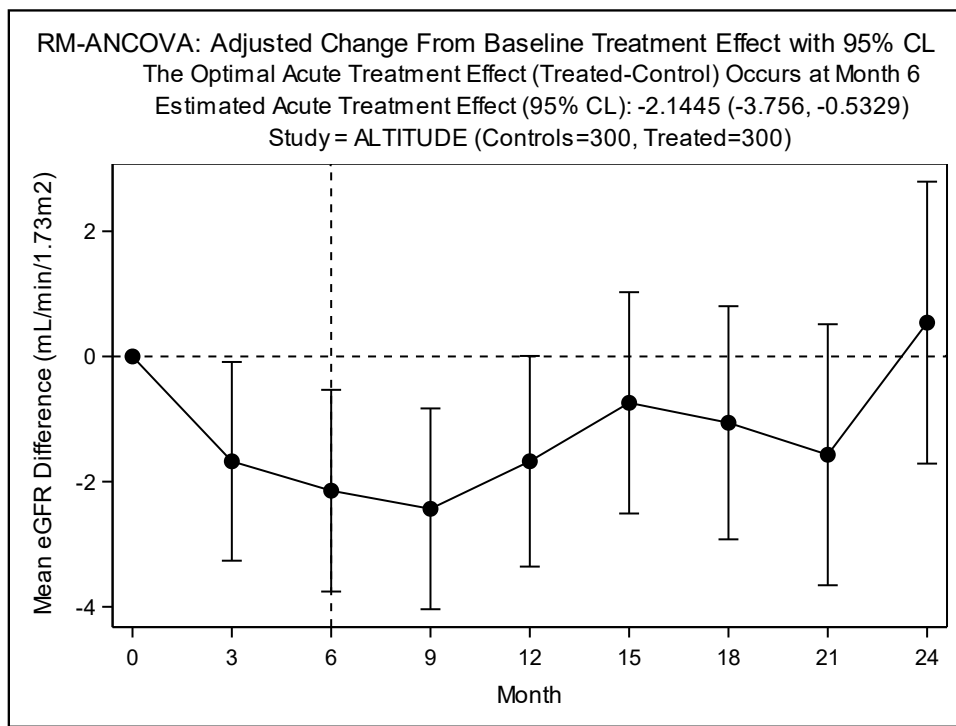
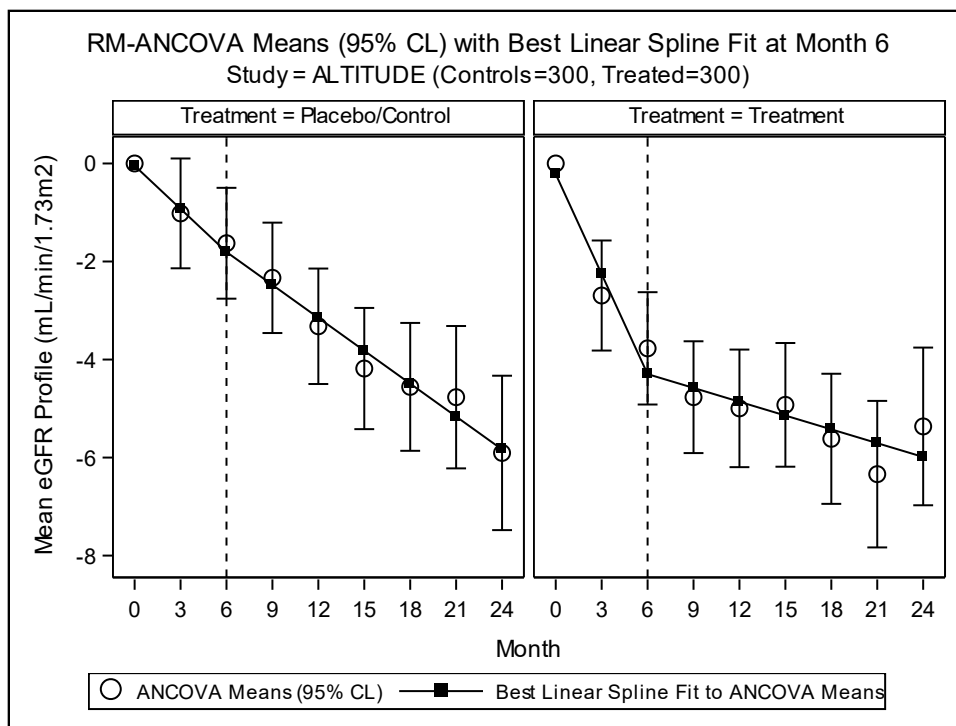
Linear Spline Mixed-Effects Model (95% CL) With Incremental Knot at Month 7
 Acute Effect (Treated-Control): -0.8181 (-3.1015, 1.4654) based on actual intercepts
 Acute Effect (Treated-Control): -1.3744 (-2.4575, -0.2913) based on equal intercepts
 Study = ALTITUDE (Controls=600, Treated=600)



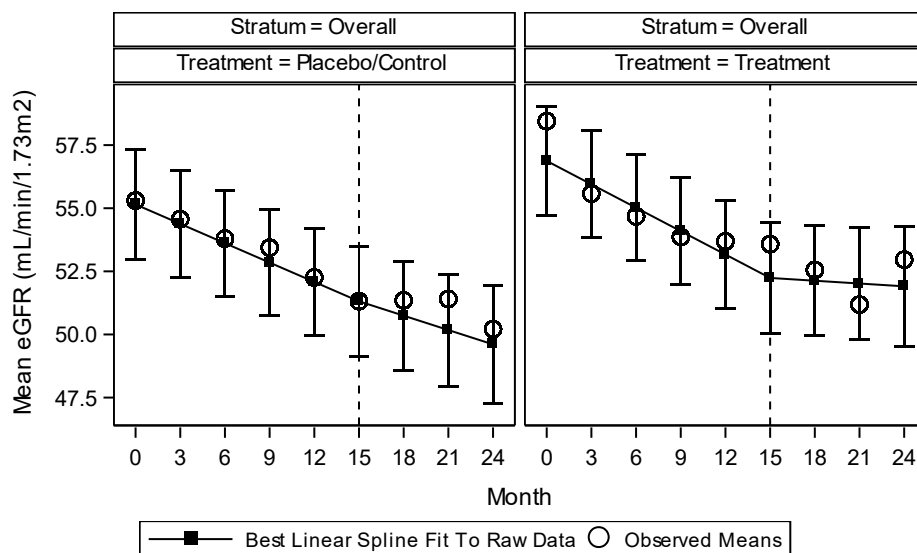
Case 2: Sample Size=300 per Treatment group



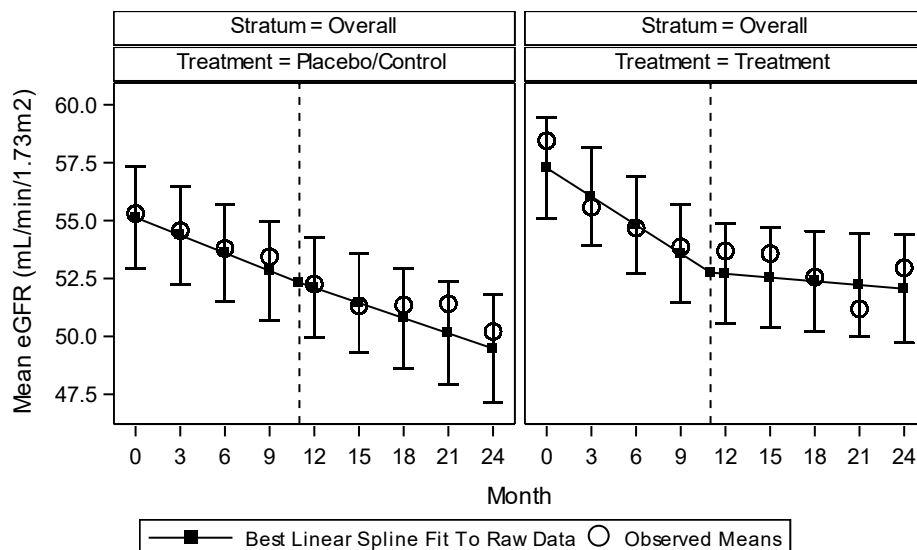




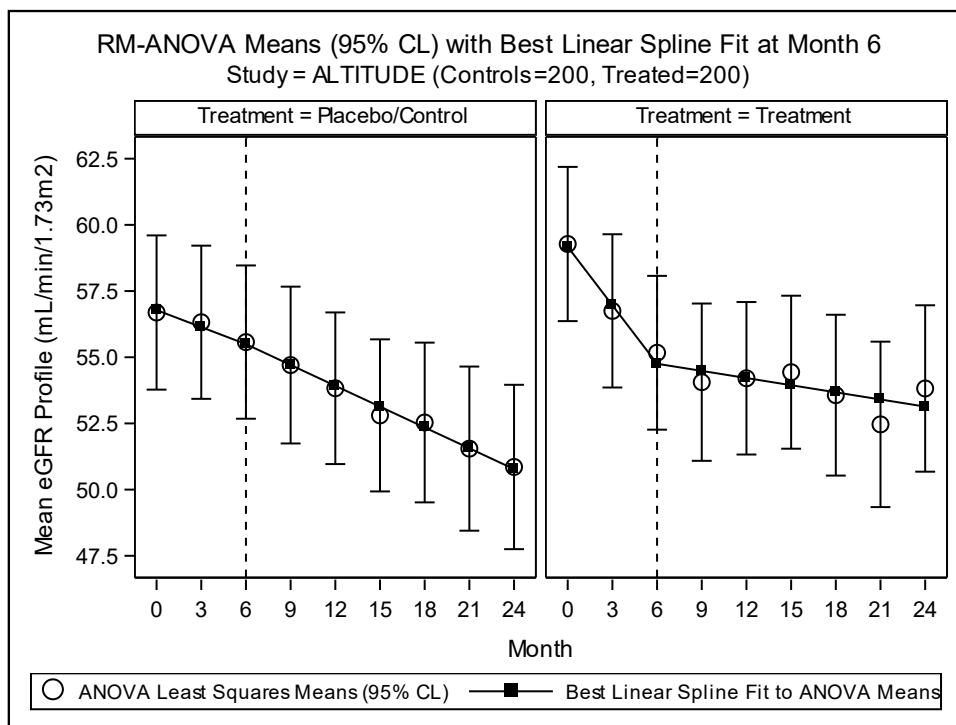
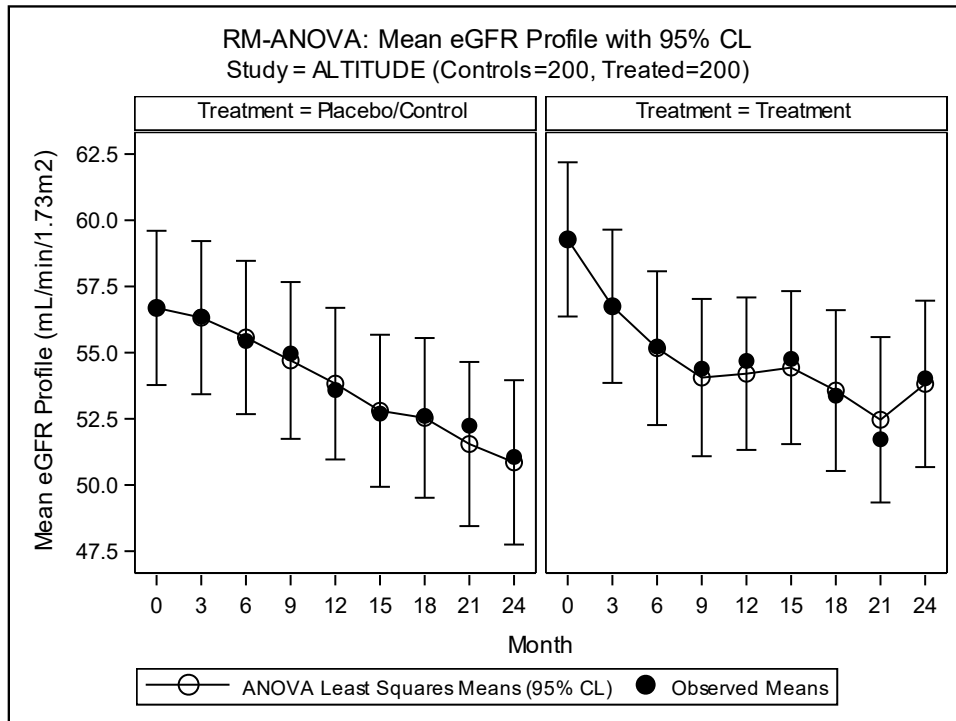
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 15
 Acute Effect (Treated-Control): 0.9268 (-2.1626, 4.0162) based on actual intercepts
 Acute Effect (Treated-Control): -0.7975 (-2.414, 0.8189) based on equal intercepts
 Study = ALTITUDE (Controls=300, Treated=300)

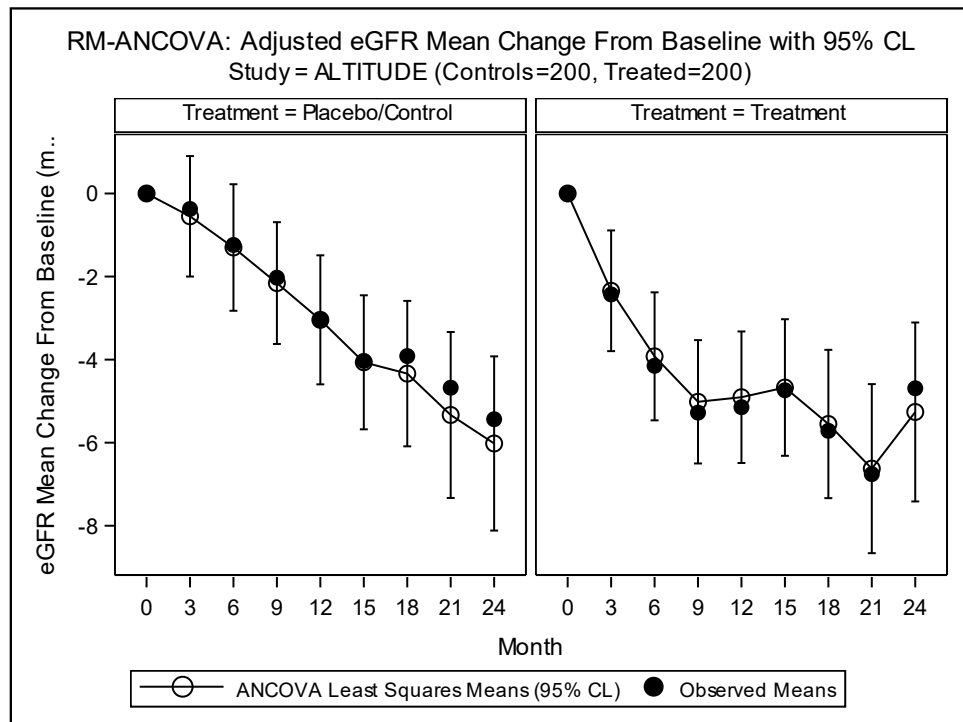
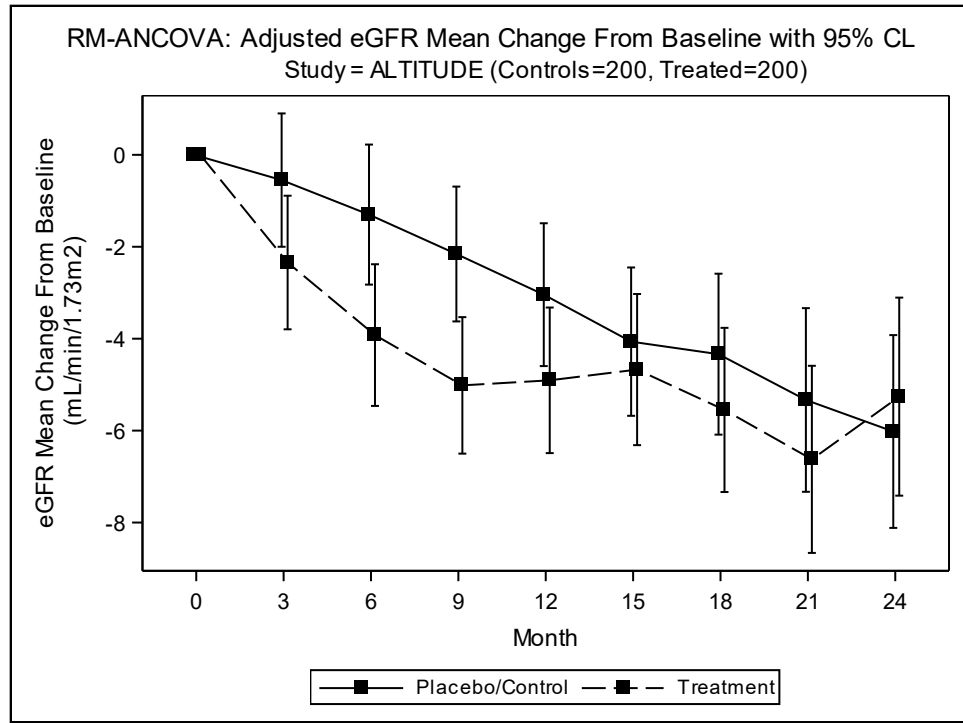


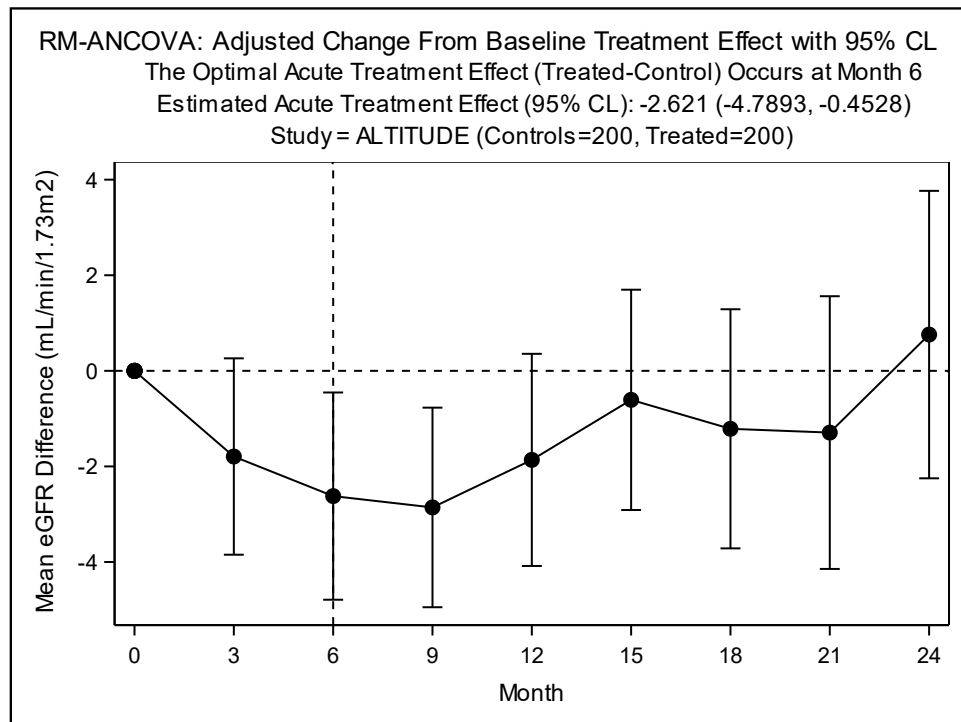
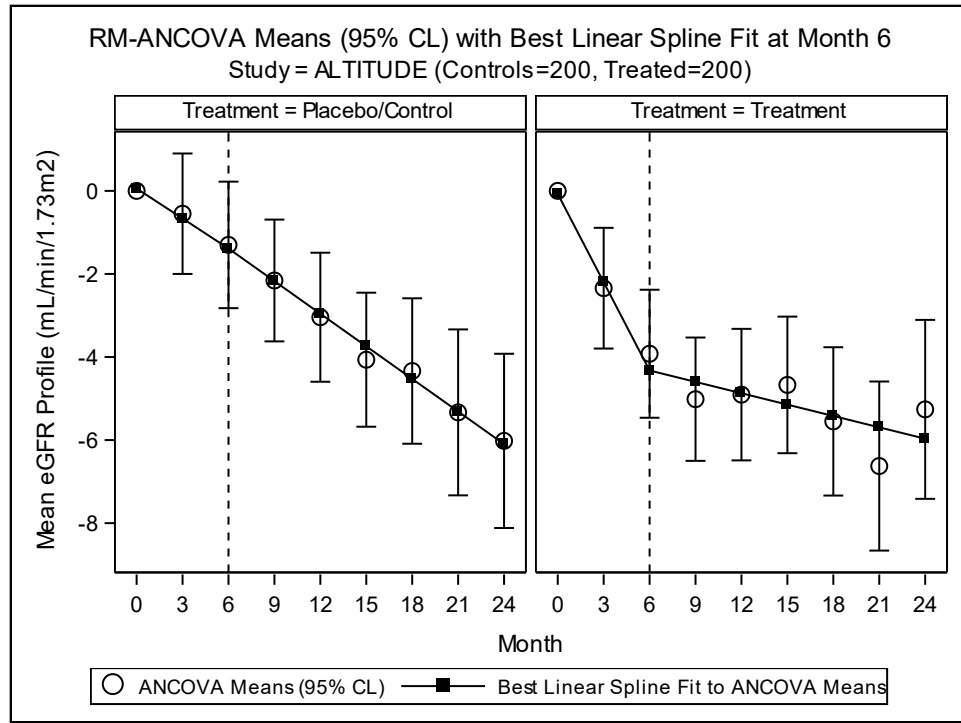
Linear Spline Mixed Model (95% CL) With Incremental Knot at Month 11
 Acute Effect (Treated-Control): 0.4366 (-2.6349, 3.5081) based on actual intercepts
 Acute Effect (Treated-Control): -1.6998 (-3.2758, -0.1238) based on equal intercepts
 Study = ALTITUDE (Controls=300, Treated=300)



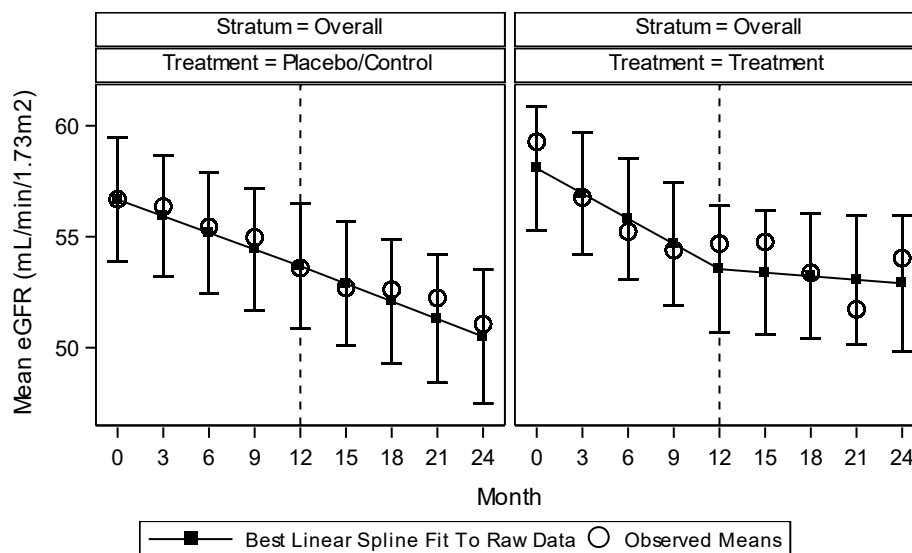
Case 3: Sample Size=200 per Treatment group



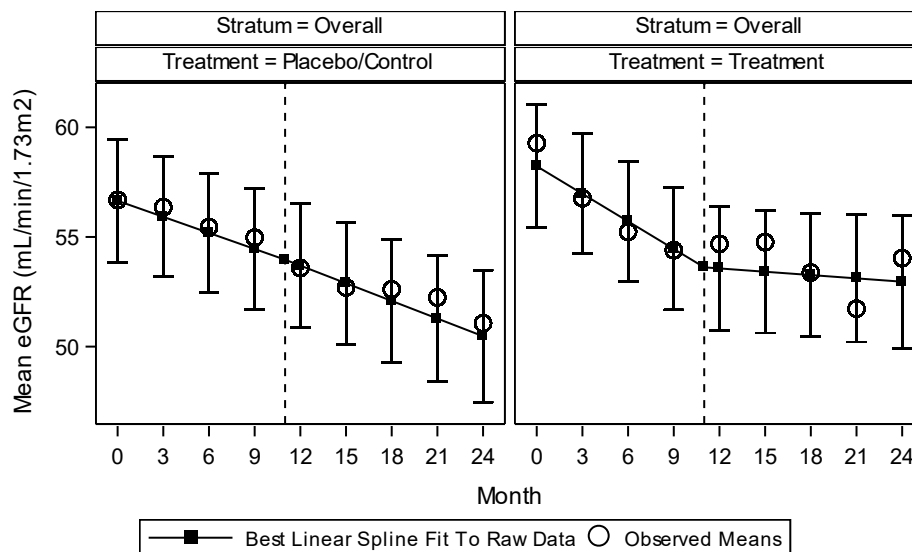




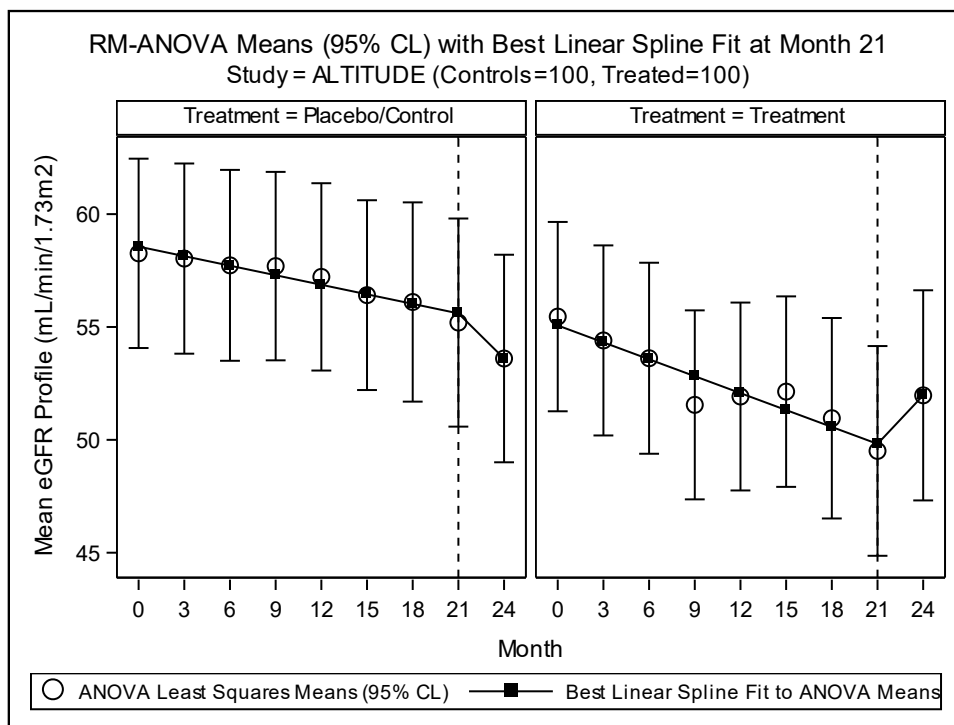
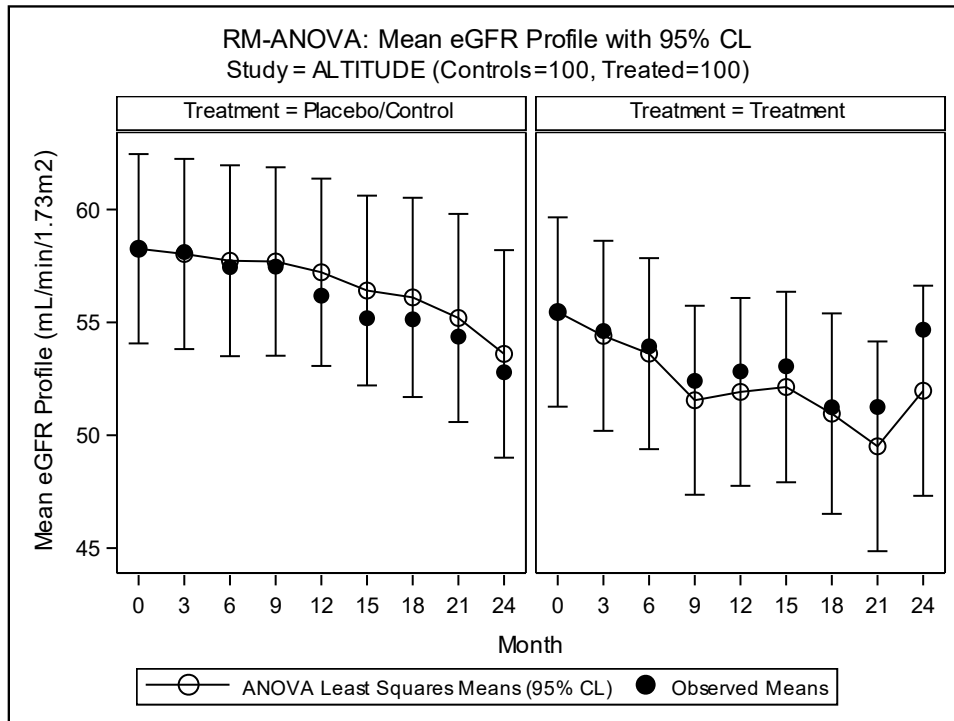
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 12
 Acute Effect (Treated-Control): -0.1364 (-4.152, 3.8791) based on actual intercepts
 Acute Effect (Treated-Control): -1.5444 (-3.6223, 0.5335) based on equal intercepts
 Study = ALTITUDE (Controls=200, Treated=200)

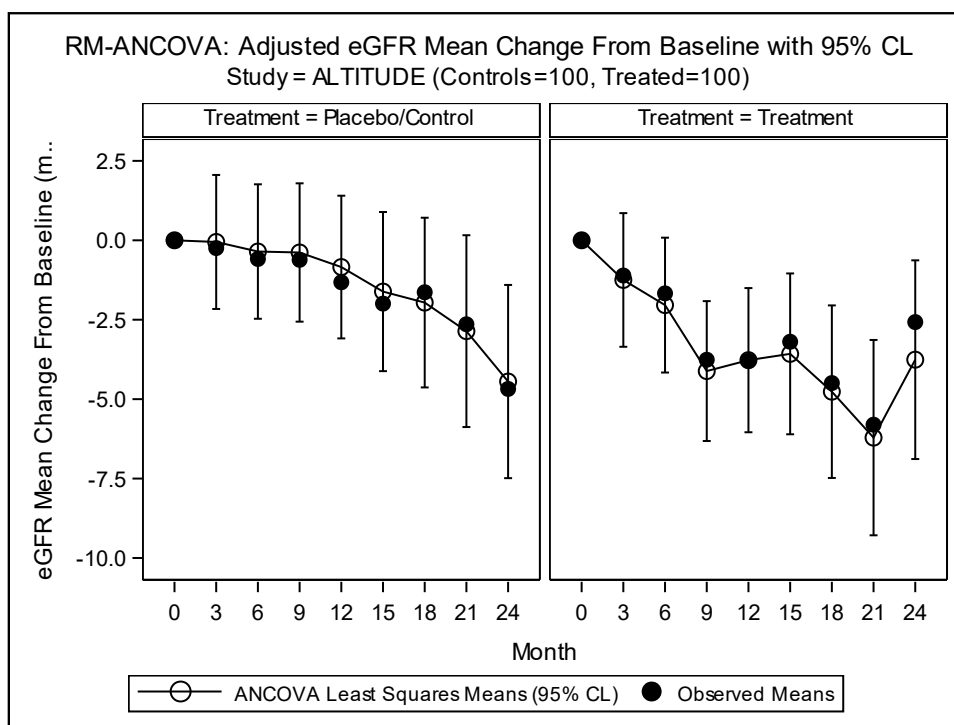
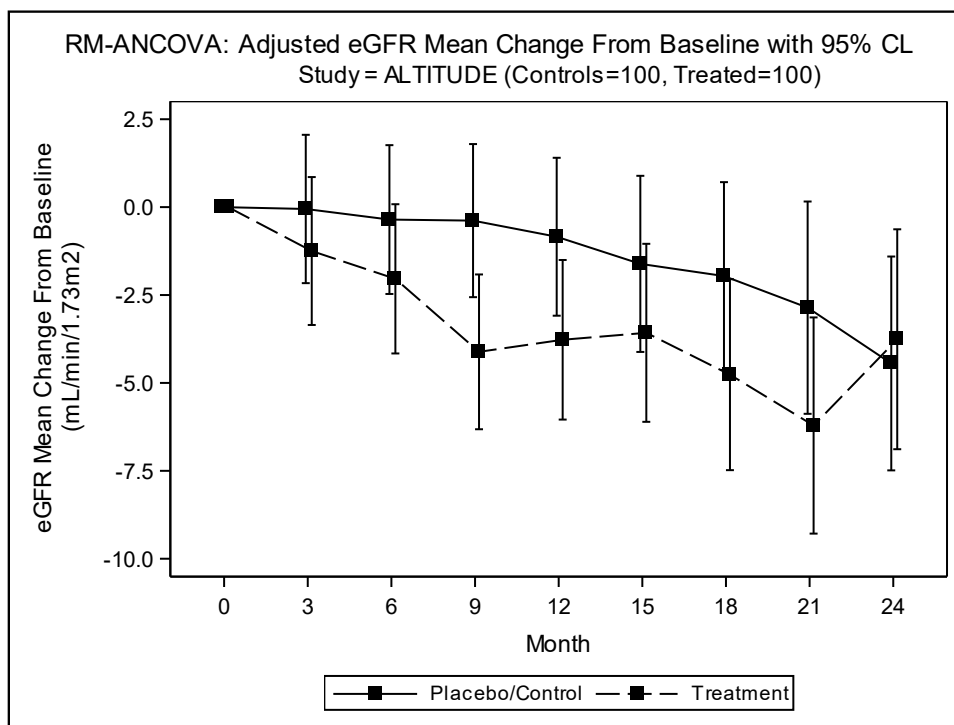


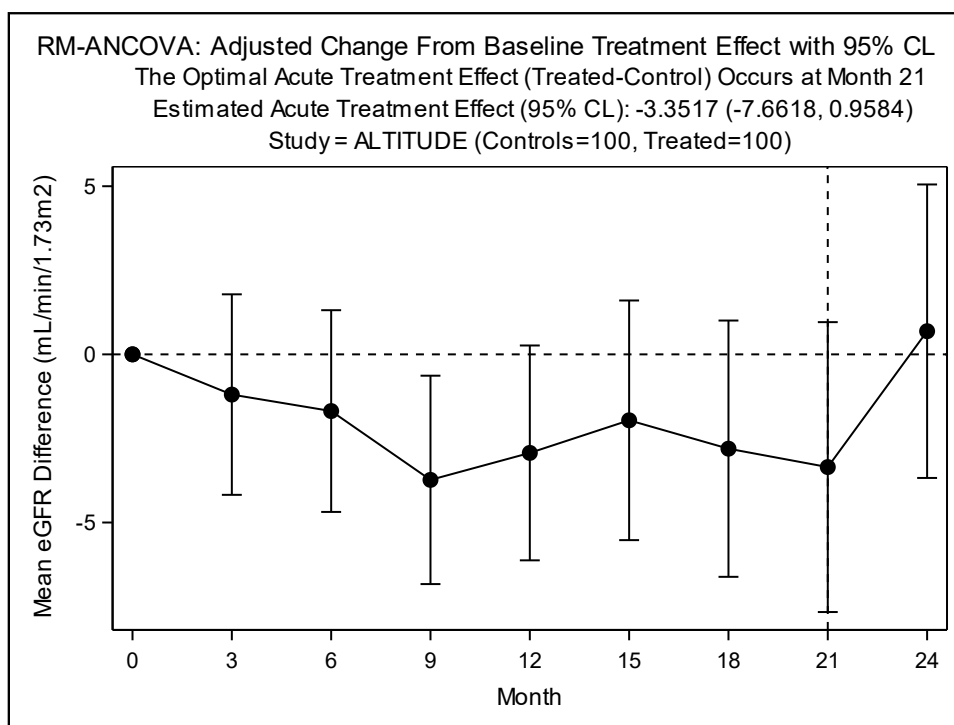
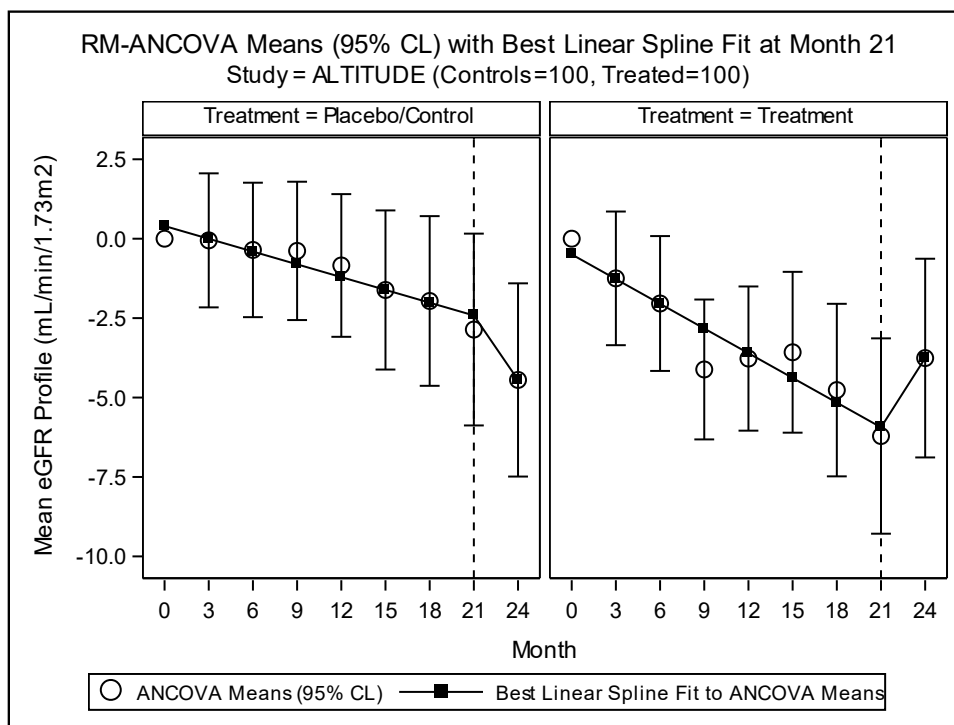
Linear Spline Mixed Model (95% CL) With Incremental Knot at Month 11
 Acute Effect (Treated-Control): -0.3473 (-4.3751, 3.6805) based on actual intercepts
 Acute Effect (Treated-Control): -1.9236 (-4.0295, 0.1823) based on equal intercepts
 Study = ALTITUDE (Controls=200, Treated=200)



Case 4: Sample Size=100 per Treatment group





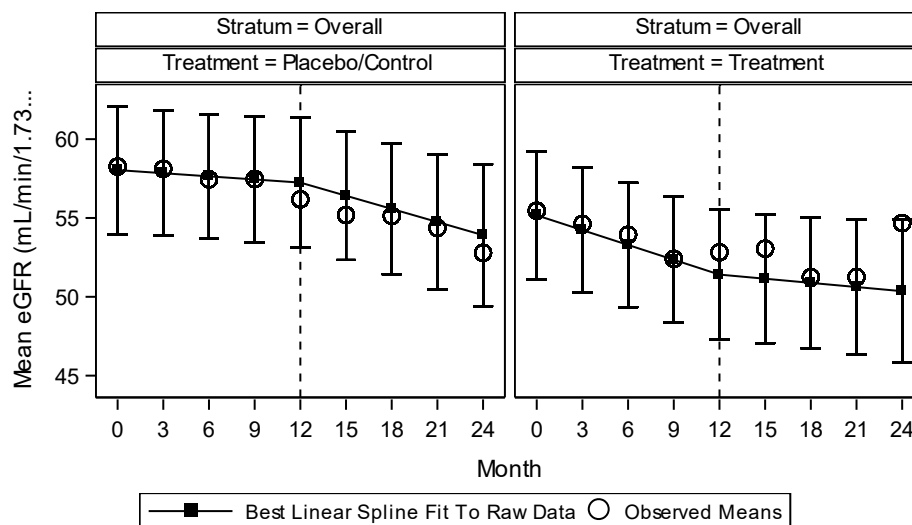


Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 12

Acute Effect (Treated-Control): -5.8348 (-11.6519, -0.0178) based on actual intercepts

Acute Effect (Treated-Control): -2.9655 (-5.9923, 0.0614) based on equal intercepts

Study = ALTITUDE (Controls=100, Treated=100)

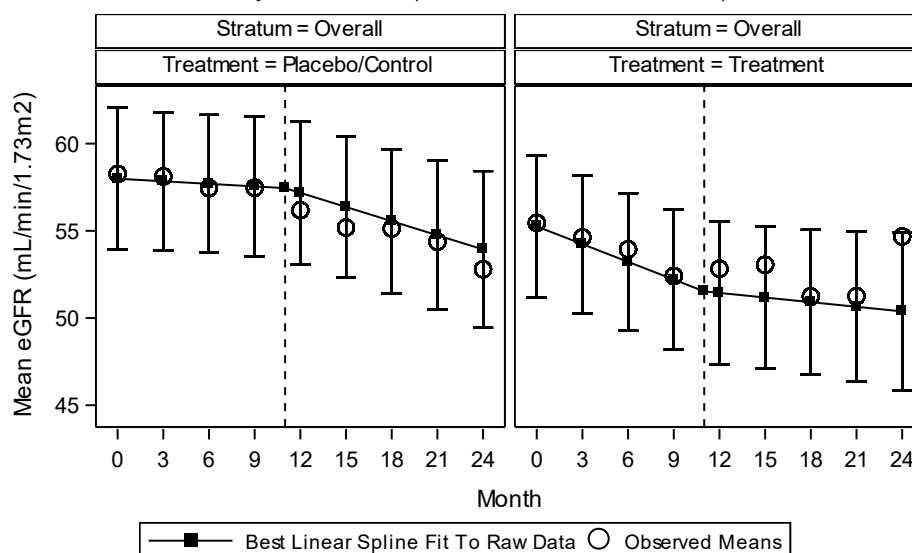


Linear Spline Mixed Model (95% CL) With Incremental Knot at Month 11

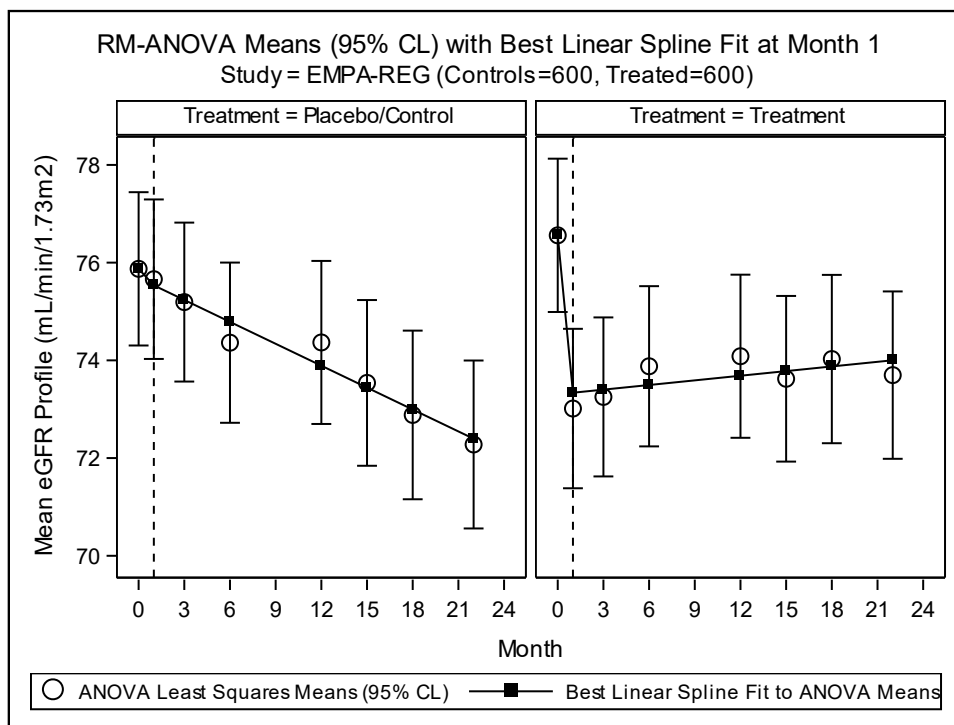
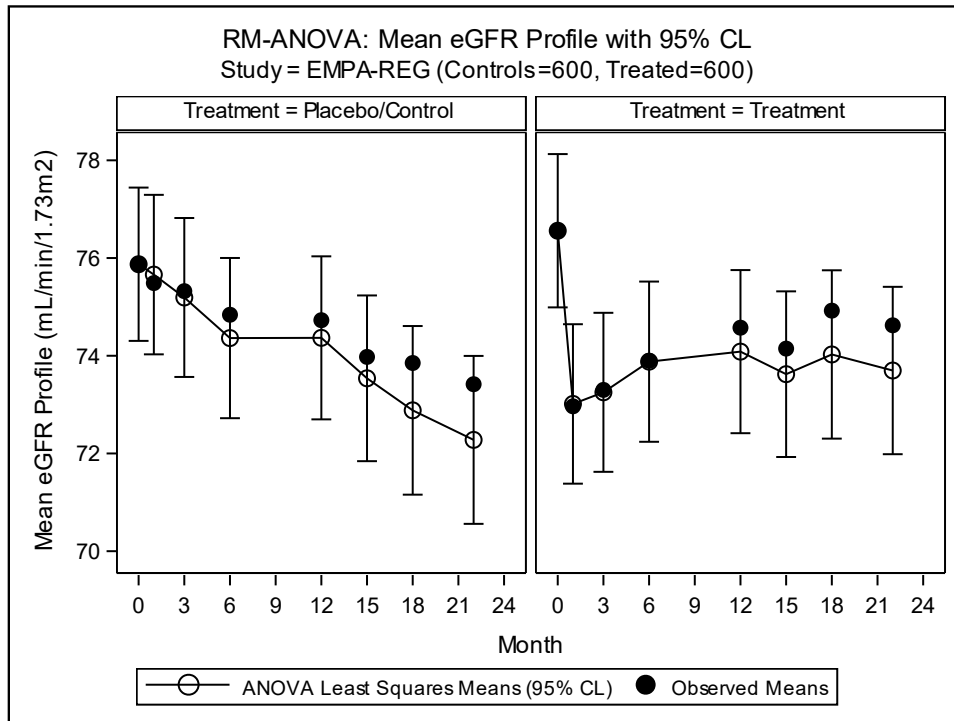
Acute Effect (Treated-Control): -5.9199 (-11.742, -0.0978) based on actual intercepts

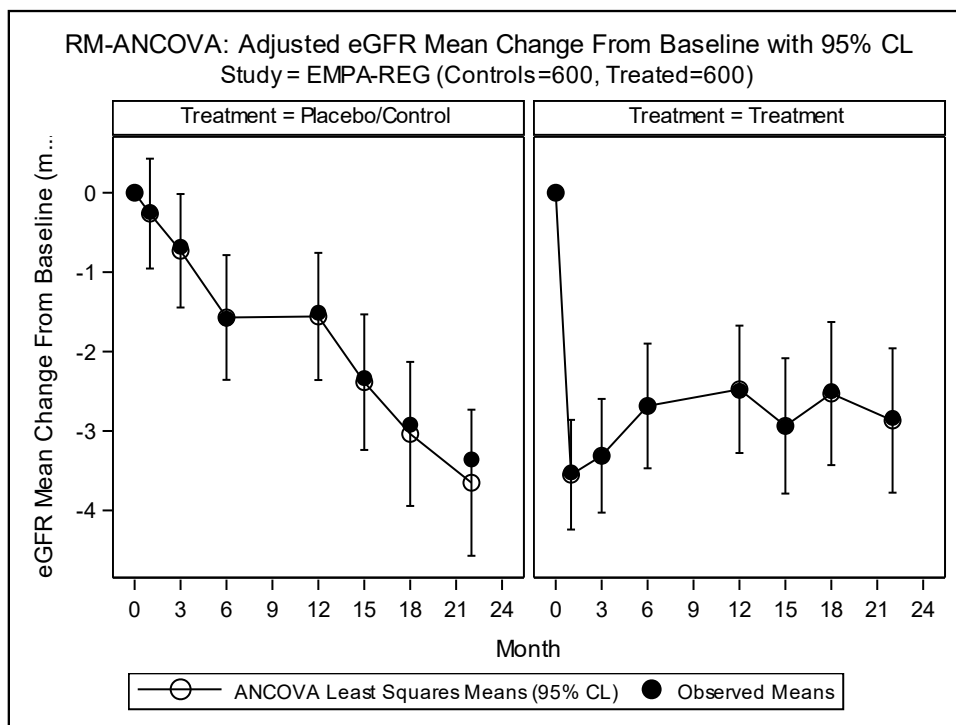
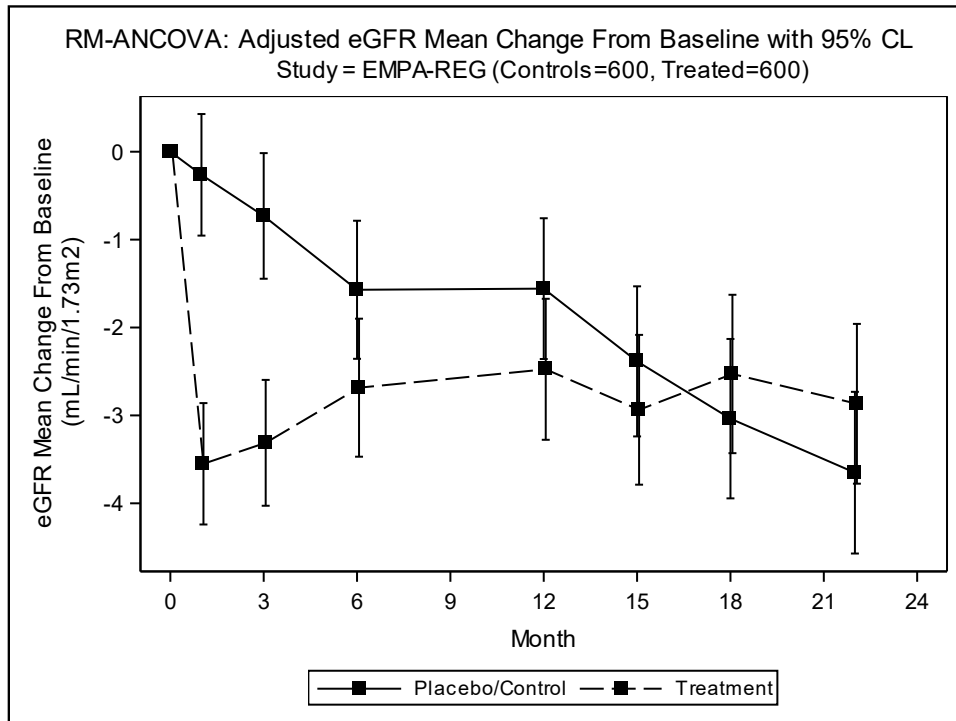
Acute Effect (Treated-Control): -3.1828 (-6.2478, -0.1178) based on equal intercepts

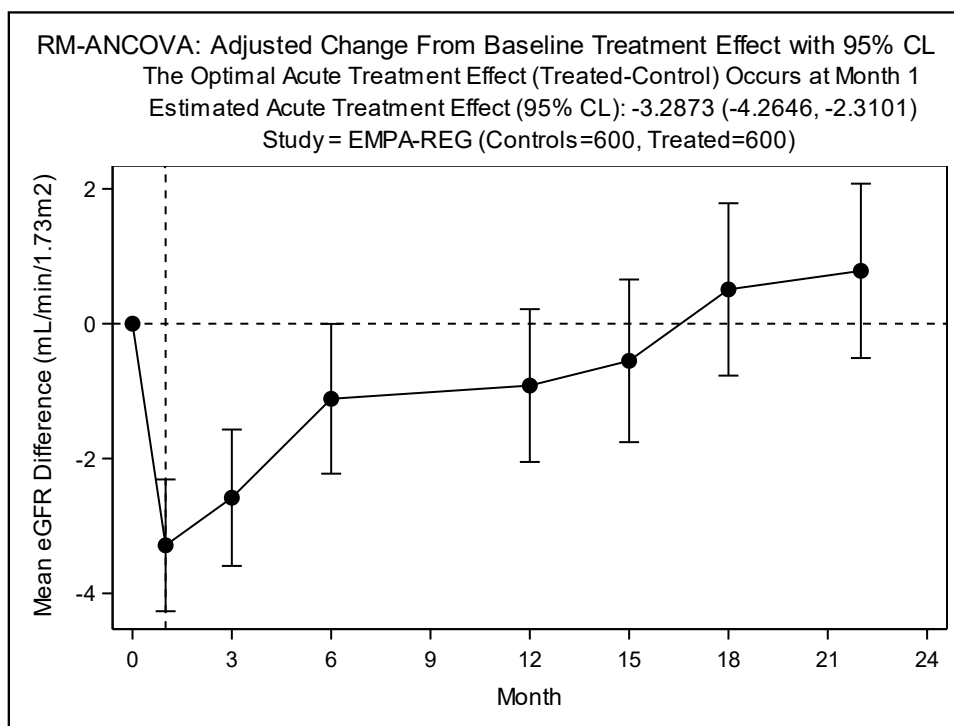
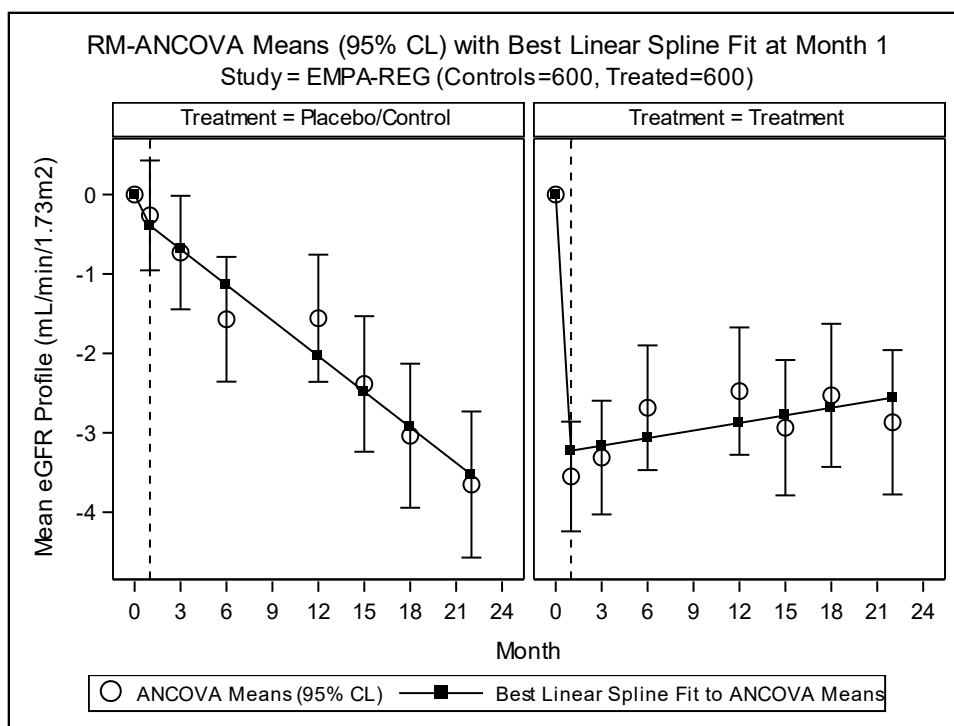
Study = ALTITUDE (Controls=100, Treated=100)



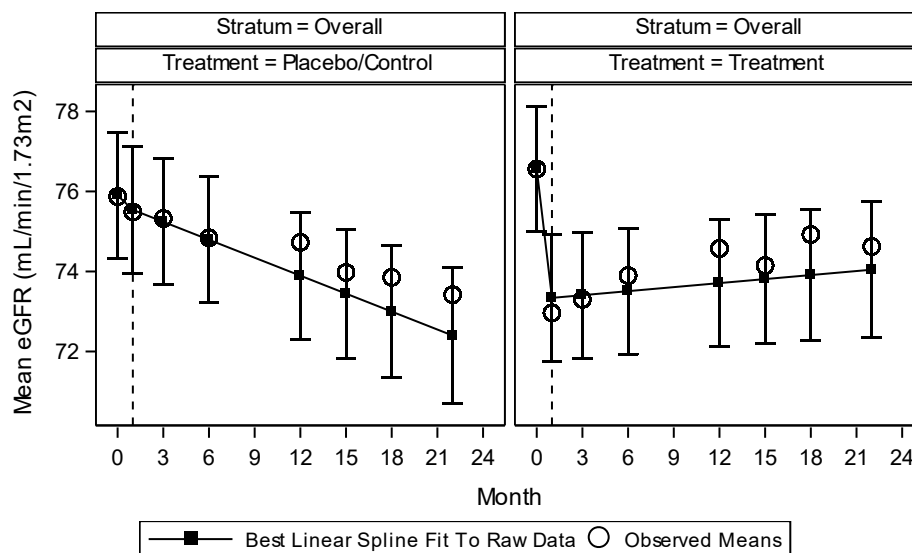
Appendix 2: EMPA-REG results



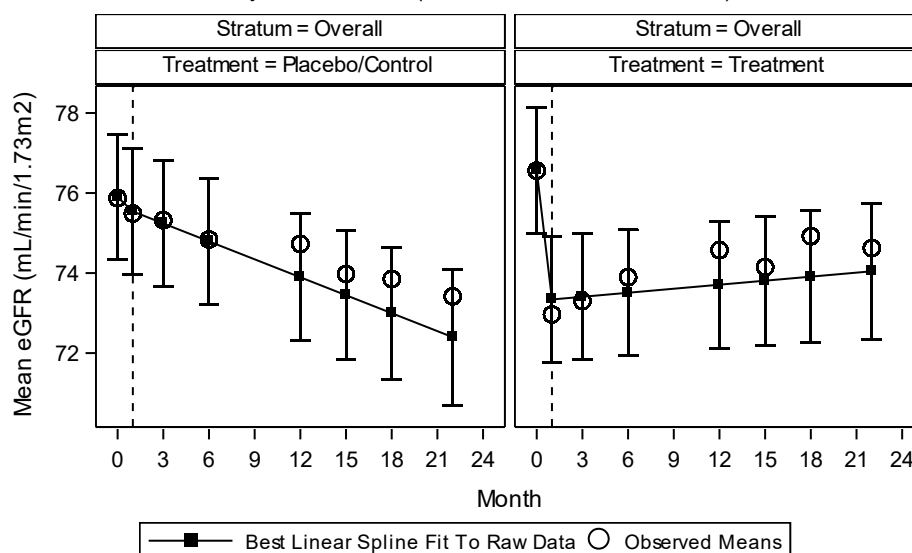




Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 1
 Acute Effect (Treated-Control): -2.1995 (-4.4378, 0.0388) based on actual intercepts
 Acute Effect (Treated-Control): -2.8643 (-3.7297, -1.9989) based on equal intercepts
 Study = EMPA-REG (Controls=600, Treated=600)

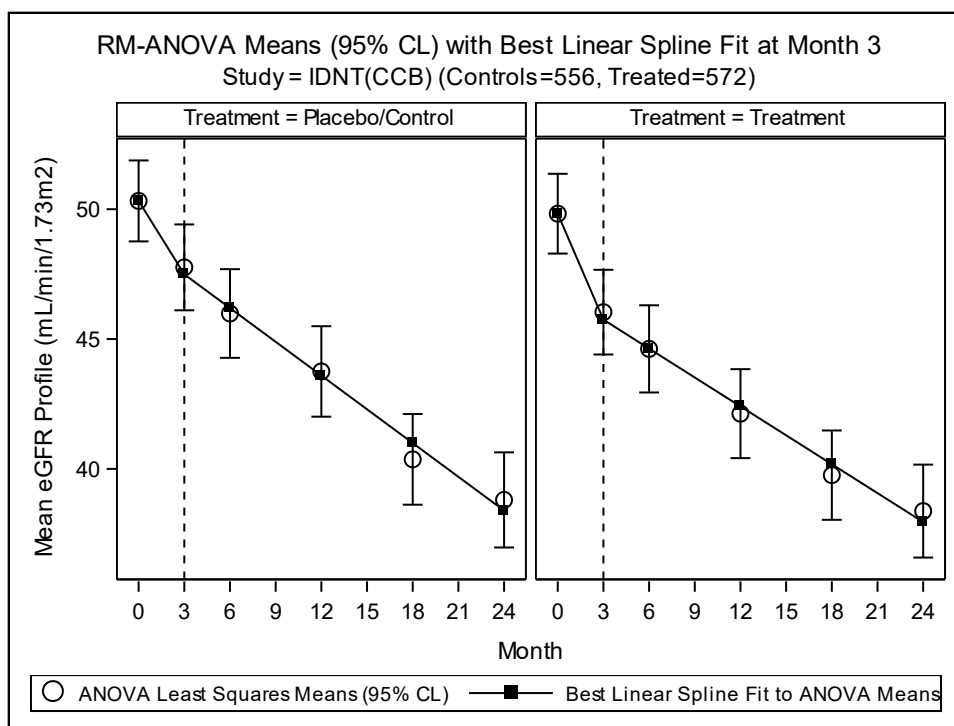
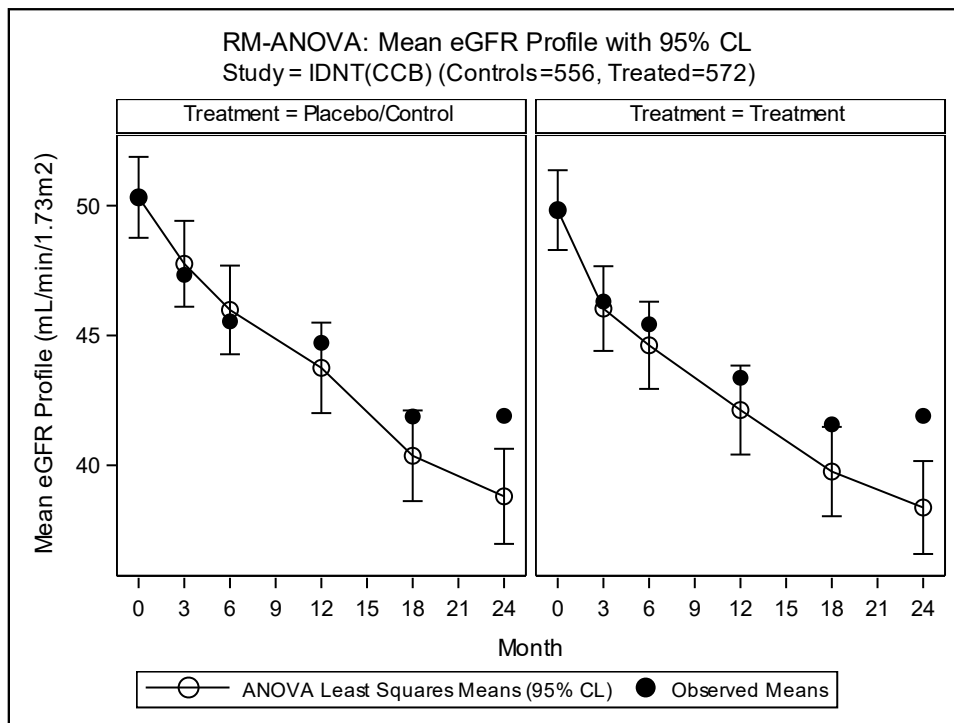


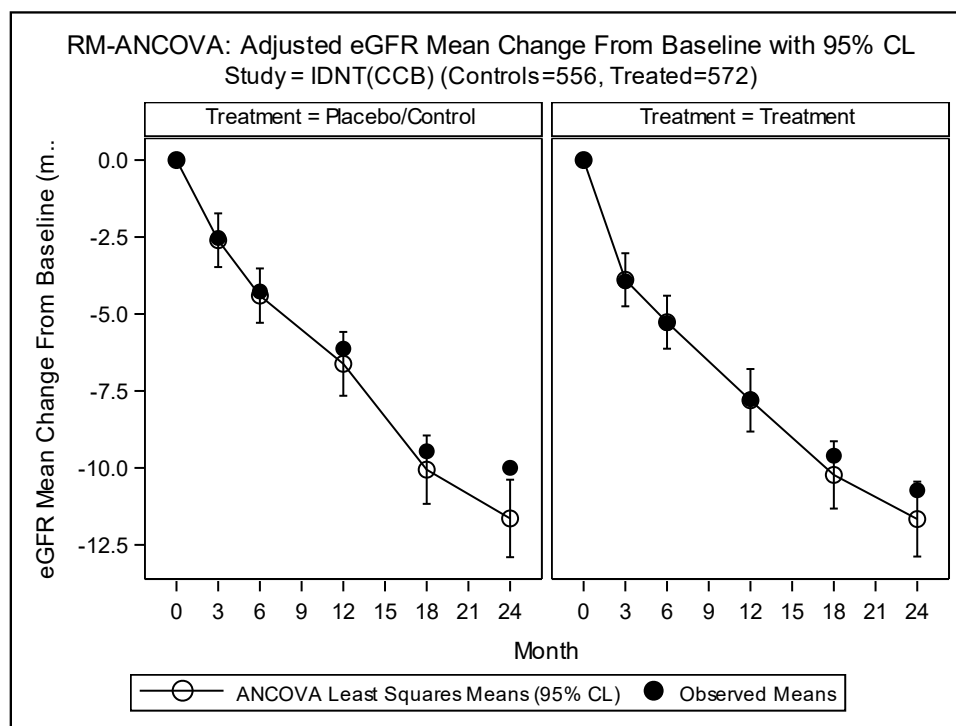
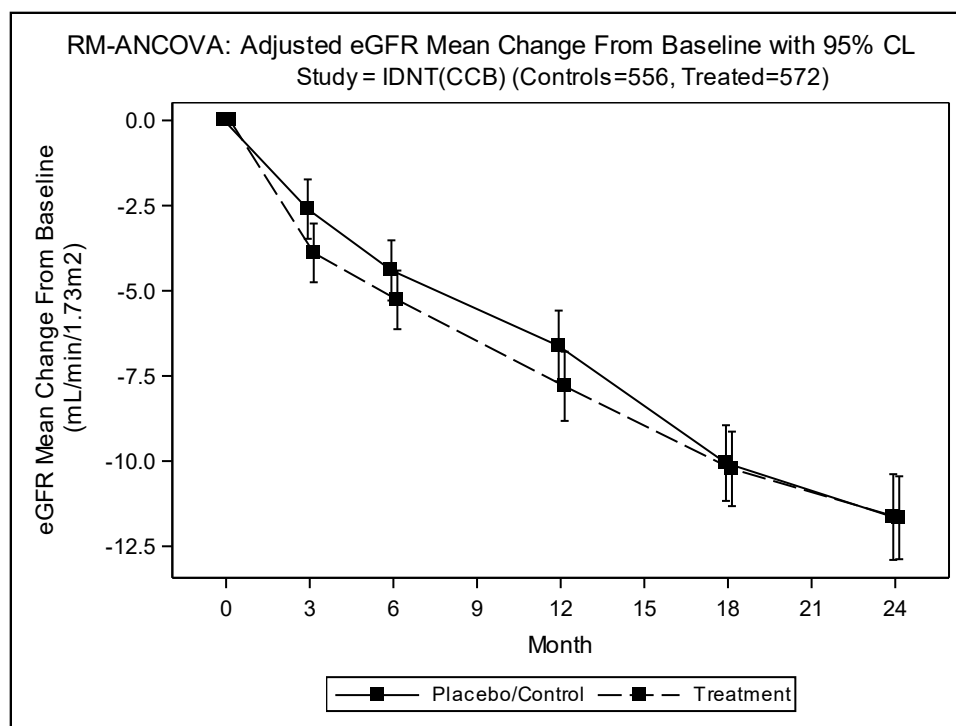
Linear Spline Mixed-Effects Model (95% CL) With Incremental Knot at Month 1
 Acute Effect (Treated-Control): -2.1995 (-4.4378, 0.0388) based on actual intercepts
 Acute Effect (Treated-Control): -2.8643 (-3.7297, -1.9989) based on equal intercepts
 Study = EMPA-REG (Controls=600, Treated=600)

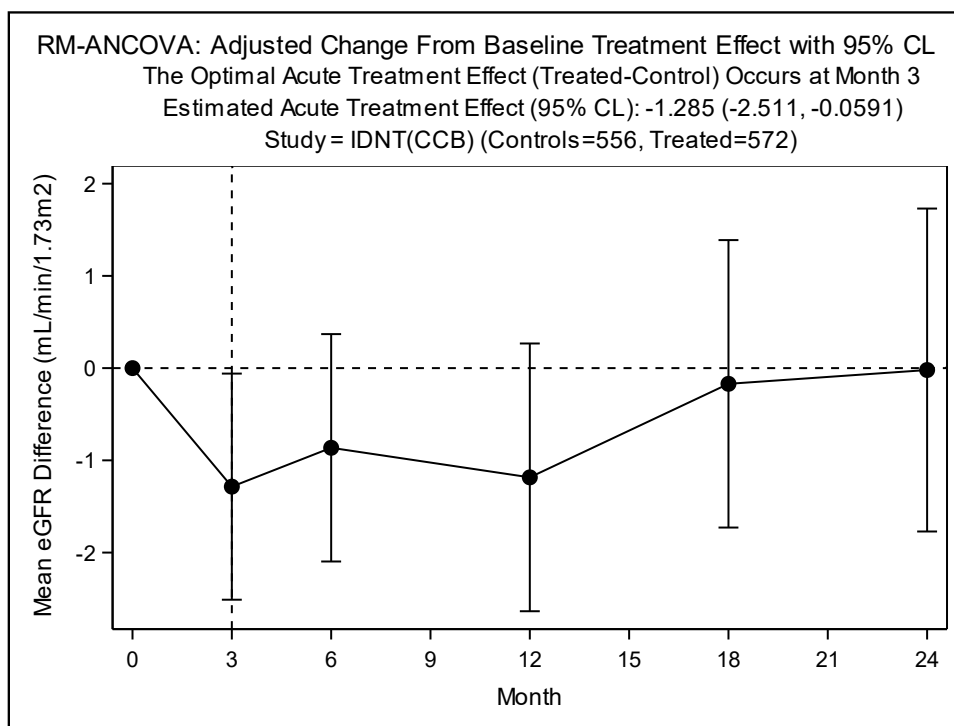
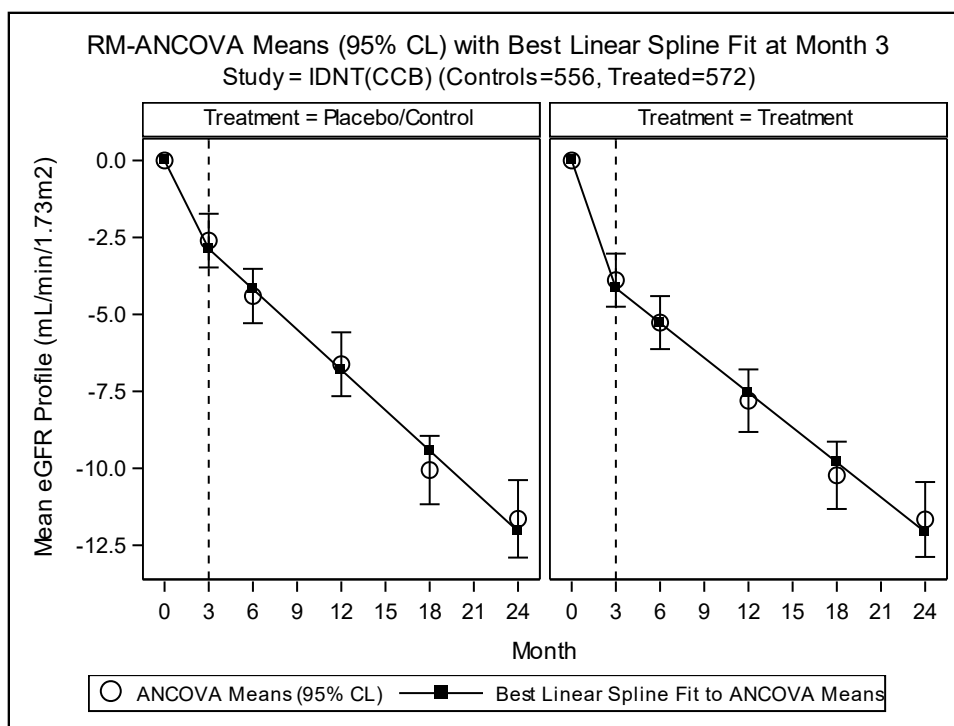


Appendix 3: IDNT(CCB) results

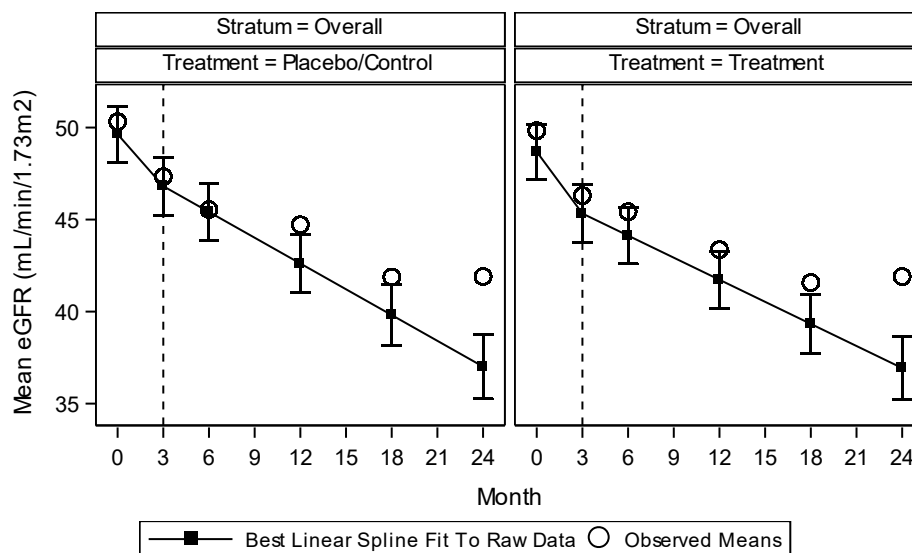
Case 1: Results for the Overall Cohort



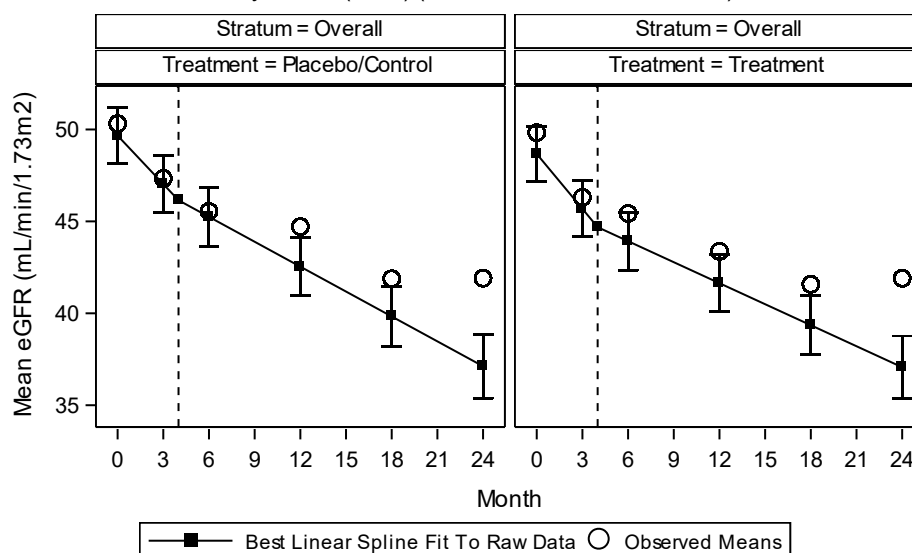




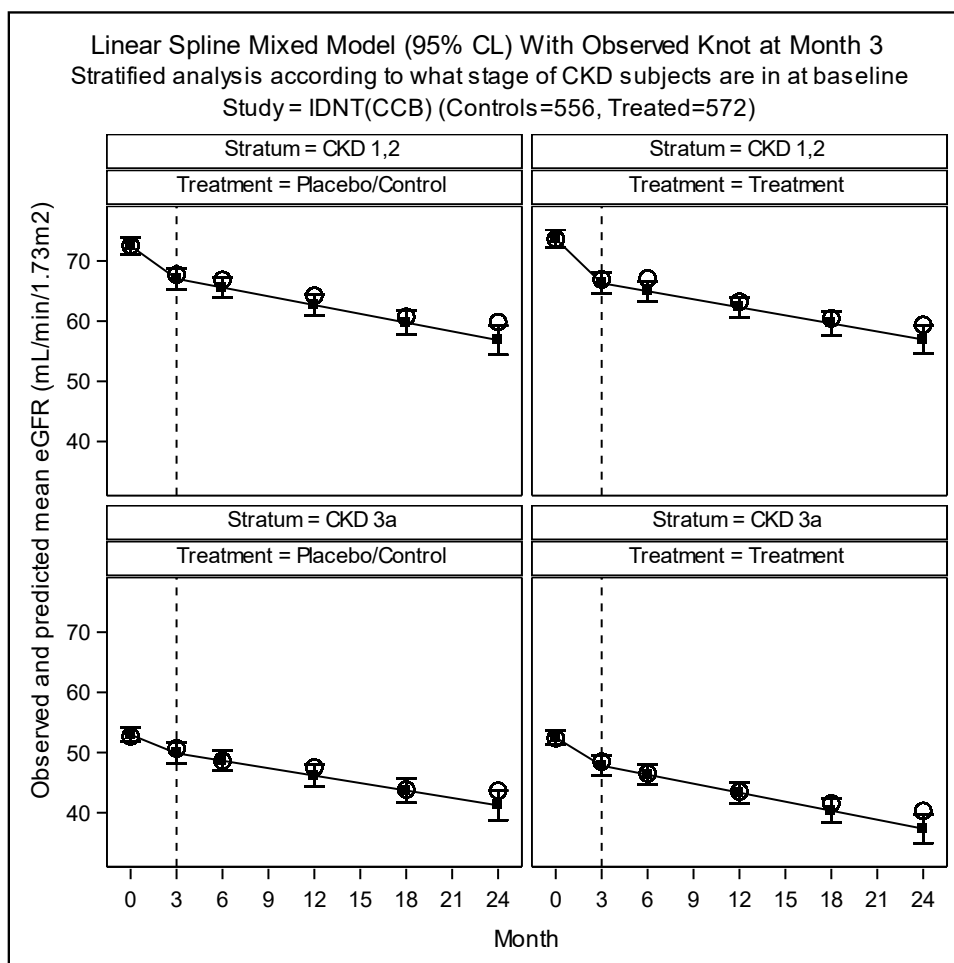
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 3
 Acute Effect (Treated-Control): -1.486 (-3.7032, 0.7312) based on actual intercepts
 Acute Effect (Treated-Control): -0.5319 (-1.493, 0.4291) based on equal intercepts
 Study = IDNT(CCB) (Controls=556, Treated=572)

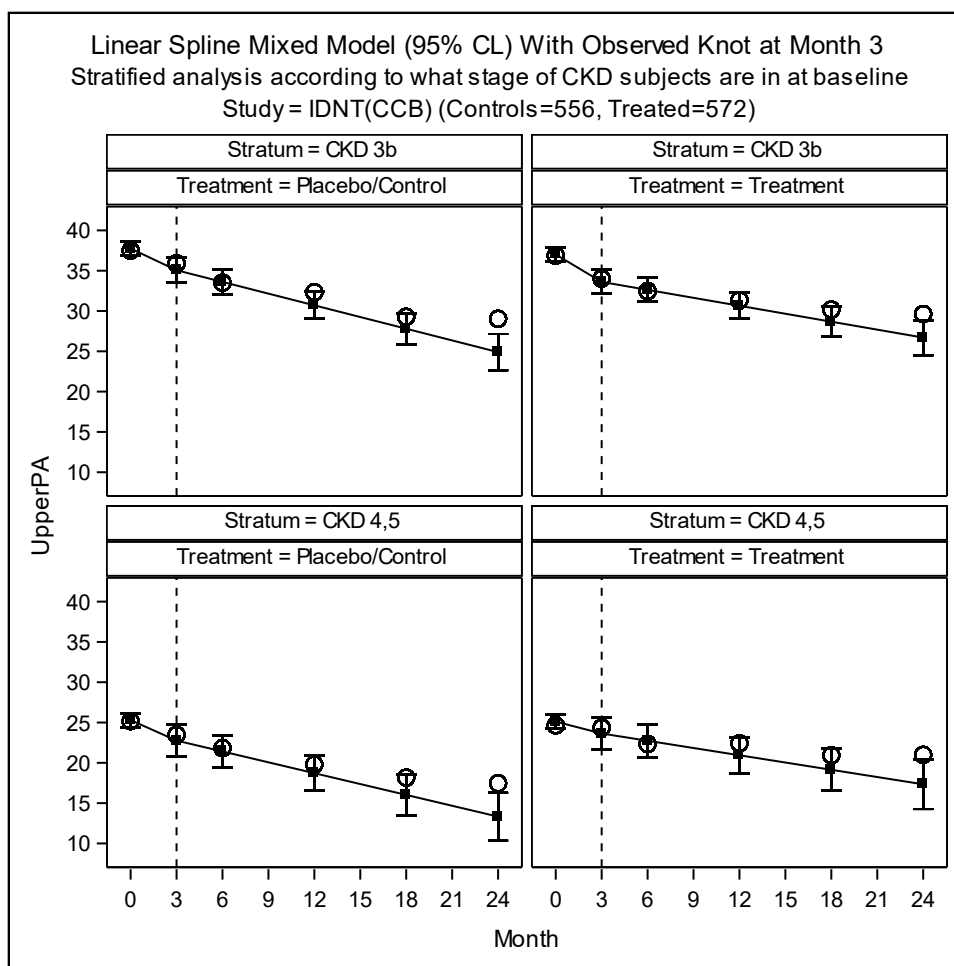


Linear Spline Mixed-Effects Model (95% CL) With Incremental Knot at Month 4
 Acute Effect (Treated-Control): -1.4641 (-3.7259, 0.7978) based on actual intercepts
 Acute Effect (Treated-Control): -0.4715 (-1.5328, 0.5899) based on equal intercepts
 Study = IDNT(CCB) (Controls=556, Treated=572)



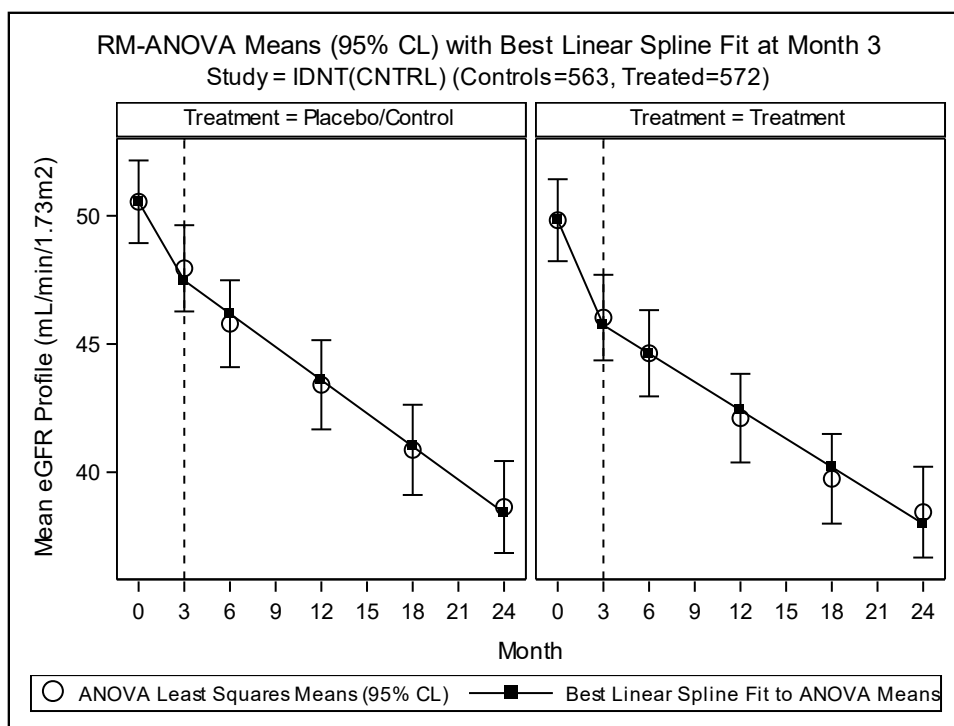
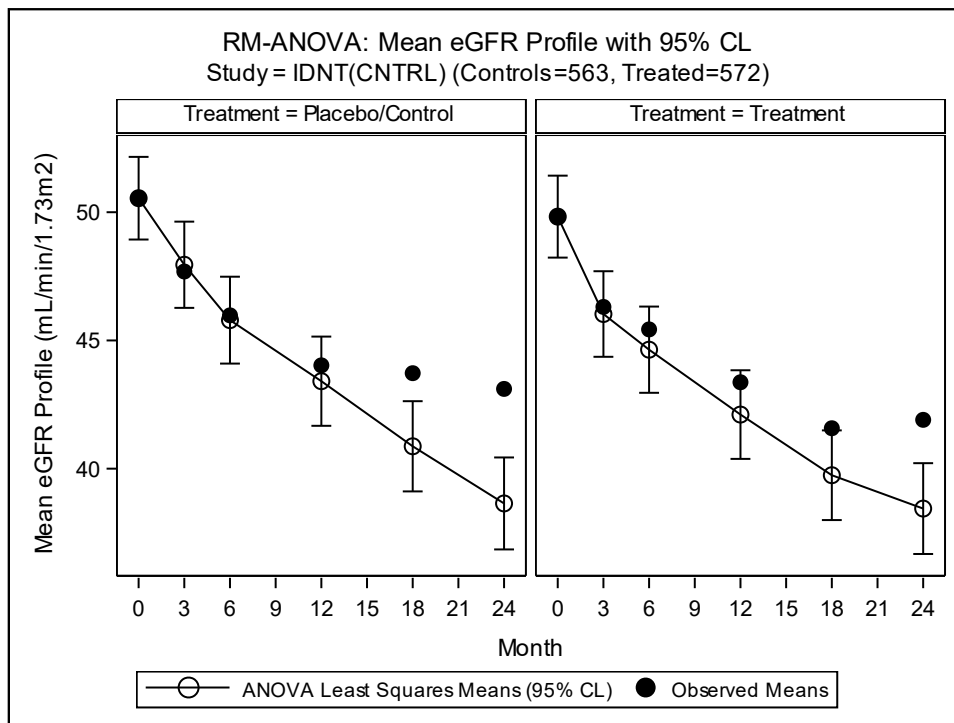
Case 2: Results When Stratified by CKD Stage

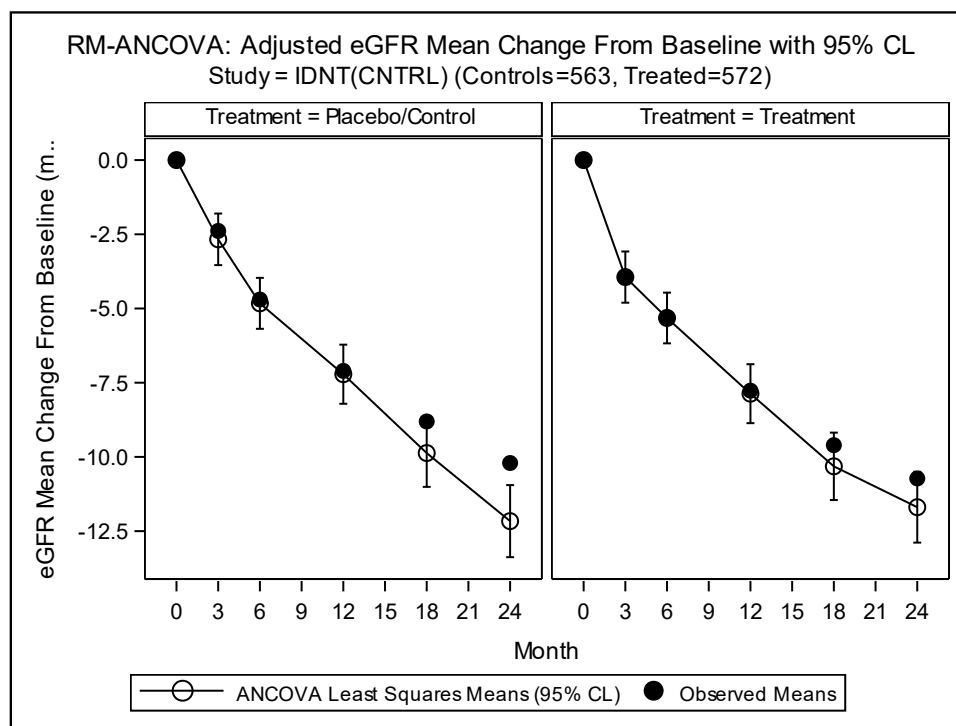
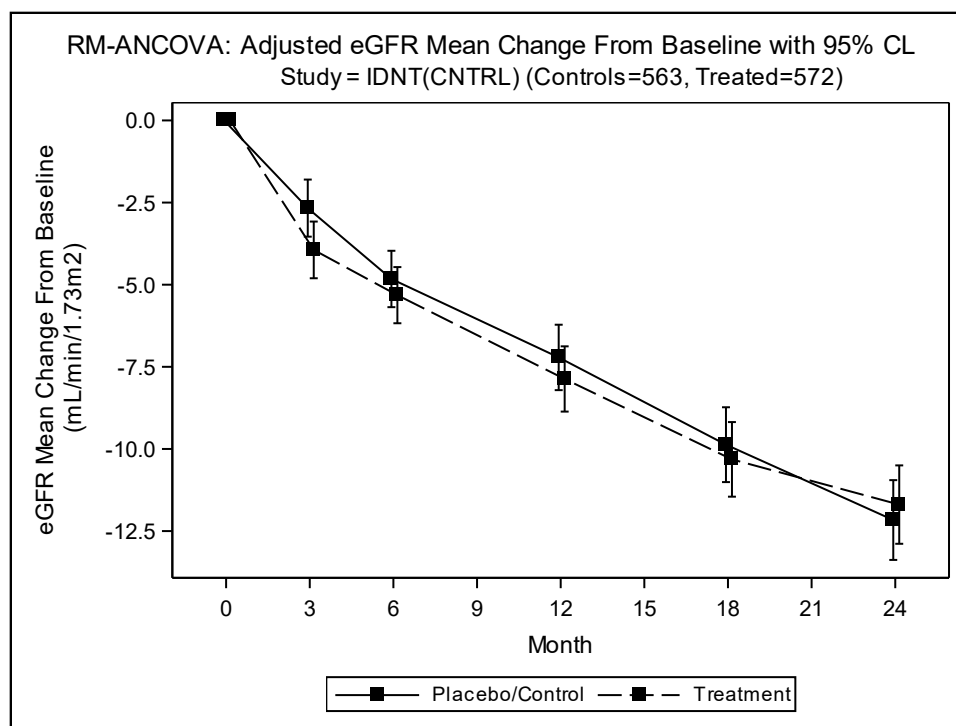


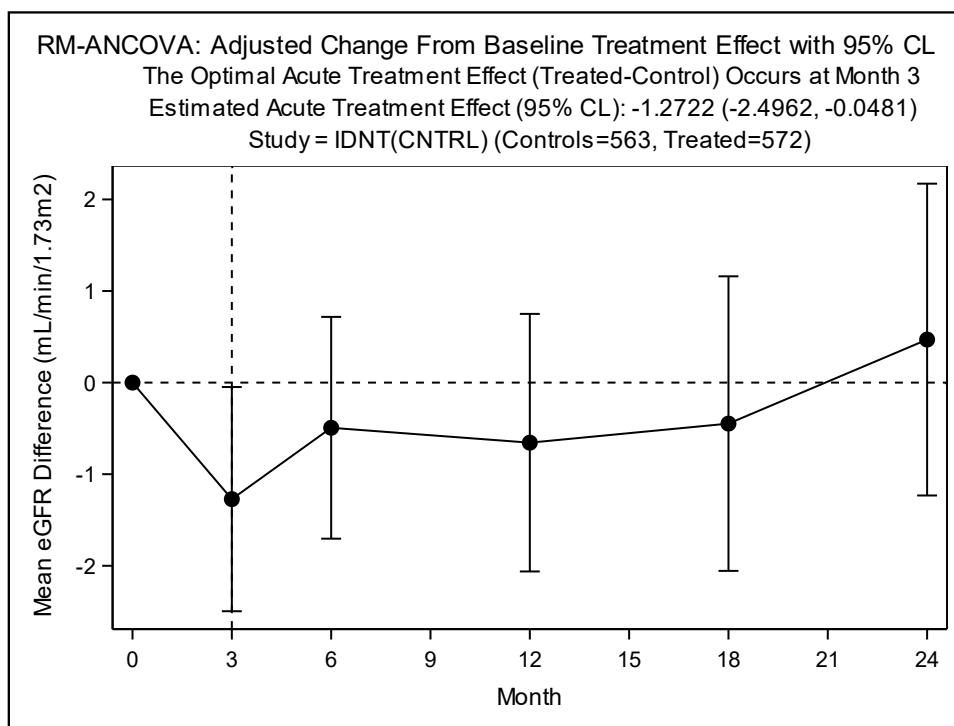
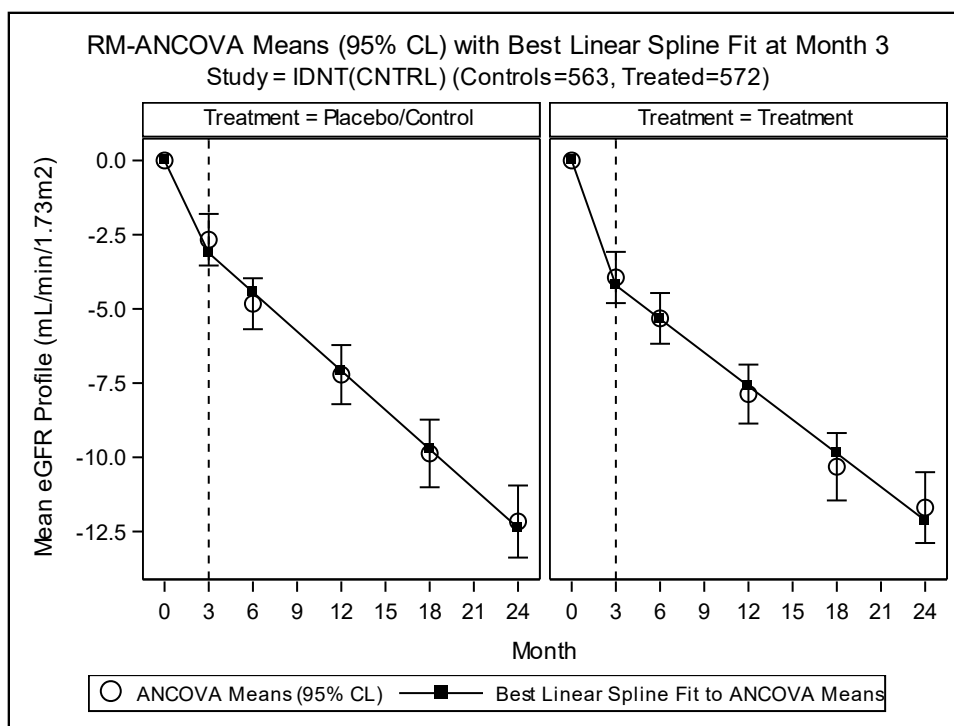


Appendix 4: IDNT(CNTRL) results

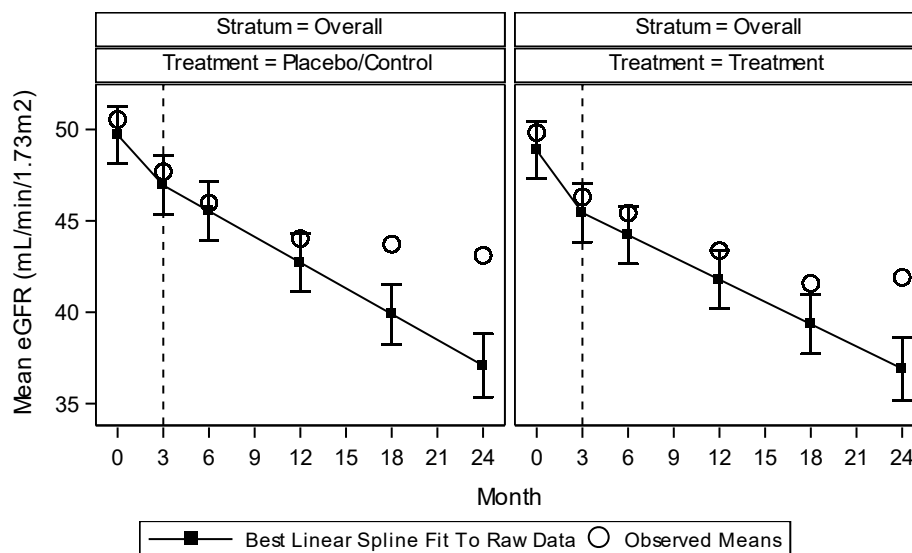
Case 1: Results for the Overall Cohort



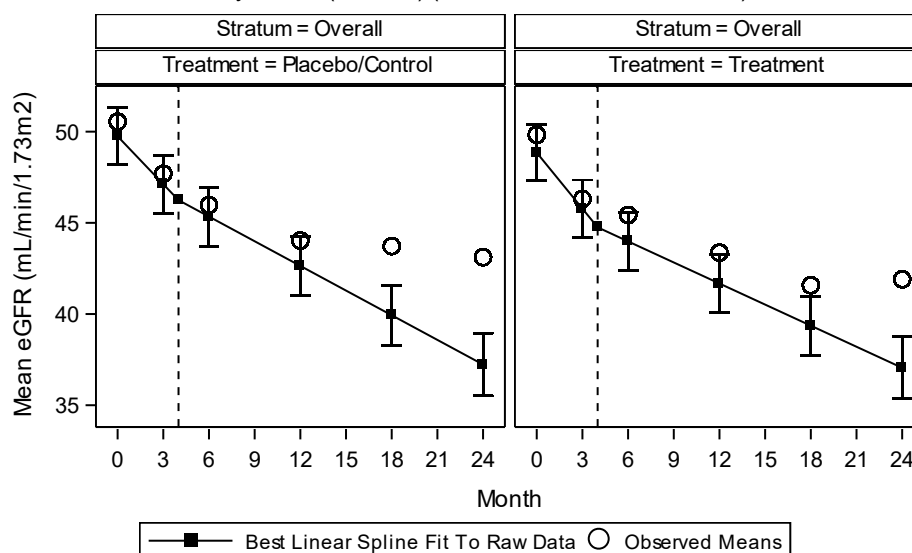




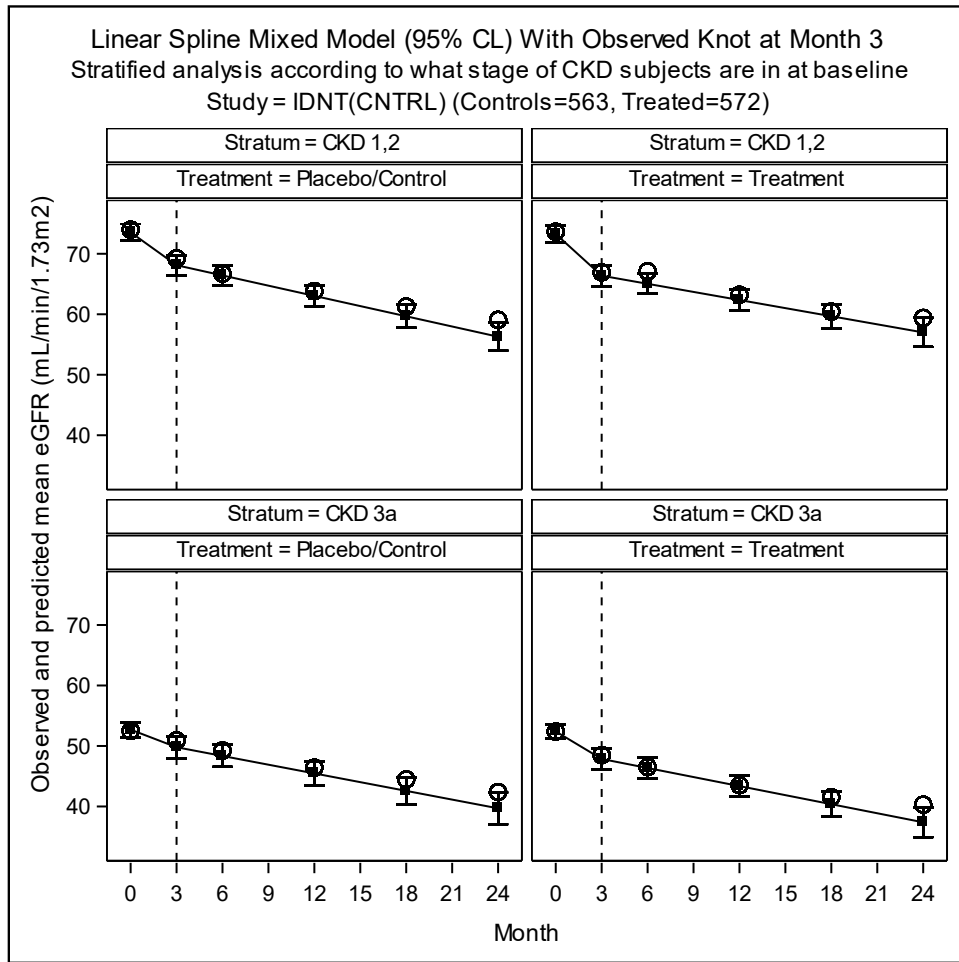
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 3
 Acute Effect (Treated-Control): -1.5024 (-3.7728, 0.768) based on actual intercepts
 Acute Effect (Treated-Control): -0.654 (-1.6193, 0.3114) based on equal intercepts
 Study = IDNT(CNTRL) (Controls=563, Treated=572)



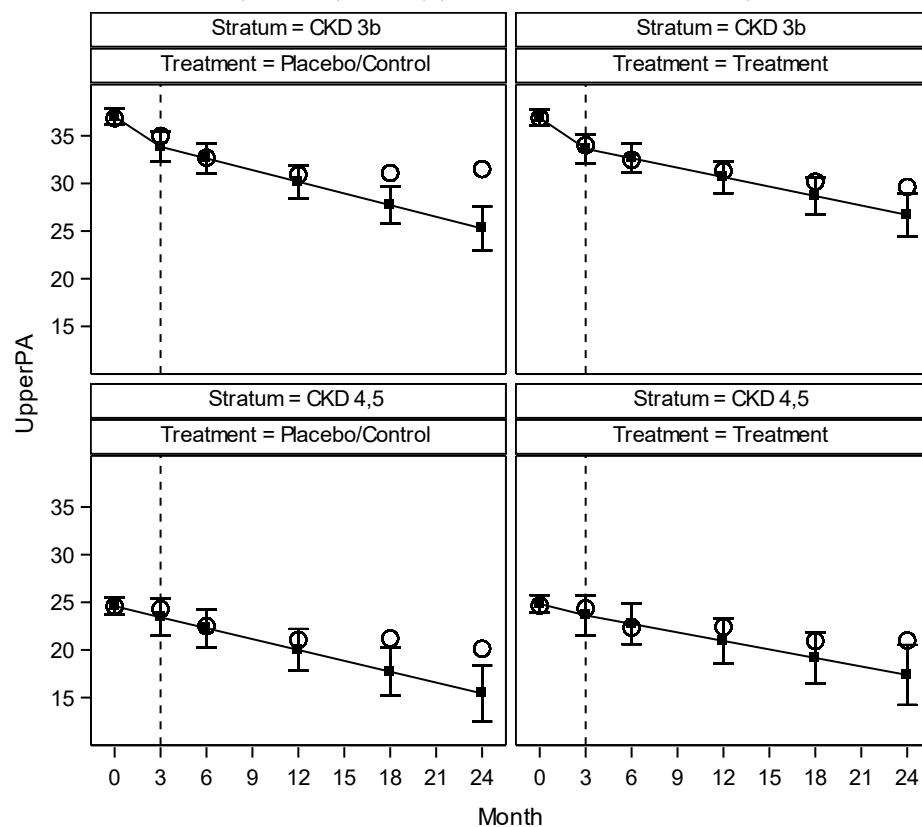
Linear Spline Mixed-Effects Model (95% CL) With Incremental Knot at Month 4
 Acute Effect (Treated-Control): -1.4903 (-3.7974, 0.8168) based on actual intercepts
 Acute Effect (Treated-Control): -0.5955 (-1.6712, 0.4801) based on equal intercepts
 Study = IDNT(CNTRL) (Controls=563, Treated=572)



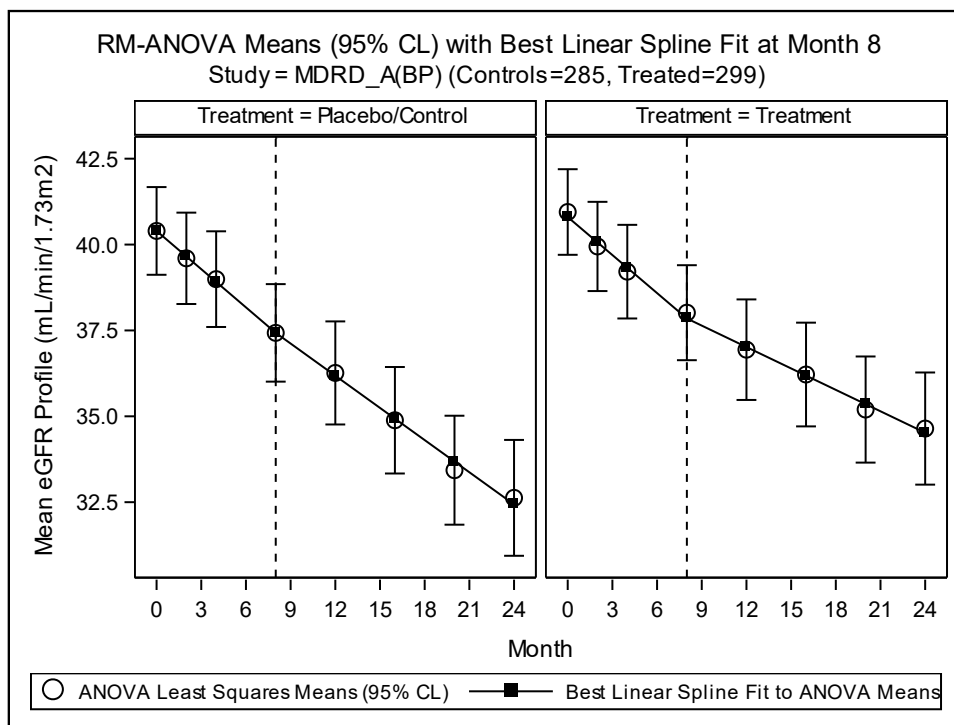
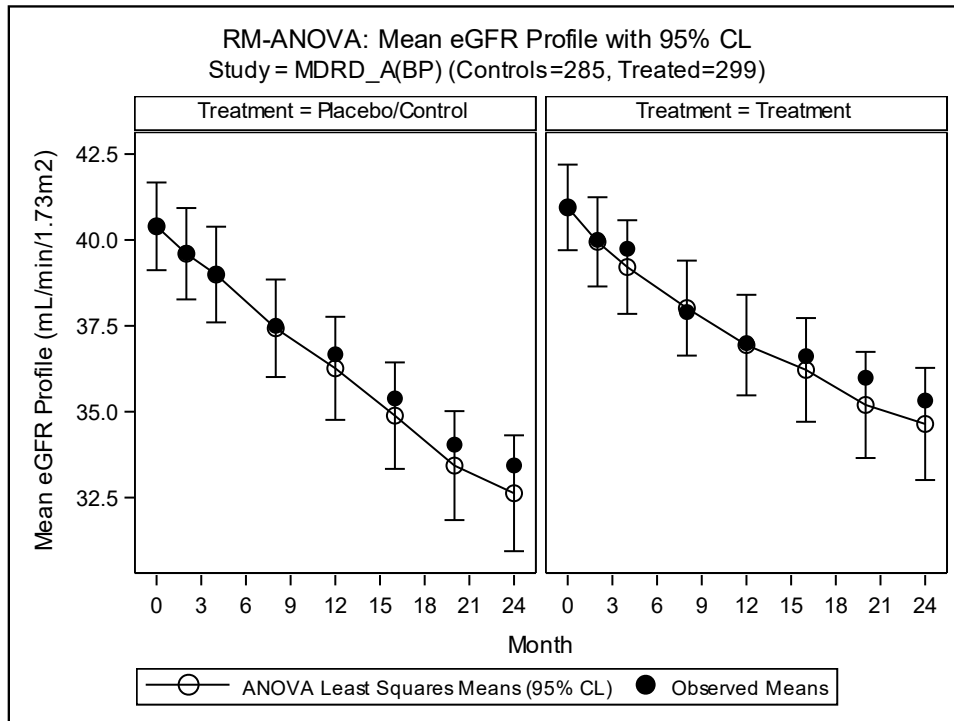
Case 2: Results When Stratified by CKD Stage

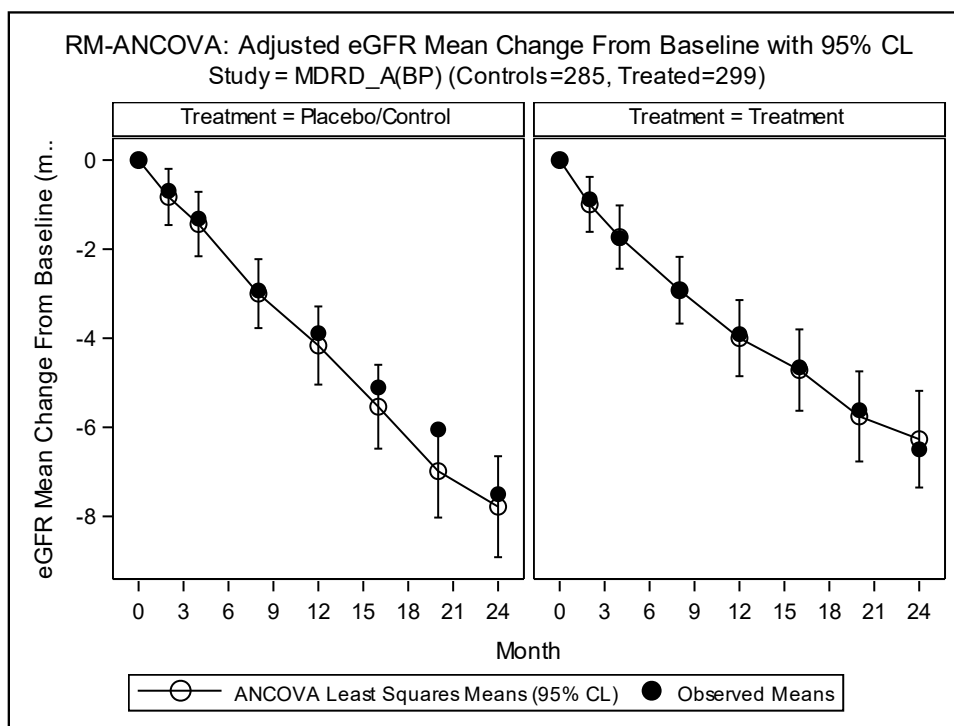
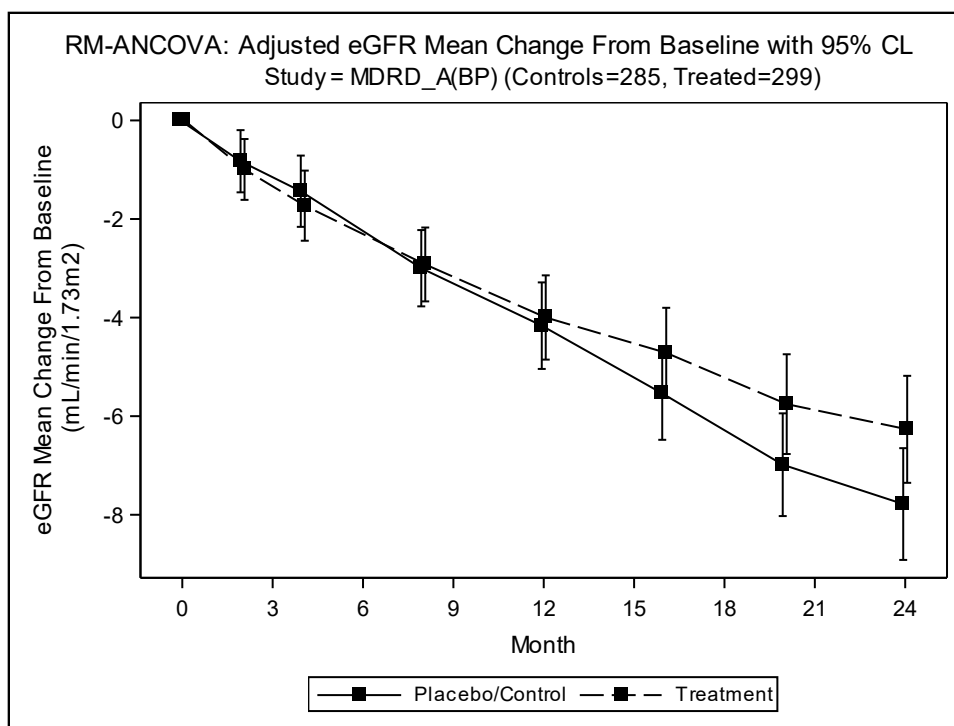


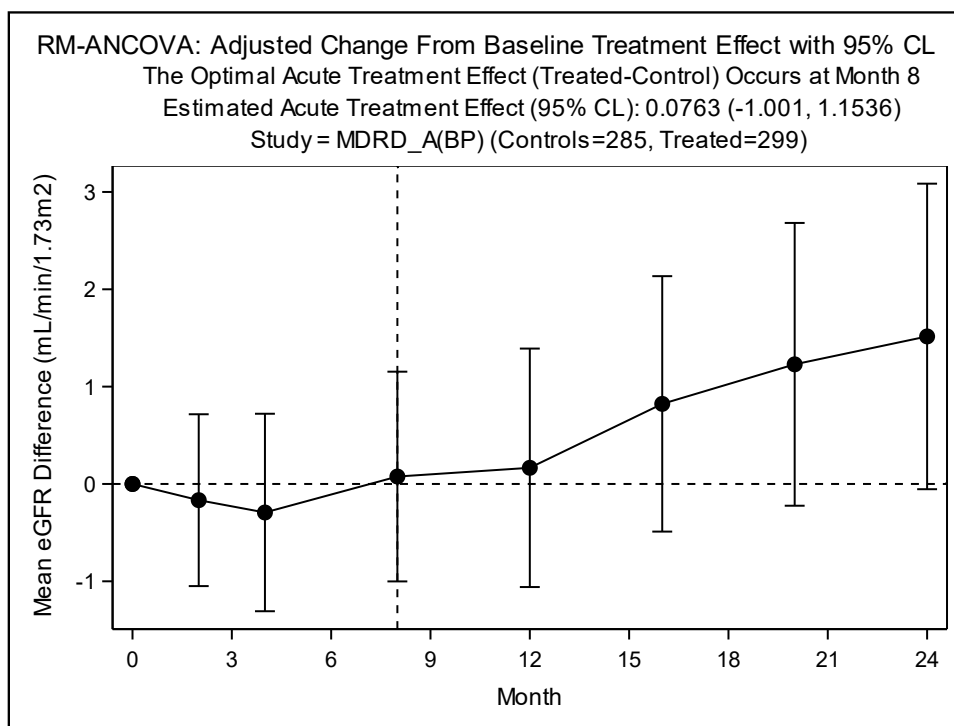
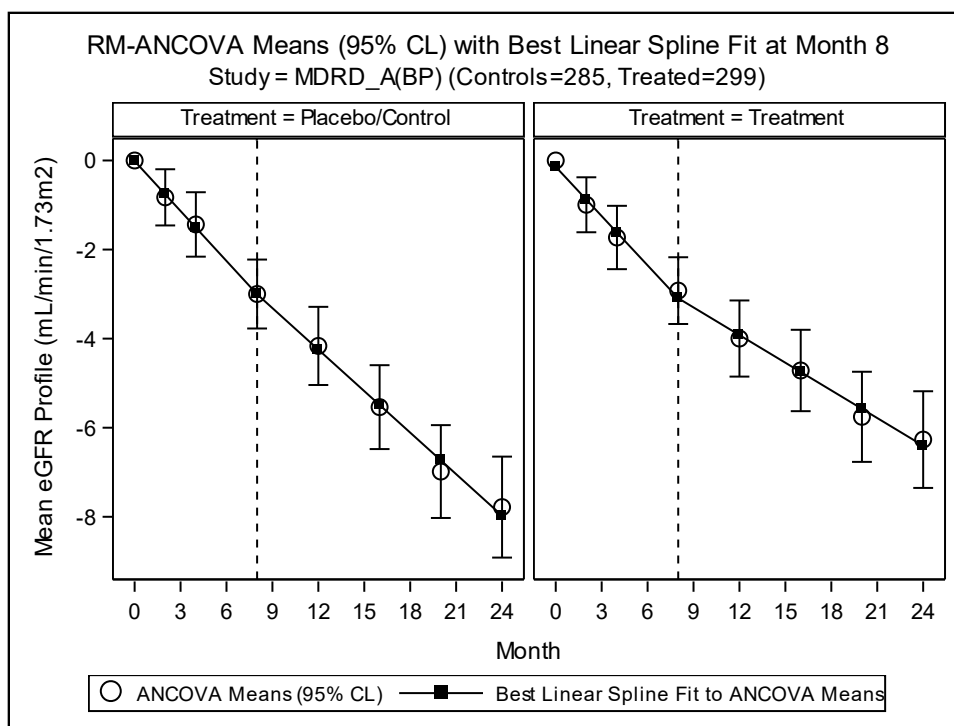
Linear Spline Mixed Model (95% CL) With Observed Knot at Month 3
Stratified analysis according to what stage of CKD subjects are in at baseline
Study = IDNT(CNTRL) (Controls=563, Treated=572)



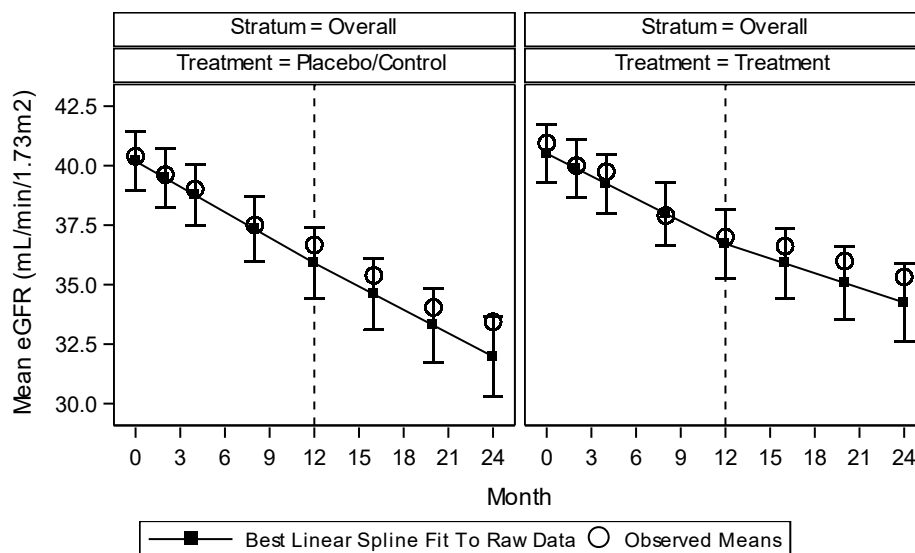
Appendix 5: MDRD-A(BP) results



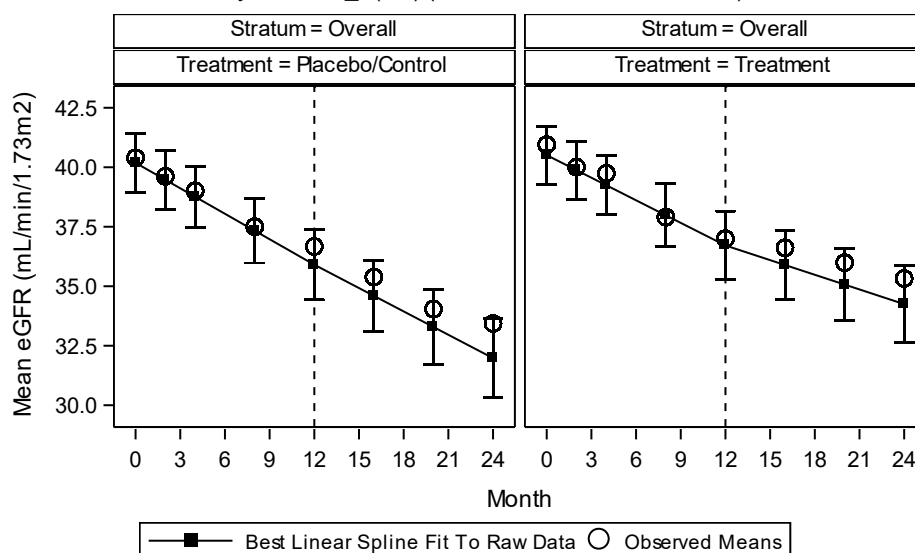




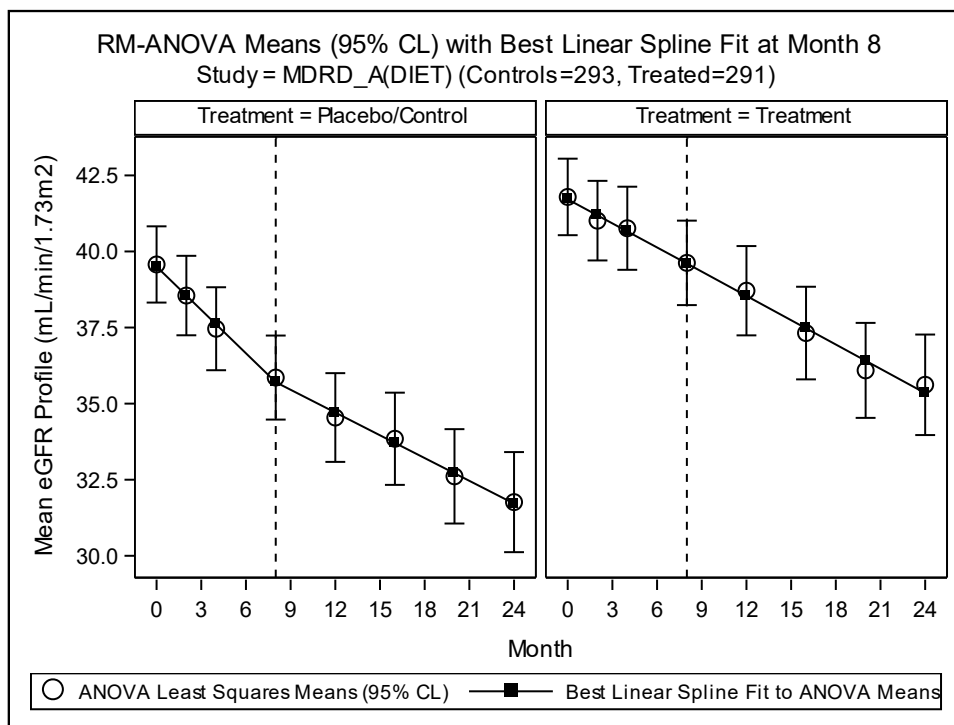
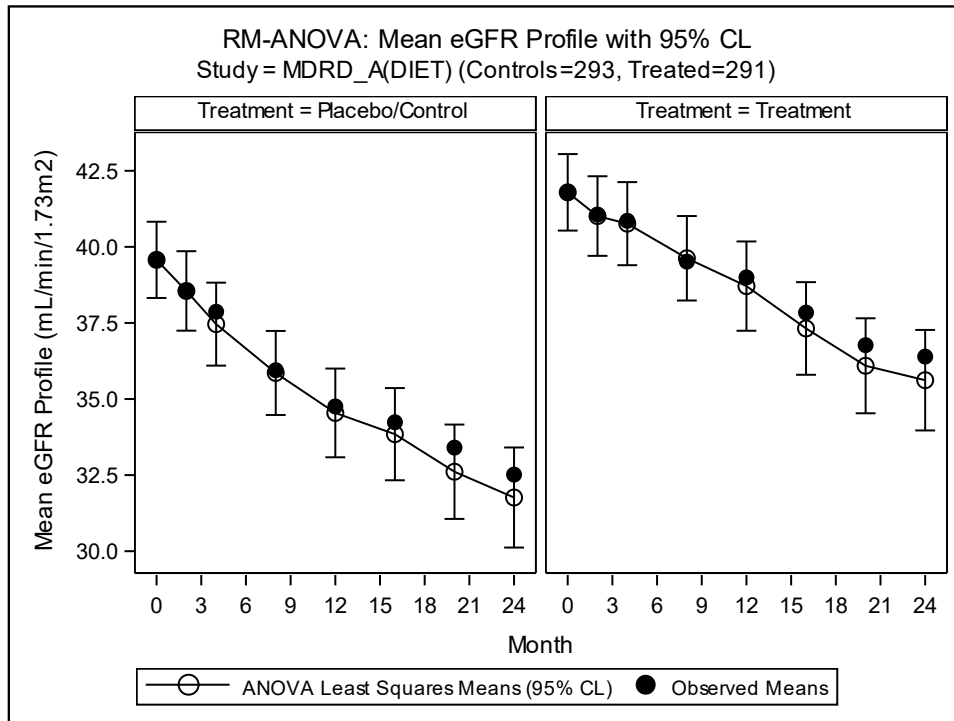
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 12
 Acute Effect (Treated-Control): 0.8092 (-1.263, 2.8815) based on actual intercepts
 Acute Effect (Treated-Control): 0.4857 (-0.651, 1.6224) based on equal intercepts
 Study = MDRD_A(BP) (Controls=285, Treated=299)

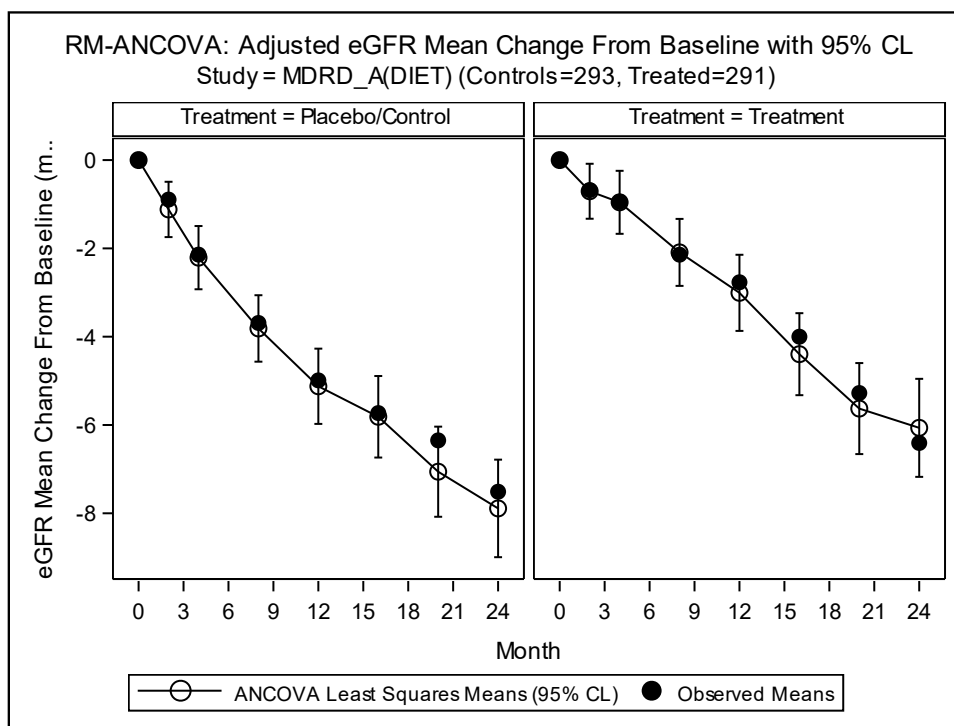
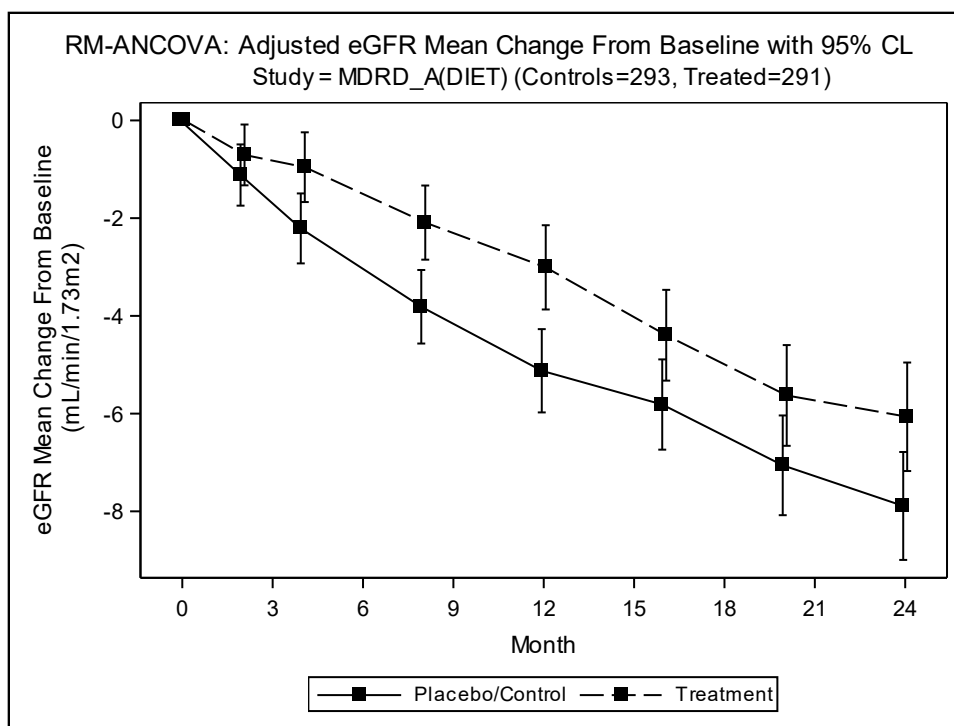


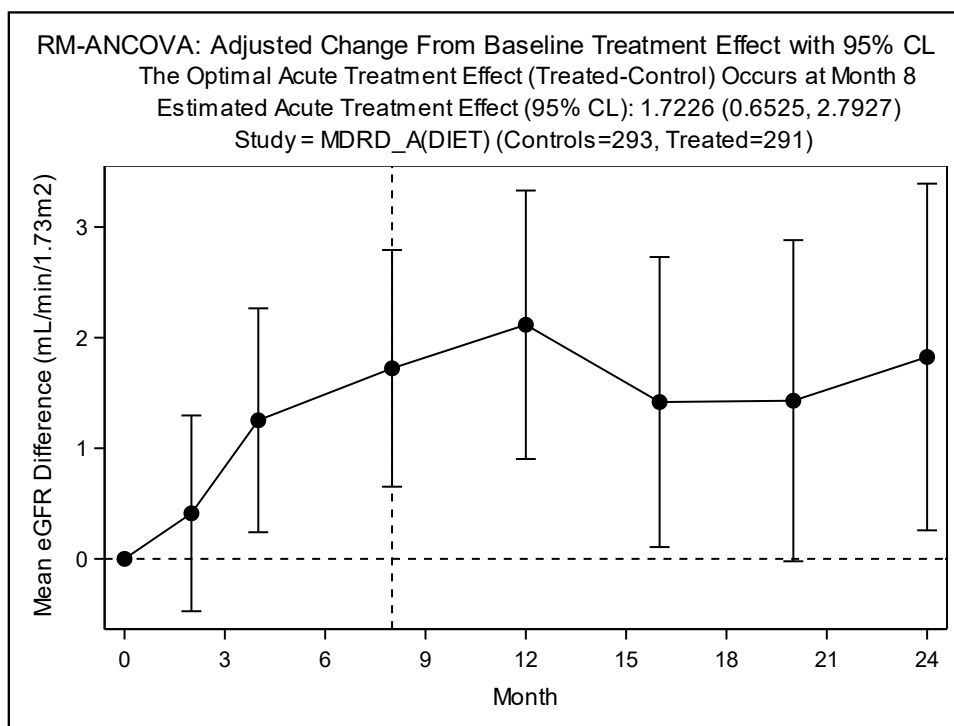
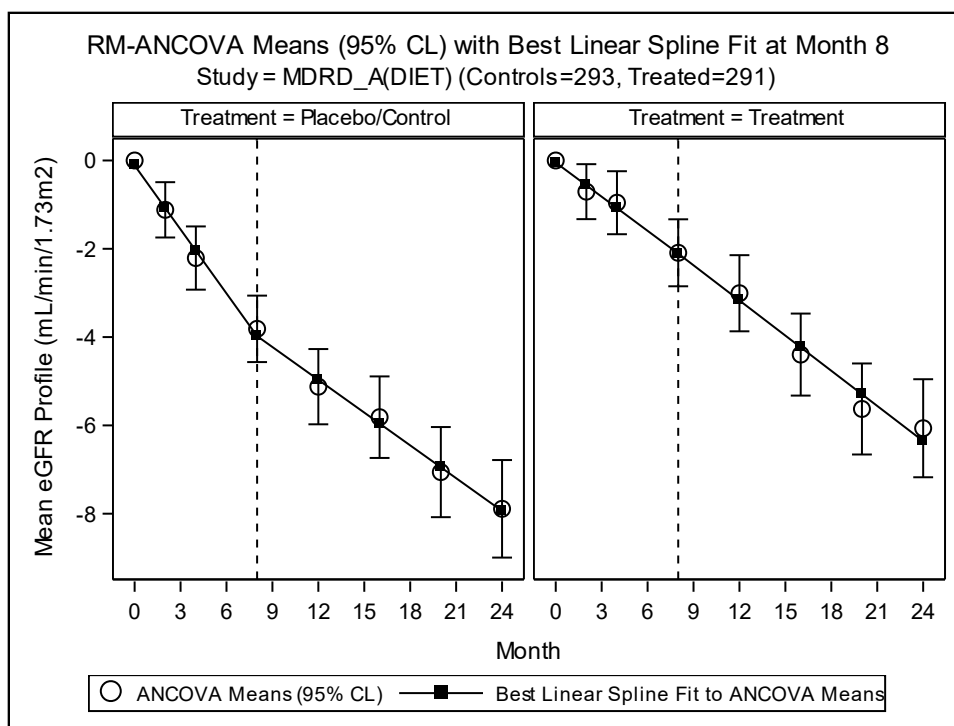
Linear Spline Mixed Model (95% CL) With Incremental Knot at Month 12
 Acute Effect (Treated-Control): 0.8092 (-1.263, 2.8815) based on actual intercepts
 Acute Effect (Treated-Control): 0.4857 (-0.651, 1.6224) based on equal intercepts
 Study = MDRD_A(BP) (Controls=285, Treated=299)



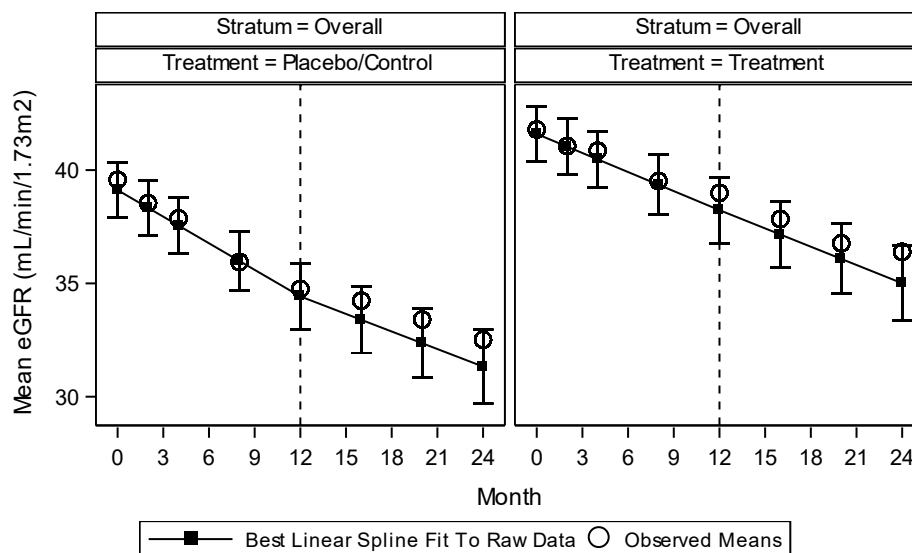
Appendix 6: MDRD-A(DIET) results



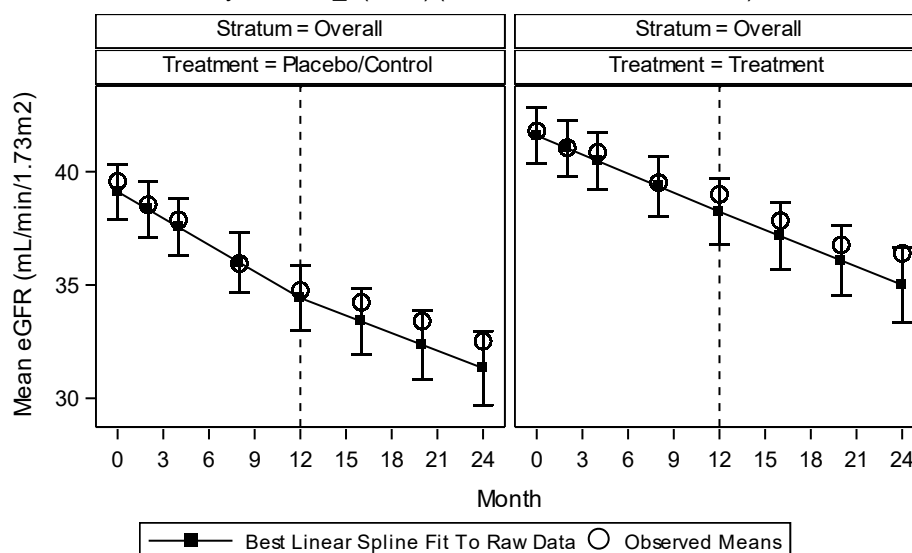




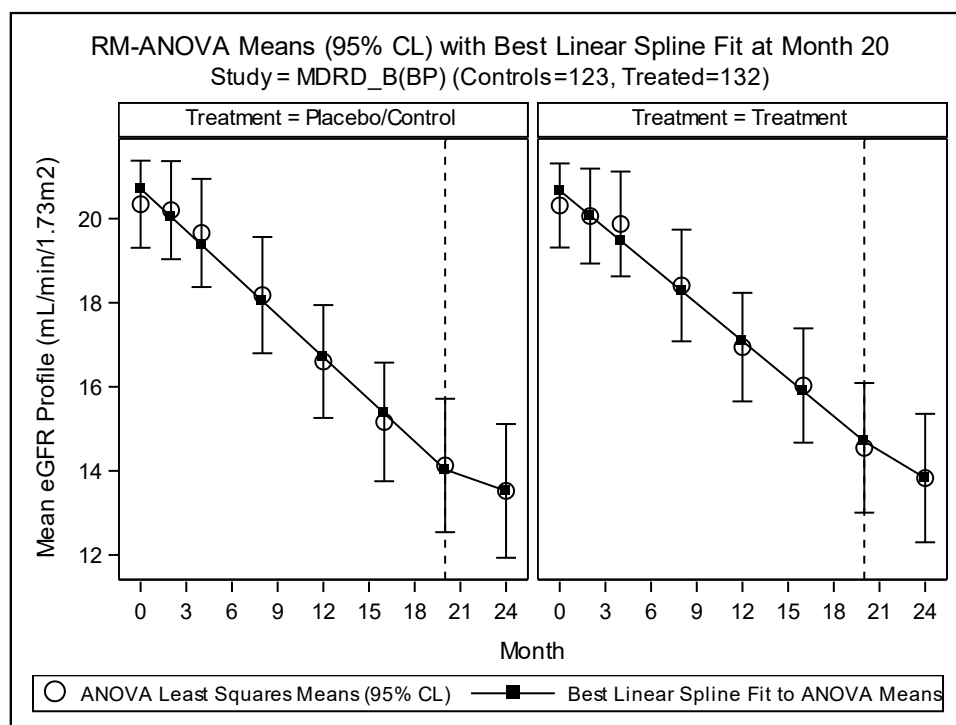
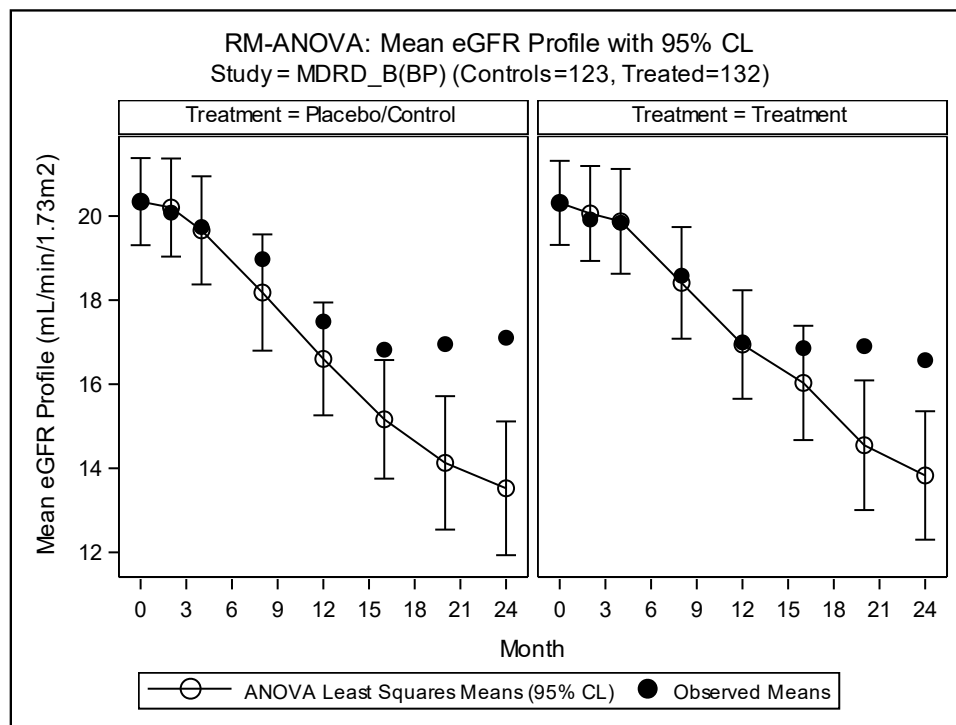
Linear Spline Mixed-Effects Model (95% CL) With Observed Knot at Month 12
 Acute Effect (Treated-Control): 3.8075 (1.758, 5.8571) based on actual intercepts
 Acute Effect (Treated-Control): 1.3289 (0.1962, 2.4617) based on equal intercepts
 Study = MDRD_A(DIET) (Controls=293, Treated=291)

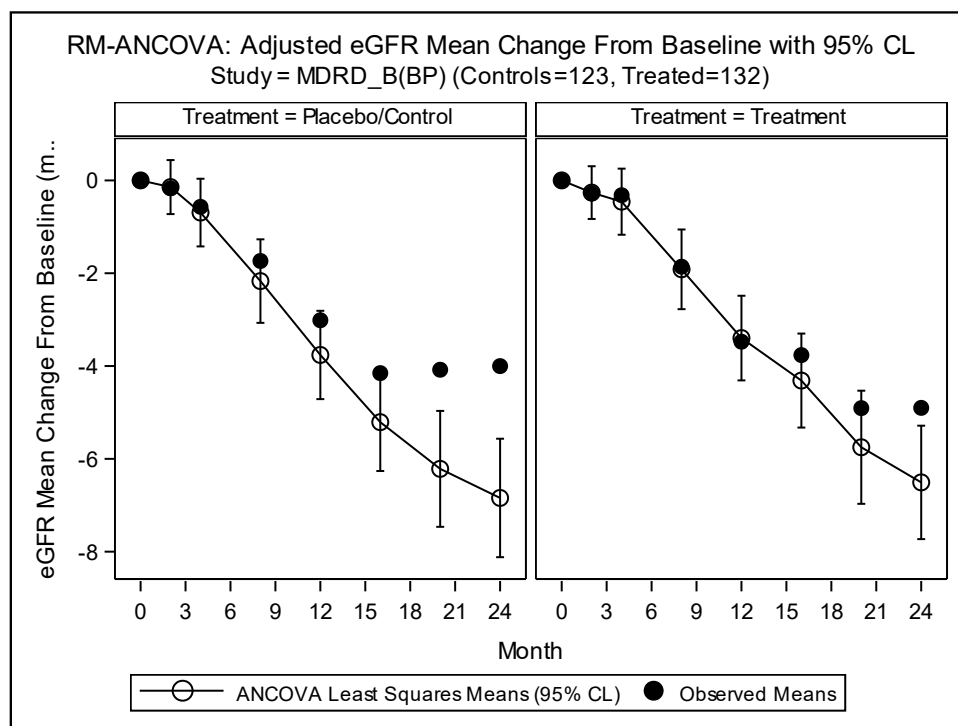
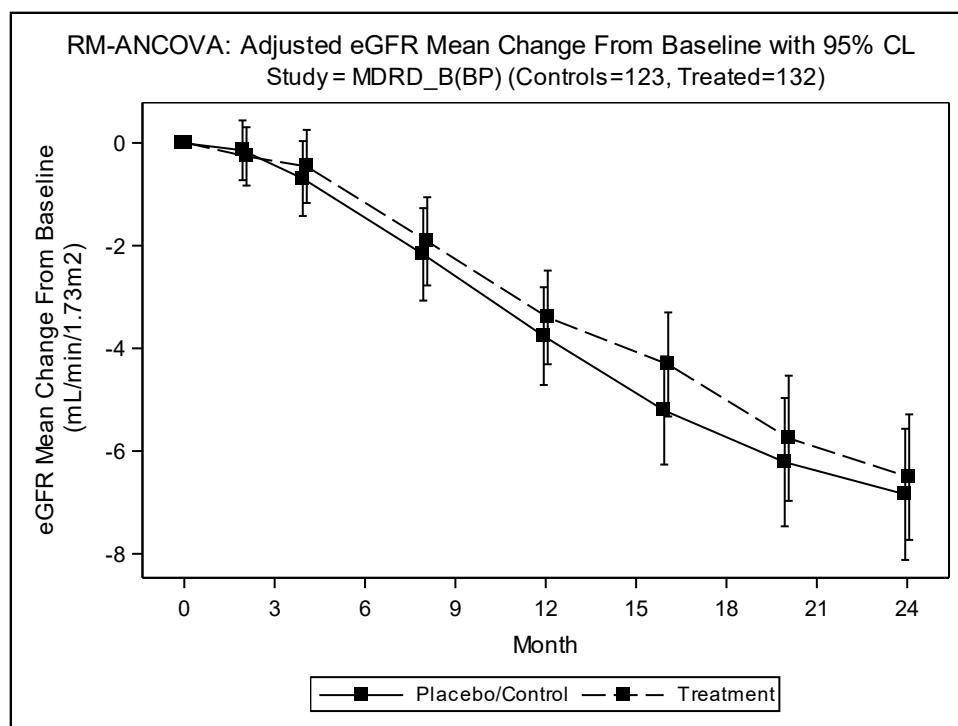


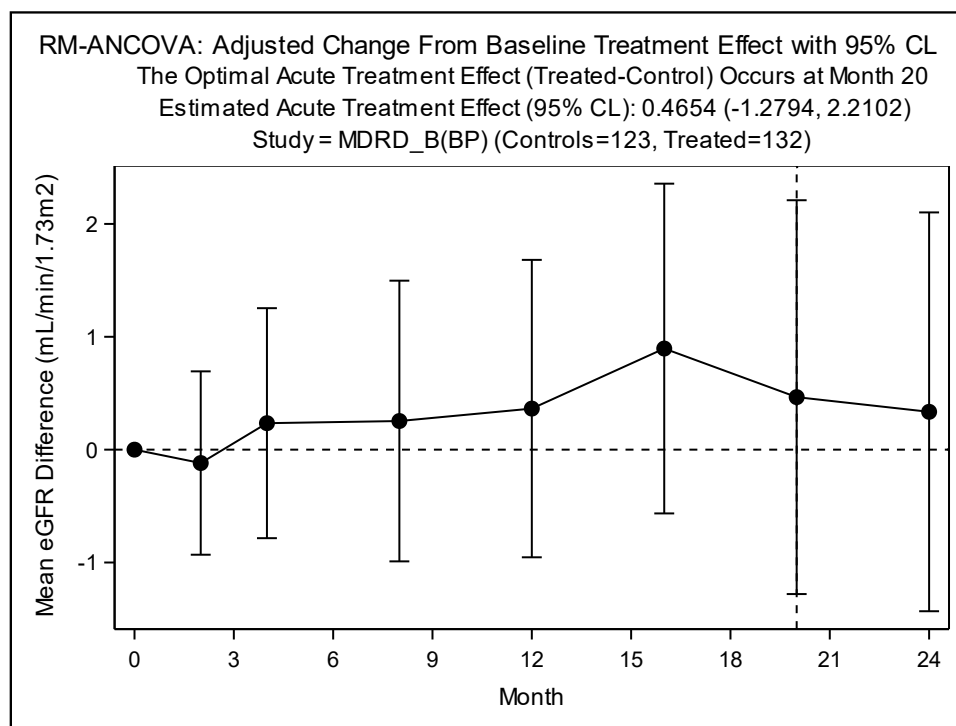
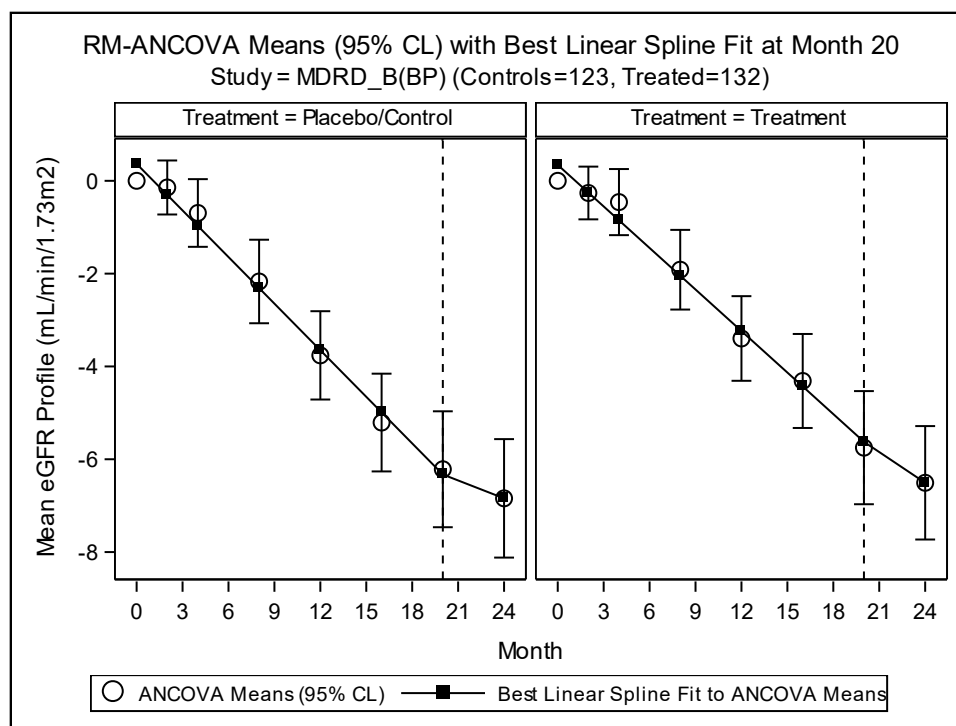
Linear Spline Mixed Model (95% CL) With Incremental Knot at Month 12
 Acute Effect (Treated-Control): 3.8075 (1.758, 5.8571) based on actual intercepts
 Acute Effect (Treated-Control): 1.3289 (0.1962, 2.4617) based on equal intercepts
 Study = MDRD_A(DIET) (Controls=293, Treated=291)

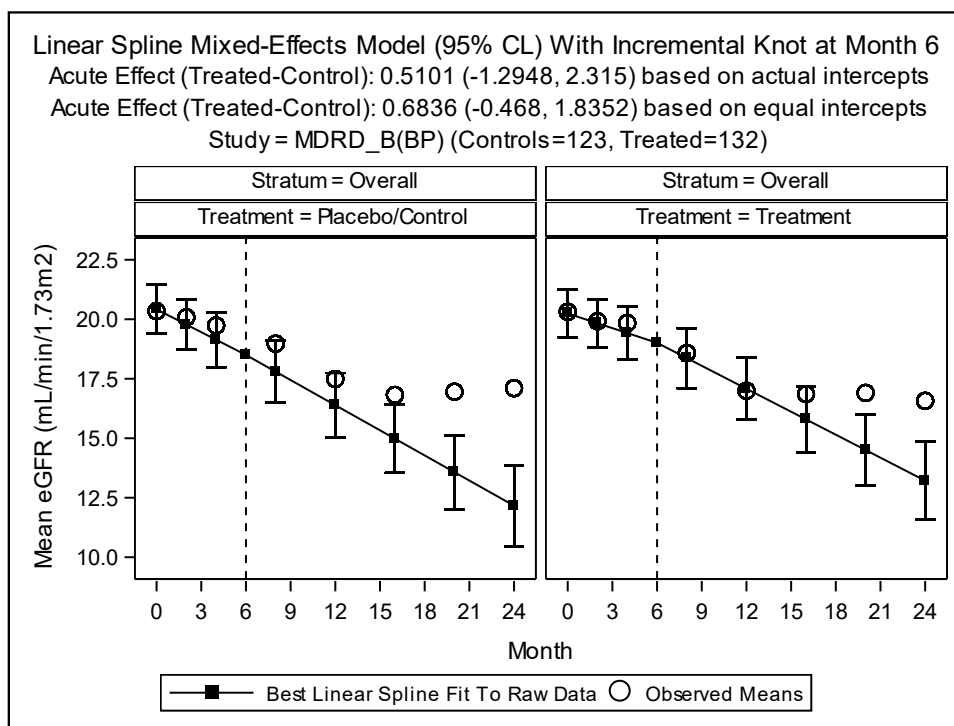
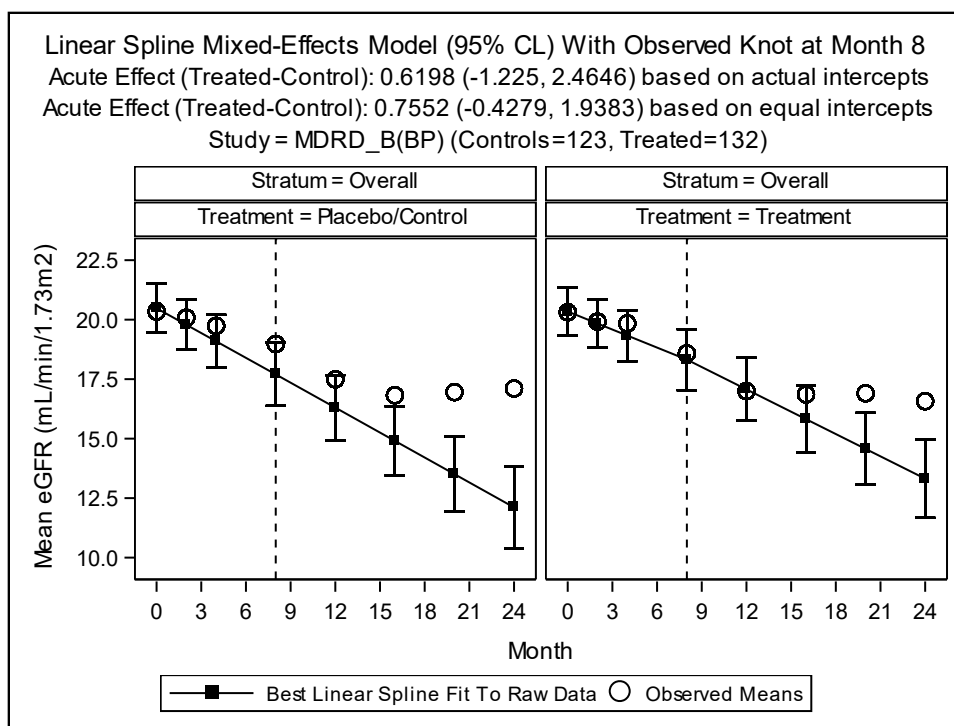


Appendix 7: MDRD-B(BP) results

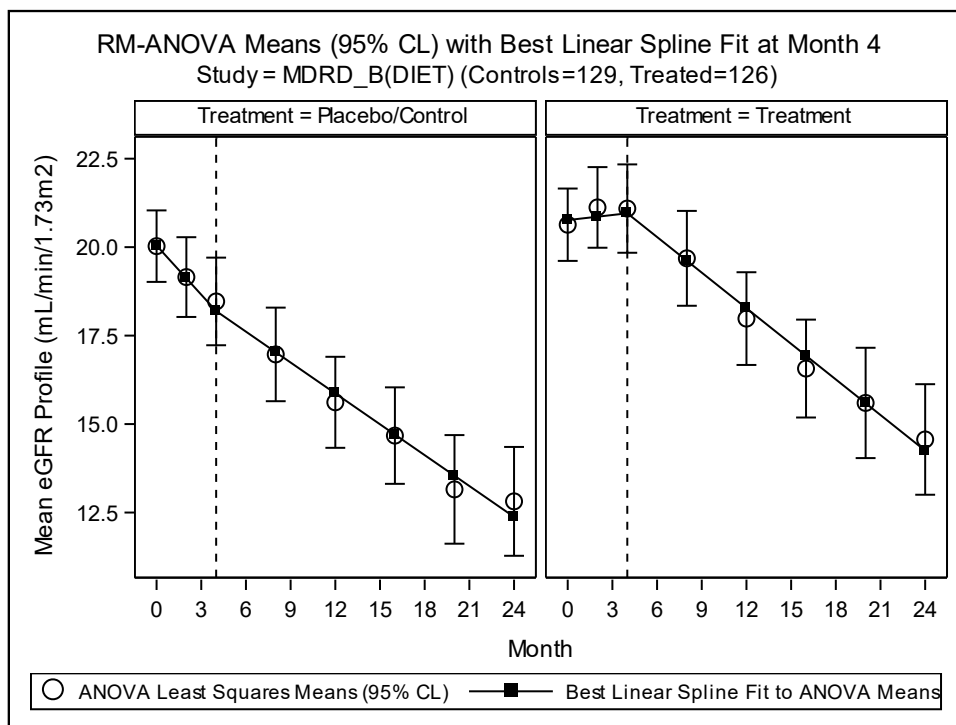
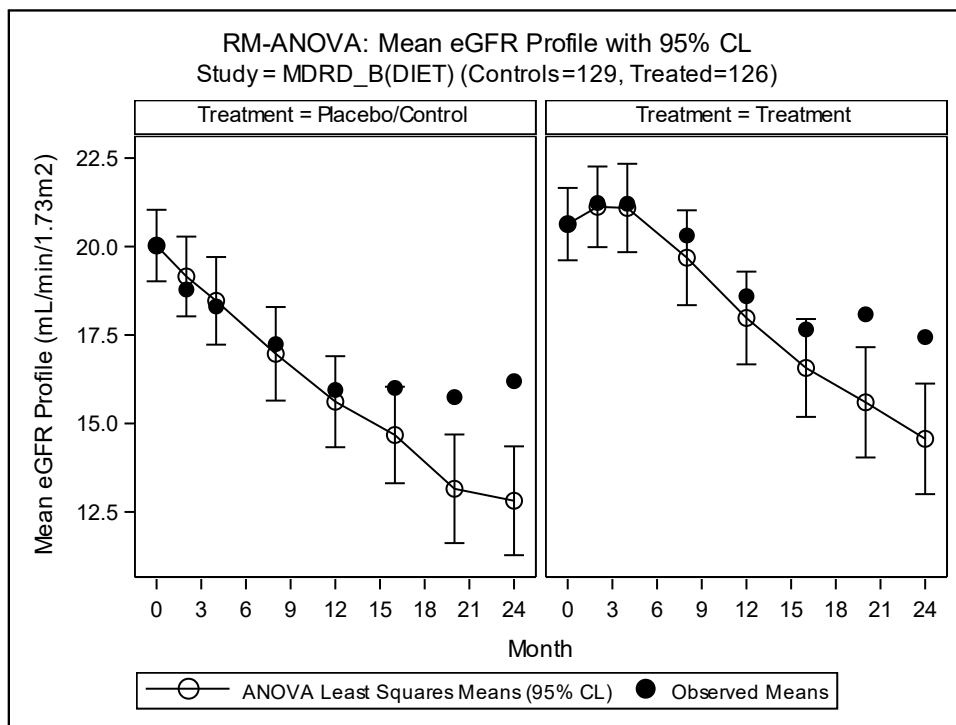


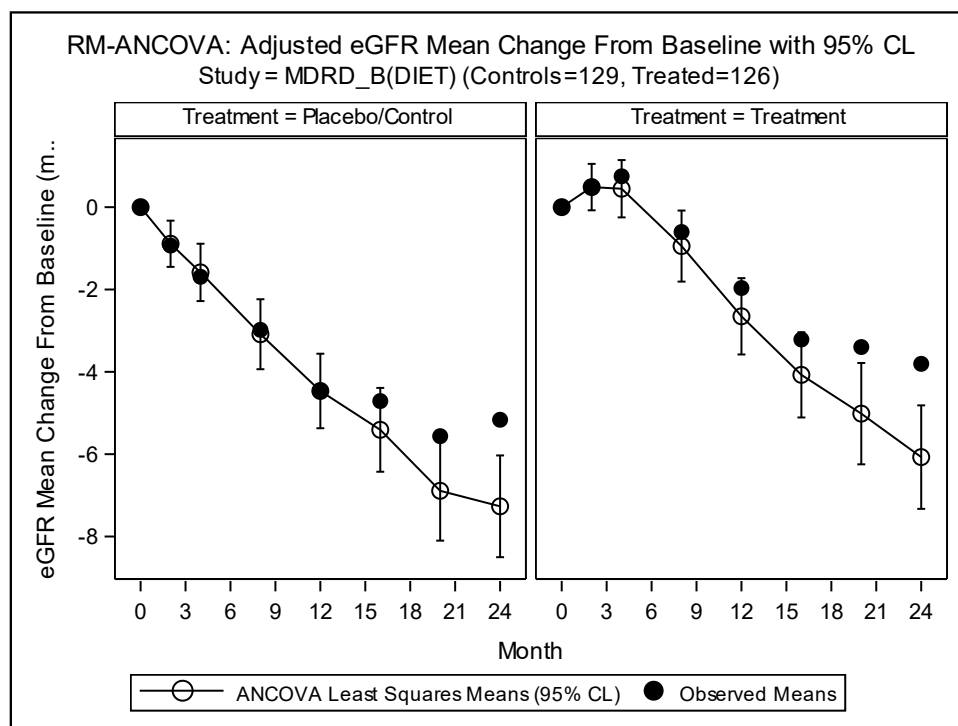
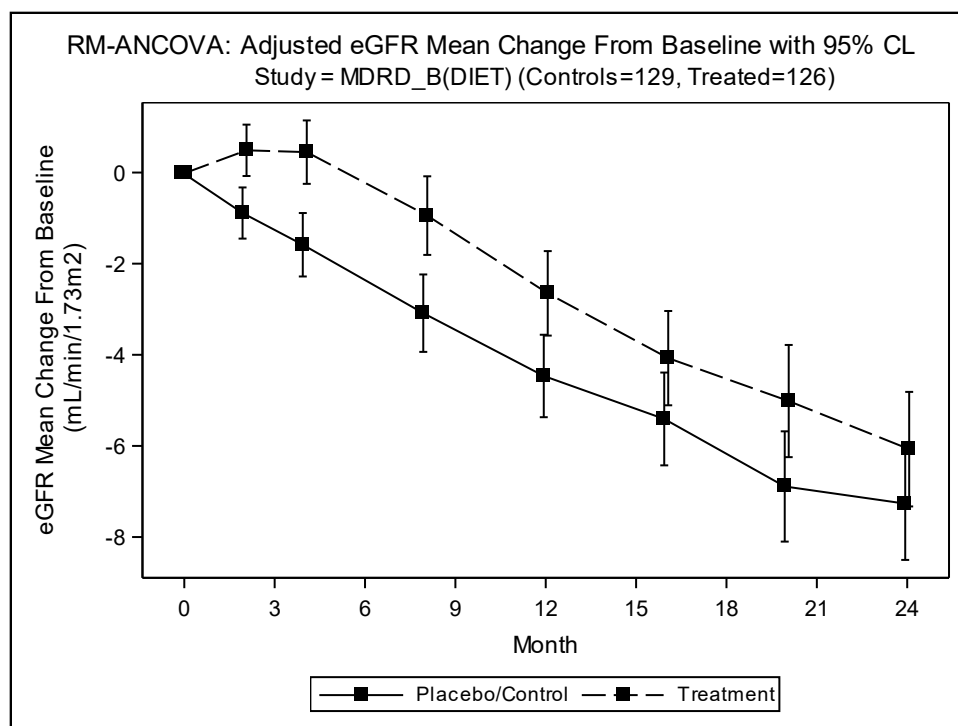


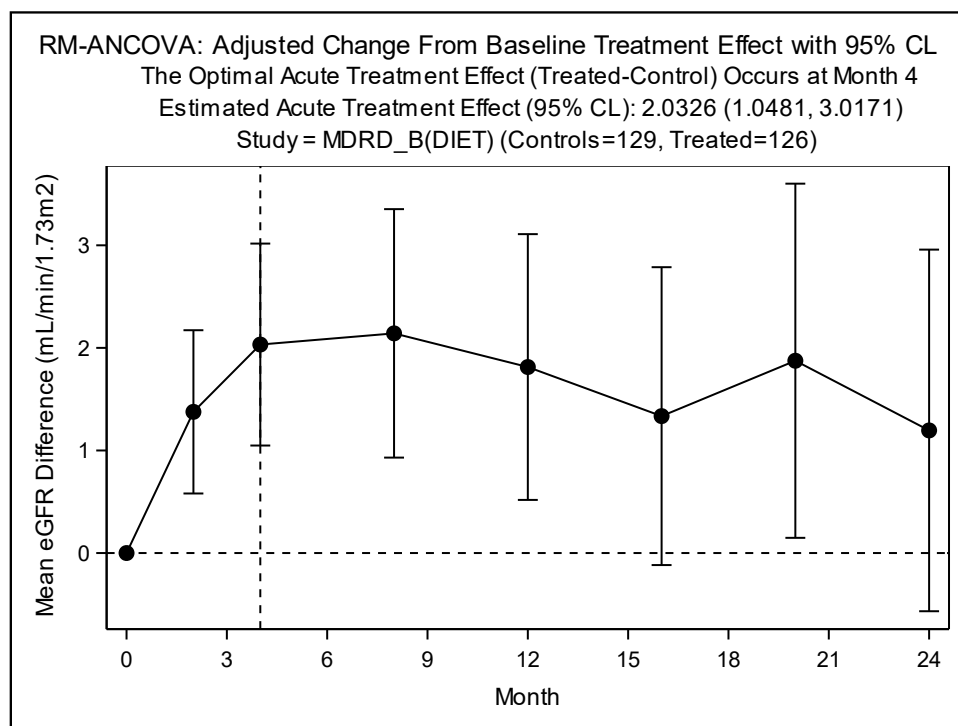
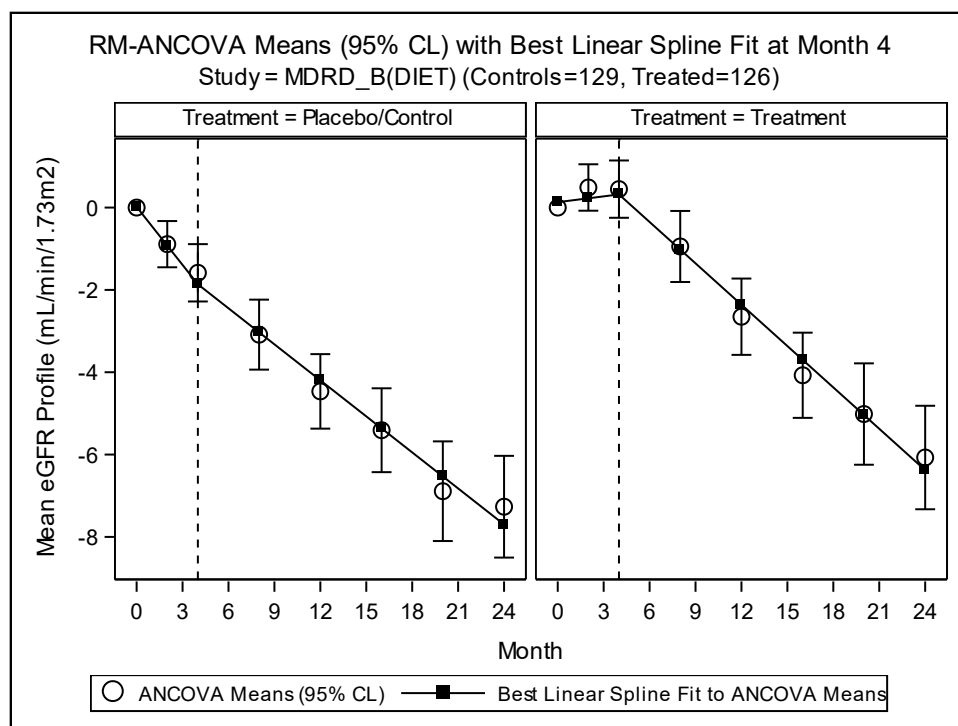


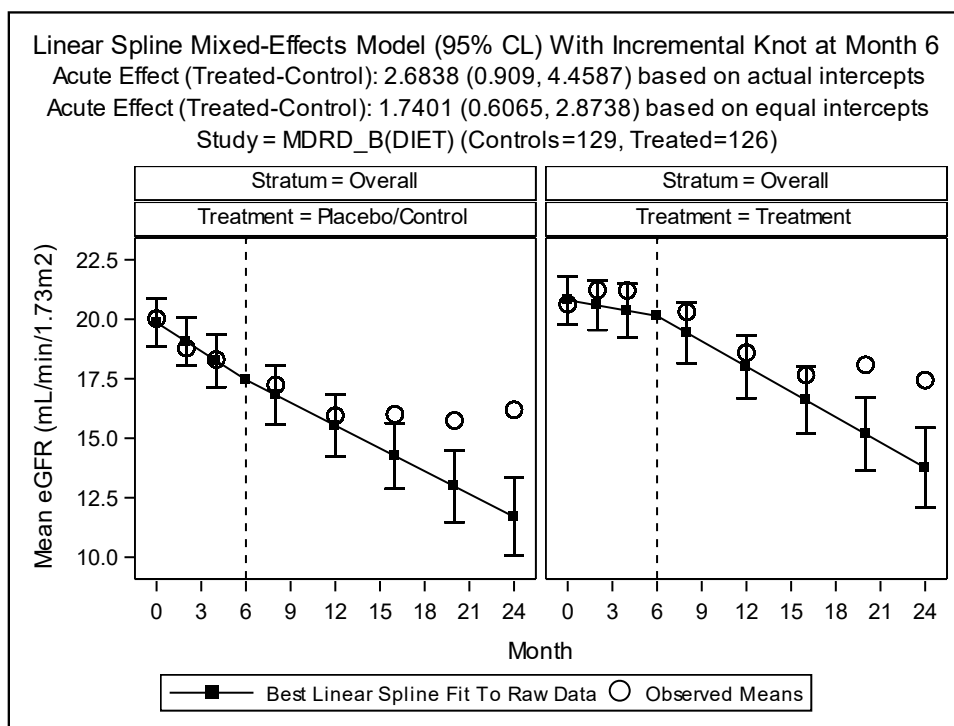
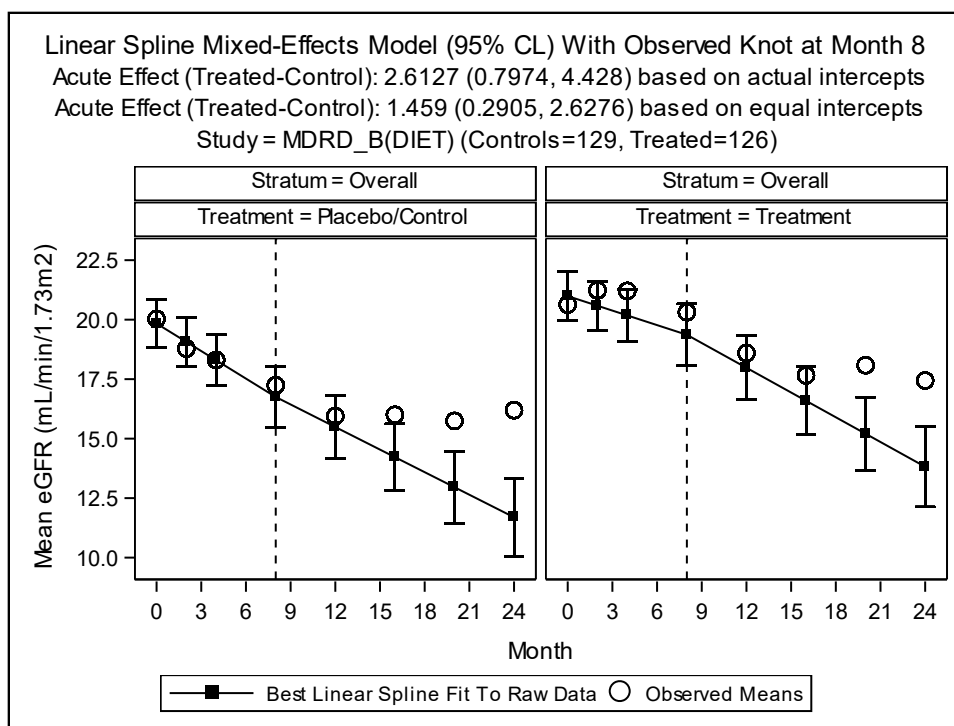


Appendix 8: MDRD-B(DIET) results

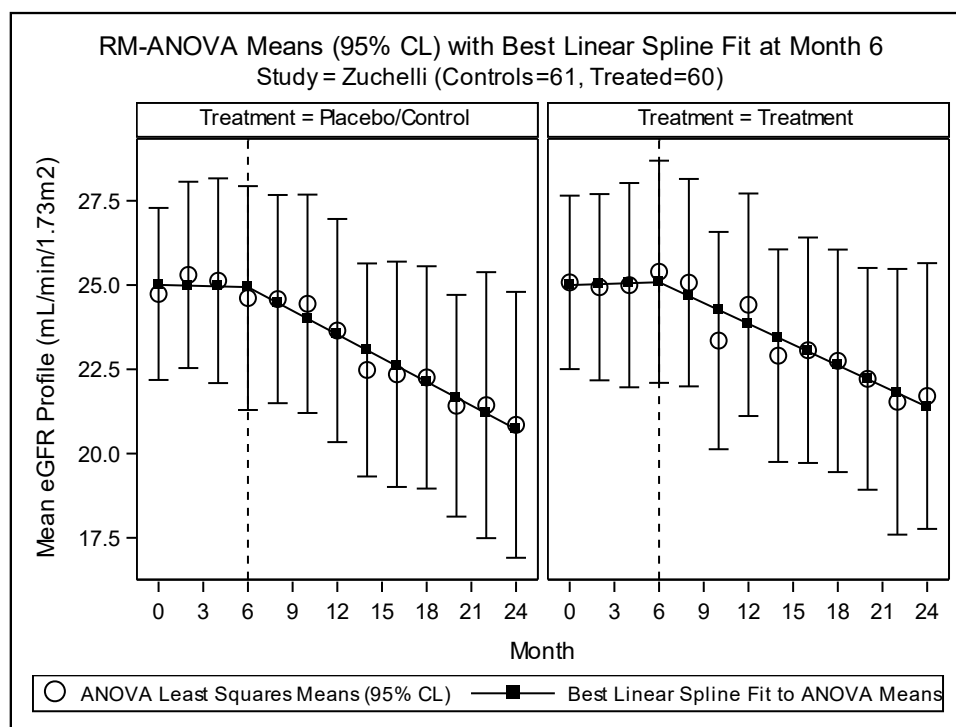
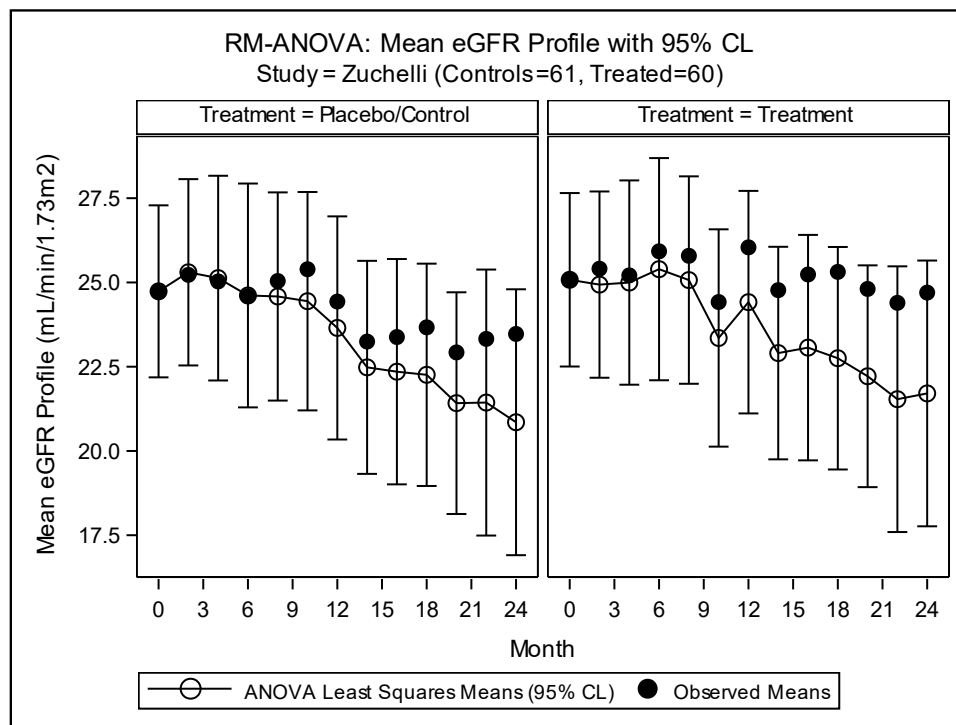


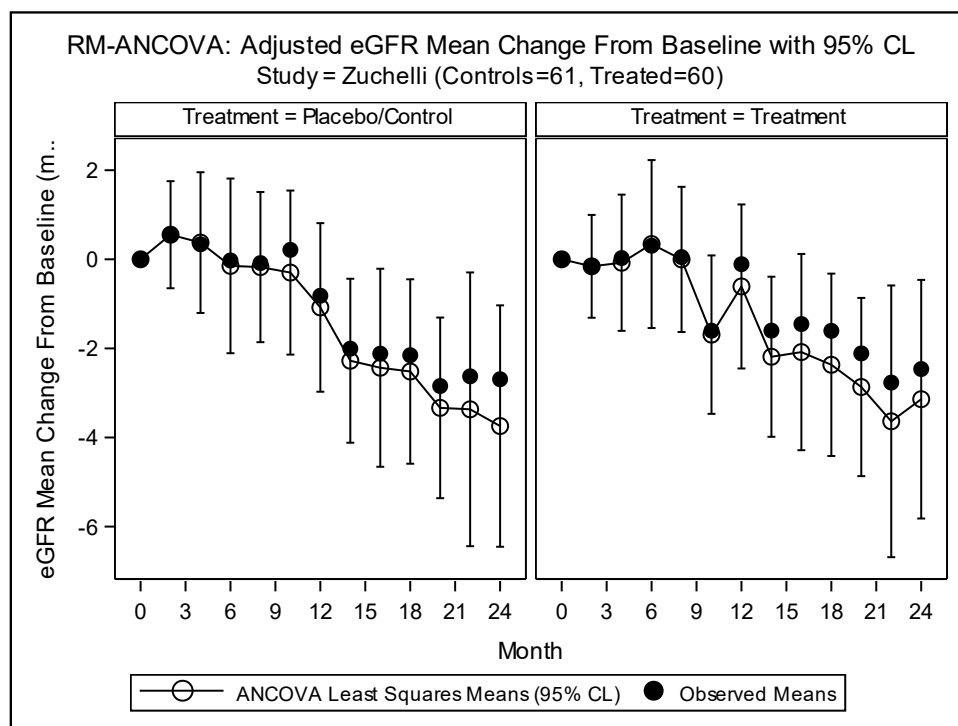
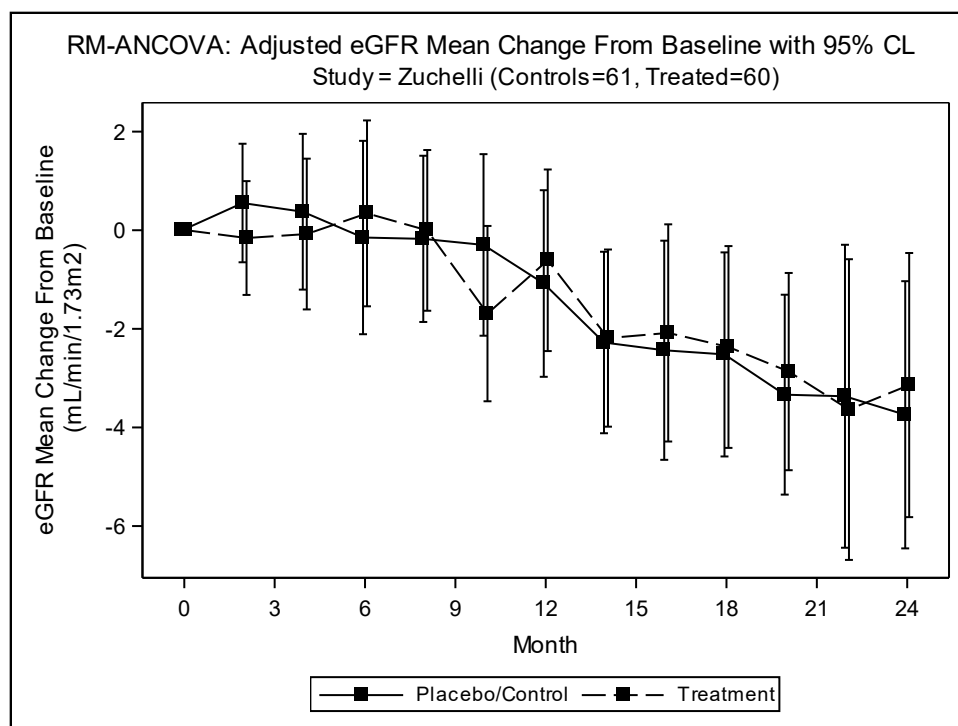


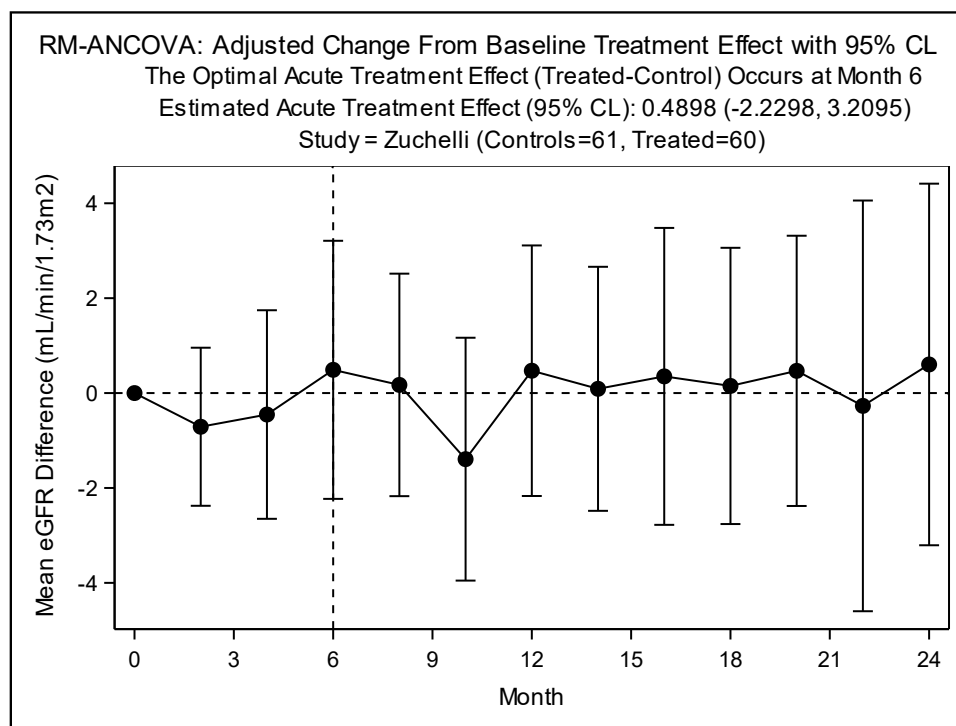
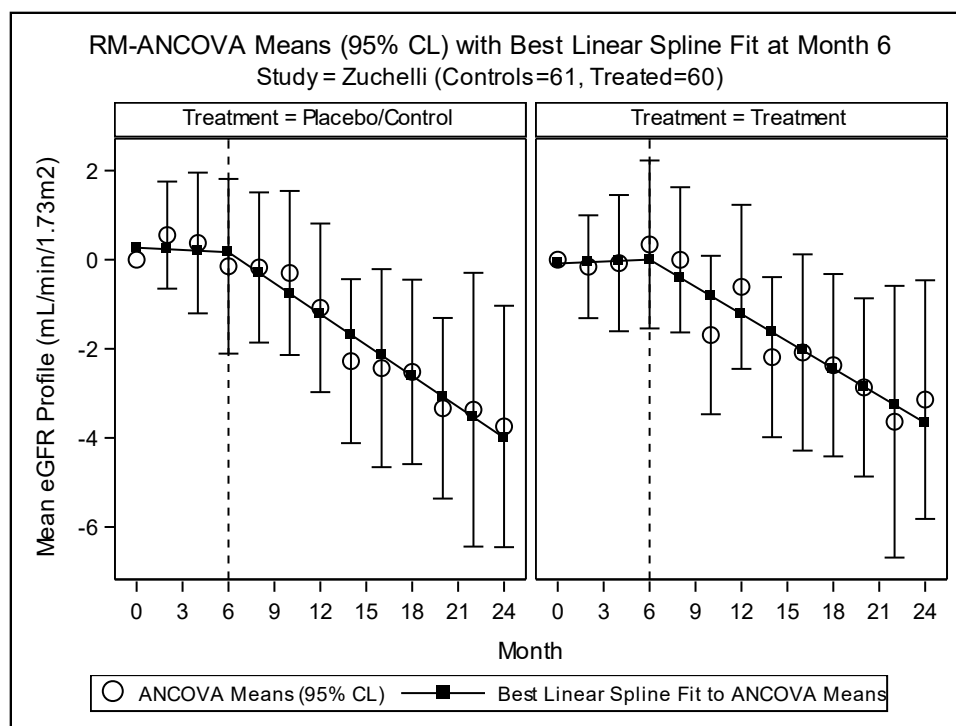


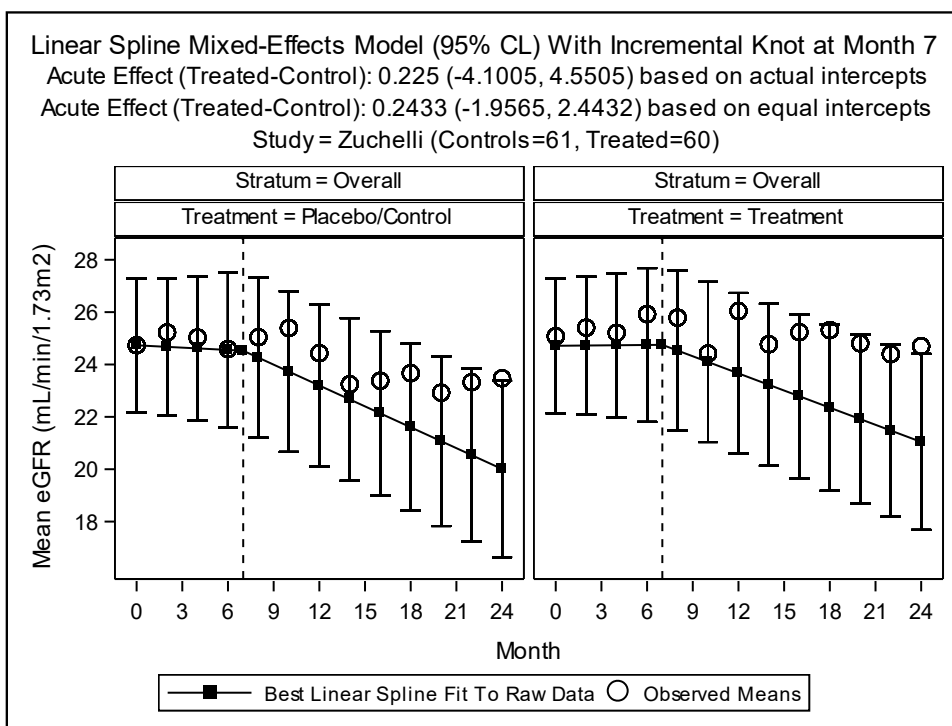
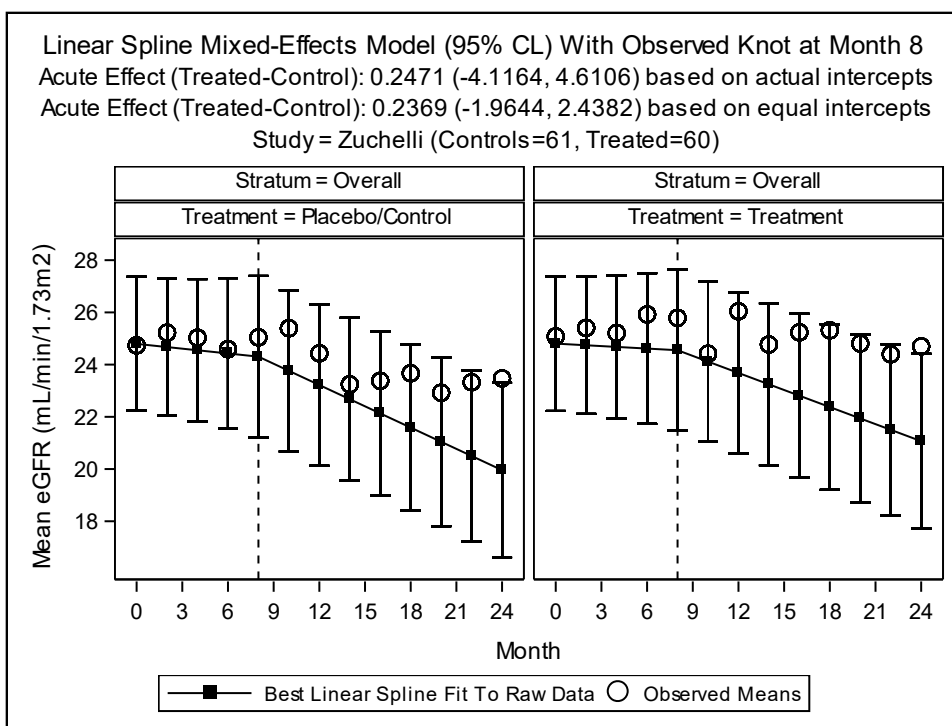


Appendix 9: Zuchelli results









Vonesh Statistical Consulting, LLC
1928 Forest Creek Lane
Libertyville, IL 60048

Phone: 847-367-1678,
Fax: 847-367-1802
Email: Edward@Vonesh-Stats.com

December 20, 2020

To: Lesley Inker, MD, MS
Tom Greene, PhD
Hocine Tighiouart, MS
Jian Ying, 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.

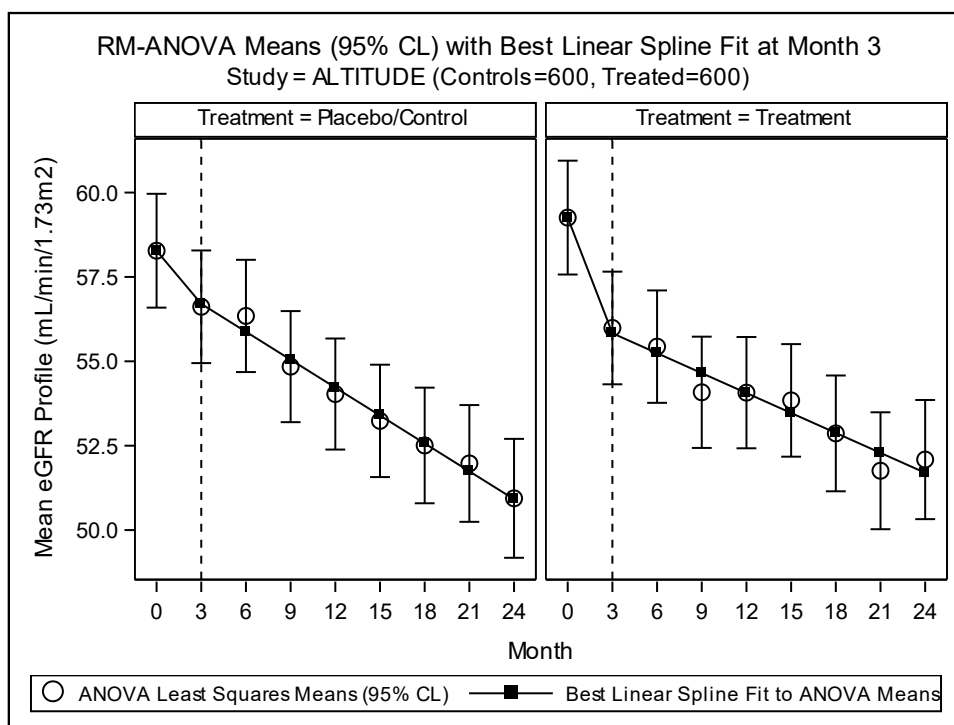
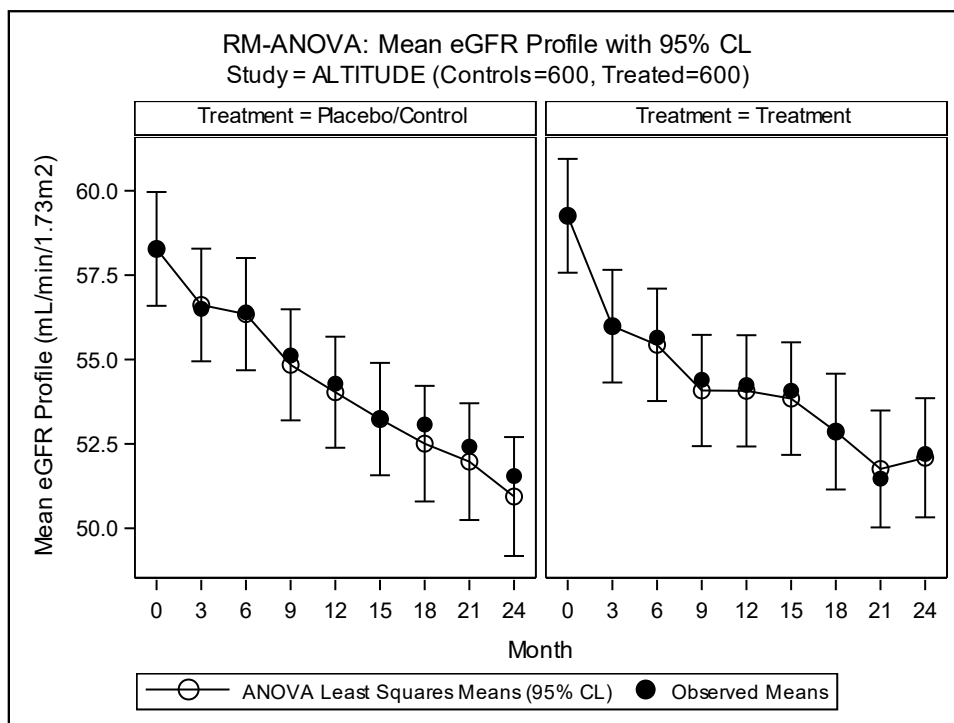
1. Introduction

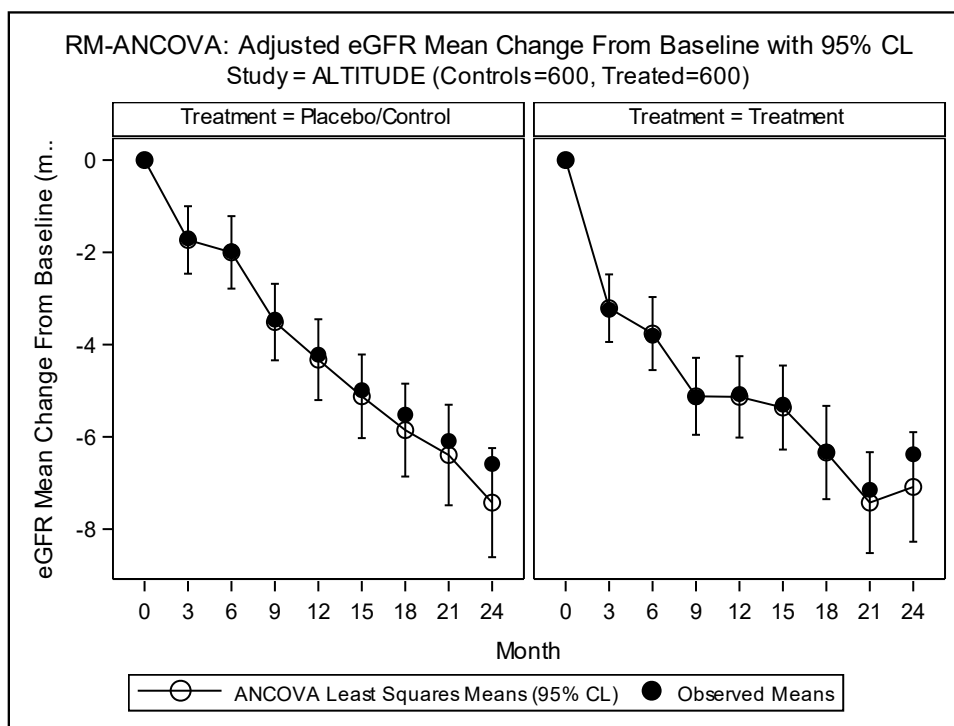
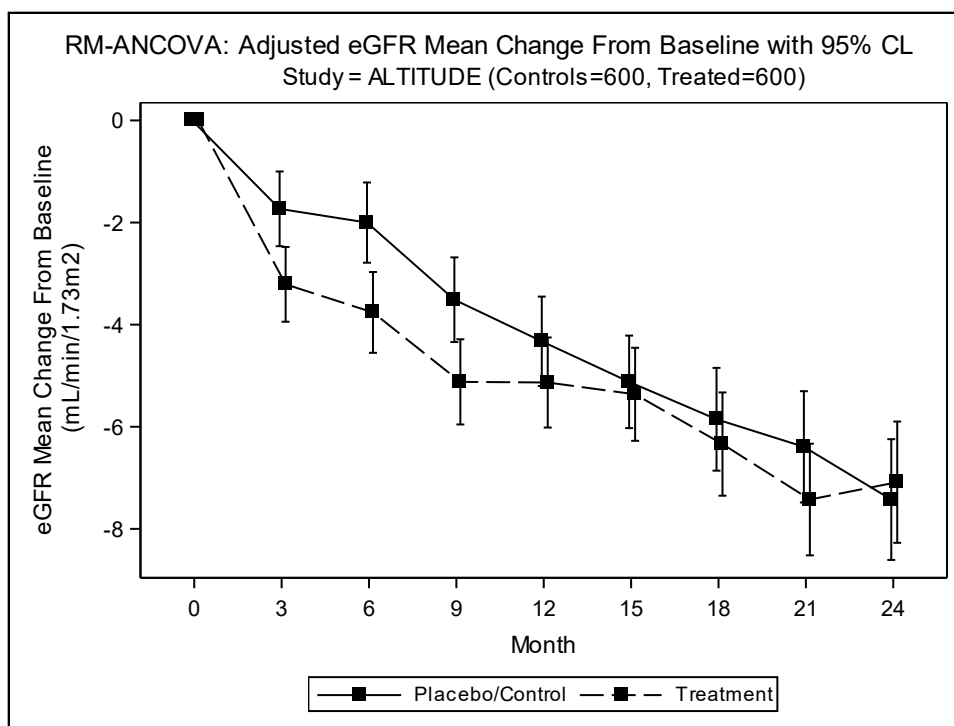
This amendment to the initial report gives results for the RM-ANOVA and RM-ANCOVA analyses in which a weighted linear spline regression model was fit to the least squares means with weights given by the squared inverse of the RM-ANOVA (or RM-ANCOVA) standard errors. This provides a best linear unbiased estimate of the linear spline parameters based on the estimated variances of the RM-ANOVA and RM-ANCOVA least square means. The results are graphically displayed in the attached Appendices 1a – 9a corresponding to the 9 selected studies. While the regression parameters from the weighted linear spline regression analysis differ slightly from those given by the unweighted linear spline regression analysis, there are no perceptible differences in the prediction lines (see attached graphs) nor has there been any change in the knot selection or in the estimated acute treatment effects which are based on the least squares mean differences and their standard errors.

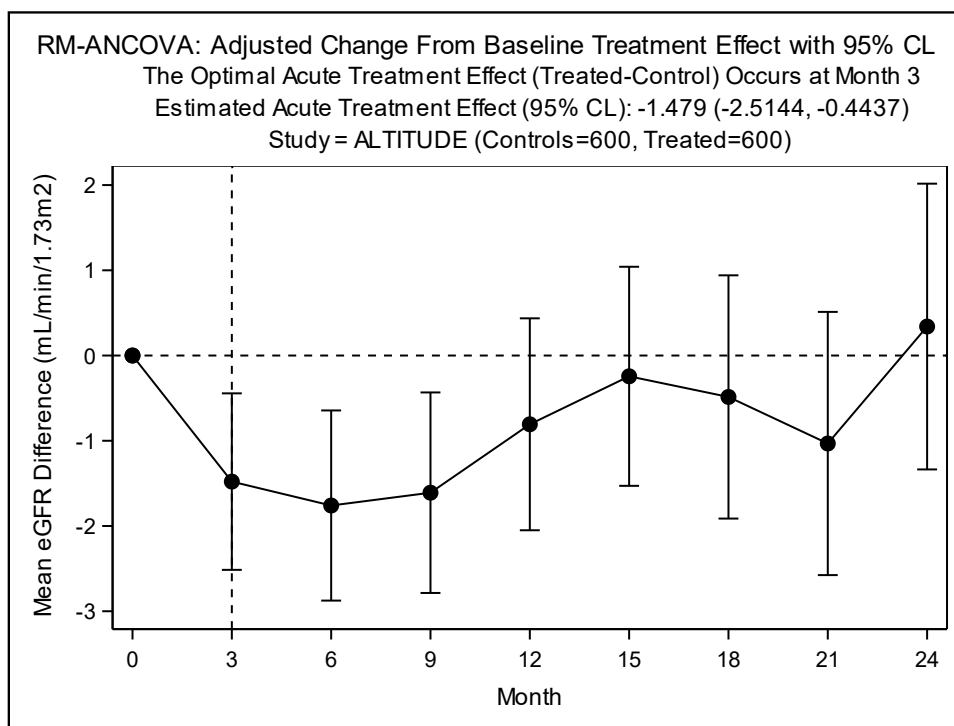
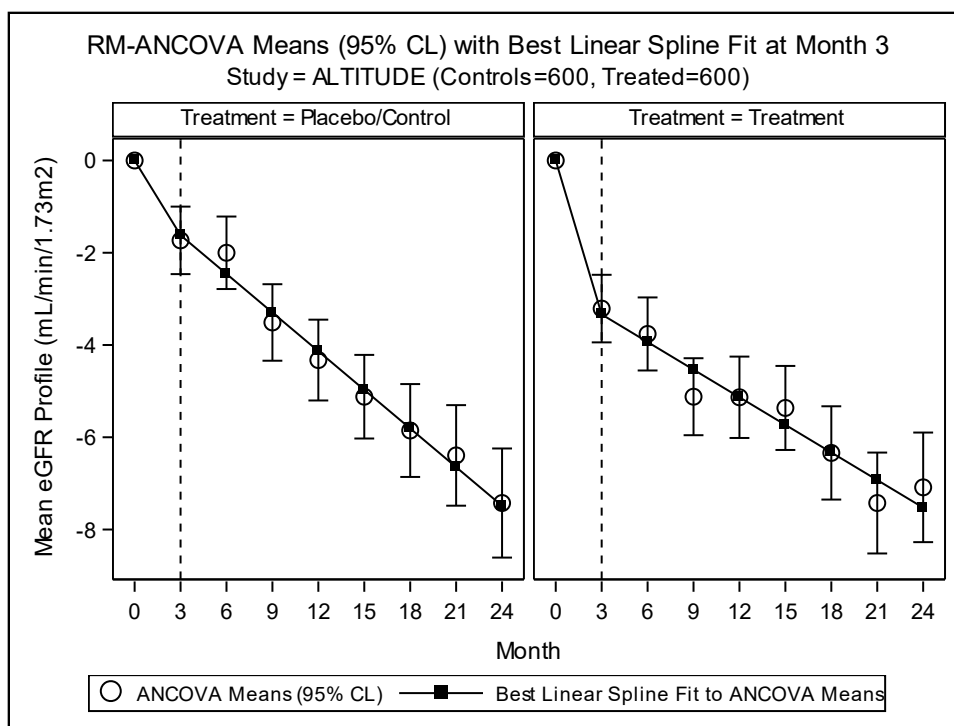
2. Conclusions

This new analysis supports the original findings as it pertains to the two-step repeated measures profile analysis using either a RM-ANOVA model or a RM-ANCOVA model. Results for the linear spline-mixed-effects analyses remain unchanged from the original report.

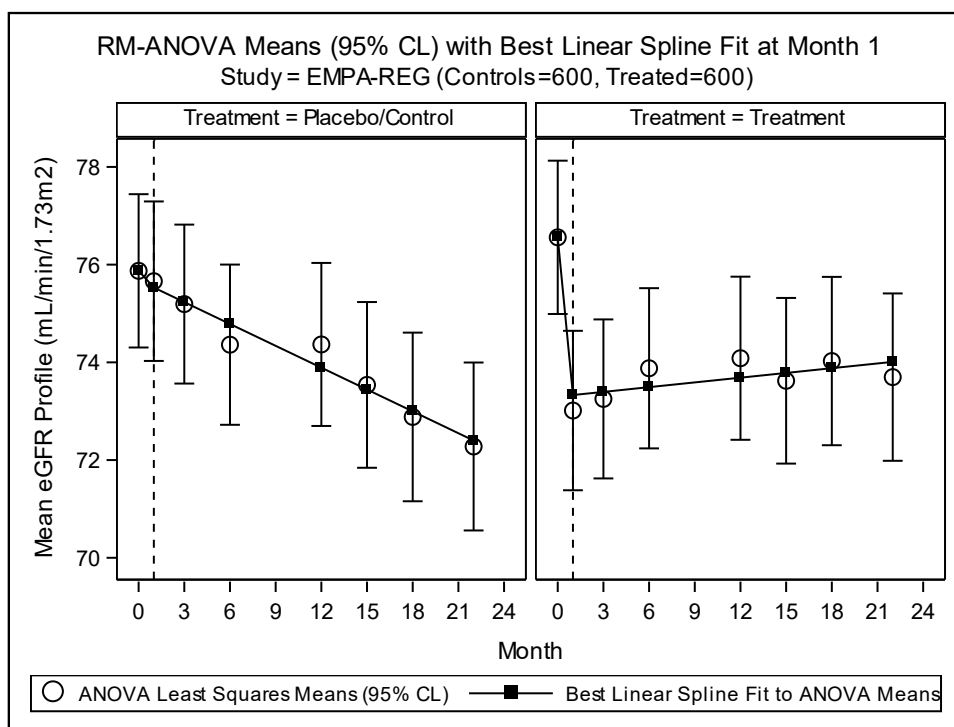
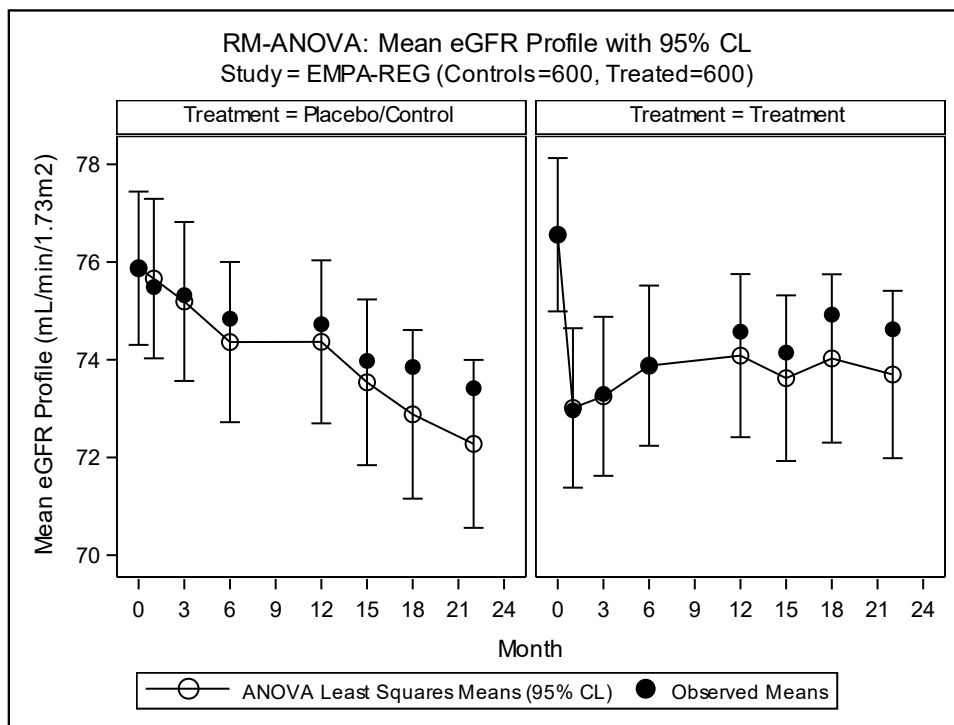
Appendix 1a: ALTITUDE results
Case 1: Sample Size=600 per Treatment group

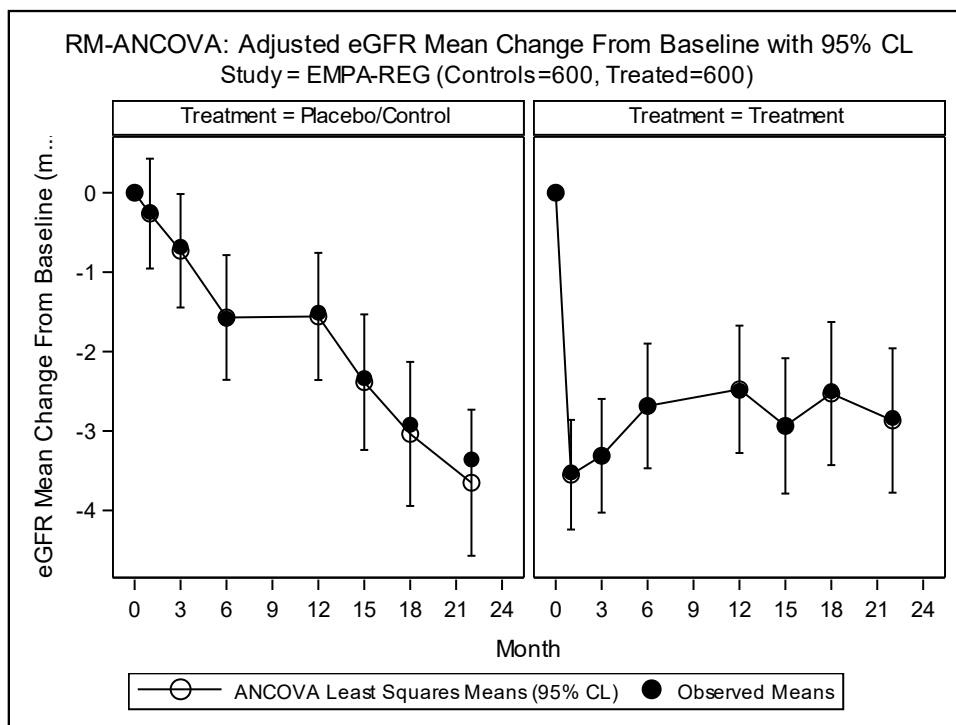
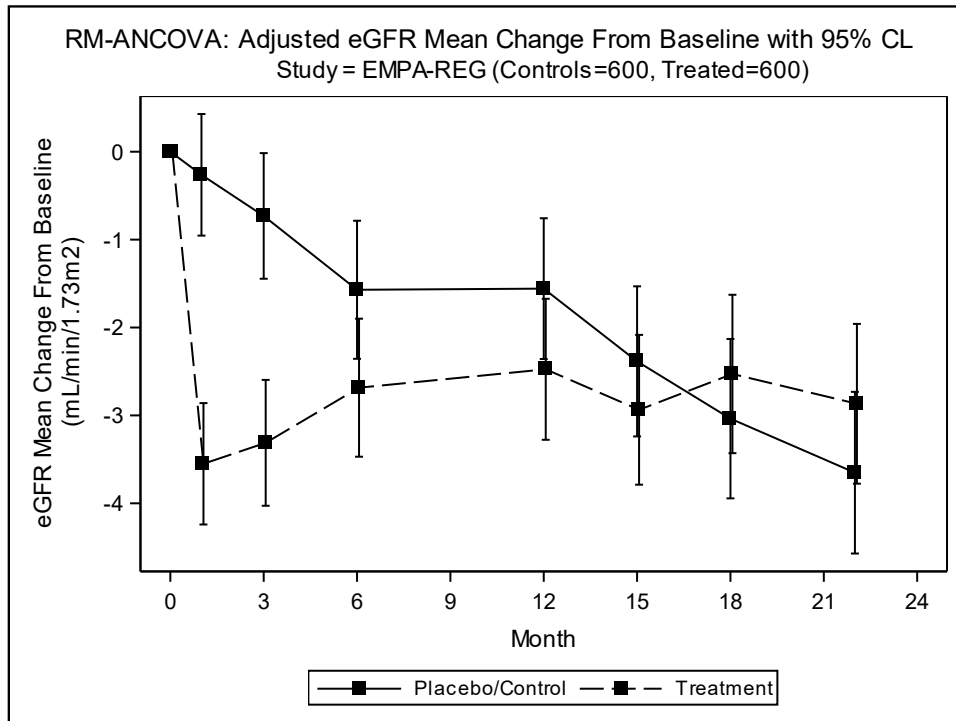


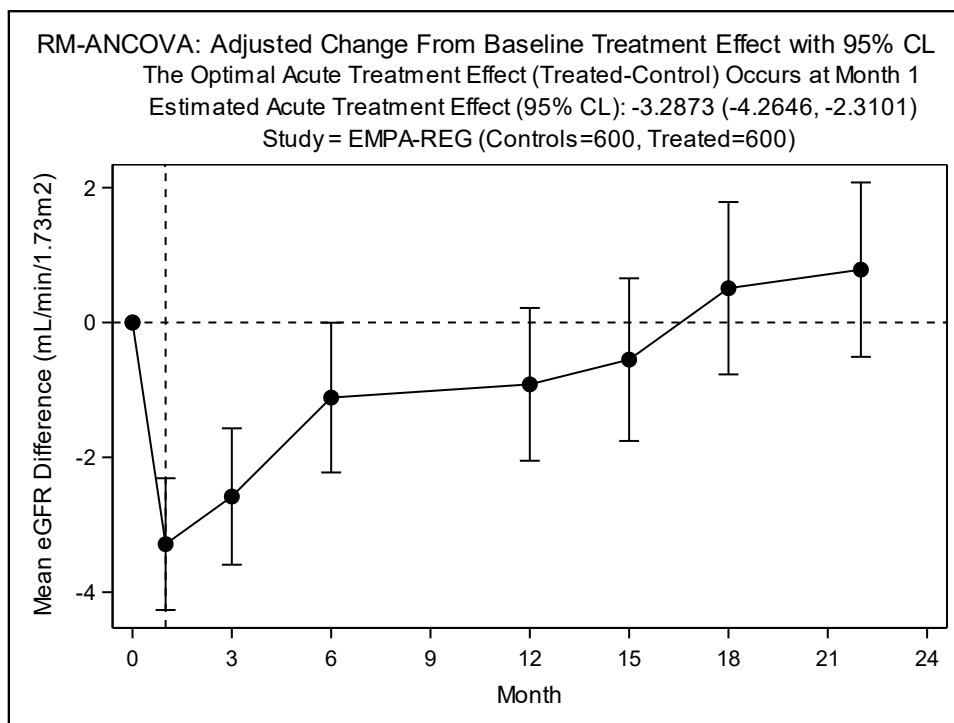
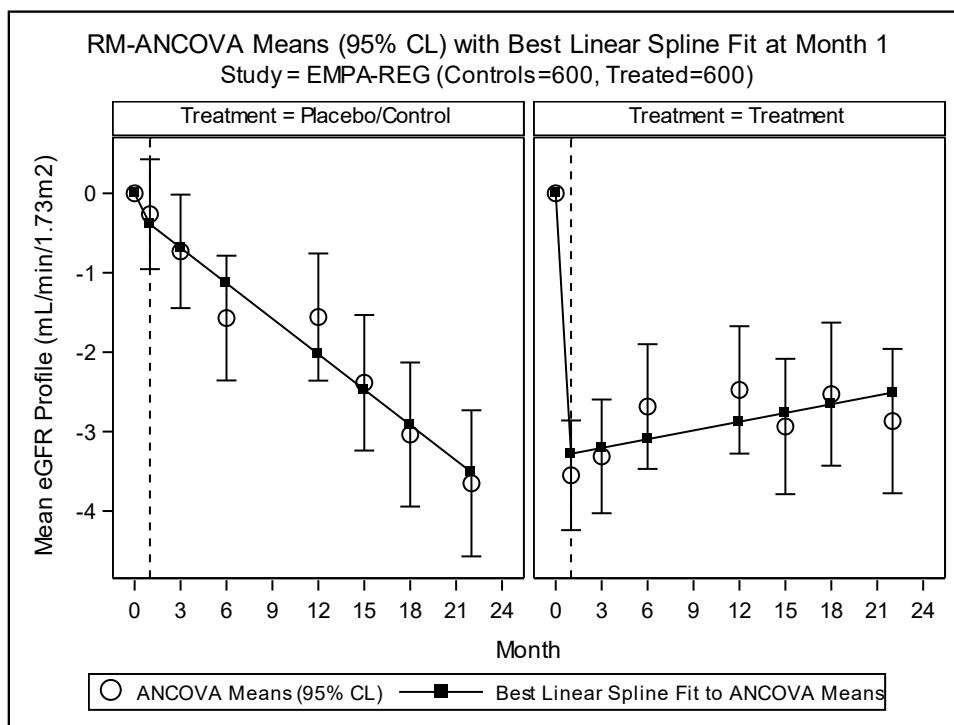




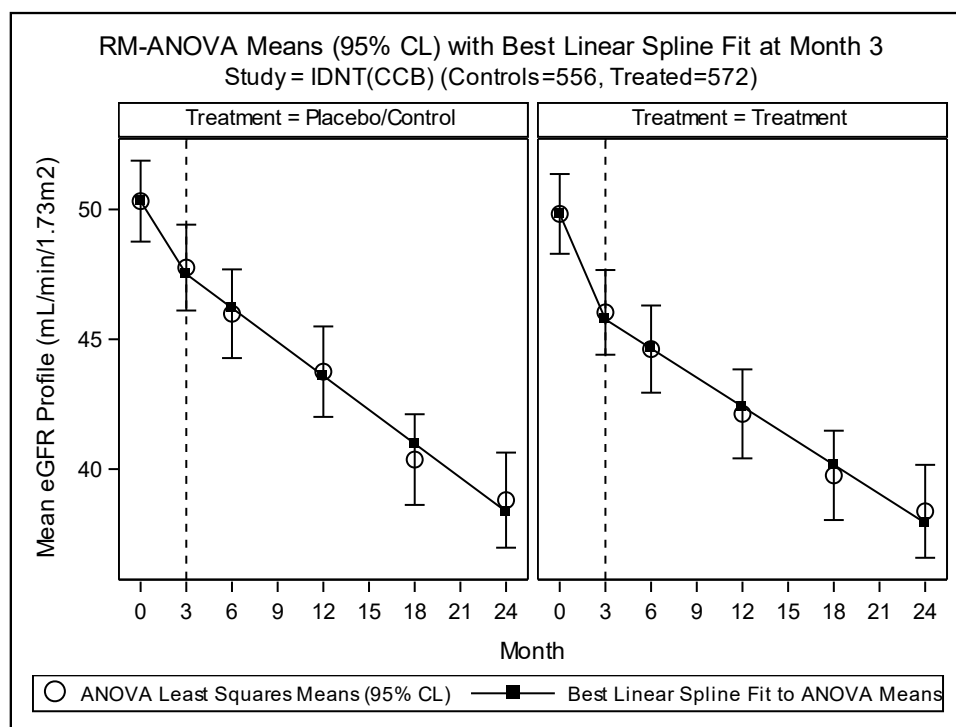
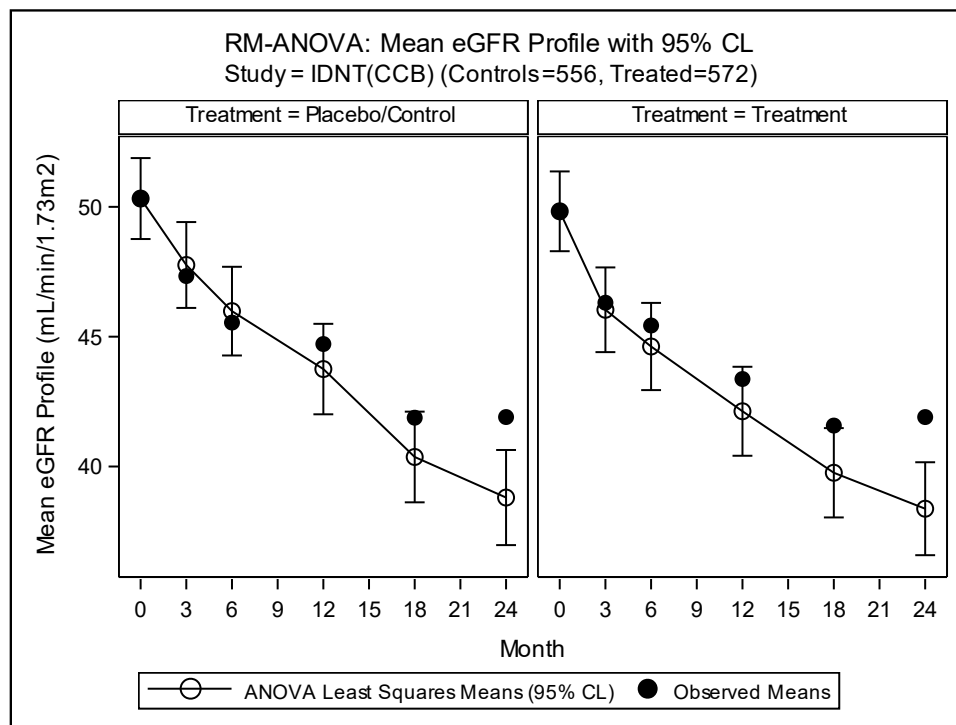
Appendix 2a: EMPA-REG results

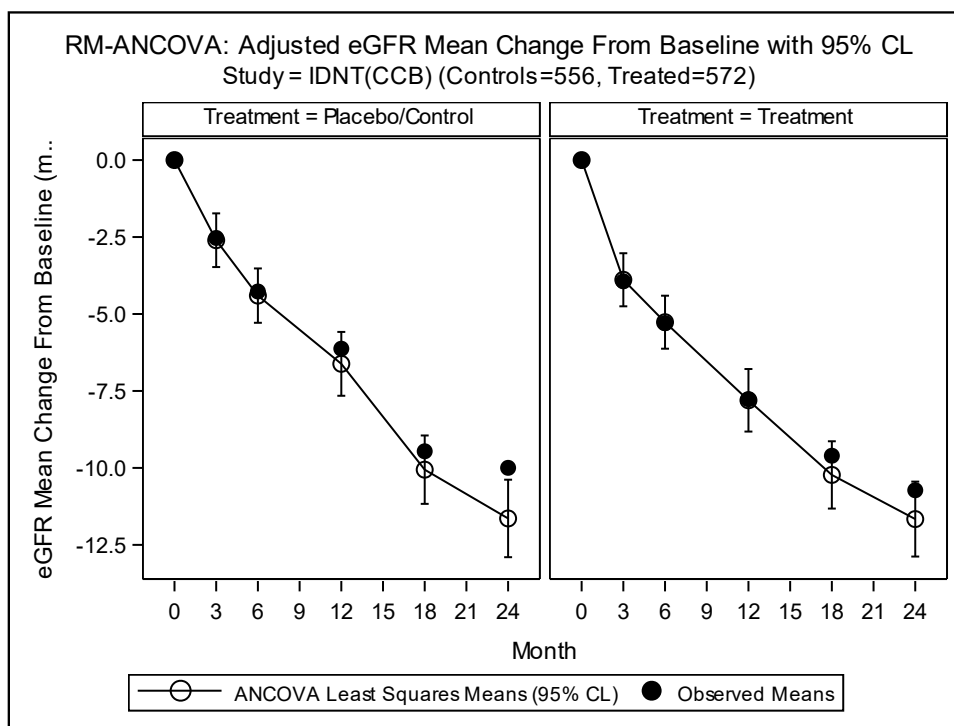
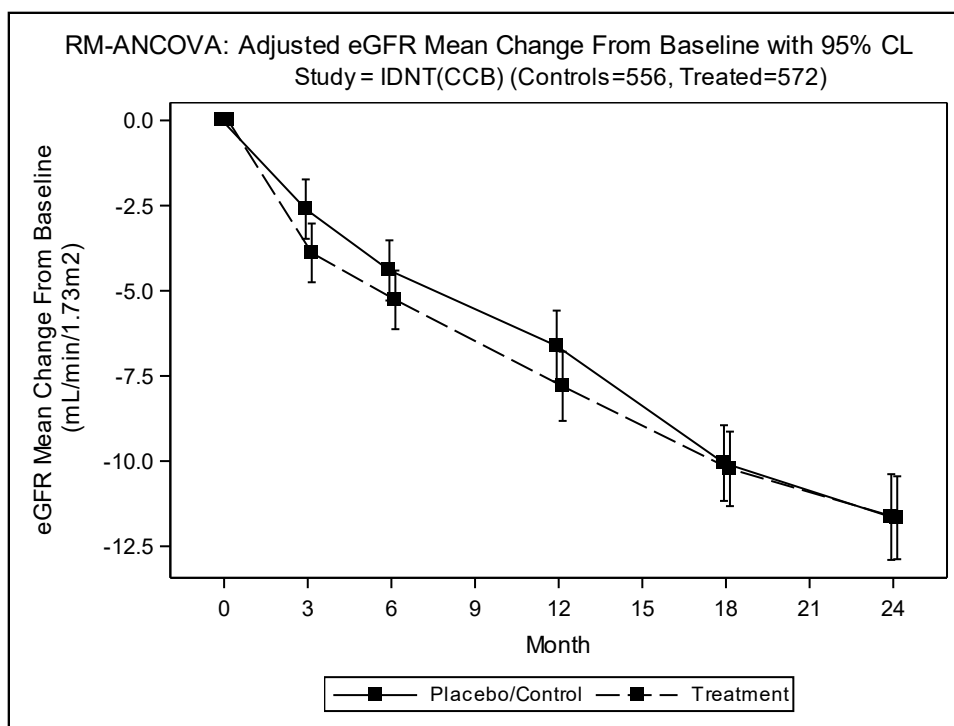


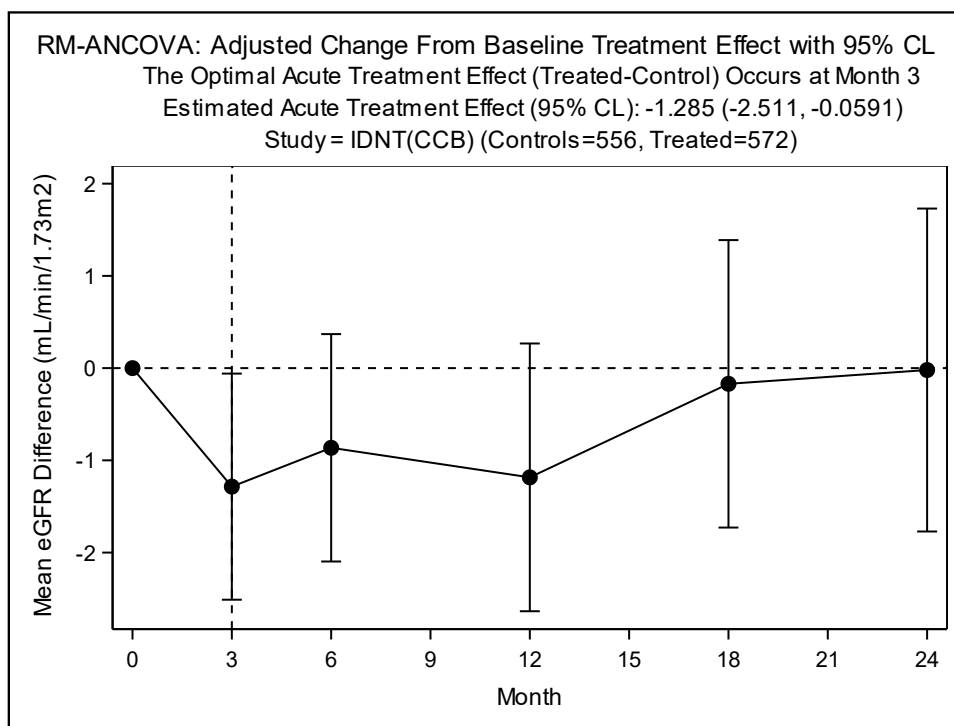
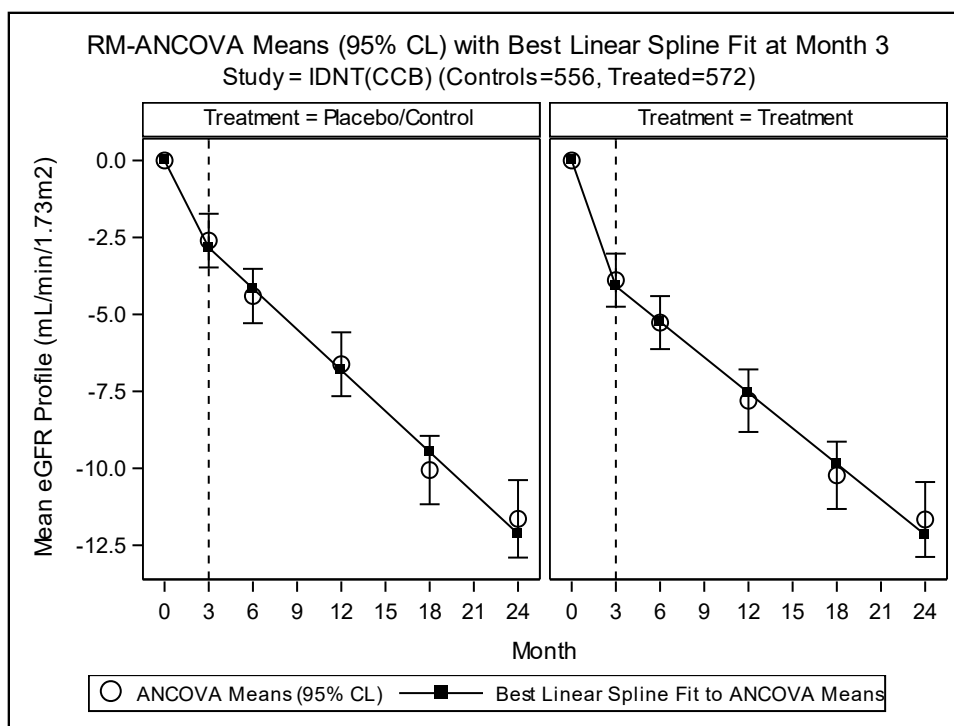




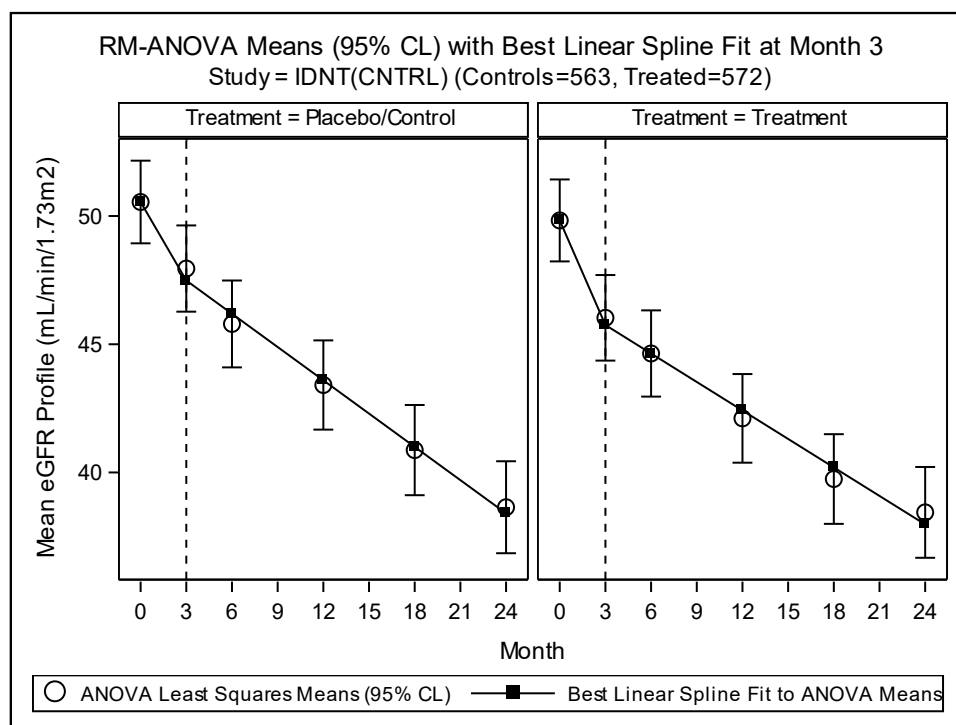
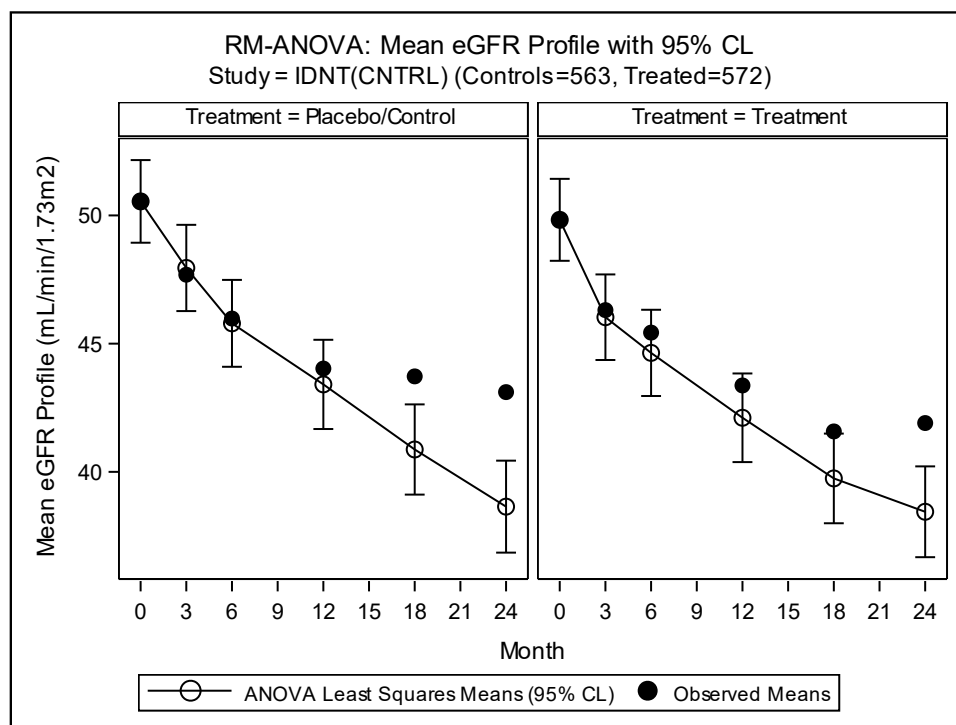
Appendix 3a: IDNT(CCB) results

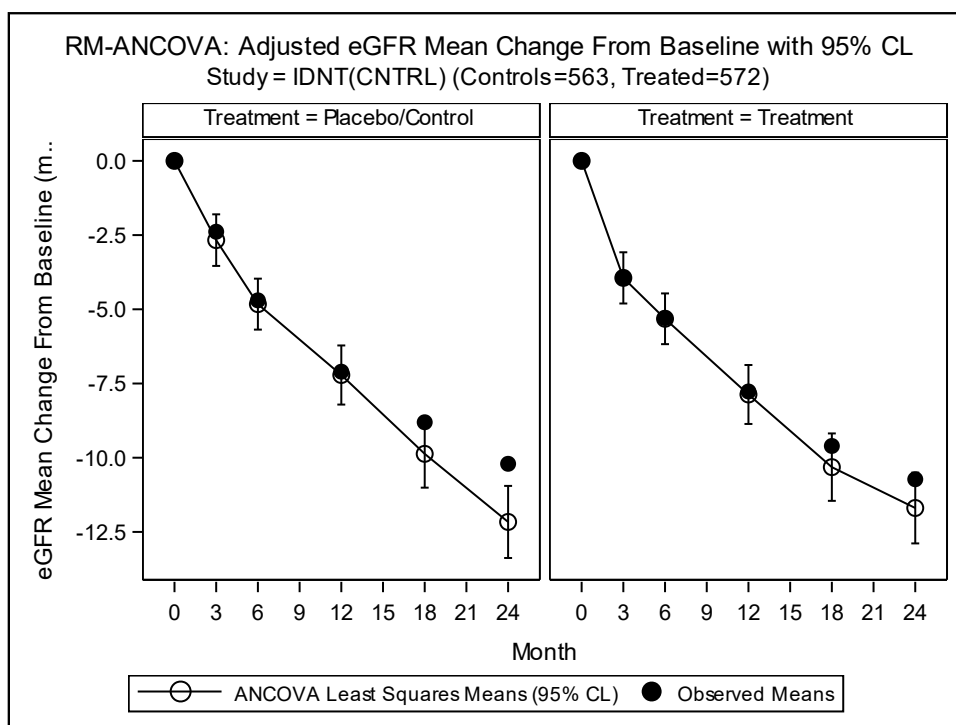
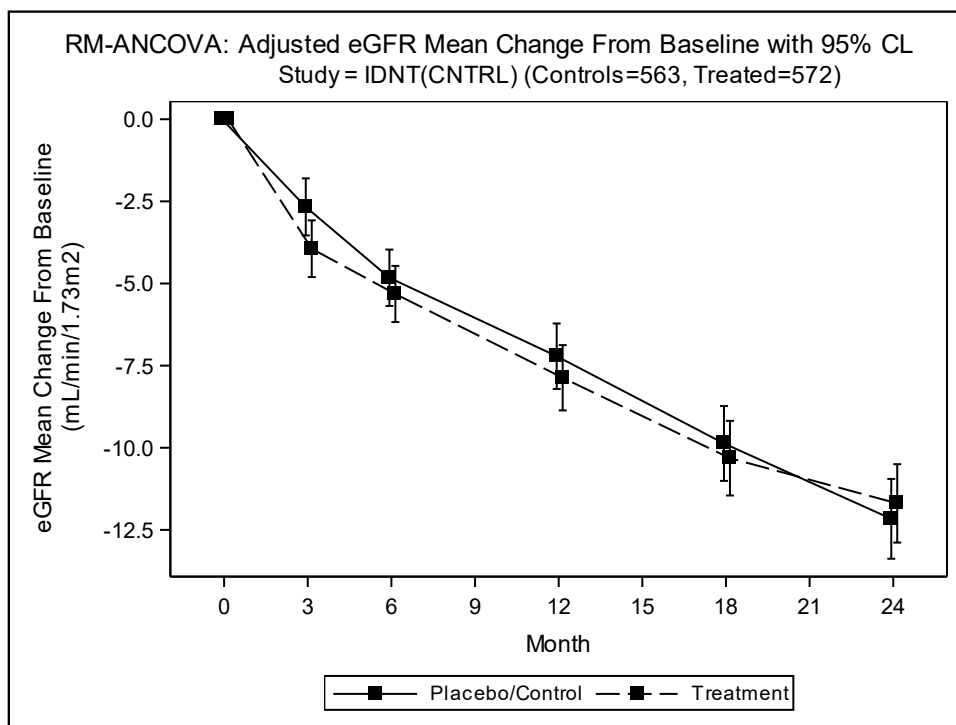


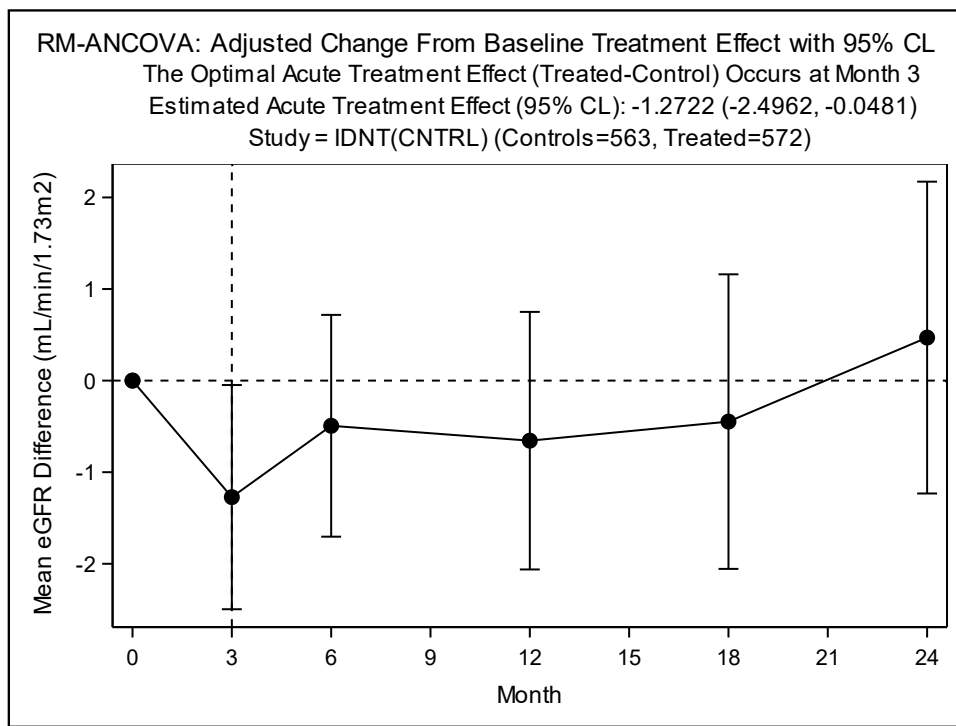
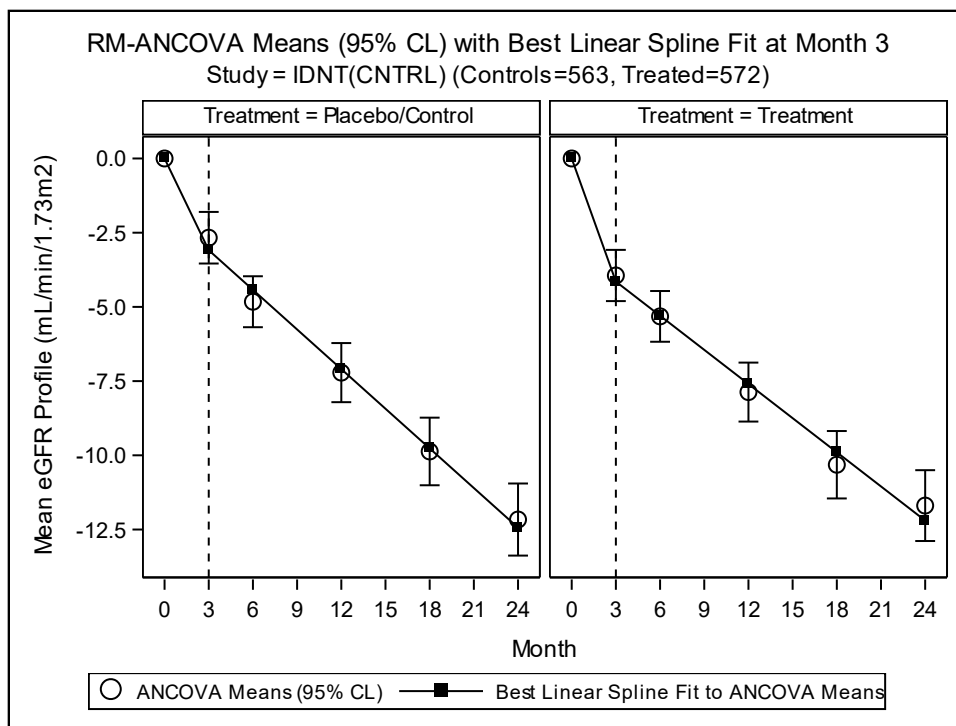




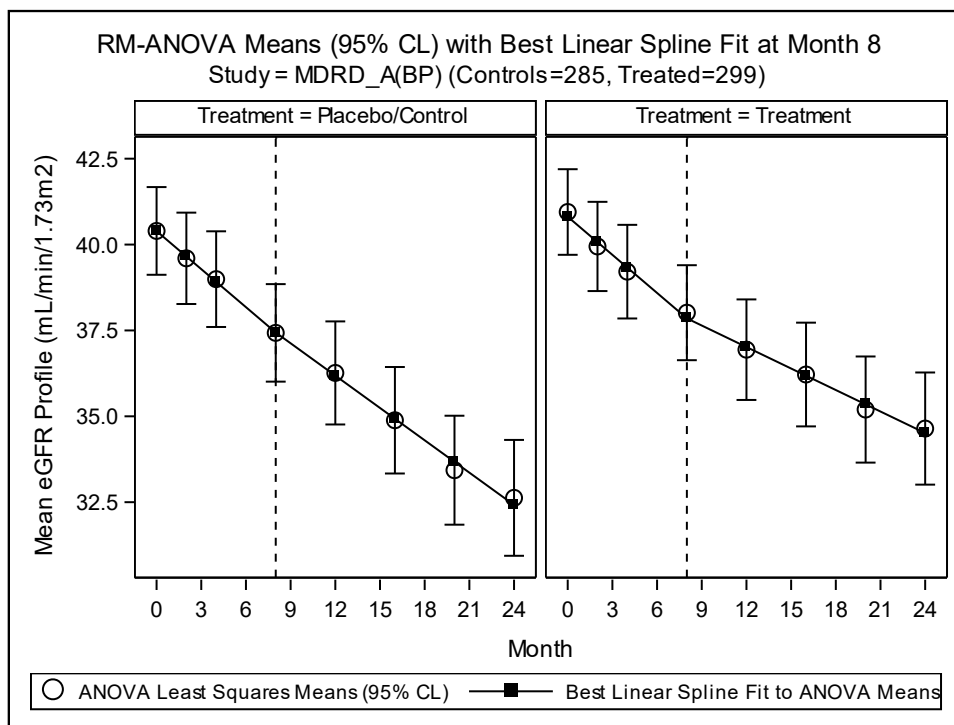
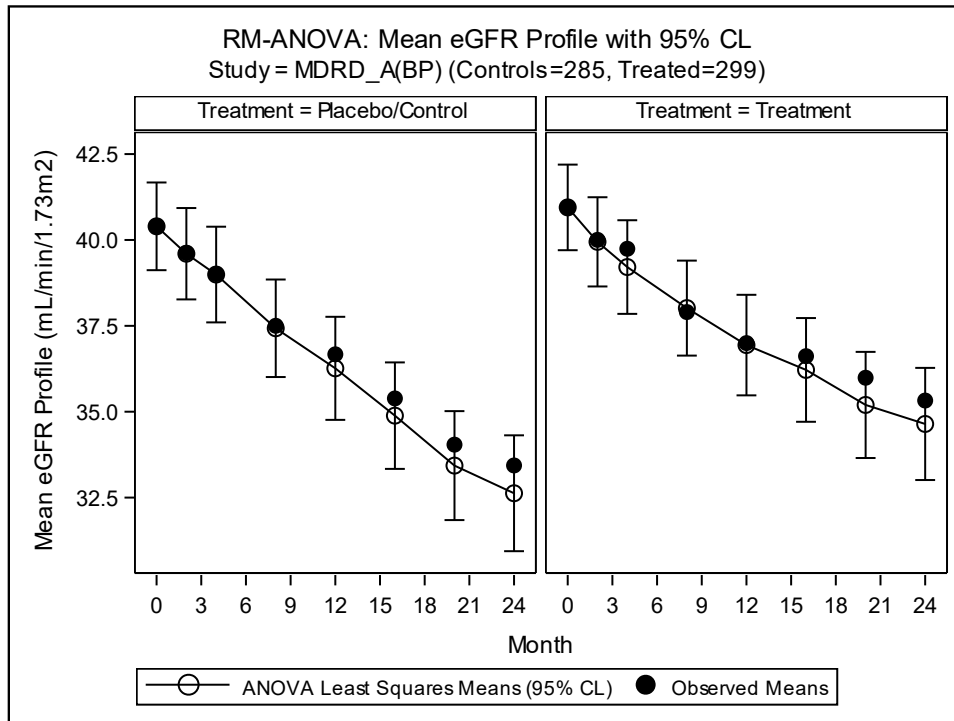
Appendix 4a: IDNT(CNTRL) results

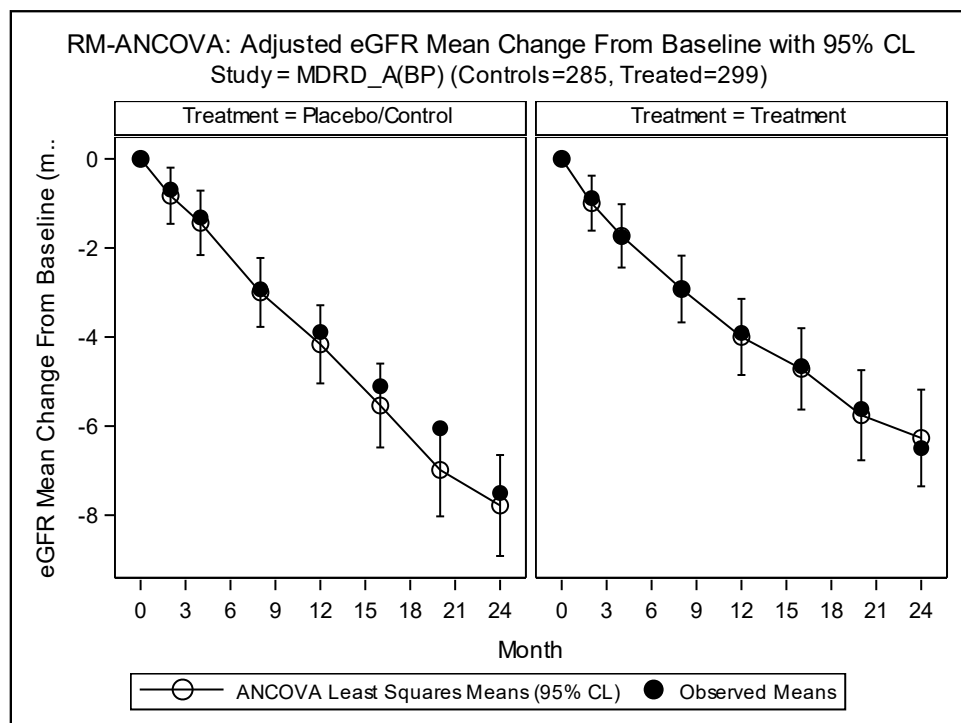
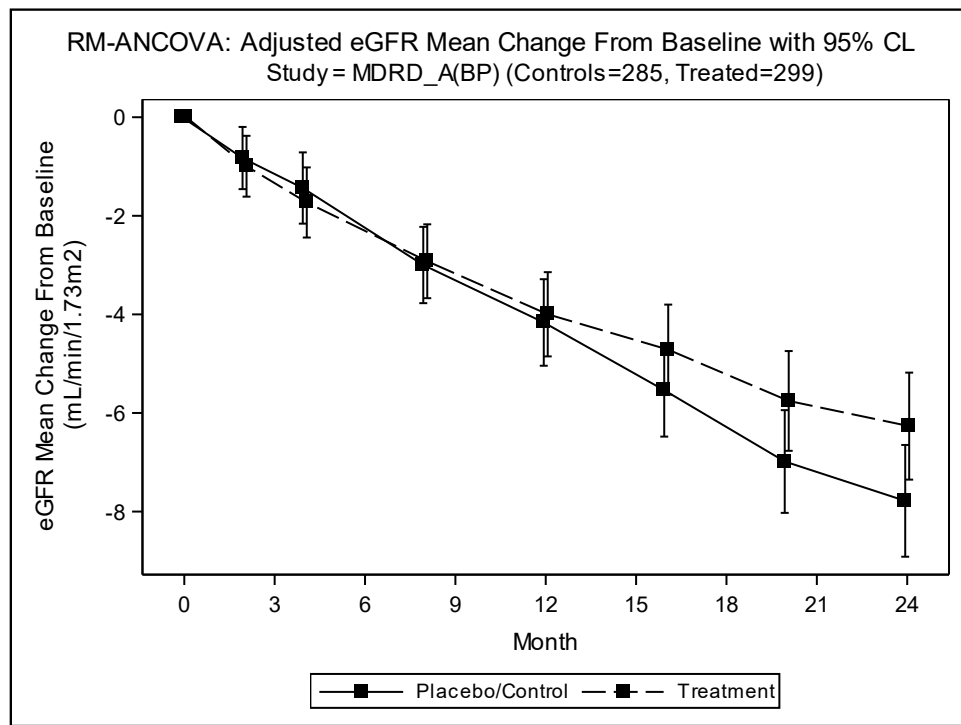


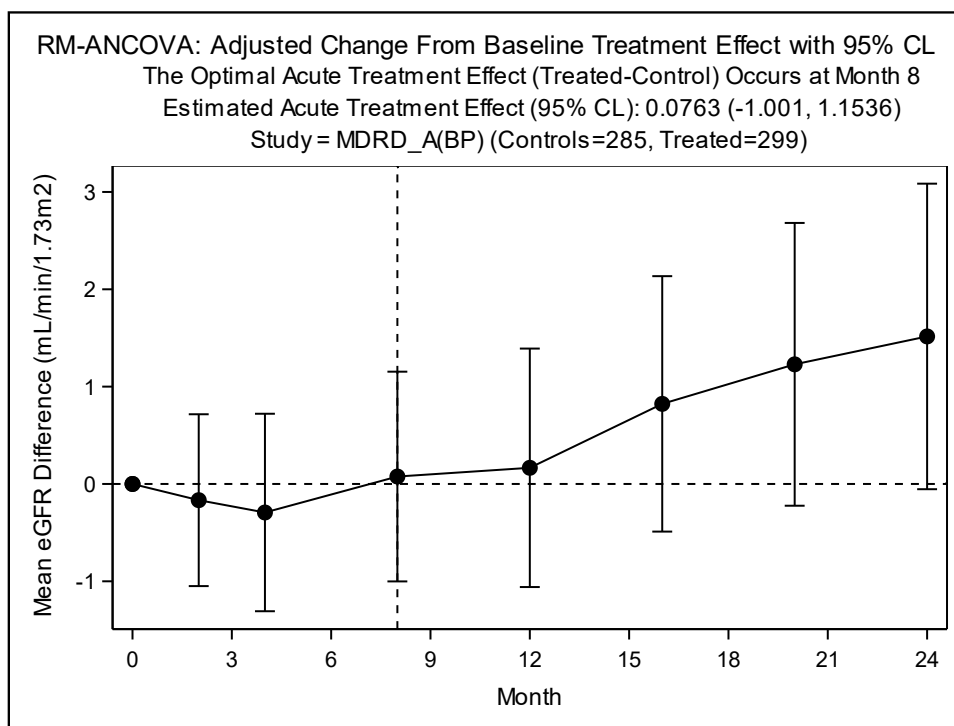
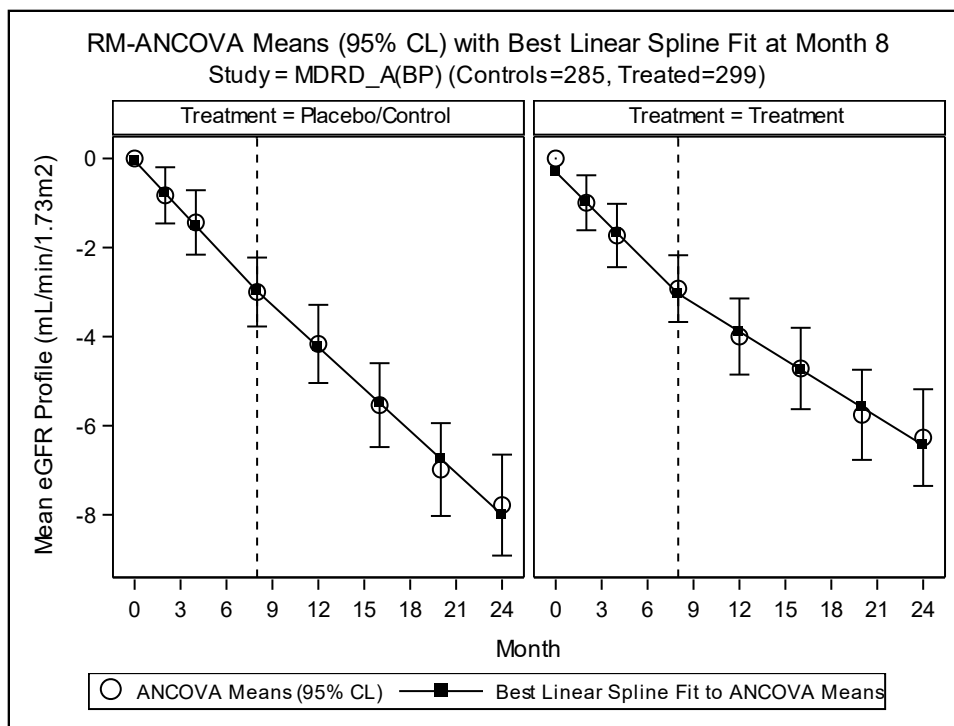




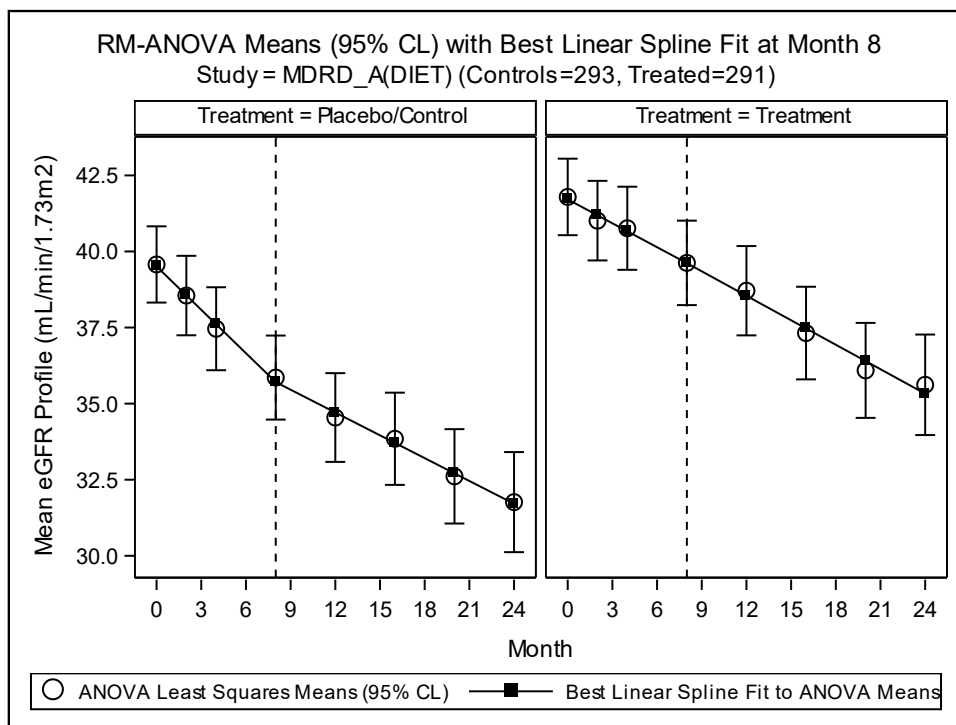
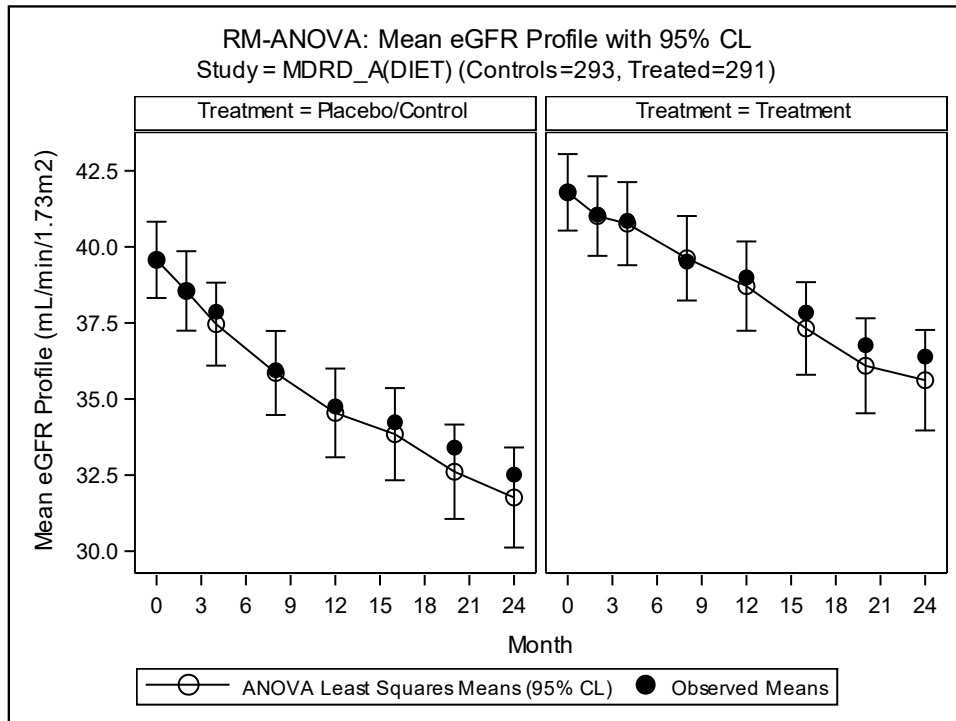
Appendix 5a: MDRD-A(BP) results

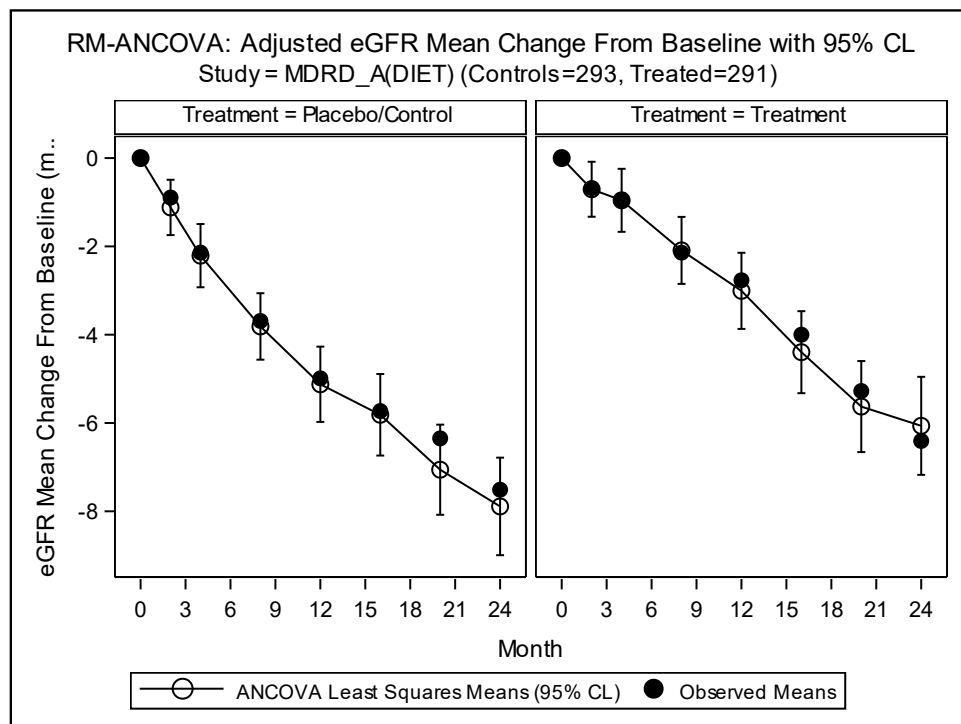
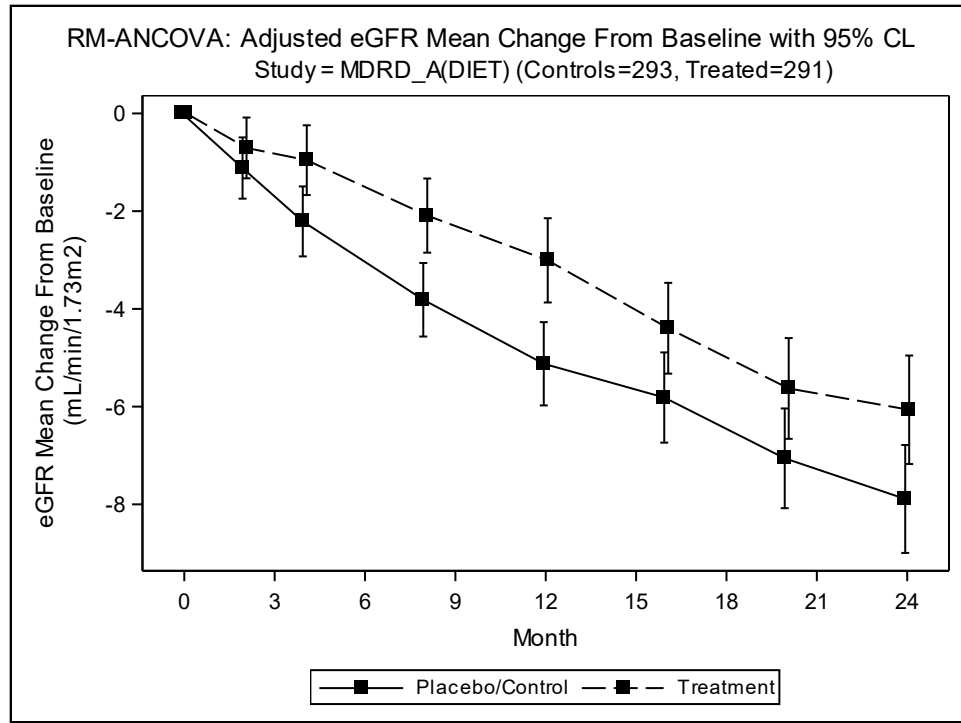


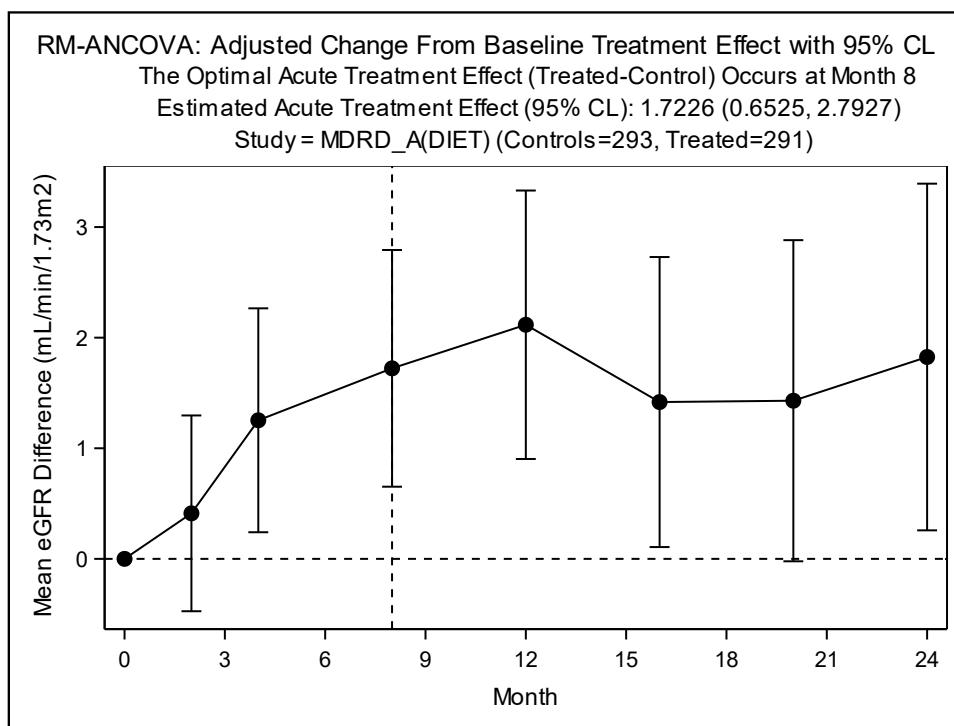
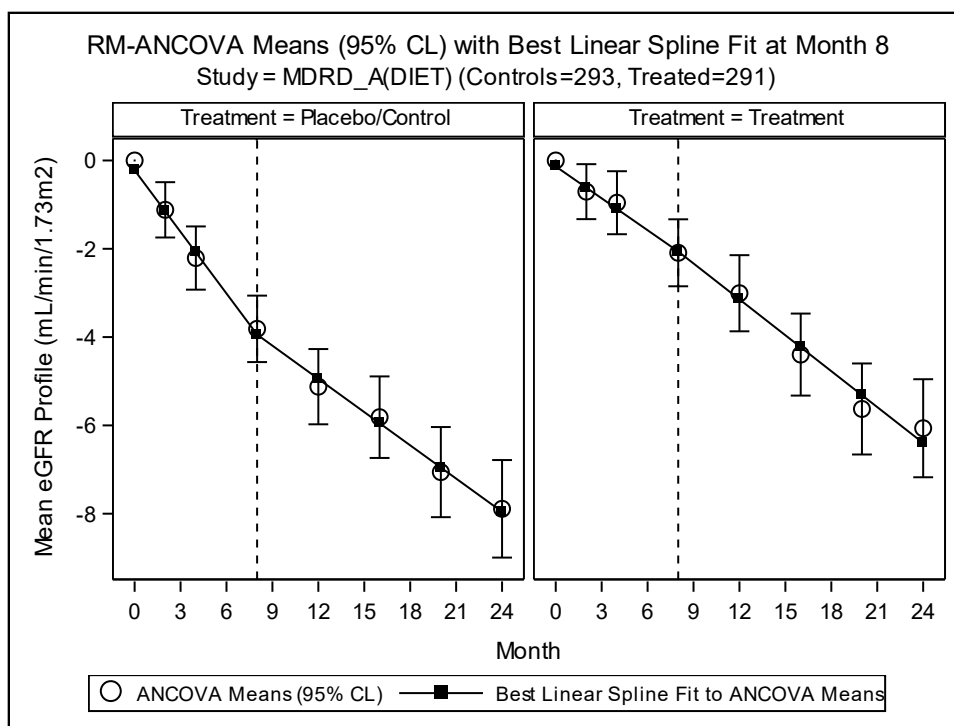




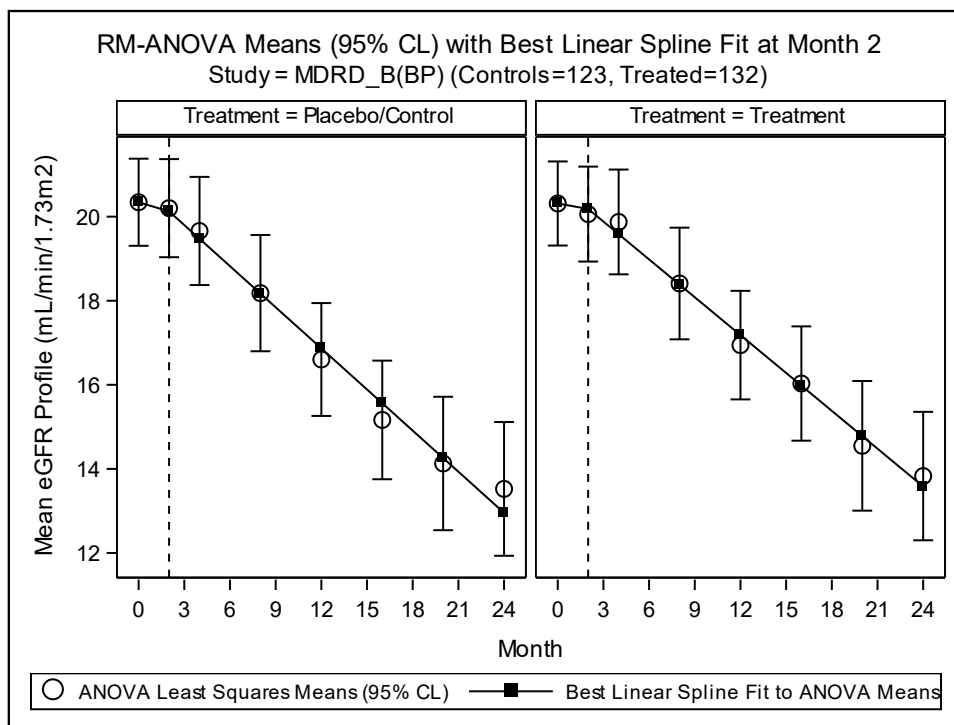
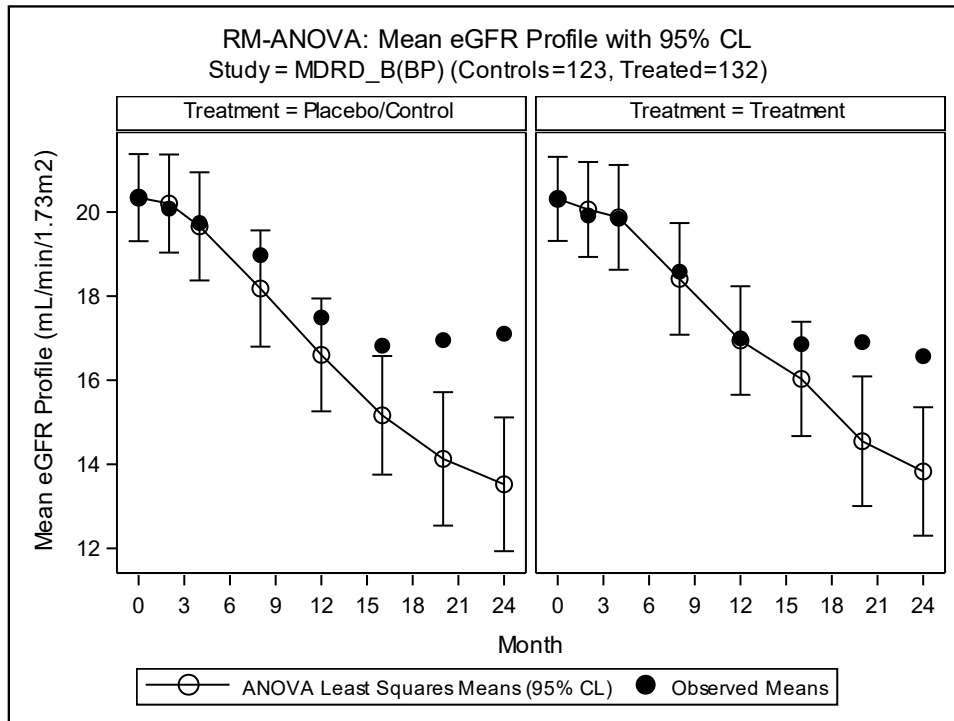
Appendix 6a: MDRD-A(DIET) results

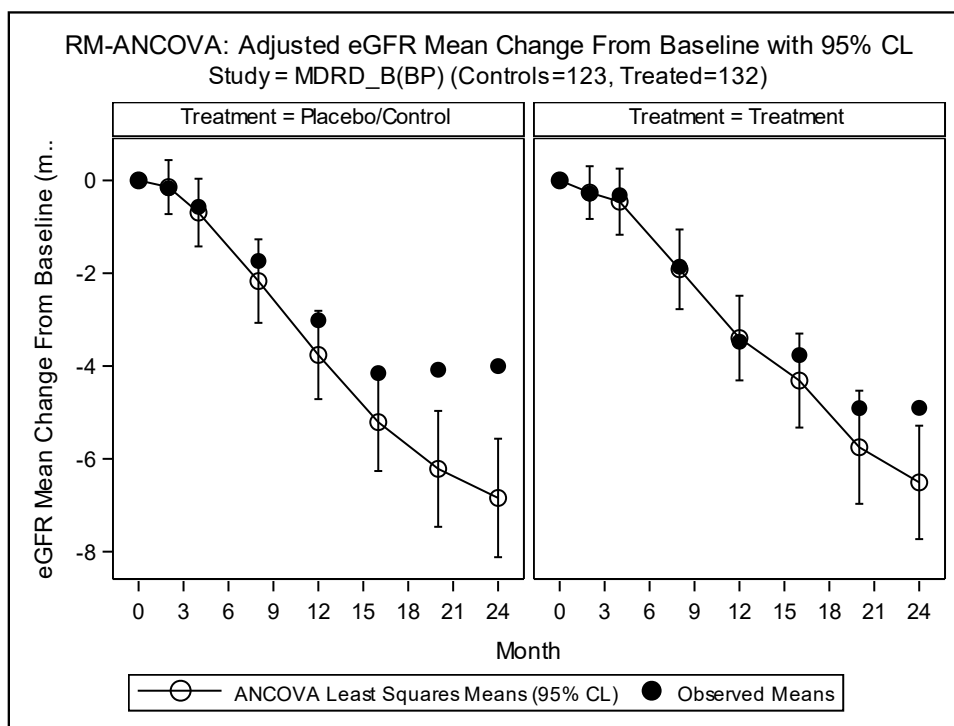
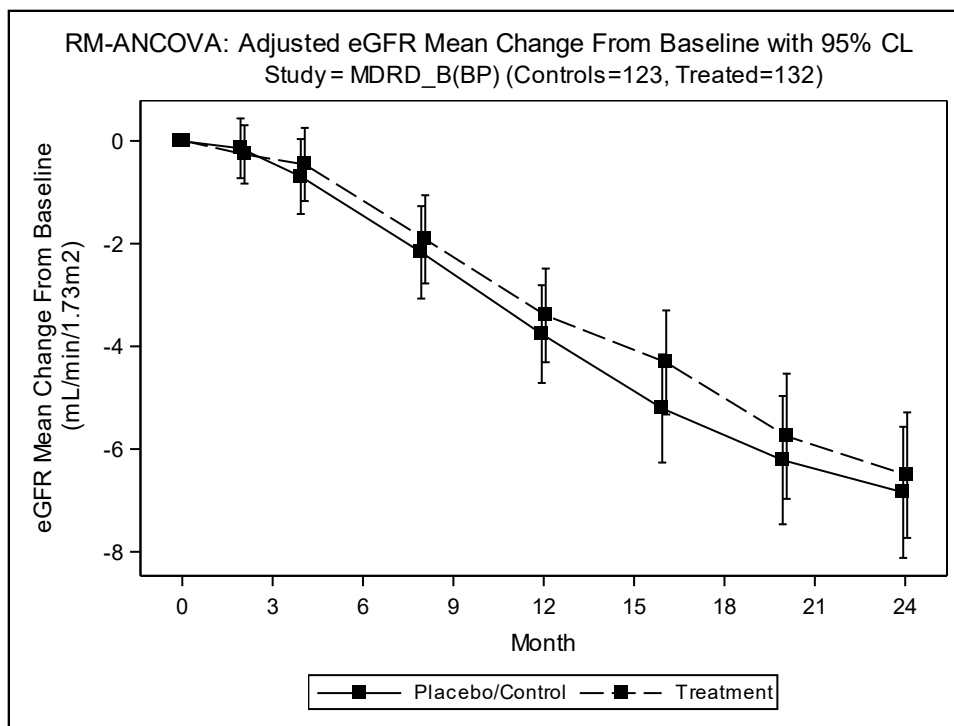


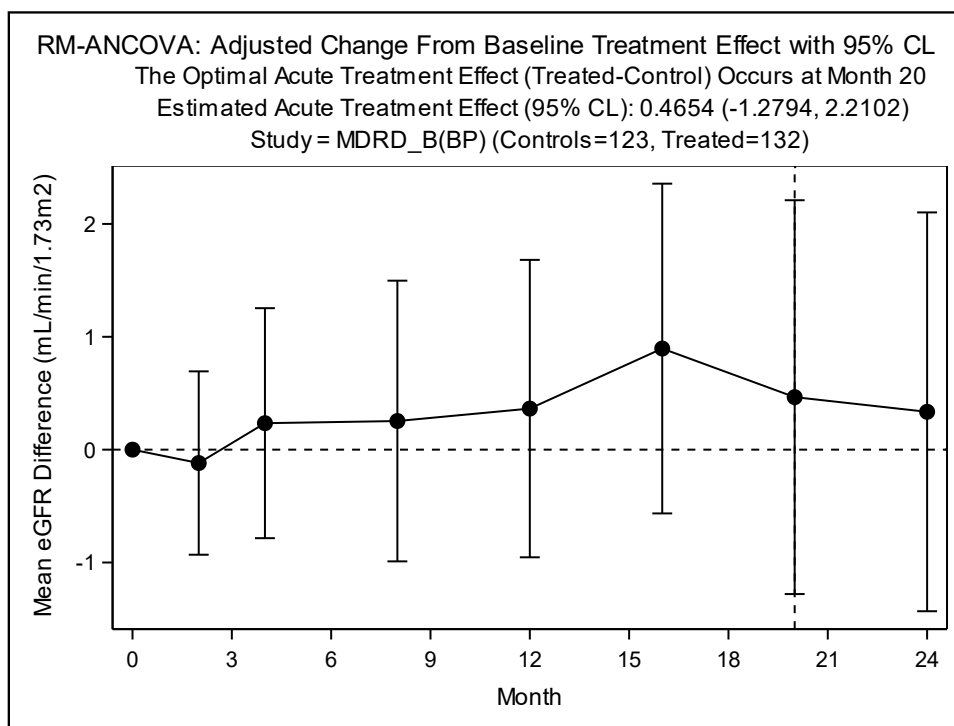
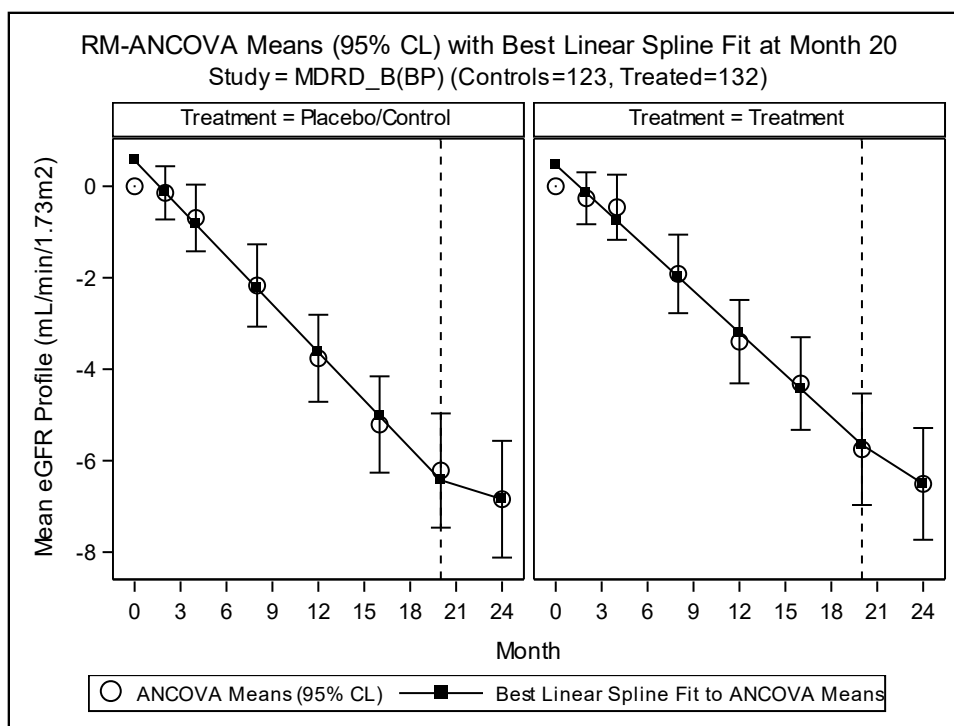




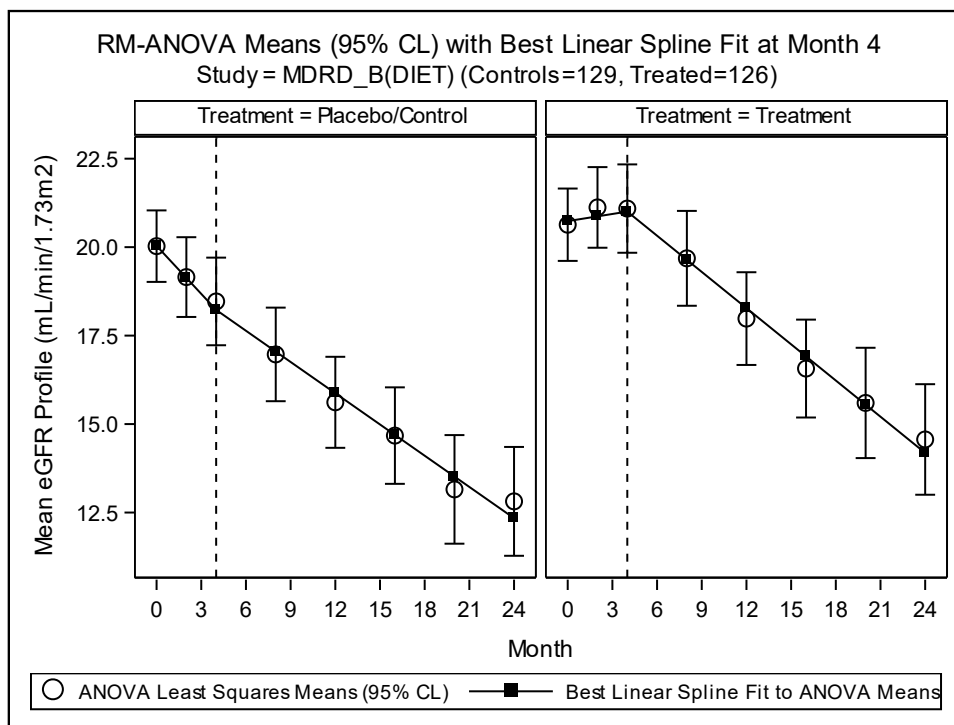
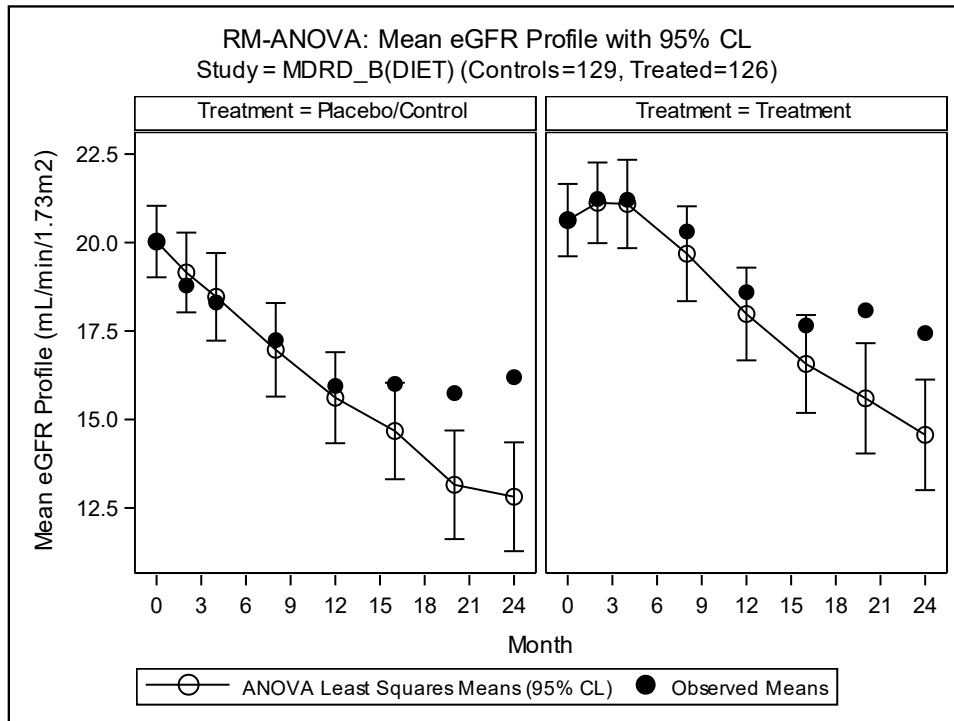
Appendix 7a: MDRD-B(BP) results

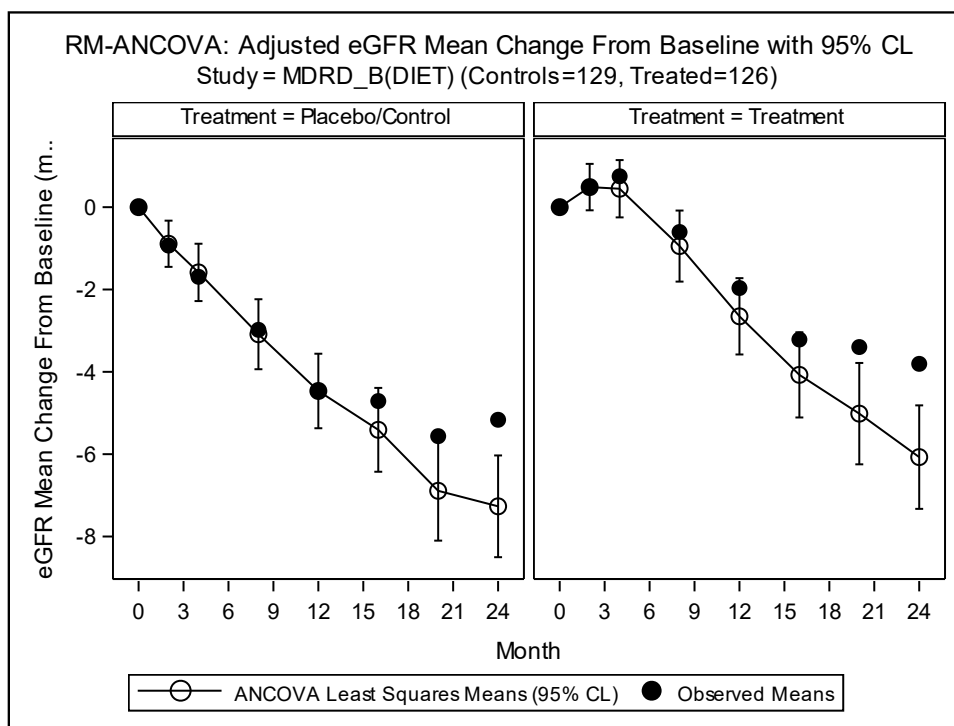
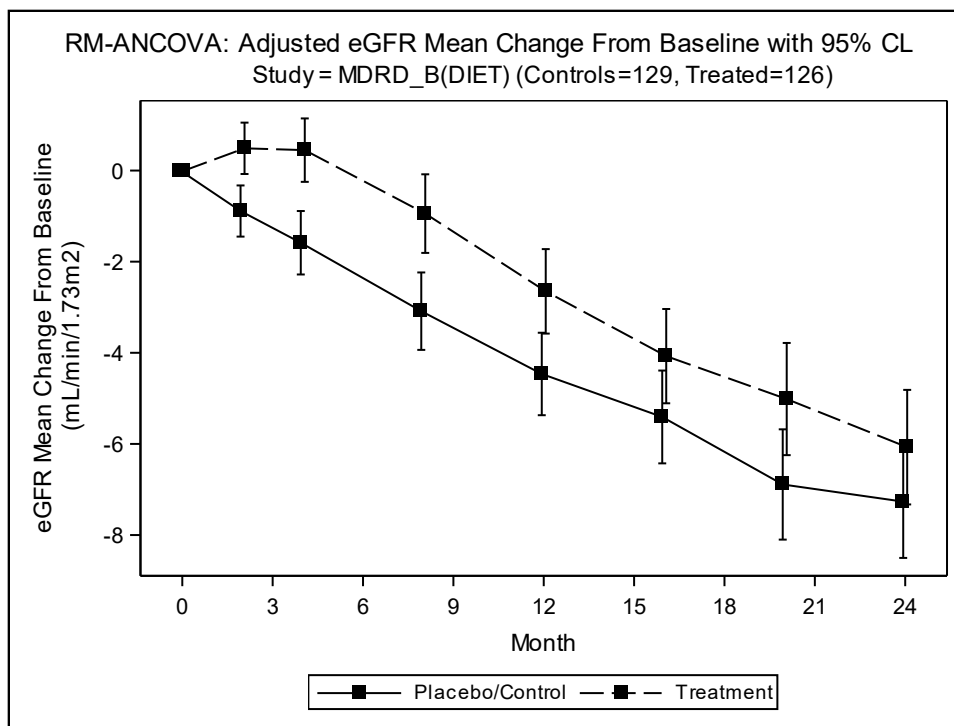


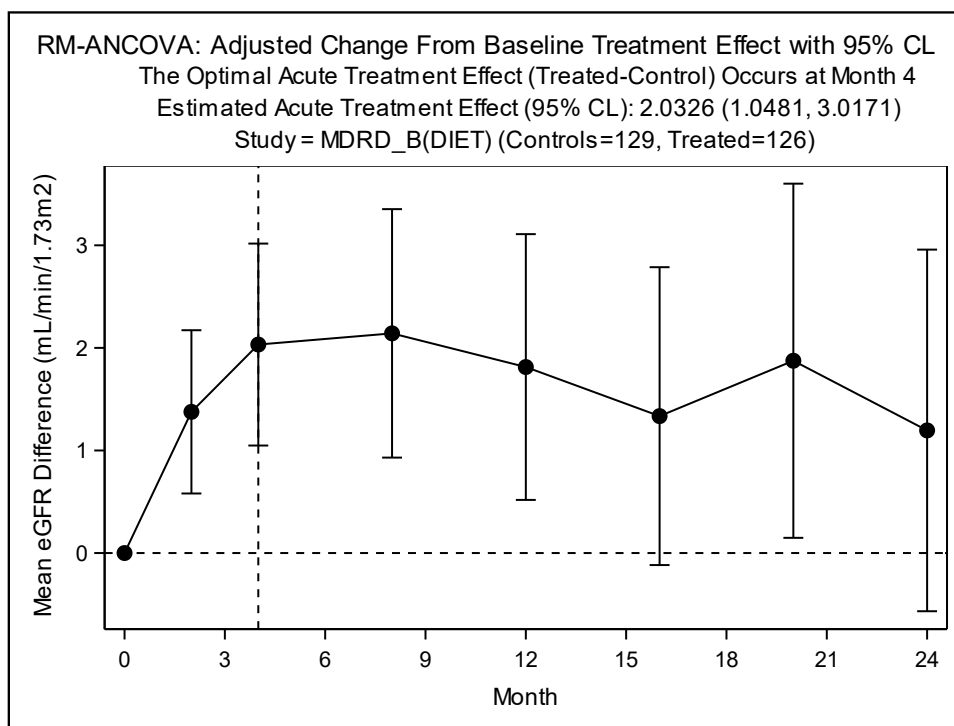
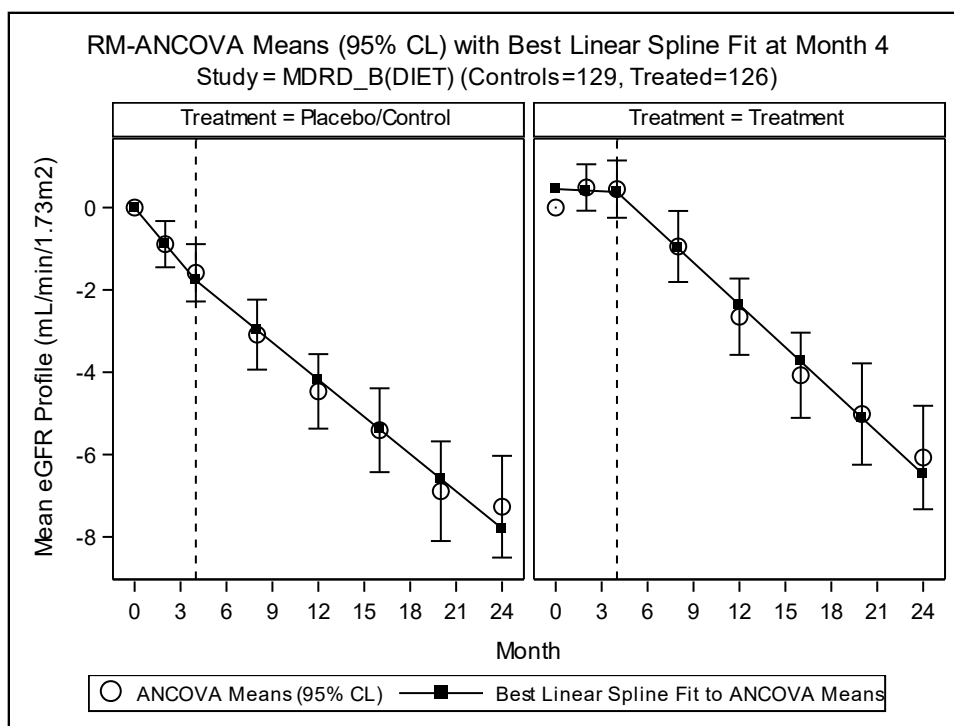




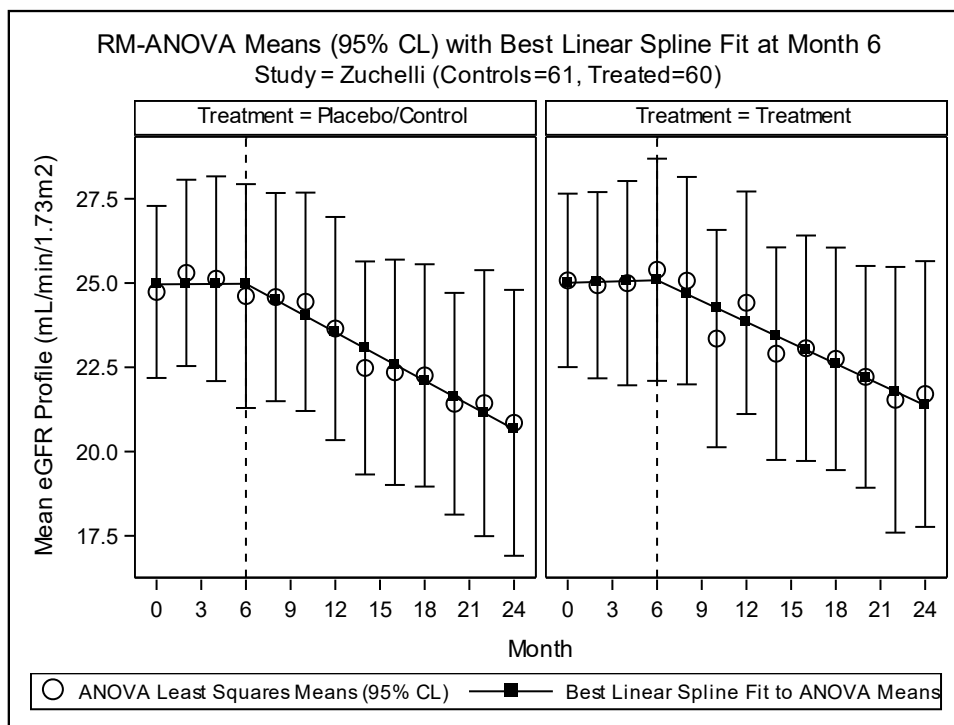
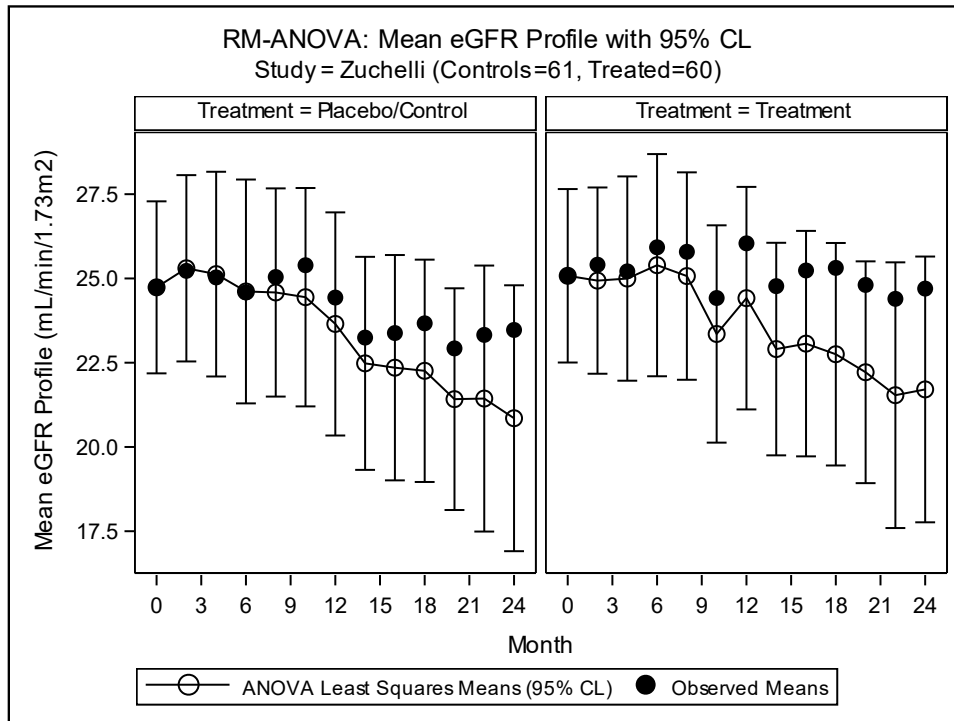
Appendix 8a: MDRD-B(DIET) results

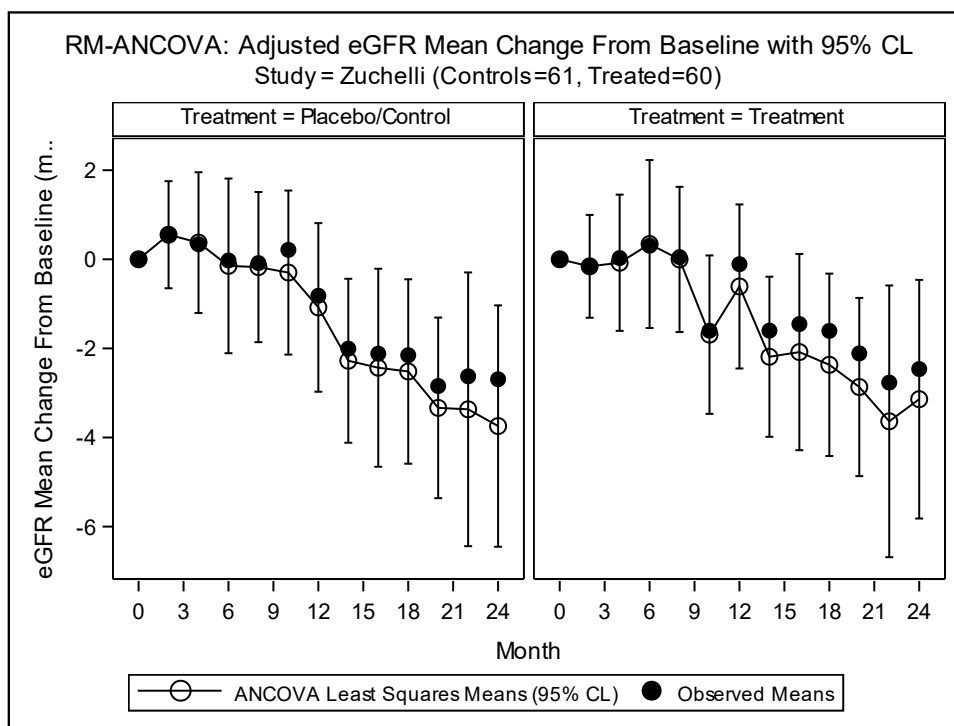
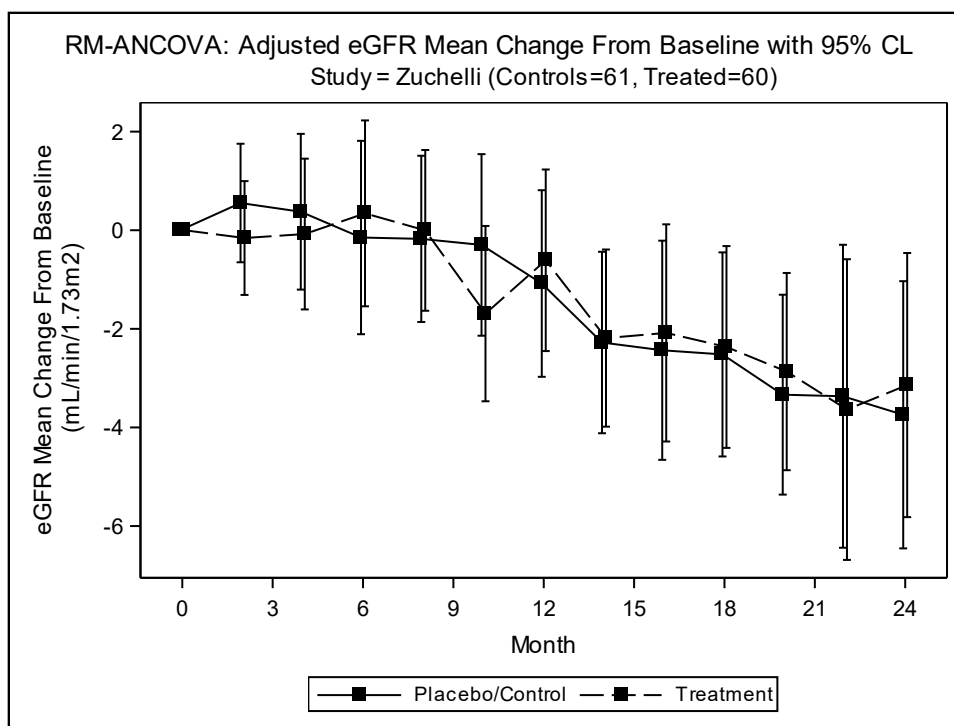


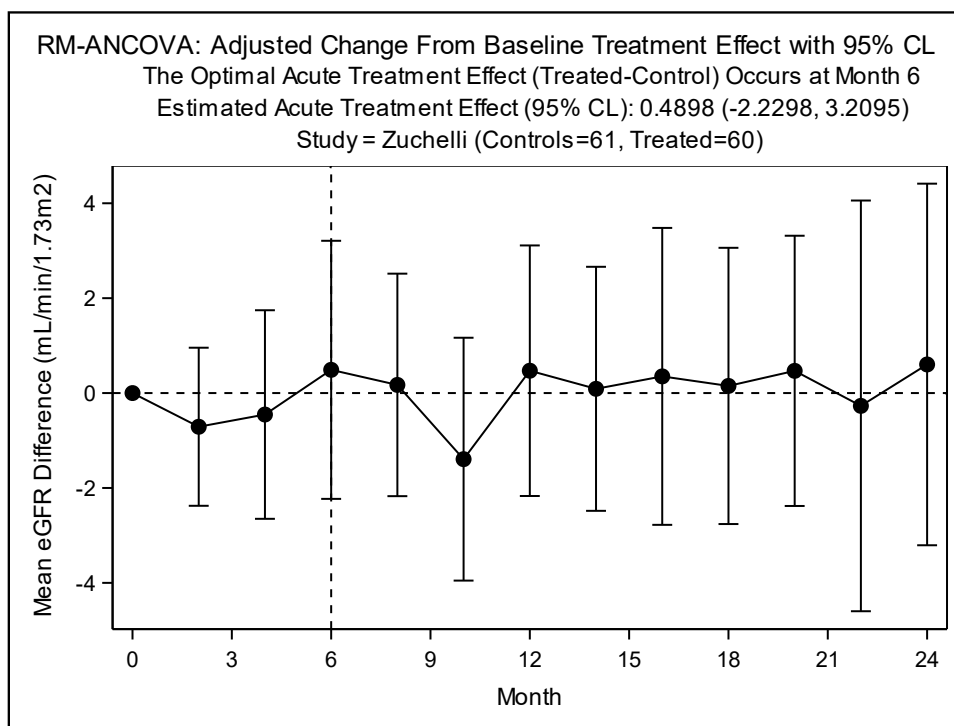
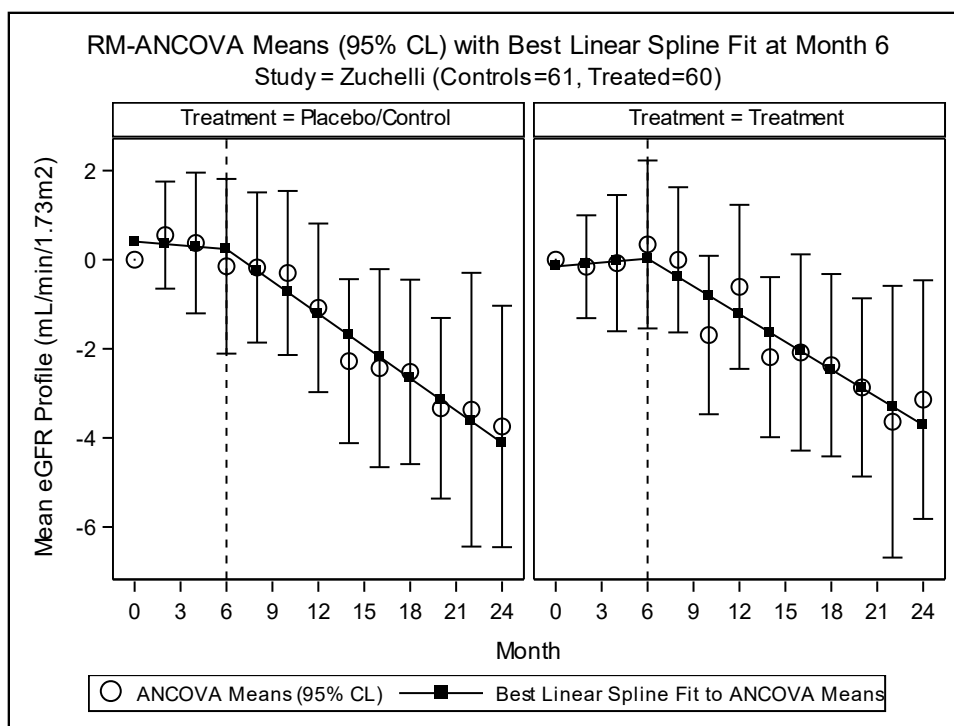




Appendix 9a: Zuchelli results







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

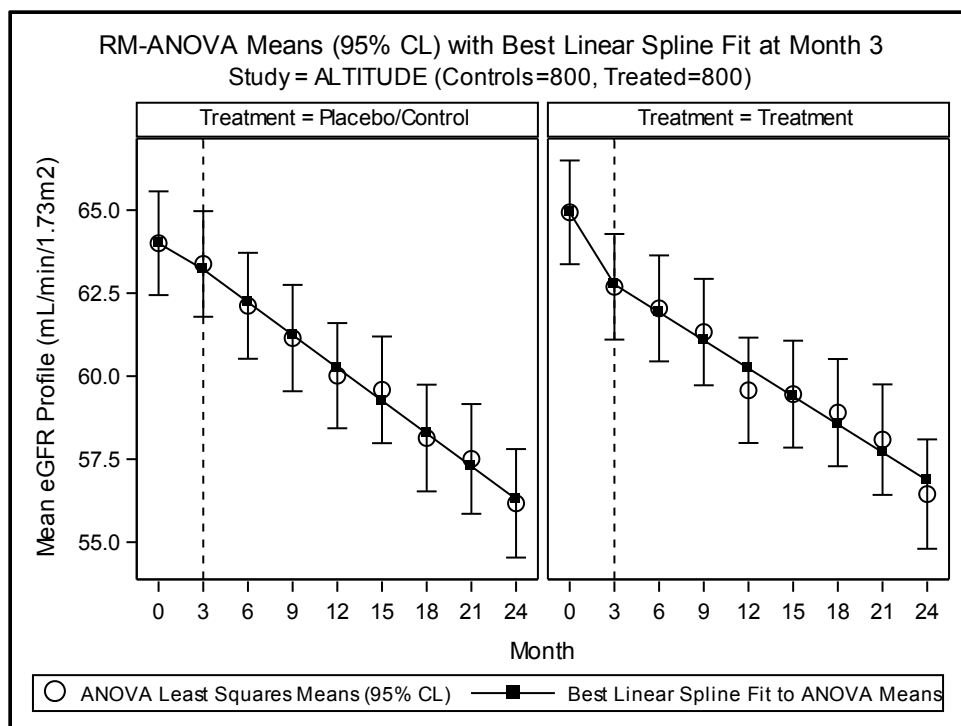
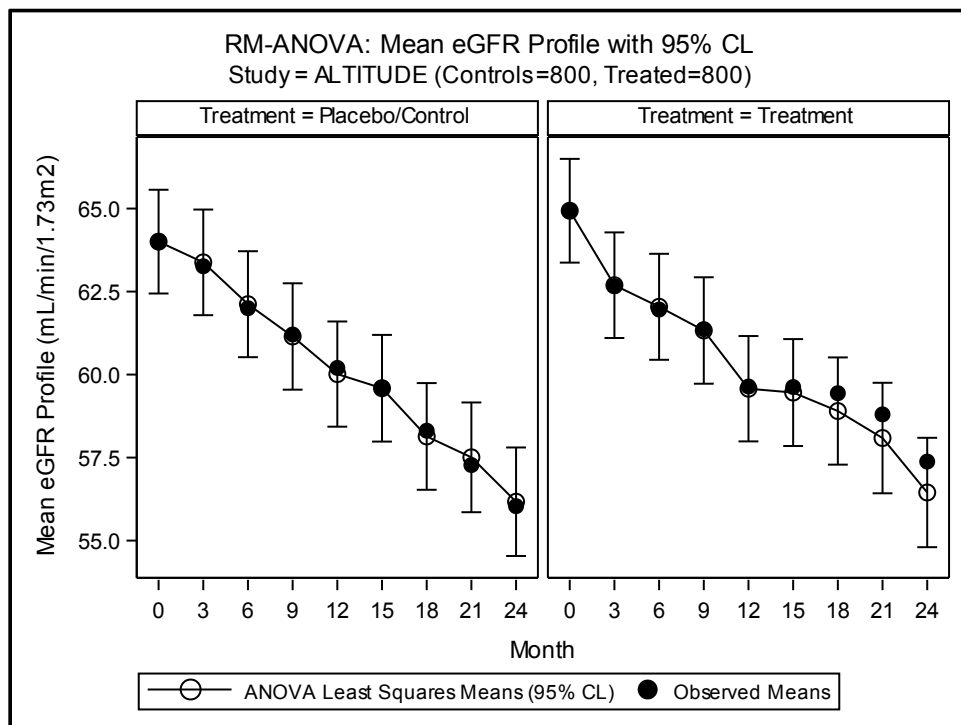
The analyses presented here illustrate the importance of using a purely random sample of subjects when estimating the timing and magnitude of acute treatment effects based on sample sizes that may be more realistic for phase 2 trials. The two large studies and one small study selected here will be used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

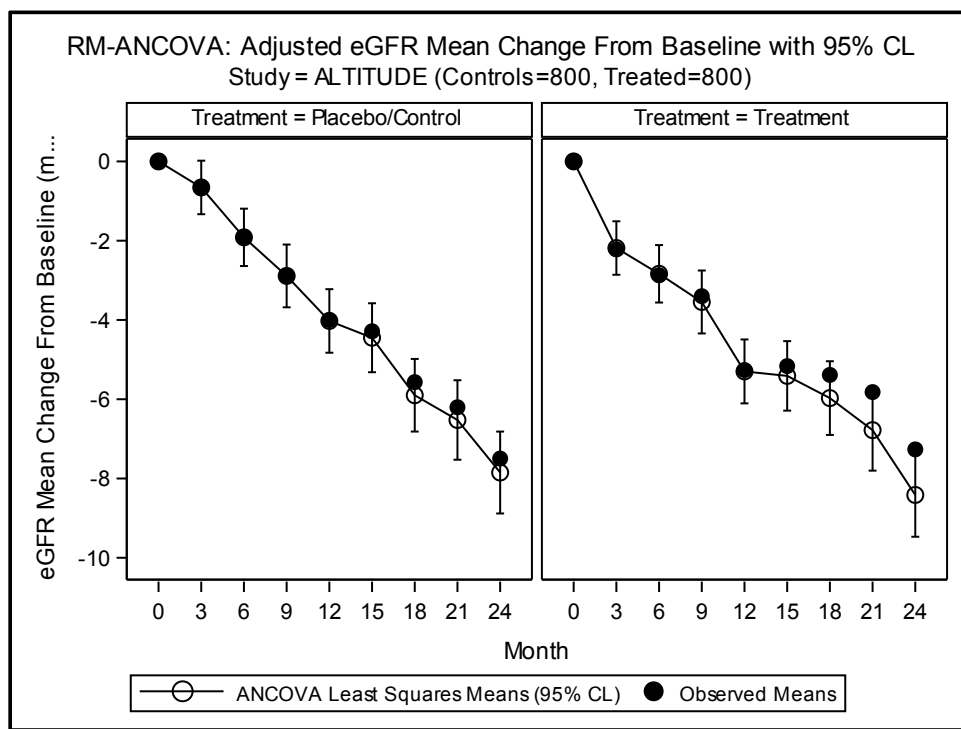
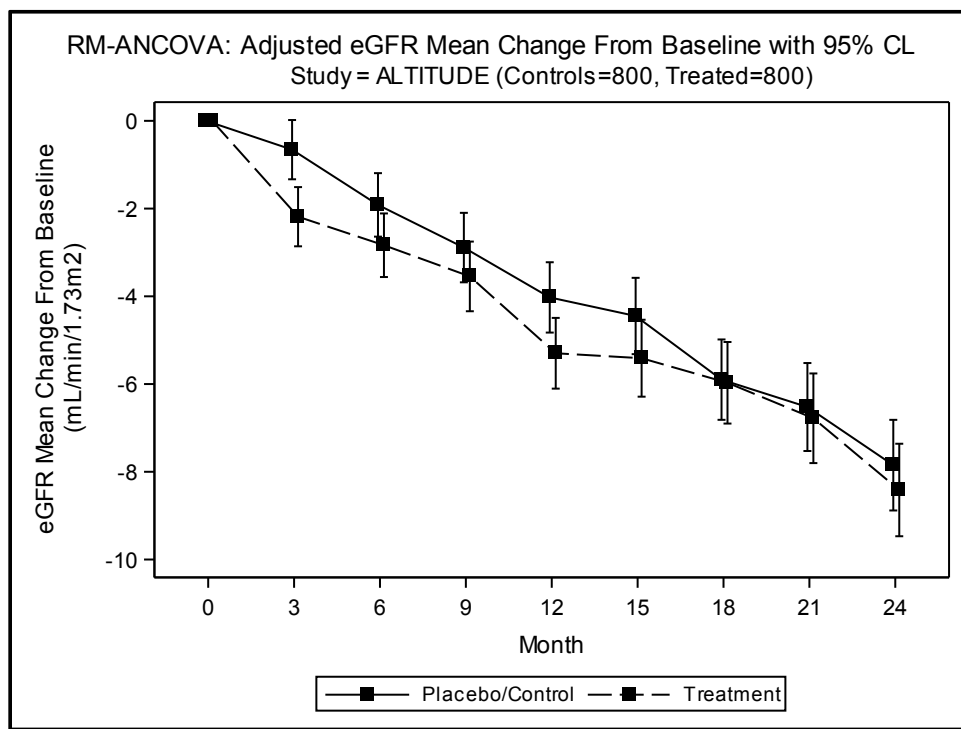
CKD trials – both large and small. There are advantages and disadvantages to using either the RM-ANCOVA modeling approach or the linear spline-mixed-effects modeling approach. By comparing how concordant or discordant results from both methods are and how well those results agree with clinical experience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

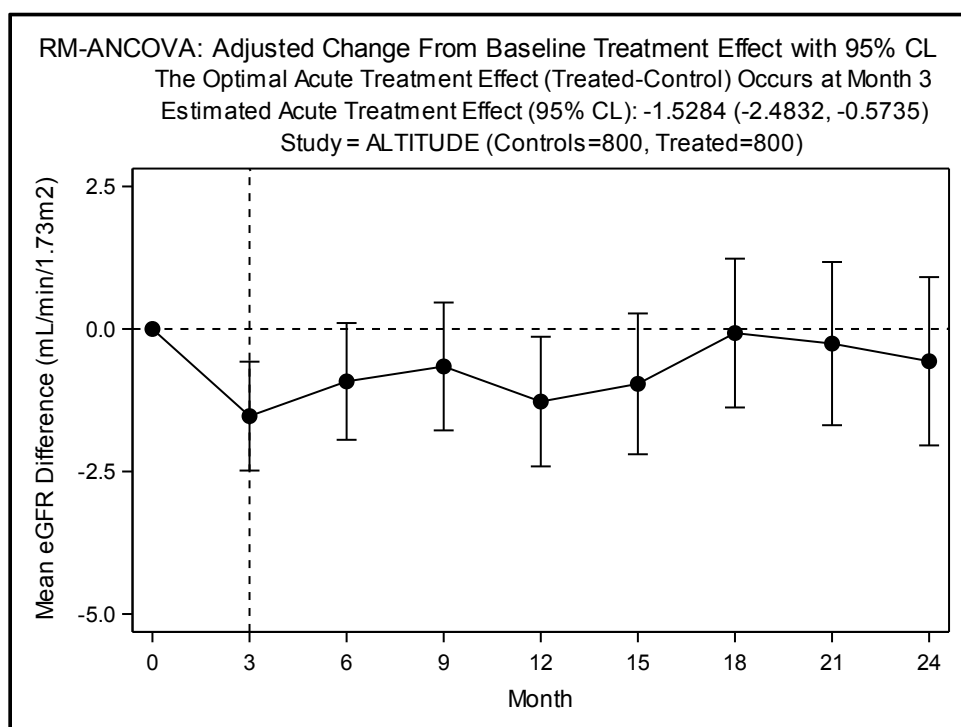
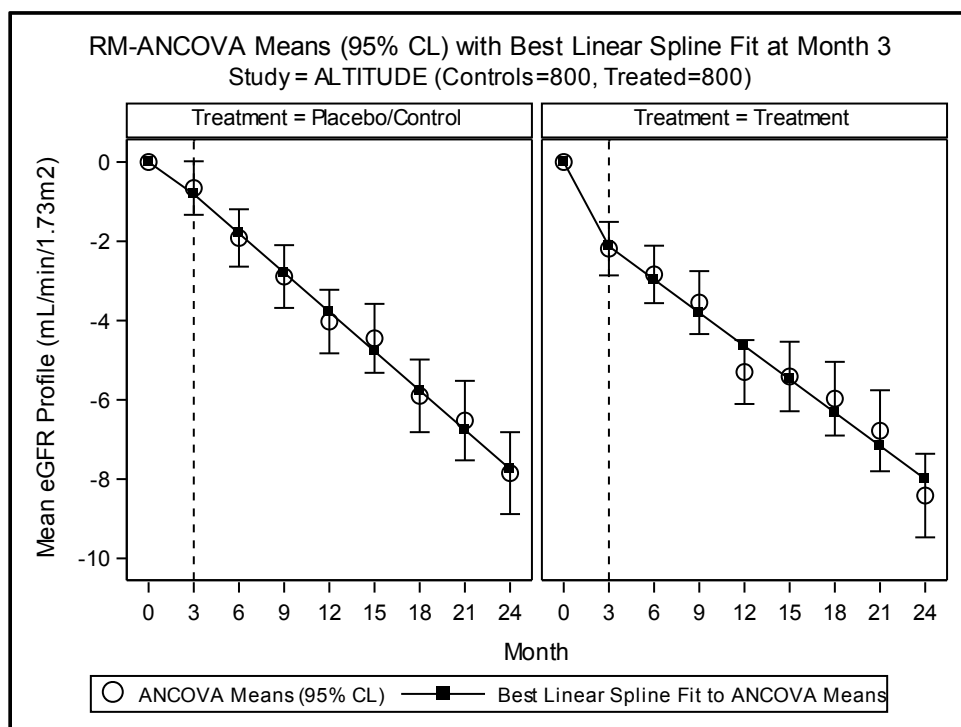
References

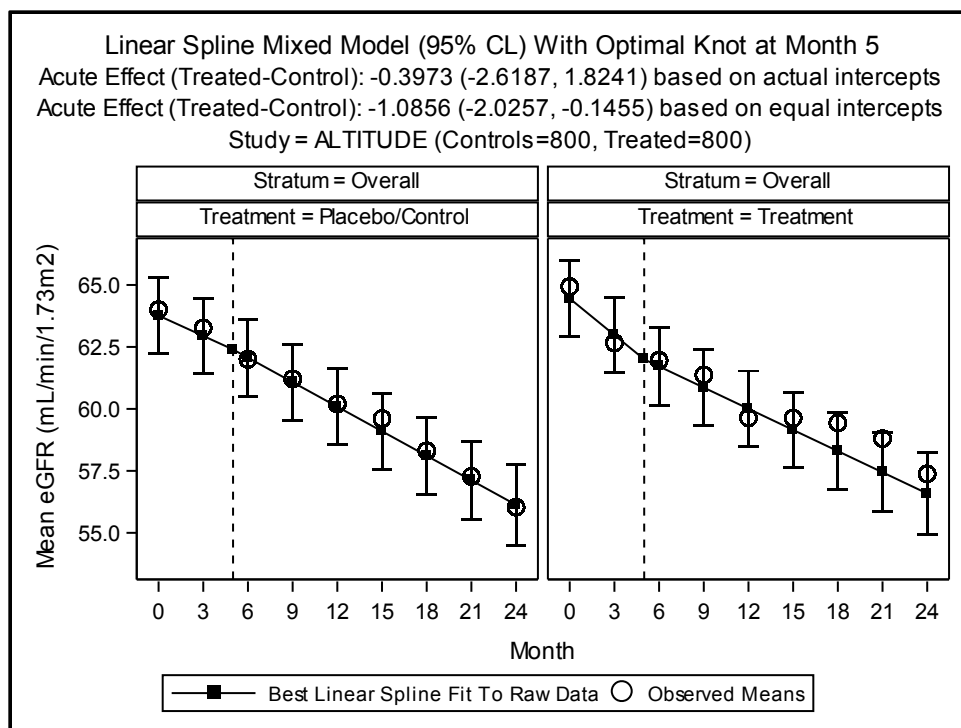
1. Vonesh, E., Tighiouart, H., Ying, J., Heerspink, H. L., Lewis, J., Staplin, N., et al. (2019). Mixed-effects models for slope-based endpoints in clinical trials of chronic kidney disease. *Statistics in Medicine* 38, 4218-4239.
2. Vonesh, E. Technical report summarizing different methods for determining 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

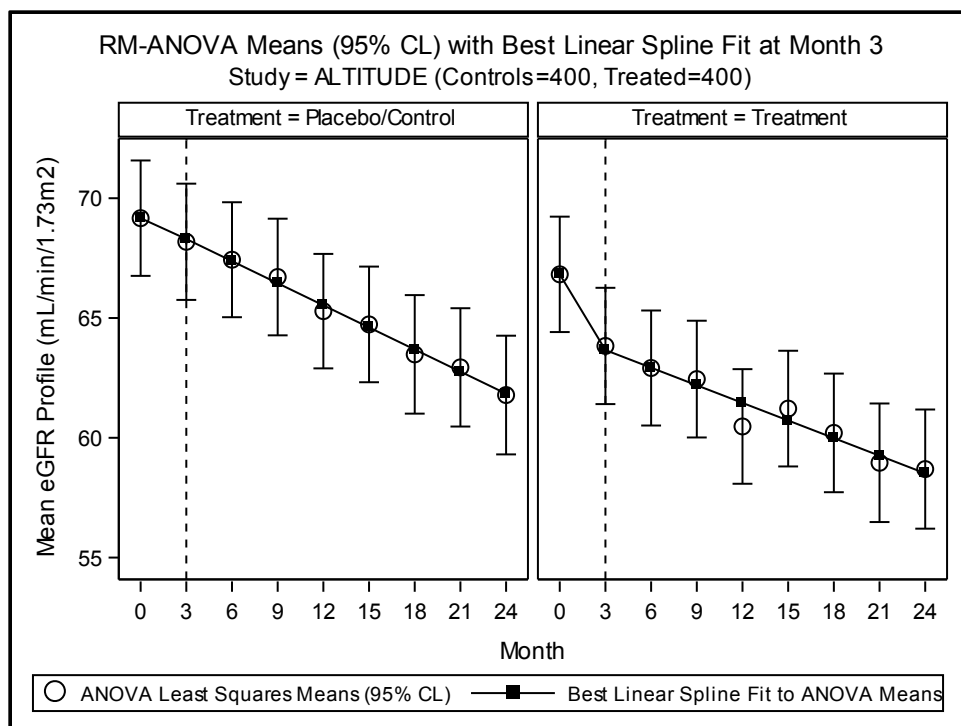
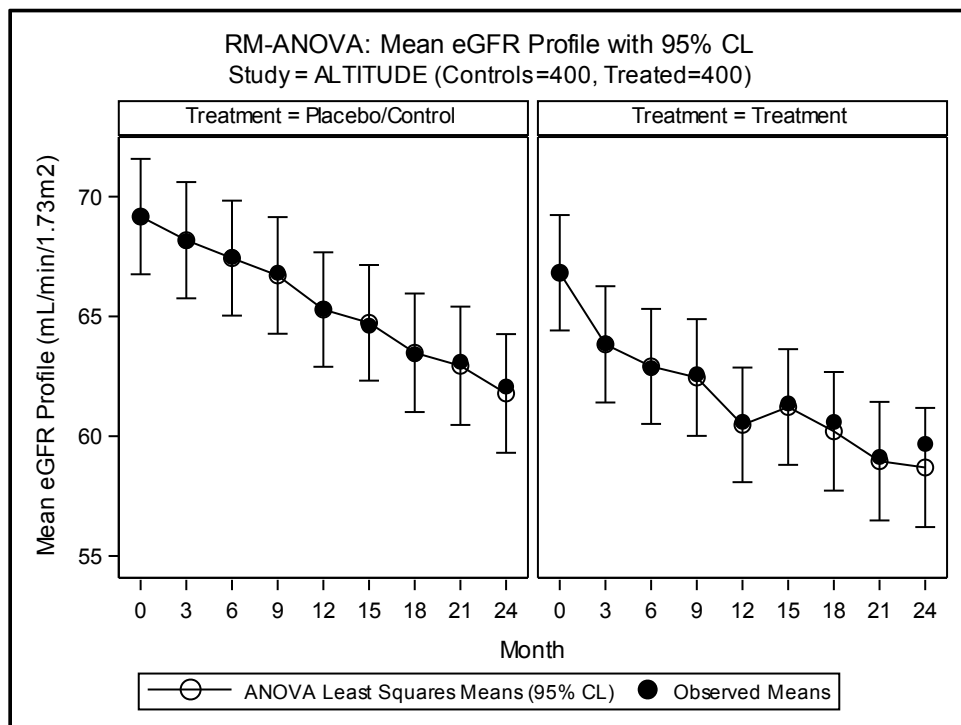


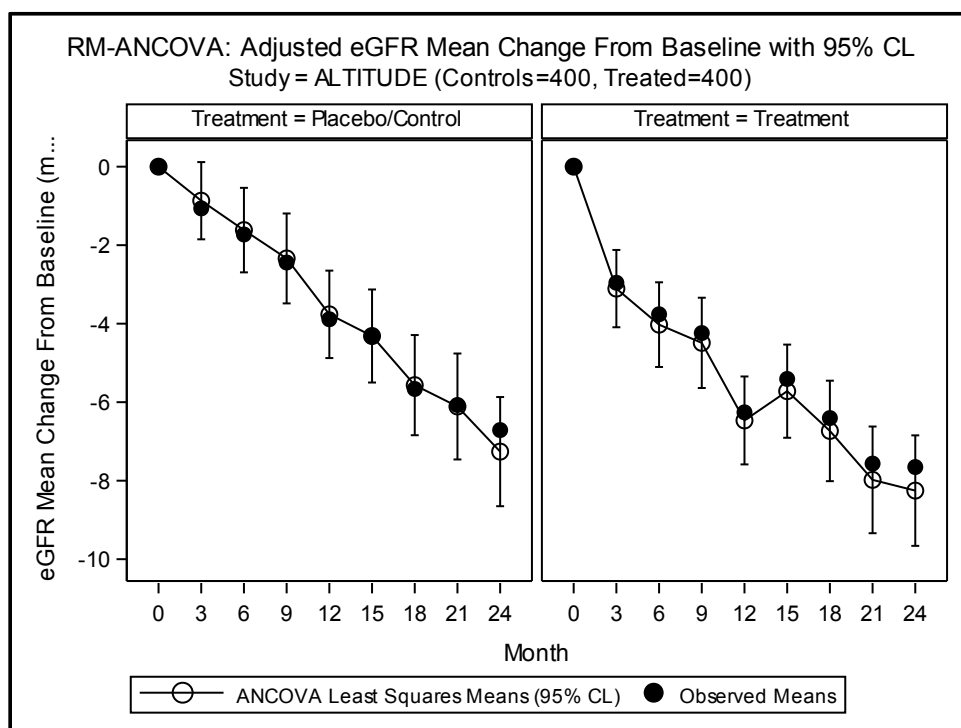
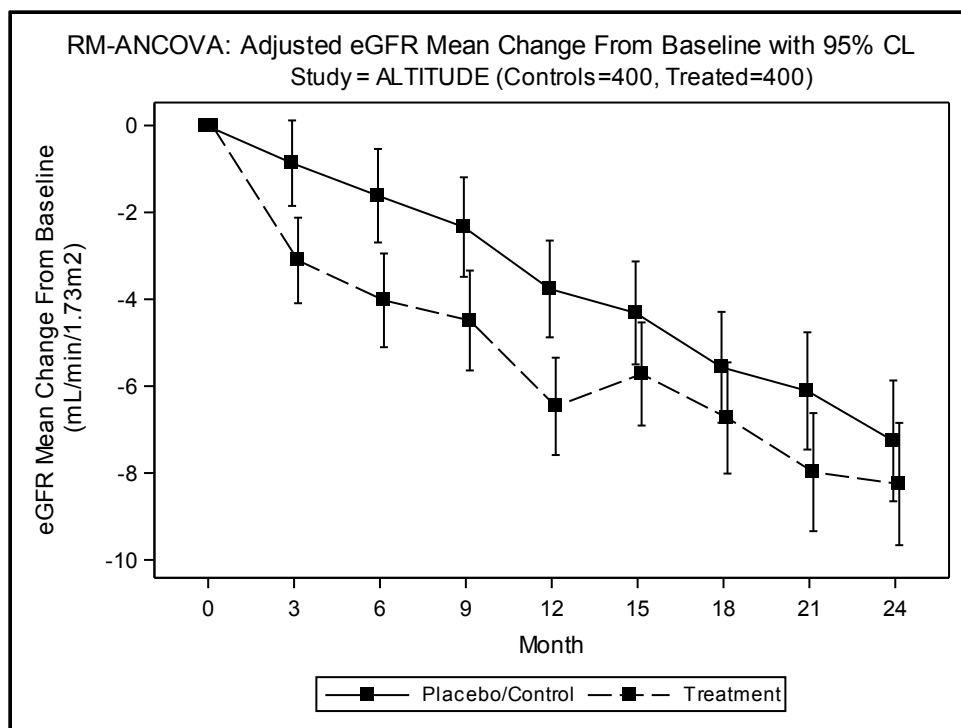


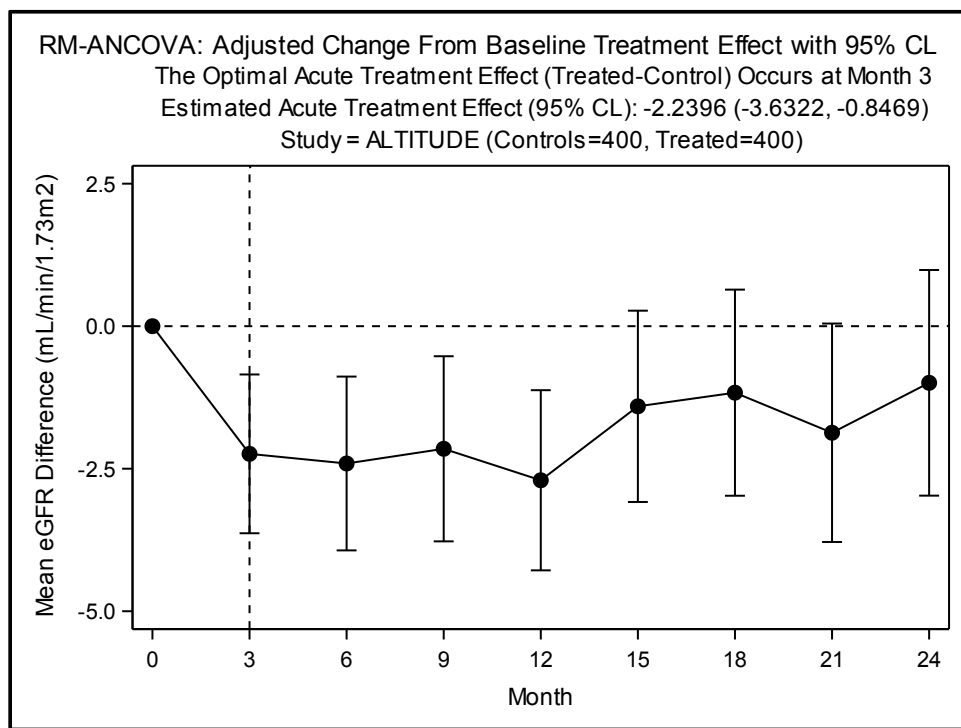
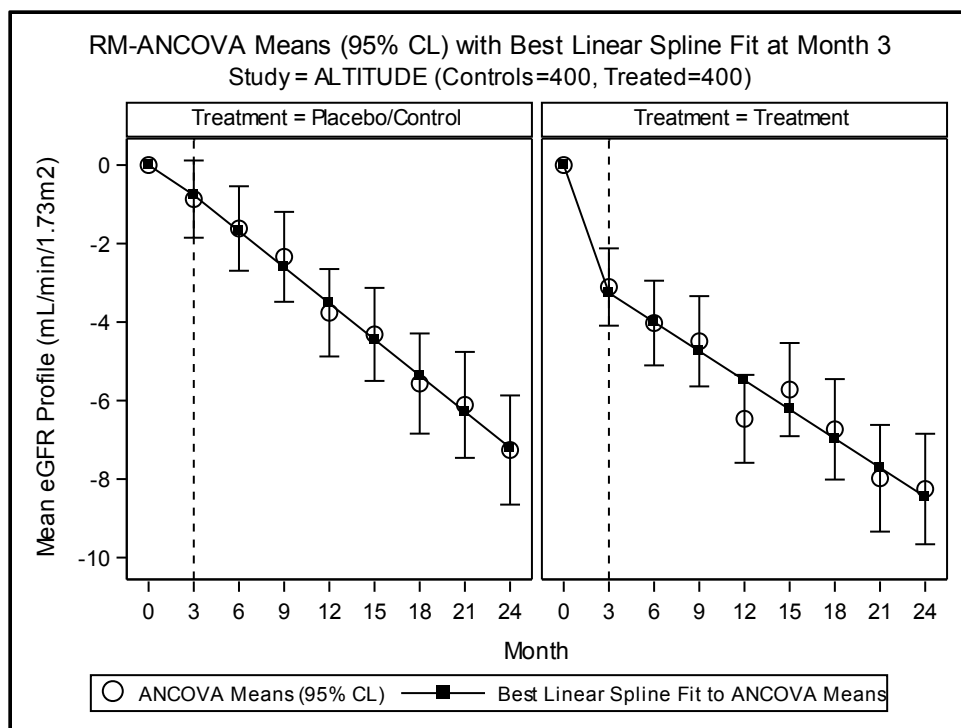


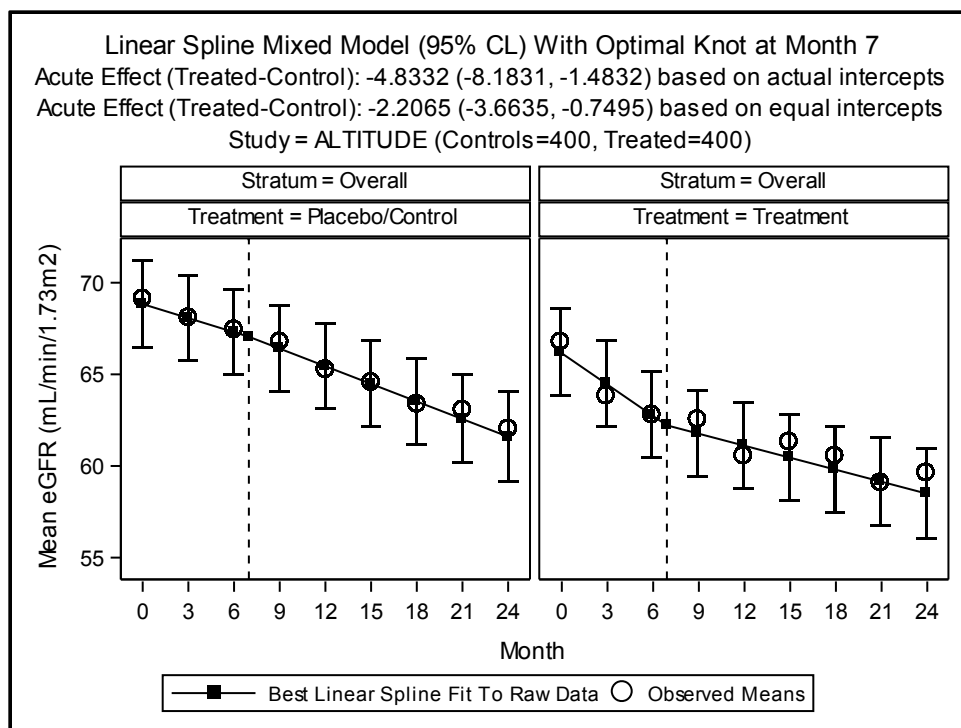


Appendix 1: ALTITUDE results Case 2: Sample Size=400 per Treatment group

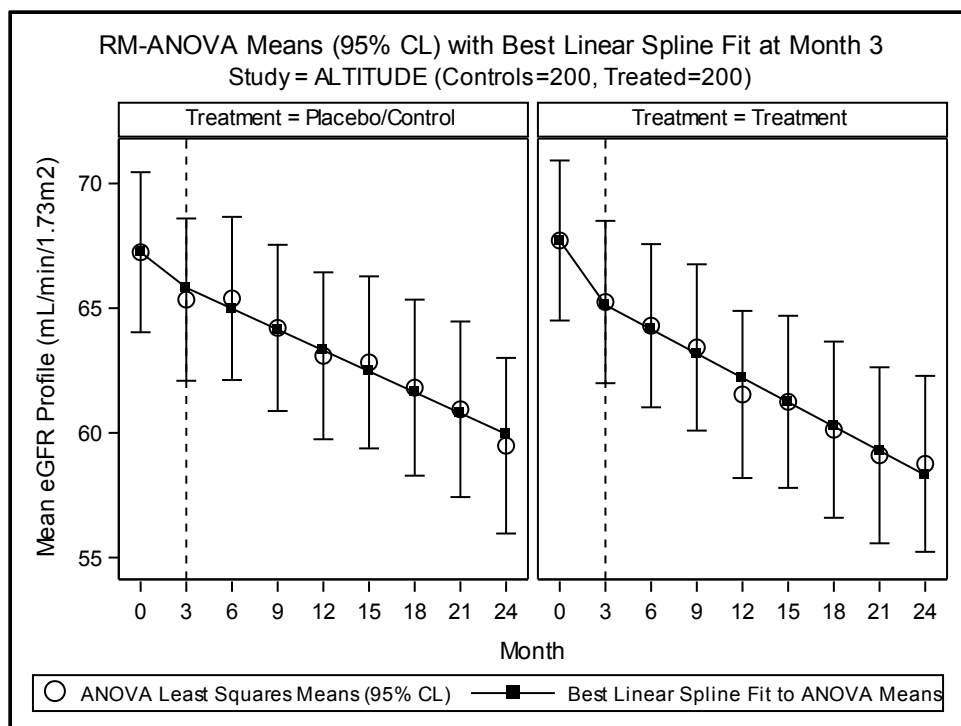
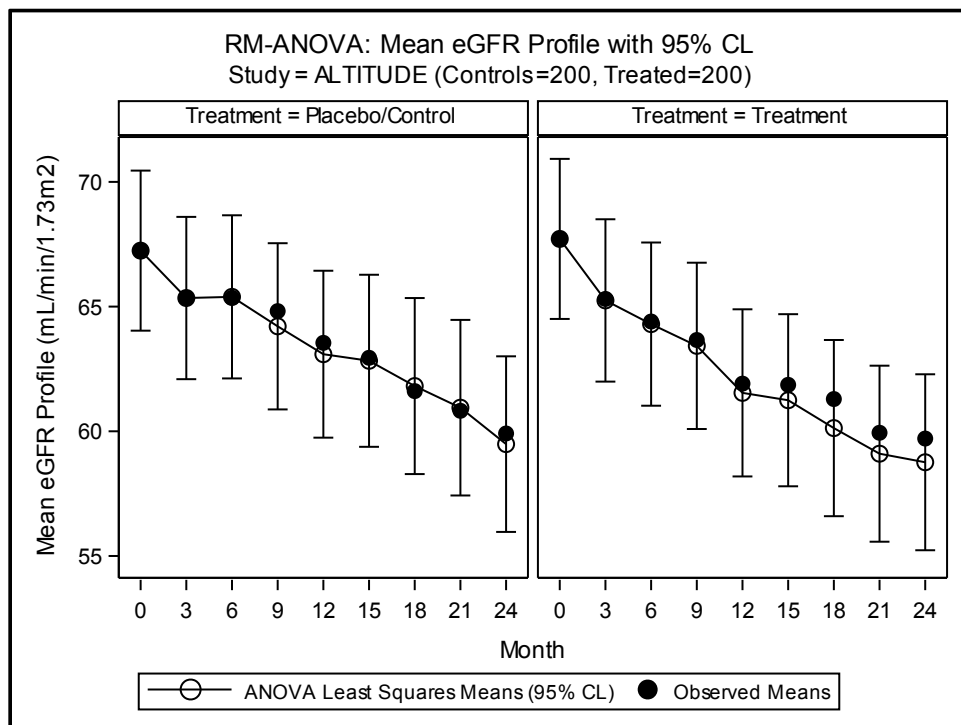


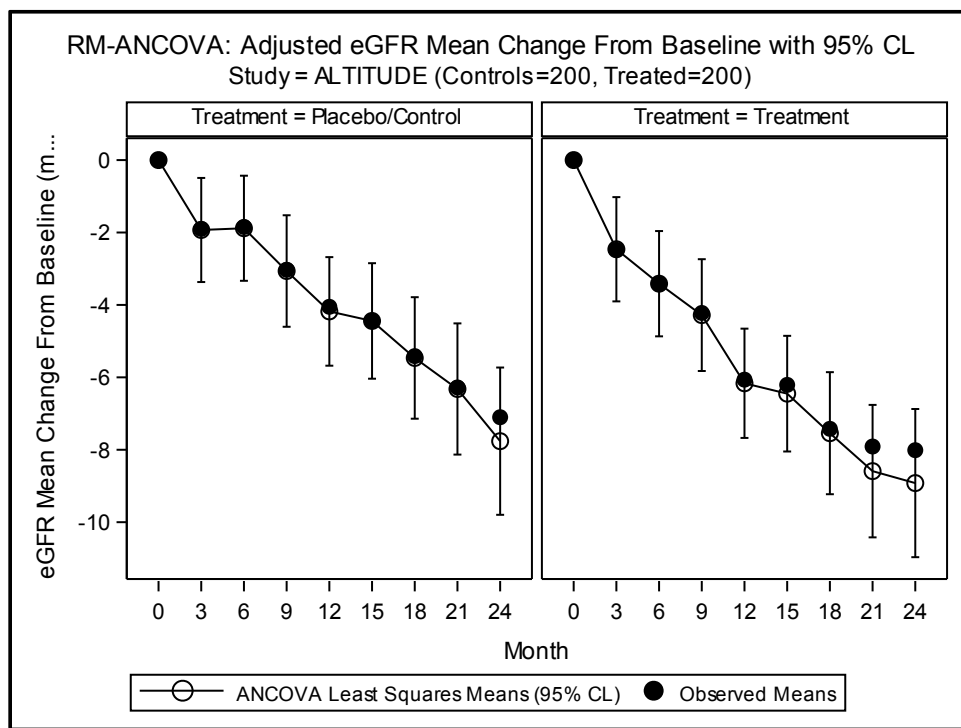
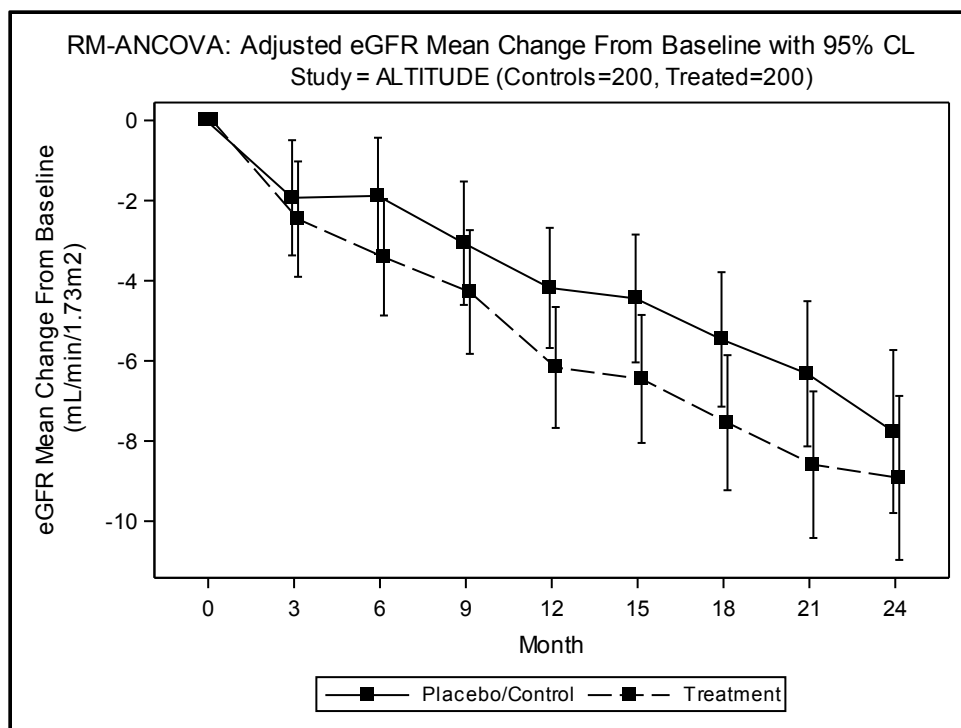


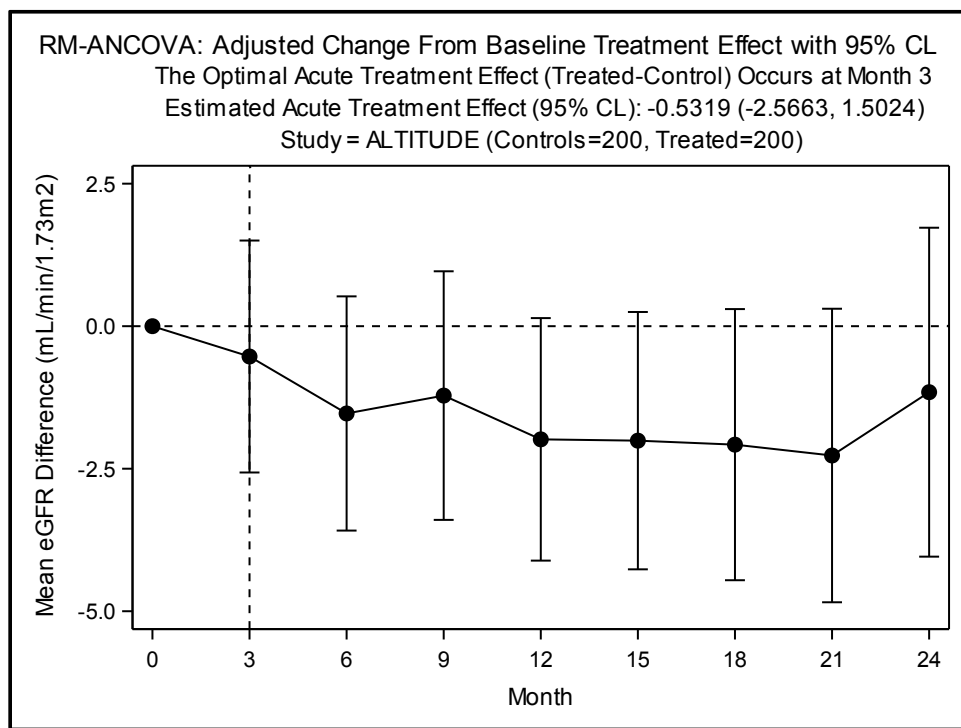
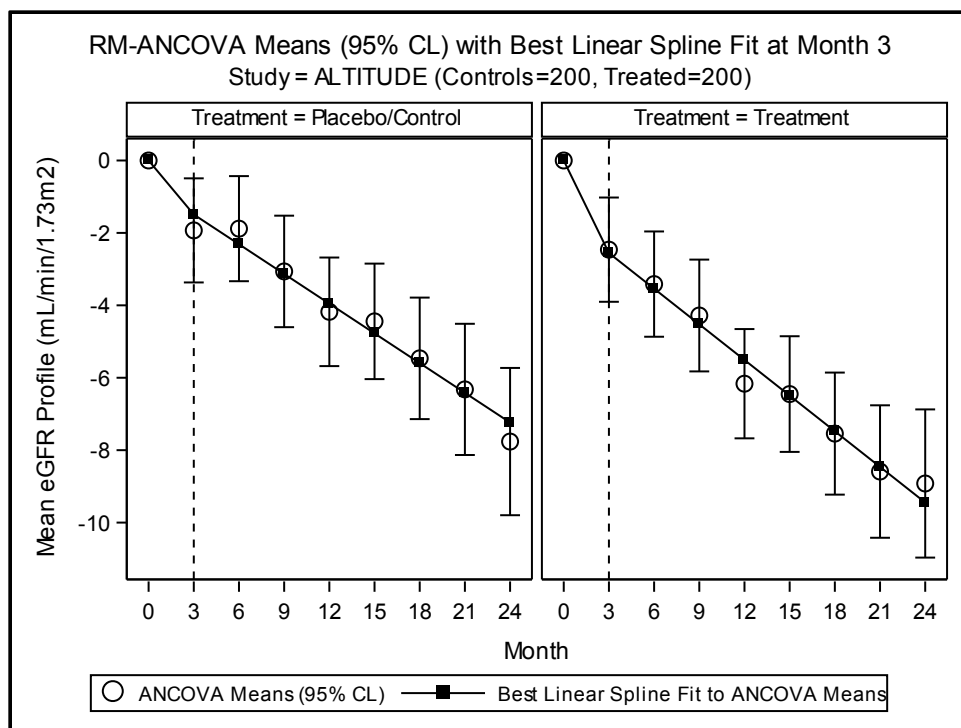


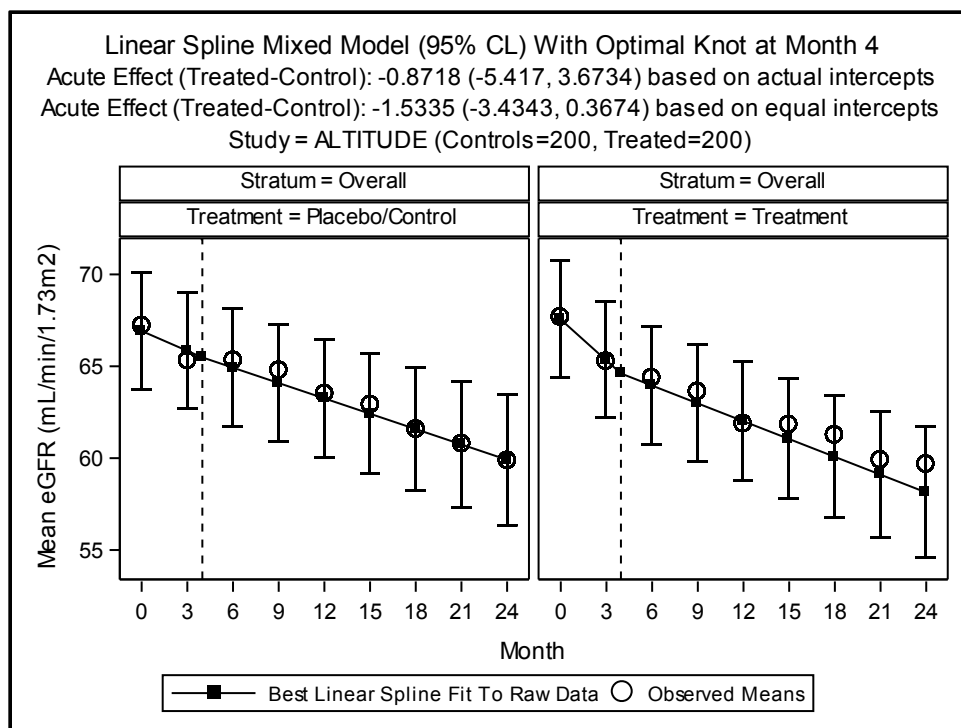


Appendix 1: ALTITUDE results Case 3: Sample Size=200 per Treatment group



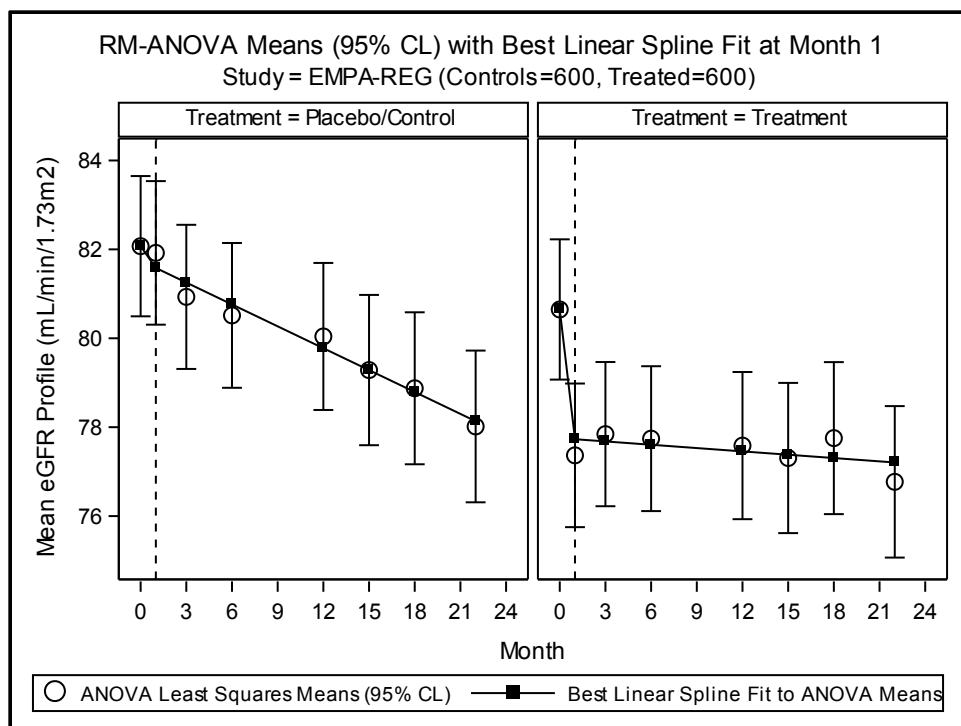
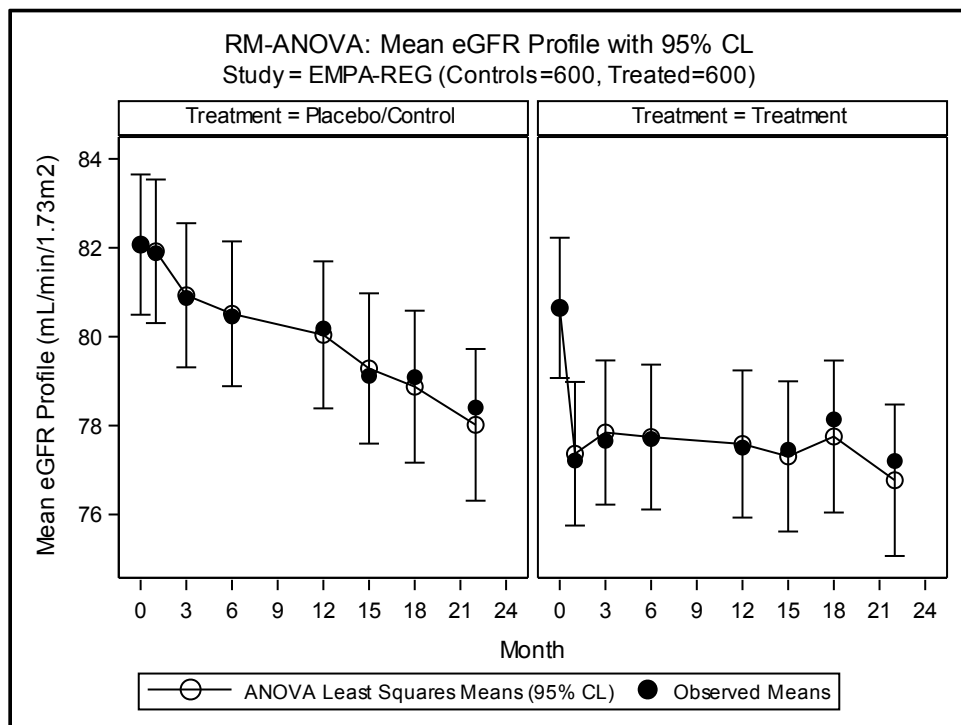


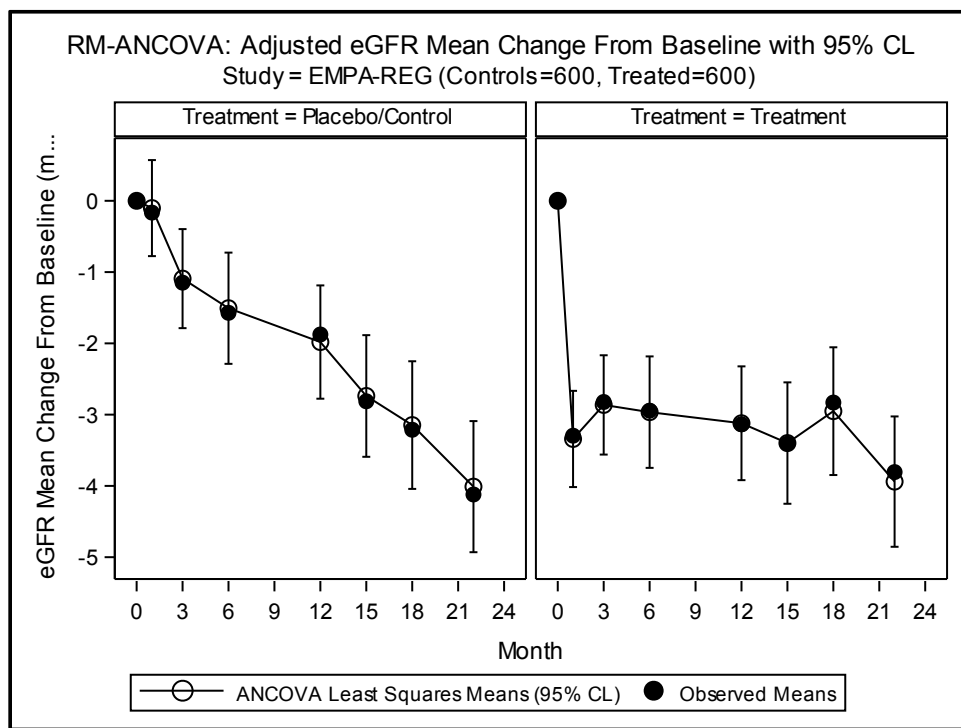
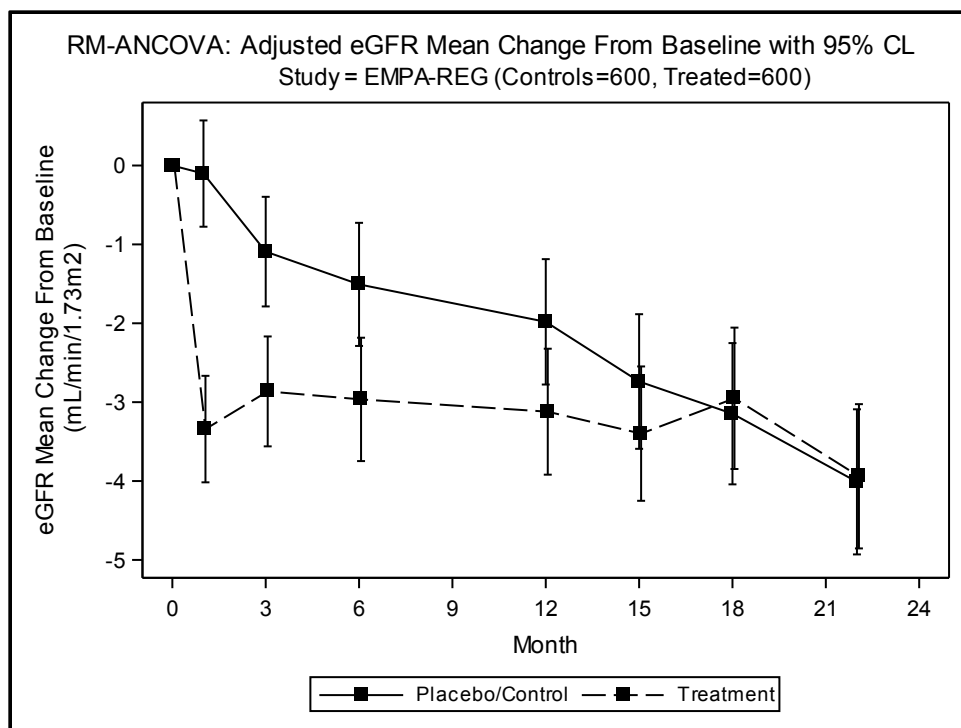


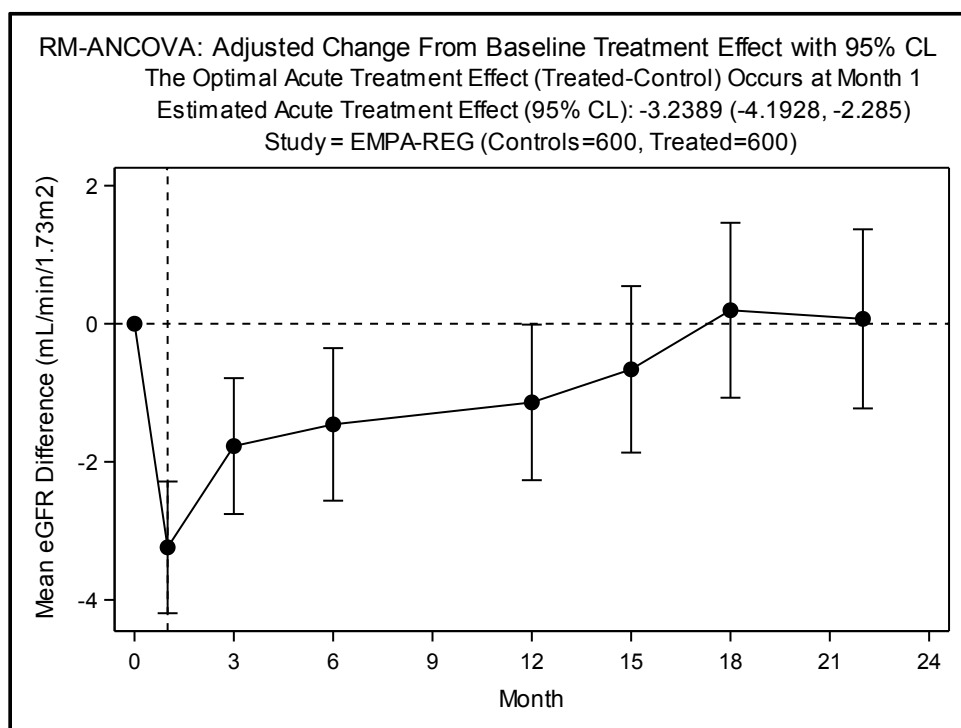
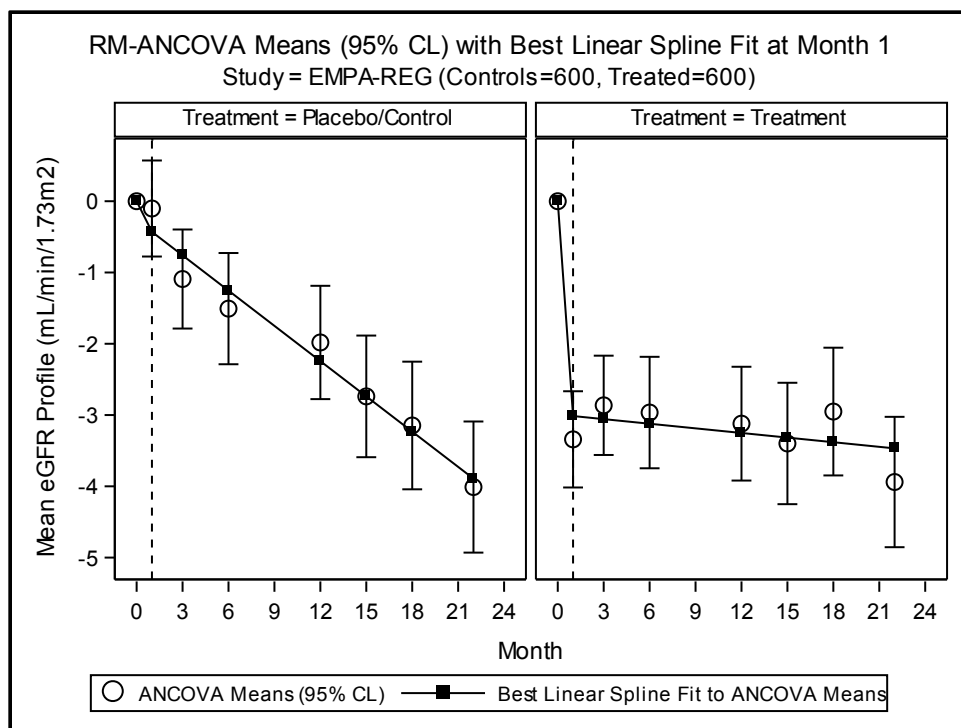


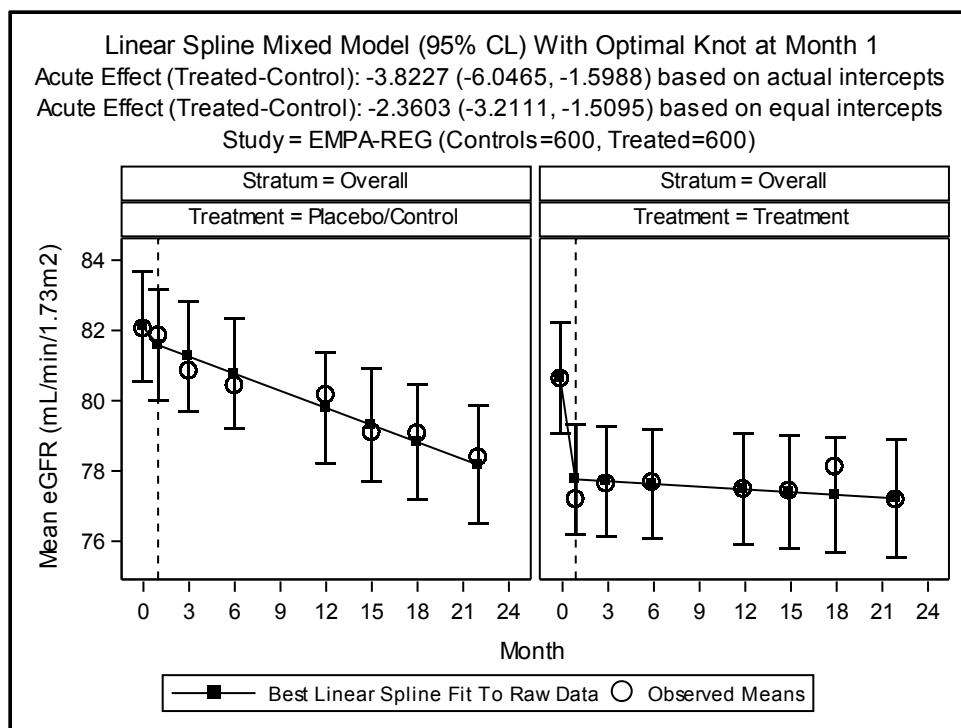
Appendix 2: EMPA-REG results

Case 1: Sample Size=600 per Treatment group



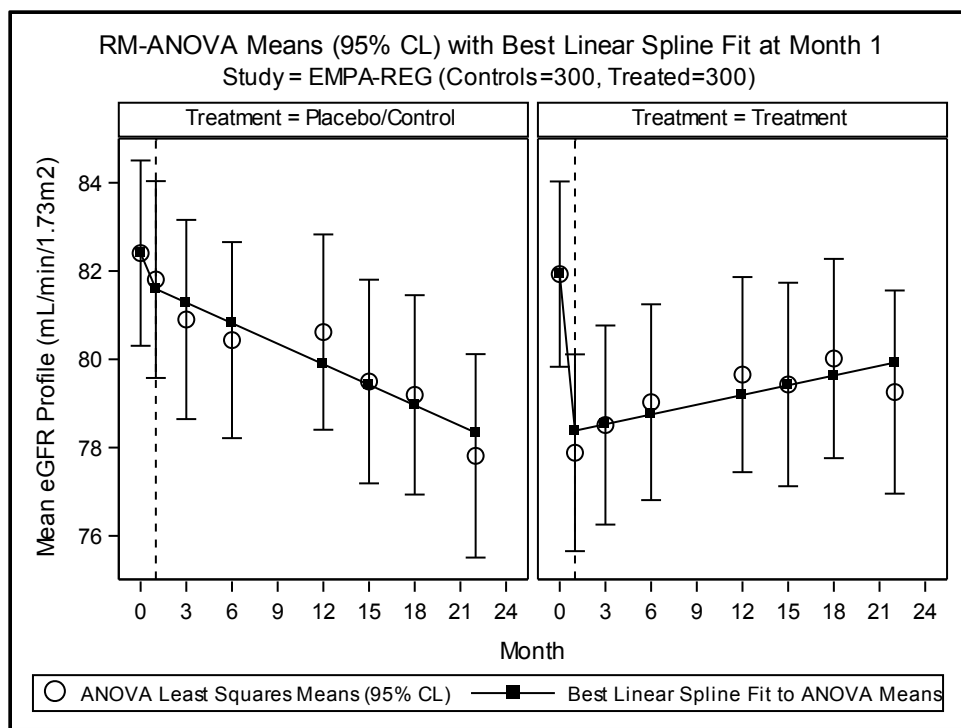
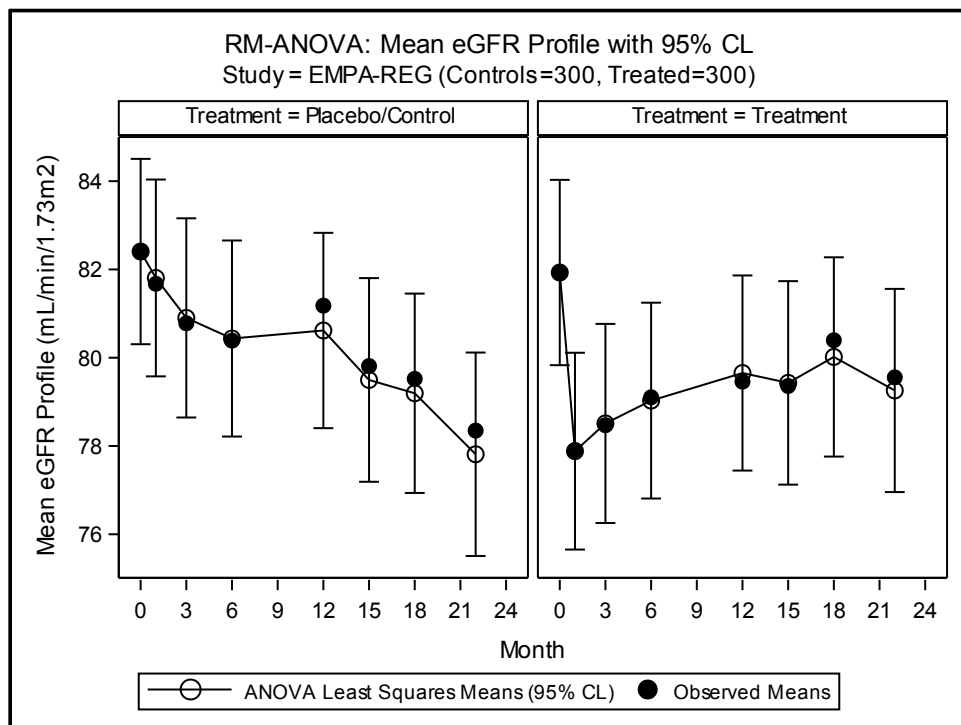


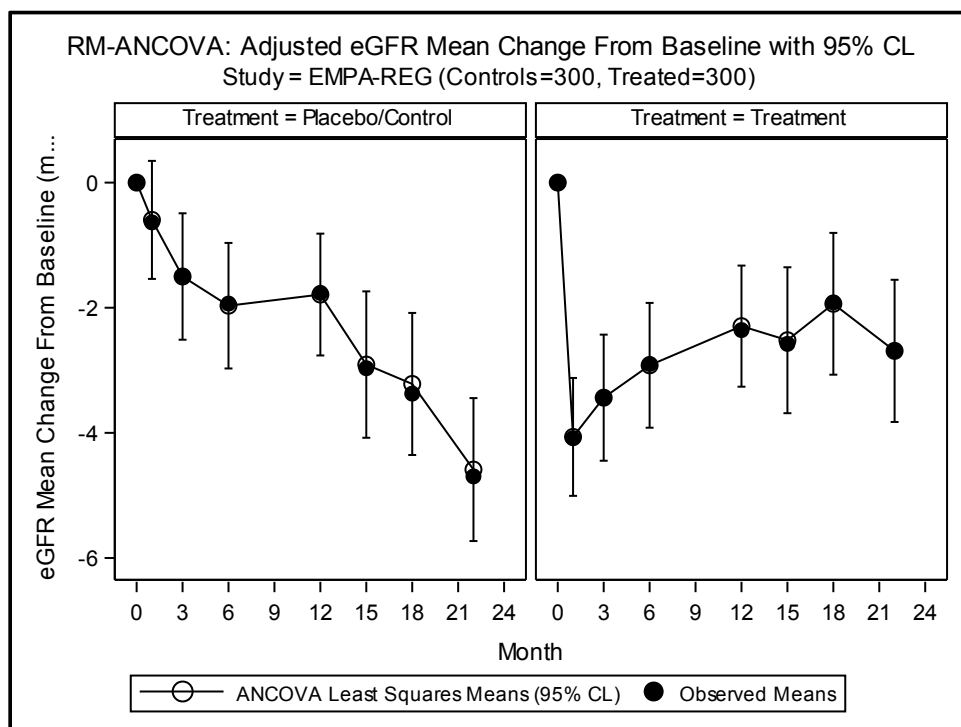
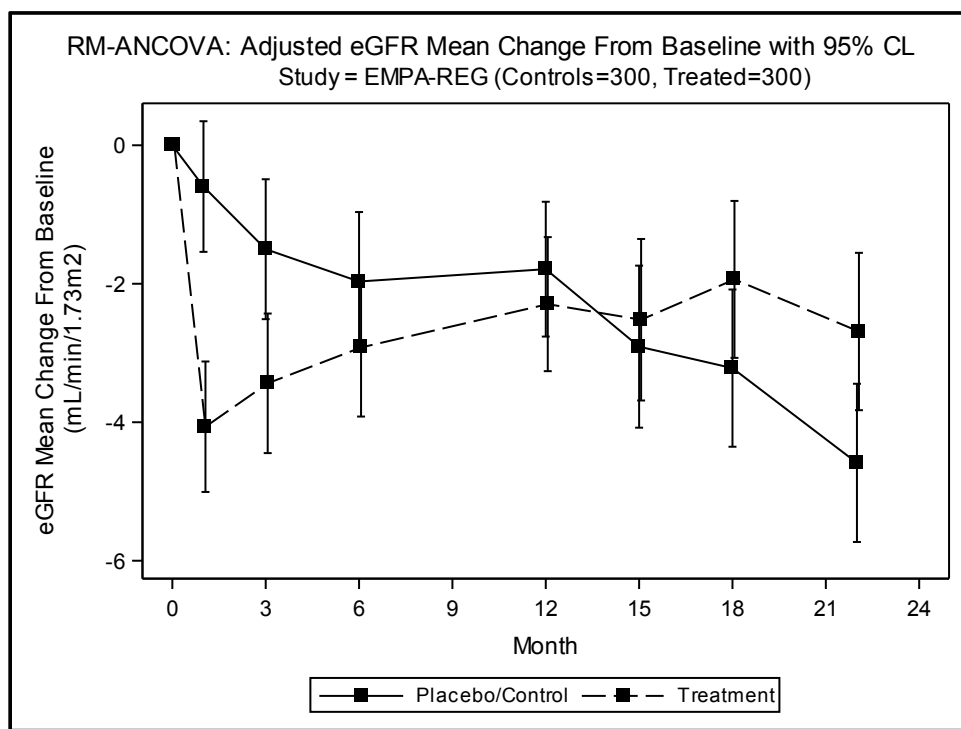


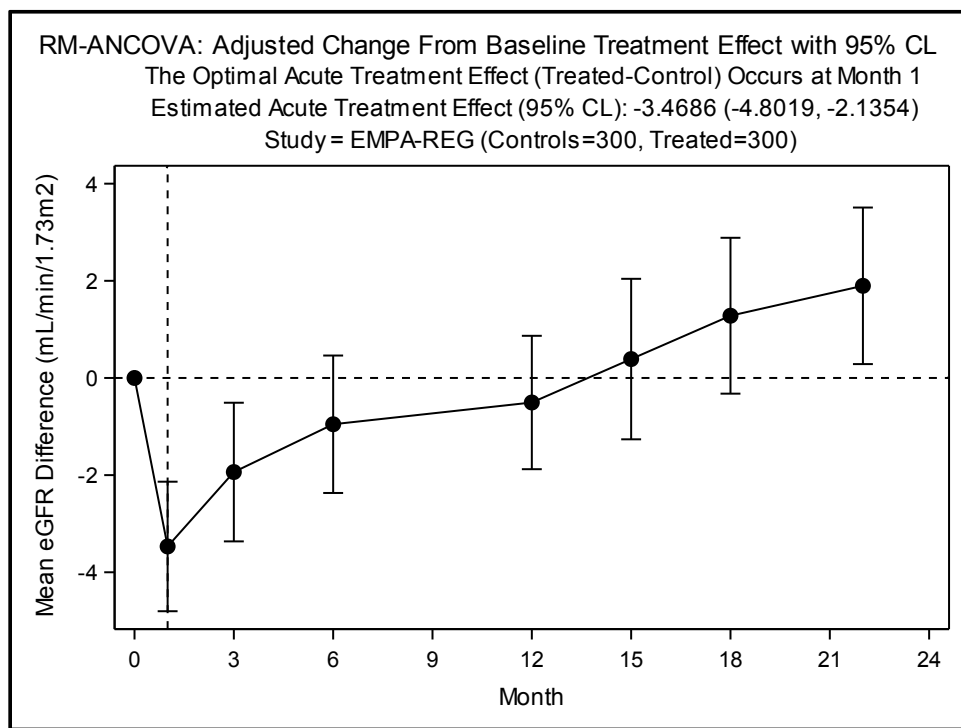
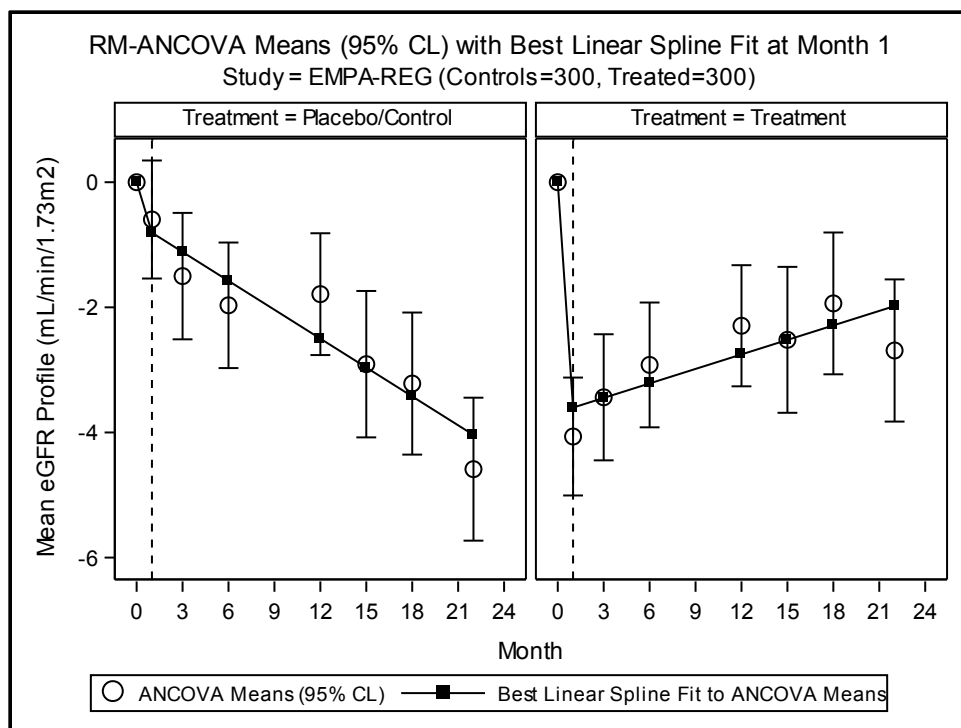


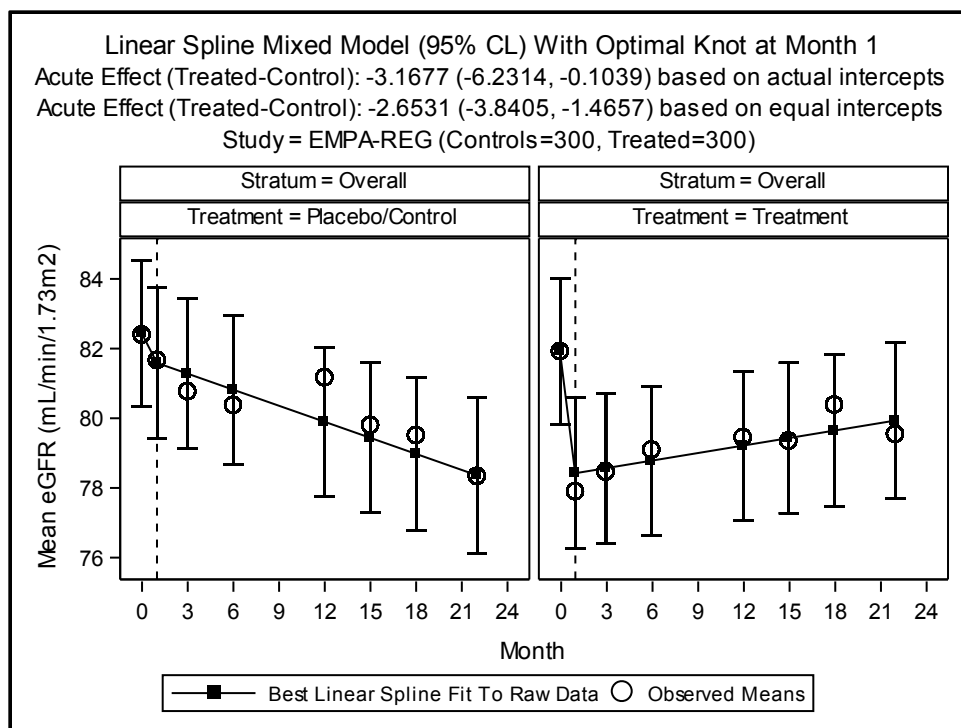
Appendix 2: EMPA-REG results

Case 2: Sample Size=300 per Treatment group



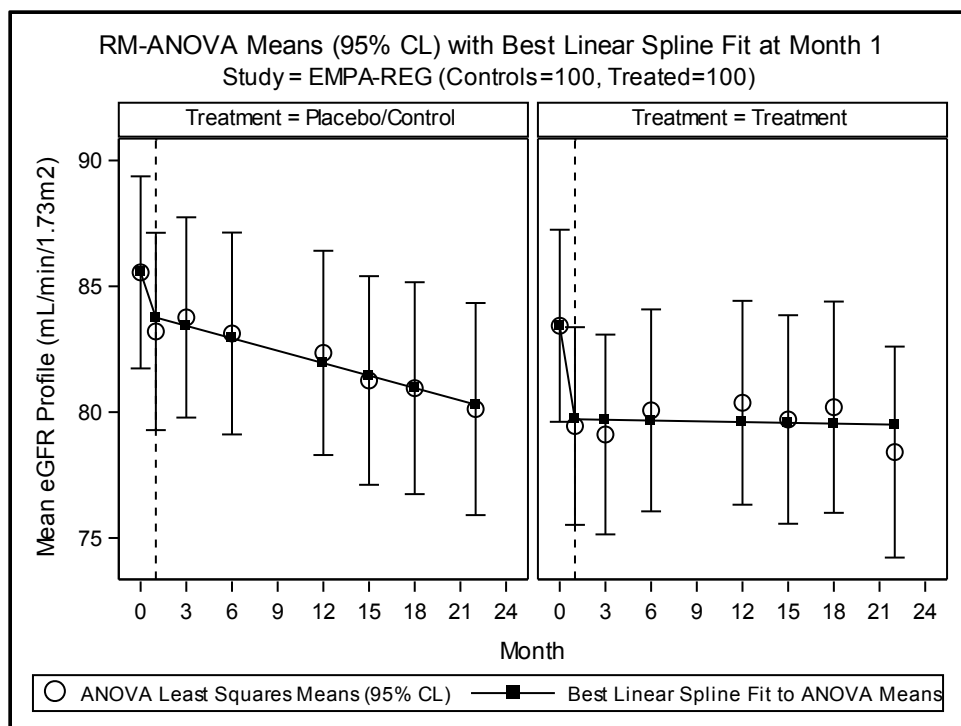
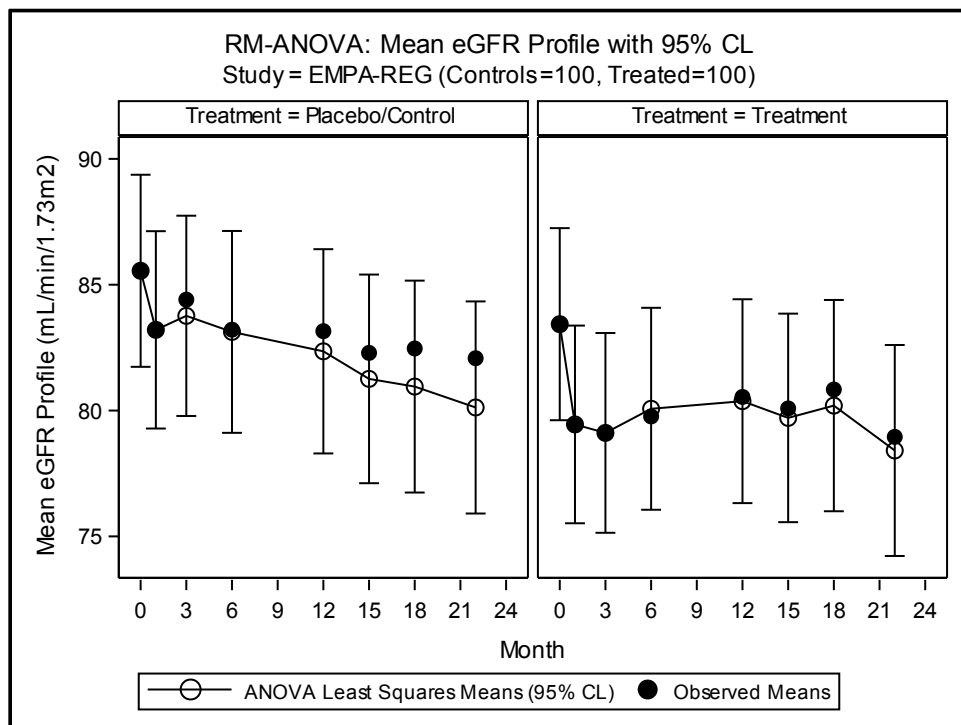


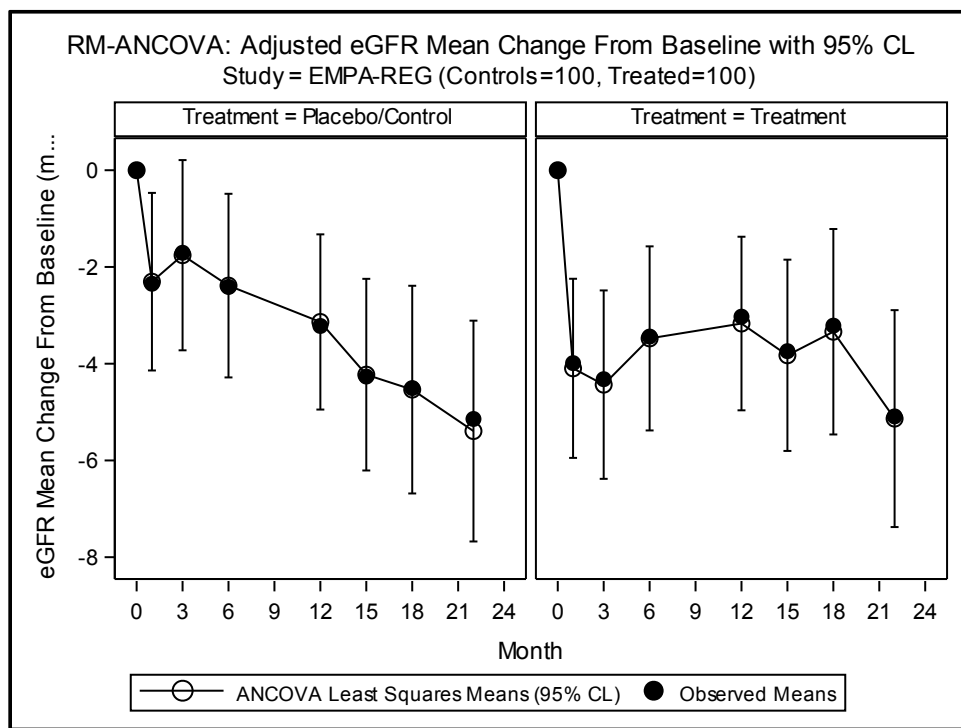
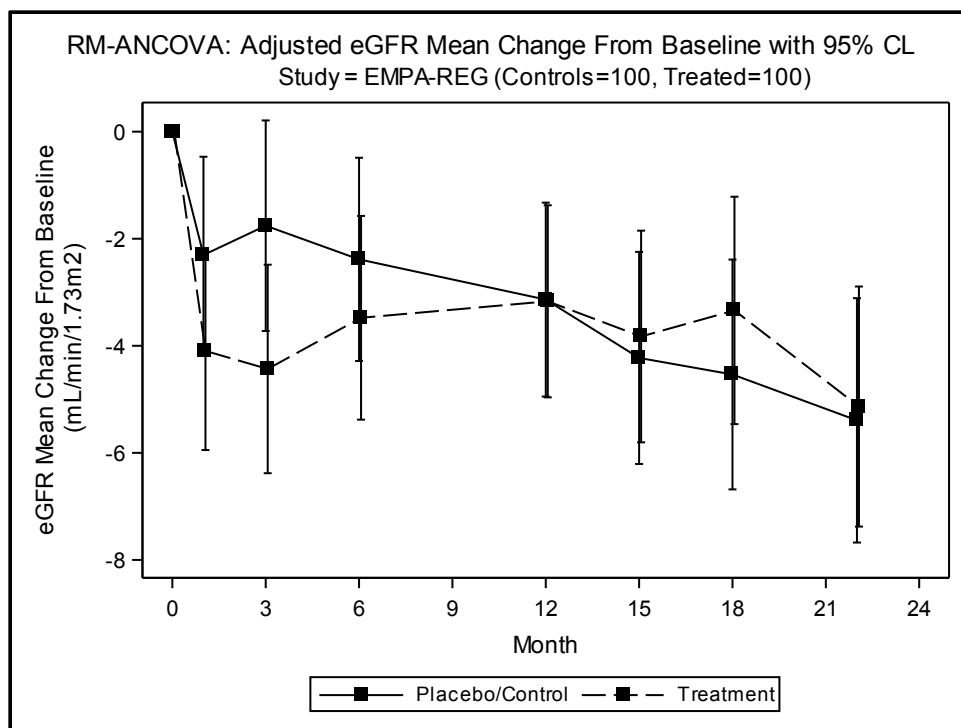


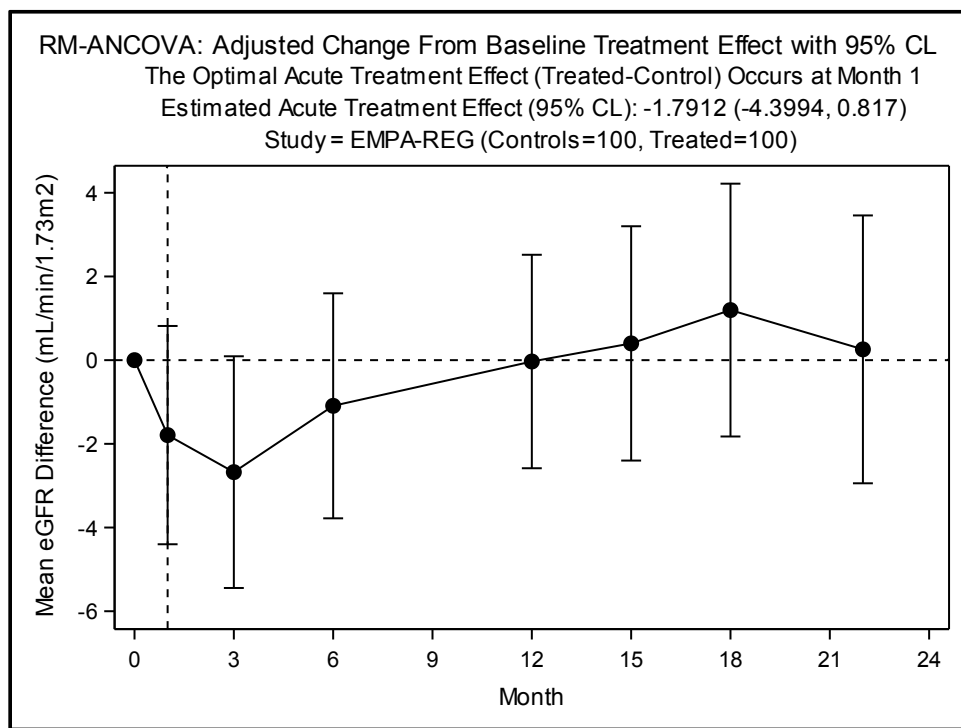
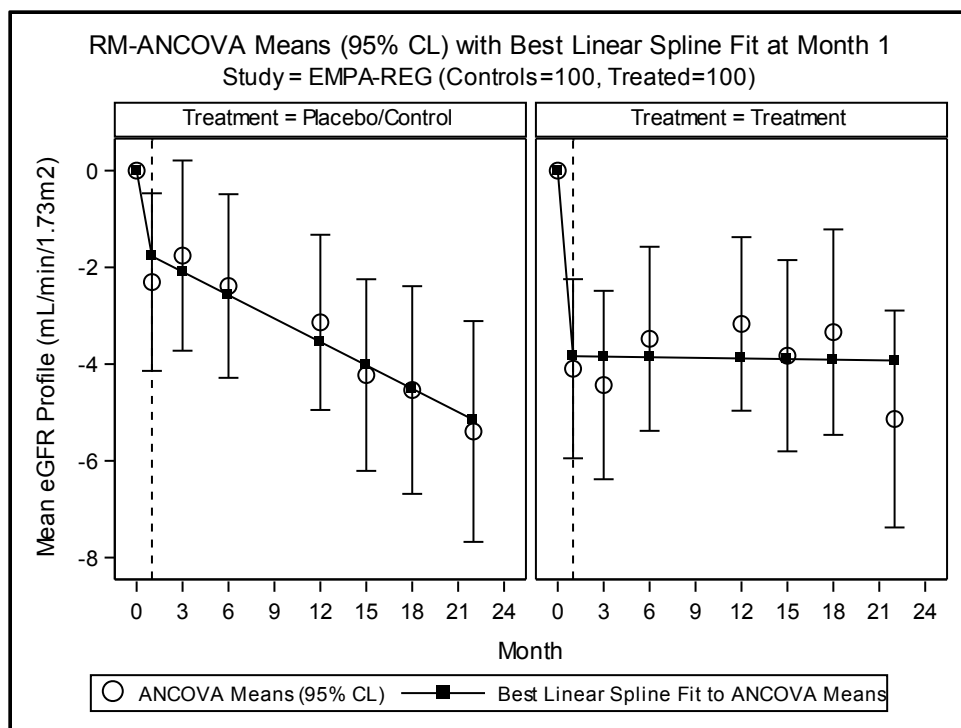


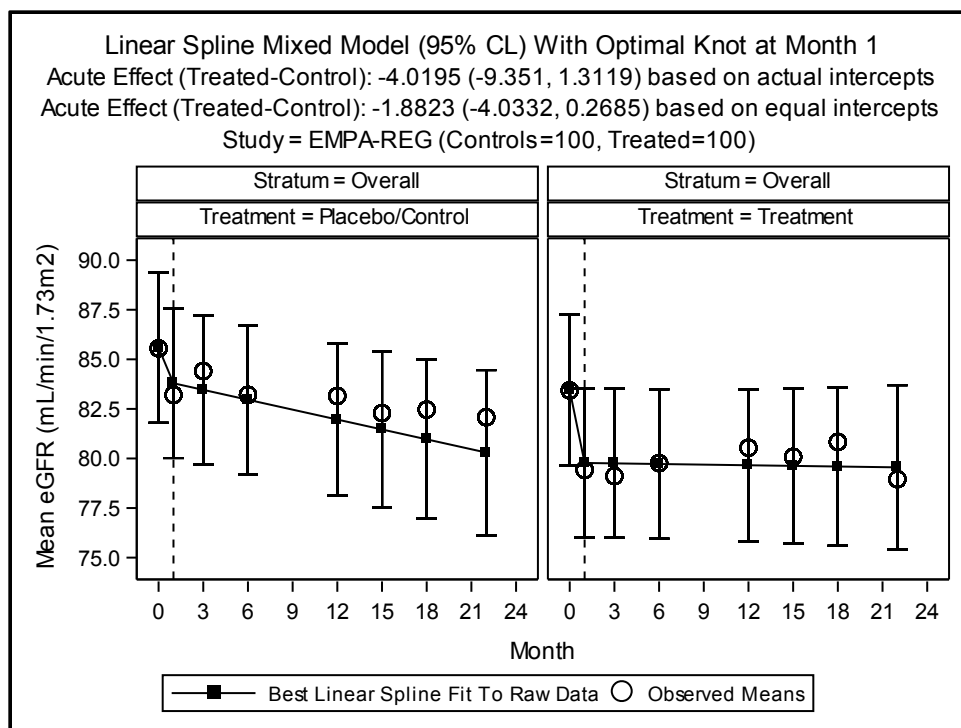
Appendix 2: EMPA-REG results

Case 3: Sample Size=100 per Treatment group









Appendix 3: Zuchelli results

Sample Size=61 Control, 60 Treatment

