NCT # NCT02273726

FibroGen, Inc.

Protocol Number: FGCL-4592-064
A Phase 3, Open-Label, Randomized, Active-Controlled Study of the Efficacy and Safety of Roxadustat in the Maintenance Treatment of Anemia in Subjects with End Stage Renal Disease (ESRD) on Stable Dialysis

Amendment 2

STATISTICAL ANALYSIS PLAN

Version: **Final 1.0**
Release Date: **14Oct2018**

Approvals

I have reviewed and accept the information in this document to be a true and accurate representation of the Statistical Analysis Plan for Study FGCL-4592-064.

Initiator:

Signature: [redacted] Date: [redacted]

Reviewed by:

Signature: [redacted] Date: [redacted]

Signature: [redacted] Date: [redacted]

Signature: [redacted] Date: [redacted]

Signature: [redacted] Date: [redacted]

Signature: [redacted] Date: [redacted]

Signature Significance

The following significance is lent to the signatures on the Approvals page of this document.

Signature	Significance
Initiator	By signing, the author is attesting that the content of the document is complete and accurate.
Reviewer	By signing, the reviewer is attesting that the document's approach and contents are compliant with the study protocol, all appropriate, regulatory requirements, and other significant guidelines. This individual(s) has reviewed the document for accuracy and completeness.

Table of Contents

List of Abbreviations

AE	Adverse Event
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ATC	Anatomical Therapeutic Class
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
CRP	C-Reactive Protein
ECG	Electrocardiogram
EDC	Electronic Data Capture
EPO	Epoetin
ESA	Erythropoiesis-stimulating agent
ESRD	End Stage Renal Disease
FAS	Full Analysis Set
FDA	US Food and Drug Administration
GCP	Good Clinical Practice
HD	Hemodialysis
hs-CRP	High Sensitivity C-Reactive Protein
ICH	International Conference on Harmonization
IDMC	Independent Data Monitoring Committee
IU	International Unit
IV	Intravenous
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
NI	Non-Inferiority
NID	Newly Initiated Dialysis
OL	Open-Label
PCS	Potentially Clinically Significant
PD	Peritoneal Dialysis
PPS	Per-Protocol Set
RBC	Red Blood Cell
QOL	Quality of Life
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Subcutaneous

SOC	System Organ Class (used in MedDRA dictionary)
TEAE	Treatment Emergent Adverse Event
TLF	Tables, Listings, and Figures
TIW	Three times weekly
US	United States
WHOART	World Health Organization Adverse Event Terminology
WHODD	World Health Organization Drug Dictionary
WHO	World Health Organization

1 INTRODUCTION

This statistical analysis plan (SAP) provides a more technical and detailed elaboration of the statistical analyses of efficacy and safety as outlined and/or specified in the amendment 2 of study protocol dated 18 August 2017. Specifications of tables, figures, and data listings are contained in a separate document.

A separate pooled-statistical analysis plan (SAP) for pre-specified analyses based on adjudicated safety data will complement this study specific SAP.

2 STUDY OBJECTIVES

2.1 Primary Objectives

Evaluate the efficacy and safety of roxadustat compared with active control (epoetin alfa) for the maintenance treatment of anemia in ESRD subjects on stable dialysis.

2.2 Secondary Objectives

- Evaluate the utilization of intravenous (IV) iron with roxadustat compared with active control (epoetin alfa)
- Evaluate the effect of roxadustat on serum lipid parameters compared with active control (epoetin alfa)

3 STUDY DESIGN

This Phase 3, multi-center, open-label, randomized, active-controlled study is designed to evaluate the efficacy and safety of roxadustat compared to epoetin alfa for the maintenance of hemoglobin levels in subjects on stable hemodialysis (HD) or peritoneal-dialysis (PD) and newly initiated dialysis (NID) originally on erythropoiesis-stimulating agent (ESA) for treatment of anemia.

A total of up to 1200 subjects are planned to be randomized to receive roxadustat or epoetin alfa (active control) in a 1:1 ratio, respectively.
For subjects randomized to roxadustat, the initial roxadustat dose will be determined using a conversion table based on the subject's previous average weekly prescribed ESA dose in the last 4 weeks prior to randomization if the subject is on epoetin or darbepoetin, and average monthly (4 weeks) prescribed ESA dose in 8 weeks prior to randomization if the subject is on Mircera® (Table 1).

If the mean qualifying screening hemoglobin value at randomization is less than the lower limit of the target hemoglobin range (i.e. <10 g/dL), the starting roxadustat dose will be increased by one dose step. If the converted initial dose exceeds the maximum dose of 3.0 mg/kg then the lower dose step should be chosen as the initial dose. Roxadustat will be dosed three times weekly (TIW) for the entire duration of the Treatment Period. Dosing frequency may only be adjusted to BIW or QW if a subject requires < 20 mg TIW (i.e., < 60 mg per week) to maintain an Hb level of approximately 11 g/dL.

Table 1. Initial Dosing of Roxadustat: Conversion Table from ESAs to Roxadustat

Epoetin (i.e., alfa, beta, theta, zeta, delta, or omega) [a] (IU/week)	Darbepoetin alfa[a, b] (µg/week)	Mircera®[c] (µg/ monthly)	Roxadustat Dose[d] (mg/dose) TIW
< 5,000	< 25	< 80	70
5,000 to 8,000	25-40	80-120	100
> 8,000 to 16,000	> 40-80	> 120-200	150
> 16,000	> 80	> 200	200

a Average weekly prescribed dose in last 4 weeks prior to randomization
b If darbepoetin is used biweekly, use half the dose to get per week dose
c Average prescribed monthly (4-wks) dose in last 8 weeks prior to randomization
d Starting dose will be one step higher if the mean hemoglobin at randomization is < 10.0 g/dL, however, not to exceed 3.0 mg/kg/dose.

Notes: The roxadustat dose steps are 20, 40, 50, 70, 100, 150, 200, 250, 300, and 400 mg.

For subjects randomized to epoetin alfa, the initial epoetin alfa dose will be determined using a conversion table based on the subject's weekly average prescribed ESA dose in the last 4 weeks prior to randomization if on epoetin or darbepoetin, and average monthly (4-week) prescribed ESA dose in 8 weeks prior to randomization if on Mircera® (Table 2).

All hemodialysis subjects randomized to Epoetin Alfa arm will receive epoetin alfa intravenously TIW during the Treatment Period. For PD and HD subjects, the route of administration should remain same as baseline. However, if the investigator decides to change the route of administration from subcutaneous (SC) to IV after randomization that should be done during the early part of the treatment phase (preferably prior to Week 16).

In case of a change in route of administration from SC to IV, the initial dose of IV epoetin alfa will be determined by the investigator per local standard of care.

Table 2. Initial Dosing of Epoetin Alfa: Conversion Table from Non-Epoetin ESAs to Epoetin Alfa

Baseline ESA	Conversion Ratio	Examples of Converted Initial Epoetin Alfa dose (approximate) IU/week[c]
Epoetin (i.e., alfa, beta, theta, zeta, delta, or omega) (IU/week) [a]	x 1	6,000 IU/week x 1 = 6,000 IU/week
Darbepoetin alfa (µg/week) [a]	x 200	40 µg/week x 200 = 8,000 IU/week
Mircera® (µg/once monthly) (i.e., 4-wks) [b]	x 70 – 80 (for lower Mircera® dose lower conversion ratio may be used)[d]	100 µg/month x 70 = 7,000 IU/week 200 µg/month x 80 = 16,000 IU/week

a Mean weekly ESA in 4 weeks prior to randomization
b Mean monthly (4-week) ESA in 8 weeks prior to randomization
c May be rounded as deemed necessary by the Investigator
d Per discretion of the Investigator

The study periods are as follows:

- **Screening Period:** Up to 6 weeks. For subjects currently taking Mircera®, the screening period may be extended up to 8 weeks.
- **Treatment Period:** Treatment duration is variable for individual subjects with a maximum treatment duration could be up to approximately 3 years from the date the last subject is randomized.
- **Post-Treatment Follow-Up Period:** 4 weeks

During the Treatment Period, subjects will attend weekly study visits from Day 1 to Week 2, followed by every two weeks study visits from Weeks 4 to 24. Following Week 24, study visits will occur every four weeks until the End of Treatment (EOT). After EOT, subjects proceed to the post-treatment Follow-up Period and will return for the End of Study (EOS) visit. The schedule of assessments for the study is provided in Appendix 1.

All subjects discontinuing study medication prematurely will be followed up for vital status, cardiovascular events, and hospitalization until study closure, unless consent to participate is withdrawn. Upon completion of ET and 4-week Follow-up (EOS) visits, these subjects will be followed up every 3-6 month interval (depending on the availability of subjects) until study closure. These visits may occur either in-person or via telephone.

3.1 Dose Adjustments

During the Treatment Period, roxadustat dose adjustment will be made according to the dose adjustment algorithm in Appendix 2, in order to maintain a hemoglobin level of approximately 11 g/dL, while closely monitoring the rate of rise of hemoglobin and hemoglobin levels.

The dose of roxadustat will remain constant during the first 4 weeks of the Treatment Period unless a dose reduction is required for predefined out of range hemoglobin elevations. Roxadustat dose adjustments are permitted from week 4 onwards, and every 4 weeks

thereafter i.e. Week 4, Week 8, Week 12, etc.; however, the dose may be adjusted between the two pre-specified windows if the following two criteria are met:

- No dose adjustment has been made in last 4 weeks
- Hemoglobin < 9.0 g/dL

In this study, a rate of rise of hemoglobin > 2 g/dL within 4 weeks or a hemoglobin level of ≥ 13 g/dl at any time would be considered as predefined out of range hemoglobin elevations. For a rate of rise of hemoglobin > 2 g/dL within 4 weeks, dose should be reduced by one dose step and for a hemoglobin value ≥ 13.0 g/dL, the dose should be on hold until hemoglobin drops < 12.0 g/dL per guidelines provided in the protocol (Appendix 2). When roxadustat will be resumed, the dose will be reduced by one dose step.

Dose adjustments or temporary dose holds for predefined out of range hemoglobin elevations can occur at any time during the Treatment Period. Any dose adjustment will reset the dose-adjustment window to every 4 weeks thereafter (e.g., dose adjustment for a qualified reason at Week 6 leads to next dose adjustment at Week 10). If a dose adjustment review is scheduled on Week 22, then the next dose adjustment review will be scheduled on Week 28, since there is no scheduled visit on Week 26.

3.2 Rescue Therapy & Emergency Procedures

Red Blood Cell Transfusion (for all subjects):

For subjects in both treatment groups, red blood cell (RBC) transfusion is allowed if rapid correction of anemia is required to stabilize the subject's condition (e.g. acute hemorrhage) or the investigator is of the opinion that the blood transfusion is a medical necessity. Study treatment may continue during or after RBC transfusion administration.

ESA use*:*
For subjects randomized to roxadustat the use of ESAs is generally prohibited. ESA rescue is restricted to no more than one 4-week cycle of use during the Treatment Period; the Investigator may initiate use of an approved epoetin (EPO) analogue if all of the following criteria are met:

- A subject's hemoglobin level has not sufficiently responded to two or more dose increases or the maximum dose limit of the study drug has been reached; and
- The subject's hemoglobin is < 8.5 g/dL on two consecutive measurements (central lab) drawn at least five days apart; and
- Clinical judgment does not suggest iron deficiency or bleeding as a cause of lack of response or rapid decline in hemoglobin; and
- Reducing the risk of alloimmunization in transplant eligible patients and/or reduction of other RBC transfusion-related risks is a goal

The subject is not allowed to be administered both the EPO analogue and roxadustat at the same time. Treatment with an EPO analogue should be started ≥ 3 days after the last dose of roxadustat, and should be stopped when hemoglobin > 9 g/dL or after a 4-week cycle has been completed, whichever comes first. If a subject requires longer than 4-week therapy due to inadequate response, the Medical Monitor should be contacted. Study treatment should be resumed after the following intervals:

- Two days after stopping epoetin
- One week after stopping darbepoetin alfa
- Two weeks after stopping methoxy polyethylene glycol-epoetin beta (Mircera®)

If more than one cycle of EPO analogue rescue is required, the Investigator should permanently discontinue roxadustat. Use of EPO analogues will be recorded in the eCRF. Roxadustat subject who receives an ESA inadvertently and does not meet the criteria above may be allowed to continue taking study medication, if considered safe by the Investigator or Medical Monitor.

For subjects randomized to epoetin alfa, the Investigator may initiate use of a different EPO analogue if clinically indicated. Use of the different EPO analogue will be recorded in the eCRF and will be considered rescue therapy.

Emergency Procedures
Therapeutic Phlebotomy:

If there are clinical concerns for a subject's high hemoglobin levels, the investigator may decide to perform a therapeutic phlebotomy in addition to a temporary study treatment dose hold – this should be documented and discussed with the study Medical Monitor.

Supplemental Iron

Oral Iron Supplementation:

In this study, all subjects will be encouraged to take oral iron as the first-line iron supplementation. The dose and frequency are at the discretion of the principal Investigator. Oral iron therapy should be started before the subject becomes iron depleted.

Intravenous Iron Supplementation:

Intravenous iron supplementation is permitted if in the opinion of the Investigator the subject has not responded adequately while taking oral iron or cannot tolerate oral iron, AND is considered iron deficient as determined by either ferritin < 100 ng/ml or TSAT < 20%.

A dose of up to 250 mg IV iron per dose administration cycle may be administered per regional package insert. A list of recommended IV irons is listed in the protocol (Appendix 6). Three to four weeks following the last dose of IV iron infusion, central lab ferritin and TSAT levels and hemoglobin level will be assessed; a dose of IV iron may be repeated only if hemoglobin response is still lacking, and either ferritin < 100 ng/ml or TSAT < 20%. If there is no hematopoietic effect with the first administration of IV iron but with accompanying increases in TSAT and ferritin, then the Investigator should exercise judgment whether additional IV iron should be given. Treatment with study medication will continue during IV iron administration. Discontinuation of IV iron supplementation is recommended once subject is no longer iron deficient (e.g. ferritin >= 100 ng/mL and TSAT >= 20%).

In Protocol Amendment 2, above IV Iron administration guidelines have been revised – intravenous iron supplementation is permitted if in the opinion of the Investigator the subject has not responded adequately and is considered iron deficient. IV iron may be administered per local standard of care as deemed necessary by the investigator.

4 STUDY ENDPOINTS

4.1 Primary Efficacy Endpoints

For US (FDA) submission: Hemoglobin change from baseline to the average level during the evaluation period defined as Week 28 until Week 52.

For Ex-US submission: Hemoglobin change from baseline to the average hemoglobin of Weeks 28 to 36, without having received rescue therapy (i.e., RBC transfusion or rescue ESA therapy) within 6 weeks prior to and during this 8-week evaluation period.

Rescue therapy for all subjects is defined as ESA use (for rescue or inadvertent use) or RBC/blood transfusion.

Hemoglobin results obtained from the central laboratory will be used to determine Hemoglobin response.

Baseline hemoglobin is defined as the mean of the at least last three available central laboratory hemoglobin values obtained prior to the first dose, three of the latest screening hemoglobin values plus the predose hemoglobin value collected on Day 1. In subjects with missing Day 1 hemoglobin value, the mean of three latest screening laboratory hemoglobin values will be considered as baseline hemoglobin value.

4.2 Secondary Efficacy Endpoints

The secondary efficacy endpoints in this study are as below for both US and Ex-US submission unless otherwise specified:

- US (FDA) submission: Proportion of subjects with mean hemoglobin level during the evaluation period defined as Week 28 until Week 52 ≥ 10.0 g/dL.
- Ex-US submission: Proportion of subjects with hemoglobin response, defined as mean hemoglobin during Weeks 28 to 36 within the target range of 10.0 to 12.0 g/dL without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
- Change from baseline in Low Density Lipoprotein (LDL) cholesterol to the average LDL cholesterol of weeks 12 to 28.
- Hemoglobin change from baseline to the average level during the Weeks 18 to 24 for subjects with baseline hs-CRP > ULN.
- Average monthly IV iron use per subject during the Treatment Period during Weeks 28 to 52 (monthly defined as a period of 4 weeks).
- Time to first RBC Transfusion during the treatment
- Change from BL in mean arterial pressure (MAP) to the MAP value averaged over weeks 20 to 28
- Time to first exacerbation of hypertension: An increase from BL of ≥ 20 mm systolic blood pressure (SBP) and SBP ≥ 170 mmHg or an increase from baseline of ≥15 mmHg diastolic blood pressure (DBP) and DBP ≥ 100 mmHg during Weeks 28 to 52.

4.3 Additional Evaluation of Efficacy

The additional efficacy evaluations in this study are:

- **Hemoglobin Maintenance:**
  - Hemoglobin long-term Maintenance: Mean change from baseline in hemoglobin averaged over 8 weeks of treatment at Weeks 44-52 without

rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
  - Mean change from baseline in hemoglobin averaged over the Week 96 to 104 of treatment, without rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
  - Change from baseline in hemoglobin at each of the selected post-dosing time points.
  - Proportion of subjects with Hb >=10 g/dL averaged over Weeks 28 to 36, 44 to 52, and 96 to 104 of treatment without rescue therapy within 6 weeks prior to and during the evaluation periods.

  - Proportion of subjects with Hb >=10 g/dL averaged over Weeks 28 to 36, 44 to 52, and 96 to 104 of treatment regardless of rescue therapy
  - Proportion of subjects with Hb within10-12 g/dL averaged over Weeks 28 to 36, 44 to 52, and 96 to 104 of treatment without rescue therapy within 6 weeks prior to and during these 8-week evaluation periods.
  - Proportion of subjects with Hb within 10-12 g/dL averaged over Weeks 28 to 36, 44 to 52, and 96 to 104 of treatment regardless of rescue therapy use.

  - Hemoglobin change from baseline to the average hemoglobin value of weeks 28 to 36, 44 to 52, and 96 to 104 of treatment regardless of the use of rescue therapy.
- **Iron Supplement**:
  - Time to and Proportion of subjects who has requirement for IV iron supplementation.
  - Mean monthly IV iron (mg) per subject during weeks 28-52, 53-104, and 105 -EOT.
  - Mean monthly oral iron (mg) use per subject during 28-52, 53-104, and 105 - EOT.
  - Proportion of subjects who has Requirement for only oral iron supplementation (Week 1 to 52).
- **Hospitalizations**
  - Time to first hospitalization (% of subjects) up to Week 52, 7 days after Last Dose. Time to first hospitalization or skilled nursing facility (% of subjects) up to Week 52, 7 days after Last Dose.
  - Number of days of hospitalizations due to AE per subject-exposure year (PEY)
  - Number of days in hospital or skilled nursing facility due to AE per patient-year exposure (PEY)
  - Number of days in medical-facility (hospital, skilled nursing facility, emergency room, or overnight observation) due to AE per patient-year exposure (PEY)

  - Hospitalization-free survival days on treatment up to Week 52, 7 days after Last Dose. The days will be compared between the 2 treatment groups using ANCOVA model with baseline Hb and the baseline stratification factors.

  - Number of days on treatment out of hospital and skilled nursing facility up to Week 52, 7 days after Last Dose. Similar analysis as above will be performed.

- **Missed Dialysis:**
  - Occurrence (number) of missed dialysis sessions per subject.
  - Proportion of subjects with missed dialysis sessions.
  - Number of days of missed dialysis sessions per patient-exposure year (PEY).
- **Rescue Therapy Use:**
  - Proportion of subjects who receive RBC/blood transfusions.
  - Number of RBC/blood packs per subject-month exposure to study medication.
  - Proportion of subjects who received rescue therapy [composite of RBC transfusions and/or rescue ESA].
  - Time to initial rescue therapy.
- **Changes in Cholesterol Levels:**
  - Change from baseline at each of the selected treatment time points (Weeks 12-28, every 8 weeks onward until Week 52) in:
    - total cholesterol,
    - low-density lipoprotein/high-density lipoprotein ratio,
    - Non-HDL cholesterol.
  - Proportion of subjects achieving LDL target of <100 mg/dL averaged over Weeks 12-28 of treatment for subjects whose baseline LDL ≥ 100.
- **Blood Pressure Effect:**
  - Proportion of subjects achieving blood pressure treatment goal in ESRD subjects (predialysis systolic BP <140 mmHg and diastolic BP<90 mmHg) averaged over Weeks 12-28.
  - Time to first exacerbation of hypertension from baseline during weeks 1 to 36 defined as:
    - An increase from baseline of ≥ 20 mmHg systolic BP (SBP) and SBP >170 mmHg

    Or,
    - An increase from baseline of ≥ 15 mmHg diastolic BP (DBP) and DBP >100 mmHg
  - Time to a treatment-emergent AE of hypertension.
  - Time to exacerbation from baseline of antihypertensive therapy.
  - Proportion of subjects with exacerbation of hypertension, defined as an increase from baseline of ≥ 20 mmHg systolic BP and sBP >170 mmHg or an increase from baseline of ≥ 15 mmHg diastolic BP and dBP>100 mmHg.
  - Proportion of subjects with a treatment-emergent AE of hypertension.
  - Proportion of subjects with exacerbation from baseline of antihypertensive therapy.
- **Vascular Access Thrombosis (HD subjects):**
  - Time to a treatment-emergent AE of vascular access thrombosis
  - Proportion subjects with a treatment-emergent AE of vascular access thrombosis
- **Health Related Quality of Life (HRQoL) and EQ-5D-5L Benefits of Anemia Therapy:**

  Change from baseline averaged over Weeks 12 to 28, and 28 to 52 of treatment for parameters listed below.

  - Vitality Sub-score of SF-36 (Appendix 3):
    - In FAS subjects with baseline Vitality Sub-score below 50.
    - In all FAS subjects.

  - Physical Functioning Sub-score of SF-36:
    - In FAS subjects with baseline Physical Functioning Sub-score below 40.
    - In all FAS subjects.
  - Anemia Subscale ("Additional Concerns") of Functional Assessment of Cancer Therapy-Anemia (FACT-An) Scores (Appendix 4):
    - In FAS subjects with baseline subscale scores below 55 (generally associated with fatigue).
    - In all FAS subjects.
  - Total FACT-An Scores:
    - In FAS subjects with baseline FACT-An scores below 135
    - In all FAS subjects.
  - EQ-5D-5L Scores and other QoL Measures/Other Component scores of SF-36: In all FAS subjects.
- **Hepcidin, Iron Indices, and HbA1c:**
  - Change from baseline in serum hepcidin at each of the selected time points
  - Change from baseline in serum ferritin at each of the selected time points, total and sub-grouped by baseline values of <400 ng/mL, and ≥400 ng/mL.
  - Change from baseline in TSAT at each of the selected time points, total and sub-grouped by baseline values of <40%, and ≥40%.
  - Change in HbA1c level at each of the selected time points, in all subjects, in subjects without history of diabetes and in subjects with history of diabetes.

4.4 Safety Assessments

Study-specific Safety will be assessed by evaluating the following:

- treatment emergent adverse events (TEAEs), treatment emergent serious adverse events (TESAEs)
- Vital signs, electrocardiogram (ECG) findings, and clinical laboratory values

Safety interpretation will also be made based on analyses of adjudicated safety data pooled across multiple studies in the roxadustat Phase 3 program as described in PSAP. The members of an independent adjudication committee blinded to treatment assignment will adjudicate the following events in multiple phase 3 studies:

- Death from any cause, MI, stroke, heart failure requiring hospitalization, unstable angina requiring hospitalization, hypertensive emergency, deep venous thrombosis, pulmonary embolism, and vascular access thrombosis.

5 GENERAL STATISTICAL CONSIDERATIONS

5.1 Sample Size Determination

At least 600 subjects and up to 1200 subjects will be randomized via IXRS to receive roxadustat or epoetin alfa in a 1:1 ratio. Study drug will be administered in an open-label manner. Randomization will be stratified by the following four factors:

- Mean qualifying screening hemoglobin (≤10.5 vs. >10.5 g/dL)
- History of cardiovascular, cerebrovascular or thromboembolic diseases (Yes vs. No)
- Mean prescribed baseline epoetin alfa dose (or, equivalent epoetin dose for non-epoetin subjects) in 4 weeks prior to randomization (≤150 vs. >150 IU/kg/week)

Up to 600 subjects will be enrolled in this study in each treatment group. During the course of this study, which is being conducted in parallel with other Phase 3 studies, up to 1200 subjects may be enrolled for safety evaluation of roxadustat in comparison to epoetin alfa including pre-specified and adjudicated safety events of interest (i.e., all-cause death, MI, stroke, congestive heart failure requiring hospitalization, unstable angina requiring hospitalization, DVT, pulmonary embolism, vascular access thrombosis, and hypertensive emergency). The final number of patients to be enrolled will be based on the enrollment rate of other studies within the same indication, in order to optimize stopping these studies at comparable time frame.

With at least 600 subjects, the study will provide at least 99% power to demonstrate statistical non-inferiority of roxadustat versus ESA in the primary endpoint for US (FDA) submission (i.e., specifically, hemoglobin change from baseline to the average level during the evaluation period defined as Week 28 until Week 52).

The study will provide at least 99% power to demonstrate statistical non-inferiority of roxadustat versus ESA in the primary endpoint outside of the United States (i.e., specifically, hemoglobin change from baseline in the averaged hemoglobin over Weeks 28 to 36). This assumes a difference (roxadustat minus ESA) of -0.30 g/dL, a non-inferiority margin for this difference of 0.75 g/dL and a standard deviation of 1.25 g/dL. Appendix 7 has description on the justification that the non-inferiority margin of 0.75 was used.

5.2 Analysis Populations

5.2.1 Intent-to-Treat (ITT) Population

The ITT population consists of all randomized subjects. If treatment received differs from the randomized treatment, the randomized treatment group will be used.

5.2.2 Full Analysis Set (FAS)

The FAS population will consist of all randomized/enrolled subjects who received at least one dose of study drug and have at least one post-dose hemoglobin assessment. If treatment received differs from the randomized treatment, the randomized treatment group will be used.

5.2.3 Per Protocol Set (PPS)

The PPS population will consist of all subjects in the FAS population who received at least 8 weeks of treatment, have at least one valid post-dose hemoglobin assessment and are without Major protocol violations. If treatment received differs from the randomized treatment, the randomized treatment group will be used.

5.2.4 Safety Population (SAF)

The Safety Population will consist of all randomized/enrolled subjects who received at least one dose of study medication. If treatment received differs from the randomized treatment, the actual treatment group will be used.

5.2.5 Major Protocol Deviations

A subject who meets any of the major protocol deviation criteria in Table 3 will be excluded from the PPS. These will be identified while data are collected and a blinded review will occur before database lock.

Table 3. Criteria for Assessing Major protocol Deviations

Number	Major Protocol Deviation
1	Violation of key* inclusion or exclusion criteria which may affect the assessment of the efficacy of the study drug.
2	Administration of wrong randomization study drug for more than 4 weeks of wrong treatment before week 24; any duration from week 24 to week 52.
3	Study drug compliance < 75% (up to Week 52).
4	Administration of prohibited concomitant medication that may impact evaluation of study drug*.
5	Significant noncompliance with study procedures, that may impact evaluation of efficacy of the study, will be evaluated case by base*.

*Subject to Medical Monitor's decision

The number and percentage of Major protocol deviations will be categorized and summarized by treatment group as deemed appropriate.

5.3 Methodology and Conventions

Continuous variables will be presented by descriptive statistics: n, mean, standard deviation, median, minimum, and maximum. Categorical variables will be tabulated by frequency count and percentage. For efficacy endpoints, the standard error and 95% confidence intervals (CI) will be presented as part of the descriptive summaries.
Safety and efficacy data will be summarized and presented by treatment group and time point in summary tables.

Lab results obtained from the central laboratory will be used for all efficacy and safety analyses. Local laboratory values, if collected in the CRF's, will be listed only in data listing. Unless otherwise stated, all statistical tests will be two-sided hypothesis tests performed at the 5% level of significance for main effects and all confidence intervals will be two-sided 95% confidence intervals.

The secondary endpoints will be tested sequentially using the fixed sequence approach for multiplicity adjustments at an alpha level of 0.05. There will be no adjustments for multiple comparisons for other tests.

All analyses will be performed using SAS® Version 9.1.3 or higher.

5.4 Additional Data Handling Rules and Presentation Specifications

The following general guidelines will apply to all statistical analyses and data presentations:

- Baseline is defined as the last available value obtained prior to the first dose of study drug, unless otherwise specified in this SAP.
- Hemoglobin baseline is defined as the mean of the at least last three available central laboratory hemoglobin values obtained prior to the first dose of study medication (three

of the latest screening hemoglobin values plus the predose hemoglobin value collected on Day 1). In subjects with missing Day 1 hemoglobin value, the mean of three latest screening laboratory hemoglobin values will be considered as baseline hemoglobin value.

- Baselines for reticulocyte count, reticulocyte hemoglobin content (CHr), hepcidin, hs-CRP, serum iron parameters (transferrin, TIBC, ferritin, sTfR, and iron), blood pressures and heart rate are defined as the mean of central lab values obtained within 6 weeks prior to the first dose of study medication.
- Randomization stratification factors derived from actual data (not the ones from the randomization system) will be used in all applicable analysis models
    - The stratification factors to be included in the analyses, where appropriate, include the randomization stratification factors and incident dialysis within ≤ 4 months vs. > 4 months of starting dialysis when randomized.
- When the actual treatment received by a subject is different from the randomized treatment assigned, the subject will be analyzed per the randomized treatment for the efficacy parameters; while they will be analyzed per actual treatment that was taken for the safety parameters for US submission.
- Unscheduled visits within an allowable window will be grouped into the closest scheduled visits based on the visit window specified in Appendix 5. For subjects who have more than one measurement at a certain scheduled visit, the last measurements will be used, with the exception of CPK, WBC, liver function tests (i.e., ALT, AST, GGT, ALP, and total bilirubin), in which the maximum measurement will be used.
- By default, US conventional units will be used for laboratory value presentations. A set of lab summary tables in SI units will alsao be provided based on TLF index.
- Age is calculated as of date that the informed consent form was signed.

age = INTCK('YEAR', Birth date, date of Informed Consent, 'C') where INTCK is a SAS function.

- Duration of treatment or days in treatment is calculated as: last dose date of study medication – first dose date of study medication + 1
- Body weight, height and temperatures will converted using the following formula:
  - kg = lb/2.2
  - cm = 2.54 x in
  - $C^o = (5/9) \times (F^o - 32)$
- The mean, standard deviation and median will be presented with adding one more decimal to raw data with rounding off. The minimum and maximum will be presented with the same number of decimals as in the raw data.
- All percentages will be rounded to one decimal place and lined up by the decimal place. The percentage will be suppressed when the count is zero.
- Any p-values will be rounded to four decimal places and will be presented as '<.0001' if they are less than 0.0001 after rounding.
- All tables and listings will have a header showing "FibroGen, Inc.", the protocol number, and Page x of y. Footer will indicate the program file path/name, run date and run time.
- For continuous variables that are recorded as "< X" or "> X", the value of "X" will be used in the calculation of summary statistics. The original values will be used for the listings.
- Decimal points will be presented as follows: N will be presented without decimal, minimum and maximum in same precision as in the database, mean and median in one

more decimal than minimum and maximum, and SD and SE in one more decimal than mean and median.

- Tables and figures will use derived analysis visit. Listings will use nominal visits, show the flag to indicate analysis visit to be used. Namely, both Nominal visit and analysis visit will be presented in the listing.
- Additional data handling conventions are detailed in Appendix 5.
- Analysis visits assigned for hemoglobin samples collection corresponding to the range of treatment days (window) during which an actual visit may have occurred are presented in Appendix 5. All scheduled and unscheduled hemoglobin samples that belong to each window will be taken into account.

6 SUBJECT ACCOUNTABILITY AND DISPOSITION

The following subject data will be summarized and presented by treatment group (roxadustat and active control) unless specified otherwise:

- Number and percentage of subjects screened (overall only) and randomized (using the screened subjects population)
- Number and percentage of subjects screened (overall only) and randomized (using the screened subject population) in the original protocol and amendments.
- Number and percentage of subjects randomized at each center, and for all centers combined, by treatment group (using the ITT population)
- Number and percentage of subjects in each analysis set, by treatment group (using the ITT population)
- Number and percentage of subjects excluded from the Per Protocol analysis set by reason for exclusion, treatment group (using the ITT population)
- Kaplan-Meier plots will be generated for premature treatment discontinuation by randomization arm showing 2 curves (one curve per treatment group).
- All subjects who prematurely discontinued during the Treatment Period will be listed by discontinuation reason for the ITT population.

7 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographic parameters and important baseline disease characteristics will be summarized by treatment group for the ITT, Safety, FAS and PPS populations. These include but may not be limited to age, age group (18 to 64, 65 to 74, ≥75), sex, ethnicity, race, region, weight, body mass, by stable dialysis and newly initiated dialysis, index (BMI), hemoglobin (the mean of the two most recent central lab results), platelets, mean hemoglobin category (≤10.5 g/dL or > 10.5 g/dL), mean weekly baseline ESA (EPO or EPO equivalent) dose per kilogram category (≤150 IU/kg/week or >150 IU/kg/week) in the 4 weeks prior to randomization, mean ESA (EPO or EPO equivalent) in 4 weeks (if on epoetin or darbepoetin) or in 8 weeks (if on Mircera) prior to randomization category (<5,000 IU/week, 5,000 to 8,000 IU/week, >8,000 to 16,000 IU/week, >16,000 IU/week), ferritin, ferritin group (<100, 100-<400, ≥400 ng/mL), TSAT and TSAT group (<20, 20-<40, ≥40%), iron deplete (ferritin <100 or TSAT <20%) vs. not, baseline C-reactive protein (CRP) group (CRP ≤ULN vs. CRP > ULN), and history of cardiovascular disease or cerebrovascular disease or thromboembolic disease (Yes or No), primary reason for CKD/ESRD (DM and HPT vs. all others).

Descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) will be presented for continuous variables. In addition, 25%-75% values of hemoglobin and platelets will be presented. Frequency distributions (number and percentage of subjects) will be presented for categorical variables.

Descriptive statistics of baseline values for other parameters will be presented in their change from baseline tables.

8 MEDICAL HISTORY

Medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA). Medical History of interest including chronic kidney disease (CKD) History inclusive of CKD cause, cardiovascular disease including dyslipidemia, cerebrovascular disease, thromboembolic history, hypertension history, diabetes history, and anemia history will be summarized by system organ class, preferred term and treatment group for the Safety Population.

9 STUDY MEDICATION

9.1 Extent of Exposure

Exposure to study medication will be summarized by treatment group in terms of treatment duration in weeks, which is calculated using the following formula: (the last dose date of study medication taken - the first dose date of study medication taken +1)/7.

Total weekly study drug exposure is defined as the total prescribed dose (in mg and mg/kg for roxadustat and IU and IU/kg for ESAEPO) of study drug administered within the week (windowed by 7-day period from Day 1).

Duration of exposure, weekly exposure and total study drug exposure will be tabulated by treatment group for the safety population.

Per dose amount and dosing frequency will also be tabulated by treatment group for the safety population and PPS population.

Patient-Exposure-Year (PEY) is defined as Last Dose Date – First Dose Date + 1.

9.1.1 Dosing Changes

Dosing changes for both treatment groups are collected in the Study Drug Administration/ Dose Adjustment Form in the eCRF. Two types of dosing changes will be calculated.

A dose-per-intake change is the change in the number of milligrams on the intake day (for example from 200 mg TIW to 250 mg TIW). A weekly-dose change is the change in the prescribed weekly dose, calculated as the dose-per-intake times the weekly frequency.

For example a change from 200 TIW to 250 TIW is a change of 600 mg to 750 mg per week which is considered as a change in dose-per-intake and a change in weekly-dose as well.

For each subject the total number of dose-per-intake changes and the weekly-dose changes will be calculated for safety and PPS population.

9.1.2 Duration of Exposure

Exposure time will be categorized according to the following categories by treatment arm (roxadustat and ESA) for safety and PPS population:

- Less than 2 weeks
- At least 2 weeks, less than 4 weeks
- At least 4 weeks, less than 26 weeks
- At least 26 weeks, less than 52 weeks
- At least 52 weeks, less than 78 weeks
- At least 78 weeks, less than 104 weeks
- At least 104 weeks, less than 130 weeks
- At least 130 weeks, less 156 weeks
- More than 156 weeks
- Unknown

9.2 Treatment Compliance

Study medication dosing compliance for a specified period is defined as the total dose (mg or IU) actually taken by a subject during that period divided by the prescribed dose expected to be taken during the same period multiplied by 100.

An overall per-dose compliance measure can be calculated by (# of actual dose administrations)/ (Total # of expected dose administrations) *100 during the subject's Treatment Period. Descriptive statistics for study medication compliance will be presented by treatment group for the entire Treatment Period of the study for safety population.

Compliance will be summarized as follows:

- Descriptive statistics will be summarized by the 2 treatment groups.
- Percent compliance will be categorized according to the following categories by treatment groups:
    - less than 50% (significant drug noncompliance)
    - at least 50%, less than 75% (moderate drug noncompliance)
    - greater or equal 75% (drug compliance)
    - unknown

10 PRIOR AND CONCOMITANT MEDICATIONS

The World Health Organization Drug Dictionary (WHODD) will be used to classify prior and concomitant medications by therapeutic class and generic name based on ATC code level 3. Prior medication is defined as any medication taken and stopped prior to the first dose of the study medication within 30 days before signing the informed consent. Concomitant medication is defined as any medication taken between the day of first dose of the study medication and the day of ET/EOT Visit date + 28 days, inclusive.

Medication start and end dates will be compared with the start date of study drug and classified as per Table 4.

In case of partial or missing dates, comparisons will be made based on the level of detail available. For example, if start date of study drug is 04Jan2013, and a medication has a start date of 04Jan2013, the medication will be classified as concomitant. A start date of Jan2013 (i.e., missing day) would also see the medication classified as concomitant.

Table 4. Classification of Prior and Concomitant Medications

End date / Start date	Before start of study drug administration	On or after start of study drug administration	Missing
Before start of study drug administration	Prior	Concomitant	Concomitant
On or after start of study drug administration	–	Treatment Emergent Concomitant	Treatment Emergent Concomitant
Missing	Prior	Concomitant	Concomitant

Both prior and concomitant medication usage will be summarized by the number and proportion of subjects in each treatment group. Subjects will only be counted one time in each unique ATC Class and generic name if multiple drugs are used by a subject.

Detailed analyses may be performed on prior and concomitant medications of special interests such as oral iron, blood pressure medications and lipids medications.

11 EFFICACY ANALYSES

Efficacy will be conducted on the ITT and FAS for US (FDA) submission. Efficacy analysis of non-inferiority on the PPS population and superiority on FAS population will be used for EU regulatory submission.

The **Efficacy Emergent Period** is defined as the evaluation period from the Analysis date of first dose intake up to 7 days after the Last Dose of study drug or EOT Visit, whichever occurs first. This period will be used as reference period for the time to event analyses related to efficacy endpoints, unless specified otherwise.

11.1 Analyses of Primary Endpoint

11.1.1 Primary Endpoint and Hypothesis

11.1.1.1 U.S. Submission

The primary efficacy endpoint for the US (FDA) submission is defined as the hemoglobin change from baseline to the average level during the evaluation period defined as Week 28 until Week 52.

The primary hypothesis to be tested for the primary efficacy analysis for US (FDA) submission is:

> H_0: Hemoglobin mean change from baseline to the average level from Week 28 to Week 52 in the roxadustat group ≤ hemoglobin mean change from baseline in the epoetin alfa group minus 0.75 g/dL
> Versus:
> H_1: Hemoglobin change from baseline to the average level of Week 28 to Week 52 in the roxadustat group > hemoglobin mean change from baseline in the epoetin alfa group minus 0.75 g/dL

A Multiple Imputation (MI) analysis of covariance (ANCOVA) model will be used. The model will contain terms for treatment group, baseline Hb measurement, and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL). The primary efficacy analysis will be based on the estimated difference between the two treatments overall mean effects throughout the evaluation period based on the pooled ANCOVA model.

This null hypothesis will be rejected if the two-sided 95% CI for the difference between the two treatment groups using MI ANCOVA model lies entirely above -0.75 g/dL. Hemoglobin values under the influence of a rescue therapy will not be censored in this analysis.

The below stratification factors will be used in efficacy analyses.

- Mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL) for other than Hb related endpoints
- History of CV, cerebrovascular, or thromboembolic diseases (yes vs. no)
- Mean prescribed weekly epoetin alfa dose (or, equivalent epoetin dose for nonepoetin subjects) in the 4 weeks prior to randomization (≤ 150 vs. > 150 IU/kg/week)
- Incident dialysis within ≤ 4 months vs. > 4 months of starting dialysis when randomized.

The following steps will be used to conduct the primary analysis of the primary endpoints for both regions:

1. Generate 200 datasets, using seed 162345 for U.S. and 347631 for Ex-U.S., where only intermittent missing hemoglobin data will be imputed for each treatment relying on non-missing data from all subjects within each treatment group using the Monte Carlo Markov Chain, MCMC) imputation model and the available non missing hemoglobin for each scheduled Week by treatment group.

   The MCMC statement in the SAS PROC MI procedure with monotone option will be used. As a result, each dataset will only have missing ending data, namely, a monotone missing data pattern.

2. For each dataset from step 1, missing ending data (hemoglobin up through end of evaluation period) will be imputed. As a result, 200 imputed complete datasets will be generated.
   - Missing data at Week 1 will be imputed using the regression imputation model with baseline stratification factor, baseline and hemoglobin from Week 1, using the SAS PROC MI procedure with the REGRESSION option in the MONOTONE statement.
   - The SAS PROC MI procedure will use data separately from each treatment subjects to impute the missing data for a specific Week (i.e. only those that need the imputation for the Week). Since subjects from the different treatment groups for that Week are excluded from the step, they will not contribute to the imputation for the Week.
   - Repeat for all other scheduled Weeks sequentially (Week 2 to the end of evaluation period). Subjects whose missing data were imputed for previous Weeks will contribute to the imputation for the current Week.
   - The regression imputation model includes an intercept and the slopes of the hemoglobin from previous Weeks.

3. Analyze each imputed complete dataset using the ANCOVA model using the mean of observed or otherwise imputed Hb values within the evaluation period (week 28 -52). The model will contain terms for baseline Hb measurement as a covariate and treatment group and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL) as fixed effects.

   Sample SAS code:

```
PROC MIXED data=xx;
  class treatment categorical covariates;
  model change_Week36 = treatment covariates / solution;
  lsmeans treatment / diff cl;
  ods output Diffs=lsdiffs LSMeans=lsm solutionF=Parms;
  by _Imputation_;
run;
```

4. Combine estimates from the results for each of the 200 ANCOVA model using SAS PROC MIANALYZE.

   Sample SAS code:

```
PROC MIANALYZE parms(classvar=full)=lsdiffs;
  class treatment categorical covariates;
  modeleffects treatment;
  ods output ParameterEstimates=MIAN_lsdiffs;
run;
```

Report the results of the least-squares mean estimates of the change from baseline in hemoglobin during the evaluation period, the estimates of treatment effect (e.g. least-squares mean CFB in hemoglobin for the treatment group minus the least-squares mean CFB in hemoglobin for the active comparator group) and the corresponding p-values during the evaluation period.

11.1.1.2 EU Regulatory Submission

The primary efficacy endpoint for Ex-U.S. submission is defined as the hemoglobin change from baseline to the average hemoglobin of Weeks 28 to 36, without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period and will be based on the PPS population based on patients enrolled during the original protocol before the amendment 1. This primary efficacy analysis will be based on the estimated difference between the two treatments overall mean effects throughout the evaluation period based on the MMRM model described below.

The primary hypothesis to be tested for the primary efficacy analysis for Ex-US submission is:

H_0: Hemoglobin mean change from baseline to the average of week 28 to week 36 in the Roxadustat group

≤ Hemoglobin mean change from baseline in the Epoetin alfa group minus 0.75 g/dL

Versus

H_1: Hemoglobin change from baseline to the average of week 28 to week 36 in the roxadustat group

> Hemoglobin mean change from baseline in the Epoetin alfa group minus 0.75 g/dL

Rescue therapy for all subjects is defined as ESA rescue or RBC transfusion. For subjects who require rescue therapy, the hemoglobin value after the initiation of the rescue therapy will be set to missing up to 6 weeks (or 8 weeks if during the evaluation period) from the last date of the rescue therapy.

The primary endpoint for EU Regulatory. will be analyzed using MMRM using baseline value as covariate and treatment group, visit (up to Week 52), interaction of treatment group

and visit, and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL) as fixed effects.

Due to the large amount of visits (up to Week 52) to be included in the model, the unstructured covariance pattern model will be applied first. If the algorithm for unstructured covariance pattern does not converge then the heterogeneous Toeplitz structure will be used. If this second model does not converge either then the (homogeneous) Toeplitz structure will be tried and finally compound symmetry as a covariance structure to achieve convergence.

The estimates for difference of change in Hb averaged over Weeks 28 to 36 between the two treatment groups will be generated from an estimate statement from Visit Week 28 to 36.

This null hypothesis will be rejected if the two-sided 95% confidence interval for the difference of least square means between the two treatment groups using the MMRM model lies entirely above –0.75 g/dL.

11.1.2 Sensitivity Analyses of Primary Endpoint

The potential impact of missing data on the estimates of the primary efficacy endpoint will be examined using the following sensitivity analyses. The results of the analyses will be summarized in Table 5. These sensitivity analyses are further detailed in each Section.

Table 5: Primary Endpoint Analysis Results

Analysis (ITT for U.S. and PPS Ex-U.S.)	Treatment Difference and 95% CI (roxadustat vs Epo)	Std. Err.	Degree of Freedom*	t-statistics	p-value
ANCOVA with Multiple Imputations	.xxx (.xxx, .xxx)	.xxx	xxx	.xxx	.xxxx
ANCOVA-MI with Hb censored** for rescue therapy (US endpoint only)	.xxx (.xxx, .xxx)	.xxx	xxx	.xxx	.xxxx
MMRM	.xxx (.xxx, .xxx)	.xxx	xxx	.xxx	.xxxx
PMM-Last Mean Carried Forward	.xxx (.xxx, .xxx)	.xxx	xxx	.xxx	.xxxx
ANCOVA-MI for subjects enrolled in original protocol	.xxx (.xxx, .xxx)	.xxx	xxx	.xxx	.xxxx

* Maximum of Degree of Freedom from each individual ANCOVA. The actual Degrees of Freedom in PMM will be calculated using $df = (m-1)\left(1+\frac{m\,\overline{U}}{(m+1)B}\right)^2$. where B is between-imputation variance and $\overline{U} = \frac{1}{m}\sum_{j=1}^{m} U_j$. with U the standard error associated with estimates, Rubin (1987).

**Rescue therapy for all subjects is defined as ESA rescue or RBC transfusion. For subjects who require rescue therapy, the hemoglobin value after the initiation of the rescue therapy will be set to missing up to 6 weeks (or 8 weeks if during the evaluation period) from the last date of the rescue therapy.

11.1.2.1 MMRM

The same analysis as Ex-U.S. primary endpoint will be repeated for the primary U.S. endpoint as one of the sensitivity analysis.

11.1.2.2 Pattern Mixture Model

PMMs provide a general and flexible framework for sensitivity analyses that allows formulating assumptions regarding missing data in a transparent and clinically interpretable manner. This is expected to address the possibility of the data being missing not at random (MNAR). All factors mentioned in the primary analysis will be included in the PMM.

The following aspects of data missingness, may affect the estimates.

- Timing and extent of missingness
- Assumed underlying mechanism for data missingness

11.1.2.2.1 Timing and Extent of Missing Data

To assess the potential effect of data missingness on the estimate of treatment effect, subjects will be classified as full data or missing data cases. Patterns of missingness will be based on non-missing hemoglobin before the end of the evaluation period.

- Full data cases are defined as subjects with non-missing hemoglobin for all scheduled weeks of the Treatment period.

- Missing data cases are defined as subjects with a missing hemoglobin on at least one scheduled Week of the treatment period. The missing data cases are further grouped into intermittent missing and monotone missing cases.

    - Intermittent missing hemoglobin cases are defined as subjects with a missing hemoglobin for at least one scheduled week of but not on consecutive scheduled weeks up to end of the evaluation period.

    - Monotone missing hemoglobin cases are defined as subjects who have consecutive scheduled Weeks with missing hemoglobin up to the end of evaluation period. A subject who is a Monotone missing case could have intermittent missing hemoglobin prior to the ending Week.

Subjects will be grouped as follows:

- Full data cases
- Intermittent missing data cases
- Monotone missing data cases

Should the incidence of Monotone missing data cases and intermittent missing data cases be relatively small, 5%, then those cases will be combined so that the groups are full data

cases and missing data cases. The summary of missing pattern in first 52 scheduled visits will be presented in a table.

11.1.2.2.2 Assumptions on Missing Data Mechanism

In addition to the extent of data missingness, the mechanism under which missing data occur may affect the estimate of the parameter of interest.

The potential impact of missing efficacy endpoints on the estimates of treatment effects will be assessed using alternative statistical models with different underlying assumptions on the missing data mechanism (missing not at random (MNAR)) (Little and Rubin, 1987).

11.1.2.2.3 PMM –Last Mean Carried Forward

A pattern-mixture model using a last mean carried forward multiple imputation method (Carpenter et al, 2013) will be used as another sensitivity analysis to explore the robustness of the ANCOVA results for the primary efficacy variables. Using this method, missing data after ending Week will be imputed based on the last non-missing mean from its own treatment group.

The steps to implement this sensitivity analysis are as follows. Parameters below refer to the parameters of the multivariate normal distribution for baseline and post baseline Hb measurement.

1. Create posterior distribution of parameters: Separately for each treatment arm, take all patients observed data and assuming MAR to fit a multivariate normal distribution with unstructured mean (i.e. a separate mean for each of the baseline plus post-baseline scheduled weeks and unstructured variance covariance matrix using a Bayesian approach with an improper prior for the mean and an uninformative Jeffreys' prior for the variance-covariance matrix (Schafer, 1997, p. 155).

2. Draw parameters: Separately for each treatment arm, draw variance-covariance matrix from the posterior distribution for the parameters using seed 453628 for U.S. and 386045 for Ex-U.S. The mean Vector would be set to the marginal mean for their randomized treatment arm at their last non-missing measurement.

3. Build joint distribution of missing data and observed data: For each subject with missing data, using the draws for the parameter to build the joint distribution of their observed and missing data.

4. Construct conditional distribution of missing data give observed data: For each patient with missing data, use their joint distribution in previous step to construct their conditional distribution of missing given observed outcome data. Sample their missing data from this conditional distribution, to create a `completed' data set, using seed 732545 for U.S. and 538529 for Ex U.S.

Repeat the above steps for 200 times and resulting in 200 imputed data sets. Then fit ANCOVA model for each imputation data set, and combine the resulting parameter estimates and standard errors using Rubin's rules (Rubin, 1987) for final inference. This sensitivity analysis will be performed for both U.S. and Ex-U.S. primary endpoints.

11.1.3 Subgroup Analyses

The analysis of the primary endpoints of US and Ex-US may be performed by gender, age group, baseline hemoglobin categories, mean prescribed weekly epoetin alfa dose categories, baseline CRP categories, cardiovascular/cerebrovascular/thromboembolic medical history, baseline iron replete status, and stable vs incident dialysis: dialysis duration >4 months vs ≤ 4 months where appropriate.

In addition, hemoglobin change from baseline to the average level during Weeks 18-24 may be analyzed by subgroups of baseline hs-CRP (<=ULN, vs. > ULN), baseline ESA dose categories, and dialysis status.

11.2 Analysis of Secondary Endpoints

11.2.1 Secondary Endpoints and Hypothesis Testing

Once the null hypothesis of the primary endpoint has been rejected, the secondary endpoints below will be tested using a fixed sequence testing procedure, in order to maintain the overall two-sided type I error of 0.05. If the claim of superiority or non-inferiority is successful and the test will progress to the next comparison in sequence as follows.

Table 6 Key Secondary Endpoints Fixed Sequence Testing Procedure

Test	Variable	Comparison
1	Proportion of subjects with mean hemoglobin level during the evaluation period defined as Week 28 until Week 52 ≥ 10.0 g/dL (U.S.) or Proportion of Hb responders in the average of weeks 28 to 36 within the target range of 10.0 to 12.0 g/dL without having received rescue therapy. (Ex-U.S.).	Non-inferiority of roxadustat versus EPO with margin of -0.15
2	LDL cholesterol change from BL to the average of weeks 12 to 28.	Superiority of roxadustat versus EPO
3	Hemoglobin change from baseline to the average level during Weeks 18 to 24 for subjects with baseline hsCRP > ULN	Non-inferiority of roxadustat versus EPO
4	Average monthly IV iron (mg) use per subject during weeks 28 to 52.	Superiority of roxadustat versus EPO
5	Time to first RBC Transfusion during the treatment.	Non-inferiority of roxadustat versus Epo with margin of 1.8
6	MAP change from BL to the average MAP of weeks 20 to 28.	Superiority of roxadustat versus EPO
7	Time to first exacerbation of hypertension: An increase from baseline of ≥ 20 mmHg systolic blood pressure (SBP) and SBP ≥ 170 mmHg or an increase from baseline of ≥15 mmHg diastolic blood pressure (DBP) and DBP ≥ 100 mmHg during Weeks 28 to 52.	Superiority and Non-inferiority of roxadustat versus EPO with margin of 1.8

11.2.2 Primary Analysis of Secondary Endpoints

Unless otherwise specified, the primary analysis for the secondary endpoints will be analyzed using MMRM described in 11.1.2.1 based on the FAS for U.S., and PPS for non-inferiority tests and FAS for superiority tests for Ex-U.S. if applicable.

11.2.2.1 Proportion of Hemoglobin Responders

- US: Proportion of subjects with mean hemoglobin level during the evaluation period defined as Week 28 until Week 52 ≥ 10.0 g/dL, and patients with mean hemoglobin level between Hb 10-12 g/dL.
- Ex-US: Proportion of subjects with mean Hb between range of 10.0 to 12.0 g/dL without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period.

The hypothesis to be tested for the efficacy analysis is:

H0: Hemoglobin response rate for subjects in the roxadustat group - hemoglobin responder rate for subjects in the Epoetin group ≤ -15%
Versus
H1: Hemoglobin response rate for subjects in the roxadustat group - hemoglobin responder rate for subjects in the Epoetin group > -15%.

A 2-sided 95% confidence interval for the difference of 2 responder rates (roxadustat – ESA) based on the Miettinen & Nurminen approach adjusting for treatment group and stratification factors will be calculated and non-inferiority will be declared if the lower bound of the 95% CI is greater than -15%.

Subjects who dropped out from the study before week 28 will be treated as non-responder.

11.2.2.2 Change from Baseline in LDL Cholesterol Averaged over Weeks 12 to 28

The change from baseline in LDL cholesterol will be analyzed and compared between the 2 treatment groups using MMRM model with baseline LDL cholesterol as a covariate and treatment group, visit, the interaction of treatment and visit, and the above mentioned stratification factors as fixed effects. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure. Data up to visit of Week 52 will be included in the model. The estimates for difference of LDL Cholesterol averaged over Weeks 12 to 28 between the two treatments groups will be generated from an estimate statement from Visit Week 12 to 28. Superiority will be declared if the upper bound of the 2-sided 95% confidence interval of the difference between roxadustat and active control does not exceed 0.

Baseline, LDL, is defined as the last available value obtained prior to the first dose of study drug.

11.2.2.3 Hemoglobin change from baseline to the average level during the Week 18 to 24 for Subjects with CRP > ULN

Change from baseline in Hemoglobin from baseline to the average level during the Week 18 to 24 will be analyzed using the ANCOVA MI as the primary endpoints. Both Non-inferiority of roxadustat vs. Epoetin and superiority will be tested. The non-inferiority margin is fixed as a difference of -0.75.

11.2.2.4 Average Monthly of IV Iron Use during Weeks 28 to 52.

The average monthly IV iron use during Weeks 28 to 52 Treatment Period will be calculated for monthly intervals. The Treatment Period will be divided in periods of 28 days and for each of these periods the monthly mean of IV iron will be calculated using the following formula:

Monthly iron use for each subject = Total IV iron in mg / [(last visit date – first dose date of study medication +1)/ 28]

The average monthly iron use will be compared between the 2 treatment groups using An ANCOVA model with baseline iron repletion status, treatment group, andstratification factors as fixed effects. Superiority will be declared if the lower bound of the 2-sided 95% confidence interval of the difference between active control and roxadustat exceeds 0.

11.2.2.5 Time to first RBC Transfusion

Time to the first RBC transfusion during treatment will be analyzed and compared between the 2 treatment groups using the Cox Proportional Hazards model adjusting for baseline Hb, treatment group and stratification factors.

Non-inferiority of roxadustat vs. Epoetin will be declared if the upper bound of the 2-sided 95% confidence interval of the hazard ratio is less than 1.8 and superiority.

11.2.2.6 Mean change in pre-dialysis mean arterial pressure (MAP) averaged over Weeks 20-28

Mean Arterial Pressure (MAP) will be calculated for each subject using the following formula:

MAP = (2/3) * DBP + (1/3) SBP.

Mean change from baseline in MAP will be analyzed and compared between the 2 treatment groups using an MMRM model with baseline MAP as a covariate, treatment group, visit, the interaction of treatment and visit, and stratification factors as fixed effects. The same strategy as that used in MMRM for Hb will be used to choose variance covariance Structure. Data up to visit of Week 52 will be included in the analyses. The estimates for difference of mean change from baseline in MAP for Week 20 to 28 between two treatment groups will be generated from an estimate statement. Superiority will be declared if the upper bound of the 2-sided 95% confidence interval of the difference between roxadustat and Epoetin (roxadustat – Epoetin) is below 0.

11.2.2.7 Blood pressure increase

- Time to first exacerbation of hypertension During Weeks 28 to 52

Time to first exacerbation during weeks 28 to 52 will be analyzed and compared between the 2 treatment groups using the Cox Proportional Hazards model adjusting for treatment group and stratification factors.

Superiority of roxadustat vs. Epoetin will be declared if the upper bound of the 2-sided 95% confidence interval of the hazard ratio is less than 1.

Subjects will be censored at the time of the last available blood pressure if an exacerbation in blood pressure does not occur.

Sensitivity analysis will be performed on proportion of subjects with increase in blood pressure at any time during the study, defined as an increase from baseline of ≥ 20 mm Hg systolic BP and sBP >170 mmHg or an increase from baseline of ≥ 15 mm Hg diastolic BP and dBP>100 mmHg. Increases from baseline in blood pressure are considered as confirmed by taking the mean of triplicates. The proportion of subjects with increase in blood pressure will be compared between the 2 treatment groups using logistic regression model adjusting for baseline predialysis systolic and diastolic BP as covariates and treatment group and stratification factors as fixed effects.

11.2.3 Secondary Analyses

The proportion of subjects who achieve a hemoglobin response may also be analyzed using logistic regression test and including baseline hemoglobin and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL). The odds ratio (ROXADUSTAT vs. EPO) and its 95% confidence interval will be provided. . The non-inferiority margin is fixed as a difference of -0.15.

11.3 Additional Efficacy Analyses

11.3.1 Hemoglobin Maintenance

11.3.1.1 Hemoglobin long-term Maintenance

Hemoglobin long-term maintenance will be assessed by:

- Mean change in hemoglobin averaged over 8 weeks of treatment at Weeks 44-52 without rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
- Mean change in hemoglobin averaged over the Week 96- 104 of treatment without rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
- Change in hemoglobin at each of the selected post-dosing time points
- Proportion of subjects with hemoglobin >= 10 g/dL averaged over 36 to 44, 44 to 52, and the Week 96- 104 without rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
- Proportion of subjects with hemoglobin within 10-12 g/dL averaged over 36 to 44, 44 to 52, and the Week 96- 104 without rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
- Hb response during Weeks 28 and 36 regardless of use of rescue therapy
- Hb change from BL to the average Hb value of Weeks 28 to 36, 44 to 52, and 96 to 104 of treatment regardless of the use of rescue therapy

The change in hemoglobin, will be analyzed using MMRM model with baseline hemoglobin as a covariate, treatment group, visit, the interaction of treatment and visit, and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL) as factors.. The same strategy as that used in MMRM in Section 11.1.2.1 will be used to choose variance covariance. Data up to visit of Week 52 will be included in the analyses. The estimates for

the difference of change from baseline in Hemoglobin for Week 28 to 36 and 44 to 52 between the two treatment groups will be generated from an estimate statement. The mean change in Hb averaged over 8 weeks of treatment at Weeks 96-104 will be analyzed using the MI analyses as the one for the primary efficacy endpoint of US-submission.

The proportion of subjects with hemoglobin within 10-11 g/dL in US and 10-12 g/dL in Ex-US will be analyzed using logistic regression model including baseline hemoglobin as a covariate and treatment group and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL) as fixed effects.

Hemoglobin change from baseline to the average hemoglobin value of the Week 96 to 104 weeks of treatment will be analyzed using an MI ANCOVA model with baseline hemoglobin as a covariate, and treatment group and stratification factors except mean qualifying screening hemoglobin (≤ 10.5 vs. >10.5 g/dL) as fixed effects, using the Hb mean from the non-missing weeks.

In this study, there are patients in whom roxadustat dose frequency needed to be reduced to BIW or QW as part of their dose reduction below the lowest possible TIW dose of 20 mg TIW during treatment, to evaluate Hb maintenance, the following summary will be provided for these subgroups of subjects treated on BIW or QW (including any frequency <QW) for longer than 8 weeks (i.e., >=56 days):

- Average Hb values over time for every 4-8 weeks after the initiation of BIW or QW
- Average weekly dose over time for every 4-8 weeks after the initiation of BIW or QW

11.3.2 Hospitalizations

Hospitalizations are collected in the Hospitalization Records Form of the eCRF. For each subject, an entry will be recorded for each hospitalization. The days of hospitalization will be calculated as the sum of all hospitalizations durations in days (Date of discharge – Date of Admission + 1). In case of missing dates, the hospitalization duration will be assumed to be 5 days.
Only hospitalizations with admission dates that occur during the Treatment Period and up to 7 days after the last study medication date will be taken into account.

The following hospitalization-related data will be analyzed and compared between the 2 treatment groups:

- Time to first hospitalization (% of subjects) up to Week 52, 7 days after last dose.
- Time to first hospitalization or skilled nursing facility (% of subjects) up to Week 52, 7 days after last dose.
- Number of days of hospitalizations
- Number of days of medical-facility
- Number of days of hospitalizations per patient-exposure year (PEY).
- Number of days in hospital or skilled nursing facility per patient-year exposure (PEY).
- Number of days of medical-facility (hospital, skilled nursing facility, emergency room, or overnight observation) per subject-exposure year (PEY).
- Number of days on treatment out of hospital and skilled nursing facility up to Week 52, 7 days after Last Dose.

Note: Sensitivity analysis may be performed by excluding elective procedures using only hospitalization due to AE

The time to hospitalization event will be compared between the two treatment groups using cox model including baseline stratification factors. The proportion of subjects hospitalized will be compared between the 2 treatment groups using logistic regression model and stratified by the baseline stratification factors. The mean number of days of hospitalization will be analyzed and compared between the 2 treatment groups using the Mantel-Haenszel mean score test adjusting for stratification factors.

11.3.3 Missed Dialyses

Only missed dialyses (defined as a dialysis session that was planned but not received at the patient's usual outpatient dialysis center) with starting dates that occur during the Treatment Period and up to 7 days after the last study medication date will be taken into account. The days of missed dialysis will be calculated as the sum of all missed dialyses in days (# of missed dialysis/weekly prescribed dialysis frequency* 7), where weekly prescribed dialysis is TIW in this study.

The following missed dialysis-related data will be analyzed and compared between the 2 treatment groups:

- Occurrence (number) of missed dialyses. The number of missed dialysis for each subject will be collected. The mean number of missed dialysis will be analyzed and compared between the 2 treatment groups using the Mantel-Haenszel mean score test adjusting for stratification factors.
- Proportion of subjects who has missed dialyses. The proportion of subjects with missed dialysis will be compared between the 2 treatment groups using CMH model adjusting for stratification factors.
- Number of days of missed dialyses per subject-exposure year (PEY). The number of days of Missed Dialysis for each subject will be calculated. The mean number of days of Missed Dialysis will be analyzed and compared between the 2 treatment groups using the Mantel-Haenszel mean score test adjusting for stratification factors.

11.3.3.1 Red Blood Cell Transfusion

For a subject receiving RBC or whole blood transfusion, the Time at Risk (time up to first use) will be calculated (in years) as:

(First use of transfusion date – First dose date of study medication + 1) / 365.25

For a subject not receiving transfusion, the Time at Risk (time until they get censored) is calculated as:

(Date of last study medication – First dose date of study medication + 1) / 365.25

The red blood cell transfusion form of the eCRF in the cumulative visit will be used to derive the number of RBC packs. The number of RBC units is collected in this form. For transfusions where the number of units is not given but the volume transfused is given, the number of units will be estimated by dividing the volume transfused by 250 mL (for transfusion of packed cells) or by dividing the volume transfused by 500 mL (for transfusion of whole blood).

The total number of RBC units/packs during the Treatment Period is calculated for each subject by the sum of the transfused units between the Analysis Date of First Dose and up to the Analysis Date of Last Dose. The following 2 endpoints will be analyzed:

- Proportion of subjects who receive RBC transfusions. The proportion of subjects who received RBC transfusion will be compared between the 2 treatments using logistic regression model adjusting for baseline hemoglobin as covariate and treatment group and stratification factors as fixed effects.
- Number of RBC packs per subject-month exposure to study medication. The mean number of RBC packs will be compared between the 2 treatment groups using ANCOVA model adjusting for baseline hemoglobin as covariate, treatment group, and stratification factors as fixed effects.

11.3.3.2 ESA Usage in roxadustat patients

For roxadustat subjects, ESA use will be recorded in the CRFs. The total number of ESA-week dose per subject will be calculated and listed.
When ESA dose were needed in the EPO arm, such dose will be captured as dose change of study drug. If another ESA were used in addition, it too will be summarized separately from roxadustat patients.
For each entry that meets the criteria below the ESA-week will be calculated as follows:

- If drug is epoetin alfa, epoetin beta, or an epoetin biosimilar (ATC code: B03XA01), then ESA-Weeks = (stop date – start date + 1) / 7;
- If drug is darbepoetin SQ or IV dose (ATC code: B03XA02), then ESA-Weeks = 2 x (stop date – start date + 1) / 7;
- If drug is Mircera IV or SQ dose (ATC code: B03XA03), then ESA-Weeks = 4 x (stop date – start date + 1) / 7.

The total number of ESA-week during the Treatment Period is calculated for each subject by the sum of the ESA-week between the analysis date of first dose and the analysis date of last dose for roxadustat only.

11.3.3.3 Use of Iron

IV Iron

Only IV iron therapy that started during the study treatment and up to the ET/EOT Visit will be taken into account for the total amount.

The average monthly IV iron use during weeks 53 to EOS will be calculated for monthly intervals. The Treatment Period will be divided in periods of 28 days and for each of these periods the monthly mean of IV iron will be used using the following formula:
Monthly iron use for each subject = Total IV iron in mg / [(last visit date – first dose date of study medication in the period+1)/ 28]

The following 2 endpoints will be analyzed:

- Proportion of subjects who receive IV iron. The proportion of subjects who received IV iron will be compared between the 2 treatments using logistic regression model adjusting for baseline iron repletion status and treatment group and stratification factors as fixed effects. (This will be performed for week 1 to week 36 as well).
- The average monthly iron use will be compared between the 2 treatment groups using ANCOVA model with baseline iron repletion status, and treatment group, and stratification factors as fixed effects.

Oral iron

The average monthly oral iron use during Weeks 28 to 52, and weeks 53 to ET/EOT will be calculated similarly to IV iron for monthly intervals.
The following 2 endpoints will be analyzed:

- Proportion of subjects who receive Oral iron. The proportion of subjects who received oral iron will be compared between the 2 treatment groups using logistic regression model adjusting for baseline iron repletion status, treatment group and stratification factors as fixed effects.
- The average monthly oral iron use will be compared between the 2 treatment groups using ANCOVA model with baseline iron repletion status, treatment group, and stratification factors as fixed effects.

In addition the proportion of subjects who received both IV and/or oral iron will be compared between the 2 treatments using logistic regression model adjusting for baseline iron repletion status and treatment group and stratification factors as fixed effects.

11.3.4 Changes in Cholesterol Levels

- Change at each of the selected scheduled testing timepoints (Weeks 12-28, every 8 weeks onward until Week 52) from baseline in:
    - Total cholesterol,
    - Low-density lipoprotein/high-density lipoprotein ratio,
    - Non-HDL cholesterol.

  The mean change in these 3 parameters at each post-dosing time point from baseline will be analyzed and compared between the 2 treatment groups using MMRM model adjusting baseline value as covariate, treatment group, visit, the interaction of treatment and visit, and stratification factors as fixed effects. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure.
- Proportion of subjects achieving LDL target of <100 mg/dL averaged over Weeks 12-28 of treatment for Subjects with Baseline LDL >= 100 mg/dL. The proportion of subjects achieving LDL target will be compared between the 2 treatment groups using logistic regression model adjusting for baseline LDL as covariate and treatment group and stratification factors as fixed effects.

11.3.5 Blood Pressure Effect

- Proportion of subjects achieving blood pressure treatment goal in ESRD subjects (pre dialysis systolic BP <140 mmHg systolic and diastolic BP<90 mmHg) averaged over Weeks 12-28. The proportion of subjects achieving blood pressure goal will be compared between the 2 treatment groups using logistic regression model adjusting for baseline pre dialysis systolic and diastolic BP as covariates and treatment group and stratification factors as fixed effects.
- Time to a treatment-emergent AE (TEAE) of hypertension. Time to a TEAE of hypertension will be analyzed using the Cox Proportional Hazards model adjusting for treatment group and stratification factors as fixed effects.
- Time to exacerbation from baseline of antihypertensive therapy. Time to intensification from baseline of antihypertensive therapy will be analyzed using the Cox Proportional Hazards model adjusting for treatment group and stratification factors.

Subject with exacerbation of hypertension is defined as meeting at least 1 of the following criteria:

- An increase from baseline of ≥ 20 mmHg predialysis systolic BP (sBP) and predialysis SBP >170 mmHg, or
- An Increase from baseline of ≥ 15 mmHg predialysis diastolic BP (dBP) and predialysis DBP >100 mmHg
- Proportion of subjects with a treatment-emergent AE of hypertension. The proportion of subjects with a TEAE of hypertension will be compared between the 2 treatment groups using logistic regression model adjusting for baseline predialysis systolic and diastolic BP as covariates and treatment group and stratification factors as fixed effects.
- Proportion of subjects with exacerbation from baseline of antihypertensive therapy. The proportion of subjects with exacerbation from baseline of antihypertensive therapy will be compared between the 2 treatment groups using logistic regression model adjusting for baseline predialysis systolic and diastolic BP as covariates and treatment group and stratification factors as fixed effects.

11.3.6 Vascular Access Thrombosis (HD subjects):

- Time to a TEAE of vascular access thrombosis will be analyzed using the Cox Proportional Hazards model adjusting for treatment group and stratification factors.
- Proportion of subjects with a TEAE of vascular access thrombosis will be analyzed using CMH model adjusting for treatment group and stratification factors.

11.3.7 Health Related Quality of Life (HRQoL) and EQ-5D-5L Benefits of Anemia Therapy

The Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) is a multi-purpose, short-form health survey with 36 questions (Appendix 3). It yields an 8-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures. It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group. Accordingly, the SF-36 has proven useful in surveys of general and specific populations, comparing the relative burden of diseases, and in differentiating the health benefits produced by a wide range of different treatments.
The SF-36 contains 36 items that measure eight dimensions: (1) physical functioning (PF); (2) role limitations due to physical health problems (RP); (3) bodily pain (BP); (4) social functioning (SF); (5) general health perceptions (GH); (6) role limitations due to emotional problems (RE); (7) vitality, energy or fatigue (VT); and (8) mental health (MH).
Item scores for each dimension are coded, summed, and transformed to a scale from 0 to 100, with higher scores indicating better self-perceived health (Table 7. The transformed items are then averaged to give the subscales. The subscales are averaged to give the composite scores. The composite scores are averaged to give the overall score. The reliability and validity of the SF-36 is well documented in a variety of different subject groups, including subjects with vascular diseases.

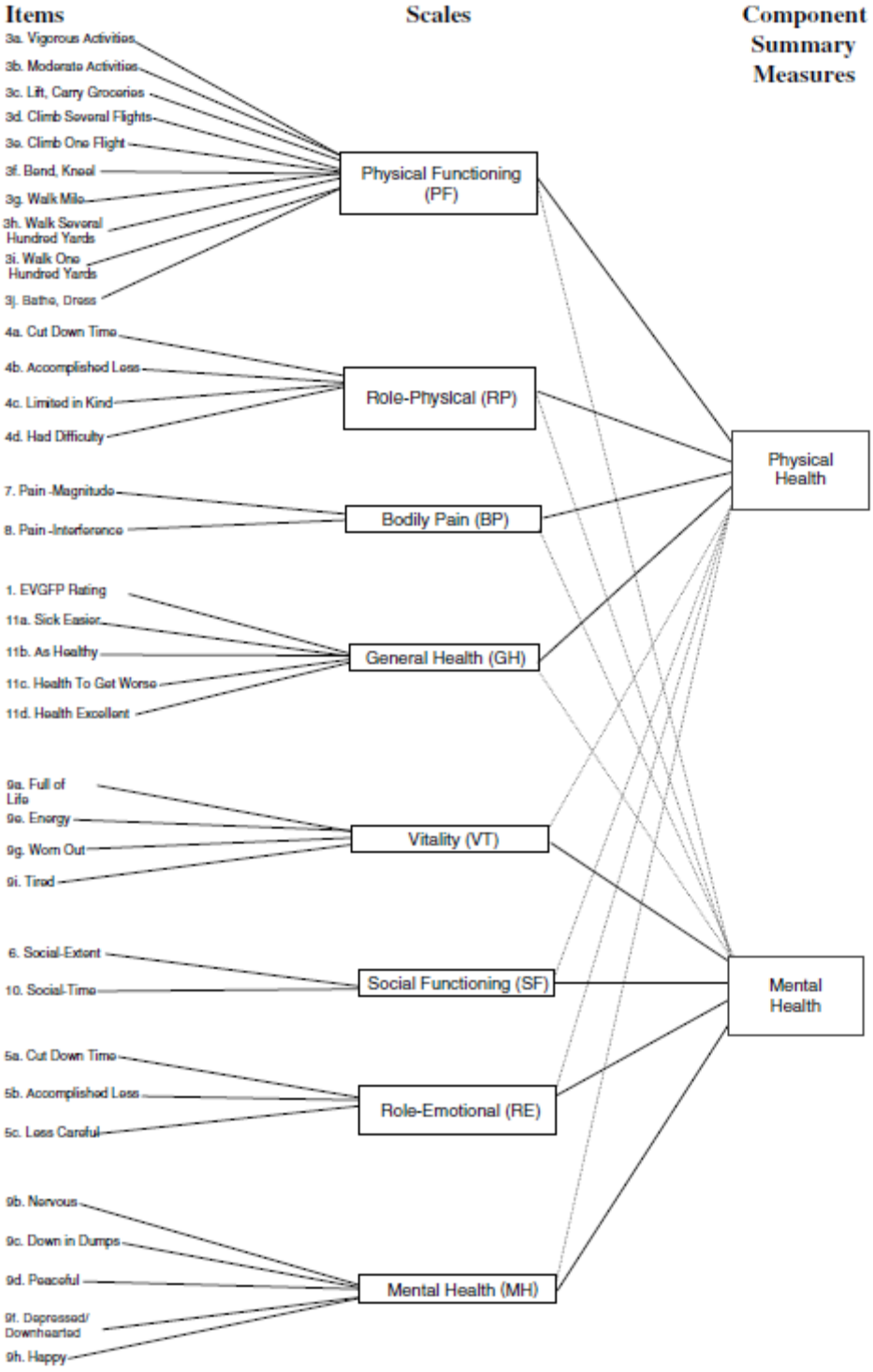
Table 7. Transformation of SF-36 items to a Scale from 0 to 100

Question	**Self-perceived health**

	0 (Poor)	**100 (Better)**
1, 2, 6, 8, 9a, 9d, 9e, 9h, 11b, 11d	5	1
3a to 3j	1	3
4a to 4d, 5a to 5c, 9b, 9c, 9f, 9g, 9i, 10, 11a, 11c	1	5
7	6	1
Programming note: Once transformed – the items should each span the range 0 to 100.		

For each of the 8 dimensions, if less than 50% of the items which constitute that dimension are missing, the dimension score will be calculated by the mean of the available non-missing items. The physical health composite summary and mental health composite summary will only be calculated if a score has been calculated for all 4 dimensions that constitute the composite summary.

Figure 3. SF-36 Model



The mean change from baseline at weeks 12, 28, and 52 weeks will be computed for each treatment group for the following endpoints. A paired t-test will be used to assess within treatment effect. Between treatment difference will be assessed using MMRM using baseline subscore as a covariate and treatment group, visit, the interaction of treatment and visit, and stratification factors as fixed effects. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure. Non-inferiority of roxadustat vs. Epoetin will be tested. The non-inferiority margin is fixed as a difference of 2 points.
Mean change in these endpoints at above-mentioned time points will be analyzed and compared between the 2 treatment groups using MMRM using the baseline value as a covariate, treatment group, visit, the interaction of treatment and visit, adjusting for the stratification factors. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure. Other than below endpoints for exploratory analyses, other QoL variables may be performed.

HRQoL benefit will be assessed using SF-36 vitality and physical functioning subscales

- **Change from Baseline in SF-36 Vitality (VT) Sub-Score**

The change from baseline in SF-36 VT sub-score to the average VT sub-score will be analyzed using MMRM with baseline VT subscore as a covariate and treatment group, visit, the interaction of treatment and visit, and stratification factors as fixed effects. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure. Data up to visit of Week 52 will be included in the analyses. The estimates for the difference of change from baseline in SF-36 VT for Week 12 to 28 between two treatment groups will be generated from an estimate statement. Non-inferiority of roxadustat vs. Epoetin will be tested.

- **Change from Baseline in SF-36 Physical Functioning (PF) Sub-Score**

Change from baseline in SF-36 PF sub-score to the average PF sub-score of weeks 12 to 28 will be analyzed using MMRM model with baseline PF subscore as a covariate, treatment group, visit, the interaction of treatment and visit, and stratification factors as fixed effects. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure. Data up to visit of Week 52 will be included in the analyses. The estimates for the difference of change from baseline in SF-36 PF for Week 12 to 28 between the two treatment groups will be generated from an estimate statement.

- Vitality Sub-score of SF-36: In subjects with baseline vitality sub-score below 50.
- Physical Functioning Sub-scores of SF-36:
    - In FAS subjects with baseline physical Functioning scores below 40.
    - In all FAS subjects.
- Anemia Subscale (“Additional Concerns”) of Functional Assessment of Cancer Therapy-Anemia (FACT-An) Scores:
    - In FAS subjects with baseline subscale scores below 55 (generally associated with fatigue).
    - In all FAS subjects.
- Total FACT-An Scores:
    - In FAS subjects with baseline FACT-An scores below 135
    - In all FAS subjects.
- EQ-5D-5L Scores: In all FAS subjects.

11.3.8 Hepcidin, Iron, and HbA1c

The mean change from baseline in the following endpoints will be analyzed using MMRM model with baseline value as a covariate, treatment group, visit, the interaction of treatment and visit, and stratification factors as factors. The same strategy as that used in MMRM for Hb will be used to choose variance covariance structure

- Change in serum hepcidin at each of the selected time points during which tested, weeks 4, 12, 20, 36, and 52.
- Change in serum ferritin at each of the selected time points it was tested, week 4, 8, 12, 20, 28, 36, 44 and 52, total and sub-grouped by baseline values of <100 ng/mL, 100 to 400 ng/mL, and >=400 ng/mL.
- Change in TSAT at each of the selected time points, week 4, 8, 12, 20, 28, 36, 44 and 52, total and sub-grouped by baseline values of < 20%, 20 to 40, and >=40%.
- Change in HbA1c level at each of the selected time points, week 12, 24, 36and 52, in subjects without history of diabetes, in subjects with history of diabetes.

12 SAFETY ANALYSES

The safety analysis will be performed using the Safety Population.

For US regulatory submission, safety analyses will be based on all patients overall and stratified by stable dialysis and newly initiated dialysis,

Safety parameters include adverse events, laboratory parameters, vital signs, ECG parameters, and physical examinations.

The cardiovascular safety assessment of roxadustat will also be based on pooled analysis of composites of adjudicated cardiovascular events pooled across multiple global Phase 3 clinical studies which includes this study, according to the pooled SAPs for meeting regional regulatory requirements.

For each safety parameter, the last assessment made prior to the first dose of study medication will be used as the BL for all analyses of that safety parameter.

12.1 Adverse Events

Adverse events will be coded using the latest MedDRA version. An AE (classified by preferred term) started during the Treatment Period will be considered a treatment-emergent adverse event (TEAE) if it was not present prior to the first dose of study medication, or it was present prior to the first dose of study medication but increased in severity during the Treatment Period up to 7days after last dose of study drug or until the administration of another anemia drug (other than the randomized treatment). An AE that occurs more than 7 days after study medication or after the administration of another anemia drug (other than the randomized treatment) will not be counted as a TEAE.

The number and percentage of subjects reporting TEAEs in each treatment group will be tabulated by system organ class and preferred term; by system organ class, preferred term, and severity; and by system organ class, preferred term, and relationship to study medication. If more than one event occurs with the same preferred term for the same subject, the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to the study medication. The distribution of TEAEs by severity and relationship to study medication will be summarized by treatment group.

The incidence of common (≥ 5% of subjects in any treatment group) TEAEs, treatment-emergent serious AEs (TESAE), and AEs leading to discontinuation of study medication will be summarized by SOC, preferred term and treatment group, sorted in decreasing frequency for the test treatment. In addition, the incidence of death and fatal SAEs (i.e., events that caused death) will be summarized separately by treatment group, SOC and preferred term. Treatment emergent adverse events of interest will also be reported in terms of incidence rate per Patient Exposure Year (PEY).

Discontinuation of study medication due to worsening anemia based on hemoglobin value (lack of efficacy) will not be counted under AEs leading to discontinuation.
Listings will be presented of subjects with serious adverse events (SAEs), subjects with adverse events leading to discontinuation, and subjects who died.

Temporal profile of TEAEs of special interest may also be plotted by treatment group showing the subjects in the y-axis and time to these TEAEs in the x-axis (Appendix 8).

12.2 Clinical Laboratory Parameters

Descriptive statistics for laboratory values (in US conventional and SI units) and changes from baseline at each assessment time point will be presented by treatment group for the following laboratory parameters collected in the study including but are not limited to the following:

- Hematology: Hemoglobin, hematocrit, RBC count, MCV, WBC count, WBC differential, and platelet counts, reticulocyte count;
- Chemistry: Alkaline phosphatase, ALT, AST, GGT, total bilirubin, LDH, lipase, total protein, albumin, fasting glucose, phosphate, uric acid, BUN, creatinine, sodium, and potassium;
- Serum iron, ferritin, TIBC, TSAT, reticulocyte hemoglobin content
- HbA1c
- Hepcidin, high-sensitivity C-reactive protein (hs-CRP)

Laboratory tests values are clinically significant (CS) if they meet either the low or high CS criteria (Appendix 6). The number and percentage of subjects with post-baseline CS values will be tabulated by treatment group. The percentages are to be calculated relative to the number of subjects with available non-CS baseline values and at least one post-baseline assessment. The numerator is the total number of subjects with at least one post-baseline CS value. In addition, shift tables will be presented by treatment group and time point. The following 3 data listings will be presented by subject:

- A listing of lab values for all lab tests at all collected time points.
- A listing of subjects with post-baseline CS values will be provided including the baseline and post-baseline values.
- A listing of all AEs for subjects with CS laboratory values will also be provided.

12.3 Vital Signs

Blood pressures and heart rate baselines for predialysis are defined as the mean of values obtained from the last 6 weeks of screening including Day 1 prior to the first dose.
Descriptive statistics for vital signs (e.g., systolic and diastolic blood pressure, MAP, pulse rate, and respiratory rate) and their changes from baseline at each visit and at the end of study will be presented by treatment group.
Vital sign values are potentially clinically significant (PCS) if they meet both the observed value criteria and the change from baseline criteria listed in Table 8. The number and percentage of subjects with post-baseline PCS values will be tabulated by treatment group. The percentages are to be calculated relative to the number of subjects with baseline and at least one post-baseline assessment. The numerator is the total number of subjects with at least one post-baseline PCS vital sign value. Shift tables may be presented. A supportive listing of subjects with post-baseline PCS values will be provided including the subject ID, study center, baseline, and post-baseline values. A listing of all AEs for subjects with PCS vital signs will also be provided.

Table 8. Criteria for Potentially Clinically Significant Vital Signs

Vital Sign Parameter	Flag	Criteria*	
		Observed Value	Change from Baseline
Systolic Blood Pressure (mmHg)	High	≥ 170	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20

Diastolic Blood Pressure (mmHg)	High	≥ 100	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Pulse Rate (bpm)	High	≥ 120	Increase of ≥ 20
	Low	≤ 50	Decrease of ≥ 20
Weight (kg)	High	-	Increase of ≥ 10%
	Low	-	Decrease of ≥ 10%
*A post-baseline predialysis value is considered as a PCS value if it meets both criteria for observed value and change from predialysis baseline, respectively.			

Additional analyses include but are not limited to

- Subgroup analyses of subjects without any change in BP meds during Treatment Period
- Proportion of subjects meeting National Kidney Foundation BP target: within sBP 120-140 mmHg/dBP 70-90 mmHg at baseline, during treatment, and 4 weeks post treatment

12.4 Electrocardiogram (ECG)

Descriptive statistics for ECG parameters (e.g., Heart Rate, PR interval, QRS interval, QT interval, and QTc interval) at baseline and changes from baseline at each assessment time point will be presented by treatment group. QTc interval will be calculated using both Bazett ($QTcB = QT/ (RR)^{1/2}$) and Fridericia ($QTcF = QT/(RR)^{1/3}$) corrections; and if RR is not available, it will be replaced with 60/HR in the correction formula.

A plot for each parameter of mean (+/- 95% CI) versus visit will be produced by treatment group (roxadustat vs. Epoetin).

ECG parameters values are potentially clinically significant (PCS) if they meet or exceed the upper limit values listed in Table 9. The number and percentage of subjects with post-baseline PCS values will be tabulated by treatment group. The percentages are to be calculated relative to the number of subjects with available non-PCS baseline and at least one post-baseline assessment. The numerator is the total number of subjects with at least one post-baseline PCS ECG value. Shift tables may be presented. A listing for all subjects with post-baseline PCS value will be provided including the subject ID, study center, baseline, and post-baseline PCS values.

In addition, a listing of all TEAEs for subjects with PCS ECG values and a listing of subjects with post-baseline significant ECG abnormalities as reported by the investigators will also be provided.

Table 9. Criteria for Potentially Clinically Significant ECG

ECG Parameter	Unit	Higher Limit
QRS interval	Msec	≥ 150
PR interval	Msec	≥ 250
QTc interval	Msec	> 500; Change from baseline > 60

12.5 Other Safety Analyses

A separate meta-analysis SAP for pre-specified, adjudicated (blinded) composite safety endpoints to assess cardiovascular, cerebrovascular and thrombo-embolic events will be developed to complement this study specific SAP.

13 ADDITIONAL AND SUBGROUP ANALYSES

Selected efficacy and safety analyses may be performed separately by sex, age group, baseline iron replete status and baseline stratification factors.

14 INTERIM ANALYSES

Safety data and dosing decisions will be monitored on an ongoing basis. Additional ongoing review of safety data will be conducted by an independent data and safety monitoring board (DSMB).

15 REFERENCES

ICH Harmonized Tripartite Guideline E3. Structure and Content of Clinical Study Reports, November 1995. (www.ich.org; Guidelines; "Efficacy" Topics)

ICH Harmonized Tripartite Guideline E9. Statistical Principles for Clinical Trials, February 1998. (www.ich.org; Guidelines; "Efficacy" Topics)

Carpenter JR, Roge JH and Kenward MG, Analysis of longitudinal trials with protocol deviation:|a Framework for Relevant, Accessible Assumptions, and Inference via Multiple Imputation. Journal of Biopharmaceutical Statistics, issue 6 (November/December) in volume 23 (2013). 1352-137

Cella, D. The functional assessment of Cancer Therapy-Anemia (FACT-An) Scale: A new tool for the assessment of outcome in cancer anemia and fatigue. Hematology Seminars 1997; 34: 13-19

Ge M, Durham LK, Meyer RD, Xie W and Thomas N. Covariate-Adjusted Difference in Proportions from Clinical Trials Using Logistic Regression and Weighted Risk Differences. *Drug Information Journal* 2011 45: 481

Haybittle, J. L. (1971), "Repeated Assessment of Results in Clinical Trials of Cancer Treatment," *British Journal of Radiology*, 44, 793-797.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., and Smith, P. G. (1976), "Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient: I. Introduction and Design," *British Journal of Cancer*, 34, 585-612.

FDA Guidance for Industry Non-Inferiority Clinical Trials, 2010 (http://www.fda.gov/downloads/Drugs/Guidances/UCM202140.pdf).

Little RJA, Rubin D.1987. Statistical Analysis with Missing Data. John Wiley and Sons.

Ratitch B, Kelly M 2011. Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures, PharmaSUG2011 – Paper SP4.

Rubin, D. B. (1987) Multiple imputation for nonresponse in surveys. New York: Wiley.

Schafer, J. L. (1997) Analysis of incomplete multivariate data. London: Chapman and Hall.

16 APPENDICES

16.1 Appendix 1 Schedule of Assessments

Study Period:	Screening (up to 6 to 8 Wks [a])			Treatment Period (≥ 52 Wks)						Follow-up Period (4 Wks)	Unscheduled Visit
					± 2 days		± 4 days	± 4 days	± 7 days	±7 days	
					Weekly (Wks 1-2)						
Visits:	S1	S2	S3	Day 1 [b]	1	2	Every 2 Wks (Wks 4-24)	Every 4 Wks (from Wk 28)	EOT/ET	EOS (EOT/ET +4 wks) [c]	
Written ICF	X										
Randomization				X							
Eligibility criteria	X			X							
Demographics	X										
Height (screening only), weight (dry weight [post dialysis] in HD subjects)	X			X				Wks 36, 52 + every 24 wks	X	X	
Medical history	X										
Physical examination		X		X			Wks 12 [d], 24 [d]	Wks 36 [d], 52 [d] + every 24 wks [d]	X	X [d]	O
Vital Signs: BP [e], HR [e], RR [f], Temp	X	X	X	X	X	X	X	X	X	X	X
Local hemoglobin Assessments					X	X	X	X	X		X
CBC with WBC diff	X			X	X	X	Wks 4, 8, 12, 20	Wks 28, 36, 44 + every 8 wks	X	X	O
Hemoglobin [g]		X	X				X [g]	X [g]			X
Reticulocyte count and CHr	X			X	X	X	Wks 4, 6, 8, 12, 16, 20	Wks 28, 36, 52 + every 16 wks	X	X	O
12-lead ECG				X			Wk 24	Wks 36, 52 + every 24 wks	X		O
Renal ultrasound [h]		X									
Serum chemistry	X			X			Wks 4, 8, 12, 20	Wks 28, 36, 44 + every 8 wks	X	X	O
LFTs						X	Wks 6, 16, 24	Wk 32			O
Serum Lipid panel	X			X			Wks 4, 8, 12, 20	Wks 28, 36, 48 + every 12 wks	X	X	O
Serum iron, ferritin, TIBC, UIBC, TSAT	X			X			Wks 4, 8, 12, 20	Wks 28, 36, 44 + every 8 wks	X	X	O
HbA1c	X			X			Wks 12, 24	Wks 36, 52 + every 52 wks	X	X	O
Vitamin B_{12}, folate	X										
HIV ELISA, HBsAg, anti-Hcardiovascular Ab	X										
Serum (hCG) pregnancy test (every 12 weeks) [i]			X				Wks 12, 24	Wks 36, 48 + every 12 wks	X		O
hs-CRP, hepcidin				X			Wks 4, 12, 20	Wks 36, 52	X	X	
Archival serum & plasma samples for biomarkers				X			Wks 4, 12, 20	Wks 36, 52 + every 24 wks	X	X	
HRQoL Questionnaires [j]				X			Wks 12	Wks 28, 52	X		
Study drug: dispensing and/or accountability [k]				X	X	X	X	X	X		

Study Period:	Screening (up to 6 to 8 Wks [a])			Treatment Period (≥ 52 Wks)						Follow-up Period (4 Wks)	Unscheduled Visit
					± 2 days		± 4 days	± 4 days	± 7 days	±7 days	
					Weekly (Wks 1-2)						
Visits:	S1	S2	S3	Day 1 [b]	1	2	Every 2 Wks (Wks 4-24)	Every 4 Wks (from Wk 28)	EOT/ET	EOS (EOT/ET +4 wks) [c]	
Dose adjustment review [l]					X	X	X	X			
AE and concomitant medication recording	X	X	X	X	X	X	X	X	X	X	X
Procedure and nondrug therapy recording	X	X	X	X	X	X	X	X	X	X	X

Abbreviations: Ab = antibody; AE = adverse event; BP = blood pressure; CBC = complete blood count; cHR = reticulocyte hemoglobin content; cardiovascular = cardiovascular; ECG = electrocardiogram; ELISA = enzyme-linked immunosorbent assay; EOT = End of Treatment; ET=Early Termination; EOS = End of Study; HbA1c = glycated hemoglobin; HBsAg = hepatitis B surface antigen; hCG = human chorionic gonadotropin; Hcardiovascular = hepatitis C virus; HD = hemodialysis; HIV = human immunodeficiency virus; HR = heart rate; HRQoL = health-related quality of life questionnaire; hs-CRP = high-sensitivity C-reactive protein; ICF = informed consent form; LFTs = liver function tests; O = optional test/assessment; PE = physical examination; RBC = red blood cell; RR = respiratory rate; TIBC = total iron binding capacity; UIBC = unsaturated iron binding capacity; TSAT = transferrin saturation; WBC = white blood cells; Wk(s) = week(s); X = mandatory test/assessment.

Notes: In subjects on HD, all lab sampling is to be performed prior to the dialysis session.

Subjects who discontinue study medication prematurely will be followed up until the end of the study, unless consent to participate is withdrawn. Upon completion of EOT and EOS visits, these subjects will be followed up every 3-6 month interval for vital status, cardiovascular events, and hospitalization (depending on the availability of subjects) until study closure. These visits may occur either in-person or via telephone.

a All screening procedures should be completed within 6 weeks. For subjects currently taking Mircera®, the screening period can be extended up to 8 weeks.

b Day 1 is a due treatment date per the existing ESA treatment schedule. All study assessments are to be performed prior to first study drug administration. In subjects on HD study assessments with the exception of HRQoL and body-weight measurement should also be performed prior to or at initiation of dialysis; HRQoL assessments should be done approximately during dialysis (preferably at the beginning) and weight measurement after dialysis.

c Subjects discontinuing study medication prematurely will complete EOT and EOS. Unless consent is withdrawn, cardiovascular events, vital status and hospitalizations will be collected every 3 to 6 months (phone or in-person clinic visit) for these subjects until study closure

d Targeted PE (general appearance, CV, respiratory and abdomen)

e BP and HR should be measured using automated calibrated instruments at pre dialysis in subjects on HD and preferably approximately at the same time in subjects on PD. Measurement to occur prior to study drug administration (if applicable); BP and HR should be assessed per guidelines provided in Appendix 5

f Respiratory rate should be measured at predialysis in subjects on HD and preferably approximately at the same time in subjects on PD. Measurement should occur prior to study drug administration (if applicable)

g Separate hemoglobin samples should be collected at all the visits where a CBC is not collected (i.e., hemoglobin at Weeks 6, 10, 14, 16, 18, 22, 24, 32, 40, 48, etc)

h Not required if result of a renal ultrasound report within 12 weeks prior to randomization is available

i Collect from female subjects of childbearing potential only

j Including SF-36, FACT-An, EQ-5D 5L, HRQoL assessments should be administered approximately at the same time of the day. Example: in HD subjects, questionnaires to be completed by subject approximately at the same time during dialysis and in PD subjects approximately at the same time of the day

k Initial dosing of roxadustat and epoetin alfa per Tables 1 and 2, respectively. Dose adjustment per Appendix 2 (roxadustat subjects) or epoetin package insert or SmPC (epoetin subjects); dose review for predefined out of range hemoglobin elevations at every visit. All subjects on HD will receive epoetin alfa TIW intravenously (IV). For peritoneal dialysis subjects, the route of administration (IV or SC) of epoetin alfa may remain the same as baseline (per discretion of the Investigator)

l Subjects randomized to roxadustat: dose adjustment begins at Week 4 refer to Appendix 2

Notes: In subjects on HD, all lab sampling is to be performed prior to the dialysis session.

.

16.2 Appendix 2 Roxadustat Dose Adjustment Rules

Following Day 1, at each study visit during the Treatment Period, hemoglobin will be measured (prior to dialysis in HD subjects) locally to determine the need for a dose adjustment or to assess for predefined out of range hemoglobin elevations. This local hemoglobin measurement may be made with the use of a point-of-care device (e.g., HemoCue®, CritLine®) or by local laboratory (i.e., Stat Lab*) testing if the point-of-care device is not available or the result is considered unreliable. In the event that the hemoglobin value of the visit as determined by the study central laboratory is significantly different from that measured locally, and per Investigator that warrants a reversal of the dose adjustment decision made earlier based on locally measured hemoglobin value, the Medical Monitor should be informed, if possible.

In this study, a rate of rise of hemoglobin > 2 g/dL in 4 weeks or a hemoglobin level of ≥ 13 g/dl at any time would be considered as excessive haematopoiesis and would require either dose reduction or temporary dose hold.

The dose of roxadustat will remain constant during the first 4 weeks of the Treatment Period unless a dose reduction is required for predefined out of range hemoglobin elevations. Roxadustat dose adjustments are permitted from Week 4 onwards, and every 4 weeks thereafter (e.g., Week 4, Week 8, Week 12); however, the dose may be adjusted between two prespecified windows (e.g., anytime between Week 4 and Week 8 visits, Week 8 and Week 12 visits) if all of the following criteria are met:

- No dose adjustment has been made in last 4 weeks
- Hemoglobin < 9.0 g/dL.

Any dose adjustment will reset the dose-adjustment window to every 4 weeks thereafter (e.g., dose adjustment for a qualified reason at Week 6 leads to next dose adjustment at Week 10). Dose adjustment must not be done sooner than 4 weeks unless discussed with and preapproved by the medical monitor. Dose adjustments for roxadustat are described in the Table 10 below.

Subjects randomized to roxadustat will take doses TIW for the entire duration of the Treatment Period. Dosing frequency may only be adjusted to BIW or QW if a subject requires < 20 mg TIW (i.e., < 60 mg per week) to maintain a hemoglobin level of approximately 11 g/dL. The Medical Monitor should be notified as soon as possible of such dose change.

Table 10. Roxadustat Dose Adjustment Rules

Change in hemoglobin over past 4 weeks (g/dL)	Hemoglobin (g/dL) at Dose Adjustment Review Visit			
	< 10.5	**10.5 to < 12.0**	**12.0 to < 13.0**	**≥ 13.0**
< -1.0	↑	↑	No change	Hold dosing, check hemoglobin at least weekly, then resume dosing when hemoglobin < 12.0 g/dL, at a dose that is reduced by one dose step
-1.0 to 1.0	↑	No change	↓	
> 1.0	No change	↓	↓	

Abbreviations: ↑ = increase; ↓ = decrease.

Notes:

Dose Increases and Reductions:

- Dose increases (↑) and reductions (↓) are preset.
- The dose steps are as follows: 20, 40, 50, 70, 100, 150, 200, 250, 300, and 400 mg.
  For example, a subject previously receiving 70 mg of roxadustat requiring a dose increase would have his dose changed to100 mg of roxadustat. A subject previously receiving 150 mg of roxadustat requiring a dose reduction would have his dose changed to 100 mg of roxadustat.
  Medical Monitor should be informed when < 20 mg /dose is required.
- The maximum dose is capped at 400 mg or 3.0 mg/kg/dose (whichever is lower).

Dose Adjustment for predefined out of range hemoglobin elevations:

- If hemoglobin increases by > 2.0 g/dL at any time within 4 weeks, the dose should be reduced by one dose step.
- Only one dose reduction for predefined out of range hemoglobin elevations is recommended within a period of 4 weeks.
- Temporary dose holds for predefined out of range hemoglobin elevations should be confirmed based on central lab hemoglobin assessments. Upon confirmation, if the hemoglobin is ≥ 13 g/dL, withhold dosing, check hemoglobin at least weekly, then resume dosing when hemoglobin < 12.0 g/dL, at a dose that is reduced by one dose step. The subsequent hemoglobin follow-up assessments should also be done by a central lab.

* A “Stat lab" is to be used only for urgent lab test that is needed for immediate decision making related to protocol, management of adverse events as determined by the Investigator.

16.3 Appendix 3 SF-36 (Version 2)

Your Health and Well-Being

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities. *Thank you for completing this survey!*

For each of the following questions, please tick the one box that best describes your answer.

1. **In general, would you say your health is:**

Excellent	Very good	Good	Fair	Poor
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

2. **<u>Compared to one year ago</u>, how would you rate your health in general <u>now?</u>**

Much better now than one year ago	Somewhat better now than one year ago	About the same as one year ago	Somewhat worse now than one year ago	Much worse now than one year ago
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

SF-36 (Version 2)

3. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
a Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports	☐ 1	☐ 2	☐ 3
b Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	☐ 1	☐ 2	☐ 3
c Lifting or carrying groceries	☐ 1	☐ 2	☐ 3
d Climbing several flights of stairs	☐ 1	☐ 2	☐ 3
e Climbing one flight of stairs	☐ 1	☐ 2	☐ 3
f Bending, kneeling, or stooping	☐ 1	☐ 2	☐ 3
g Walking more than a mile	☐ 1	☐ 2	☐ 3
h Walking several hundred yards	☐ 1	☐ 2	☐ 3
i Walking one hundred yards	☐ 1	☐ 2	☐ 3
j Bathing or dressing yourself	☐ 1	☐ 2	☐ 3

SF-36 (Version 2)

4. **During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of your physical health?**

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a Cut down on the amount of time you spent on work or other activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b Accomplished less than you would like	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c Were limited in the kind of work or other activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d Had difficulty performing the work or other activities (for example, it took extra effort)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

5. **During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?**

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a Cut down on the amount of time you spent on work or other activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b Accomplished less than you would like	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c Did work or other activities less carefully than usual	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

SF-36 (Version 2)

6. **During the <u>past 4 weeks</u>, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbours, or groups?**

Not at all	Slightly	Moderately	Quite a bit	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

7. **How much <u>bodily</u> pain have you had during the <u>past 4 weeks</u>?**

None	Very mild	Mild	Moderate	Severe	Very severe
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

8. **During the <u>past 4 weeks</u>, how much did <u>pain</u> interfere with your normal work (including both work outside the home and housework)?**

Not at all	A little bit	Moderately	Quite a bit	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

SF-36 (Version 2)

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks…

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a Did you feel full of life?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b Have you been very nervous?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c Have you felt so down in the dumps that nothing could cheer you up?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d Have you felt calm and peaceful?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
e Did you have a lot of energy?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
f Have you felt downhearted and low?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
g Did you feel worn out?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
h Have you been happy?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
i Did you feel tired?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

10. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

All of the time	Most of the time	Some of the time	A little of the time	None of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

SF-36 **(Version 2)**

11. How TRUE or FALSE is <u>each</u> of the following statements for you?

	Definitely true	Mostly true	Don't know	Mostly false	Definitely false
a I seem to get ill more easily than other people	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b I am as healthy as anybody I know	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c I expect my health to get worse	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d My health is excellent	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Thank you for completing these questions!

16.4 Appendix 4 FACT-An (Version 4)

Below is a list of statements that other people with your illness have said are important.
Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	PHYSICAL WELL-BEING	Not at all	A little bit	Some-what	Quite a bit	Very much
GP1	I have a lack of energy	0	1	2	3	4
GP2	I have nausea	0	1	2	3	4
GP3	Because of my physical condition, I have trouble meeting the needs of my family	0	1	2	3	4
GP4	I have pain	0	1	2	3	4
GP5	I am bothered by side effects of treatment	0	1	2	3	4
GP6	I feel ill	0	1	2	3	4
GP7	I am forced to spend time in bed	0	1	2	3	4

	SOCIAL/FAMILY WELL-BEING	Not at all	A little bit	Some-what	Quite a bit	Very much
GS1	I feel close to my friends	0	1	2	3	4
GS2	I get emotional support from my family	0	1	2	3	4
GS3	I get support from my friends	0	1	2	3	4
GS4	My family has accepted my illness	0	1	2	3	4
GS5	I am satisfied with family communication about my illness	0	1	2	3	4
GS6	I feel close to my partner (or the person who is my main support)	0	1	2	3	4
Q1	*Regardless of your current level of sexual activity, please answer the following question. If you prefer not to answer it, please mark this box ☐ and go to the next section.*					
GS7	I am satisfied with my sex life	0	1	2	3	4

FACT-An (Version 4)

Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	EMOTIONAL WELL-BEING	Not at all	A little bit	Some-what	Quite a bit	Very much
GE1	I feel sad	0	1	2	3	4
GE2	I am satisfied with how I am coping with my illness	0	1	2	3	4
GE3	I am losing hope in the fight against my illness	0	1	2	3	4
GE4	I feel nervous	0	1	2	3	4
GE5	I worry about dying	0	1	2	3	4
GE6	I worry that my condition will get worse	0	1	2	3	4

	FUNCTIONAL WELL-BEING	Not at all	A little bit	Some-what	Quite a bit	Very much
GF1	I am able to work (include work at home)	0	1	2	3	4
GF2	My work (include work at home) is fulfilling	0	1	2	3	4
GF3	I am able to enjoy life	0	1	2	3	4
GF4	I have accepted my illness	0	1	2	3	4
GF5	I am sleeping well	0	1	2	3	4
GF6	I am enjoying the things I usually do for fun	0	1	2	3	4
GF7	I am content with the quality of my life right now	0	1	2	3	4

FACT-An (Version 4)

Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	ADDITIONAL CONCERNS	Not at all	A little bit	Some-what	Quite a bit	Very much
HI7	I feel fatigued	0	1	2	3	4
HI12	I feel weak all over	0	1	2	3	4
An1	I feel listless ("washed out")	0	1	2	3	4
An2	I feel tired	0	1	2	3	4
An3	I have trouble starting things because I am tired	0	1	2	3	4
An4	I have trouble finishing things because I am tired	0	1	2	3	4
An5	I have energy	0	1	2	3	4
An6	I have trouble walking	0	1	2	3	4
An7	I am able to do my usual activities	0	1	2	3	4
An8	I need to sleep during the day	0	1	2	3	4
An9	I feel lightheaded (dizzy)	0	1	2	3	4
An10	I get headaches	0	1	2	3	4
B1	I have been short of breath	0	1	2	3	4
An11	I have pain in my chest	0	1	2	3	4
An12	I am too tired to eat	0	1	2	3	4
BL4	I am interested in sex	0	1	2	3	4
An13	I am motivated to do my usual activities	0	1	2	3	4
An14	I need help doing my usual activities	0	1	2	3	4
An15	I am frustrated by being too tired to do the things I want to do	0	1	2	3	4
An16	I have to limit my social activity because I am tired	0	1	2	3	4

16.5 Appendix 5 Data Handling Conventions

16.5.1 Visit Time Windows

The table below presents the visits assigned for efficacy and safety analyses corresponding to the range of treatment days (window) during which an actual visit may have occurred.

Table 11. Visit Time Windows

Derived Visit	Scheduled Visit Day[a]	Window
Baseline, Week 0	Day 1	Days ≤ 1
Week 1	Day 7*(Week #)+1	Days [Day 2, 10]
Week 2	Day 7*(Week #)+1	Days [Scheduled Day -3 Scheduled Day ± 6]
Weeks 4 prior to 24	Day 7*(Week #)+1	Days [Scheduled Day -7, Scheduled Day ± 6]
Week 24	Day 7*(Week #)+1	Days [Scheduled Day -7 Scheduled Day ± 13]
>=Week 28	Day 7*(Week #)+1	Days [Scheduled Day -14, Scheduled Day ±13]
ET	Earlier Termination, match to a closest scheduled visit in protocol if patient had not been off drug for more than 7 days.	
EoT	Last assessment between Day 2 and EOT visit day, match to a closest scheduled visit in protocol if patient had not been off drug for more than 7 days.	
EoS (FU-4Wk)	Final visit for the Study 15 – 31 days after the last dose (excluding long term follow-up for early termination).	

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.

Table 12: Analysis Visit Windows for QoL

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Day 1	Day 1	Day 1	Baseline
Week 12	Day 7 * (Week #) + 1	Target day -14, Target Day + 27	Week 12
Week 28	Day 7 * (Week #) + 1	Target Day – 28, Target Day +83	Week 28
Week 52	Day 7 * (Week #) + 1	Target Day – 84, Target Day + 83	Week 52
EOT Visit		Last assessment between Day 2 and EOT visit day, remapped to the closest next scheduled visit for HRQoL collection.	Week 12, 28, 52 and > 1 year

[a]: Relative to Day 1 (first dose date of study medication)

Table 13. Analysis Visit Windows for Lipid Panel

Derived Visit	Scheduled Visit Day [a]	Window
Baseline, Week 0	Day 1	Day 1
Week 4	Day 7*(Week #)+1	[Day 2, Day +42]
Week 8	Day 7*(Week #)+1	[Scheduled Day-14, Scheduled Day +13]
Weeks 12	Day 7*(Week #)+1	[Scheduled Day -14, Scheduled Day +27]
Week 20, 28	Day 7*(Week #)+1	[Scheduled Day -28, Scheduled Day +27]
Weeks 36	Day 7*(Week #)+1	[Scheduled Day -28, Scheduled Day +41]
Week 48 to xx	Day 7*(Week #)+1	[Scheduled Day -42, Scheduled Day +41]
Week xx	.	Last assessment between Day 2 and EOT visit day, match to a closest scheduled visit in protocol if patient had not been off drug for more than 7 days
FU-4Wk		Final visit in the Study 15 – 31 days after Last Dose.

[a]: Relative to Day 1 (first dose date of study medication)

Visit Day is calculated by (visit date – first dose date of study medication + 1). If a subject has ≥ 2 actual visits within the same window, the last visit with non-missing value will be used for analysis.

Repeated or Unscheduled Assessments of Safety Parameters
If a subject has repeated assessments prior to the start of study medication, then the results from the final assessment made prior to the start of study medication will be used as baseline. If end of study assessments are repeated or unscheduled, the last post-baseline assessment will be used as the end of study assessment for generating summary statistics. However, all post-baseline assessments will be used for PCS value determination and all assessments will be presented in the data listings.

Missing Date of Study Medication
When the last date of study medication during the study treatment phase is missing, all efforts should be made to obtain the date from the investigator. If it is still missing after all efforts, then the last visit date during the Treatment Period will be used in the calculation of treatment duration.

Missing Severity Assessment for Adverse Events

If severity is missing for an AE started prior to the first study medication, then a severity of "Mild" will be assigned. If the severity is missing for an AE started on or after the first study medication dosing, then a severity of "Severe" will be assigned. The imputed values for severity assessment will be used for incidence summary, while the actual missing values will be presented in data listings.

Missing Relationship to Study Drug for Adverse Events
If the relationship to the study medication is missing for an AE started after baseline, a causality of "Related" will be assigned. The imputed values for relationship to study medication will be used for incidence summary, while the actual values will be presented in data listings.

16.5.2 Missing Date Imputation for Adverse Events

- **Incomplete Start Date**

The following imputation rules only apply to the case where the start date is incomplete (i.e., partial missing) for adverse events.
Missing day and month
- If the year is same as the year of first day on study medication, then the day and month of the start date of study medication will be assigned to the missing fields.
- If the year is not the same as the year of first day on study medication, then January 1 will be assigned to the missing fields.

Missing month only
Treat day as missing and replace both month and day according to the above procedure.
Missing day only
If the month and year are same as the year and month of first day on study medication, then the start date of study medication will be assigned to the missing day.
If the month and year are not the same as the year and month of first day on study medication, then the first day of the month will be assigned to the missing day.
If the stop date is complete and the imputed start date as above is after the stop date, the start date will be imputed by the stop date.

- **Incomplete Stop Date**

If needed, the following imputation rules apply to the case where the end date is incomplete (i.e., partially missing) for adverse events. Other partial end date will not be imputed.

Missing day and month, or Missing month only

December 31 will be assigned to the missing fields.

Missing day only

The last day of the month will be assigned to the missing day.

16.5.3 Missing Date Imputation for Prior or Concomitant Medications

For prior or concomitant medications, including rescue medications, incomplete (i.e., partial missing) start date and/or stop date will be imputed. When the start date and the stop date are both incomplete for a subject, impute the start date first.

- **Incomplete Start Date**

The following rules will be applied to impute the missing numerical fields. If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

Missing day and month

If the year of the incomplete start date is the same as the year of the first dose date of study medication, then the day and month of the first dose date will be assigned to the missing fields.

If the year of the incomplete start date is prior to the year of the first dose date of study medication, then December 31 will be assigned to the missing fields.

If the year of the incomplete start date is after the year of the first dose date of study medication, then January 1 will be assigned to the missing fields.

Missing month only

Treat day as missing and replace both month and day according to the above procedure.

Missing day only

If the month and year of the incomplete start date are the same as the month and year of the first dose date of study medication, then the day of the first dose date will be assigned to the missing day.

If either the year is before the year of the first dose date of study medication or if both years are the same but the month is before the month of the first dose date of study medication, then the last day of the month will be assigned to the missing day.

If either the year is after the year of the first dose date of study medication or if both years are the same but the month is after the month of the first dose date of study medication, then the first day of the month will be assigned to the missing day.

Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the last dose date of study medication is missing, replace it with the last visit date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

Missing day and month

If the year of the incomplete stop date is the same as the year of the last dose date of study medication, then the day and month of the last dose date will be assigned to the missing fields.

If the year of the incomplete stop date is prior to the year of the last dose date of study medication, then December 31 will be assigned to the missing fields.

If the year of the incomplete stop date is after the year of the last dose date of study medication, then January 1 will be assigned to the missing fields.

Missing month only

Treat day as missing and replace both month and day according to the above procedure.

Missing day only

If the month and year of the incomplete stop date are the same as the month and year of the last dose date of study medication, then the day of the last dose date will be assigned to the missing day.

If either the year is before the year of the last dose date of study medication or if both years are the same but the month is before the month of the last dose date of study medication, then the last day of the month will be assigned to the missing day.

If either the year is after the year of the last dose date of study medication or if both years are the same but the month is after the month of the last dose date of study medication, then the first day of the month will be assigned to the missing day.

16.5.4 Missing Date Imputation for last dose date

Imputed last dose date = earliest date of (last drug dispense date + number of days of drug dispensed, date of death, date of EOT/EOS visit, and other dates as appropriate).

16.5.5 Character Values of Clinical Laboratory Parameters

If the reported value of a clinical laboratory parameter cannot be used in a statistical summary table due to, for example, that a character string is reported for a parameter of the numerical type, coded value needs to be appropriately determined and used in the statistical analyses. However, the actual values as reported in the database will be presented in data listings.

Example for Coding of Special Character Values for Clinical Laboratory Parameters

Lab Test	Possible Lab Results (in SI unit)	Coded Value for Analysis
Chemistry: ALT	< 5	5
Chemistry: AST	< 5	5
Chemistry: Bilirubin, Total	< 2	2
Urinalysis: Glucose	= OR > 55, >= 55, > 0	Positive
	<= 0, Negative	Negative
Urinalysis: Ketones	= OR > 8.0, >=8.0, > 0	Positive
	<= 0, Negative	Negative
Urinalysis: pH	> 8.0, >= 8.0	8.0
	>= 8.5,	8.5
Urinalysis: Protein	= OR > 3.0, >=3.0, > 0	Positive
	<= 0	Negative

16.6 Appendix 6 Ranges of Potentially Clinically Significant Lab Values

Parameter	SI Unit	Lower Limit	Higher Limit
CHEMISTRY			
Alanine Aminotransferase (ALT)	U/L		≥3 * ULN
Alkaline Phosphatase	U/L		≥3 * ULN
Aspartate Aminotransferase (AST)	U/L		≥3 * ULN
GGT	U/L		≥3 * ULN
Calcium	mmol/L	<0.8*LLN	>1.2 * ULN
Creatinine	µmol/L		> 1.5x Day 1
Potassium	µmol/L	<0.75*LLN	>1.2 * UNL
Sodium	mmol/L	<0.9*LNL	>1.1 * UNL
Total Bilirubin	µmol/L		>1.5 * UNL
Total Protein	µmol/L	<0.9*LNL	>1.1 * UNL
Urea (BUN)	mmol/L		>1.5X Day 1
HEMATOLOGY			
Neutrophils	10^9/L	≤1	
Platelet Count	10^9/L	≤ 100	≥700
White Blood Cell Count	10^9/L	≤2.5	≥15
LLN: Lower limit of normal, value provided by the laboratory ULN: Upper limit of normal, value provided by the laboratory			

16.7 Appendix 7 Justification of the Non-Inferiority Margin

Non-Inferiority (NI) Clinical Trials states that "the NI trial concept depends on how much is known about the size of the treatment effect the active comparator will have in the NI study compared to no treatment" and this effect "must be assumed, based on an analysis of past studies of the control".

Pursuant to this, we sought out studies that compared ESA therapies currently approved for use in the United States to placebo (or no therapy) in the treatment of anemia. Criteria to be considered relevant to estimate the potential treatment effect included:

1- Randomized trial design
2- Prospective follow up
3- Treatment groups which included epoetin-alfa (or other recombinant erythropoeitin derivatives) and either placebo or no treatment
4- Inclusion of treatment naïve adult patients with ESRD related anemia.
5- Post randomization monitoring of hemoglobin following randomization

Studies were identified through a search of the bibliographies of peer-reviewed meta-analyses examining the effect of ESAs on outcomes.

1. Palmer et al. Meta-analysis: Erythropoiesis-Stimulating Agents in Patients with Chronic Kidney Disease. *Ann Intern Med.* 2010;153:23-33.
2. Phrommintikul A et al. "Mortality and target haemoglobin concentrations in anaemic patients with chronic kidney disease treated with erythropoietin: a meta-analysis." *The lancet* 369.9559 (2007): 381-388.
3. Koulouridis, Ioannis, et al. "Dose of erythropoiesis-stimulating agents and adverse outcomes in CKD: a metaregression analysis." *American Journal of Kidney Diseases* 61.1 (2013): 44-56.

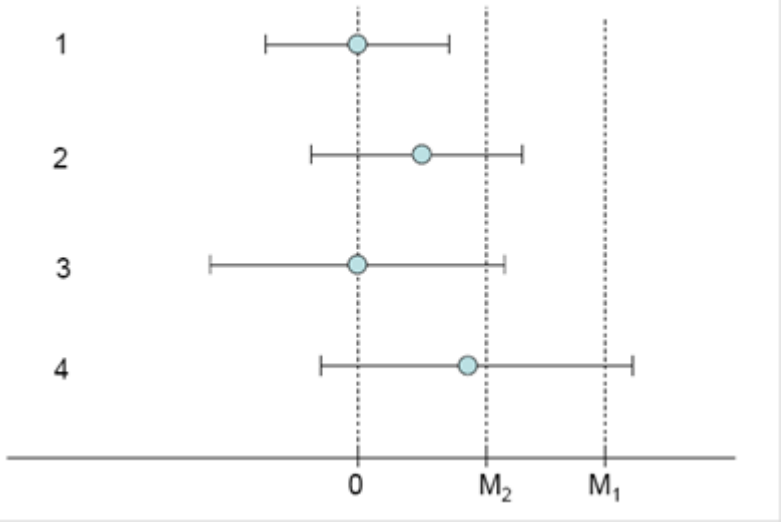
Identification of all appropriate trials was confirmed through a literature search (www.pubmed.gov) using combinations of the terms "ESA", "epoetin", "CKD", "ESRD", "placebo", and "anemia".

The approach taken for this NI margin is the fixed margin method described in the U.S. Food and Drug Administration's 2010 guidance "Non-Inferiority Clinical Trials." In this guidance, the fixed margin approach is summarized as follows:
"Briefly, in the fixed margin method, the margin M_1 is based upon estimates of the effect of the active comparator in previously conducted studies, making any needed adjustments for changes in trial circumstances. The NI margin is then pre-specified and it is usually chosen as a margin smaller than M_1 (i.e., M_2), because it is usually felt that for an important endpoint a reasonable fraction of the effect of the control should be preserved.

The NI study is successful if the results of the NI study rule out inferiority of the test drug to the control by the NI margin or more. It is referred to as a fixed margin analysis because the past studies comparing the drug with placebo are used to derive a single fixed value for M_1, even though this value is based on results of placebo-controlled trials (one or multiple trials versus placebo) that have a point estimate and confidence interval for the comparison with placebo. The value typically chosen is the lower bound of the 95% CI (although this is potentially flexible) of a placebo-controlled trial or meta-analysis of trials. This value becomes the margin M_1, after any adjustments needed for concerns about constancy.

The fixed margin M_1, or M_2 if that is chosen as the NI margin, is then used as the value to be excluded for C-T in the NI study by ensuring that the upper bound of the 95% CI for C-T is $<M_1$ (or M_2). This 95% lower bound is, in one sense, a conservative estimate of the effect size shown in the historical experience. It is recognized, however, that although we use it as a "fixed" value, it is in fact a random variable, which cannot invariably be assumed to represent the active control effect in the NI study." This approach is shown schematically in Figure 1.

Figure 1 Active Control – Test Drug differences (Point estimate, 95% CI)



1. C-T point estimate = 0 and upper bound of 95% CI < M2, indicating test drug is effective (NI demonstrated).
2. Point estimate of C-T favors C and upper bound of 95% CI < M1 but > M2, indicating effect > 0 but unacceptable loss of the control effect.
3. Point estimate of C-T is zero and upper bound of 95% CI < M1 but it is slightly greater than M2. Judgment could lead to conclusion of effectiveness.
4. C-T point estimate favors C and upper bound of 95% CI > M1, indicating there is no evidence of effectiveness for test drug."

Table 14 below displays the mean change in hemoglobin for recombinant human (EPO) based on 3 publications, [Canadian et al, 1990], [Bennett et al, 1991] and [Nissenson et al, 1995]. The 3 studies are all randomized, double-blinded, randomized, placebo controlled study to evaluate the effect EOP (epoetin alfa or beta) in anemia patient with ESRD with PD or HD.

Using the data in Table 14 below, the weighted mean of the point estimate of treatment effect in mean change from baseline for EPO (mean change in hemoglobin or hemoglobin equivalence in EPO group – mean change in placebo group) was calculated, along with its 95% confidence interval.

Table 14. Historical studies with mean change and SD in hematocrit or hemoglobin available in ESRD: EPO versus Placebo

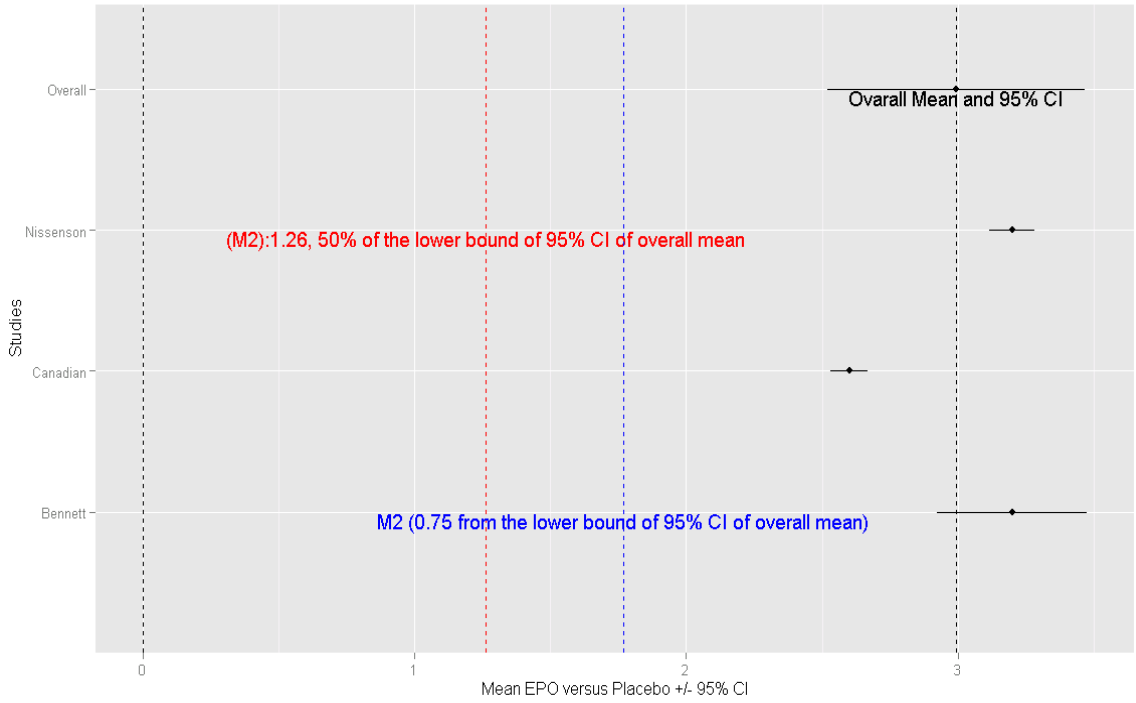
Nissenson et al, 1995 (Erythropoietin in Peritoneal Dialysis)	Endpoint	EPO Mean (SD)	Placebo Mean (SD)	EPO-Placebo Mean (Std Err)
		sample size = 78	Sample size = 74	
	Baseline in hematocrit	23.8 (3.8) %	23.8 (3.3) %	
	12 week Follow up in haematocrit	33.7 (4.8) %	24.1 (3.8) %	
	HMG equivalent Mean Change*	3.3(0.33) g/dL	0.1(0.17) g/dL	3.2 (0.042) g/dL
Bennett et al, 1991 (Epoetin Beta - Hemodialysis)		sample size = 90	sample size = 41	
	Baseline	7.1 (0.1*sqrt(90)) g/dL	6.8 (0.2*sqrt(41)) g/dL	
	12 week Follow up	11.1 (0.2*sqrt(90)) g/dL	7.6 (0.3*sqrt(41)) g/dL	
	HMG Mean Change*	4.0 (0.95) g/dL	0.8(0.64) g/dL	3.2 (0.141) g/dL
Canadian et al, 1990 (Erythropoietin - hemodialysis patients)		sample size = 40	sample size = 40	
	Baseline	7.1 (0.9) g/dL	6.9 (1.0) g/dL	
	6 month follow up	10.2 (1.0) g/dL	7.4(1.2) g/dL	
	HMG Mean Change*	3.1(0.1) g/dL	0.5 (0.2) g/dL	2.6 (0.035) g/dL

*Hematocrit in % was converted to hemoglobin in g/dL by dividing by 3. Due to the information are limited from the reference, the mean change in hemoglobin and the corresponding standard deviation were derived where the baseline and follow up hemoglobin were assumed to be independent which is very conservative.

The weighted mean point estimate for mean change from baseline in hemoglobin for EPO (mean change from baseline in hemoglobin for EPO group – mean change from baseline in hemoglobin for placebo group) was estimated to be 2.859 g/dL by treating the studies as fixed effect, with a 95% confidence interval of (2.806, 2.911) and 2.993 g/dL by treating the studies as random effect, with a 95% confidence interval of (2.520, 3.466) from meta-data analysis. The lower bound of this 95% confidence interval by treating the study as random effect, 2.520 is taken to be the NI margin M_1. To ensure that not more than 50% of the effect of EPO was lost, giving an NI margin (M_2) of 1.26.

However, a very conservative M_2 margin was chosen, 0.75, which is 29.76% of M_1, thus ensuring preservation of 70.24% of M_1 in this study. See Figure 2 for the visual effect of NI margin based on EPO group versus placebo group in mean change from baseline in hemoglobin.

Figure 2 Active Control – Placebo differences (Point estimate, 95% CI) in History Studies



16.8 Appendix 8 Temporal Profile of TEAE's of Special Interest