FibroGen, Inc.

A Phase 3, Randomized, Double-Blind, Placebo Controlled Study of the Efficacy and Safety of Roxadustat (FG-4592) for the Treatment of Anemia in Chronic Kidney Disease Patients not on Dialysis

Protocol Number: FGCL-4592-060 (Amendment 3)

STATISTICAL ANALYSIS PLAN

Version: **Final 1.0**

Release Date: **September 13, 2018**

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-4592-060 (Amendment 3).

Initiator:

Signature: _____ Date: _____

Reviewed by:

Signature: _____ Date: _____

Signature: _____ Date: _____

Signature: _____ Date: _____

Signature: _____ Date: _____

Signature: _____ Date: _____

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
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Class
BIW	Twice Weekly
BP	Blood Pressure
CHr	Reticulocyte Hemoglobin Content
CKD	Chronic Kidney Disease
CMH	Cochran-Mantel-Haenszel
CPK	Creatine Phosphokinase
CRF	Case Report Form
CRP	C-Reactive Protein
CTCAE	Common Terminology Criteria for Adverse Events
DB	Double-Blind
DSMB	Data and Safety Monitoring Board
ECG	Electrocardiogram
eGFR	estimated Glomerular Filtration Rate
EH	Excessive Hematopoiesis
EQ5D	European Quality of Life Questionnaire in 5 Dimensions

ESA	Erythropoiesis-Stimulating Agent
FACT-An	Functional Assessment of Cancer Therapy – Anemia
GGT	Gamma-Glutamyl Transferase
Hb	Hemoglobin
HDL	High-Density Lipoprotein
HRQoL	Health Related Quality of Life
ICH	International Conference on Harmonization
ICH E3	Structure and Content of Clinical Study Reports
ICH E8	General Considerations for Clinical Trials
ICH E9	Statistical Principles for Clinical Trials
IPCW	Inverse Probability Censoring Weighted
ITT	Intent-To-Treat
IV	Intravenous
LDL	Low-Density Lipoprotein
LLN	Lower Limit of Normal, value provided by the laboratory
LOCF	Last Observation Carried Forward
MAP	Mean Arterial Pressure
MAR	Missing At Random
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MNAR	Missing Not At Random
PCS	Potentially Clinically Significant

PD	Pharmacodynamics
PEY	Patient-Exposure-Year
PF	Physical Functioning subscale component of SF-36
PFY	Patient-Followup-Year
PK	Pharmacokinetics
PMM	Pattern Mixture Model
PPS	Per Protocol Set
QW	Once Weekly
RBC	Red Blood Cell
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SF-36	The 36-Item Short Form Health Survey
SQ	Subcutaneous
sTfR	Soluble Transferrin Receptor
TEAE	Treatment Emergent Adverse Event
TIBC	Total Iron Binding Capacity
TIW	Three Times Weekly
TLF	Tables, Listings, and Figures
TSAT	Transferrin Saturation
UACR	Urine Albumin-to-Creatinine Ratio
ULN	Upper Limit of Normal, value provided by the laboratory
WHODD	World Health Organization Drug Dictionary
WHO	World Health Organization

SOC	System Organ Class (used in MedDRA dictionary)

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 final study protocol. Specifications of tables, data listings, and figures (TLF) are contained in a separate document. A separate Pooled Safety SAP that describes the analyses for specific pre-specified and adjudicated cardiovascular and cerebrovascular events pooled across multiple Phase 3 studies will complement this study-specific SAP.

2 STUDY OBJECTIVES

The primary objectives of this study are to:

- Evaluate the efficacy of roxadustat for the treatment of anemia (correction and maintenance of Hb) in CKD subjects not on dialysis.
- Evaluate the safety of roxadustat administered over a minimum of 52 weeks of treatment.

The secondary objectives in this study are to:

- Evaluate the effect of roxadustat in CKD anemia on:
    - Serum lipid parameters
    - HRQoL
    - Blood pressure
    - The need for anemia rescue therapy: red blood cell (RBC) or blood transfusion, or ESA, or intravenous (IV) iron
    - Time to achieve Hb response
    - Renal function

3 STUDY DESIGN

This is a Phase 3, multi-center, randomized, double-blind, placebo-controlled study in anemic subjects with Stage 3, 4, or 5 CKD who are not on dialysis.

This study will consist of three study periods:

- **Screening Period:** Up to 6 weeks
- **Treatment Period:** Variable for individual subjects. In order to complete the Treatment Period simultaneously for all study subjects, the minimum treatment duration may be less than 52 weeks, with a maximum treatment duration of up to 3 years after the last subject is randomized.
- **Post-Treatment Follow-Up Period:** 4 weeks

3.1 IN ORIGINAL PROTOCOL AND AMENDMENT 1

Following the Screening Period, subjects are randomized to 1 of 6 treatment arms, defined by their dose frequency during maintenance (see Table 1). The randomization will result in a 2:2:2:1:1:1 ratio of subjects receiving roxadustat or placebo, respectively.

Randomization stratification factors include screening Hb values, cardiac/cerebrovascular/thromboembolic medical history, and estimated glomerular filtration rate (eGFR), and geographic region.

The Investigator, study site staff, subject, and the sponsor, are blinded to study drug, but not to dose or dosing frequency. Additionally, all efforts will be made to keep subjects blinded to study Hb values.

Study drug will be dosed TIW for anemia correction until subjects achieve a Hb value of ≥11 g/dL and ≥1 g/dL increase from baseline at two consecutive visits. Once these criteria are met, a dose frequency conversion to QW or BIW or a continuation of TIW will occur according the subject's assigned treatment arm at randomization, defining entry into the maintenance phase.

Table 1. Study Drug Dose Frequencies

Treatment Arms	Correction	Maintenance	n (Roxadustat)	n (Placebo)
1A	TIW	QW	100*	0
1P	TIW	QW	0	50*
2A	TIW	BIW	100	0
2P	TIW	BIW	0	50
3A	TIW	TIW	100	0
3P	TIW	TIW	0	50

A=Active, P= Placebo

*These numbers are based on a sample size of 450 and would increase if overall sample size increases.

3.2 UNDER AMENDMENT 2 & 3

Following the Screening Period, subjects will be randomized in a 2:1 ratio to receive roxadustat (approximately 300-800 subjects) or placebo (approximately 150- 400 subjects). Randomization is stratified by the following factors:

- Screening Hb values (≤ 8 g/dL vs > 8 g/dL)
- History of cardiovascular, cerebrovascular or thromboembolic diseases (Yes vs No)
- eGFR (< 30 mL/min/1.73 m^2 vs ≥ 30 mL/min/1.73 m^2)
- Geographic region (US vs Ex-US).

Subjects who were randomized after the implementation of Amendment 2 or those who were previously randomized but had not been changed to BIW or QW for maintenance treatment will be dosed TIW throughout the study; subjects who have already converted to BIW or QW dosing regimens as a result of being enrolled under previous Study FGCL-4592-060 protocol versions and were converted to BIW or QW during maintenance phase will continue respective maintenance dosing frequencies. Starting doses of study drug will be dose adjusted according to anemia correction guidelines in the protocol until subjects achieve a Hb value of ≥ 11 g/dL and ≥ 1 g/dL increase from baseline by central lab (single occurrence). Once these criteria are met, subjects will enter into the maintenance phase of anemia treatment and will be dose-adjusted according to maintenance guidelines.

Study drug dose will remain constant during the first 4 weeks of the Treatment Period, unless a dose reduction is required for Excessive Hematopoiesis (EH). Dose adjustments are permitted starting at Week 4, and at intervals of every 4 weeks thereafter.

4 STUDY ENDPOINTS AND ASSESSMENTS

4.1 PRIMARY EFFICACY ENDPOINT

There are two separate regionally-based primary efficacy endpoints in this study depending upon whether the data are being filed to support submission to the US FDA or to Ex-US health authorities, such as the European EMA.

The **primary endpoint for the US (FDA)** submission is defined as the change in Hb from baseline to the average level during the evaluation period (defined as Week 28 until Week 52). Hb values under the influence of rescue therapy will not be censored unless otherwise specified.

The **primary endpoint for the Ex-US** submission is defined as the proportion of subjects who achieve a Hb response at two consecutive visits at least 5 days apart during the first 24 weeks of treatment, without rescue therapy (i.e., blood/RBC transfusion, ESA, or IV iron).
A Hb response is defined, using central laboratory values, as the following:

- Hemoglobin ≥ 11 g/dL and Hb increase from baseline by ≥ 1 g/dL in subjects with baseline Hb > 8 g/dL, or
- An increase in Hb by ≥ 2 g/dL in subjects with baseline Hb ≤ 8.0 g/dL

Both scheduled and unscheduled on-treatment Hb values from the central laboratory will be taken into account.

The first date of the two consecutive visits will be used as the date of response. The second date of the two consecutive visits will be used when evaluating the presence or absence of rescue therapy.

Subjects who discontinued or received rescue therapy before or on the date of the second consecutive Hb value that fulfills the definition of response, will be classified as non-responders.

4.2 SECONDARY ENDPOINTS

The secondary efficacy endpoints in this study are:

- For Ex-US submission only: Hb maintenance: Mean change from baseline (CFB) in Hb averaged over 8 weeks of treatment at Weeks 28 to 36, without rescue therapy within 6 weeks prior to and during this 8-week evaluation period.
- Mean CFB in Hb during the evaluation period (defined as Week 28 until Week 52) in subjects with baseline CRP > ULN
- Proportion (%) of subjects with Hb Level >=10 g/dL between Week 28 to 36, without use of rescue therapy within 6 weeks prior to and during these 8-week evaluation periods.
- Hb maintenance by dose frequency: TIW, BIW, QW
- Mean CFB in LDL cholesterol averaged over Weeks 12 to 28
- Time to (and proportion of subjects who received) rescue therapy (composite of blood/RBC transfusion, ESA use, and IV iron) in the first 52 weeks of treatment
- Mean CFB in SF-36 Vitality subscore averaged over Weeks 12 to 28
- Progression of CKD: rate of change in eGFR over time adjusted by baseline eGFR, censored at dialysis or kidney transplant
- Time to (and proportion of subjects who received) blood/RBC transfusion in the first 52 weeks of treatment
- Mean CFB in the SF-36 PF subscore averaged over Weeks 12 to 28
- Blood pressure effect
    a. Mean CFB in mean arterial pressure (MAP) averaged over Weeks 20 to 28, measured as the mean of the triplicate measurements 5 minutes apart
    b. Time to (and proportion of subjects with) worsened hypertension (defined as systolic BP ≥ 170 mmHg and an increase from baseline of ≥ 20 mmHg, OR diastolic BP ≥ 110 mmHg and an increase from baseline of ≥ 15 mmHg, confirmed by the mean of triplicate measurements 5 minutes apart)

4.3 ADDITIONAL EVALUATION OF EFFICACY

Hemoglobin correction and maintenance:

- For Ex-US submission only: Mean Hb level averaged over 8 weeks of treatment at Weeks 28 to 36, without rescue therapy within 6 weeks prior to and during this 8-week evaluation period

- For Ex-US submission only: Time to first Hb response as defined by the primary endpoint for Ex-US submission
- Change from baseline in Hb at each of the selected post dosing time points: Weeks 1-4, 5-8, 9-12, 13-20, 21-28, 28-36, 44-52, 96-104
- Proportion of Hb values within 10.0 to 12.0 g/dL in weeks 28-36, 44-52, and 96-104 without use of rescue therapy within 6 weeks prior to and during these 8-week evaluation periods.
- Change from baseline in mean Hb level Weeks 28 -36 weeks in subjects with baseline hs-CRP > ULN, and in subjects with baseline hs-CRP<=ULN;
- Change from baseline in mean Hb level Weeks 28 -36 weeks in subjects who were iron-replete (ferritin >=100 ng/mL and TSAT>=20%) at baseline and in subjects who were not iron-replete at baseline (ferritin <100 ng/mL or TSAT<20%).

Hospitalization

- Hospitalization-free survival days on treatment up to Week 52, 7 days after Last Dose
- Number of days on treatment out of hospital and skilled nursing facility up to Week 52, 7 days after Last Dose
- Number of days of hospitalization per patient-year exposure (PEY)
- Number of days in hospital or skilled nursing facility per patient-year exposure (PEY)
- Number of days in medical-facility (hospital, skilled nursing facility, emergency room, or overnight observation) per patient-year exposure (PEY)
- 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

Note: Elective procedures may be excluded from analyses by using only hospitalization due to AE. Dialysis overnight stays are not counted as hospitalization.

Rescue therapy use

- Time to (and proportion of subjects who received) rescue therapy (composite of blood/RBC transfusion, ESA use, and IV iron) in the first 24 weeks of treatment, and in treatment period
- Number of blood/RBC packs per patient-month exposure to study medication

- Time to (and proportion of subjects who receive) ESA rescue therapy in the first 24 weeks of treatment, and in treatment period
- Number of ESA-week dose per PEY: 1-3 doses of epoetin alfa or beta or biosimilar thereof (in Europe) administered within 1 week = 1 ESA-week; 1 darbepoetin subcutaneous (SQ) or IV dose = 2 ESA-week; 1 Mircera IV or SQ dose = 4 ESA-week
- Time to (and proportion of subjects who receive) IV iron therapy in the first 24 weeks of treatment, and in treatment period
- IV iron(mg) per patient-month (4 weeks) exposure to study medication
- Exposure adjusted incidence rate of rescue therapy (composite, and individual component) per PEY, overall and sub-grouped by baseline Hb (<=8 g/dL, >8 g/dL)

Changes in cholesterol levels

- Change from baseline at each of the selected treatment time points (averaged over Weeks 12-28, and Week 36, 44 and 52) in:
    - Total cholesterol
    - Low-density lipoprotein/high-density lipoprotein ratio
    - Non-HDL cholesterol
- Proportion of subjects achieving LDL target of < 100 mg/dL for average LDL over Weeks 12 to 28

Blood pressure effect

- Proportion of subjects achieving blood pressure treatment goal in CKD subjects (systolic BP < 130 mmHg systolic and diastolic BP < 80 mmHg) based upon the average blood pressures over Weeks 12 to 28 of treatment

Health-Related Quality of Life (HRQoL) and European Quality of Life in Five Dimensions, Five Levels (EQ-5D-5L) benefits of anemia therapy in subjects with CKD anemia

- Mean CFB averaged over Weeks 12 to 28 of treatment in:
    - Vitality Subscale of SF-36
        - In subjects with baseline VT subscore below 50
    - Physical Functioning (PF) Subscale of SF-36
        - In subjects with baseline PF subscore below 50
    - Physical Component Scores of SF-36
        - In subjects with baseline physical component scores below 40
        - In all subjects

- Other component scores of SF-36
  - In all subjects
- Anemia Subscale ("Additional Concerns") of Functional Assessment of Cancer Therapy-Anemia (FACT-An) Scores
  - In subjects with baseline subscale scores below 55 (generally associated with fatigue.)
  - In all subjects
- Total FACT-An Scores
  - In subjects with baseline FACT-An scores below 135
  - In all subjects
- EQ-5D-5L Scores
  - In all subjects

Hepcidin, CHr, Iron Parameters, HbA1c, and CKD Progression

- Change from baseline in serum hepcidin at each scheduled time point up to Week 52
- Change from baseline in CHr at each scheduled time point up to Week 52
- Change from baseline in serum ferritin at each scheduled time point up to Week 52, in all subjects and sub-grouped by baseline values of < 100 ng/mL, 100 to 400 ng/mL, and > 400 ng/mL
- Change from baseline in TSAT at each scheduled time point, in all subjects and sub-grouped by baseline values of < 20%, 20% to 40%, and > 40%
- Change from baseline in HbA1c level at each scheduled time point in all subjects, and sub-grouped by subjects with history of diabetes and those without diabetes medical history
- Time to (and proportion of subjects requiring) renal replacement therapy (i.e., chronic dialysis (duration>30 days), kidney transplant or dropouts due to dialysis/kidney transplant), or death adjusted by baseline eGFR as a covariate
- Time to (and proportion of subjects having) doubling of serum creatinine compared to baseline, adjusted by baseline serum creatinine as a covariate, censored at dialysis or kidney transplant
- Time to (and proportion of subjects with) doubling of serum creatinine, renal replacement therapy (i.e., chronic dialysis (duration>30 days), kidney transplant or dropouts due to dialysis/kidney transplant), or death adjusted by baseline eGFR as a covariate
- Proportion of subjects with a 30%, 40% and 50% eGFR decrease, censored at dialysis or kidney transplant

- eGFR at the time of dialysis initiation (all treatment groups combined) by region and overall to show comparability among regions of nephrology practice.

4.4 SAFETY ASSESSMENTS

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

- AEs, SAEs, and clinically significant laboratory values
- Vital signs, ECG parameters, and clinical laboratory values
- Excessive hematopoiesis

Pooled safety interpretation will also be determined based on analyses of specific cardiovascular-related endpoints derived from adjudicated events pooled across multiple studies in the roxadustat Phase 3 program. The members of an independent adjudication committee blinded to treatment assignment will adjudicate the following events in multiple phase 3 studies:
All cause death, MI, stroke, congestive heart failure requiring hospitalization, unstable angina requiring hospitalization, hypertensive emergency, deep venous thrombosis, pulmonary embolism, and vascular access thrombosis.

Various region-specific pooled analyses of composites of these adjudicated events, pooled across multiple studies will be conducted. The analyses of the adjudicated events will be detailed in the region-specific pooled SAP.

5 GENERAL STATISTICAL CONSIDERATIONS

5.1 SAMPLE SIZE DETERMINATION

A total of at least 450 and up to 1200 planned subjects will be randomized in a 2:1 ratio to receive either roxadustat (at least 300, but at most 800 subjects) or placebo (at least 150, but at most 400 subjects) in a double-blind manner.

The study is sufficiently powered for both regionally-based primary efficacy endpoints. A minimum of 450 subjects are planned to be randomized to receive roxadustat or placebo (2:1 with approximately a minimum of 300 roxadustat vs 150 placebo) in a double-blind manner in order to support the primary endpoint(s) of the study.

For US (FDA) Submission Primary Efficacy Endpoint

A sample size of 450 will have > 99% power to detect a 0.75 g/dL difference in mean Hb values between the two treatment groups, assuming that the common standard deviation is 1.2 g/dL, using an analysis of variance (ANOVA) test with a 0.05 two-sided significance level.

For the comparisons of individual treatment arms for each of the 3 dose frequencies of roxadustat for Hb maintenance versus pooled placebo, a sample size of 35 roxadustat and 150 placebo will have 90% power to detect a 0.75 g/dL difference in mean Hb values between the

two treatment groups, assuming that the common standard deviation is 1.2 g/dL, using an ANOVA test with a 0.05 two-sided significance level.

For Ex-US Submission Primary Efficacy Endpoint

Based on a two-sided test at the alpha = 0.05 level of significance, 450 subjects will have > 95% power to demonstrate a statistically significant difference between roxadustat and placebo, assuming that the proportion of subjects with a Hb response in the roxadustat group is at least 65% and the proportion of subjects with a Hb response in the placebo group is at most 25%.

During the course of this Phase 3 study, which is being conducted in parallel with other Phase 3 studies, up to 1200 subjects may be enrolled in this study to support the overall safety evaluation of roxadustat across pooled multiple studies in the Phase 3 program, including adjudicated composite safety endpoints of interest. The study will stop enrollment at the Sponsor's discretion if:

1) The minimum of 450 subjects has been achieved to support the primary endpoint in this individual study, and

2) Once the total number of patients needed from this study is estimated to contribute sufficient adjudicated safety events during the planned treatment duration to accumulate across the nondialysis Phase 3 program to support pooled safety analyses of these events across multiple studies in the same CKD-nondialysis patient population in this Phase 3 program.

5.2 ANALYSIS POPULATIONS

The following analysis populations will be used for the statistical analysis:

5.2.1 Intent-to-Treat Population (ITT)

The ITT population will include all randomized/enrolled subjects. All efficacy data will be analyzed using the ITT population for US FDA. If actual treatment received differs from the randomized treatment arm, the randomized treatment arm will be used for analysis for the ITT.

ITT1 is a subset of ITT which includes all subjects randomized/enrolled prior to Amendment 2. ITT2 is a subset of ITT which includes all randomized/enrolled subjects under Amendment 2 or later.

5.2.2 Safety Population

The Safety population will include all subjects who took any dose of study medication. All safety data will be analyzed using the safety population. If actual treatment received differs from the randomized treatment arm, the actual treatment arm will be used for analysis for the Safety population.

5.2.3 Full Analysis Set (FAS)

The FAS consists of all randomized/enrolled subjects who received at least one dose of study drug and have baseline and at least one post-dose Hb assessment. If actual treatment received differs from the randomized treatment arm, the randomized treatment arm will be used for analysis for the FAS except for analyses by maintenance dosing frequency which will be presented as-treated. This analysis set is primarily used for EX-US submissions.

FAS1 is a subset of FAS subjects who are randomized/enrolled prior to Amendment 2. FAS2 is a subset of FAS subjects who are randomized/enrolled under Amendment 2 or later.

5.2.4 Per Protocol Set (PPS)

All randomized/enrolled subjects who receive at least 8 weeks of study treatment, have valid corresponding Hb measurements, and are without major protocol violations. The PPS population will be utilized for supportive analysis of primary endpoints.

PPS1 is a subset of PPS subjects who are randomized/enrolled prior to Amendment 2. PPS2 is a subset of PPS subjects who are randomized/enrolled under Amendment 2 or later.

All primary and secondary efficacy analyses will be based on the ITT (for US FDA) or FAS (for EMA) population as the primary analysis set respectively. Primary and secondary analyses will be repeated using the FAS or ITT respectively as supportive sensitivity analyses. Additional efficacy analyses will use the FAS as specified in the TLF (Tables, Listings and Figures) index.

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 presented by frequency count and percentage.

Lab results obtained from the central laboratory, rather than local laboratories, will be used for all efficacy and safety analyses. Local laboratory values, if collected in the CRF's, will be listed.

The treatment groups will be presented as follows:

Roxadustat QW(1A), Roxadustat BIW(2A), Roxadustat TIW(3A), Pooled Roxadustat (1A+2A+3A), Pooled Placebo (1P+2P+3P). Most safety and efficacy tables will be presented by Pooled Roxadustat vs. Pooled Placebo. Selected safety and efficacy tables will present randomized treatment arms of QW, BIW and TIW if the patient received the actual maintenance dosing frequency for QW, BIW and TIW. In these tables, Roxadustat subjects who were randomized to BIW or QW under Amendment 1 but never switched (e.g., due to Amendment 2) to that maintenance frequency will not be presented in BIW or QW group, but in the TIW and pooled Roxadustat groups.

The statistical comparisons will be done between roxadustat against Placebo as deemed appropriate, where the groups before amendment 2 are pooled

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 primary and 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.
- Hb baseline is defined as the mean of the last four available values obtained prior to the first dose, including the Day 1 pre-dose value.
- eGFR and urine albumin-to-creatinine ratio (UACR) baseline will use the mean of all available central lab values prior to the first dose of study drug, including the Day 1 pre-dose value.
- Baselines for reticulocyte count, reticulocyte hemoglobin content (CHr), hepcidin, hs-CRP, serum iron parameters (transferrin, TIBC, TSAT, Ferritin, sTfR, and iron), blood pressures and heart rate are defined as the mean of values obtained within two weeks prior to the first dose.
- Randomization stratification factors derived from actual data (not the ones from the randomization system) will be used in all applicable analysis models. A table will show the discrepancies between the two.
- Analysis visits (instead of the nominal visits from CRF) derived from visit dates and visit time windows will be used in all analyses, tables, listings and figures. Unscheduled visits within an allowable window will be grouped into the closest scheduled visits based on the visit window specified in Appendix 1. For subjects who have more than one measurement at a scheduled visit, the last measurement 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 also 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 – first dose date +1
- Body weight, height and temperatures will converted using the following formula:

  - kg = lb/2.2
  - cm = 2.54 x in
  - $C^o = (5/9) \times (F^o - 32)$

- eGFR will be calculated using the following Modification of Diet in Renal Disease (MDRD) equation:

  eGFR (in mL/min per $1.73m^2$) = 175 x $(S_{Cr}$ in mg/dL$)^{-1.154}$ x (Age in years$)^{-0.203}$ x (0.742 if female) x (1.212 if African American)

  where S_{Cr} = serum creatinine concentration

- Mean Arterial Pressure (MAP) will be derived using the following equation:

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

- For continuous variables that are recorded as "<X" or ">X", the value of "X" will be used in the calculation of summary statistics.
- 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 0.0001before rounding.
- All tables and listings will have a header showing "FibroGen, Inc.", the protocol number, date of the database transfer, and Page x of y. A footer will show the program file path/name, output file path/name, run date and run time.
- Additional data handling conventions are detailed in Appendix 1.

6 SUBJECT ACCOUNTABILITY AND DISPOSITION

The number of subjects included in each study population (i.e., ITT, Safety, FAS, and PPS) will be summarized by treatment group. The table will be repeated for subjects randomized/enrolled under Amendment 2 or later.

The number and percentage of subjects completed (through treatment and through end of study) and prematurely discontinued during the double-blind period and overall (including pre-dose and follow-up periods) will be presented for each treatment group and overall for the randomized population. Reasons for premature discontinuation from treatment as recorded on the disposition CRF will be summarized (number and percentage) by treatment group for the randomized population. Percentage of premature study drug discontinuations will be compared between treatment groups (roxadustat overall vs. placebo) using *Fisher's Exact Test*.

All subjects who prematurely discontinued from treatment or study will be listed by discontinuation reason and point of discontinuation (pre-dose, double blind treatment, or follow-up period) for the randomized population.

7 PROTOCOL DEVIATIONS

Important protocol deviations of interest may include, but are not limited to, the following:

- Subjects who did not meet inclusion/exclusion criteria
- Subjects who received disallowed concomitant medications or non-drug therapy
- Subjects who missed more than 25% doses of prescribed study medication during the overall treatment period.

The number and percentage of important protocol deviations will be categorized and tabulated as appropriate. A subset of pre-specified major protocol deviations will exclude some patients in the PPS analyses. These will be identified prior to database lock and unblinding. Considerations will be given according to the following table.

Table 2. 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 1 week
3	Study drug compliance < 75% (up to Week 52)
4	Administration of prohibited concomitant medication that may impact evaluation of efficacy of the study drug*
5	Significant noncompliance with study procedures that may impact evaluation of efficacy of the study drug will be evaluated case by case*

* Subject to Medical Monitor's decision

8 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographic parameters and important baseline and disease characteristics including randomization stratification factors will be summarized by treatment group for the ITT, ITT1, ITT2, Safety, FAS, FAS1, FAS2, and PPS populations. These include but may not be limited to age and age group, sex, ethnicity, race, weight, body-mass index (BMI), Hemoglobin, Ferritin and Ferritin group (<=100 vs. >100 ng/mL), TSAT and TSAT group (<=20% vs. >20%), iron repletion status ([TSAT $\geq$20% and ferritin $\geq$100 ng/mL] vs. others), baseline C-reactive protein (CRP) group (CRP $\leq$ULN vs. CRP > ULN), eGFR and eGFR group (<10, 10-<15, 15-<30, 30-<45, 45-<60, and >=60 mL/min/1.73m^2).

Descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) will be presented for continuous variables. In addition, 25%-75% values of Hb 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.

Comparability of baseline characteristics among treatment groups will be tested using analysis of variance (ANOVA) model for continuous variables and Chi-Square test for categorical variables for guidance only.

9 MEDICAL HISTORY

Medical History of interest including but not limited to Chronic Kidney Disease (CKD) History inclusive of CKD Cause, Cardiovascular Disease, Cerebrovascular Disease, Thrombosis History, Diabetes History, and Anemia History will be summarized by treatment group for the Safety Population.

10 STUDY MEDICATION

10.1 EXTENT OF EXPOSURE

Exposure to study medication will be summarized in terms of treatment duration, which is calculated as the number of days from the date of first double-blind medication taken to the date of last dose taken, inclusively.

Total weekly study drug exposure is defined as the total prescribed dose (in mg or mg/kg) of study drug administered within the week (per CRF dosing week).

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

Per dose amount and dosing frequency for every four-week period will also be tabulated by treatment group for the Safety population.

Patient-Exposure-Year (PEY) is defined as (Last Dose Date – First Dose Date + 1)/365.25.
Patient-Followup-Year (PFY) is defined as (Last Dose Date – First Dose Date + 28)/365.25.

10.2 DOSE ADMINISTRATION

Summary of weekly dose (in mg and mg/kg) immediately prior to Hb correction target achieved, weekly dose and dose per intake after Hb correction target achieved (in the 1st, 4th, 8th, 12th weeks, and Weeks 49-52) will be tabulated. Number of dose adjustments by type and overall before and after Hb correction(within 8 weeks and throughout treatment period up to Week 52) will be summarized.

10.3 TREATMENT COMPLIANCE

Study medication compliance for a specified period is defined as the total dose (mg) actually taken by a patient during that period divided by the prescribed dose expected to be taken during the same period multiplied by 100. Descriptive statistics for study medication compliance will be presented by treatment group for the whole double-blind period of the study.

11 PRIOR AND CONCOMITANT MEDICATIONS

The World Health Organization Drug Dictionary (WHODD) Version 15 will be used to classify prior and concomitant medications by therapeutic class and preferred term 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. Concomitant medication is defined as any medication taken between the day of first dose of the study medication and the day of last study medication date + 28 days, inclusive.

Medication start and end dates will be compared with the start date of study drug and classified as per Table 3. In case of partial or missing dates, comparisons will be made based on the level of details available. For example, if start date of study drug infusion 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 3. Classification of Prior and Concomitant Medications

Start Date \ End 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	Prior	(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 receiving each drug within each therapeutic class using the safety population. Multiple drug usage by a patient will be counted only once.

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

12 EFFICACY ANALYSES

Hemoglobin results obtained from the central laboratory will be used for all efficacy analyses.

Baseline Hb is defined as the mean of the last four central laboratory Hb values prior to the first dose of study treatment. Hemoglobin values after a rescue therapy will not be excluded for the US (FDA) submission primary efficacy analysis. Hemoglobin values within 6 weeks after a rescue therapy will be excluded for the Ex-US submission primary efficacy endpoint as well as some sensitivity efficacy analyses.

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.

12.1 PRIMARY ANALYSIS

For US (FDA) Submission

The primary US FDA efficacy endpoint will be analyzed using the ITT population as the primary analysis population.

The primary efficacy endpoint for the US (FDA) submission is the Hb CFB to the average level during the evaluation period, defined as Weeks 28 to 52.

The rationale for choosing the specified evaluation period as the primary endpoint for US (FDA) submission is that one would expect to have achieved Hb correction by Week 28, and averaging it to Week 52 would be an accurate reflection of a longer term therapeutic effect of roxadustat.

The hypothesis to be tested for the primary efficacy analysis is:

H_0: Hb CFB to the average of Weeks 28 to 52 in the roxadustat group = Hb CFB to the average of Weeks 28 to 52 in the placebo group
Versus:
H_1: Hb CFB to the average of Weeks 28 to 52 in the roxadustat group ≠ Hb CFB to the average of Weeks 28 to 52 in the placebo group

The efficacy endpoint will be defined as each patient's Hb change from baseline to the average level during the evaluation period, defined as Week 28 to 52 of the treatment period. A Multiple Imputation Analysis of Covariance model (MI ANCOVA) will be used to test the primary hypothesis. The model will contain terms for baseline Hb and baseline eGFR measurements as covariates and treatment arm and the other randomization stratification factors except screening Hb (≤8 vs >8) and eGFR factor as fixed effects.

The MI ANCOVA model will be used to compare the pooled roxadustat and pooled placebo groups: Pooled roxadustat (TIW+BIW+QW) vs. pooled placebo. Efficacy by maintenance dosing frequency (i.e., TIW, BIW, QW vs. pooled placebo) will also be evaluated using FAS.

H_0 will be tested at the two-sided alpha = 0.05 level of significance and will be rejected if the $p < 0.05$ from the test.

The following steps will be used to conduct the primary analysis.

1. Generate 200 datasets, using seed 126345, where 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 with treatment, baseline hemoglobin, randomization stratification factors, and the available non missing hemoglobin for each scheduled Week.

   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, or a monotone missing data pattern.

   Sample SAS code:

```
PROC MI data=xx out=xx seed=xx nimpute=200;
  by treatment;
  var treatment covariates Week1 … Week 52;
  mcmc impute=mono;
run;
```

2. For each dataset from step 1, missing ending data (hemoglobin up through end of evaluation period) will be imputed using seed 472794. 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 subject 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 and the stratification factors.

3. Analyze each imputed dataset using the analysis of covariance model (ANCOVA) using the mean of all observed or imputed Hb values within the evaluation period. The model will contain terms for baseline Hb and eGFR measurements as covariates and treatment arm and the other randomization stratification factors except baseline Hb (≤8 vs >8) and eGFR 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 of each of the 200 ANCOVA runs 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 placebo group) and the corresponding p-values and 95% CIs during the evaluation period.

For Ex-US Submission

The primary EX-US efficacy endpoint will be analyzed using the FAS population as the primary analysis population.

The primary efficacy endpoint for the Ex-US submission of proportion of responders will be compared using Cochran-Mantel-Haenszel (CMH) adjusting for the randomization stratification factors comparing the pooled roxadustat (TIW+BIW+QW) to pooled placebo (TIW+BIW+QW). The hypothesis to be tested for the primary efficacy analysis is:

> H_0: Hb responder rate in the roxadustat group = Hb responder rate in Placebo group
> Versus:
> H_1: Hb responder rate in the roxadustat group ≠ Hb responder rate in Placebo.

H_0 will be tested at the two-sided alpha = 0.05 level of significance and will be rejected if the $p < 0.05$ from the test.

The SAS procedure will be similar to the following:

```
proc freq;
  tables covariates*Treatment*Response / cmh;
run;
```

The 95% CIs based on CMH adjusted odds ratio will be reported. In addition, the 95% CIs of the responder rate and rate difference between treatment groups based on the exact method of Clopper-Pearson will be calculated and presented.

12.2 SENSITIVITY ANALYSES OF THE PRIMARY ENDPOINT

For US (FDA) Submission

As sensitivity analyses, the primary analysis of the U.S. (FDA) primary endpoint will be repeated using the ITT2, PPS, and FAS.

For Ex-US Submission

As sensitivity analyses, the primary analysis of the primary endpoint for Ex-U.S. will be repeated using the FAS2, PPS, and ITT.

In addition, the impact of missing data on the analysis of the U.S. primary efficacy endpoint will be examined using the following additional sensitivity analyses as in the ITT population summarized in Table 4. These sensitivity analyses are further detailed in Sections 12.2.1 to 12.2.3.

Table 4. U.S. Primary Endpoint Analysis Results

Hb Value Censoring due to Rescue	**Analysis ITT Population**	**Treatment Difference and 95% CI (pooled roxadustat vs pooled placebo)**	**Std. Err.**	**Degree of Freedom***	**t-statistics**	**p-value**
No censoring	ANCOVA with Multiple Imputations	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
	MMRM	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
	PMM-Last Mean Carried Forward	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
	PMM-Baseline Carried Forward (Roxa. Only)	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
	PMM-Baseline Carried Forward (Both Groups)	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
	PMM-Jump to Control	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
Set Hb to missing for 6 weeks due to Rescue Medications	ANCOVA with Multiple Imputations	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx
	MMRM	.xx (.xx, .xx)	.xxx	xxx	.xxx	.xxxx

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

12.2.1 Change in Hb - MMRM Model

As sensitivity analysis, a mixed model of repeated measures (MMRM) approach will be used. The model will contain terms for treatment arm, baseline Hb and eGFR measurements, visit, treatment*visit, and the other randomization stratification except baseline Hb (≤8 vs >8) and eGFR as fixed effects. The primary efficacy analysis will be based on the estimated difference in the overall mean effect between the two treatment groups throughout the Evaluation Period based on the MMRM model.

The MMRM model will be used to compare the pooled roxadustat vs. pooled placebo. H_0 will be tested at the two-sided alpha = 0.05 level of significance and will be rejected if the $p < 0.05$ from the test.

Due to the large amount of visits (up to week 52) to include 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 instead. If this second model also does not converge, then the (homogeneous) Toeplitz structure will be tried, and finally compound symmetry will be used as the covariance structure to achieve convergence. If none of them converge, first order autoregressive (AR (1)) as a covariance structure will be used to achieve convergence.

12.2.2 Change in Hb - Pattern Mixture Model (PMM)

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).

The following aspects of missing data may affect the estimates.

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

12.2.2.1 Timing and Extent of Missing Data

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

- Full data cases are defined as subjects without non-missing hemoglobin for all scheduled weeks of the evaluation period.

- Missing data cases are defined as subjects with a missing hemoglobin value (s) on at least one scheduled assessment in of the evaluation 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 value for at least one scheduled assessment in the evaluation period but not on consecutive scheduled assessments up to the end of evaluation period.

- Monotone missing hemoglobin cases are defined as subjects who have missing Hb values in consecutive scheduled Hb assessments 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 (e.g., <xx%), then those cases will be combined so that the groups are full data cases and missing data cases. The summary of missing pattern in evaluation period will be presented in a table.

12.2.2.2 Assumptions on Missing Data Mechanism

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

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

12.2.2.3 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. Parameter below refers to the parameter of the multi-normal distribution for baseline and post baseline Hb measurements.

1. Create posterior distribution of parameter: Separately for each treatment arm, take all patients observed data and assuming Missing At Random (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. 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 given observed data: For each patient with missing data, use their joint distribution in previous step to construct their conditional distribution of missing data given observed data. Sample the missing data from this conditional distribution, to create a `completed' data set.

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.

12.2.2.4 PMM – Baseline Carried Forward (Roxadustat Only and Both Groups)

The analysis is the same as PMM – Last Mean Carried Forward except imputing the missing data. The imputation data will be generated similarly as last mean carried forward method described above but instead of using post-baseline observed data, only baseline data will be used. The similar analyses will be conduct in two scenarios.

- The baseline carried forward imputation will be performed for the roxadustat treatment group only, while for active control group, the imputation data will be generated using the last mean carried forward described above.

- The baseline carried forward imputation will be performed for the both treatment groups.

Similarly, the Rubin's method will be then used to combine the estimates and the differences between the least square mean differences between the two treatment groups from each of the ANCOVA analysis. This sensitivity analysis will be performed for U.S. primary endpoints only.

12.2.2.5 PMM - Jump to Control

A pattern-mixture model using a last mean carried forward jump to control multiple imputation method (Carpenter et al, 2013) will also be used as another sensitivity analysis similar to the PMM-Last Mean Carried Forward except for step 3, where the joint distribution of the patient's

observed and missing data are considered to be multivariate normal with mean and covariance matrix from the control (placebo) treatment arm.

12.2.3 Hb Censoring After Rescue Therapy

Because the rescue use may potentially confound the effect of roxadustat, analysis will also be performed where Hb values will be censored for 6 weeks (or 8 weeks if during an evaluation period) after rescue therapy use to examine its potential impact on the robustness of the efficacy findings.

All sensitivity analyses for Change in Hb described in the above section will be applied to these scenarios using the ITT Population.

12.2.4 Hb Responder: Logistic Regression

The primary analysis for EX-US endpoint will also be repeated using logistic regression instead of Cochran–Mantel–Haenszel (CMH) test and including in this model the randomization stratification factors excluding baseline Hb and eGFR. Baseline Hb and baseline eGFR will be included as continuous covariates instead of as class factors. The odds ratio (pooled roxadustat versus pooled placebo) and its 95% confidence interval will be provided.

The SAS procedure will be similar to the following:

```
proc logistic;
  class treatment region cv_history;
  model response = treatment region cv_history baseline_Hb baseline_eGFR;
run;
```

This model will also be repeated using logistic regression and adjusting by sex and age, in addition to the factors mentioned above.

The analyses in this section will use the FAS population.

12.3 ANALYSES OF SECONDARY EFFICACY ENDPOINTS

The secondary efficacy endpoints will be analyzed using the FAS population as the primary analysis population for superiority, and PPS the primary analysis set for non-inferiority.

Secondary endpoints will be tested using a fixed sequence approach to adjust for multiple endpoints comparing between pooled roxadustat vs. pooled placebo, unless otherwise specified. If a null hypothesis is rejected, the claim of roxadustat over placebo will be declared successful and the test will progress to the next comparison in sequence.

1. Ex-US submission only: Mean CFB in Hb averaged over 8 weeks of treatment at Weeks 28 to 36, without rescue therapy within 6 weeks prior to and during this 8-week evaluation period for the two treatment groups will be compared using an MMRM model with baseline Hb and eGFR as covariates and treatment group, visit, treatment*visit, and the other

randomization stratification factors as fixed effects. The same strategy as that used in MMRM for primary endpoint 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 Hb CFB averaged over Weeks 28 to 36 between the two treatment groups will be generated from an estimate statement for Weeks 28 to 36. Superiority will be declared if the lower bound of the 2-sided 95% confidence interval of the difference between roxadustat and placebo exceeds 0. Missing values will be handled by the MMRM model.

2. Mean CFB in Hb averaged over Weeks 28 to 52 in a subgroup of subjects with baseline CRP>ULN for the two treatment groups will be compared using MI ANCOVA model with baseline Hb as a covariate and treatment group, visit, treatment*visit, and the above-mentioned stratification factors as fixed effects. Data up to visit of Week 52 will be included in the model. The estimates for difference of Hb averaged over Weeks 28 to 52 between the two treatment groups will be generated from an estimate statement from Visit Week 28 to 52. Superiority will be declared if the lower bound of the 2-sided 95% confidence interval of the difference between roxadustat and placebo exceeds 0.
3. Proportion of subjects with Hb (averaged Weeks 28 to 36) >=10 g/dL, without use of rescue therapy within 6 weeks prior to and during this 8-week evaluation period, will be compared using Cochran-Mantel-Haenszel (CMH) adjusting for the randomization stratification factors.
4. To assess the effect of maintenance dosing frequencies, pairwise comparisons of roxadustat TIW (3A), BIW and QW (2A and 1A from the original and amendment 1 versions of the protocol) vs. pooled placebo (1P+2P+3P) will be performed in subjects treated with roxadustat who have achieved Hb response (achieved Hb >=11 g/dL and Hb>=1 g/dL at two consecutive visits) within the first 24 weeks of treatment, and were treated in the maintenance phase at dose frequencies of TIW, BIW, or QW after Hb response. The Hb levels of these patients will be analyzed by dose frequency during Hb maintenance phase vs pooled placebo for the following two endpoints in a fixed sequence:
    - Average Hb level over Weeks 28 to 36 in the roxadustat Hb maintenance dose frequency subgroup (e.g., TIW) vs pooled placebo using the same method for the US primary endpoint.
    - Proportion of subjects with Hb (averaged over Weeks 28 to 36) >=10 g/dL in the roxadustat Hb maintenance dose frequency subgroup (e.g., TIW) vs. pooled placebo using the same method for the ex-US primary endpoint.
        1) Roxadustat TIW vs. pooled placebo,
        2) Roxadustat BIW vs. pooled placebo,
        3) Roxadustat QW vs. pooled placebo.
5. Mean CFB in LDL cholesterol averaged over Weeks 12 to 28 for the two treatment groups will be compared using MMRM models with baseline LDL cholesterol as a covariate and treatment group, visit, treatment*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 treatment groups will be generated from an estimate statement from Visit Week 12 to 28. Superiority will be declared if the lower bound of the 2-sided 95% confidence interval of the difference between roxadustat and placebo exceeds 0.

6. Time to rescue therapy (composite of blood/RBC transfusion, ESA use, and IV iron) in the first 52 weeks of treatment will be compared using Inverse Probability of Censoring Weighted (IPCW) model detailed in Appendix A or Cox Proportional Hazards model as deemed appropriate. The Cox regression model will adjust for baseline Hb, baseline eGFR and other randomization stratification factors. Hazard ratio and its associated 95% CI will be computed between the roxadustat group versus pooled placebo. Superiority will be declared if the upper bound of the 2-sided 95% CI of the hazard ratio does not exceed one. Proportion of subjects who received rescue therapy over the treatment period will be derived using Kaplan-Meier approach.
7. Mean CFB in the SF-36 Vitality subscore averaged over Weeks 12 to 28 for the two treatment groups will be compared using an MMRM model as outlined above. A subgroup of subjects with baseline vitality subscore below 50 will be analyzed as a sensitivity analysis.
8. Progression of CKD measured by rate of change in eGFR over time adjusted by baseline eGFR, censored at chronic dialysis or kidney transplant will be analyzed using random slope and intercept model. Annualized eGFR slope over time is estimated by a random slope and intercept model using all available eGFR values (one baseline and all post-treatment values up to End of Treatment Period or start of chronic dialysis or kidney transplant) adjusted by Baseline Hb, Region, CV history at Baseline and the interaction terms Baseline eGFR by time point and Baseline Hb by time point). All assessments collected after initiation of chronic dialysis or kidney transplant will be excluded from the analysis.
9. Time to blood/RBC transfusion in the first 52 weeks of treatment will be compared using Inverse Probability of Censoring Weighted (IPCW) model detailed in Appendix A or Cox Proportional Hazards model as deemed appropriate. The Cox regression model will adjust for baseline Hb, baseline eGFR and other randomization stratification factors. Hazard ratio and its associated 95% CI will be computed between the roxadustat group versus pooled placebo. Superiority will be declared if the upper bound of the 2-sided 95% CI of the hazard ratio does not exceed one. Proportion of subjects who received blood/RBC transfusion over the treatment period will be derived using Kaplan-Meier approach.
10. Mean CFB in the SF-36 PF subscore averaged over Weeks 12 to 28 for the two treatment groups will be compared using an MMRM model as outlined above. A subgroup of subjects with baseline PF subscore below 50 will be analyzed as a sensitivity analysis.
11. Mean CFB in MAP averaged over Weeks 20 to 28 for the two treatment groups will be compared using an MMRM model as outlined above with a non-inferiority margin of 5.

12. Time to (and proportion of subjects with) protocol defined worsened hypertension for the pooled roxadustat vs pooled placebo will be compared using the approach in #4 above. Hazard ratio and its associated 95% CI will be computed between the roxadustat group versus pooled placebo. Non-inferiority will be declared if the upper bound of the 2-sided 95% CIs of the odds ratio does not exceed 1.3. Superiority will be declared if the upper bound of the 2-sided 95% CI of the hazard ratio does not exceed one. The incidence rate per PEY by treatment group will also be reported.

In MMRM above, due to large amount of visits to include 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 instead. If this second model also does not converge, then the (homogeneous) Toeplitz structure will be tried, and finally compound symmetry as a covariance structure to achieve convergence. If none of them converge, first order autoregressive (AR (1)) as a covariance structure will be used to achieve convergence.

12.4 ADDITIONAL EFFICACY ANALYSES

The additional efficacy analyses will use the FAS population.

Additional efficacy analyses of continuous endpoints will use an MMRM model using baseline Hb and the parameter value as covariates and the other randomization stratification factors except baseline Hb (≤8 vs >8), as fixed effects as outlined above. An ANCOVA model will be used for number of days hospitalized per PEY, rescue therapy used per PEY, etc. where MMRM model is not appropriate.

Additional analyses of proportions will use the CMH model adjusting for the randomization stratification factors. A CMH adjusted odds ratio and its associated 95% CI will be computed to compare between the roxadustat TIW group and pooled placebo. A two-sided 95% CI of the proportion based on the exact method of Clopper-Pearson will be computed for each treatment group.

Additional efficacy analyses of time-to-event endpoints will use the Cox Regression model. Median time to event and associated 95% CI will be estimated using Kaplan-Meier method.

Proportion of Hb values within 10.0-12.0 g/dL will be calculated in two ways:

<u>Proportion of Hb values:</u>

The following proportion in percentage for each subject will be defined:

(*Number of Hb values within 10.0-12.0 g/dL / Total number of protocol-planned Hb values)* * *100* in weeks 28 to 36, 44 to 52 and 96 to 104 without use of rescue therapy within 6 weeks prior to and during this 8 week evaluation period. All scheduled and unscheduled hemoglobin

values that belong to each period will be taken into account using the analysis windows defined in Table 10.3‑1.

Percentage of time:

The percentage of time each patient has Hb values within 10.0-12.0 g/dL will be calculated (as a percentage of the total of the length of time between the first and last Hb assessment during the evaluated period). The percentage of time will be calculated via linear interpolation. That is, if the change in Hb category (for instance from within 10.0-12.0 g/dL to > 12.0 g/dL) occurs between two visits V0 and V1, the day of change will be calculated by:

$$x = x_0 + (y - y_0)\frac{(x_1 - x_0)}{(y_1 - y_0)}$$

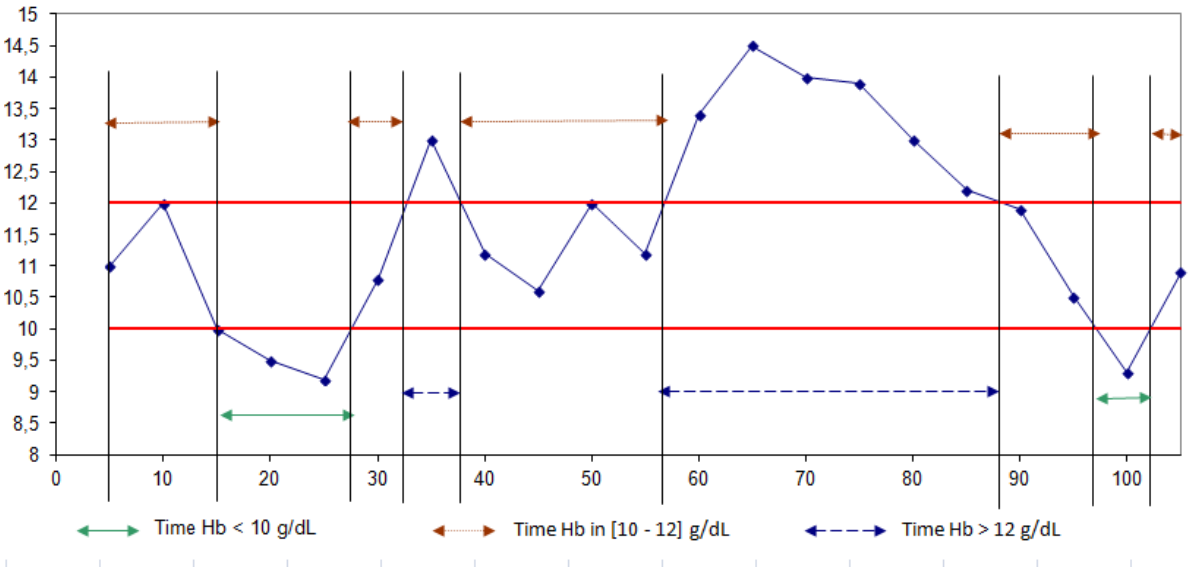
Where x_1 and x_0 are the dates when Hb was measured at V0 and V1 respectively,

y_0 and y_1 are the Hb value at the respective visits V0 and V1 and y is the level of the Hb

boundary (i.e. 12.0, 13.0 or 14.0g/dL).

Figure 11 shows visually how the linear interpolation will calculate the total number of days that a subject is in each Hb category for an example subject:

Figure 1: Example of Linear Extrapolation



In case that several Hb values are on the same day the average of these values will be used to represent the Hb of that day in the above formula. This calculation will provide the day that the change in Hb value occurs. The number of days that the Hb value has been in each category will

be determined and the percentage calculated based on the length of time between the first and last Hb assessment during the evaluated period, i.e.:

Date of Last Hb assessment during the evaluated period – Date of first assessment during the evaluated period.

No imputation will be performed if no Hb value is available in relevant time windows.

Rate of change in eGFR over time will be analyzed using the random effect model adjusted by baseline eGFR.

To evaluate Hb maintenance by other dosing frequencies, the following summary will be provided for the 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

13 SAFETY ANALYSES

The safety analysis will be performed using the Safety Population. Safety parameters include adverse and serious adverse events, laboratory parameters, vital signs, and ECG parameters. For each safety parameter, unless otherwise specified in section 5.4, the last assessment made prior to the first dose of double-blind study medication will be used as the baseline for all analyses.

13.1 ADVERSE EVENTS

Adverse events will be coded using MedDRA.

13.1.1 Proportion of Subjects with TEAE

An AE (classified by preferred term) occurring after the first dose of study medication and up to 28 days after the last dose of study medication 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 double-blind treatment period. An AE that occurs more than 28 days after the last dose of study medication will not be counted as a TEAE. TEAEs up to 3 days and 7 days after the last dose of study medication may also be evaluated.

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 (>=3 in a separate table); and by system organ class, preferred term, and Investigator-determined relationship (at least possibly related in a separate table) to study medication. If more than one event occurs with the same preferred term for the same patient, the patient 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 number and percentage of subjects reporting common (≥ 5% of subjects in any treatment group) TEAEs, treatment-emergent serious AEs (TESAE), renal TESAEs and AEs leading to discontinuation of study medication will be summarized by SOC, preferred term and treatment group, sorted in decreasing frequency overall. In addition, the number and percentage of death and fatal SAEs (i.e., events that caused death) will be summarized by treatment group, SOC and preferred term.

Sensitivity analysis of TEAE during study medication treatment will consist of a summary of TEAE and TESAEs reported after the first dose of study medication up until 72 hours(to cover > 4 half-lives of roxadustat post treatment) and 7 days after the last dose of study medication.

Treatment emergent major cardiovascular events (death, stroke, myocardial infarction) stratified by baseline eGFR (<10, 10-<15, 15-<30, and >=30 mL/min/1.73 m^2) will be tabulated.

Listings will be presented of subjects with serious adverse events (SAEs), subjects with adverse events leading to discontinuation, and subjects who died.

13.1.2 Exposure Adjusted Incidence Rate

Since the number and percentage of subjects with adverse events might be influenced by differential discontinuation rates in the treatment arms, in addition to the frequency tables above, the exposure adjusted incidence rate EAIR (per 100 patient years at risk) and the cumulative incidence at 6 months, 12, 18 months & 24 months with the 95% confidence interval, using Greenwood's formula, will be reported by treatment arm (pooled roxadustat and pooled placebo) for each of the following event types of special interest: serious TEAEs, renal SAEs, dialysis, deaths, Investigator-determined related serious TEAEs, AEs leading to discontinuation of study drug, TEAE that are severe and/or NCI CTC Grades 3 or higher.

The exposure adjusted incidence rate (per 100 patient years at risk) will be calculated as follows:

$$\frac{\text{Number of subjects with event}}{\text{Total cumulative time at risk (years)}} \times 100$$

Where Total cumulative time at risk is the sum of individual time at risk.

The Time at Risk for a subject with the event of interest will be calculated (in years) as:

(First event date – analysis Date of first dose + 1) / 365.25

Where 'Analysis Date of first dose' is defined as the date of initial dose of the study drug.

The Time at Risk for a subject without the event of interest is calculated as

(Last day of contact – Date of first dose + 1) / 365.25

Risk of each of the TEAE categories of special interest above will be compared (pooled roxadustat vs. pooled placebo) using Cox Regression model. Hazard ratio and its 95% will be calculated for the frequency of pooled roxadustat as relative to pooled placebo for each of the TEAE categories of special interest.

A dot-and-forest plot will be produced showing the above TEAE categories in the y-axis and the incidence rates (pooled roxadustat and pooled placebo), the Cox proportional hazard ratio (pooled roxadustat vs. pooled placebo) and its 95% CI in the x-axis. The system organ classes in the y-axis will be sorted by the hazard ratio.

In addition, cumulative incidence plots for subjects experiencing each of the TEAE categories above will be produced by treatment arm and combined group (pooled roxadustat and pooled placebo). Temporal profile of TEAEs of special interest may also be plotted by combined treatment group (pooled roxadustat vs. pooled placebo) showing the subjects in the y-axis and time to these TEAEs in the x-axis (Appendix 7).

Due to suspected differential dropout rates between roxadustat vs. placebo and the resulted biases on event rates, selected TEAEs and TESAEs will be summarized by PEY censoring for start of chronic dialysis or renal transplant, and also stratified by baseline eGFR category (<10, 10-<15, 15-<30, and >=30 mL/min/1.73m^2).

13.1.3 Hospitalization-Free Survival On Treatment

Hospitalization-free survival days, and number of days on treatment out of hospitalization and skilled nursing facility be compared between roxadustat vs. placebo using ANCOVA with baseline Hb, baseline eGFR and other randomization stratification factors as baseline covariates.

13.1.4 Pooled Analyses

The analytical methods for the composite safety endpoints of interest will be described in a region-specific pooled SAP to reflect the nature of the pooling of these endpoints across multiple studies in the Phase 3 program and the region-specific safety endpoints.

13.2 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 laboratory parameters collected in the study that include but are not limited to the following:

Hematology: Hemoglobin, hematocrit, RBC count, MCV, WBC count, WBC differential, and platelet counts;

Chemistry: Alkaline phosphatase, ALT, AST, total bilirubin, CPK, LDH, total protein, albumin, fasting glucose, phosphate, uric acid, BUN, creatinine, sodium, and potassium; eGFR, HbA1C; UACR

To assess potentially clinically significant (PCS) laboratory abnormalities (Appendix 2), the number and percentage of subjects with post-baseline lab values outside a pre-defined range (low or high) or limit of change will be tabulated by treatment group. Laboratory test values will be considered PCS if they meet the criteria listed in an appendix. The percentages are to be calculated relative to the number of subjects with available non-PCS baseline values and at least one post-baseline assessment. The numerator is the total number of subjects with at least one post-baseline PCS value. Shift tables may be presented. A supportive listing of subjects with post-baseline PCS values will be provided including the patient ID, study center, baseline, and post-baseline values. A listing of all AEs for subjects with PCS laboratory values will also be provided.

13.2.1 Liver Function Tests

A matrix scatter plot of Liver Enzymes and Bilirubin will be showing the maximum ALT, AST, ALP and total bilirubin during the treatment period crossed against each other on log scales. Different dots will be used for roxadustat and placebo.

Individual displays of abnormal Liver Enzymes (AST or ALT >=3 x ULN) and Bilirubin parameters (T. Bili >=2 x ULN) with corresponding other liver enzymes and bilirubin values during the treatment periods will be listed.

Time to first occurrence of abnormal Liver Enzymes or Bilirubin will be plotted.

For subjects who require further liver function investigations, additional information will be collected and listed.

13.3 VITAL SIGNS

Blood Pressures and Heart Rate baselines are defined as the mean of values obtained from the last 2 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 13.3 – 1* below. 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 patient ID, study center, baseline, and post-baseline values. A listing of all AEs for subjects with PCS vital signs will also be provided.

Table 13.3 – 1. 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	≥ 110	Increase of ≥ 15
	Low	≤ 45	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%

*Except for body weight, a post-baseline value is considered as a PCS value if it meets both criteria for observed value and change from baseline.

Additional analyses may include but are not limited to subgroup analyses of blood pressure in patients without any change in BP meds during treatment period.

13.4 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.

ECG parameters values are potentially clinically significant (PCS) if they meet or exceed the upper limit values and changes (where applicable) listed in Table 13.4-1 below. 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 patient 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 13.4 – 1. 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

13.5 EXCESSIVE HEMATOPOIESIS

Time to initial onset of potential Excessive Hematopoiesis, regardless any use of rescue therapy, will be analyzed based on Hb central lab

The presence of potential EH will be defined as either:

- Hb increase by >2.0 g/dL between any 2 visits within 4 weeks of treatment, or
- at least one Hb value >13 g/dL during the treatment period.

Each of these two criteria will be analyzed separately. Time to first occurrence of potential EH (for either criterion separately) regardless the use of rescue therapy during the treatment period will be defined in weeks as :(First event date – Analysis date of first dose + 1) / 7

where 'First event date' is defined as first date of occurrence of the criterion met during the treatment period.

For a subject without potential EH, the time to censoring will be calculated (in weeks) as:

(Date of last hemoglobin assessment during the treatment period – Analysis date of first dose + 1) / 7

14 ADDITIONAL AND SUBGROUP ANALYSES

Primary analysis will be performed separately by sex, age group, baseline CKD stage, baseline iron repletion status ([TSAT ≥20% and ferritin ≥100 ng/mL] vs. others), baseline CRP group (CRP ≤ULN vs. CRP > ULN), eGFR (≤15 vs. >15 mL/min/1.73m^2) and other randomization stratification factors, as needed.

In addition, secondary endpoints will be analyzed by subgroups using pooled data from multiple studies, detailed in an Integrated Summary of Efficacy Statistical Analysis Plan.

15 INTERIM ANALYSIS

The study will have no interim analysis with statistical inference. Safety data and dosing decisions will be monitored on an ongoing basis. Ongoing review of safety data will be completed by an independent DSMB.

16 REFERENCES

ICH Harmonized Tripartite Guideline E 3. Structure and Content of Clinical Study Reports, November 1995. (www.ich.org; Guidelines; "Efficacy" Topics)

ICH Harmonized Tripartite Guideline E 8. General Considerations of Clinical Trials, July 1997. (www.ich.org; Guidelines; "Efficacy" Topics)

ICH Harmonized Tripartite Guideline E 9. 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

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.

Webster K, Cella D and Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. Heath and Quality of Life Outcomes 2003; 1: 79

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

17 APPENDIX 1: DATA HANDLING CONVENTIONS

17.1 VISIT TIME WINDOWS

Table 17.1 – 1 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 17.1 – 1. Analysis Visit Windows

Analysis Visit	Scheduled Visit Day [a]	Window
Baseline, Week 0	Day 1	Days ≤ 1
Week 1	Day 7*(Week #)+1	Days [2, 10]
Week 2	Day 7*(Week #)+1	[Scheduled Day -3, Scheduled Day +6]
Week 4-22	Day 7*(Week #)+1	[Scheduled Day -7, Scheduled Day +6]
Week 24	Day 7*(Week #)+1	[Scheduled Day -7, Scheduled Day +13]
Week 28-xx	Day 7*(Week #)+1	[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 in the Study 15 – 31 days after Last Dose (excluding long term follow-up for early termination).	

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

Table 17.1 – 2. Analysis Visit Windows for HRQoL

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 + 55]	Week 12
Week 28	Day 7 * (Week #) + 1	[Target Day – 56, Target Day + 55]	Week 28
Week 52	Day 7 * (Week #) + 1	[Target Day – 56, Target Day + 83]	Week 52
Week 76	Day 7 * (Week #) + 1	[Target Day – 84, Target Day + 83]	Week 76
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, 76 or >1.5 years

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

Analysis Visit windows, as depicted in below, will be used for the Lipid Panel, including LDL cholesterol efficacy study assessment:

Table 17.1 – 3. Analysis Visit Windows for Lipid Panel

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Day 1	Day 1	Day 1	Baseline
Week 4	Day 7 * (Week #) + 1	Day 2 to Day 42	Week 4
Week 8	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 13]	Week 8
Week 12	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 27]	Week 12
Week 20,28,36,44,	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 27]	Week 20,28,36,44
Week 52	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 41]	Week 52
Q12Wk	Day 7 * (Week #) + 1	[Target Day – 42, Target Day + 41]	Week 64, 72, 88, 100…
ET/EOT Visit		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	Week xx
EOS (FU-4Wk) Visit		Final visit in the study 15-31 days after Last Dose.	FU-4Wk

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

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

17.2 REPEATED OR UNSCHEDULED ASSESSMENTS OF SAFETY PARAMETERS

If a patient has repeated assessments prior to the start of double-blind study medication, then the results from the final assessment made prior to the start of study medication will be used as baseline, except where otherwise specified. If post-baseline assessments are repeated or unscheduled, the last post-baseline assessment within a visit window will be used as the study visit 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.

17.3 MISSING DATE OF STUDY MEDICATION

When the last date of double-blind 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.

17.4 MISSING SEVERITY ASSESSMENT FOR ADVERSE EVENTS

If severity is missing for an AE started prior to the first double blind study medication, then a severity of "Mild" will be assigned. If the severity is missing for an AE started on or after the first double blind 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.

17.5 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 double-blind study medication will be used for incidence summary, while the actual values will be presented in data listings.

17.6 MISSING DATE INPUTATION FOR ADVERSE EVENTS

- **Incomplete Start Date**

The following imputation rules apply to the case where the start date is incomplete (i.e., partially missing) for adverse events.

Missing day and month

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

If the year is not the same as the year of first day on double-blind 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 double-blind study medication, then the start date of double-blind 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 double-blind study medication, then the first day of the month will be assigned to the missing day.

- **Incomplete Stop Date**

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.

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.

Table 17.6-1 Imputation of the Analysis Adverse Event Start Date

Reported Date	Date of First Drug Intake	Analysis Date (Derived)
--/MM/YYYY	DD/MM/YYYY	
--/02/2008	14/02/2008	14/02/2008*
--/02/2008	14/02/2007	01/02/2008
--/02/2008	14/02/2009	01/02/2008
--/--/YYYY	DD/MM/YYYY	
--/--/2008	14/02/2008	14/02/2008
--/--/2008	14/02/2007	01/01/2008
--/--/2008	14/02/2009	01/01/2008
DD/--/---- --/MM/---- --/--/----		No imputation

Table 17.6-2 Imputation of the Analysis Adverse Event Stop Date

Reported Date	Analysis Date (Derived) *
--/MM/YYYY	31/MM/YYYY or 30/MM/YYYY or 29/MM/YYYY or 28/MM/YYYY
--/--/YYYY	31/12/YYYY
DD/--/----, or --/MM/----, or --/--/----	No imputation

*Death has to be taken into consideration when calculating this.

17.7 MISSING DATE IMPUTATION FOR PRIOR OR CONCOMITANT MEDICATIONS

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

- **Incomplete Start Date**

The following rules will be applied to impute the missing start date. 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 double-blind 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 not the same as the first dose date of double-blind 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 double-blind study medication, then the day of the first dose date will be assigned to the missing day.

If the month and year of the incomplete start date are the same as the first dose date of double-blind 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 stop date, if needed. If the last dose date of double-blind study medication is missing, impute 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 replaced with 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 double-blind 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 not the same as the year of the last dose date of double-blind study medication, then December 31 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 double-blind study medication, then the day of the last dose date will be assigned to the missing day.

If the month and year of the incomplete stop date are not the same as the month and year of the last dose date of double-blind study medication, then the last day of the month will be assigned to the missing day.

17.8 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).

17.9 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.

Table17.9 – 1. Example for Coding of Special Character Values for Clinical Laboratory Parameters

Lab Test	Possible Lab Results (in SI unit)	Coded Value for Analysis
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

18 APPENDIX 2: RANGES OF POTENTIALLY CLINICALLY SIGNIFICANT LAB VALUES

Parameter	SI Unit	Lower Limit	Higher Limit
CHEMISTRY			
Alanine Aminotransferase (ALT)	U/L		≥3 * ULN
Alkaline Phosphatase (ALP)	U/L		≥3 * ULN
Aspartate Aminotransferase (AST)	U/L		≥3 * ULN
Gamma-Glutamyl Transferase (GGT)	U/L		≥3 * ULN
Calcium	mmol/L	<0.8*LLN	>1.2 * ULN
Creatine Phosphokinase(CPK)	U/L		>10 * ULN
Creatinine	µmol/L		> 1.5 x Baseline Value
Potassium	µmol/L	<0.75*LLN	>1.2 * ULN
Sodium	mmol/L	<0.9*LLN	>1.1 * ULN
Total Bilirubin	µmol/L		>1.5 * ULN
Total Protein	µmol/L	<0.9 * LLN	>1.1 * ULN
Urea (BUN)	mmol/L		>1.5x Baseline Value
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			

19 APPENDIX 3: SF-36 V2

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 v2

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

		Yes, limited a lot	Yes, limited a little	No, not limited at all
a	<u>Vigorous activities,</u> such as running, lifting heavy objects, participating in strenuous sports	☐ 1	☐ 2	☐ 3
b	<u>Moderate activities,</u> 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 <u>several</u> flights of stairs	☐ 1	☐ 2	☐ 3
e	Climbing <u>one</u> flight of stairs	☐ 1	☐ 2	☐ 3
f	Bending, kneeling, or stooping	☐ 1	☐ 2	☐ 3
g	Walking <u>more than a mile</u>	☐ 1	☐ 2	☐ 3
h	Walking <u>several hundred yards</u>	☐ 1	☐ 2	☐ 3
i	Walking <u>one hundred yards</u>	☐ 1	☐ 2	☐ 3
j	Bathing or dressing yourself	☐ 1	☐ 2	☐ 3

SF-36 v2

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

	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 <u>amount of time</u> you spent on work or other activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b <u>Accomplished less</u> than you would like	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c Were limited in the <u>kind</u> of work or other activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d Had <u>difficulty</u> performing the work or other activities (for example, it took extra effort)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

5. **During the <u>past 4 weeks</u>, how much of the time have you had any of the following problems with your work or other regular daily activities <u>as a result of any emotional problems</u> (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 <u>amount of time</u> you spent on work or other activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b <u>Accomplished less</u> than you would like	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c Did work or other activities <u>less carefully than usual</u>	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

6. **During the past 4 weeks, 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 bodily pain have you had during the past 4 weeks?**

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

8. **During the past 4 weeks, how much did pain 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 v2

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

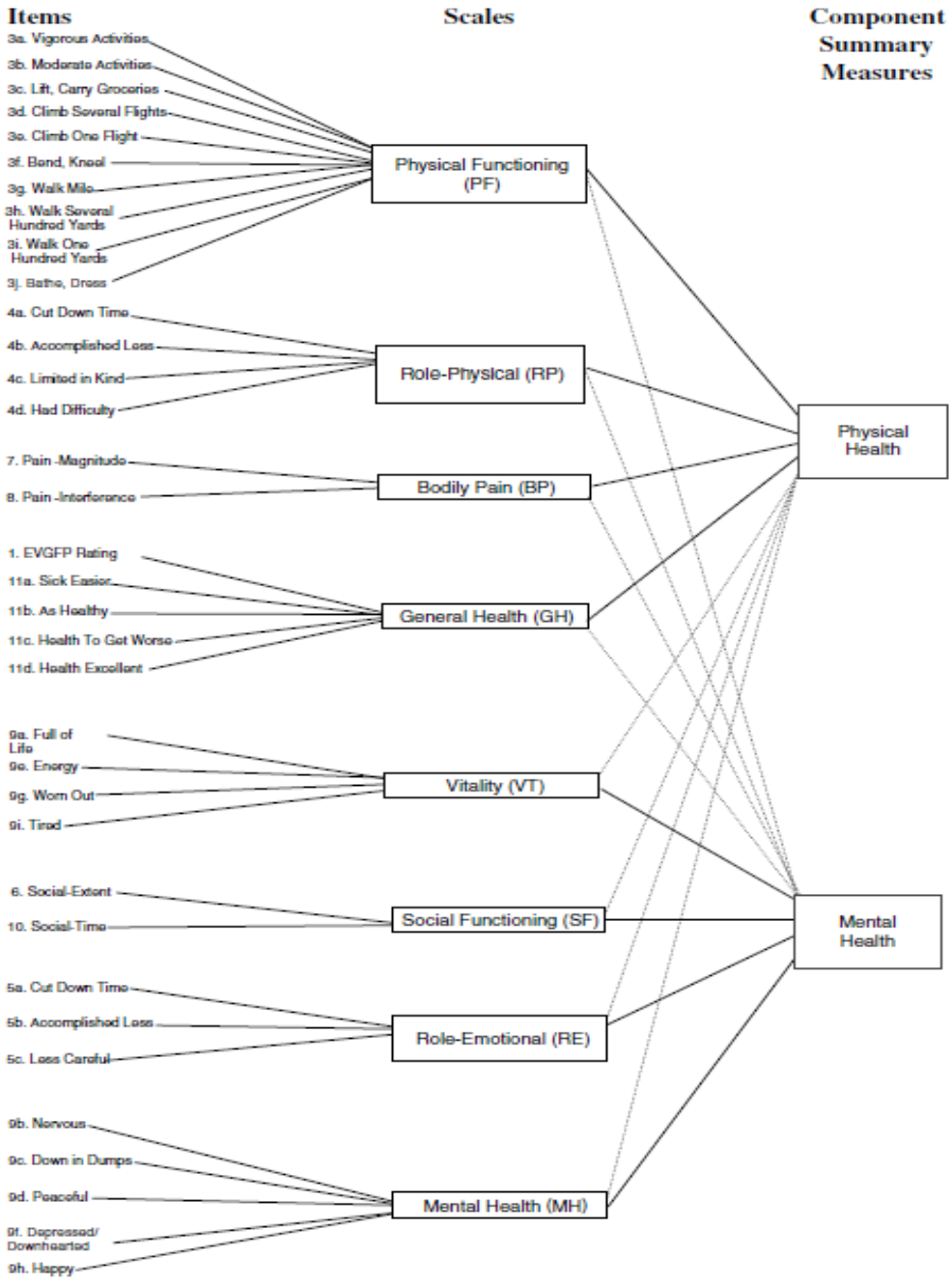
SF-36 v2

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!

SF-36 Model


Items
Scales
Component Summary Measures
3a. Vigorous Activities
3b. Moderate Activities
3c. Lift, Carry Groceries
3d. Climb Several Flights
3e. Climb One Flight
3f. Bend, Kneel
3g. Walk Mile
3h. Walk Several Hundred Yards
3i. Walk One Hundred Yards
3j. Bathe, Dress
Physical Functioning (PF)
4a. Cut Down Time
4b. Accomplished Less
4c. Limited in Kind
4d. Had Difficulty
Role-Physical (RP)
7. Pain -Magnitude
8. Pain -Interference
Bodily Pain (BP)
Physical Health
1. EVGFP Rating
11a. Sick Easier
11b. As Healthy
11c. Health To Get Worse
11d. Health Excellent
General Health (GH)
9a. Full of Life
9e. Energy
9g. Worn Out
9i. Tired
Vitality (VT)
6. Social-Extent
10. Social-Time
Social Functioning (SF)
Mental Health
5a. Cut Down Time
5b. Accomplished Less
5c. Less Careful
Role-Emotional (RE)
9b. Nervous
9c. Down in Dumps
9d. Peaceful
9f. Depressed/ Downhearted
9h. Happy
Mental Health (MH)

20 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

FACT-An Scoring Guidelines

Instructions:* 1. Record answers in "item response" column. If missing, mark with an X
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
4. Add subscale scores to derive total scores (TOI, FACT-G & FACT-An).
5. **The higher the score, the better the QOL.**

Subscale	Item Code	Reverse item?		Item response	Item Score
PHYSICAL WELL-BEING (PWB)	GP1	4	-	_____	=_____
	GP2	4	-	_____	=_____
	GP3	4	-	_____	=_____
	GP4	4	-	_____	=_____
Score range: 0-28	GP5	4	-	_____	=_____
	GP6	4	-	_____	=_____
	GP7	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____
=**PWB subscale score**

SOCIAL/FAMILY WELL-BEING (SWB)	GS1	0	+	_____	=_____
	GS2	0	+	_____	=_____
	GS3	0	+	_____	=_____
	GS4	0	+	_____	=_____
Score range: 0-28	GS5	0	+	_____	=_____
	GS6	0	+	_____	=_____
	GS7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____
=**SWB subscale score**

EMOTIONAL WELL-BEING (EWB)	GE1	4	-	_____	=_____
	GE2	0	+	_____	=_____
	GE3	4	-	_____	=_____
	GE4	4	-	_____	=_____
Score range: 0-24	GE5	4	-	_____	=_____
	GE6	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 6: _____
Divide by number of items answered: _____
=**EWB subscale score**

FUNCTIONAL WELL-BEING (FWB)	GF1	0	+	_____	=_____
	GF2	0	+	_____	=_____
	GF3	0	+	_____	=_____
	GF4	0	+	_____	=_____
Score range: 0-28	GF5	0	+	_____	=_____
	GF6	0	+	_____	=_____
	GF7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____
=**FWB subscale score**

Subscale	**Item Code**	**Reverse item?**		**Item response**	**Item Score**
ANEMIA SUBSCALE (AnS)	HI7	4	-	_____	=_____
	HI12	4	-	_____	=_____
	An1	4	-	_____	=_____
	An2	4	-	_____	=_____
Score range: 0-80	An3	4	-	_____	=_____
	An4	4	-	_____	=_____
	An5	0	+	_____	=_____
	An6	4	-	_____	=_____
	An7	0	+	_____	=_____
	An8	4	-	_____	=_____
	An9	4	-	_____	=_____
	An10	4	-	_____	=_____
	B1	4	-	_____	=_____
	An11	4	-	_____	=_____
	An12	4	-	_____	=_____
	BL4	0	+	_____	=_____
	An13	0	+	_____	=_____
	An14	4	-	_____	=_____
	An15	4	-	_____	=_____
	An16	4	-	_____	=_____

Sum individual item scores:_____
Multiply by 20: _____
Divide by number of items answered: _____
=**An Subscale score**

To derive a FACT-An Trial Outcome Index (TOI):
Score range: 0-136

_____ + _____ + _____ =_____=**FACT-An TOI**
(**PWB** score) (**FWB** score) (**AnS** score)

To Derive a FACT-G total score:
Score range: 0-108

_____ + _____ + _____ + _____=_____=**FACT-G Total score**
(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score)

To Derive a FACT-An total score:
Score range: 0-188

_____ + _____ + _____ + _____ + _____ =_____=**FACT-An Total score**
(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score) (**AnS** score)

*For guidelines on handling missing data and scoring options, please refer to the Administration and Scoring Guidelines in the manual or on-line at www.facit.org.

FACIT-Fatigue Subscale Scoring Guidelines

Instructions:* 1. Record answers in "item response" column. If missing, mark with an X
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
4. The higher the score, the better the QOL.

Subscale	Item Code	Reverse item?		Item response	Item Score
FATIGUE	HI7	4	-	_____	=_____
SUBSCALE	HI12	4	-	_____	=_____
	An1	4	-	_____	=_____
	An2	4	-	_____	=_____
Score range: 0-52	An3	4	-	_____	=_____
	An4	4	-	_____	=_____
	An5	0	+	_____	=_____
	An7	0	+	_____	=_____
	An8	4	-	_____	=_____
	An12	4	-	_____	=_____
	An14	4	-	_____	=_____
	An15	4	-	_____	=_____
	An16	4	-	_____	=_____

Sum individual item scores:_____
Multiply by 13: _____
Divide by number of items answered: _____
=Fatigue Subscale score

21 APPENDIX 5: EQ-5D 5L V2

Under each heading, please tick the ONE box that best describes your health TODAY

MOBILITY

- I have no problems in walking about ☐
- I have slight problems in walking about ☐
- I have moderate problems in walking about ☐
- I have severe problems in walking about ☐
- I am unable to walk about ☐

SELF-CARE

- I have no problems washing or dressing myself ☐
- I have slight problems washing or dressing myself ☐
- I have moderate problems washing or dressing myself ☐
- I have severe problems washing or dressing myself ☐
- I am unable to wash or dress myself ☐

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

- I have no problems doing my usual activities ☐
- I have slight problems doing my usual activities ☐
- I have moderate problems doing my usual activities ☐
- I have severe problems doing my usual activities ☐
- I am unable to do my usual activities ☐

PAIN / DISCOMFORT

- I have no pain or discomfort ☐
- I have slight pain or discomfort ☐
- I have moderate pain or discomfort ☐
- I have severe pain or discomfort ☐
- I have extreme pain or discomfort ☐

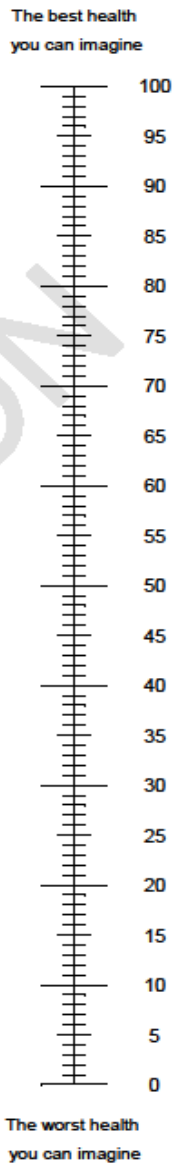
ANXIETY / DEPRESSION

- I am not anxious or depressed ☐
- I am slightly anxious or depressed ☐
- I am moderately anxious or depressed ☐
- I am severely anxious or depressed ☐
- I am extremely anxious or depressed ☐

EQ-5D 5L v2

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



3

22 APPENDIX 6: MEDICATION WHO DRUG DICTIONARY CODES

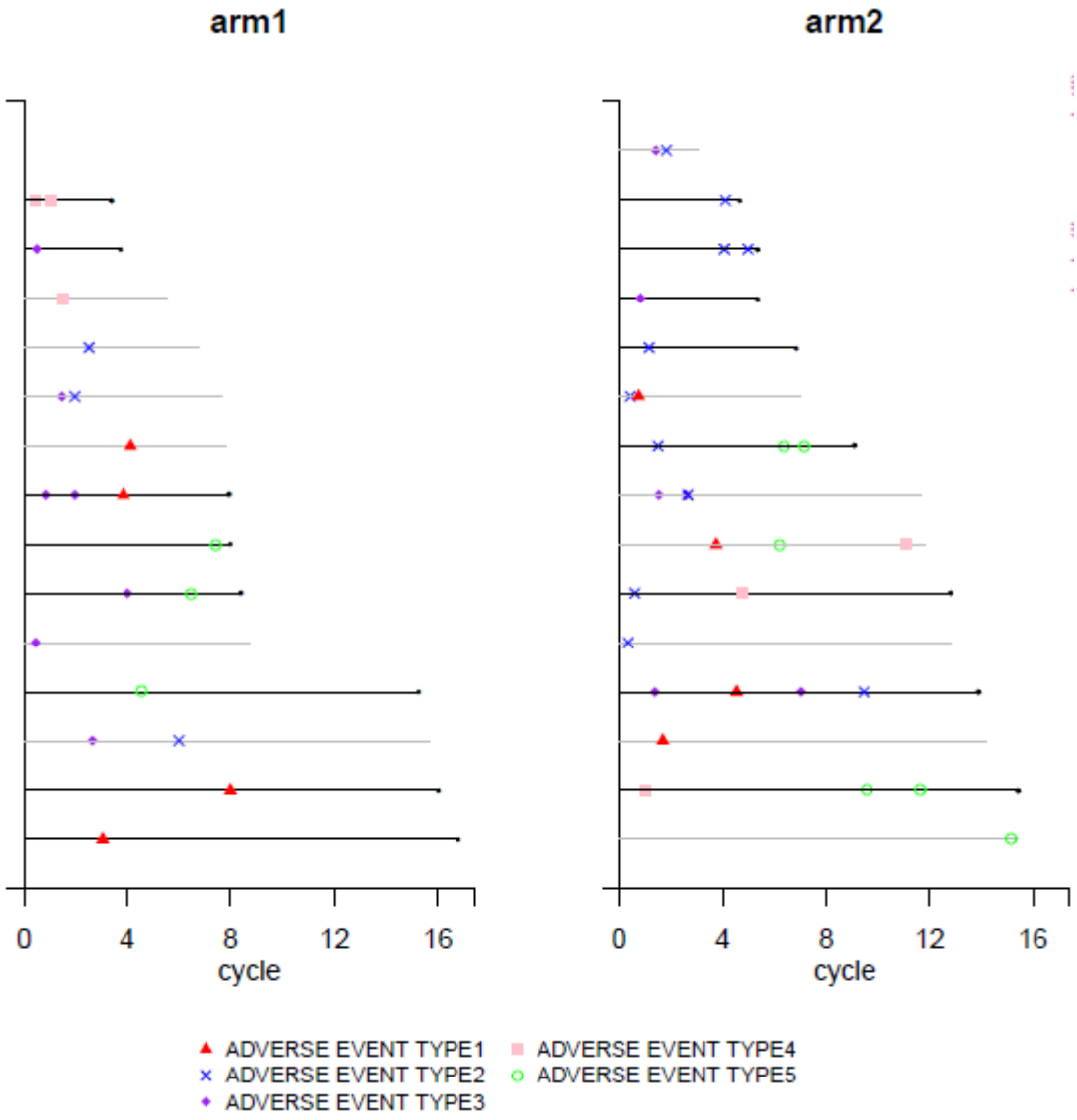
Name	Code
ESA except darbepoetin alfa	ATC level 4 = B03XA
Darbepoetin alfa	ATC level 4 = B03XA
IV Iron	ATC level 4 = B03AC
Blood/RBC transfusion	ATC level 4 = B05AX
Any investigational drug	WHODD drug code = 99999701001
Hypoxia-inducible factor HIF-PHI	ATC level 4 = B03XA
Iron-chelating agents	ATC level 4 = B03AA, B03AD, B03AE, A12CX
Androgens	ATC level 3 = G03B and G03E
Dapsone	ATC level 4 = J04BA
Acetaminophen/paracetamol	ATC level 4 = N02BE, N02AA, R05X

The following WHO-DRL codes will be classified as ESA: '00909301001', '00928301001', '02198701001', '07973701001', '01703101001'.

The following WHO-DRL code where route is INTRAVENOUS will be classified as IV IRON: '00023501001' and '90135401001'.

The following WHO-DRL code will be classified as blood/RBC transfusion: '01186901001'.

23 APPENDIX 7: TEMPEROAL PROFILE OF TEAE'S OF SPECIAL INTEREST



24 APPENDIX 8: INVERSE PROBABILITY CENSORING WEIGHTED MODEL

Background

Differential dropouts between treatment arms with informative censoring, may cause biased results from conventional survival methodology using Cox regression or other non-parametric survival analysis methods. The inverse probability of censoring weighted (IPCW) approach can be used under these situations to better estimate the treatment difference.

IPCW is an analysis method of time-to-event endpoint adjusting time-dependent variables as covariates and contribution of subjects in the analysis model using the inverse of their censoring probability. The basic idea of using this IPCW estimator is to correct for time-variant censored subjects by giving more weight to subjects who are not censored. This way, the final survival analysis model is fit as if censoring is absent. Subjects carry more weight in the analysis model if subjects stay longer in the trial, in other words, subjects who are censored sooner are less representative in IPCW estimators

Step 1: Model the censoring mechanism.

In the censoring model, the event of interest is censoring, hence the subjects who are lost to follow-up have an 'event'. Subjects that are not censored, i.e. those who experience the original event of interest, are now considered ‘censored' since their censoring time is not observed. The probability of being censored will be estimated.

Step 2: Estimate the Product-Limit estimator and Cox proportional hazards estimator

Estimate the P-L estimator and Cox regression estimators using time to censoring for each subject j at each time point t, $K^{0}_{j}(t)$ and $K^{Z}{}_{j}(t)$. The Cox Model will include eGFR categories (e GFR < =10, 10<= eGFR < 15, 15<= eGFR < 30, and eGFR >= 30), and Dialysis initiation (Yes, No) as time varying covariates. Covariates including baseline hemoglobin, and other randomization stratification factors will also be included in the model.

Step 3: Calculate the unstablized and stabilized IPCW (SIPCW) weights for each of the subjects, j, $Wunstab_j(t) = 1/K^{z}_{j}(t)$, $Wstab_j(t) = K^{0}_{j}(t)/K^{z}_{j}(t)$.

Step 4: Estimate the survival and/or Cox model for time-to-event in the absence of censoring with the IPCW weight, SIPCW.