FibroGen Inc.
409 Illinois Street
San Francisco, CA 94158 USA

Protocol Number	**FGCL-4592-082, Amendment 03**
Title	**A Phase 3 Randomized Double-Blind Placebo-Controlled Study Investigating the Efficacy and Safety of Roxadustat (FG-4592) for Treatment of Anemia in Patients with Lower Risk Myelodysplastic Syndrome (MDS) with Low Red Blood Cell (RBC) Transfusion Burden (LTB)**
Document	**Statistical Analysis Plan**
Version	**1.0**
Release Date	**April 10, 2023**

Approvals

I have reviewed and accepted the information in this document to be a true and accurate representation of the Statistical Analysis Plan for Study FG-4592-082.

Initiator:

Signature: [redacted] Date: _____

Reviewed by:

Signature: [redacted] Date: _____

Signature: [redacted] Date: _____

Signature: [redacted] Date: _____

Signature: [redacted] Date: _____

Signature: [redacted] Date: _____

Signature: [redacted] Date: _____

Signature: [redacted] 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
ANCOVA	Analysis of Covariance
ATC	Anatomical/Therapeutic/Chemical
ALT	Alanine Aminotransferase
AML	Acute Myeloid Leukemia
AST	Aspartate Aminotransferase
BSC	Best Supportive Care
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus disease 2019
CRF	Case Report Form
CRO	Contract Research Organization
CRP	C-Reactive Protein
CTCAE	Common Terminology Criteria for Adverse Events
DB	Double Blind
DSMB	Data Safety Monitoring Committee
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EPO	Gene Encoding Erythropoietin
EQ-5d-5L	European Quality Of Life-5 Dimensions-5 Levels
EOT	End of Treatment
ET	Early Termination
FAS	Full Analysis Set
GFR	Glomerular Filtration Rate
PF	Physical Function
Hb	Hemoglobin
HRQoL	Health-related Quality of Life
ICH	International Conference on Harmonization
INR	International Normalized Ratio
IPSS-R	International Prognostic Scoring System – Revised
IRT	Item Response Theory
IWG	International Working Group
IWRS	Interactive Web Response System
LSM	Least-Squares Mean
LTB	Low Red Blood Cell Transfusion Burden
MDS	Myelodysplastic Syndromes
MedDRA	Medical Dictionary For Regulatory Activities
MCID	Minimal Clinically Important Difference
MMRM	Mixed Model Repeated Measures
NCI	National Cancer Institute
OL	Open Label
PEW	Patient-Exposure-Week
PD	Pharmacodynamics
PI	Principal Investigator
PK	Pharmacokinetics
PPS	Per-Protocol Set
pRBC	packed Red Blood Cell
PROMIS	Patient Reported Outcomes Measurement Information System

PT	Preferred Term
QoL	Quality of Life
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SMQ	Standardized MedDRA Query
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TI	Transfusion Independence
TIBC	Total Iron Binding Capacity
TSAT	Transferrin Saturation
ULN	Upper Limit of Normal
UIBC	Unsaturated Iron Binding Capacity
WHO	World Health Organization
WHODD	WHO Drug Dictionary

1. INTRODUCTION

This document describes prospectively the statistical analyses and presentation of data for the clinical trial conducted under Protocol FGCL-4592-082, A Phase 3 Randomized Double-Blind Placebo-Controlled Study Investigating the Efficacy and Safety of Roxadustat (FG-4592) for Treatment of Anemia in Patients with Lower Risk Myelodysplastic Syndrome (MDS) with Low Red Blood Cell (RBC) Transfusion Burden (LTB), Amendment 03, dated 30 January 2020. It contains definitions of analysis populations, derived variables, and statistical methods for the analysis of efficacy and safety.

The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to any data analysis prior to database lock. This SAP will be finalized prior to database lock of the primary efficacy analysis of the study after all patients have completed 28 weeks of double-blind treatment or terminated before this time point.

This study consists of three components – an open-label lead-in dose-finding component followed by a randomized, double-blind component and an open-label high Erythropoietin (EPO) component. Data from patients in the open-label component and an open-label high EPO will be analyzed separately from the double-blind randomized component of the study. This SAP is based on protocol version 3.0 dated 30 January 2020.

2. STUDY OBJECTIVES

2.1. Open-Label Component Objective

Optimize the starting dose for the double-blind component

2.2. Open-label (OL) High-Erythropoietin Component Objective

This is an "exploratory cohort" for patients with higher EPO levels (serum Erythropoietin levels >400 mIU/mL) to further understand treatment response to Roxadustat in patients with higher than 400 mIU/mL endogenous EPO levels.

2.3. Double-Blind Component

2.3.1. Primary Objective

The primary objective of this study is to evaluate the efficacy of roxadustat in the treatment of anemia in patients with lower risk MDS who have a low burden of RBC transfusion.

2.3.2. Secondary Objectives

The secondary objectives in this study are to:

- Evaluate the safety of roxadustat
- Evaluate the impact of roxadustat on RBC transfusion requirements
- Evaluate pharmacokinetics (PK) and pharmacodynamics (PD) of roxadustat in MDS patients
- Evaluate effect of roxadustat on quality of life (QoL) parameters

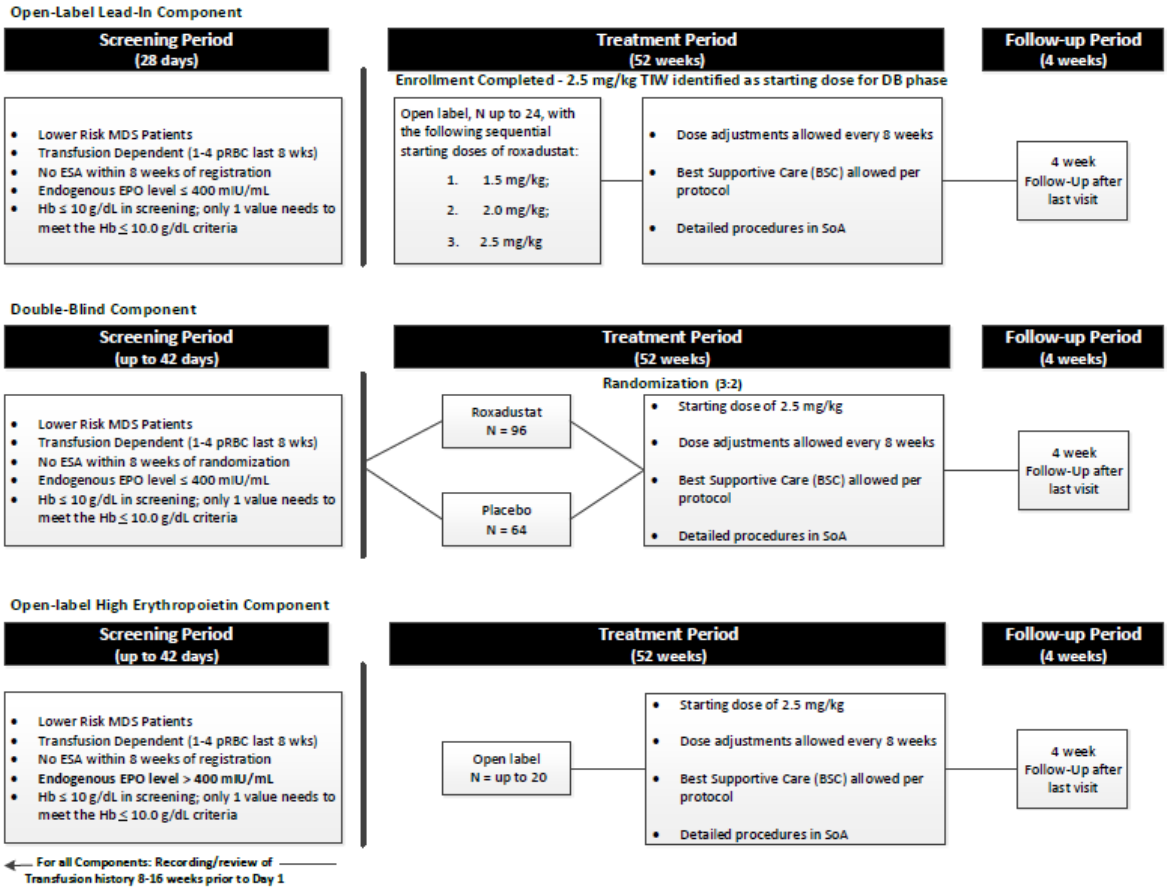
3. INVESTIGATIONAL PLAN

3.1. Overall Study Design

This is a Phase 3 Randomized, Double-Blind, Placebo-Controlled Study Investigating the Efficacy and Safety of Roxadustat (FG-4592) for Treatment of Anemia in Patients with Lower Risk Myelodysplastic Syndrome (MDS) with Low Red Blood Cell (RBC) Transfusion Burden (LTB). Patients are planned to be accrued from approximately 100 global centers.

The study consists of three components: (1) an open-label lead-in component to optimize the starting dose; (2) a double-blind, randomized 2-arm component to evaluate the efficacy and safety of roxadustat (FG-4592) versus matching Placebo along with best supportive care (BSC); 3) an open-label (OL) high-erythropoietin component to evaluate safety and efficacy of roxadustat (transfusion independence) in patients with high endogenous erythropoietin in order to facilitate decision making whether similar patients with elevated endogenous serum erythropoietin levels should be allowed to participate in the DB component. The first two components will be carried out sequentially. The third component will be carried out concurrent to double blind enrollment, refer to Figure 1 for the study design schema from the protocol.

Figure 1: Study Schema



Open-label (OL) component (Open-Label, Dose-Finding, Lead-In Component): Dose evaluation based on efficacy. Up to 24 patients will be enrolled in sequential dose level cohorts prior to start of the double-blind component of the study. All patients will receive roxadustat in an open-label manner. The first 8

may receive a starting dose of 1.5 mg/kg. The next 8 may receive a starting dose of 2.0 mg/kg. The last 8 may receive a starting dose of 2.5 mg/kg.

The initial open-label patients may remain in the study for up to 52 weeks of treatment based on investigator discretion, adequate safety, and clinical benefit.

The initial open-label patients will undergo all procedures as the randomized patients in the double-blind study with the exception of not completing the Patient Reported Outcomes (PRO) assessments.

Double-blind (DB) component: 160 eligible patients will be randomized in the double-blind component of the study. When all the randomized patients complete 28 weeks of treatment or are discontinued from study medication before this time, the analysis of the primary efficacy endpoints will be performed for a topline read-out. When all the randomized patients complete 52 weeks of treatment or are discontinued from study medication before this time, the full analysis of primary and secondary efficacy endpoints will be performed.

In this DB component, 160 eligible patients will be randomized in a 3:2 assignment ratio to roxadustat or matching placebo with the following stratification factors:

- Endogenous serum EPO level: ≤ 200 mIU/mL vs. > 200 mIU/mL and ≤ 400 mIU/mL
- IPSS-R: low risk / very low risk vs. intermediate risk classification
- RBC transfusion burden: 1 pRBC/8-weeks over 16 consecutive weeks vs. 2-4 pRBC/8-weeks

NOTE: 1-pRBC randomization will be limited to no more than 30% of the total randomized patient population

OL high-erythropoietin component： Based on the Amendment 03 of the protocol dated 30 January 2020, concurrent to DB enrollment, up to 20 patients with high serum erythropoietin levels will be enrolled in an OL exploratory cohort. Unless otherwise mentioned, patients enrolled in this high erythropoietin cohort will follow all DB study procedures including the same starting dose (2.5 mg/kg TIW), same visit schedule, etc., except that the patients will all be administered roxadustat in an OL manner. Safety and efficacy assessments (transfusion independence) will be made on an ongoing basis to facilitate decision making on whether patients with elevated serum erythropoietin levels should be allowed to participate in the DB component of the study.

The study will consist of the following visits:

Baseline Assessments: To be obtained ≤ 42 days prior to starting study drug for open-label component and double-blind component respectively.

Randomization (double-blind component only**):** Randomization into the double-blind component to occur on Day 1 (Week 1) prior to first dose.

Treatment: The maximum treatment duration for patients in both the open-label and double-blind components is 52 weeks.

Week 28 visit: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit). Primary analysis of efficacy will be conducted when all the randomized (DB) patients complete 28 weeks of treatment or are discontinued from study medication before this time. **End-of-Treatment (EOT/ET) Visit:** Patients who complete 52-week treatment period or discontinue study medication early (Early Termination before week 52) will have an EOT/ET visit.

Post-Treatment Follow-Up Period: Week 56 (Post-EOT Week 4)

Post-Treatment Follow-Up Period for ET Patients: Patients who discontinue early will have visits every 8 weeks until Week 52.

End of Study (EOS) Visit: All patients who complete 52 weeks (of treatment with study medication) will undergo a 4 week follow-up period after the last dose of study medication and will return for the End of Study (EOS) visit. Patients who discontinue study medication early (for any reason) will remain in the study through the entire 52 week study period. Patients will have a 4 week follow-up visit after stopping study medication/ET visit and then be followed up at a reduced visit schedule of every 8 weeks until the end of the study.

3.2. Sample Size Determination

For the open-label component, the sample size is not determined based on statistical consideration. Up to 24 patients will be enrolled in sequential dose level cohorts. All patients will receive roxadustat in an open-label manner.

For the OL high-erythropoietin component, the sample size is not determined based on statistical consideration. Up to 20 patients will be enrolled in an OL exploratory cohort. All patients will receive roxadustat in an open-label manner.

For the double-blind component, 160 patients are planned to be randomized in a 3:2 ratio to roxadustat or matching placebo. This sample size will provide at least 90% power to demonstrate a statistically significant difference between roxadustat and placebo at the significance level of 0.05 using a two-sided Fisher's test, and assuming the response rates of the primary endpoint are 35% and 5% in roxadustat and placebo, respectively.

In total, approximately up to 204 patients will participate in this study.

Due to the slow enrollment in the double-blind component, an unplanned blinded sample size estimation was performed. Based on this, 120 patients randomized in a 3:2 ratio to roxadustat or matching placebo will provide at least 90% power to demonstrate a statistically significant difference between roxadustat and placebo at the significance level of 0.05 using a two-sided Fisher's exact test, and assuming the response rates of the primary endpoint are 56% and 13% in roxadustat and placebo, respectively. The response rate on roxadustat was based on the response rates from open-label and double-blind components in a blinded fashion. The response rate on placebo was based on the MEDALIST study.

3.3. Randomization and Treatment Assignment

For the open-label component, eligible patients will be sequentially assigned to the cohorts, starting with 1.5mg/kg for the first 8 patients, then 2.0mg/kg for the next 8 patients and at 2.5 mg/kg for the last 8 patients. No randomization is needed.

For the OL high-erythropoietin component, eligible patients will be assigned to an OL exploratory cohort, starting with 2.5 mg/kg TIW. No randomization is needed.

For the double-blind component, patient randomization will be performed by an Interactive Web Response System (IWRS) to ensure a 3:2 assignment ratio between two treatment arms within each stratum. The randomization will be stratified by the following 3 factors:

- Endogenous serum EPO level: ≤ 200 mIU/mL vs. > 200 mIU/mL and ≤ 400 mIU/mL
- IPSS-R risk category: low risk / very low risk **vs.** intermediate risk classification
- RBC transfusion burden: 1 pRBC/8-weeks over 16 consecutive weeks **vs.** 2- 4 pRBC/8-weeks

NOTE: 1-pRBC randomization will be limited to no more than 30% of the total randomized patient population.

3.4. Blinding

For open-label component and OL high-erythropoietin component, blinding is not applicable.

For double-blind component, any data displays reviewed before database lock will be blinded. Only Sponsor personnel who are not directly involved with the study sites/patients will have access to the unblinded data. The blind will be maintained for study patients, principal Investigator (PI), site staff, site monitoring personnel, sponsor personnel and Contract Research Organization (CRO) personnel who are directly involved in the study conduct with study sites / patients for the remainder of the study (to Week 52 plus a 4-week follow-up period).

A statistician who is not otherwise affiliated with the study will prepare the randomization schedule and the schedule will not be available to study staff until study unblinding. The patients, clinical site staff, FibroGen Medical Monitor and Clinical Operations staff are blinded to whether a patient is randomized to roxadustat or placebo.

Independent Data and Safety Monitoring Board (DSMB) members who review the safety data periodically might be unblinded.

Primary analysis of efficacy for the double-blind component will be conducted when all randomized (DB) patients complete 28 weeks of treatment or are discontinued from study medication before this time. Only the sponsor unblinding team who are not directly involved with the double-blind component will have access to the interim unblinded data, and will no longer participate in blinded trial management or analysis until after the final database lock of double-blind component. Sponsor personnel, sites, and patients who are directly involved in the study conduct will remain blinded to study treatment in the double-blind component until the final study database lock and unblinding. A detailed Data Access Plan for the primary analysis summarizes blinding and unblinding plan for interim analysis.

Treatment assignments will be unblinded after the completion of the study.

4. STUDY ENDPOINTS

4.1. Efficacy Endpoint for Open-Label Component and OL High-Erythropoietin Component

For the Open-Label and OL high-erythropoietin component, the primary efficacy endpoint is the proportion of patients who were TI for≥8 consecutive weeks during the first 28 treatment weeks. Secondary endpoints include TI ≥ 56 consecutive days anytime in the first 52 weeks of treatment, proportion of patients who achieved ≥ 50% reduction in number of RBC transfusion over any 8 weeks compared to baseline, proportion of patients who achieved TI for > 20 weeks.

4.2. Efficacy Endpoint for Randomized, Double-Blind Component

Following the framework in ICH E9 guideline Addendum R1 (Final version Adopted on 20 November 2019) (E9 (R1)), this section details estimand and selected strategies in handling intercurrent events for each primary and secondary endpoint. How to analyze primary and second endpoints are specified in Section 12.3.

4.2.1. Definition of Estimand for Primary Efficacy Endpoint

Population of interest: Male and female patients aged ≥ 18 years and with diagnosis of primary MDS (confirmed histopathologically by bone marrow aspirate and biopsy prior to treatment Day 1), classified by the IPSS-R as very low, low, or intermediate risk with <5% bone marrow blasts and with RBC transfusion requirement of either a. 2 to 4 pRBC during the 8-weeks prior to registration/randomization or b. 1 pRBC during the 8-weeks prior to registration/randomization. There is no restriction on prior use of recombinant erythropoietins or analogues (erythropoiesis-stimulating agents (ESAs)), except that the patient must not have received any ESA within the 8 weeks prior to Day 1 registration/randomization. ESA-naïve patients were allowed to be randomized into the study. Other key eligibility criteria include Hb ≤10.0 g/dL during Screening, , EPO levels ≤400 mIU/mL, No del(5q) cytogenic abnormality, Body weight ≥45 kg and ECOG performance status of 0, 1 or 2 during screening.

Endpoint of interest: RBC TI (transfusion independence) response defined as if patients achieved TI ≥ 56 consecutive days in the first 28 weeks of treatment.

Population-level summary for the endpoint: odds ratio of patients who achieved transfusion TI ≥ 56 consecutive days in the first 28 weeks of treatment between treatment and placebo arm.

Handling of intercurrent events:

Treatment discontinuation (such as: due to Death, Disease progression, lack of efficacy, impaired Liver function, Patient decision, Investigator decision, AE/SAE, etc.) uses ‘on treatment strategy’ in which to assess the response to treatments of interest while the patient is exposed to treatment All administered RBC transfusion records from the first dose to the end of treatment will be used to define transfusion independence.

4.2.2. Definition of Estimand for Secondary Efficacy Endpoints

Population of interest: The same population as defined in Primary Efficacy Endpoint.

Table 1: Other Estimand Attributes for Secondary Efficacy Endpoints

Table Endpoint of interest	Population-level summary for the endpoint	Handling of intercurrent events	
		Treatment discontinuation*	**Death**
RBC TI (transfusion independence) response defined as if patients achieved TI ≥ 56 consecutive days in the first 52 weeks of treatment.	Odds ratio of patients who achieved transfusion TI ≥ 56 consecutive days in the first 52 weeks of treatment between treatment and placebo arm.	While on treatment strategy: response to treatment prior to the occurrence of treatment discontinuation	While on treatment strategy: response to treatment prior to death.
RBC TI (transfusion independence) response defined as if patients achieved TI ≥ 56 consecutive days anytime during the study	Odds ratio of patients who achieved transfusion TI ≥ 56 consecutive days anytime during the study between treatment and placebo arm	Treatment-policy strategy: regardless of these intercurrent events occurred;	Treatment-policy strategy: regardless of these intercurrent events occurred.
Reduction in number of pRBC transfusions over 8 weeks compared to their baseline for any 8-week period during the study	Odds ratio of patients who achieved ⩾50% reduction in number of pRBC transfusions over 8 weeks compared to their baseline for any 8-week period during the study between treatment and placebo arm	Treatment-policy strategy: regardless of these intercurrent events occurred;	Treatment-policy strategy: regardless of these intercurrent events occurred.
Patient-exposure-weeks (PEW) of TI over the first 28 weeks of the treatment period	Difference in PEW of TI over the first 28 weeks of the treatment period between treatment and placebo arm	While on treatment strategy: treatment effect prior to the occurrence of treatment discontinuation	While on treatment strategy: Treatment effect while on treatment
Number of pRBC packs transfused compared to baseline over the first 28 weeks of the treatment period	Difference in mean of pRBC transfused change from baseline over the first 28 weeks of the treatment period between treatment and placebo arm	While on treatment strategy: treatment effect prior to the occurrence of treatment discontinuation	While on treatment strategy: Treatment effect while on-treatment

RBC TI (transfusion independence) response defined as if patients achieved TI ≥ 140 consecutive days anytime during the study	Odds ratio of patients who achieved transfusion TI ≥ 140 consecutive days anytime during the study between treatment and placebo arm	Treatment-policy strategy: regardless of these intercurrent events occurred;	Treatment-policy strategy: regardless of these intercurrent events occurred.
Change from baseline in the PROMIS-SF v2.0 Physical Function 10b score at week 9	Difference in mean change from baseline in the PROMIS-SF v2.0 Physical Function 10b score at week 9 between treatment and placebo arm	**Due to reasons other than AE/SAE:** Treatment-policy strategy: regardless of these intercurrent events occurred; **Due to AE/SAE:** Hypothetical strategy: A scenario is envisaged after the intercurrent event occurred;	Treatment-policy strategy: regardless of these intercurrent events occurred.
Change from baseline in the PROMIS-SF v1.0 Fatigue 13a score at week 9	Difference in mean change from baseline in the PROMIS-SF v1.0 Fatigue 13a scores at week 9 between treatment and placebo arm	**Due to reasons other than AE/SAE:** Treatment-policy strategy: regardless of these intercurrent events occurred; **Due to AE/SAE:** Hypothetical strategy: A scenario is envisaged after the intercurrent event occurred;	Treatment-policy strategy: regardless of these intercurrent events occurred.
Change from baseline in EQ-5D-5L VAS score at week 9	Difference in mean change from baseline in EQ-5D-5L VAS score at week 9 between treatment and placebo arm	**Due to reasons other than AE/SAE:** Treatment-policy strategy: regardless of these intercurrent events occurred; **Due to AE/SAE:** Hypothetical strategy: A scenario is envisaged after the intercurrent event occurred;	Treatment-policy strategy: regardless of these intercurrent events occurred.

⁎Such as due to Disease progression, Liver function, Patient decision, Investigator decision, AE/SAE, etc.

4.2.3. Exploratory Endpoints

The exploratory efficacy endpoints include:

- Maximum Duration of TI
- Proportion of patients who achieve a mean Hb increase of ≥ 1.0 g/dL and ≥ 1.5 g/dL (averaged over any 8 weeks in those achieved TI) compared to pre-transfusion Hb at baseline
- Impact of baseline endogenous erythropoietin levels on roxadustat treatment response and dose requirement
- Effect on hepcidin and iron indices metabolism (e.g., serum iron, ferritin, TIBC, TSAT, and UIBC)
- Effect on cholesterol and lipid parameters

4.3. Safety Assessments

For both open-label and double-blind components, safety will be assessed by evaluating the following over 52 weeks:

- AE and SAE reporting
- N (%) of patients with progression to acute myeloid leukemia (AML)
- Clinical laboratory measures including Hb and liver chemistries (AST, ALT, Total bilirubin, and alkaline phosphatase)
- Vital signs (heart rate, blood pressure, respiratory rate, and temperature) and ECG
- Evaluation of PK/PD data at the end of study

An independent Data and Safety Monitoring Board (DSMB) assesses the safety profile during study conduct.

5. GENERAL STATISTICAL CONSIDERATIONS

5.1. General Presentation of Summaries and Analyses

Continuous variables will be presented by descriptive statistics: n, mean, standard deviation (standard error, as needed), median, minimum, and maximum. Categorical variables will be presented by counts of patients and percentage. Unless otherwise specified, the denominator for all percentages will be the number of patients in the population of interest. For plasma concentrations and PK variables, coefficient of variation (CV%) and geometric mean will be additionally provided.

Any p-values will be rounded to four decimal places and will be presented as '<.0001' if they are less than 0.0001 after rounding.

For continuous variables that are recorded as "< X" or "> X", the value of "X" will be used in the calculation of summary statistics. The original values will be used for the listings.

In this study, eGFR will be calculated by the central lab using the Cockcroft-Gault (CG) formula as described in protocol Appendix L.

Decimal points will be presented as follows: N will be presented without decimal, CV with one decimal, minimum/maximum in same precision as in the database, mean/median in one more decimal than minimum/maximum, and SD in one more decimal than mean/median.

Analyses will be conducted using SAS® Version 9.3 or higher. All tables and listings will have a header showing "FibroGen, Inc." and the protocol number. A footer will show the program file name and path, run date, run time, and output file name and path.

The secondary endpoints will be tested sequentially, the family-wise type I error rate will be controlled using the method described in Section 12.1.

All individual data will be presented in individual patient data listings.

Data from the open-label and double-blind component of the study will be analyzed separately. Patients enrolled in open-label component may remain in the study up to 52 weeks and will undergo all procedures as will the randomized patients (except Health-Related Quality of Life (HRQoL) assessments including PROMIS and EQ-5D-5L), but the data will be included in the analysis of open-label component only.

For the double-blind component, data will be analyzed for the first 28 weeks of treatment and for the end of the 52-week study treatment period respectively.

Patients from the OL high-erythropoietin component will undergo the same procedures as will patients from the OL component with the exception of completing the HRQoL questionnaires. Data from the OL high-erythropoietin component will be analyzed along with the OL component. HRQoL data from the OL high-erythropoietin component will be listed. The final analysis will be done for the end of the 52-week study treatment period.

5.2. Analysis Population for Open-Label Component and OL High-Erythropoietin Component

5.2.1. Full Analysis Set (FAS)

The full analysis set consists of all patients enrolled into this study who receive at least one dose of study medication, and at least one corresponding on-treatment Hb assessment.

5.2.2. Safety Analysis Set (SAF)

The safety analysis set consists of all patients who have received at least one dose of study drug. Analyses of the safety data will be based on the safety analysis set.

5.3. Analysis Population for Double-Blind Component

5.3.1. Full Analysis Set (FAS)

The full analysis set consists of all patients randomized to this study who receive at least one dose of study medication, and at least one corresponding on-treatment Hb assessment. Patients will be included in the treatment group to which they are randomized for the FAS analysis.

5.3.2. Safety Analysis Set (SAF)

The safety analysis set consists of all patients who are randomized to this study and have received at least one dose of study drug. Analyses of the safety data will be based on the SAF. Patients will be included in the treatment group based on treatment they actually received.

5.3.3. Per-Protocol Set (PPS)

The Per-Protocol Set (PPS) consists of patients who have received at least 8 weeks of treatment with corresponding on-treatment Hb assessments and are without major protocol deviations that affect the assessment of study endpoints. The PPS population will be used for supportive and/or sensitivity analyses of the efficacy endpoints.

5.4. Methods of Data Transformation and Handling of Missing Data

All data collected and recorded in the study database will be eligible for inclusion in the analysis. No substitutions for missing data will be made unless otherwise specified in this SAP, except for dates that occur prior to a patient being enrolled in the study (e.g., medications received before screening). For these date imputations, refer to Appendix 19.3.

Some lab data are collected using local labs (in lieu of the ICON central lab) due to central lab kit shortages caused by the COVID-19 pandemic. In these cases, local lab values and reference ranges collected from CRFs will be presented in the listings along with central lab data when appropriate and feasible. Local lab values will be flagged in datasets and data listings to be differentiated from central lab values. For summary analyses of Hemoglobin efficacy and laboratory safety in tables and figures, if central lab is not available at a scheduled visit and local lab is available, the local lab will be used for substitution.

5.5. Baseline Measures and Change from Baseline

For patients who achieved ≥ 50% reduction in number of pRBC transfusions over any 8 weeks (day 1 to week 52) compared to their baseline, the baseline will be calculated based on the transfusions within 8 or 16 weeks prior to the first dose.

For subjects with transfusion records within 16 weeks prior to first dose: Baseline number of transfusions (pRBC/8-weeks) = total number of packs of red blood cells within 16 weeks prior to first dose/2.

For subjects with transfusion records within only 8 weeks prior to first dose: Baseline number of transfusions (pRBC/8-weeks) = total number of packs of red blood cells within 8 weeks prior to first dose.

For other measurements, baseline value will be defined as the last measure taken before any study drug administration unless otherwise specified.

Change from baseline is defined as the difference between the baseline value obtained and the respective post-baseline time point for the continuous variables. Negative values represent decreases from baseline; positive values reflect increases from baseline.

Percent change from Baseline is defined as the change from Baseline divided by the Baseline value，multiplied by 100:

100 x (Post-Baseline value minus Baseline value) / (Baseline value)

5.6. Randomization Errors

Patients who were randomized from an incorrect stratum will be identified by comparing the strata recorded on the RTSM system (Randomization and Trial Supply Management)/randomization CRF page with the strata derived from screening data. A listing of patients who were randomized from an incorrect stratum will be presented. The randomized stratum (from RTSM system/randomization CRF) rather than actual stratum (derived from screening data) will be used for all the efficacy analyses. A sensitivity analysis will be performed using the actual stratum on the primary endpoint if the discrepancy rate is greater than 5% in the final result.

5.7. Other Key Definitions and Computations

Study Day

Study day will be calculated in reference to the date of the first dose (Study Day 1, which refers to the enrollment/randomization in the protocol). For assessments conducted on or after the first dose date, study day is calculated as (assessment date – first dose date + 1). For assessments conducted before the first dose date, study day is calculated as (assessment date – first dose date). There will be no Study Day 0.

In the listings, study days will be presented along with assessment date where applicable.

Date of First Dose and Date of Last Dose of Study Drug

The date of the first dose of study drug is defined as the date a patient receives the first dose of study drug (roxadustat or placebo). The date of the last dose of study drug is defined as the date a patient receives the last dose of study drug (roxadustat or placebo).

Treatment Day

Treatment day will be calculated in reference to the date of the first dose of study drug. Treatment Day 1 corresponds to the date a patient receives the first dose of study drug. For assessments conducted on or after the date of the first dose of study drug, treatment day will be calculated as (assessment date – date of first dose of study drug + 1). There will be no Treatment Day 0.

Treatment Period

The treatment period is defined as the period of time from the date and time of the first dose of study drug through the date of last dose.

6. PATIENT ENROLLMENT AND DISPOSITION

6.1. Eligibility Criteria

Eligibility criteria will be summarized for all screened subjects in open-label component and double-blind component. The data will be summarized with respect to:

- number of subjects screened
- number (%) of subjects screen-failed
- number (%) of subjects for each failed inclusion/exclusion criterion

Subject level inclusion criteria not met/exclusion criteria met listings will be provided.

6.2. Subject Accountability and Disposition

For open-label component, patient disposition will be summarized by dose level cohort and overall. For OL high-erythropoietin component, the summaries will be presented along with the cohorts from open-label component. For double-blind component, patient disposition will be summarized by treatment group and overall.

The data will be summarized with respect to:

- number of patients screened,
- number (%) of patients screen-failed (and by major reason),
- number of patients enrolled (defined as patients who signed informed consent form and received treatment assignment),
- number (%) of patients randomized, not treated (double-blind component),
- number (%) of patients randomized, treated (double-blind component),
- number (%) of patient completed study treatment (Note: for 28-week analysis, the summary will use the data in the first 28 weeks of treatment period),
- number (%) of patients prematurely discontinued from study treatment by specific reason as recorded on the CRF,
- number (%) of patients prematurely discontinued from the study by specific reason as recorded on the CRF.

Percent of screen-failed will be based on the number of patients screened. All other percentages for patients will be based on the number of patients enrolled.

In addition, the number and percentage of patients included in the FAS, PPS and SAF will be summarized. Percentage will be based on the number of patients enrolled.

Number and percentage of subjects with early discontinuation of study medication by demographics and baseline characteristics will be summarized in the SAF.

Consort diagrams describing the disposition of patients in the SAF will be provided.

All patients who prematurely discontinued from treatment or study will be listed by discontinuation reason.

7. PROTOCOL DEVIATIONS

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

- Patients who did not meet inclusion/exclusion criteria
- Patients who received disallowed concomitant medications or non-drug therapy during study treatment period
- Patients who took IV iron while on study

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 study unblinding. Considerations will be given according to the following table. The number and percentage of major protocol deviations will be categorized and tabulated as appropriate.

Table 2: Criteria for Assessing Major Protocol Deviations Table

Number	Major Protocol Deviation
1	Violation of inclusion or exclusion criteria which may affect the assessment of the efficacy of the study drug
2	Administration of prohibited concomitant medication or non-drug therapy
3	Significant noncompliance with study procedures, e.g., a patient skipping consecutive scheduled assessments, will be evaluated case by base
4	Developed criteria for discontinuation of study drug but did not discontinue study drug treatment
5	Received the wrong treatment or incorrect dose resulting in any adverse event
6	Were incorrectly stratified by the study investigators

8. DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

For categorical data, frequencies and percentages will be provided and, for continuous data, descriptive statistics, including sample size (n), mean, median, standard deviation, and range of values (i.e., minimum and maximum values) will be provided. All information to be summarized will also be presented in listings.

Comparability of baseline characteristics among treatment groups will be tested using analysis of variance (ANOVA) model for continuous variables and Chi-Square test (or Fisher's exact test if cell count < 5) for categorical variables for reference only. The p-values are exploratory and no formal statistical inference will be made.

For open-label component, the summaries will be presented by dose level cohort and overall using safety population. For OL high-erythropoietin component, the summaries will be presented along with the cohorts from open-label component using safety population. For double-blind component, the summaries will be presented by treatment group and overall using FAS, Safety and PPS. The following baseline characteristics will be summarized and listed:

- age
- age group (< 65 / >= 65 years)
- sex
- ethnicity
- race
- weight
- baseline Serum EPO level
- baseline Serum EPO level category [<200 (IU/L), 200-400 (IU/L), >400 (IU/L)]
- serum EPO level category based on actual strata [<200 (IU/L), 200-400 (IU/L)]
- pRBC-transfusion burden at baseline
- baseline pRBC-transfusion burden group (1 pRBC/8-weeks vs. 2-4 pRBC/8-weeks over 16 consecutive weeks)
- pRBC-transfusion burden based on actual strata (1 pRBC/8-weeks vs. 2-4 pRBC/8-weeks over 16 consecutive weeks)
- baseline ECOG performance status
- MDS duration (based on medical history)
- WHO Classification of MDS
- pre-transfusion hemoglobin (Hb) at baseline
- baseline platelet counts
- cytogenetic actual category
- Anticoagulants use at baseline (Yes/No) - if applicable.
- bone marrow blast
- IPSS-R risk category (low risk / very low risk **vs.** intermediate risk classification)

- eGFR
- baseline C-reactive protein (CRP)
- CRP group (CRP ≤ULN vs. CRP > ULN)
- Ferritin and Ferritin group (< 100 ng/mL, >= 100 ng/mL and < 400 ng/mL, >= 400 ng/mL)
- TSAT and TSAT group (< 20%, >= 20% and < 40%, >= 40%),
- iron repletion status ([TSAT <20% or ferritin < 100 ng/mL] vs. others)

MDS duration will be calculated from the earliest date documented on the medial history CRF to Day 1.

9. MEDICAL HISTORY

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA, version 19.1). A summary table will be presented by system organ class (SOC) and preferred term (PT).

For open-label component, the summaries will be presented by dose level cohort using safety population. For OL high-erythropoietin component, the summaries will be presented along with the cohorts from open-label component using safety population. For double-blind component, the summaries will be presented by treatment group using FAS.

10. EXTENT OF EXPOSURE TO STUDY MEDICATION

Extent of exposure, treatment compliance and dose modification will be summarized for the safety population.

For open-label component, the summaries will be presented by dose level cohort. For double-blind component, the summaries will be presented by treatment group. For extent of exposure and treatment compliance, the summaries will be presented up to Week 28 and Week 52/EOT separately. For OL high-erythropoietin component, the summaries will be presented along with the cohorts from open-label component.

10.1. Extent of Exposure

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

Duration (in weeks) = (date of last dose - date of first dose + 1) / 7

Total actual dose (in mg or mg/kg) for dose interval k is calculated as

*dose strength in interval k * (last day of dose interval k - first day of dose interval k + 3 days) * (frequency for dose interval k/7)*

The first and last day of dose interval will be obtained from IP dosing form. The first day of dose interval k is the first date of a new doing level. For example, the dose of 50 mg TIW starts on 01/07/2017, ends on 02/10/2017, the total dose is 50*(33+3)*(3/7) = 771.4 mg. The average weekly dose is calculated as the sum of all actual doses over the treatment period (in mg) divided by actual exposure duration (in weeks) during the period. The average weekly dose per weight is calculated as the sum of all actual doses over the treatment period per weight (in mg/kg) divided by actual exposure duration (in weeks) during the period.

Duration of exposure and average weekly exposure over time will be tabulated for the safety population. Exposure duration by category (<2 weeks, 2 weeks - <8 weeks, 8 weeks - <=28 weeks) will be summarized for Week 28 analysis. Subject exposure over time (probability of subjects still on treatment) up to Week 28 and up to Week 52 will be plotted by treatment groups.

10.2. Treatment Compliance

Study medication compliance will be calculated as the total amount of doses actually administered during that period divided by the total amount of doses prescribed expected to be taken during the same period (e.g. Day1 to Last dose) multiplied 100.

Total prescribed dose (in mg or mg/kg) for dose interval k is calculated as

*dose strength in interval k * (last day of dose interval k - first day of dose interval k + 3 days) * (frequency for dose interval k/7)*

The first and last day of dose interval will be obtained from Study Drug Dispensation form. The first day of dose interval is the first day of a new prescribed dosing level. The last day of dose interval is the day before the initial day of the next new prescribed dosing level, if it is the last dosing level, the last day of this dosing level will be the last treatment day.

If there is a dose interruption period and the reason is “PI decision”, the amount of doses in the interruption period should be excluded from both the numerator and the denominator in the compliance calculation.

Descriptive statistics for study medication compliance will be presented by treatment group. A frequency summary of treatment compliance category (<75%, >=75% to < 85%, >=85% to <= 100% and >100%) will also be provided.

Date of first/last dose, duration of treatment, dates of discontinuation of treatment and study, and reason for study drug discontinuation will be listed for patients who discontinued the study drug prematurely.

10.3. Dose Modification

Dose Escalation

Following commencement of study treatment, patients are evaluated for response after 8 weeks (beginning at Week 9) and, if the patient has required one or more RBC transfusion in that period then the dose is increased by one dose level based on the body weight category. After each dose adjustment, the patient remains on the dose level for a minimum of 8 weeks before further adjustment is permitted based on the next scheduled visit.

Dose Reduction

The down titration of the study treatment is allowed at any time but must be performed if any of the following criteria are met. Hb levels should be retested for confirmation every 2 weeks after a dose reduction; however, it is per investigator's discretion:

a. Transfusion independent for ≥ 56 days consecutively and Hb ≥ 12 g/dL; or
b. Transfusion independent for ≥ 56 days consecutively and the rate of rise of Hb ≥ 2.5 g/dL over 4 weeks.
c. In the instance where the patient achieved TI in the initial 8-weeks of the study, or has undergone treatment with the maximum allowable dose, i.e., 400 mg TIW or 3.5mg/kg for at least 8-weeks, patient may undergo dose reduction by 1-dose level. If there were no difference in erythroid response (same transfusion burden and Hb was unchanged) at this dose level for 8-weeks compared to another lower dose level, investigator may contact Sponsor's Medical Monitor to discuss another potential dose reduction.

Dose adjustment rules for patients who have TI for ≥ 56 Days (Maintenance Phase) can be referenced to Section 19.1.

The number of dose escalations will be summarized descriptively.

The number and percentage of patients who exceed maximum dose will be summarized. The maximum dose of study medication is capped at 3.5 mg/kg per dose or 400 mg, whichever is lower.

The number and percentage of patients with a prescribed dose increase and decrease will be summarized. The number and percentage of patients with an actual dose hold, increase and decrease will be summarized.

The summaries above will be presented by every cumulative 4-weeks, e.g., up to 4 weeks, up to 8 weeks, …, up to 52 weeks and up to 52 weeks and beyond.

Dose modifications by primary investigator or by patient collected on the eCRF will be listed. Hb level meeting the dose adjustment rules will be combined in the listing.

11. PRIOR AND CONCOMITANT MEDICATIONS

For open-label component, the summaries will be presented by dose level cohort.

For OL high-erythropoietin component, the summaries will be presented along with the cohorts from open-label component.

For double-blind component, the summarizes will be presented by treatment group.

The World Health Organization Drug Dictionary (WHODD) Version 2015 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 administration 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

End date / Start Date	Before start of study drug administration	On or after start of study drug administration	Missing
Before start of study drug administration	Prior	Concomitant	Concomitant
On or after start of study drug administration	N/A	Concomitant	Concomitant
Missing	Prior	Concomitant	Concomitant

Both prior and concomitant medication usage will be summarized by the number and proportion of patients receiving each drug within each therapeutic class using the safety population. Multiple drug usage in the same therapeutic class and preferred term by a patient will be counted only once.

A summary for prior ESA use and ESA Refractory, will be presented for double-blind component.

Detailed analyses may be performed on prior and concomitant medications of special interests such as oral iron, statin, phosphate binders and anticoagulants based on TLF shells.

Procedures and non-drug therapies will be summarized by the number and percentage. Detailed analyses may be performed on procedures and non-drug therapies of special interest such as bone marrow transplant.

12. EFFICACY ANALYSES

12.1. Multiple Comparisons/Multiplicity

For open-label component and OL high-erythropoietin component, there will be no statistical inference intended; hence, no multiplicity adjustment is needed.

For double-blind component, the statistical test results of the primary and secondary endpoints in the first 28-week treatment period (or at Week 28) and in the 52-weeks of treatment period (or at Week 52) will be used to determine treatment efficacy of roxadustat. To control the family-wise type I error at the alpha level of 0.05, a sequential gatekeeping procedure will be used to test the primary and secondary efficacy endpoints. Specifically, if the test of the primary endpoint is significant at the level of 0.05, the first secondary endpoint will be tested at the level of 0.05. Other secondary endpoints will be tested conditioning on the previous endpoint achieving the statistical significance, in the order listed in Section 12.3.3. Statistical testing will stop at the first endpoint that fails to achieve the significance level. All statistical tests will be two-sided at the significance level of 0.05.

12.2. Analysis of Open-Label and OL High-Erythropoietin Data

For the Open-Label and OL high-erythropoietin component, the primary efficacy endpoint is the proportion of patients who were TI for≥8 consecutive weeks during the first 28 treatment weeks. Secondary endpoints include TI ≥ 56 consecutive days anytime in the first 52 weeks of treatment, proportion of patients who achieved ≥ 50% reduction in number of RBC transfusion over any 8 weeks compared to baseline, proportion of patients who achieved TI for > 20 weeks. These estimands are defined in Sections 12.3.1 and 12.3.3.2.

Because of the small sample sizes for the OL component and OL high erythropoietin component, descriptive analyses with frequency tables will be provided.

A frequency table of patients with at least a 50% reduction in the number of pRBC transfusions over any 8-week (56 consecutive days) period during the study as compared with the baseline will be provided by dose level cohort. Similar strategies as described in Section 12.3.3.2 will be implemented for potential intercurrent events.

A frequency table of patients who achieved TI ≥ 56 consecutive days (e.g., 8 weeks) since first dose in the first 28-weeks of treatment will be provided by dose level cohort. TI is defined as the absence of any intravenous RBC transfusion (packed cell or whole blood) during any consecutive 56 days during the treatment period (e.g., Days 1 to 56, Days 2 to 57, Days 3 to 58). Similar strategies as described in Section 12.3.1.1 will be implemented for potential intercurrent events.

A by-patient listing of transfusion records will also be provided.

12.3. Analysis of Double-Blind Efficacy Endpoints

12.3.1. Analysis of Primary Efficacy Endpoint

The primary efficacy endpoint for this study is proportion of patients who achieved TI ≥ 56 consecutive days (e.g., 8 weeks) since first dose in the first 28-week of treatment. TI is defined as the absence of any intravenous RBC transfusion (packed cell or whole blood) during any consecutive 56 days during the treatment period (e.g., Days 1 to 56, Days 2 to 57, Days 3 to 58). All RBC transfusions from 8 to16 weeks prior to first dose until end of study will be recorded on CRF. TI will be estimated in the duration beginning with the first dose date (Day 1) and ending with the date of end of treatment or last visit date at

week 28, whichever comes earlier. RBC transfusions after end of treatment will be recorded on the CRF until the last visit at week 52 from first dose but will not be used to define TI.

The endpoint will be also analyzed for the end of the

52-week study treatment period as a secondary endpoint.

12.3.1.1. Primary Analysis

The RBC TI response rate and its two-sided 95% confidence interval based on the exact method of Clopper-Pearson will be presented in a frequency table by treatment group using the FAS.

The primary objective of the study will be evaluated to demonstrate a statistically significant difference between roxadustat and placebo group with respect to the RBC transfusion independence rate based on the Cochran-Mantel-Haenszel (CMH) chi-square test adjusting for the stratification factors (EPO level, IPSS-R risk category and RBC transfusion burden). A CMH-adjusted odds ratio (roxadustat versus placebo) and its associated 95% CI will be computed to compare between the roxadustat group and placebo group.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation (such as: due to Disease progression, Liver function, Patient decision, Investigator decision, AE/SAE, etc.): using while on treatment strategy: Patients who have discontinued treatment early for any reasons prior to week 8 will be classified as non-responders. For patients who have discontinued treatment early for any reasons after week 8, TI will be estimated in the duration beginning with the first dose date (Day 1) and ending with the date of end of treatment or last visit date at week 28, whichever comes earlier.

2. Death: using While on treatment strategy: Treatment effect while on treatment is of interest. Patients who died prior to week 8 will be classified as non-responders. For patients who died after week 8, the RBC TI response will be based on RBC transfusion records up to the date of end of treatment or last visit date at week 28 or death, whichever comes earlier.

The number of periods of 8 weeks (56 consecutive days) that meet the criterion of TI per TI responder will be summarized descriptively.

Hypothesis testing:

The odds ratio (OR) in treatment response (roxadustat vs. placebo) is given by:

$OR = (n_1*(N_2-n_2))/n_2*(N_1-n_1))$

Where:

N_1 = the number of patients in roxadustat group

N_2 = the number of patients in placebo group

n_1 = the number of patients who achieved TI $\geqslant$ 56 consecutive days (e.g., 8 weeks) since first dose in the first 28-weeks of treatment in roxadustat group

n_2 = the number of patients who achieved TI $\geqslant$ 56 consecutive days (e.g., 8 weeks) since first dose in the first 28-weeks of treatment in placebo group

The null and alternative hypotheses that will be tested at 2-sided significant level 5% are:

H_0: OR = 1 versus Ha: OR $\neq$ 1

The p-value from the stratified CMH chi-square test will be the confirmatory p-value for the test of the null hypothesis that the proportion of patients achieving RBC transfusion independence is equal between the two treatment groups.

SAS code for Cochran-Mantel-Haenszel (CMH) test

The two-sided CMH test will be performed using SAS procedure PROC FREQ with the CMH option in the TABLES statement. Baseline EPO level, IPSS-R risk category and RBC transfusion burden will be the stratification variables. The variables <treatment> and <response> determine the rows and columns of the tables.

Sample code:

```
PROC FREQ DATA = <input_dataset sorted> NOPRINT;
    TABLES <strata_var>* <treatment_var> * <response_var> / CMH EXACT RISKDIFF;
RUN;
```

Handling of missing values in primary analysis

Patients who have discontinued treatment early for any reasons prior to week 8 will be classified as non-responders. If significant differential treatment discontinuation prior to week 8 is observed between the two treatment groups, additional sensitivity analyses may be explored.

Sensitivity Analysis

Sensitivity Analysis 1: PPS Analysis

The primary analysis will be repeated using PPS as a sensitivity analysis.

Sensitivity Analysis 2: Logistic Regression Analysis

The primary analysis for primary endpoint will also be repeated using a multivariate logistic regression. The following variables/factors will be used in the model:

- treatment group (roxadustat vs. placebo);
- age group (< 65 vs. >=65 years);
- centrally confirmed IPSS-R category (Intermediate vs. Low-risk/very low risk);
- transfusion burden (1 pRBC/8-weeks over 16 consecutive weeks vs. 2-4 pRBC/8-weeks);
- baseline hemoglobin
- baseline EPO level (≤200 mIU/mL vs. >200 mIU/mL and ≤400 mIU/mL);

In addition, MDS duration, baseline platelet counts, and bone marrow blasts will be also investigated as continuous variables.

A Wald test based on the estimated regression parameter for treatment group assignment will provide the basis for the hypothesis testing. The odds ratio (roxadustat versus placebo) and its 95% confidence interval will be provided. An example of the SAS code is displayed below:

```
PROC LOGISTIC;
    CLASS <treatment_var> <factors>;
    MODEL <response_var> = <treatment_var> <factors>;
RUN;
```

An adjusted difference in proportions and its 95% Confidence Interval from the same logistic regression model will be presented using the method proposed by Ge et al. (2011). See Appendix 19.7 for SAS sample code.

A swimming plot will be provided to display the RBC transfusions and TI periods during treatment period. The RBC TI response rates over longer periods (at least 8 consecutive weeks, at least 12 consecutive weeks, at least 16 consecutive weeks, and at least 20 consecutive weeks) during treatment will be plotted by treatment groups, in bar charts, with p-values calculated by the Cochran-Mantel-Haenszel (CMH) chi-square test adjusting for the stratification factors (EPO level, IPSS-R risk category and RBC transfusion burden).

The analyses in this section will use the FAS population.

12.3.2. Sequential Gatekeeping Procedure

Once the primary hypothesis has been rejected for the primary endpoint, the primary endpoint at week 52 and the secondary endpoints in the order specified in the Table 4 will be tested using a sequential gatekeeping procedure, in order to maintain the overall two-sided type I error of 0.05. If p-value from a test is < 0.05, the claim of superiority will be considered successful, and the test will progress to the next comparison in sequence as follows.

Table 4: Sequential Gatekeeping Procedure

Test	Variable	Comparison
1	Proportion of patients who achieved TI ≥56 consecutive days (e.g., 8 weeks) since first dose in the first 28-weeks of treatment	Superiority of roxadustat versus placebo
2	Proportion of patients who achieved TI ≥56 consecutive days (e.g., 8 weeks) since first dose in the 52-weeks of treatment	Superiority of roxadustat versus placebo
3	Proportion of patients who achieved TI ≥56 consecutive days (e.g., 8 weeks) anytime during the study	Superiority of roxadustat versus placebo
4	Proportion of patients who achieved ≥50% reduction in number of pRBC transfusions over 8 weeks compared to their Baseline for any 8-week period during the study	Superiority of roxadustat versus placebo
5	Cumulative number of patient-exposure-weeks (PEW) of TI over the first 28 weeks of the treatment period	Superiority of roxadustat versus placebo
6	Number of pRBC transfused change from baseline over the first 28 weeks of the treatment period	Superiority of roxadustat versus placebo
7	Proportion of patients who achieved TI for ≥20 consecutive weeks (140 consecutive Days)	Superiority of roxadustat versus placebo

8	Mean change from baseline in the PROMIS-SF v2.0 Physical Function 10b score at week 9	Superiority of roxadustat versus placebo
9	Mean change from baseline in the PROMIS-SF v1.0 Fatigue 13a score at week 9	Superiority of roxadustat versus placebo
10	Mean change from baseline in EQ-5D-5L VAS score at week 9	Superiority of roxadustat versus placebo

12.3.3. Analysis of Secondary Efficacy Endpoints

All secondary endpoints measured for treatment periods will be analyzed both for the first 28 weeks of treatment and for 52-week study treatment period. Others will be analyzed for the time during the study. The secondary efficacy endpoints will be analyzed based on the FAS population unless it is specified otherwise. Analyses using the PPS will be used as sensitivity analyses.

12.3.3.1. Proportion of Patients who achieved TI ⩾56 consecutive days anytime during the study

The proportion of patients who achieved TI for ≥ 56 consecutive days will be summarized using the FAS and PPS and will be compared between the two treatment groups. TI is defined as the absence of any intravenous RBC transfusion (packed cell or whole blood) during any consecutive 56 days anytime during the study. TI will be estimated in the duration beginning with the first dose date (Day 1) and ending with the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier. RBC transfusions after end of treatment will be recorded on CRF until the last visit at week 52 from first dose and will be used to define TI.

The p-value obtained will be reported for the test of the null hypothesis that there is no difference between the two treatment groups.

Similar methods will be used as for the primary endpoint.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation due to any reasons: using Treatment-policy strategy: regardless of these intercurrent events occurred; the TI will be estimated using RBC transfusion records in the whole study duration.

2. Death: using Treatment-policy strategy: regardless of these intercurrent events occurred; Patients who died prior to week 8 will be classified as non-responders. TI will be estimated using RBC transfusion records for the study duration up to death.

12.3.3.2. Proportion of Patients Who Achieved ≥ 50% Reduction in Number of pRBC Transfusions over 8 Weeks Compared to Their Baseline for Any 8 Week Period during the Study (Factor of Erythroid Response)

See Section 5.5 for the definition of number of pRBC (packs of red blood cells) transfusions at baseline. Percent change from baseline for any of the post-baseline 8-week period will be rounded to one decimal place (i.e., xx.x %) and then compared to 50% to determine whether the required percentage of reduction is achieved for that period. A frequency table of patients with at least a 50% reduction in the number of pRBC transfusions over any 8-week period in the duration beginning with the first dose date (Day 1) and ending with the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier as compared with the baseline will be provided. The response rate and its two-sided 95% confidence intervals based on the exact method of Clopper-Pearson will be presented in the table by treatment group.

Treatment comparison will be based on the Cochran-Mantel-Haenszel (CMH) chi-square test, controlling for stratification factors (Serum EPO level, IPSS-R risk category and pRBC transfusion burden).

Two-sided 95% confidence interval (CI) of the odds ratio will also be presented.

The p value from the stratified CMH chi-square test will be provided for the test of the null hypothesis that the proportion of patients with at least a 50% reduction is equal between the two treatment groups.

The number of periods of 8 weeks (e.g., consecutive 56 days) per responder will be summarized descriptively.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation due to reasons: using Treatment-policy strategy: regardless of these intercurrent events occurred; the RBC transfusion reduction will be estimated using RBC transfusion records in the whole study duration.

2. Death: using Treatment-policy strategy: regardless of these intercurrent events occurred;. Patients who died prior to week 8 will be classified as non-responders. RBC transfusion will be estimated for any 8-week period in the study duration up to death.

PPS analysis will be used as sensitivity analysis.

12.3.3.3. Cumulative Number of Patient-Exposure-Weeks (PEW) of TI over the first 28 Weeks of the Treatment Period

The cumulative number of patient-exposure-weeks per patient of TI over the first 28-weeks of the treatment period will be summarized and analyzed using the FAS and PPS population.

The exposure time will be measured using patient-exposure-week (PEW). PEW of TI periods over the first 28-weeks will be added up to a cumulative number of weeks. For a patient with at least one TI response period over the first 28-weeks, the last TI response period will end with the date of a subsequent RBC transfusion, visit date at week 28, date of the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier. For a patient with no TI response period over the first 28 weeks, the cumulative number of PEW will be set to zero.

The cumulative number of PEW over the first 28 weeks of treatment period will be summarized using descriptive statistics by treatment group. Treatment comparison will be based on a non-parametric ANCOVA (stratified rank ANCOVA) described in Stokes, Davis, and Koch 2000. The analysis will use baseline hemoglobin as covariate, treatment arm as fixed effects, controlling for stratification factors (baseline endogenous EPO level, IPSS-R risk category and RBC transfusion burden). See Section 5.5 for the definition of number of pRBC (packs of red blood cells) transfusions at baseline.

Since it is expected that the sample sizes are not large for all strata, the following steps and SAS codes are to be used:

Step 1: Compute the ranks of the response variable and covariate across the two treatment groups.

Sample SAS code:

```
PROC RANK TIES = mean OUT = ranks;
        BY <strata_var>;
        VAR <baseline_var> <response_var>;
RUN;
```

Step 2: With the ranks of response variable, an ANCOVA model will be fitted including the baseline ranks and stratification factor as covariates.

Sample SAS code:

```
PROC MIXED DATA = ranks;
        CLASS <strata_var>;
        MODEL <R_response_var> = <R_Baseline_var> <strata_var>/OUTP = res;
    RUN;
```

Step 3: CMH mean score statistics will be used to compare the two treatment groups using the values of the residuals obtained from step 2 as scores. The p-value obtained will be reported.

Sample SAS code:

```
PROC FREQ DATA = res;
        TABLES <strata_var>* <treatment_var>*resid/CMH2;
RUN;
```

The p-value from the stratified CMH chi-square will be provided for the test of the null hypothesis that there is no difference in two treatment groups.

Week 28 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation (such as: due to Disease progression, Liver function, Patient decision, Investigator decision, AE/SAE, etc.): using While on treatment strategy: PEW of TI will be estimated in the duration beginning with the first dose date (Day 1) and ending with the date of end of treatment or last visit date at week 28/52, whichever comes earlier.

2. Death: using While on treatment strategy: Treatment effect while on-treatments of interest. PEW of TI will be based on RBC transfusion records up to the date of end of treatment or last visit date at week 28/52 or death, whichever comes earlier.

In addition, the cumulative number of Patient-Exposure-Weeks of TI at week 28 and week 52 will be also summarized using the box-whisker plot by treatment groups.

12.3.3.4. Number of pRBC Transfused compared to Baseline over the first 28 Weeks of the Treatment Period

Descriptive statistics of the number (pRBC/8 weeks) of pRBC transfused and number (pRBC/8 weeks) of pRBC transfused change from baseline will be presented by treatment group using FAS and PPS. See Section 5.5 for the definition of number of pRBC transfusions at baseline.

Treatment comparison will be performed for the number (pRBC/8 weeks) of pRBC transfused change from baseline based on a non-parametric ANCOVA (stratified rank ANCOVA) using baseline hemoglobin and baseline pRBC transfusions (pRBC/8 weeks) as covariate, treatment arm as fixed effects, controlling for stratification factors baseline endogenous EPO level and IPSS-R risk category. The p-value obtained will be reported for the test of the null hypothesis that there is no difference in two treatment groups.

Week 28 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively.

Similar methods will be used as in Section 12.3.3.3. Similar Strategies will be implemented for potential intercurrent events.

12.3.3.5. Proportion of Patients Who Achieved TI for ≥20 consecutive Weeks (140 consecutive Days) anytime during the study

The proportion of patients who achieved TI for ≥ 20 Weeks (140 Days) will be summarized using the FAS and PPS and will be compared between the two treatment groups. TI is defined as the absence of any intravenous RBC transfusion (packed cell or whole blood) during any consecutive 140 days anytime during the study. TI will be estimated in the duration beginning with the first dose date (Day 1) and ending with the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier. RBC transfusions after end of treatment will be recorded on CRF until the last visit at week 52 from first dose and will be used to define TI.

The p-value obtained will be reported for the test of the null hypothesis that there is no difference between the two treatment groups.

Similar methods will be used as in Section 12.3.3.1. Similar strategies will be implemented for potential intercurrent events.

12.3.3.6. Mean Change from Baseline in Physical Functioning Subscale of PROMIS

The PROMIS PF item measures self-reported, current capability to carry out activities that require physical actions, ranging from self-care (activities of daily living) to more complex activities that require a combination of skills, often within a social context. The questionnaire will be administrated at the Week 1 Day 1, Week 9, Week 17, Week 28 and Week 52 (or EOT). The PF 10-item short form which contains 10 questions will be used in this study, and each item is scored on a 5-point rating scale, with higher scores indicating better functioning. To find the total raw score, sum the values of the response to each question, with the lowest possible raw score is 10; the highest possible raw score is 50.

PROMIS items are not scored as sums, but rather on a standardized T-score metric (Appendix 19.5) using IRT. The standardized norm-based T-scores will be used in the summary. The mean change from Baseline in the PROMIS-SF v2.0 Physical Function 10b score will be summarized and analyzed. A mixed model for repeated measures (MMRM) with treatment, visit, treatment*visit, and stratification factors as fixed effects, the baseline score value and baseline*visit interaction as covariate will be used to assess treatment difference. The LSM and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model. The LSM for the difference will be considered clinically meaningful if it exceeds the distribution-based minimal clinically important difference (MCID) and will be marked in the results, see Section 19.8.1 for details of MCID. An unstructured variance-covariance structure will be used to model the within-patient errors. This variance-covariance matrix will be estimated across treatment groups. If the model fails to converge, then the covariance structure that converges and with the highest REML log likelihood will be used.

Missing values will be handled by the MMRM model, utilizing data from all visits and all patients.

A sensitivity analysis will be conducted using Analysis of Covariance (ANCOVA) model with Last Observation Carries Forward (LOCF) imputation, see Section 19.8.3 for details.

Week 9 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively, the p-value obtained will be reported.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation due to reasons other than AE/SAE: using Treatment-policy strategy: regardless of these intercurrent events occurred;

2. Treatment discontinuation due to AE/SAE: using Hypothetical strategy: A scenario is envisaged in which the patient suffers from the worst quality of life after discontinuation treatment due to AE/SAE. If a patient discontinued treatment due to AE/SAE, all the missing PROMIS-SF v2.0 Physical Function 10b score afterward will be substituted with the worst possible scores before being used in descriptive summary and statistical models.

3. Death: using Treatment-policy strategy: regardless of these intercurrent events occurred. If a patient died, the PROMIS-SF v2.0 Physical Function 10b scores up to death will be used for analysis.

Line plot of mean and mean change from baseline over time with the associated standard error will be provided. A cumulative distribute plot will be presented to show the difference of patients achieving meaningful change in PROMIS-SF v2.0 Physical Function 10b score between two treatment groups.

A by-patient listing will also be provided.

The analysis will be conducted for FAS and PPS.

12.3.3.7. Mean Change from Baseline in Fatigue Subscale of PROMIS

Fatigue will be measured in the study using the 13-item fatigue scale of the FACIT Measurement System, each item is scored on a 5-point rating scale ranging from 1 to 5 depicting “not at all” to “very much”, with higher scores indicating better functioning. The questionnaire is designed for self-completion by patients and will be administered at baseline, Week 9, Week 17, Week 28 and Week 52(or EOT).

PROMIS items are not scored as sums, but rather on a standardized T-score metric (Appendix 19.6) using IRT. The standardized norm-based T-scores will be used in the summary. The mean change from Baseline in the PROMIS-SF v1.0 Fatigue 13a score over time will be summarized and analyzed using the FAS and PPS Population. The analysis window is defined in Appendix 19.3.1.

A mixed model for repeated measures (MMRM) with treatment, visit, treatment*visit, and stratification factors as fixed effects, the baseline score value and baseline*visit interaction as covariate will be used to assess treatment difference. The LSM and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model. The LSM for the difference will be considered clinically meaningful if it exceeds the distribution-based minimal clinically important difference (MCID) and will be marked in the results, see Section 19.8.1 for details of MCID. An unstructured variance-covariance structure will be used to model the within-patient errors. This variance-covariance matrix will be estimated across treatment groups. If the model fails to converge, then the covariance structure that converges and with the highest REML log likelihood will be used.

Missing value will be handled by the MMRM model, utilizing data from all visits and all patients.

A sensitivity analysis will be conducted using Analysis of Covariance (ANCOVA) model with Last Observation Carries Forward (LOCF) imputation, see Section 19.8.2 for details.

Week 9 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively, the p-value obtained will be reported.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation due to reasons other than AE/SAE: using Treatment-policy strategy: regardless of these intercurrent events occurred;

2. Treatment discontinuation due to AE/SAE: using Hypothetical strategy: A scenario is envisaged in which the patient suffers from the worst quality of life after discontinuation treatment due to AE/SAE. If a patient discontinued treatment due to AE/SAE, all the missing PROMIS-SF v1.0 Fatigue 13a score afterward will be substituted with the worst possible scores before being used in descriptive summary and statistical models.

3. Death: using Treatment-policy strategy: regardless of these intercurrent events occurred. If a patient died, the PROMIS-SF v1.0 Fatigue 13a scores up to death will be used for analysis.

Line plot of mean and mean change from baseline over time with the associated standard error will be provided. A cumulative distribute plot will be presented to show the difference of patients achieving meaningful change in PROMIS-SF v1.0 Fatigue 13a score between two treatment groups. A by-patient listing will also be provided.

The analysis will be conducted for FAS and PPS.

12.3.3.8. Mean Change from Baseline in EQ-5D assessment

The EQ-5D questionnaire is designed for self-completion by patients and will be administered at baseline, Week 9, Week 17, Week 28 and Week 52(or EOT). The EQ-5D-5L descriptive system, which will be applied in the study, comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems, slight problem, moderate problems, severe problems, and unable to/extreme problems. The questionnaire also includes a visual analogue scale, where the patient will be asked to rate current health status on a scale of 0-100, with 0 being the worst imaginable health state and 100 being the best imaginable health.

Descriptive statistics (frequency counts per answer) by treatment group will be presented for each of the EQ-5D dimension at each study visit.

The mean change from Baseline in EQ-5D-5L VAS score will be also summarized and analyzed. Baseline is defined as the last assessment before first intake of dose of study treatment. The analysis windows are defined in Appendix 19.3.1.

A mixed model for repeated measures (MMRM) with treatment, visit, treatment*visit, and stratification factors as fixed effects, the baseline score value and baseline*visit interaction as covariate will be used to assess treatment difference. The analysis will include the patients with both a baseline and post-baseline value at the specific visit. The least-squares mean (LSM) and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model. An unstructured variance-covariance structure will be used for the within-patient variation. This variance-covariance matrix will be estimated across treatment groups. If the model fails to converge, then the covariance structure that converges and with the highest REML log likelihood will be used.

Missing value of this endpoint will be handled by the MMRM model, utilizing data from all visits and all patients.

Week 9 and week 52 are the analysis endpoint for 28-week and 52-week analysis, respectively. The p-value obtained will be reported.

This analysis will take into account the potential intercurrent events as follows:

1. Treatment discontinuation due to reasons other than AE/SAE: using Treatment-policy strategy: regardless of these intercurrent events occurred;

2. Treatment discontinuation due to AE/SAE: using Hypothetical strategy: A scenario is envisaged in which the patient suffers from the worst quality of life after discontinuation treatment due to AE/SAE. If

a patient discontinued treatment due to AE/SAE, all the missing EQ-5D-5L VAS scores afterward will be substituted with the worst possible scores before being used in descriptive summary and statistical models.

3. Death: using Treatment-policy strategy: regardless of these intercurrent events occurred. If a patient died, the EQ-5D-5L VAS scores up to death will be used for analysis.

Line plot of mean and mean change from baseline over time on EQ-5D assessment will be provided. A cumulative distribute plot will be presented to show the difference of patients achieving meaningful change in EQ-5D-5L VAS score between two treatment groups. A by-patient listing will also be provided.

The analysis will be conducted for FAS and PPS.

12.3.4. Analysis of Exploratory Efficacy Endpoints

P values generated from treatment comparison for exploratory endpoints are exploratory, no multiplicity adjustment will be needed. Nominal p-values are used primarily as measures of strength of association rather than formal criteria to claim statistical significance. The nominal p-values will also be explained in table footnotes. All analysis will be performed using the FAS.

12.3.4.1. Time to transfusion independence (TI)

Time to transfusion independence (in weeks) is measured from the day of first study drug to the first day of TI response anytime during the study or a censor date. Section 12.3.3.1 for the detail of definition of TI response anytime during the study. If a patient has more than one at least 56-day period with transfusion free, that is, more than one TI response period, only the first will be used for analysis.

- for those who have a TI response period, time to transfusion independence will be calculated as duration = first day of TI response - day of first study drug +1 where the first day of TI response is defined as the first day of at least 56-day transfusion-free period
- for those who do not have a TI response period, time to transfusion independence will be calculated as duration = date of the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier - day of first study drug + 1.

Median time to transfusion independence will be estimated using Kaplan-Meier (K-M) method and its associated 2-sided 95% CIs for each treatment group will be calculated using the method of Brookmeyer and Crowley. The estimated survival probability and its associated 2-sided 95% CIs for each treatment group up to 52 weeks by every 4-weeks will be estimated by Kaplan-Meier product limit method. Kaplan-Meier curves will also be presented by treatment group.

Sample SAS codes are presented as below:

```
PROC LIFETEST DATA=<input dataset> PLOTS=SURVIVAL;
        TIME value*censor (1);
STRATA treatment;
SURVIVAL OUT=xx;
RUN;
```

Treatment comparison will be performed using the stratified log-rank test stratified by randomization factors.

Sample SAS codes are presented as below:

PROC LIFETEST DATA=<input dataset>;

TIME value*censor (1);

STRATA <stratum_var>/GROUP=treatment;

RUN;

Time to transfusion independence will be compared between treatment groups using the Cox regression model, controlling for baseline hemoglobin and the stratification factors. Two-sided 95% CI of the hazard ratio will be presented for the treatment difference.

Sample SAS codes are presented as below:

PROC PHREG DATA=<input dataset>;

MODEL value*censor (1)= treatment base_hb <stratum_var>;

RUN;

Model checking: The proportional hazards assumption will be checked graphically using a log-cumulative hazard plot against log-survival time. This plot will give approximately parallel lines if the proportional hazards assumption between treatment arms subgroups holds. If the proportional hazard assumption is violated, then stratified Cox Model (with stratification-factors in STRATA statement) will be used instead.

12.3.4.2. Time to 50% reduction in pRBC transfusion

Time to 50% reduction in pRBC transfusion (in weeks) is measured from the day of first study drug to the first day of response of 50% reduction of pRBC transfusion or a censor date. See Section 12.3.3.1 for the detail of definition of the response.

- for those who have a response, time to 50% reduction in pRBC transfusion will be calculated as duration = first day of response - day of first study drug +1
- for those who do not have a response, time to 50% reduction in pRBC transfusion will be calculated as duration = date of the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier - day of first study drug + 1.

Median time to 50% reduction in pRBC transfusion will be estimated using Kaplan-Meier (K-M) method and its associated 2-sided 95% CIs for each treatment group will be calculated using the method of Brookmeyer and Crowley. The estimated survival probability and its associated 2-sided 95% CIs for each treatment group up to 52 weeks by every 4-weeks will be estimated by Kaplan-Meier product limit method. Kaplan-Meier curves will also be presented by treatment group.

Sample SAS codes are presented as below:

PROC LIFETEST DATA=<input dataset> PLOTS=SURVIVAL;

TIME value*censor (1);

STRATA treatment;

SURVIVAL OUT=xx;

RUN;

Treatment comparison will be performed using the stratified log-rank test stratified by randomization factors.

Sample SAS codes are presented as below:

PROC LIFETEST DATA=<input dataset>;

TIME value*censor (1);

STRATA <stratum_var>/GROUP=treatment;

RUN;

Time to 50% reduction in pRBC transfusion will be compared between treatment groups using the Cox regression model, controlling for baseline hemoglobin and the stratification factors. Two-sided 95% CI of the hazard ratio will be presented for the treatment difference. Sample SAS codes are presented as below:

PROC PHREG DATA=<input dataset>;

MODEL value*censor (1)= treatment base_hb <stratum_var>;

RUN;

Model checking: The proportional hazards assumption will be checked graphically using a log-cumulative hazard plot against log-survival time. This plot will give approximately parallel lines if the proportional hazards assumption between treatment arms subgroups holds. If the proportional hazard assumption is violated, then stratified Cox Model (with stratification-factors in STRATA statement) will be used instead.

12.3.4.3. Maximum Duration of transfusion independence (TI)

Duration of TI (in weeks) is measured from the first day of the TI response anytime during the study to the date of the first RBC transfusion after this period or a censor date. See Section 12.3.3.1 for the detail of definition of TI response. If a patient has more than one at least 56-day period with transfusion free, i.e., more than one TI response period, then the one with the maximum duration will be used for analysis.

- for those who receive a subsequent RBC transfusion, the duration of TI will not be censored and will be calculated as duration = last day of TI response - first day of TI response +1 where the last day of response is defined as 1 day before the first subsequent RBC transfusion which is given at 56 days or more after the response starts.
- for those who do not receive a subsequent RBC transfusion, the end day of the TI response will be censored as in Section 12.3.3.1 and duration of the TI response will be calculated as duration = date of the end of study or treatment discontinuation due to AE/SAE or death, whichever comes earlier - first day of TI response + 1.

A Frequency summary with regard to patients who had a transfusion after the TI response with the maximum duration of the TI response, patients who maintained transfusion independence until the censored date will be provided by treatment group.

Median maximum duration of TI will be estimated using Kaplan-Meier (K-M) method and its associated 2-sided 95% CIs for each treatment group will be calculated using the method of Brookmeyer and Crowley. The estimated survival probability and its associated 2-sided 95% CIs for each treatment group up to 52 weeks by every 4-weeks will be estimated by Kaplan-Meier product limit method. Kaplan-Meier curves will also be presented by treatment group.

Sample SAS codes are presented as below, the ‘input dataset’ will consist of TI responders only:

PROC LIFETEST DATA=<input dataset> PLOTS=SURVIVAL;

TIME value*censor (1);

STRATA treatment;

SURVIVAL OUT=xx;

RUN;

Treatment comparison will be performed using the stratified log-rank test stratified by randomization factors.

Sample SAS codes are presented as below:

PROC LIFETEST DATA=<input dataset>;

TIME value*censor (1);

STRATA <stratum_var>/GROUP=treatment;

RUN;

Maximum Duration of TI will be compared between treatment groups using the Cox regression model, controlling for baseline hemoglobin and the stratification factors. Two-sided 95% CI of the hazard ratio will be presented for the treatment difference. Sample SAS codes are presented as below:

PROC PHREG DATA=<input dataset>;

MODEL value*censor (1)= treatment base_hb <stratum_var>;

RUN;

Model checking: The proportional hazards assumption will be checked graphically using a log-cumulative hazard plot against log-survival time. This plot will give approximately parallel lines if the proportional hazards assumption between treatment arms subgroups holds. If the proportional hazard assumption is violated, then stratified Cox Model (with stratification-factors in STRATA statement) will be used instead.

Summary statistics for duration of response (in weeks) as well as a frequency summary of duration of Response by different time points (e.g. ⩾ 8 weeks, ⩾ 16 weeks, etc.) will also be provided.

12.3.4.4. Proportion of patients who achieve a mean Hb increase of ≥ 1.0 g/dL and ≥ 1.5 g/dL (averaged over 8 weeks in those that achieved TI) compared to pre-transfusion Hb at baseline)

Moving average (MA) of change from baseline in Hb will be calculated for any 8-week window (e.g., Days 1 to 56, Days 2 to 57, Days 3 to 58, etc.), starting at Week 1 day 1 post-dose.

Baseline Hb value is defined as the mean of screening central laboratory Hb value and week 1 Day 1 Hb value prior to treatment. For each post-baseline 8-week window, a minimum of 3 evaluable Hb assessments is required for 28-week analysis (only one assessments every 4 weeks after Week 31.); otherwise, the mean is set to missing in the corresponding window. A patient is considered a responder of this endpoint during the treatment period if

- Transfusion independent for ≥ 56 days consecutively;
- The moving averages of change from baseline in Hb is ≥ 1.0 g/dL (or ≥ 1.5 g/dL) during the treatment period.

The following summaries will be provided based on FAS:

- Analysis of average change from baseline in Hb by every 8-week period (e.g., 56 consecutive days)
- Number and percentage of TI responder and non-responder

Baseline Hb, absolute Hb value and change from baseline at each scheduled post-baseline visit

The following summaries will be provided for TI responders in FAS:

- The proportion of patients who achieve a mean Hb increase of ≥ 1.0 g/dL averaged over any 8 weeks
- The proportion of patients who achieve a mean Hb increase of ≥ 1.5 g/dL averaged over any 8 weeks

The denominator for the proportions is FAS population.

Similar methods will be used as in Section 12.3.1.1.

Mean (±SE) Hb (g/dL) from Baseline to Week 28 and to Week 52 will be plotted by treatment groups.

12.3.4.5. Impact of baseline endogenous erythropoietin levels on roxadustat treatment response and dose requirement

TI response rate (RBC TI at least ≥ 8 weeks) will be analyzed by baseline endogenous EPO level (< 200 mIU/mL **OR** > 200 mIU/mL and ≤ 400 mIU/mL) using the methods described in Section 12.3.1. A summary table will be provided. See Section 12.3.3.1 for the detail of definition of TI response.

Average weekly exposure data will be analyzed by baseline endogenous EPO level and treatment group.

The average weekly dose will be analyzed by baseline endogenous EPO level (< 200 mIU/mL **OR** > 200 mIU/mL and ≤ 400 mIU/mL) and treatment group.

In order to compare patients with baseline endogenous EPO level < 200 mIU/mL to patients with baseline endogenous EPO level > 200 mIU/mL and ≤ 400 mIU/mL), the ANCOVA with baseline Hb value as covariate and baseline endogenous EPO level as class category.

RBC TI response analysis as primary efficacy endpoint will be repeated for each level of baseline EPO：≤ 100, ≤ 200, ≤ 300 and, ≤ 400 mIU/mL.

Proportion of patients who achieve a mean Hb increase of ≥ 1.0 g/dL and ≥ 1.5 g/dL (averaged over 8 weeks in those that achieved TI) compared to pre-transfusion Hb at baseline will be summarized for each level of baseline EPO：< 200 mIU/mL **OR** > 200 mIU/mL and ≤ 400 mIU/mL.

12.3.4.6. Potential Effect on Hepcidin and Iron Metabolism

The effect on Hepcidin and Iron Metabolism will be assessed using the mean change from baseline over time in Hepcidin and Iron parameters.

Hepcidin Levels

Hepcidin levels will be measured at screening, week 1 day 1, week 5, week 13, week 21, week 25, week 28, week 39 and week 52 (or EOT).

Absolute value and mean change from baseline in hepcidin at each time point will be summarized using descriptive statistics. The analysis windows are defined in Appendix 19.3.1. The summary will include:

- Change from baseline in hepcidin
- Change from baseline in hepcidin stratified by TI vs. Non-TI (TI will be defined by week 28 as for the primary endpoint)
- Correlation between change from baseline in hepcidin and change from baseline in Hb. A scatter plot will be provided with CFB in Hepcidin on x-axis and CFB in Hb on y-axis. Different colors

will be used to indicate two treatment groups. A fitted linear regression line will be added to the scatter plot as well as the spearman correlation coefficient and p-value.

Treatment comparison will be based on a MMRM model with treatment group, visit, treatment*visit as fixed effect, baseline hepcidin value as covariate, controlling for stratification factors. The LSM and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model. Missing data will be handled by the MMRM model.

Week 28 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively; the p-values will be reported.

The analysis will be conducted for FAS population.

Iron Parameters

Iron biomarkers will be measured at screening, week 1 Day 1, week 3, week 5, week 9, week 13, week 17, week 21, week 25, week 28, week 39 and week 52 (or EOT).

Absolute value and mean change from baseline in iron parameters (Transferrin, TIBC, UIBC, TSAT and Ferritin, serum iron) at each time point will be summarized using descriptive statistics. The analysis windows are defined in Appendix 19.3.1.

Treatment comparison will be based on a MMRM model with treatment, visit, treatment*visit, and stratification factors as fixed effects, the baseline value as covariate. The LSM and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model.

Week 28 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively; the p-values will be reported.

The analysis will be conducted for FAS population.

12.3.4.7. Effect on Cholesterol and Lipid Parameters

The effect on cholesterol and lipid parameters will be assessment by mean change from baseline over time in cholesterol and lipid parameters.

Absolute value and mean change from baseline in total cholesterol at each time point will be summarized using descriptive statistics. The definitions of analysis widows are defined in Appendix 19.3.1.

The total cholesterol will be analyzed similarly using the MMRM method. The LSM and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model.

Other Serum Lipid Panel (non-fasting) parameters includes LDL, HDL, LDL/HDL ratio and triglycerides (TG), summary of mean change from baseline over time will use the same method as that for total cholesterol.

Week 28 and week 52 are the analysis endpoint for 28-week and 52-week analysis respectively; the p-values for other visits are for reference only.

For the analysis of triglycerides data, due to the nature of the skewness for the triglycerides data, the stratified non-parametric ANCOVA as that described in Section 12.3.3.3 with baseline value as the covariate will be used for the analyses.

12.4. Examination of Subgroups

Subgroup analyses (based on the FAS) will be performed based on the following criteria using the same approach as for the primary efficacy endpoint:

- by gender: male, female
- by age: < 65, >= 65 years
- by race: white, non-white
- by baseline endogenous EPO level: ≤ 200 mIU/mL, > 200 and ≤ 400 mIU/mL (defined as an exploratory endpoint in Section 12.3.4.5)
- by IPSS-R risk category: low risk / very low risk, intermediate risk
- by RBC transfusion burden: 1 pRBC/8-weeks over 16 consecutive weeks, 2- 4 pRBC/8-weeks
- by baseline iron repletion status ([TSAT < 20% or ferritin < 100 ng/mL] vs. others),
- by baseline CRP group (CRP ≤ ULN vs. CRP > ULN)
- by Prior ESA use: (ESA Naïve vs. ESA Experienced)
- by region (Section 12.5)

Subgroup analysis will also be performed based on anticoagulant use at baseline if applicable. For the mentioned subgroup analyses, un-stratified tests/models can be used to evaluate the treatment differences if the sample size within a level of a subgroup does not allow for a comparison. If the subgroup variable is the same as the stratified variable, the corresponding stratified variable will be dropped from analysis. If the strata contain few or no patients, the stratification variable will not be included in the analysis.

Patients with missing baseline values used to define the subgroups will be considered in a subgroup category missing, otherwise data will be excluded from the respective subgroup analyses.

12.5. Subgroup Analyses by Region

As it is expected to have approximately 100 global sites in this study, sensitivity analyses for the primary and secondary endpoints will be performed adding region and region by treatment interaction term in the model in order to examine whether the treatment effect varies by region and if results adjusted by region are consistent with those from the corresponding primary analysis and secondary analyses. Sites will be pooled by country into geographic region (defined as North America, Europe and Asia-pacific) to enable these analyses.

In addition, the consistency of the treatment effect across region will also be assessed by performing a subgroup-type analysis with respect to the primary and secondary endpoint, with region treated as subgroups. A listing of TI response rates will be provided by individual study site.

For NDA filing in China, the above analyses by additional region subgroup (defined as East Asia, non-East Asia) will also be provided.

13. PHARMACOKINETICS / PHARMACODYNAMICS ANALYSES

The methods described in Appendix 19.2 will be used.

14. SAFETY ANALYSES

All safety analyses will be conducted using the safety population. Percentages will be calculated based on the number of patients in the safety population.

For open-label component, the summaries will be presented by dose level cohort. For OL high-erythropoietin component, the summaries will be presented along with the cohorts from open-label component. For double-blind component, the summarizes will be presented by treatment group.

14.1. Adverse Events

A treatment-emergent AE (TEAE) is defined as an AE that begins on or after the first dose of study treatment and within 28 days of the end date of the last dose of study treatment. An adverse event with unknown onset date will be counted as TEAE unless the event is resolved prior to the first dose of study medication.

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA, version 19.1). CTCAE Version 4.03 will serve as guidance for specific severity assessment criteria for AEs. Any adverse event recorded prior to the start of study treatment will be listed together with all other adverse events. Only treatment-emergent adverse events will be summarized.

If a patient has more than one TEAE coded to the same preferred term, the patient will be counted only once for that preferred term. Similarly, if a patient has more than one TEAE within a system organ category, the patient will be counted only once in that system organ category.

In this analysis, if a patient reports multiple occurrences of TEAEs within one system organ class or preferred term, the most closely related occurrence will be presented. TEAEs with unknown relationship to study treatment will be counted as “possibly related” in the summary table but will be listed with a missing relationship in data listing.

In this analysis, if a patient reports multiple occurrences of TEAEs within one system organ class or preferred term, the most severe occurrence will be presented in the table summarized by maximum severity. TEAEs with unknown severity will be counted as “Severe” in the summary table but will be listed with a missing severity in data listing.

The overview table by treatment and overall will be summarized showing the number and percentage of patients with at least one AE in any of the following categories:

- TEAEs
- drug related (Possibly Related and Related) TEAEs
- all grade 3 or greater TEAEs
- Serious TEAEs
- TEAEs leading to discontinuation of study treatment
- TEAEs leading to death
- TEAE of special interest

In summary, the following sets of adverse events will be summarized separately in separate tables by SOC (sorted alphabetically) and PT (sorted by descending) that present the number and percentage of patients with:

- any TEAE (number of patients who have at least one adverse event)

- TEAE by maximum severity
- TEAE with relationship to study medication suspected to be related to study drug
- all grade 3 or greater TEAE
- treatment-emergent serious AEs (TESAE)
- TEAEs leading to discontinuation of study treatment,
- TEAEs leading to death
- TEAE of special interest
- TESAE leading to discontinuation of study treatment

AE of special interest consists of the following categories: Secondary hypothyroidism, Thyroid disorders (Hypothyroidism and Hyperthyroidism), Dermatitis Exfoliative Generalised, Deep vein thrombosis, Pulmonary embolism, Cardiovascular events (cardiac failure, arrythmia, myocardial infarction), Seizures, Serious infections.Number and percentage of patients with any TEAE of special interest will summarized up to 28 weeks, 52 weeks and 52 weeks and beyond. Exposure adjusted incidence rate (EAIR) of TEAE of special interest per 100 patient exposure years will also be presented. See Section 14.1.2 for EAIR calculation.

In addition, most frequently occurring adverse events (PT $\geqslant$ 5% of patients in any dose cohort/treatment group) and most frequently occurring non-serious adverse events (PT $\geqslant$ 5% of patients in any dose cohort/treatment group) for each PT will be summarized in descending frequency according to its incidence in either dose cohort/treatment group.

In addition, the incidence rate * 100 / PEY will be presented along with the number and percentage of patients in double-blind tables, where Patient-Exposure-Year (PEY) = (Last Dose Date - First Dose Date + 1)/365.25.

All AEs will be presented in a listing, including patient identifier, treatment group, system organ class, preferred term, start and stop dates, seriousness, severity, relationship to study treatment, Action taken with regard to study drug and outcome. Additionally, listings of SAEs, Death and AEs leading to permanent discontinuation of study drug and AEs of special interest will be presented. The listing of AEs prior to start of treatment will be provided for all patients enrolled.

Listings of death report will be provided in the SAF with death date, primary cause of death and verbatim, system organ class and preferred term of AE leading to death.

14.1.1. AE Subgroup Analysis

The overview table of TEAE by the following subgroups will be provided:

- by gender: male, female
- by age: < 65, >= 65 years
- by race: white, non-white
- by baseline endogenous EPO level: $\leqslant$ 200 mIU/mL, > 200 and $\leqslant$ 400 mIU/mL (defined as an exploratory endpoint in Section 12.3.4.3)
- by IPSS-R risk category: low risk / very low risk, intermediate risk
- by RBC transfusion burden: 1 pRBC/8-weeks over 16 consecutive weeks, 2- 4 pRBC/8-weeks

- by region: North America, Europe and Asia-pacific

In addition, for NDA filing in China, the overview table of TEAE, TEAE table and TESAE table by the following subgroups will be provided:

- by additional race subgroup: Asian, non-Asian
- by additional region subgroup: East Asia, non-East Asia

14.1.2. Exposure Adjusted Incidence Rate

Exposure adjusted incidence rate EAIR (per 100 patient years at risk) will be calculated regarding selected TEAEs of interest.

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

$$\frac{\text{Number of subjects with event}}{\text{Person - years at risk}} \times 100$$

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

1. 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.

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

   (Censoring date – Date of first dose + 1) / 365.25

   Censoring date will be calculated as:

   Min (death date, last contact date if the subject is lost to FU, visit date of week 52 + 28, last dose date + 28)

14.2. Disease Progression of AML

The number and percentage of patients in double-blind component who progressed to AML will be presented in a frequency table using safety population.

Listings by patient with progressive disease for OL high-erythropoietin component and double-blind component will also be provided.

14.3. Clinical Laboratory Data

Summary tables for clinical laboratory data will be performed for double-blind component only.

The results of laboratory assessments, which include CBC, Serum chemistries (sodium, potassium, chloride, bicarbonate, BUN, creatinine, glucose, calcium, magnesium, phosphorus, lactate dehydrogenase (LDH), albumin) will be provided by a central laboratory. All summaries will be based on the standard units provided by the central laboratory.

Descriptive statistics for laboratory values (in SI units) and changes from baseline at each assessment time point will be presented by treatment group for laboratory parameters collected in the study.

To assess potentially clinically significant (PCS) laboratory abnormalities (Table 5), the number and percentage of patients with post-baseline lab values outside a pre-defined range (low or high) or limit of change will be tabulated. Laboratory test values will be considered PCS if they meet the criteria listed in the table. The number and percentage of patients with post-baseline PCS values will be tabulated. The

percentages are to be calculated relative to the number of patients with available non-PCS baseline values and at least one post-baseline assessment. The numerator is the total number of patients with at least one post-baseline PCS value. Shift tables may be presented.

Patient incidence for each abnormality will be calculated based on patients with a baseline value and at least one post-baseline value for criteria requiring baseline or patients with at least one post-baseline value for criteria not requiring baseline. If differential dropout is observed between the two treatment arms, then the exposure adjusted incidence rate EAIR (per 100 patient years at risk) will be calculated regarding each abnormality.

A supportive listing of patients with post-baseline PCS values will be provided including the patient ID, study center, baseline, and post-baseline values. Values from unscheduled time points will not be used in post-baseline analysis but will be listed.

In addition, for serology testing and other lab test (e.g. PT, PTT, Vitamin B_{12} and folate) will be provided in data listings by patient. Local hemoglobin data will be listed along with prescribed dose adjustment for reference.

Table 5: Range 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
Calcium	mmol/L	<0.8*LLN	>1.2 * ULN
Creatinine	µmol/L		> 1.5x Baseline Value
Potassium	mmol/L	<0.75*LLN	>1.2 * ULN
Sodium	mmol/L	<0.9*LLN	>1.1 * ULN
Total Bilirubin	µmol/L		>1.5 * ULN
Urea (BUN)	mmol/L		>1.5x Baseline Value
CBC			
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			

14.3.1. Liver Function Tests

The frequency of the following categories will be summarized and compared across dose level cohorts (for open-label component and OL high-erythropoietin component) or treatment groups (for double-blind component).

- Aminotransferase (i.e. AST ≥ 3x ULN or ALT ≥ 3x ULN) or/and Total bilirubin (TBL) ≥ 2x ULN
- ALT or AST > 8 x ULN

- ALT or AST > 5 x ULN for more than 2 weeks
- ALT or AST > 3 x ULN and international normalized ratio (INR) > 1.5 (if INR testing is applicable/evaluated).
- ALT or AST > 3× ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash and/or eosinophilia (> 5%).

A scatter plot of the maximum aminotransferase (i.e., ALT or AST) level against maximum TBL at post-baseline will also be presently with dotted reference lines at 3 x ULN for ALT or AST, and 2 x ULN for total bilirubin. In each plot, TBL will be in the vertical axis. Different dots will be used for roxadustat and placebo.

Individual displays of abnormal Liver Enzymes (ALT or AST >= 3 x ULN) and Bilirubin parameters (>= 2 x ULN) with corresponding other liver enzymes and bilirubin values during the treatment periods will be listed. The timepoints for subjects falling into Hy's Law Range (ALT or AST > 3 x ULN, and total bilirubin > 2 x ULN), Cholestasis Range (total bilirubin > 2 x ULN), and Temple's Corollary Range (ALT or AST > 3 x ULN) will be marked.

14.4. Excessive Hematopoiesis

The presence of potential EH will be defined if Hb increases by >2.0 g/dL within 4 weeks of treatment during the treatment period. A listing of the patients with potential EH will be provided.

The following summaries will be presented for double-blind component:

- Proportion of patients with potential excessive erythropoiesis.
- Proportion of patients with Hb exceeds 12 g/dl.

14.5. 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 and their changes from baseline at each visit and at the end of study will be presented.

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 6 below. The number and percentage of patients with post-baseline PCS values will be tabulated by dose level cohort (for open-label component) /treatment group (for double-blind component). The percentages are to be calculated relative to the number of patients with baseline and at least one post-baseline assessment. The numerator is the total number of patients with at least one post-baseline PCS vital sign value. Shift tables may be presented.

Patient incidence for each abnormality will be calculated based on patients with a baseline value and at least one post-baseline value for criteria requiring baseline or patients with at least one post-baseline value for criteria not requiring baseline. If differential dropout is observed between the two treatment arms, then the exposure adjusted incidence rate EAIR (per 100 patient years at risk) will be calculated regarding each abnormality.

Line plots of mean and mean change from baseline of vital sign parameters over time with the associated standard error will be provided.

A supportive listing of patients with post-baseline PCS values will be provided including the patient ID, study center, baseline, and post-baseline values. Values from unscheduled time points will not be used in post-baseline analysis, but will be listed.

Table 6: 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.

14.6. Electrocardiogram (ECG)

ECG findings are summarized in shift tables that include cross-tabulation of 'Normal', 'Abnormal NCS', and 'Abnormal CS'. Baseline value is the last non-missing measurement on or prior to the first dose. Patients with a baseline value and at least one post-baseline value will be included in the summary.

ECG data are presented in the data listing.

14.7. ECOG Performance Status

For ECOG performance status, the shift from baseline to the worst and best during the treatment will be displayed in cross tabulations. Baseline value is the last non-missing measurement on or prior to the first dose. Patients with a baseline value and at least one post-baseline value will be included in the summary.

14.8. Pregnancy Test

Findings from the pregnancy test will only be listed if applicable.

15. DATA SAFETY MONITORING BOARD

Safety data and dosing decisions will be monitored on an ongoing basis. In addition, ongoing review of safety data will be completed by an independent Data Safety Monitoring Board (DSMB) on a pre-defined periodic basis. The role and function of the DSMB will be defined in a DSMB charter.

16. INTERIM ANALYSIS

A preliminary analysis of the open-label component was conducted at the end of the open-label component of the study In the double-blind component, the primary efficacy analysis will be conducted when all randomized patients have completed 28 weeks of treatment or early discontinued. A table of content (ToC) lists a subset of tables, figures, and listings for the Week 28 primary analysis in Appendix 19.9. All other planned analyses outlined in this SAP will not be conducted until the end of study. The Week 28 analysis will be conducted by a small unblinded group not directly involved in the day-to-day study conduct. The study team will remain blinded until the final database lock and study unblinding. A detailed Data Access Plan summarizes the maintenance of blinding for the rest of the study.

No other formal interim analysis is planned.

17. CHANGES FROM THE PROTOCOL-SPECIFIED ANALYSES

3. Exploratory analyses on time to transfusion independence and time to 50% reduction in pRBC transfusion are added in this SAP due to clinical interest, although they are not in the protocol.

4. The 52 weeks analysis on the primary endpoint is regarded as a secondary endpoint and been included in the sequential gatekeeping procedure.

5. Based on results from prior roxadustat studies, measuring HRQoL 2-4 weeks after Hb “peaks” or stabilizes is shown to be most informative. In order to capture the treatment effect on Physical Function/fatigue before it wears off, Week 9 will be used as the primary time-point for HRQoL hypothesis testing in the sequential gatekeeping procedure.

6. Based on the re-estimated sample size of 120, 140 patients were enrolled in the double-blind component of the study.

18. REFERENCE

Stokes, Davis, Koch (2000) Categorical Data Analysis Using the SAS System, Page 174-179.

Rombach, I., Rivero-Arias, O., Gray, A. M., Jenkinson, C., & Burke, Ó. (2016). The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation, 25(7), 1613–1623. https://doi.org/10.1007/s11136-015-1206-1

ICH (International Conference on Harmonization) Harmonized Tripartite Guideline E9 (R1). Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials, Final Version adopted on 20 November 2019

Fenaux P, Platzbecker U, Mufti GJ, et al. Luspatercept in Patients with Lower-Risk Myelodysplastic Syndromes. N Engl J Med. 2020;382(2):140-151. doi:10.1056/NEJMoa1908892

Ge, M., Durham, L. K., Meyer, R. D., Xie, W., & Thomas, N. (2011). Covariate-Adjusted Difference in Proportions from Clinical Trials Using Logistic Regression and Weighted Risk Differences. Drug Information Journal, 45(4), 481–493.

19. APPENDIX

19.1. Dose Adjustment Rules for Patients who have TI for ≥ 56 Days

A modified dose adjustment algorithm will apply to patients who have attained TI ≥ 56 days (Maintenance Phase). The down titration of the dose is allowed at any time.

Change in Hb over past 4 weeks (g/dL)	**Hemoglobin Values (g/dL)**			
	<10.5	**10.5 to 12.9**	**>13.0**	**≥13.5**
< -1.0 (or had RBC transfusion within 4 weeks)	↑	↑	No change	Hold, then resume dosing when: Hb <12 g/dL, at a dose that is reduced by one dose steps **Patients are to return weekly to monitor Hb until dosing can be resumed**
-1.0 to 1.0	↑	No change	↓	
>1.0	No change	↓	↓	

Dose Adjustment for Excessive Hematopoiesis:

At any time during the Treatment Period if Hb increases by > 2.0 g/dL within 4 weeks, the dose should be reduced by one dose step.

19.2. Pharmacokinetics / Pharmacodynamics Analyses

Blood samples will be collected at the study site for determination of roxadustat plasma concentrations and erythropoietin levels at 1-3 and 3-5 hours after dosing at clinic on Day 1 (Week 1) as well as 4-6 hours at Week 5 and 8-10 hours at Week 21 after self-administration at home or at the clinic.

Population PK/PD Analysis will be done outside of the SAP, which will focus on safety and efficacy.

19.3. Data Handling Conventions

19.3.1. Analysis Windows

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

The analysis windows will be only applied to those assessments up to the day before the first week of follow-up visit. Assessments obtained afterwards are considered follow-up assessments. For follow-up assessments, only follow-up assessments on scheduled follow-up visits (i.e. nominal follow-up visits) will be summarized unless otherwise specified.

If more than one assessment is available in the same window, the last assessment is used for analysis. Assessments that do not fall in any windows will not be included in analyses unless otherwise specified. However, all post-baseline assessments will be used for PCS value determination and all assessments will be presented in the data listings.

Table 7.1 Analysis Visit Windows (General)

Analysis Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1	Up to Day 1 Pre-Dose
Week 3	Day 7*(Week # -1)+1	[Day 1 Post-Dose, Scheduled Day +7)
Weeks 5, 7, 9, 11…27	Day 7*(Week # -1)+1	[Scheduled Day -7, Scheduled Day +7)
Week 28[b]	Day 7*(Week #)+1	[Scheduled Day -7, Scheduled Day +7)
Weeks 31	Day 7*(Week # -1)+1	[Scheduled Day -7, Scheduled Day +14)
Week 35, 39, 43	Day 7*(Week # -1)+1	[Scheduled Day -14, Scheduled Day +14)
Week 47	Day 7*(Week # -1)+1	[Scheduled Day -14, Scheduled Day +21)
Week 52[c]	Day 7*(Week #)+1	[Scheduled Day -21, Up to Week 4 Follow up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.2 Analysis Visit Windows for Lipid Parameters

Derived Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1(first dose day)	Up to Day 1 Pre-Dose
Week 5	Day 7 * (Week # -1) +1	[Day 1 Post-dose, Scheduled Day + 56)
Week 21	Day 7 * (Week # -1) +1	[Scheduled Day - 56, Scheduled Day + 28)
Week 28[b]	Day 7 * (Week #) +1	[Scheduled Day - 28, Scheduled Day + 35)
Week 39	Day 7 * (Week # -1) +1	[Scheduled Day - 35, Scheduled Day + 49)
Week 52[c]	Day 7 * (Week #) +1	[Scheduled Day -49, Up to Week 4 follow-up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.3 Analysis Visit Windows for Serum Iron, Ferritin, TIBC, UIBC, TSAT, Transferrin, Vital Sign, Serum Chemistry

Derived Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1(first dose day)	Up to Day 1 Pre-Dose
Week 5	Day 7 * (Week # -1) +1	[Day 1 post-dose, Scheduled Day + 14)
Weeks 9, 13, 17, 21	Day 7 * (Week # -1) +1	[Scheduled Day - 14, Scheduled Day + 14)
Week 25	Day 7 * (Week # -1) +1	[Scheduled Day - 14, Scheduled Day + 14)
Week 28[b]	Day 7 * (Week #) +1	[Scheduled Day - 14, Scheduled Day + 35)
Week 39	Day 7 * (Week # -1) +1	[Scheduled Day - 35, Scheduled Day + 49)
Week 52[c]	Day 7 * (Week #) +1	[Scheduled Day - 49, Up to Week 4 follow-up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.4 Analysis Visit Windows for Hepcidin

Derived Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1(first dose day)	Up to Day 1 Pre-Dose
Week 5	Day 7 * (Week # -1) +1	[Day 1 post-dose, Scheduled Day + 28)
Week 13	Day 7 * (Week # -1) +1	[Scheduled Day - 28, Scheduled Day + 28)
Week 21	Day 7 * (Week # -1) +1	[Scheduled Day - 28, Scheduled Day + 14)
Week 25	Day 7 * (Week # -1) +1	[Scheduled Day - 14, Scheduled Day + 14)
Week 28[b]	Day 7 * (Week #) +1	[Scheduled Day - 14, Scheduled Day + 35)
Week 39	Day 7 * (Week # -1) +1	[Scheduled Day - 35, Scheduled Day + 49)
Week 52[c]	Day 7 * (Week #) +1	[Scheduled Day - 49, Up to Week 4 follow-up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.5 Analysis Visit Windows for ECG

Derived Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1(first dose day)	Up to Day 1 Pre-Dose
Week 28[b]	Day 7 * (Week #) +1	[Day 1 post-dose, Scheduled Day + 84)
Week 52[c]	Day 7 * (Week #) +1	[Scheduled Day - 84, Up to Week 4 follow-up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.6 Analysis Visit Windows for CBC

(Only in Subjects Enrolled before Effective Day of Protocol Amendment 3)

Analysis Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1	Up to Day 1 Pre-Dose
Week 2	Day 7*(Week # -1)+1	[Day 1 Post-Dose, Scheduled Day +3]
Week 3,4	Day 7*(Week # -1)+1	[Scheduled Day -3, Scheduled Day +3]
Week 5	Day 7*(Week # -1)+1	[Scheduled Day -3, Scheduled Day +7)
Weeks 7, 9, 11…27	Day 7*(Week # -1)+1	[Scheduled Day -7, Scheduled Day +7)
Week 28[b]	Day 7*(Week #)+1	[Scheduled Day -7, Scheduled Day +7)
Weeks 31	Day 7*(Week # -1)+1	[Scheduled Day -7, Scheduled Day +14)
Week 35, 39, 43	Day 7*(Week # -1)+1	[Scheduled Day -14, Scheduled Day +14)
Week 47	Day 7*(Week # -1)+1	[Scheduled Day -14, Scheduled Day +21)
Week 52[c]	Day 7*(Week #)+1	[Scheduled Day -21, Up to Week 4 Follow up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.7 Analysis Visit Windows for EQ-5D-5L, PROMIS Physical Function and PROMIS Fatigue Function

Derived Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1(first dose day)	Up to Day 1 Pre-Dose
Week 9	Day 7 * (Week # -1) +1	[Day 1 post-dose, Scheduled Day + 28)
Weeks 17	Day 7 * (Week # -1) +1	[Scheduled Day - 28, Scheduled Day + 42)
Week 28[b]	Day 7 * (Week #) +1	[Scheduled Day - 42, Scheduled Day + 84)
Week 52[c]	Day 7 * (Week #) +1	[Scheduled Day - 84, Up to Week 4 follow-up)

[a]: Relative to the first study medication date. For example, Day 1 = the first dose date of study medication.
[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

Table 7.8 Analysis Visit Windows for ECOG

Derived Visit	Visit Scheduled Day[a]	Window
Baseline	Day 1(first dose day)	Up to Day 1 Pre-Dose
Week 5	Day 7 * (Week # -1) +1	[Day 1 post-dose, Scheduled Day + 28)
Week 13	Day 7 * (Week # -1) +1	[Scheduled Day - 28, Scheduled Day + 14)
Weeks 17	Day 7 * (Week # -1) +1	[Scheduled Day - 14, Scheduled Day + 28)
Weeks 25	Day 7 * (Week # -1) +1	[Scheduled Day - 28, Scheduled Day + 14)
Week 28[b]	Day 7 * (Week #) +1	[Scheduled Day - 14, Scheduled Day + 84)
Week 52[c]	Day 7 * (Week #) +1	[Scheduled Day - 84, Up to Week 4 follow-up)

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

[b]: Week 28 visit should be conducted at the end of 28 weeks of treatment (2 weeks after Week 27 Visit).
[c]: Week 52 Visit should be scheduled at the end of 52 weeks treatment (6 weeks after Week 47 visit).

19.3.2. Missing Date of Study Medication

When the last date of study medication during the study treatment period 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.

19.3.3. Missing Severity Assessment for Adverse Events

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

19.3.4. 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 "possibly related" will be assigned. The imputed values for relationship to study medication will be used for incidence summary, while the actual values will be presented in data listings.

19.3.5. Missing Date Information for Adverse Events

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

Missing day and month

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

If the year is prior to the year of first day on study medication, then December 31 will be assigned to the missing fields.

If the year is after the year of first day on study medication, then January 1 will be assigned to the missing fields.

Missing month only

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

Missing day only

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

If the month and year are before the year and month of first day on study medication, then the last day of the month will be assigned to the missing day.

If the month and year are after the year and month of first day on study medication, then the first day of the month will be assigned to the missing day.

If the stop date is complete and the imputed start date as above is after the stop date, the start date will be imputed by the stop date.

19.3.6. Missing Date Information for Prior or Concomitant Medications

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

Incomplete Start Date

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

Missing day and month

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

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

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

Missing month only

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

Missing day only

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

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

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

Incomplete Stop Date

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

Missing day and month

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

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

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

Missing month only

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

Missing day only

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

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

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

19.3.7. Missing Start Date for Study Disease (Myelodysplastic Syndrome)

When start date of study disease (Myelodysplastic syndrome) is incomplete, the following imputation rules will be adopted:

Missing day and month

January 1 will be assigned to the missing fields.

Missing month only

January will be assigned to the missing month.

Missing day only

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

19.4. 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 ❑

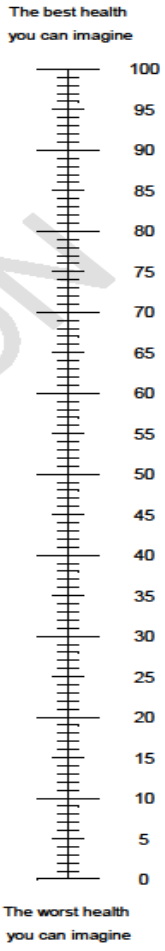
PAIN / DISCOMFORT
- I have no pain or discomfort ❑
- I have slight pain or discomfort ❑
- I have moderate pain or discomfort ❑
- I have severe pain or discomfort ❑
- I have extreme pain or discomfort ❑

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

19.5. PROMIS Physical Function 10b Short Form

Physical Function – Short Form 10b

Please respond to each question or statement by marking one box per row.

		Without any difficulty	With a little difficulty	With some difficulty	With much difficulty	Unable to do
PFA11	Are you able to do chores such as vacuuming or yard work?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFA56	Are you able to get in and out of a car?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFA21	Are you able to go up and down stairs at a normal pace?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFA53	Are you able to run errands and shop?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFA9	Are you able to bend down and pick up clothing from the floor?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFB28r1	Are you able to lift 10 pounds (5 kg) above your shoulder?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
		Not at all	**Very little**	**Somewhat**	**Quite a lot**	**Cannot do**
PFA1	Does your health now limit you in doing vigorous activities, such as running, lifting heavy objects, participating in strenuous sports?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFA5	Does your health now limit you in bathing or dressing yourself?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFB3	Does your health now limit you in putting a trash bag outside?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
PFB44	Does your health now limit you in doing moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

Instructions: You can use the following conversion table to translate the total raw score into a T-score for each participant. These conversions are accurate ONLY when all questions on the short form have been answered.

PROMIS-SF v2.0 Physical Function 10b Scoring Table

Adult v2.0 – Physical Function 10b		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
10	13.8	3.9
11	17.2	3.1
12	19.3	2.8
13	21.0	2.6
14	22.4	2.4
15	23.6	2.3
16	24.7	2.2
17	25.7	2.1
18	26.6	2.0
19	27.4	2.0
20	28.2	1.9
21	28.9	1.9
22	29.6	1.9
23	30.3	1.8
24	31.0	1.8
25	31.7	1.8
26	32.3	1.8
27	32.9	1.8
28	33.5	1.8
29	34.2	1.8
30	34.8	1.8
31	35.4	1.8
32	36.0	1.8
33	36.7	1.8
34	37.3	1.8
35	37.9	1.8
36	38.6	1.8
37	39.3	1.8
38	40.0	1.8
39	40.7	1.9
40	41.5	1.9
41	42.3	2.0
42	43.2	2.0
43	44.2	2.1
44	45.2	2.2
45	46.5	2.4
46	48.1	2.8
47	50.0	3.2
48	52.5	3.7
49	55.0	4.0
50	61.3	6.1
*SE = Standard Error on T-score metric		

19.6. Fatigue – Short Form 13a

Please respond to each question or statement by marking one box per row.

	During the past 7 days…	Not at all	A little bit	Somewhat	Quite a bit	Very much
HI7	I feel fatigued	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
HI12	I feel weak all over	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN1	I feel listless ("washed out")	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN2	I feel tired	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN3	I have trouble starting things because I am tired	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN4	I have trouble finishing things because I am tired	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN5	I have energy	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
AN7	I am able to do my usual activities	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
AN8	I need to sleep during the day	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN12	I am too tired to eat	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN14	I need help doing my usual activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN15	I am frustrated by being too tired to do the things I want to do	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
AN16	I have to limit my social activity because I am tired	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Instructions: You can use the following conversion table to translate the total raw score into a T-score for each participant. These conversions are accurate ONLY when all questions on the short form have been answered.

PROMIS-SF v1.0 Fatigue 13a Scoring Table

Fatigue 13a (FACIT-Fatigue) - Adult v1.0 *Short Form Conversion Table*					
Raw Score	T-score	SE*	Raw Score	T-score	SE*
13	30.3	4.7	40	60.8	1.8
14	35.0	3.5	41	61.4	1.8
15	38.0	3.0	42	62.0	1.8
16	40.3	2.8	43	62.6	1.8
17	42.1	2.6	44	63.2	1.8
18	43.7	2.5	45	63.8	1.8
19	45.0	2.3	46	64.4	1.8
20	46.3	2.2	47	65.0	1.8
21	47.3	2.1	48	65.6	1.8
22	48.3	2.0	49	66.2	1.9
23	49.3	2.0	50	66.9	1.9
24	50.1	1.9	51	67.5	1.9
25	51.0	1.9	52	68.2	1.9
26	51.7	1.9	53	68.9	2.0
27	52.5	1.9	54	69.6	2.0
28	53.2	1.9	55	70.4	2.0
29	53.9	1.8	56	71.2	2.1
30	54.6	1.8	57	72.0	2.2
31	55.3	1.8	58	72.9	2.3
32	55.9	1.8	59	73.9	2.4
33	56.6	1.8	60	75.0	2.5
34	57.2	1.8	61	76.2	2.7
35	57.8	1.8	62	77.5	2.9
36	58.4	1.8	63	79.1	3.1
37	59.0	1.8	64	81.2	3.3
38	59.6	1.8	65	83.5	3.4
39	60.2	1.8			

*SE = Standard Error on T-score metric

19.7. An Adjusted Difference in Proportions using Logistic Regression

The method was proposed by Ge et al. (2011). Please ensure the model built in this macro is the same with logistic regression model which is used for odds ratio calculation.

```
%macro LR (
data=, /* input data set */
var1=, /* continuous covariates in the logistic regression */
var2=, /* categorical covariates in the logistic regression */
p1=, /* number of continuous covariates in the logistic regression */
p2=, /* number of categorical covariates in the logistic regression */
resp=, /* binary response variable in the logistic regression */
ntrt= ); /* position of the treatment variable in the categorical covariates */
/* delete the observations with missing values in either numeric or categorical covariates */
```

```
data newdata;
set &data;
if cmiss(of &var1 &var2)>0 then delete;
run;
/* extract the design matrix and variance-covariance matrix of the regression and store them in
data1 and parms respectively */
/* cases when there is no continuous covariates */
%if &p1=0 %then
%do;
proc transreg data=newdata dummy noprint;
model identity(&resp) = class(&var2/zero=first);
output out = data1(drop=_: &var2 &resp) replace;
run;
proc logistic data=newdata desc covout outest=parms(drop=_:)
noprint;
class &var2 / param=ref ref=first;
model &resp = &var2;
run;
%end;
/* cases when both categorical and continuous covariates exist */
%else
%do;
proc transreg data=newdata dummy noprint;
model identity(&resp) = class(&var2/zero=first) identity(&var1);
output out = data1(drop=_: &var2 &resp) replace;
run;
proc logistic data=newdata desc covout outest=parms(drop=_:)
noprint;
class &var2 / param=ref ref=first;
model &resp = &var2 &var1;
run;
%end;
/* estimate the proportion difference and standard error to construct the 95% confidence
interval */
```

```
proc iml;
use parms;
read all into params;
close parms;
use data1;
read all into xmat;
close data1;
p = nrow(params)-2;
cov = params[2:(p+2),];
psize = nrow(xmat);
ntrt = &ntrt+1;
xt = xmat;
xt[,ntrt] = J(psize,1,1);
axT = xt*params[1,]`;
xc = xmat;
xc[,ntrt] = J(psize,1,0);
axC = xc*params[1,]`;
pderT = J(psize,1,0);
pderC = J(psize,1,0);
pderivT = J(psize,(p+1),0);
pderivC = J(psize,(p+1),0);
gt = J(1,(p+1),0);
gc = J(1,(p+1),0);
do k=1 to psize;
pderT[k] = exp(axT[k])/(1+exp(axT[k]));
pderC[k] = exp(axC[k])/(1+exp(axC[k]));
pderivT[k,] = pderT[k]*(1-pderT[k])*xt[k,];
pderivC[k,] = pderC[k]*(1-pderC[k])*xc[k,];
end;
difb=sum(pderT-pderC)/psize;
pT=sum(pderT)/psize;
pC=sum(pderC)/psize;
do k=1 to (p+1);
gt[k] = sum(pderivT[,k])/psize;
```

```
gc[k] = sum(pderivC[,k])/psize;
end;
se=sqrt(gt*cov*gt`+ gc*cov*gc`-2*gt*cov*gc`);
upperb=difb+1.96*se;
lowerb=difb-1.96*se;
store difb se pT pC lowerb upperb;
create difb var {"difb" "se" "pT" "pC" "lowerb" "upperb"}; **output statistics;
append;
close difb;
quit;
%mend;
/* print out the LR estimation results */
```

19.8. PRO Appendix for PROMIS

19.8.1. Distribution-based Minimal Clinically Important Difference (MCID)

The distribution-based minimal clinically important difference (MCID), ½ SD and 1 SEM will be provided based on the pooled PROMIS Physical Function scores and Fatigue scores of two treatment arms at baseline for the FAS and the PPS, respectively.

½ SD = 1/2 * Standard Deviation of the pooled norm-based T-scores of two treatment arms at baseline

1 SEM = [(1 – Cronbach's alpha) ^ (1/2)] * SD

Where SD is the same as the definition for ½ SD and Cronbach's alpha refers to the Standardized Cronbach's alpha calculated using the following SAS code. For example, when calculating Cronbach's alpha for PROMIS Physical Function, the individuals' 10 item scores that constituted the Physical Function score will be used in the VAR clause.

Sample SAS code:

```
ODS OUTPUT CronbachAlpha=CronbachAlpha;
PROC CORR DATA = PF NOMISS ALPHA;
VAR pfa11 pfa56 pfa21 pfa53 pfa9 pfb28r1 pfa1 pfa6 pfb3 pfb44;
RUN;
```

19.8.2. Sensitivity Analysis of PROMIS

Sensitivity analysis will be conducted using Analysis of Covariance (ANCOVA) model for PROMIS Physical Function T-scores and Fatigue T-scores using Last Observation Carries Forward (LOCF) imputation for the FAS and the PPS, respectively. The LOCF imputation should be performed after any data handling for the intercurrent events as described in Section 12.3.3.6 and Section 12.3.3.7.

An ANCOVA model with treatment and stratification factors as fixed effects, the baseline score value as covariate will be used to assess treatment difference. The LSM and the corresponding 95% confidence intervals for the differences between roxadustat and the placebo group (roxadustat – placebo) will be presented along with the p-values for the treatment differences resulting from the model. The LSM for the difference will be considered clinically meaningful if it exceeds the MCID and will be marked in the results, see Section 19.8.1 for details of MCID.

19.9. Scope of Week 28 Primary Analysis

The following tables, figures and listing will be provided in Week 28 Primary Analysis for the double blind component. The date when last patient in the double blind component finished Week 28 visit is the cutoff date for Week 28 database lock. For efficacy TFLs, only primary endpoint at Week 28 will be analyzed. Data for adverse events up to the cutoff date for Week 28 database lock will be summarized. For other safety TFLs, data up to last subject's Week 28 will be included.

Table

Table 14.1.1.2 Subject Accountability and Disposition Double Blind Component (All Subjects)

Table 14.1.2.2 Important Protocol Deviations Double Blind Component (Full Analysis Set)

Table 14.1.3.1.2 Demographics and Baseline Characteristics Double Blind Component (Safety Analysis Set)

Table 14.1.3.1.3 Demographics and Baseline Characteristics Double Blind Component (Per Protocol Set)

Table 14.1.3.1.4 Demographics and Baseline Characteristics Double Blind Component (Full Analysis Set)

Table 14.1.4.2 Medical History Double Blind Component (Full Analysis Set)

Table 14.1.5.1.2 Exposure and Compliance Double Blind Component (Safety Analysis Set)

Table 14.1.5.2.2 Dose Modification Double Blind Component (Safety Analysis Set)

Table 14.1.6.2b Prior ESA Use and ESA Refractory Double Blind Component (Safety Analysis Set)

Table 14.2.1.2 Primary Efficacy Endpoint Double Blind Component (Full Analysis Set)

Table 14.2.2.1 Sensitivity Analysis I Double Blind Component (Per Protocol Set)

Table 14.2.2.2.1 Sensitivity Analysis III Double Blind Component (Full Analysis Set)

Table 14.2.5.1.3 Primary Efficacy Endpoint by Randomization Endogenous EPO Level Double Blind Component (Full Analysis Set)

Table 14.2.5.1.4 Primary Efficacy Endpoint by Randomization IPSS-R Risk Category Double Blind Component (Full Analysis Set)

Table 14.2.5.1.5 Primary Efficacy Endpoint by Randomization RBC Transfusion Burden Double Blind Component (Full Analysis Set)

Table 14.2.5.1.11 Primary Efficacy Endpoint by Prior ESA Use Double Blind Component (Full Analysis Set)

Table 14.3.1.1.2 Overall Summary of Treatment-Emergent Adverse Events Double Blind Component (Safety Analysis Set)

Table 14.3.1.2.2 TEAEs by System Organ Class and Preferred Term Double Blind Component (Safety Analysis Set)

Table 14.3.1.3.1 TEAEs by System Organ Class and Preferred Term and Maximum Severity Double Blind Component (Safety Analysis Set)

Table 14.3.1.4 TEAEs Related to Study Treatment Double Blind Component (Safety Analysis Set)

Table 14.3.1.5 TEAEs with CTCAE with Grade 3 or Higher Double Blind Component (Safety Analysis Set)

Table 14.3.1.6.2 Serious TEAEs Double Blind Component (Safety Analysis Set)

Table 14.3.1.7 TEAEs Leading to Discontinuation of Study Treatment Double Blind Component (Safety Analysis Set)

Table 14.3.1.8 TEAEs Leading to Death Double Blind Component (Safety Analysis Set)

Table 14.3.1.9 Most Frequently Occurring TEAEs Double Blind Component (Safety Analysis Set)

Table 14.3.1.11 Most Frequently Occurring Non-Serious TEAEs Double Blind Component (Safety Analysis Set)

Table 14.3.1.12 TEAEs of Special Interest Double Blind Component (Safety Analysis Set)

Table 14.3.1.13 TEAEs of Special Interest by SMQ Double Blind Component Double Blind Component (Safety Analysis Set)

Table 14.3.1.14 Serious TEAEs Leading to Discontinuation of Study Treatment Double Blind Component (Safety Analysis Set)

Table 14.3.2 Disease Progression to AML Double Blind Component (Safety Analysis Set)

Figure

Figure 14.2.1.1 Cumulative Number of PEW of TI at Week 28 Double Blind Component (Full Analysis Set)

Figure 14.2.1.2 Subject Exposure over Time up to Week 28 Double Blind Component (Full Analysis Set)

Figure 14.2.6.1 Mean (±SE) Hb (g/dL) from Baseline to Week 28 Double Blind Component (Full Analysis Set)

Figure 14.2.6.2 Mean (±SE) Hb Change (g/dL) from Baseline to Week 28 Double Blind Component (Full Analysis Set)

Figure 14.2.11.1 Swimming Plot of RBC Transfusion During the First 28 Weeks of Treatment Double Blind Component (Full Analysis Set)

Figure 14.2.10.2 Consort Diagrams Double Blind Component (Safety Analysis Set)

Listing

Listing 16.2.2.1.2 Subject Accountability and Disposition Double Blind Component (All Subjects)

Listing 16.2.2.2.2 Inclusion Criteria Not Met/Exclusion Criteria Met Double Blind Component (All Subjects)

Listing 16.2.2.3.2 Analysis Populations Double Blind Component (All Subjects)

Listing 16.2.2.3.3 Subjects Randomized from Incorrect Strata Double Blind Component (Full Analysis Set)

Listing 16.2.2.4.2 Important Protocol Deviations Double Blind Component (All Subjects)

Listing 16.2.3.1.2 Demographics Double Blind Component (Safety Analysis Set)

Listing 16.2.3.2.2.1 Baseline Characteristics I Double Blind Component (Safety Analysis Set)

Listing 16.2.3.2.2.2 Baseline Characteristics II Double Blind Component (Safety Analysis Set)

Listing 16.2.4.1.2 Medical History Double Blind Component (Full Analysis Set)

Listing 16.2.4.2.2b Prior ESA Use and ESA Refractory Double Blind Component (Safety Analysis Set)

Listing 16.2.5.2.2 Exposure Double Blind Component (Safety Analysis Set)

Listing 16.2.5.3.2 Treatment Compliance Double Blind Component (Safety Analysis Set)

Listing 16.2.6.1.2 RBC Transfusion Double Blind Component (Safety Analysis Set)

Listing 16.2.7.1.2 Adverse Events Double Blind Component (Safety Analysis Set)

Listing 16.2.7.2.2 Serious Adverse Events Double Blind Component (Safety Analysis Set)

Listing 16.2.7.3.2 AEs Leading to Death Double Blind Component (Safety Analysis Set)

Listing 16.2.7.4.2 AEs Leading to Permanent Discontinuation of Study Drug Double Blind Component (Safety Analysis Set)

Listing 16.2.7.5.2 Listing of Death Double Blind Component (Safety Analysis Set)

Listing 16.2.7.6.2 Treatment-Emergent Adverse Events in Subjects with any Liver Abnormalities Double Blind Component (Safety Analysis Set)

Listing 16.2.7.7.1 AEs of Special Interest Double Blind Component (Safety Analysis Set)

Signature Page for VV-CLIN-013690 v2.0

Approval Task	[redacted] 20-Apr-2023 19:38:02 GMT+0000
Approval Task	[redacted] 20-Apr-2023 20:51:58 GMT+0000
Approval Task	[redacted] 20-Apr-2023 21:15:57 GMT+0000
Approval Task	[redacted] 20-Apr-2023 21:20:13 GMT+0000
Approval Task	[redacted] 20-Apr-2023 21:41:24 GMT+0000
Approval Task	[redacted] 20-Apr-2023 22:00:32 GMT+0000
Approval Task	[redacted] 21-Apr-2023 00:06:03 GMT+0000

Signature Page for VV-CLIN-013690 v2.0