STATISTICAL ANALYSIS PLAN

Final Version 3.0, dated 19 September 2018

A Phase 3, Randomized, Open-Label, Active-Controlled Study to Evaluate the Efficacy and Safety of Roxadustat in the Treatment of Anemia in Chronic Kidney Disease Patients Not on Dialysis

ISN: 1517-CL-0610
EudraCT number: 2013-000951-42

Sponsor name
Astellas Pharma Europe B.V.
Sylviusweg 62
2333 BE Leiden
The Netherlands

Table of Contents

I. LIST OF ABBREVIATIONS AND KEY TERMS

List of Abbreviations

Abbreviations	Description of abbreviations
AE	Adverse Event
α1-AGP	Alpha 1-Acid Glycoprotein
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AnS	Anemia Subscale
anti-HCV Ab	Anti-hepatitis C Virus
APEB	Astellas Pharma Europe BV
Apo	Apolipoproteins
ASC	Analysis Set Classification
ASP1517	= FG-4592 (codename of investigational product) or roxadustat (international nonproprietary name)
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BIW	Twice weekly
BL	Baseline
BL Hb	Baseline Hemoglobin (please refer to key definitions for information)
BMI	Body Mass Index
BP	Blood Pressure
BUN	Blood Urea Nitrogen
CBC	Complete Blood Count
CHMP	Committee for Medicinal Products for Human Use
CHOIR	Correction of Hemoglobin and Outcomes in Renal Insufficiency
CHr	Reticulocyte Hemoglobin Content
CI	Confidence Interval
CKD	Chronic Kidney Disease
CL/F	Apparent total body clearance
CMH	Cochran-Mantel-Haenszel
C_{max}	Maximum concentration
CRF	Case Report Form
CRO	Contract Research Organization
CRP	C Reactive Protein
CSE	Composite Safety Endpoint
CSR	Clinical Study Report
CV	Cardiovascular
DBP	Diastolic Blood Pressure
DILI	Drug-induced Liver Injury
dL	Deciliter
DS	Data Science
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
eCRF	Electronic CRF
EDC	Electronic Data Capture
eGFR	Estimated Glomerular Filtration Rate
EH	Excessive Hematopoiesis

Abbreviations	Description of abbreviations
EoS	End of Study
EoT	End of Treatment
EPO	Erythropoietin
EQ-5D 5L	Health Related Quality of Life Questionnaire Consisting of Five Levels
ESA	Erythropoiesis Stimulating Agent
ESRD	End Stage Renal Disease
EU	European Union
EudraCT	Clinical trial database regulated by European Community
EWB	Emotional Well being
FACT-An	Functional Assessment of Cancer Therapy-Anemia
FACT-G	Functional Assessment of Cancer Therapy-General
FDA	Food and Drug Administration
FAS	Full Analysis Set
FG-4592	= ASP1517 (codename of investigational product) or roxadustat (international nonproprietary name)
FWB	Functional Well-being
g	Gram
GGT	Gamma Glutamyl Transferase
Hb	Hemoglobin
HbA1c	Hemoglobin A1c; Glycated hemoglobin
HBsAG	Hepatitis B Surface Antigen
Hct	Hematocrit
HD	Hemodialysis
HDF	Hemodiafiltration
HDL	High-density Lipoprotein
HEENT	Head, Eyes, Ears, Neck and Throat
HIF	Hypoxia-inducible Factor
HIF-PH	HIF Prolyl Hydroxylase
HIV	Human Immunodeficiency Virus
HRQoL	Health-Related Quality of Life
hs-CRP	High Sensitivity C-reactive protein
ICH	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICH E3	Guidance for Industry Structure and Content of Clinical Study Reports
ICH E9	Statistical Principles for Clinical Trials
ICH E14	Guidance for Industry – Clinical Evaluation of QT/QTc
IEC	Independent Ethics Committee
IERC	Independent Event Review Committee
IND	Investigational New Drug (Application)
INN	International Nonproprietary Name
INR	International Normalized Ratio
IRT	Interactive Response Technology
ISN	International Study Number
IV	Intravenous(ly)
Kg	Kilogram
LDL	Low-density Lipoprotein
LA-CRF	Liver Abnormality Case Report Form
LFT	Liver Function Tests
LLN	Lower Limit of Normal

Abbreviations	Description of abbreviations
LOCF	Last Observation Carried Forward
LSO	Last Subject Out
MACE	Major Adjudicated Cardiovascular Event (death, non-fatal myocardial infarction and/or stroke)
MACE+	MACE including, in addition, hospitalizations for either unstable angina and/or chronic heart failure
MAP	Mean Arterial Pressure
MDRD	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MI	Myocardial Infarction
mL	Milliliters
mmHg	Millimeters of Mercury
Mg	Microgram
MMRM	Mixed Model of Repeated Measures
MSAP	Meta-Analysis Statistical Analysis Plan
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NDD-CKD	Nondialysis-dependent chronic kidney disease
O	Optional
PCS	Physical Component Score
PD	Protocol Deviation
PEY	Patient-exposure year
PF	Physical Functioning
PGIC	Patients' Global Impression of Change
PK	Pharmacokinetic
PKAS	Pharmacokinetic Analysis Set
PPS	Per Protocol Set
PT	Preferred Term
PWB	Physical Well-being
QoL	Quality of Life
QRS	QRS interval
QTc	QT Interval corrected for heart rate
QTcB	QTc calculated according to Bazett's formula
QTcF	QTc calculated according to Fridericia's formula
QW	Once weekly
RBC	Red Blood Cell
RR	Respiratory Rate
RR Interval	Interval between successive Rs of the ECG
r-HuEPO	Recombinant Human Erythropoietin
RRT	Renal Replacement Therapy
SAE	Serious AE
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SBP	Systolic Blood Pressure
SC	Subcutaneous(ly)
SD	Standard Deviation
SI	International System of Units
SF-36	Short Form 36

Abbreviations	Description of abbreviations
SF-36 PCS	SF-36 Physical Component Score
SF-36 PF	SF-36 Physical Functioning
SF-36 MCS	SF-36 Mental Component Score
SF-36 VT	SF-36 Vitality
SmPC	Summary of Product Characteristics
SMQ	Standardized MedDRA Query
SOC	System Organ Class
SOP	Standard Operating Procedure
SQ	Subcutaneous
sTfR	Soluble Transferrin Receptor
SWB	Social Well-being
TEAE	Treatment-Emergent Adverse Event
TIBC	Total Iron-Binding Capacity
TIW	Thrice Weekly
TLF	Tables, Listings and Figures
TSAT	Transferrin Saturation (also known as FeSAT, iron saturation)
ULN	Upper Limit of Normal
USRDS	United States Renal Data System
VAS	Visual Analogue Scale
VT	Vitality
WBC	White Blood Cell
WHO-DD	World Health Organization Drug Dictionary
WHO-DRL	World Health Organization Drug Reference List
Wk(s)	Week(s)
WPAI:ANS	Work Productivity and Activity Impairment questionnaire: Anemic Symptoms

List of Key Terms

Terms	Definition of terms
Adverse Event	An adverse event is as any untoward medical occurrence in a subject administered roxadustat or darbepoetin alfa, and which does not necessarily have a causal relationship with this treatment.
Baseline	1) Observed values/findings which are regarded as calibrated zero status in the present study; 2) Time when 'Baseline' is observed.
Baseline Hb value	Baseline Hb is defined as the mean of all available central laboratory Hb values collected before or including the date of first study drug intake (pre-dose) [i.e. 4 latest values for patients enrolled under protocol version 1.0 and 2.0, and 3 latest values for those under protocol version 3.0].
Discontinuation	The act of concluding participation, prior to completion of all protocol-required elements, in a trial by an enrolled subject. Four categories of discontinuation are distinguished: a) dropout: Active discontinuation by a subject (also a noun referring to such a discontinued subject); b) investigator-initiated discontinuation (e.g., for cause); c) loss to follow-up: cessation of participation without notice or action by the subject; d) Sponsor-initiated discontinuation. Note that subject discontinuation does not necessarily imply exclusion of subject data from analysis. "Termination" has a history of synonymous use, but is now considered non-standard.
Efficacy Emergent Period	Evaluation period from the Analysis date of first dose intake up to End of Treatment visit or last non-missing Hb assessment (for subjects who died during the treatment period).
Endpoint	Event or outcome that can be measured objectively to determine whether the intervention being studied is beneficial. Primary and secondary variables supporting objectives of the study are called endpoints.
Enroll	To register or enter into a clinical trial; transitive and intransitive. Informed consent precedes enrollment, which precedes randomization.
Hb Correction	Hemoglobin level ≥11.0 g/dL and an Hb increase from baseline BL Hb ≥1.0 g/dL as assessed by the central laboratory at two consecutive study visits separated by at least 5 days in the Correction Period.
Hb Response	Hemoglobin level ≥11.0 g/dL and an Hb increase from baseline BL Hb ≥1.0 g/dL in case BL Hb was >8.0 g/dL, **or** an increase in Hb from BL Hb by ≥2.0 g/dL in case BL Hb was ≤8.0 g/dL, as assessed by the central laboratory at two consecutive study visits separated by at least 5 days.
Intervention	The drug, device, therapy or process under investigation in a clinical trial which has an effect on outcome of interest in a study: e.g., health-related quality of life, efficacy, safety, pharmacoeconomics.
Investigational period	Period of time where major interests of protocol objectives are observed, and where the test drug or comparative drug (sometimes without randomization) is usually given to a subject, and continues until the last assessment after completing administration of the test drug or comparative drug.

Terms	Definition of terms
Post study follow-up	Period of time from EOS visit to projected week 108 or until consent withdrawn. This period is only applicable to subjects who prematurely discontinued treatment. These subjects will be followed up every 6 months for vital status, serious adverse events (SAEs), cardiovascular and thromboembolic adverse events (AEs).
Pre-investigational period	Period of time before entering the investigational period, from the time of starting a subject enrolling into study until just before the first study drug intake.
Randomization	The process of assigning trial subjects to treatment or control groups using an element of chance to determine assignments in order to reduce bias (after randomization subjects receive either roxadustat or darbepoetin alfa from day 1 until End of Treatment (EOT)).
Roxadustat	International Nonproprietary Name (INN) of ASP1517/FG-4592 investigational product
Re-screening	Process of repeating screening, *See also screening and screen failure* If subject fails screening they may be re-screened once if deemed appropriate; all screening procedures will be repeated
Safety Emergent Period	Evaluation period from the Analysis date of first drug intake up to the minimum between [(Analysis Date of last dose + 28 days + x), EOS visit, Date of death], with x corresponding to additional days based on the last dosing frequency.
Screening	1) Process for retrieving candidates for the study. 2) Process for checking the eligibility of subjects usually done during the “pre-investigational period”
Screen failure	Screened subject, but did not fulfil protocol inclusion and/or exclusion criteria and failed to receive randomized or open label study treatment, or decided not to participate anymore (withdrew consent) prior to completing pre-investigational period
Screening Hb value	Mean of all available central laboratory Hb values collected during screening period [i.e. 3 latest values for patients enrolled under protocol version 1.0 and 2.0, and 2 latest values for those under protocol version 3.0].
Serious Adverse Event	An adverse event is considered “serious” if, in the view of either the investigator or Sponsor, it results in any of the following outcomes: results in death, is life threatening, results in persistent or significant disability/incapacity or substantial disruption of the ability to conduct normal life functions, results in congenital anomaly, or birth defect, requires in-subject hospitalization or leads to prolongation of hospitalization, or a medically important event.
Study period	Period of time from first subject screened to end of the last scheduled visit of the last subject randomized
Subject	An individual who participates in a clinical trial
Time to event	Time from a defined starting point (analysis date of first dose intake) to the

Terms	Definition of terms
	time of occurrence of the event of interest.
Time to censoring	Time from a defined starting point (analysis date of first dose intake) to the time of end of observation period in case the event did not occur.
Time to event analysis	Time to event analyses are statistical methods, such as survival analysis, that take into account 2 types of timing: the time to occurrence of an event (if an event occurred) and the time to censoring (if an event did not occur during the time we observed the subject). For time to censoring, we only know the total number of days in which the event didn't occur until the subject ceased to be followed (censored).
Treatment period	It is the period of time - between first dose of the test drug and last dose - where major interests of protocol objectives are observed, and where roxadustat (study drug) or darbepoetin alfa (comparative drug) is given to a subject. The treatment period consists of the Correction and Maintenance period.
Variable	Any quantity that varies; any attribute, phenomenon or event that can have different qualitative or quantitative values

1 INTRODUCTION

This Statistical Analysis Plan (SAP) contains a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and includes detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

The SAP version 1.0, dated 26 May 2014 was based on protocol version 1.0, dated 14 August 2013. This SAP is an amended version due to the protocol amendments (Amendment 1, dated 18 May 2015 and Amendment 2, dated 31 March 2016) as well as further harmonization with the other non-dialysis studies in this clinical program.

The SAP is finalized and signed prior to accumulation of substantial amount of data (open-label study) to ensure lack of bias. If needed, amendments to the approved SAP may be made prior to database hard lock. Amendments will be version controlled.

All details of the Pharmacokinetics Analysis Set (PKAS) will be described in a separate analysis plan, and a separate PKAS modeling report will be written.

This statistical analysis is coordinated by the responsible biostatistician of APGD. Any changes from the analyses planned in the SAP will be justified in the Clinical Study Report (CSR).

Prior to database hard lock, a final review of data and TLFs meeting will be held to allow a review of the clinical trial data and to verify the data that will be used for analysis set classification. If required, consequences for the statistical analysis will be discussed and documented. A meeting to determine analysis set classifications may also be held prior to database hard lock.

2 FLOW CHART AND VISIT SCHEDULE

Figure 1 Flow Chart

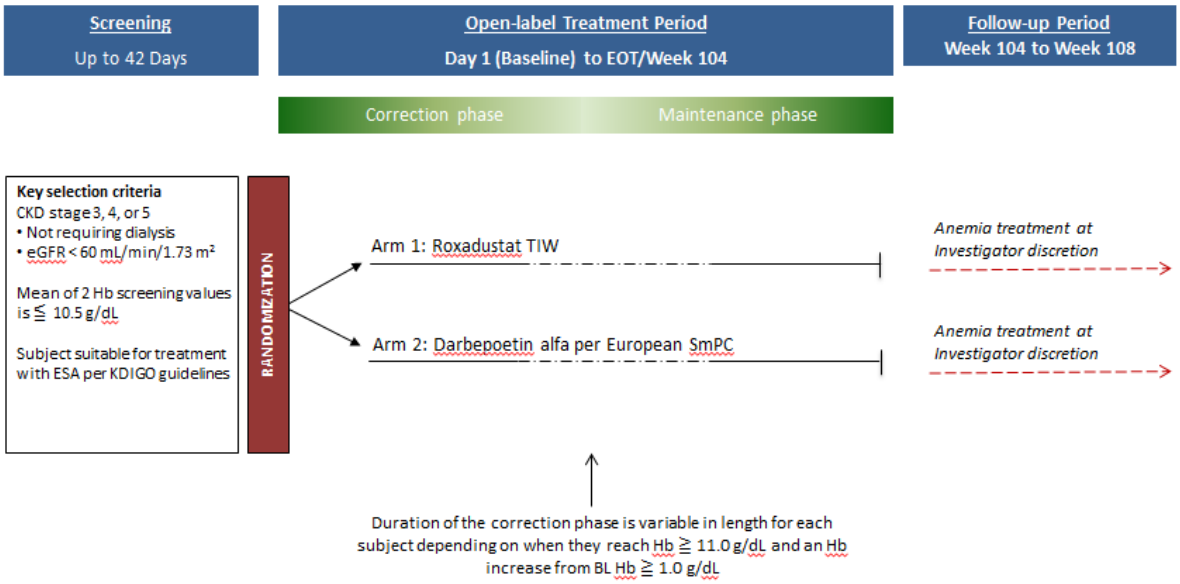


Table 1 Schedule of Assessments

Assessments	Screening Period		Treatment Period (Correction Period and Maintenance Period)					Follow-up Period		Unscheduled Visits	Post study Follow-up [c]
Visit / Week	Up to6 weeks [a]		Day 1[b]	Weekly (wks 1 - 2) ± 2 days	Every 2 Weeks (wks 4 - 24) ± 2 days	Every 4 Weeks (wks 28 - 100) ± 3 days	EOT [c] (wk 104) ± 3 days	EOT[c] + 2 wks ± 3 days	EOS [c] (EOT + 4 wks) ± 3 days		Every 6 months until projected wk 108
	S1	S2									
Written informed consent	X										
Randomization			X								
Eligibility criteria	X		X								
Demographics	X										
Medical history	X										
Physical examination	X		X		wks 12[d], 24[d]	wks 36[d], 52[d], 76[d]	X		X[d]	O[d]	
Height[e], weight	X		X		wks 12, 24	wks 36, 52, 76	X		X	O	
Blood pressure[f], heart rate[f], respiratory rate[g]	X	X	X	X	X	X	X		X	O	
CBC with WBC diff, red cell indices and platelets	X		X	X	wks 4, 8, 12, 20	wks 28, 36, 44, 52, 60, 68, 76, 84, 92, 100	X		X	O	
Hemoglobin[h]		X[i]			X	X	X	X		O[il]	
HemoCue assessment[j]			X	X	X	X				O	
Reticulocyte count and Hb in reticulocytes (CHr)	X		X	X	wks 4, 6, 8, 12, 16, 20	wks 28, 36, 44, 52, 60, 68, 76, 84, 92, 100	X		X	O	
12-lead ECG	X		X		wks 12, 24	wks 36, 52, 76	X			O	
Serum chemistry	X		X		wks 4, 8, 12, 20	wks 28, 36, 44, 52, 60, 68, 76, 84, 92, 100	X	X	X	O	
LFTs[k]				wk 2	wks 6, 16					O	
Renal ultrasound[l]	X									O	
Serum lipid panel[m]	X		X		wks 4, 8, 12, 20	wks 28, 36, 44, 52, 68, 84	X		X	O	
Serum iron, ferritin, TIBC, TSAT	X		X		wks 4, 8, 12, 20	wks 28, 36, 44, 52, 60, 68, 76, 84, 92, 100	X		X	O	
HbA1c	X		X		wk 12	wks 28, 36, 44, 60, 84	X		X	O	
Vitamin B_{12}, folate	X										
HIV immunoassay, HBsAg, anti-HCV antibody	X										
Serum pregnancy test[n]	X				wks 12, 24	wks 36, 48, 60, 72, 84, 96	X			O	

Table continued on next page

Assessments	**Screening Period**		**Treatment Period** (Correction Period and Maintenance Period)					**Follow-up Period**		**Unscheduled Visits**	**Post study Follow-up [c]**
Visit / Week	**Up to6 weeks [a]**		**Day 1[b]**	**Weekly (wks 1 - 2) ± 2 days**	**Every 2 Weeks (wks 4 - 24) ± 2 days**	**Every 4 Weeks (wks 28 - 100) ± 3 days**	**EOT [c] (wk 104) ± 3 days**	**EOT[c] + 2 wks ± 3 days**	**EOS [c] (EOT + 4 wks) ± 3 days**		**Every 6 months until projected wk 108**
	S1	**S2**									
eGFR[o]	X		X		wk 20	wks 36, 52, 68, 84	X		X	O	
hs-CRP			X		wks 4, 12, 20	wks 36, 52	X		X		
Archival serum samples for biomarkers[p]			X		wks 4, 12, 20	wks 52, 76					
Blood sample for PK[q]				wks 2 to 8							
Genotyping[r]				O							
QoL questionnaires[s]			X		wks 8, 12	wks 28, 36, 52, 76	X			O	
Urinary testing[t]			X		wks 12, 24	wks 36, 52, 64, 76, 88	X			O	
Archival urine samples for biomarkers			X		wk 24	wks 52, 76					
Study drug dispensing[u]			├	─	─	─ ┤				O	
Dose adjustment review[v]				X	X	X				O	
Hospitalization recording[w]	X	X	◄	─	─	─	─	─	─ ►		X
AE recording	X	X	◄	─	─	─	─	─	─ ►		
Concomitant medication, procedure and non-drug therapy recording	X	X	◄	─	─	─	─	─	─ ►		
Vital status, SAEs, cardiovascular and thromboembolic AEs[w]											X

S1/S2 = Screening visit 1and 2; EOT = End of Treatment; EOS = End of Study; Wk(s) = Week(s); X = mandatory test/assessment; O = optional test/assessment.

a The requirement of a 4-day separation between the screening Hb values defines the minimum duration of the screening period. The maximum duration of the screening period is 6 weeks. Sites are recommended to schedule the two screening visits in the shortest time span possible.

b All study assessments are to be performed prior to first study drug administration.

c In case of premature treatment discontinuation or withdrawal during the treatment period, subjects will complete the EOT visits (EOT visit and EOT + 2 weeks visit) and EOS visit. Thereafter, subjects who have taken at least one dose of study drug will continue to be followed up at a 6-monthly frequency for vital status, SAEs, cardiovascular and thromboembolic AEs until their projected date of completion of the follow-up period (i.e., projected week 108 date) or until consent withdrawn.

d Targeted physical examination only (e.g., respiratory and cardiovascular).

e Height measurement only required at first screening visit.

Footnotes continued on next page

f Blood pressure and heart rate measured singly during the screening period, and in triplicate at all other visits. It is recommended during the treatment period that measurements should occur prior to study drug administration if study medication is taken on same day of visit; except for visits where subjects are instructed to take study medication at home for PK sampling purpose. For subjects requiring dialysis, BP and heart rate will be recorded prior to and after dialysis (hemodialysis [HD]/hemodiafiltration [HDF] subjects only).

g Respiratory rate measured singly during screening period and all other visits. It is recommended during the treatment period that measurement should occur prior to study drug administration except for visits where subjects are instructed to take study medication at home for PK sampling purposes. For subjects requiring dialysis, respiratory rate will be recorded prior to dialysis (HD/HDF subjects only).

h Separate Hb should be collected at all the visits where Complete Blood Count (CBC) is not collected (i.e., Hb at weeks 6, 10 until the end of the study).

i An additional (third) Hb value may be collected if necessary. Only for subjects who are switched from protocol version 2.0 to 3.0 during (re) screening can a fourth Hb value be collected, as was applicable under protocol version 2.0.

i1 If during an unscheduled visit, Hb needs to be assessed, this should always be done with the HemoCue AND a central laboratory Hb assessment.

j If during an unscheduled visit Hb needs to be assessed, this should always be done with the HemoCue AND a central laboratory Hb assessment. This HemoCue assessment should be done on the blood sample that is collected for Hb analysis for the central laboratory.

k In addition to LFTs collected as part of Serum Chemistry, LFTs will separately be collected at the indicated weeks.

l Renal ultrasound examination within 12 weeks prior to randomization. Not required if results of a previous renal ultrasound (or other renal imaging modality such as CT scan or MRI) within 12 weeks prior to randomization is available and ruling out renal cell carcinoma. If other imaging modality, a conclusive report on the kidney should be available.

m Fasting whenever possible.

n Collect from female subjects of childbearing potential only.

o Using Modification of Diet in Renal Disease (MDRD) formula; calculated by the central laboratory.

p At day 1, week 20, 52, and 76, two equal volume samples should be collected.

q Sampling roxadustat will be done at 6 time points over 1 to 3 visits between weeks 2 and 8. See Section 5.6 of the protocol. At each pharmacokinetic visit, one additional sample will be collected for determination of Alpha 1-Acid Glycoprotein (α1-AGP) and albumin concentration.

r Optional assessment for subjects treated with roxadustat. A separate informed consent form must be signed before genotyping sample is collected. Sample collection can be done at any timepoint throughout the treatment period of the study.

s Including SF-36, FACT-An, EQ-5D 5L, PGIC and WPAI:ANS. PGIC will not be performed on day 1. Questionnaires to be completed by the subject preferably prior to any study assessments. When subjects need dialysis therapy, Quality of Life (QoL) questionnaires will be completed on the day of first dialysis (preferably before the dialysis is started), 4 weeks later and 12 weeks later.

t Ideally, the sample should be from the first morning void. Urinary testing includes qualitative testing with dipstick testing (for protein, pH, glucose) and quantitative assessment of albumin and creatinine for calculation of albumin/creatinine ratio.

u Dosing of darbepoetin alfa per EU SmPC.

v Subjects randomized to roxadustat: dose adjustment review from week 4 onward, and every 4 weeks thereafter until EOT (wks 4, 8, 12 etc), except in the event of excessive hematopoiesis or Hb ≥ 13.0 g/dL. If next dose adjustment interval falls on a non-visit study week, the dose adjustment review should be performed at the next scheduled visit.

w Telephone or in-person follow-up call with subject.

3 STUDY OBJECTIVE(S) AND DESIGN

3.1 Study Objective(s)

3.1.1 Primary Objective

The primary objective of this study is to evaluate the efficacy of roxadustat compared to darbepoetin alfa in the treatment of anemia in NDD-CKD subjects.

3.1.2 Secondary Objectives

The secondary objectives of this study are to:

- Evaluate the safety of roxadustat compared to darbepoetin alfa in the treatment of anemia in NDD-CKD subjects.
- Evaluate the HRQoL benefit of roxadustat compared to darbepoetin alfa in the treatment of anemia in NDD-CKD subjects.

3.2 Study Design

There are two protocol substantial amendments (Amendment 1, dated 18 May 2015 and Amendment 2, dated 31 March 2016).

The main protocol changes on the Amendment 1 are:

a) Changes in the study dosing regimen:
   - Dosing frequency changed from three times weekly (TIW), twice weekly (BIW) and once weekly (QW) to TIW only;
   - Initial study drug dose changed from 70, 100 and 150 mg to 70 and 100 mg only;
   - Maximum dose reduced from 3.5 mg/kg to 3.0 mg/kg and maximum absolute dose reduced from 400 mg to 300 mg;
   - Dose step of 120 mg was removed.

b) As a result of a), number of treatment arms were changed from 4 to 2 (roxadustat vs darbepoetin) and allocation ratio from 2:1 to 1:1

c) Reduction in visit schedule by removal of 12 study visits, resulting in total of 39 visits.

The main protocol changes on the Amendment 2 are:

a) Change in hemoglobin (Hb) eligibility criteria:
   - The number of Hb values assessed during the screening period is reduced from three to two and the mean Hb entry threshold is increased from ≤ 10.0 g/dL to ≤ 10.5 g/dL. This results in a change in inclusion criterion 4 and texts throughout the protocol that refer to the mean Hb entry threshold.

b) Update of the iron criteria: ferritin level ≥ 100 ng/mL and transferrin saturation (TSAT) level ≥ 20% at screening are removed.

c) Exclusion Criterion 21 extended to exclude patients who are likely to initiate renal replacement therapy including dialysis within the first year of the study in the opinion of the investigator.

3.2.1 General

This is a phase 3, multi-center, randomized, open-label, active-controlled study. This study is planned to recruit approximately 570 subjects from approximately 200 study centers globally.

The study is planned to provide key efficacy and safety data for the approval of roxadustat in the treatment of anaemia associated with CKD.

3.2.2 Study Population

The study population consists of subjects with CKD stages 3, 4, and 5 (eGFR < 60 mL/min/1.73 m^2) who are anemic and not on dialysis. Anemia is defined by mean Hb $\leq$10.5 g/dL upon repeated screening measurements. Anemia of non-renal origin is to be excluded. Washout periods of at least 12 weeks for any prior ESA, at least 6 weeks for any IV iron treatment, and at least 8 weeks for any RBC transfusion prior to randomization have been mandated in order to exclude a potential impact of these extraneous anemia treatments on the assessment of efficacy.

3.2.3 Description of Study

Subjects assigned to the roxadustat treatment arms will be administered roxadustat orally as a combination of tablets of different strengths. Subjects assigned to the darbepoetin alfa treatment will be administered darbepoetin alfa subcutaneously (SC) or intravenously (IV) by the investigator or a qualified member of the site staff or, after 36 weeks of treatment, by the subjects themselves or caregiver, e.g. relative, if well trained and willing to self-administer. Study treatment administration is implemented in an open-label manner.

The study will consist of three study periods as follows:

- Screening Period: up to 6 weeks
- Treatment Period: 104 weeks
- Follow-up Period: 4 weeks

Subjects that have discontinued treatment prior to their projected week 104 will continue to be followed for vital status, SAEs, cardiovascular and thromboembolic AEs in a post study follow-up period.

During the course of the study, visits and assessments will be performed as defined in the Schedule of Assessments.

Screening Period

During the Screening Period (up to 42 days duration), Hb levels will be assessed in a central laboratory for matching the inclusion criterion (mean of two Hb values must be $\leq$10.5 g/dL, see inclusion criterion 4). Subjects meeting all of the inclusion criteria and none of the exclusion criteria will be randomized at day 1, which marks the end of screening period and start of the treatment period. If a subject fails screening, the subject may be rescreened once (immediately or later) if deemed appropriate by the investigator. The subject must be

re-consented. A new rescreening period will start and all screening procedures must be repeated.

Treatment Period (Correction and Maintenance Period)

After subjects have been confirmed eligible for study participation, subjects are randomized to 1 of 2 treatment arms (1:1 ratio) as illustrated in Table 2. The randomization will occur through an Interactive Response Technology (IRT). The randomization will result in an 1:1 ratio of subjects receiving roxadustat administered orally or darbepoetin alfa administered by SC or IV injection.

Table 2 Treatment Arms and Dosing Frequency

Treatment Arms	Study Treatment	Dosing Frequency in Treatment Period	
		Correction Period	Maintenance Period
1	Roxadustat	TIW	TIW†
2	Darbepoetin alfa	Dosing per EU SmPC	

SmPC: Summary of Product Characteristics; TIW: three times weekly

† Subjects who were randomized under protocol v1.0 to QW and BIW will be converted to the TIW.

In protocol version 1 subjects were randomized in a ratio of 2:1 receiving roxadustat versus darbepoetin alfa. The expected number of subjects randomized on protocol version 1, protocol versions 2 and 3 and the total number of randomized subjects per treatment arm are represented in the following table.

Table 3 Expected Number of Subjects Randomized

Treatment Arms	Study Treatment	Expected Number of Protocol v1 Subjects Randomized (Ratio 2:1)	Expected Number of Protocol v2 and 3 Subjects Randomized (Ratio 1:1)†	Total Expected Number of Subjects Randomized per Treatment Arm
1	Roxadustat	100	210	310
2	Darbepoetin alfa	50	210	260
Total		150	420	570

† The number of subjects under protocol v2 and v3 will depend on the number of subjects randomized under protocol v1.

During the treatment period, subjects will attend weekly study visits from day 1 to week 2, followed by every other week study visits from weeks 4 to 24. From week 24 onwards, visits will occur every four weeks until the end of treatment (EOT).

Subjects will receive study treatment (roxadustat or darbepoetin alfa) for 104 weeks.

Correction Period

The aim of the correction period is to correct Hb levels to ≥ 11.0 g/dL and a Hb increase from BL Hb ≥ 1.0 g/dL as measured at two consecutive study visits separated by at least 5 days (as assessed by central laboratory). Dosing and dose adjustments of roxadustat will follow

prespecified dose adjustment rules. Dosing and dose adjustments of darbepoetin alfa will be per EU approved Summary of Product Characteristics (SmPC).

Once Hb is corrected the subjects will enter into the maintenance period.

Maintenance Period

After achieving Hb correction, the aim of the maintenance period is to treat to a Hb target level of 11.0 g/dL by maintaining the Hb levels between 10.0 g/dL and 12.0 g/dL. Dose adjustments of roxadustat will follow prespecified dose adjustment rules. Dose adjustments of darbepoetin alfa will be per EU approved Summary of Product Characteristics (SmPC).

Follow-up Period

After the end of the treatment period, subjects proceed to the 4-week follow-up period.

Post study Follow-up (*for premature treatment discontinued subjects only*)

Subjects that have prematurely discontinued study treatment will complete the EOT visits (EOT visit and EOT +2 weeks visit) and EOS visit. Thereafter, these subjects who have taken at least one dose of study drug will continue to be followed up every 6 months for vital status, SAEs, cardiovascular and thromboembolic AEs until their projected date of completion of the follow-up period (i.e., projected week 108 date) or until consent withdrawn.

3.2.4 Comparator

Darbepoetin alfa supplied as a solution for injection in a pre-filled syringe, and is administered by SC or IV injection. Once the route of administration is chosen (IV or SC), it is recommended to keep the same administration route for a subject throughout the study. It is administered in combination with IV or oral iron to maintain iron repletion. Darbepoetin alfa will be provided in pre-filled syringes of the following strengths: 20, 30, 40, 60 and 100 µg.

3.2.5 Interim analysis

An interim analysis will be performed when all subjects have completed 36 weeks of treatment. More details will be included in section 7.10

3.3 Randomization

A randomized design has been chosen in order to ensure a balanced allocation of study subjects to the treatment arms and to minimize bias in therapeutic management and in outcomes assessment.

Darbepoetin alfa will be administered by SC or IV injection, whereas roxadustat is administered orally. An open label design was chosen since the investigational and comparator drug are provided via different routes of administration and have a different requirement for iron supplementation.

Randomization and treatment assignments will be performed via IRT prepared on behalf of the Sponsor (under the responsibility of the Data Science (DS) Department of APEB).

Specific procedures for randomization through the IRT are contained in the study procedures manual. In protocol version 1 subjects were randomized in a ratio of 2:1 and in protocol versions 2 and 3 in a ratio 1:1 receiving roxadustat versus darbepoetin alfa.

Approximately 570 subjects will be randomized to receive roxadustat TIW or darbepoetin alfa as shown in Table 4 assuming 150 subjects will be randomized under the first version of the protocol.

Randomization will be stratified by the following four factors:

- Region: region A (Western Europe and Israel) versus region B (Central and Eastern Europe) *;
  ** Assignment to region will be determined based on health care comparability.*
- Screening Hb values (Hb ≤8.0 g/dL versus >8.0 g/dL);
- History of cardiovascular, cerebrovascular or thromboembolic diseases (Yes versus No);
- Screening eGFR (<30 mL/min/1.73 m^2 versus ≥30 mL/min/1.73 m^2).

4 SAMPLE SIZE

Approximately 570 subjects will be randomized to receive roxadustat or darbepoetin alfa as follows:

Table 4 Sample Size Calculations per Treatment Arm

Treatment Group	Protocol version 1 (Ratio 2:1)		Protocol versions 2 and 3 (Ratio 1:1)		Total	
	Randomized	PPS	Randomized†	PPS	Randomized	PPS
Roxadustat	100	80	210	168	310	248
Darbepoetin alfa	50	40	210	168	260	208
Total	150	120	420	336	570	456

† The number of subjects under protocol v2.0 and v3.0 will depend on the number of subjects randomized under protocol v1.

The sample size calculation is based on the primary endpoint: Percentage of Hb responders during the first 24 weeks of treatment without rescue therapy. Assuming that the Per Protocol Analysis Set (PPS) will consist of 80% of subjects in the Full Analysis Set (FAS), a number of 570 randomized subjects in the FAS will lead to approximately 456 subjects in the PPS.

Two hundred and forty eight (248) subjects for the roxadustat treatment group and 208 subjects for the darbepoetin alfa treatment group will provide at least 98% test power to demonstrate statistically non-inferiority of roxadustat versus darbepoetin alfa in the primary endpoint assuming that the proportion of subjects with response in both groups is the same and at least 80% and a non-inferiority margin for the difference of proportions of 15% (see justification below). The power for the sensitivity analysis of post-amendment 1 data (336 subjects) will be at least 93%.

Note that the actual number of patients randomized under protocol v2.0 and v3.0 will lead a slight increase in power (according to the actual recruitment at the time of the SAP amendment).

Justification of the non-inferiority margin

In general an appropriate choice of margin should provide both:

a) Assurance that the test drug has a clinically relevant effect greater than zero (placebo). This aspect of the choice of margin is discussed in EMEA Guideline on the choice of the non-inferiority margin; section III [EMEA, 2005].
b) Assurance that the test product is not substantially inferior to the reference in EMA guideline. This aspect of the choice of margin is discussed in EMEA Guideline on the choice of the non-inferiority margin; section IV [EMEA, 2005].

Two randomized pivotal placebo controlled studies are planned in this population (FGCL-4592-060 and 1517-CL-0608) that will directly prove superiority over placebo. Since direct comparisons provide always a higher level of evidence than indirect tests, the main criterion to fix the NI-margin in the 1517-CL-0610 study is bullet b above, i.e., to prove that the difference between roxadustat and darbepoetin alfa is unimportant.

From a clinical point of view, a true response rate for roxadustat of 15% lower than the true response rate for darbepoetin alfa is regarded an unimportant loss in efficacy since roxadustat is anticipated to have a number of advantages over ESAs that have been shown during the phase II program: less need for IV iron supplementation and no increase of blood pressure.

No regulatory guideline exists in this therapeutic indication recommending a non-inferiority margin. However, the selected non-inferiority margin of 15% is in line with the FDA guideline in the urinary tract infections indication, where a high responder rate was also expected and a non-inferiority margin of 15% was recommended [FDA, 1998].

5 ANALYSIS SETS

In accordance with ICH recommendations in guidelines E3 and E9, the following analysis sets will be used for the analyses.

Detailed criteria for analysis sets will be laid out in Classification Specifications (CS) and the allocation of subjects to analysis sets will be determined prior to database hard lock.

5.1 All Randomized

The All Randomized consists of all randomized (as documented in the IRT subjects and corresponds to the Intent-To-Treat (ITT) set defined in the protocol.

5.2 Full Analysis Set (FAS)

The Full Analysis Set (FAS) consists of all randomized subjects who received at least one dose of study drug and have at least one post-dose Hb value. Subjects will be assigned to their planned treatment provided by the IRS.

Criteria for FAS exclusion is defined as follows:

- No study drug taken, or
- No Hb value post-dose.

5.3 Per Protocol Set (PPS)

The Per-Protocol Set includes all FAS subjects who do not meet any of the reasons to exclude a complete subject from PPS listed in Table 5. This PPS will be used for the primary endpoint analysis, for selected secondary analyses and for all disposition, demography and baseline characteristics.

Table 5 Criteria for excluding a complete subject from PPS

Number	Reasons for exclusion from PPS
1	Subject who receives less than 2 weeks of study treatment.
2	Patient without a valid corresponding Hb. A valid corresponding Hb is defined as an Hb value from the central laboratory that is measured at least 2 weeks after the first dose and was either before the last study drug intake or at maximum three days after the last drug intake.
3	Prescribed study drug compliance during treatment < 75% during the first 24 weeks or until EOT, whatever comes first.
4	Violation of inclusion or exclusion criteria which may affect the assessment of the efficacy of the study drug during the first 24 weeks or until EOT, whatever comes first.
5	Administration of wrong randomization study drug for more than one week during the first 24 weeks or until EOT, whatever comes first.
6	Administration of prohibited concomitant medication affecting efficacy listed in Appendix 12.1 of the protocol during the first 24 weeks or until EOT, whatever comes first.
7	Administration of rescue therapy significantly deviating from the protocol during the first 24 weeks or until EOT, whatever comes first.

More information on the derivation of these criteria can be found in the Classification Specifications.

5.4 Safety Analysis Set (SAF)

The safety analysis set consists of all randomized subjects who received at least one dose of study drug. Subjects will be assigned to their actual treatment received during the trial.

The SAF will be used to describe demographic and baseline characteristics and all safety and tolerability related variables.

5.5 Pharmacokinetics Analysis Set (PKAS)

The PKAS will be defined in a separate analysis plan. Results of the population PK analysis will not be reported in the Clinical Study Report but in a separate population PK report.

6 ANALYSIS VARIABLES

6.1 Efficacy Endpoints

The **Efficacy Emergent Period** will be defined as the evaluation period from the Analysis date of first dose intake up to EOT Visit. In case a subject died during the treatment period, the end will be set to the last non-missing Hb assessment. This period will be used as reference period for the time to event analyses related to efficacy endpoints, unless specified

otherwise. More details on the derivation of the date of End of Efficacy Emergent Period are provided in section 7.11.6

6.1.1 Primary Efficacy Endpoint

Primary definition of Hb Response (Yes/No)

Hb ≥11.0 g/dL and Hb increase from baseline by ≥ 1.0 g/dL, for subjects with baseline Hb > 8.0 g/dL; or Hb increase from baseline by ≥ 2.0 g/dL, for subjects with baseline Hb ≤ 8.0 g/dL at two consecutive visits [dates] (with available data), separated by at least 5 days, during the first 24 weeks of treatment without having received rescue therapy (i.e., RBC transfusion for all subjects or darbepoetin alfa for roxadustat treated subjects) prior to Hb response.

Alternative definitions of Hb Response (Yes/No) used for sensitivity analyses:

- ***[regardless use of rescue therapy]***: similar to the primary definition except that all Hb values will be used regardless the use of rescue therapy during the first 24 weeks.
- ***[Modified definition, without use of rescue therapy]***: Since the latest protocol amendment 2.0 allows subjects with screening Hb between 10.0 and 10.5 to be enrolled, this alternative Hb response will be defined as Hb ≥11.0 g/dL only.

Baseline Hb is defined as the mean of all available central laboratory Hb values collected before or including the date first study drug intake (pre-dose) [up to 4 latest available values, the number of Hb values considered depends on the protocol version according to subjects will be randomized].

All scheduled and unscheduled Hb values from the central laboratory will be taken into account.

The first date of the two consecutive visits will be used as the date of response. Analysis visits will be used to define the week of response (see Section 7.11.4).

6.1.2 Key Secondary Efficacy Endpoints

The primary analysis set will be PPS for the non-inferiority tests and FAS for the superiority tests.

The key secondary efficacy endpoints in this study are listed in Table 6

Table 6 Key Secondary Efficacy Endpoints

Number	Endpoint
1	Hb change from BL to the average Hb in weeks 28-36, without having received rescue therapy (i.e. RBC transfusion for all subjects, or darbepoetin alfa for roxadustat treated subjects) within 6 weeks prior to and during this 8-week evaluation period
2	Change from BL in Low–Density Lipoprotein (LDL) cholesterol to the average LDL cholesterol of weeks 12-28
3	Time to first IV iron use during weeks 1 to 36 #
4	Change from BL in SF-36 Physical Functioning (PF) sub-score to the average PF sub-score in weeks 12 to 28
5	Change from BL in SF-36 Vitality (VT) sub-score to the average VT sub-score in weeks 12 to 28
6	Change from BL in Mean Arterial Pressure (MAP) to the average MAP value in weeks 20-28
7	Occurrence and time to first occurrence of hypertension (defined as [systolic BP ≥170 mmHg AND systolic BP increase from BL ≥20 mmHg] or [diastolic BP ≥110 mmHg AND diastolic BP increase from BL ≥15 mmHg]) during weeks 1 to 36

The key secondary endpoint, Mean monthly IV iron use (mg) per subject during weeks 1 to 36 (monthly defined as a period of 4 weeks) is updated as Time to first IV iron use during weeks 1 to 36. This is due to the small number of subjects who took IV iron during the first 36 weeks.

6.1.2.1 Hb change from BL to the average Hb in weeks 28-36

All available Hb values obtained from the central laboratory will be considered (i.e., both scheduled and unscheduled Hb values), but only Hb values within analysis visit windows at weeks 28, 32 and 36, will be used for the calculation of the average of weeks 28 to 36 (see Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

Similarly to the primary endpoint, this key secondary endpoint will be derived using different rules regarding the use of rescue therapy:

[without rescue therapy]: in case a subject requires rescue therapy within 6 weeks prior to and during this 8-week evaluation period, the reported Hb values collected during weeks 28 to 36 analysis windows falling within any rescue therapy period will be set to missing. Imputation of missing data will be performed using either MMRM or ANCOVA analyses, see Section 7.11.4 for more details).

[regardless rescue therapy]: all Hb values will be considered for the analysis, regardless of the use of rescue therapy.

Baseline Hb is defined in section 6.1.1

6.1.2.2 Change from BL in Low–Density Lipoprotein (LDL) to the average LDL cholesterol of weeks 12-28

The analysis will be done on all values (fasted and non-fasted) of Day 1 and scheduled visits up to analysis Week 28 within Efficacy Emergent Period.

All available scheduled LDL values will be considered (regardless the fasting status), but only LDL values in analysis visit windows at weeks 12, 20 and 28, will be selected for the calculation of the average LDL cholesterol of weeks 12-28 (see Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

For missing LDL imputation rules, refer to section 7.11.1. Baseline LDL is defined as the LDL value on Day 1. If this value is missing, the latest value prior to first study drug administration will be used.

This analysis will also be repeated for fasted values only as a sensitivity analysis.

6.1.2.3 Time to first IV iron use during weeks 1 to 36 and Mean monthly IV iron use per subject during weeks 1 to 36 period

The use of IV iron is recorded in the *Concomitant Medication* form of the eCRF.. All medications are coded with WHO-DD. Records selected will be those coded as ATC 3rd level =IRON PREPARATIONS and route is INTRAVENOUS.

Only use of IV Iron that was ongoing or started during the Efficacy Emergent Period will be taken into account with the note that prior IV iron treatment ending on Day 1 will not be counted. Subjects without a relevant concomitant medication record will be assumed that they used no IV iron, thus set to 0 mg.

The Mean Monthly IV iron use per subject (in mg) during the first 36 weeks is defined by the following formula:

$$\frac{\begin{array}{c}Total\ of\ IV\ iron\ use\ (mg)\ from\ \text{Analysis date of first dose intake} \\ to\ Min(Analysis\ Date\ of\ Week\ 36\ visit,\ \ Date\ of\ End\ of\ Efficacy\ Emergent\ Period)\end{array}}{\begin{array}{c}((Min(Analysis\ Date\ of\ Week\ 36\ visit,\ \ Date\ of\ End\ of\ Efficacy\ Emergent\ Period) \\ -Day\ 1\ Date) + 1)/\ 28\end{array}}$$

Monthly is defined as a period of 4 weeks or 28 days.

Additional variables for IV iron use consist of:

- Received IV Iron as binary variable (Yes/No) during the first 36 weeks, where "Yes" is defined as having at least one record selected during Min (Analysis date of Week 36, end of efficacy emergent period).
- Time to first use of IV Iron during the first 36 Weeks will be calculated (in years) as

  (Date of first use of IV Iron – Analysis Date of first dose intake + 1) / 365.25

Where 'Analysis Date of first dose intake' is defined in section 6.5.4. The date of censoring for a subject without the event of interest is defined as the Minimum [Analysis date of Week 36 visit, end of Efficacy Emergent Period].

6.1.2.4 Change from BL in SF-36 Physical Functioning (PF) sub-score to the average SF-36 PF sub-score value in weeks 12 to 28

For details on the derivation of the SF-36 PF subscore, refer to section 6.1.3.15.1.

All available SF-36 PF sub-scores will be considered i.e., both scheduled and unscheduled SF-36 PF values, but only SF-36 PF subscores in the analysis visit windows at weeks 12 and 28 will be selected for the calculation of the average PF sub-score of weeks 12 to 28 (see Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

For missing SF-36 PF imputation rules, refer to Section 7.11.1

Baseline subscore is the assessment from Day 1 visit (i.e. the last one before first dose intake).

6.1.2.5 Change from BL in SF-36 Vitality (VT) sub-score to the average SF-36 VT sub-score value in weeks 12 to 28

For details on the calculation of the SF-36 VT scale subscore, see section 6.1.3.15.1.

Similar rules as for Section 6.1.2.4 will be used for the calculation of the values of the SF-36 VT scale sub-score.

6.1.2.6 Change from Baseline in Mean Arterial Pressure (MAP) to the average MAP value in weeks 20-28

Blood pressure will be measured singly for the two visits during the screening period and in triplicate with a 2-minute interval for all other visits during the study. During the study, for systolic blood pressure (SBP) and diastolic blood pressure (DBP), the average of the triplicates will be calculated for each visit. If less than three readings are available, all available values will be used in the calculation of the average.

MAP will be derived for each visit from the above averaged SBP and the DBP using the following equation:

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

All available MAP values will be considered, i.e. both scheduled and unscheduled MAP values, but only values in analysis visit windows at weeks 20, 22, 24 and 28 will be selected for the calculation of the average MAP during weeks 20-28 (see Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses). For subjects who started dialysis, only assessments collected pre-dialysis will be used. Post-dialysis ones will only be listed and excluded from this analysis.

For missing data imputation rules, refer to section 7.11.1 Baseline assessment is the assessment on Day 1 (average of the three readings, if less than three readings are available,

the non-missing readings will be used in the calculation of the average). If the baseline assessment is missing, then the latest available value prior to first drug administration will be used.

6.1.2.7 Occurrence and time to first occurrence of hypertension during week 1 to 36

Occurrence of hypertension (i.e. increased blood pressure) is a binary variable (Yes/No), defined as:

- systolic BP ≥170 mmHg AND systolic BP increase from BL ≥20 mmHg, or
- diastolic BP ≥110 mmHg AND diastolic BP increase from BL ≥15 mmHg

during the period of week 1 to 36. The date of occurrence of hypertension is defined as the first date where SBP criterion or DBP criterion is met, whichever occurs first. All available blood pressure values within the Safety Emergent Period up to and including the analysis date of Week 36 will be taken into account. Baseline assessment is the assessment from Day 1. If this value is missing, then the latest available value from the screening period will be used. For subjects who started dialysis, only assessments collected pre-dialysis will be used. Post-dialysis ones will only be listed and excluded from this analysis.

The time to occurrence of hypertension for a subject with the event of interest will be calculated (in years) as:

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

Where 'Analysis Date of first dose intake' is defined in section 6.5.4. and where 'First event date' is the first date of the occurrence of hypertension.

The time to censoring for a subject without the event of interest is calculated as:

(Minimum [Analysis date of Week 36 visit, Date of last vital signs assessment during the Safety Emergent Period] – Analysis date of first dose intake + 1) / 365.25

Refer to sections 6.2 and 7.11.5 for more details regarding the Safety Emergent Period.

6.1.3 Additional Secondary Efficacy Endpoints

The additional efficacy endpoints are listed in Table 7

Table 7 Additional Secondary Efficacy Endpoints

Number	Endpoint
	Hb correction and maintenance
1	Hb change from BL to the average Hb value of weeks 28 to 52 regardless of rescue therapy.
2	Time (weeks) to achieve the first Hb response as defined by the primary endpoint.
3	Hb averaged over weeks 28 to 36, 44 to 52, 72 to 80, 96 to 104, without use of rescue therapy within 6 weeks prior to and during these evaluation periods.
4	Hb value and Hb change from BL Hb to each post-dosing time point.
Table continued on next page	

Number	Endpoint
5	Hb change from BL Hb to the average Hb value of weeks 28 to 36, 44 to 52, 72 to 80, 96 to 104 regardless of the use of rescue therapy.
6	Proportion of Hb values within 10.0 to 12.0 g/dL and ≥10.0 g/dL in weeks 28 to 36, 44 to 52, 72 to 80, and 96 to 104 without use of rescue therapy within 6 weeks prior to and during this 8-week evaluation period
7	Occurrence and time to first potential Excessive Hematopoiesis (EH) defined as Hb increase by >2.0 g/dL between any two visits within 4 weeks of treatment.
	Hospitalizations
8	Occurrence (number) of hospitalizations, number of days of hospitalization per patient-year exposure and time to first hospitalization.
	Rescue Therapy Use
9	Occurrence and time to first use of RBC transfusions, number of RBC packs per subject, volume of RBC transfused per subject.
10	Occurrence and time to first use of rescue therapy (composite of RBC transfusions [all subjects] and darbepoetin alfa use [roxadustat treated subjects only])
	Iron supplementation
11	Occurrence of iron supplementation. Mean monthly IV iron (mg) per subject during weeks 37-52 and weeks 53-104 (monthly defined as a period of 4 weeks). Time to first IV Iron during the efficacy emergent period. Use of oral iron.
	Change in Cholesterol Levels, Apolipoproteins
12	Change from BL to each post-dosing study visit in Total cholesterol, LDL/High-density Lipoprotein (HDL) ratio, Non-HDL cholesterol, Apolipoproteins A1 and B and ApoB/ApoA1 ratio.
13	Occurrence of average LDL cholesterol <100 mg/dL (2.59 mmol/L), calculated over weeks 12 to 28, and weeks 36 to 52 of treatment.
	Blood Pressure Effect
14	Occurrence of achieved antihypertensive treatment goal (SBP <130 mmHg systolic and DBP<80 mmHg) based on the mean SBP and mean DBP calculated over weeks 12 to 28 and 36 to 52 of treatment with study drug.
	HRQoL
15	Change from BL to the average value of weeks 12 to 28 and 36 to 52 (SF-36 Physical Component Score (PCS), Anemia Subscale ("Additional Concerns") of Functional Assessment of Cancer Therapy (FACT-An) Score, Total FACT-An and Trial Outcome Index (TOI) Scores, EQ-5D 5L VAS Score and Work Productivity and Activity Impairment (WPAI:ANS).
16	Patient Global Impression of Change (PGIC) at last assessment.
	Iron status, HbA1c, and CKD progression
17	Changes from BL to each study visit (when measured) in Serum ferritin, TSAT, iron, HbA1c level, fasting blood glucose, eGFR (including eGFR slope over time), Urine albumin/creatinine ratio. Time to (and proportion of subjects) with serum creatinine having doubled during the study and Proportions of subjects with ESRD. Occurrence and time to chronic dialysis start.

6.1.3.1 Hb change from BL to the average Hb value of weeks 28 to 52 regardless of the use of rescue therapy.

The same rules as defined in Section 6.1.2.1, apply for weeks 28 to 52 and regardless use of rescue therapy.

6.1.3.2 Time (weeks) to achieve the first Hb response as defined by the primary endpoint

6.1.3.2.1 Time (weeks) to achieve the first Hb response, without rescue therapy, as defined by the primary endpoint

Hb response is defined in section 6.1.1

For a subject without rescue therapy before Hb response, the time to achieve Hb response will be calculated (in weeks) as:

(First event date – Analysis date of first dose intake + 1) / 7

where 'First event date' is defined as 'Date of the first value that meets the criteria for response ' and 'Analysis date of first dose intake' is defined in section 6.5.4

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

(Min[Date of End of Efficacy Emergent Period, Date of initiation of rescue therapy, Analysis date of Week 24 visit] – Analysis date of first dose intake+ 1) / 7

With date of End of Efficacy Emergent Period defined in section 7.11.6

6.1.3.2.2 Time (weeks) to achieve the first Hb response, regardless of rescue therapy

This variable is similar to the one defined in section 6.1.3.2.1 but regardless of the use of rescue therapy.

Time to censoring will then be calculated (in weeks) as:

(Minimum [Analysis date of Week 24 visit, Date of End of Efficacy Emergent Period] – Analysis date of first dose intake + 1) / 7

With date of End of Efficacy Emergent Period defined in section 7.11.6

6.1.3.3 Hb averaged over weeks 28 to 36, 44 to 52, 72 to 80, 96 to 104, without use of rescue therapy within 6 weeks prior to and during the evaluation periods

All scheduled and unscheduled hemoglobin values that belong to each period will be taken into account for calculating the average using the analysis windows (defined in Table 31, Section 7.11.4).

In addition, the averages over weeks 28-36, 44-52, 72-80 and 96-104 will be categorized into the following categories:

- <10 g/dL,
- 10-12 g/dL,
- >12 g/dL,

- ≥10.0 g/dL.

In case a subject requires rescue therapy within 6 weeks prior to and during this 8-week evaluation period, all Hb assessments collected during that period will be set to missing.

6.1.3.4 Hb value and Hb change from BL to each post-dosing time point

All scheduled and unscheduled hemoglobin values that belong to each window will be taken into account using one value per analysis windows, as defined in Section 7.11.4

Baseline Hb is defined in section 6.1.1

At each planned visit, Hb will also be categorized into the following categories:

- <10 g/dL,
- 10-12 g/dL,
- >12 g/dL,
- ≥10 g/dL

6.1.3.5 Hb change from BL Hb to the average Hb value of weeks 28 to 36, 44 to 52, 72 to 80, 96 to 104 regardless of the use of rescue therapy

This variable is similar to the one defined in section 6.1.2.1, applied for these time points and regardless of the use of rescue therapy.

Proportion of Hb values within 10.0 to 12.0 g/dL and ≥10 g/dL in weeks 28 to 36, 44 to 52, 72 to 80, and 96 to 104 is calculated as per section 6.1.3.3

6.1.3.6 Categorical analysis of Hb values

The following endpoints will be analyzed: proportion of Hb values within 10.0-12.0 g/dL and ≥10 g/dL by time intervals, the percentage of time with Hb values falling in each Hb interval (<10.0 g/dL, within 10.0-12.0 g/dL, ≥10 g/dL, > 12.0 g/dL, > 13.0 g/dL and > 14.0 g/dL) during the Efficacy Emergent Period and the potential Excessive Hematopoiesis (EH).

Proportion of Hb values:

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

- Number of Hb values within 10.0-12.0 g/dL / Total number of Hb values*100
- Number of Hb values with ≥10 g/dL / Total number of Hb values*100

in weeks 28 to 36, 44 to 52, 72 to 80 and 96 to 104 and overall treatment period without use of rescue therapy within 6 weeks prior to and during this 8 week evaluation period. All scheduled and unscheduled hemoglobin values that belong to each period will be taken into account using the analysis windows defined in Section 7.11.4

Percentage of time:

The percentage of time each patient has a Hb value <10.0 g/dL, within 10.0-12.0 g/dL, ≥10 g/dL, > 12.0 g/dL, > 13.0 g/dL or > 14.0 g/dL will be calculated (as a percentage of the total of the length of time between the first and last Hb assessment during the evaluated period). The percentage of time will be calculated via linear interpolation. That is, if the

change in Hb category (for instance from within 10.0-12.0 g/dL to > 12.0 g/dL) occurs between two visits V0 and V1, the day of change will be calculated by:

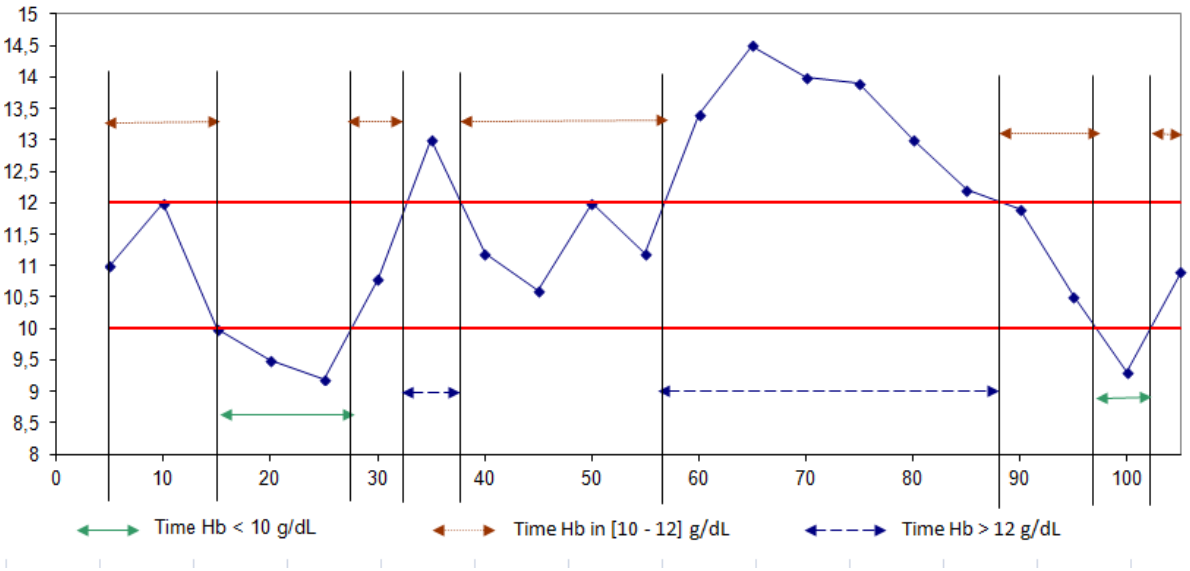
$$x = x_0 + (y - y_0)\frac{(x_1 - x_0)}{(y_1 - y_0)}$$

Where x_1 and x_0 are the dates when Hb was measured at V0 and V1 respectively, y_0 and y_1 are the Hb value at the respective visits V0 and V1 and y is the level of the Hb boundary (i.e 12.0, 13.0 or 14.0 g/dL).

Percentage of time each subject has Hb value ≥10.0 g/dL will be derived as 100% - percentage time for Hb values < 10 g/dL.

Figure 2 shows visually how the linear interpolation will calculate the total number of days that a subject is in each Hb category for an example subject:

Figure 2 Example of Linear Extrapolation



In case that several Hb values are on the same day the average of these values will be used to represent the Hb of that day in the above formula. This calculation will provide the day that the change in Hb value occurs. The number of days that the Hb value has been in each category will be determined and the percentage calculated based on the length of time between the first and last Hb assessment during the evaluated period, i.e.:

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

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

In case a subject requires rescue therapy within 6 weeks prior to and during these 8-week evaluation periods, refer to Section 7.11.1 for imputation rules.

6.1.3.7 Occurrence and time to first potential Excessive Hematopoiesis (EH)

Potential Excessive Hematopoiesis (EH), regardless use of rescue therapy, based on Hb central lab will be defined as Hb increase by >2.0 g/dL between any two visits within 4 weeks of treatment during the Efficacy Emergent Period.

Time to first occurrence of potential EH regardless the use of rescue therapy during the Efficacy Emergent Period will be defined in weeks as:

(First event date – Analysis date of first dose intake + 1) / 7

where 'First event date' is defined as first date of occurrence of the criterion met during the Efficacy Emergent Period.

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

Min[(Date of last Hb assessment during the Efficacy Emergent Period, End of Efficacy Emergent Period]] – Analysis date of first dose intake + 1) / 7

Refer to Section 6.1 and 7.11.6 for the definition of the Efficacy Emergent Period.

6.1.3.8 Occurrence (number) of hospitalizations, number of days of hospitalization per patient- exposure -year and time to first hospitalization

The occurrence and the number of hospitalizations per subject during the Efficacy Emergent Period will be calculated.

The number of days of hospitalization per patient- exposure-year (PEY) will be calculated as:

[Sum of the durations of all hospitalizations in days [Min((Date of discharge, End of Efficacy Emergent Period) – Date of admission + 1)] / [(Duration of Efficacy Emergent Period in days / 365.25)].

When hospitalization is ongoing, the date of end of the Efficacy Emergent Period will be used for the derivation of the hospitalization duration. In case of missing dates the hospitalization duration will be imputed by the average duration per stay derived from the subjects with non-missing duration within the same treatment group.

If the date of admission of a hospitalization record is the same as the date of discharge of the previous record, for example because two records are created to illustrate that the subject is moved from one hospital to another hospital or from a standard care to the intensive care unit (ICU), then the date of the transfer should not be counted twice and thus the hospitalizations duration for the later period is calculated as (Date of discharge – Date of admission). In such case hospitalization occurrence will also be counted only once.

Hospitalizations will also be described by reason for admission (admission for anemia or other reasons).

Time to first hospitalization in years will be defined in years as:

(First event date during the Efficacy Emergent Period – Analysis date of first dose intake + 1)/365.25

With 'First event date' defined as 'Date of first Admission' and 'Analysis date of first dose intake defined in section 6.5.4.

For a subject without hospitalization, the time to censoring will be calculated as:

[Date of End of Efficacy Emergent Period – Analysis date of first dose intake + 1) / 365.25

With date of End of Efficacy Emergent Period is defined in section 7.11.6

6.1.3.9 Occurrence and time to first use of RBC transfusions, number of RBC packs per subject, volume of RBC transfused per subject

The blood transfusion form of the eCRF in the cumulative visit will be used to derive all variables related to the use of RBC transfusions.

Only RBC transfusions that started during the study treatment and up to the End of Efficacy Emergent Period will be taken into account. Medication Onset Date is the date of the first use of RBC transfusion.

Volume of blood transfused and the total number of RBC packs (for each subject, the sum of blood volume and packs transfused during the Efficacy Emergent Period) will be calculated.

For RBC transfusions, when the number of units is not given but the volume transfused is given, the number of units will be estimated as volume transfused/250 mL (for transfusion of packed cell units) or volume transfused/500 mL (for transfusion of full blood).

For RBC transfusions, when the volume transfused is not given but the number of RBC units is given, the volume transfused will be estimated as number of RBC units times 250 mL (for transfusion of packed cell units) or as number of RBC units times 500 mL (for transfusion of full blood).

For subjects with use of RBC transfusion, the time to use of RBC transfusion is calculated as:

(First event date – Analysis date of first dose intake + 1) / 365.25

With 'First event date' defined as 'Date of first RBC transfusion' during the Efficacy Emergent Period and 'Analysis Date of first dose intake' defined in section 6.5.4

For a subject without use of RBC transfusion, the time to censoring is calculated as:

(Date of End of Efficacy Emergent Period – Analysis date of first dose intake + 1] / 365.25

With date of End of Efficacy Emergent Period defined in section 7.11.6

6.1.3.10 Occurrence and time to first use of rescue therapy (composite of RBC transfusions [all subjects] and darbepoetin alfa use [roxadustat treated subjects only])

Rescue therapy is defined as RBC transfusion for all subjects and darbepoetin alfa (SC or IV) for roxadustat-treated subjects. The latter information is derived from the Dosing CRF form.

Only rescue medication that started during the study treatment and up to the End of Efficacy Emergent Period will be taken into account and considered as use of rescue medication. Medication Onset Date is the date of the first use of rescue medication.

For a subject with use of rescue therapy, the time to use of rescue therapy will be calculated (in years) as:

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

With 'First event date' defined as 'Date of first dose of rescue medication' during the Efficacy Emergent Period and 'Analysis Date of first dose intake' defined in section 6.5.4.

For a subject without use of rescue therapy, the time to censoring is calculated as:

[(Date of End of Efficacy Emergent Period – Analysis date of first dose intake + 1] / 365.25

With date of End of Efficacy Emergent Period defined in section 7.11.6

Occurrence of rescue therapy will be defined as binary variable (RBC or ESA for roxadustat subjects) in case of any event during the efficacy emergent period.

6.1.3.11 Occurrence of iron supplementation

The occurrence of IV Iron use and mean monthly IV iron use per subject (in mg) for weeks 37 to 52 and 53-104 will be calculated similarly as in section 6.1.3.2. Analysis visits will be used as indicated in Table 31, section 7.11.4. Records selected will be those coded as ATC 3rd level = IRON PREPARATIONS and route is INTRAVENOUS.

Time to first IV Iron will be derived similarly to time to first use of rescue therapy, in section 6.1.3.10, with first event date corresponding to the Date of first dose of IV Iron during the efficacy emergent period. Censoring rules will be the same as for use of rescue therapy.

The occurrence of oral iron use will be also derived by the relevant periods (week 1-36, week 37-52, week 53-104) as a yes or no use of oral iron. Records selected will be those coded as ATC 3rd level = IRON PREPARATIONS and route is ORAL. This will be combined with IV iron use for these periods.

6.1.3.12 Change from BL to each post-dosing assessments in lipid parameters

For each sample the following will be calculated:

- LDL/HDL ratio (LDL Cholesterol divided by HDL Cholesterol)
- Non-HDL cholesterol (Total Cholesterol minus HDL Cholesterol)

Change from baseline to each post-dosing study visit will be calculated for the following lipid parameters:

1. Total cholesterol
2. LDL cholesterol
3. HDL cholesterol
4. Low-density lipoprotein (LDL) / high-density lipoprotein (HDL) ratio
5. Non-HDL cholesterol
6. Triglycerides
7. Apolipoproteins A1 and B (ApoA1 and ApoB)
8. ApoB/ApoA1 ratio.

All available data will be summarized descriptively for all parameters above, regardless of fasting status.

No imputation will be performed in case of a missing value. If several values are available in the same window, one value will be used, as defined in Section 7.11.4.

Baseline assessment is the assessment from Day 1 visit. If this value is missing, then the latest screening period value will be used as baseline.

6.1.3.13 Occurrence of mean LDL cholesterol <100 mg/dL (2.59 mmol/L), calculated over weeks 12 to 28, and weeks 36 to 52 of treatment

The evaluation period is defined as the average of all planned LDL cholesterol values in weeks 12-28 (visit at 12, 20 and 28 weeks) and weeks 36 to 52 (visit at 36, 44 and 52 weeks), as defined in Section 7.11.4. The occurrence of mean LDL cholesterol <100 mg/dL over weeks 12 to 28, and weeks 36 to 52 will then be defined as a binary variable (Yes/No), where ”Yes” is defined as mean LDL cholesterol <100 mg/dL over weeks 12 to 28, and weeks 36 to 52 respectively.

No imputation will be performed in case of a missing value.

This endpoint will be reported on fasting values and regardless of fasting status.

6.1.3.14 Occurrence of achieved antihypertensive treatment goal (SBP <130 mmHg systolic and DBP<80 mmHg) based on the mean SBP and mean DBP calculated over weeks 12 to 28 and 36 to 52 of treatment with study

Occurrence of achieved antihypertensive treatment goal (SBP< 130 mmHg and DBP< 80 mmHg) based on the mean SBP and mean DBP is calculated over an evaluation period defined as the average of all available values in weeks 12-28, similarly as in Section 6.1.2.6 (analysis windows defined in Section 7.11.4). Occurrence of achieved antihypertensive treatment goal will then be defined as a binary variable (Yes/No), where “Yes” is defined as SBP< 130 mmHg and DBP< 80 mmHg. It is calculated similarly for weeks 36 to 52.

No imputation will be performed in case of a missing value.

6.1.3.15 Change from BL to the average value of weeks 12 to 28 and 36 to 52 in Quality of Life scores

All study subjects will be required to complete Quality of Life (QoL) questionnaires as indicated in the schedule of assessments:

- SF-36
- FACT-An
- EQ-5D 5L
- WPAI:ANS

The next sections provide further details on how to derive these instruments, some derivations (SF-36 scales) will be provided by an external vendor (QualityMetric).

6.1.3.15.1 Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36)

The Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) is a multi-purpose, short-form health survey with 36 questions (see Appendix 1: SF-36 v2). It yields an 8-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures. It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group. Accordingly, the SF-36 has proven useful in surveys of general and specific populations, comparing the relative burden of diseases, and in differentiating the health benefits produced by a wide range of different treatments. The SF-36 contains 36 items that measure eight dimensions or scales: (1) physical functioning (PF); (2) role limitations due to physical health problems (RP); (3) bodily pain (BP); (4) social functioning (SF); (5) general health perceptions (GH); (6) role limitations due to emotional problems (RE); (7) vitality, energy or fatigue (VT); and (8) mental health (MH) (see Appendix 1: SF-36 v2). In addition, two summary measures, defined as the Physical Component Score (SF-36 PCS) and Mental Component Score (SF-36 MCS) will be provided.

Scoring of each dimension and the summary measure will be performed by QualityMetric using QualityMetric Health Outcomes(tm) Scoring Software 4.5.

Change from baseline to the average value in weeks 12-28 and 36-52 will be calculated for the Physical Component Scores of SF-36 (SF-36 PCS), following the same rules as defined in section 6.1.2.4

The evaluation period score is defined as the average of available SF-36 score values of weeks 12-28 (visit at week 12 and week 28) and weeks 36-52 (visit at week 36 and week 52).

For imputation rules, refer to section 7.11.1. Baseline always refers to the assessment at day 1 to be performed prior to first study drug administration.

The number and percent of subjects with an increase from baseline of <3 / ≥ 3 points and of <5 / ≥ 5 points will be calculated for the average of available SF-36 score values of weeks 12-28 (visit at week 12 and week 28) and weeks 36-52 (visit at week 36 and week 52) for the following: Vitality Score (SF-36 VT), Physical Functioning score (SF-36 PF) and Physical Component score (SF-36 PCS).

In addition, the eight dimensions and the two summary measures and their associated change from baseline will be reported by visit.

6.1.3.15.2 Functional Assessment of Cancer Therapy –Anemia (FACT-An)

The Functional Assessment of Cancer Therapy – General (FACT-G; version 4) contains 27 items that cover four dimensions of well-being: physical (PWB) – 7 items, functional (FWB) – 7 items, social/family (SWB) – 7 items, and emotional (EWB) – 6 items.

The 'additional concerns' section contains 20 items: 13 fatigue specific items plus 7 additional items related to anemia were developed for use in conjunction with the FACT-G (Cella 1997). The 13 fatigue items plus the seven additional items related to anemia comprise the Anemia Subscale (AnS). Administration of the FACT-G plus the Anemia Subscale (AnS)

is referred to as the FACT-An. The FACT-An has a recall period of the 'past seven days'. Respondents are asked to provide responses, (i.e., 'Not at all', 'A little bit', 'Somewhat', 'Quite a bit' and 'Very much'), to a list of statements which are either positively or negatively phrased. For all FACT-An scales, a higher score indicates better QoL (see Appendix 2: FACT-An (Version 4)

Each individual item is scored from 0 (Not at all) to 4 (Very much), and then the total score is obtained by summation of the resulted scores. Note that some scores need to be reversed for the derivation of the total score (see section 10.2.2).

If there are missing items, subscale scores will be standardized as long as there are more than 50% of items answered. This is done by multiplying the sum of the subscale by the number of items in the subscale, then dividing by the number of items actually answered. This can be done on the scoring guide or by using the formula below:

Prorated subscale score = [Sum of item scores] x [N of items in subscale] / [N of items answered]

The total score is then calculated as the sum of the un-weighted subscale scores, as long as overall item response rate is greater than 80%. In addition, total score should only be calculated if ALL of the component subscales have available scores.

The FACT-An instrument will be scored according to Appendix 2: FACT-An (Version 4) The following 9 scores will be calculated:

1. PWB subscale score
2. SWB subscale score
3. EWB subscale score
4. FWB subscale score
5. AnS subscale score
6. FACT-An TOI score
7. FACT-G total score
8. FACT-An total score
9. FACT-An Fatigue Score

Change from baseline to the average value in weeks 12-28 and 36-52 will be reported for the two scores (Anemia subscale 'Additional concerns' of FACT-An score and FACT-An total score). In addition, the score and change from baseline will be reported for each visit for all 9 scores.

The number and percent of subjects with an increase from baseline of <3 / ≥ 3 points in the Anemia Subscale score and a change from baseline of <7 / ≥ 7 points in the Total FACT-An score will be calculated for periods weeks 12-28 and weeks 36-52.

No imputation will be performed in case of a missing value.

Baseline always refers to the assessment at day 1 to be performed prior to first study drug administration.

6.1.3.15.3 EQ-5D 5L

The EQ-5D 5L is an international standardized non-disease specific (i.e. generic) instrument for describing and valuing health status, and a multi-dimensional measure of health-related QoL (see Appendix 3: EQ-5D 5L v2).

It includes two main components: (1) a VAS scale rating perception of overall health and (2) 5 qualitative domains: Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression. For more details and description of the questionnaire, refer to Appendix 3: EQ-5D 5L v2

Change from baseline to the average value in weeks 12-28 and 36-52 will be calculated for EQ-5D 5L VAS.

Frequency distributions will be described for each visit for:

1. EQ-5D 5L Mobility Score
2. EQ-5D 5L Self-Care Score
3. EQ-5D 5L Usual Activities Score
4. EQ-5D 5L Pain/Discomfort Score
5. EQ-5D 5L Anxiety/Depression Score

No imputation will be performed in case of a missing item. Baseline always refers to the assessment at day 1 to be performed prior to first study drug administration.

6.1.3.15.4 Work Productivity and Activity Impairment (WPAI: ANS)

The objective of the Work Productivity and Activity Impairment questionnaire: Anemic Symptoms v2 (WPAI: ANS) is to measure work and activity impairment during the past seven days due to anemia. It is self-assessed. The WPAI: ANS consists of 6 questions, including asking if the subject is working, how many hours the person missed work due to anemic symptoms, how many hours the subject actually worked and how the anemic symptoms impacted the productivity and ability to do daily activities (see Appendix 4: WPAI:ANS V2.0).

For subjects who are currently employed, the following four items and the change from baseline will be calculated:

- Percent work time missed due to anaemic symptoms: 100 x Q2/(Q2+Q4)
- Percent impairment while working due to anaemic symptoms: 100 x Q5/10
- Percent overall work impairment due to anaemic symptoms: 100 x [Q2/(Q2+Q4)+[(1-Q2/(Q2+Q4))x(Q5/10)]]
- Percent activity impairment due to anaemic symptoms: 100 x Q6/10

For these variables, no imputations will be performed in case one of a missing item. Baseline always refers to the assessment at day 1 to be performed prior to first study drug administration.

6.1.3.16 Patients' Global Impression of Change (PGIC)

The Patients' Global Impression of Change (PGIC) is a subject-rated instrument that measures change in subjects' overall status since the start of the study on a 7-point qualitative scale ranging from 1 (very much improved) to 7 (very much worse) (see Appendix 5: Patient Overall Impression of Change).

Data will be reported qualitatively by assessment as follows:

- Reported subject status,
- Combined Categories as binary:
  - Very Much Improved + Much Improved (yes/no)
  - Very Much Improved + Much Improved + Minimally Improved (yes/no)

No imputations will be performed in case of a missing item, except that the score at the last post-baseline assessment will be derived too.

6.1.3.17 Iron, HbA1c and CKD progression parameters

Changes from baseline to each study visit (see analysis windows in Section 7.11.4) will be calculated for these exploratory parameters:

1. Serum ferritin
2. TSAT
3. Serum Iron
4. HbA1c level (including by diabetes in the history)
5. Fasting blood glucose (including by diabetes in the history)
6. eGFR (including eGFR slope over time)
7. Serum Creatinine (log transformed)
8. Albumin/creatinine ratio in urine

Serum creatinine, eGFR and albumin/creatinine ratio in urine will be log transformed and any assessment occurring after the initiation of acute or chronic dialysis will be excluded for the summaries.

For all variables listed above, baseline assessment is the assessment from Day 1 visit. If this value is missing, then the screening period value, if collected for that parameter, will be used. 'Analysis date of first dose intake' is defined in Section 6.5.4 and Date of End of Safety Emergent Period is defined in section 7.11.5

CKD Progression exploratory variables

- ***Annualized eGFR slope over time:***

The annualized eGFR slope over time will be estimated in each treatment arm by the random slope and intercept model using all data available from each patient (one baseline and all post-treatment values up to start of acute or chronic dialysis). Change from Baseline to each study visit will be calculated.

- ***Doubling serum creatinine***:

Occurrence of serum creatinine being doubled compared with baseline at any moment during the Safety Emergent Period (see section 6.2) will be calculated.

The ratio of the maximum serum creatinine (on-treatment) by baseline value will be determined using the formula below:

Maximum on-treatment during Safety Emergent Period serum creatinine value / baseline value

Time to first occurrence of serum creatinine doubled compared with baseline will be derived as follows:

(First event date – Analysis date of first dose intake + 1) / 365.25

where 'First event date' is defined as first date of assessment where serum creatinine value is doubled compared to baseline during the Safety Emergent Period.

For a subject without event, the time to censoring will be calculated (in years) as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

- ***Time to Doubling serum creatinine or chronic dialysis or kidney transplant:***

First occurrence of serum creatinine being doubled compared with baseline during the Safety Emergent Period (see Section 6.2) will be calculated. In addition, first occurrence of chronic dialysis or kidney transplant during the Safety Emergent Period will be derived. The endpoint is defined as time to doubling serum creatinine or chronic dialysis or kidney transplant whatever comes first:

(First event date – Analysis date of first dose intake + 1) / 365.25

where 'First event date' is defined as first date of assessment where serum creatinine value is doubled compared to baseline or chronic dialysis or kidney transplant whatever comes first.

For a subject without event, the time to censoring will be calculated (in years) as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

- ***Occurrence of End Stage Renal Disease (ESRD):***

Defined as at least one of the following:

- Underwent >30 days of dialysis therapy
- Received kidney transplant
- Planned kidney transplant
- Physician recommended renal replacement therapy and subject refused therapy
- Began dialysis and died < 30 days later

- ***Occurrence and time to chronic dialysis or kidney transplant:***

Occurrence of chronic dialysis or kidney transplant during the Safety Emergent Period (see section 6.2) will be derived and time to first occurrence (i.e dialysis or kidney transplant, whichever occurs first) in years will be defined in years as:

(First event date – Analysis date of first dose intake) /365.25

where 'First event date' defined as Date of chronic dialysis or date of kidney transplant, whichever comes first.

For a subject without event, the time to censoring will be calculated as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

- ***Time to chronic dialysis or kidney transplant or Death***

Occurrence of chronic dialysis or kidney transplant (whichever occurring first) during the Safety Emergent Period (see Section 6.2) will be calculated.

Time to occurrence for a subject who died during the Safety Emergent Period (see Section 6.2) will be calculated.

For a subject, the time to chronic dialysis or renal transplant or occurrence for a subject who died will be calculated (in years) as:

(First event date – Analysis date of first dose intake + 1) / 365.25

where 'First event date' defined as 'first occurrence of chronic dialysis or kidney transplant, occurrence of subject who died (whichever occurring first)'.

For a subject without event, the time to censoring will be calculated (in years) as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

- ***Time to CKD progression (composite of doubling serum creatinine, chronic dialysis or kidney transplant or Death)***

First occurrence of serum creatinine being doubled compared with baseline during the Safety Emergent Period (see Section 6.2) will be calculated.

Occurrence of chronic dialysis or kidney transplant (whichever occurring first) during the Safety Emergent Period (see Section 6.2) will be calculated.

Time to occurrence for a subject who died during the Safety Emergent Period (see Section 6.2) will be calculated.

For a subject, the time to CKD progression will be calculated (in years) as:

(First event date – Analysis date of first dose intake + 1) / 365.25

where 'First event date' defined as 'First occurrence of serum creatinine being doubled compared with baseline, first occurrence of chronic dialysis or renal transplant, occurrence of subject who died (whichever occurring first)'.

For a subject without event, the time to censoring will be calculated (in years) as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

- ***Time to at least 40% decrease in eGFR from baseline***

First occurrence of at least 40% decrease in eGFR from baseline during the Safety Emergent Period (see Section 6.2) will be calculated.

For a subject, the time to at least 40% decrease in eGFR from baseline will be calculated (in years) as:

(First event date – Analysis date of first dose intake + 1) / 365.25

Where 'First event date' is defined as 'First occurrence of 40% decrease in eGFR from baseline.

For a subject without event, the time to censoring will be calculated (in years) as:

(min[Date of End of Safety Emergent Period, first occurrence of chronic dialysis date] – Analysis date of first dose intake + 1) / 365.25

- ***Time to at least 40% eGFR decrease from baseline or chronic dialysis or kidney transplant***

Event is defined as occurrence of at least 40% decrease in eGFR from baseline, or start of chronic dialysis or a kidney transplant.

First occurrence of at least 40% decrease in eGFR from baseline during the Safety Emergent Period (see Section 6.2) will be calculated.

For a subject, the time to at least 40% decrease in eGFR from baseline, chronic dialysis or renal transplant will be calculated (in years) as:

(First event date – Analysis date of first dose intake + 1) / 365.25

Where 'First event date' defined as 'First occurrence of 40% decrease in eGFR from baseline, first occurrence of chronic dialysis or renal transplant (whichever occurring first).

For a subject without event, the time to censoring will be calculated (in years) as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

6.1.4 Other exploratory variables: hs-CRP (High Sensitivity C-Reactive Protein) and sTFR (Soluble Transferrin Receptor)

The variables hs-CRP and sTFR will be collected from the central laboratory on the following visits: Day 1, weeks 4, 12, 20, 36, 52, EOT and EOS. Absolute values and changes from baseline to each study visit will be calculated. Baseline assessment is the assessment from Day 1 visit. If Day 1 assessment is missing, change from baseline will not be reported. Analysis windows are defined in Section 7.11.4

6.2 Safety Variables

Safety will be assessed by evaluation of the following variables:

- Treatment-emergent adverse events (TEAEs; frequency, severity, seriousness, and relationship to study drug),
- Vital signs (systolic and diastolic blood pressure, pulse rate, respiratory rate and weight),
- Clinical laboratory variables (hematology, biochemistry including liver function tests, and urinalysis),
- Local 12-lead electrocardiogram (ECG).
- Pre-specified adjudicated cardiovascular and cerebrovascular events.

The **Safety Emergent Period** will be defined as the evaluation period from the Analysis date of first drug intake up to the minimum between [(Analysis Date of last dose + 28 days + x), EOS visit date, Date of death], with x corresponding to additional days based on the last dosing frequency. This period will also be used to identify the minimum or maximum values collected on-treatment, defined as values collected from Day 2 up to the end of the Safety Emergent Period.

6.2.1 Adverse Events

6.2.1.1 Treatment emergent adverse event (TEAE)

TEAE is defined as an adverse event observed after starting administration of the test drug/comparative drug. If the adverse event occurs on Day 1 and the onset check box is marked “Onset after first dose of study drug” or the onset check box is left blank, then the adverse event will be considered treatment emergent. If the adverse event occurs on Day 1 and the onset check box is marked “Onset before first dose of study drug”, then the adverse event will not be considered treatment emergent. If a subject experiences an event both during the pre-investigational period and during the investigational period, the event will be considered as TEAE only if it has worsened in severity (i.e. it is reported with a new start date). Only adverse events starting during the Safety Emergent Period will be counted as TEAE. If an AE is reported with both start and stop dates are completely missing and the adverse event occurs on Day 1 check box is missing, then this will be considered as a TEAE.

A drug-related TEAE is defined as any TEAE with at least possible relationship to study treatment as assessed by the investigator or with missing assessment of the causal relationship.

Severity of AEs will be graded according to National Cancer Institute - Common Terminology Criteria for Adverse Events (NCI-CTCAE) Version 4.0.

For AE onset date imputation rules, refer to section 7.11.2

All adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.0.

6.2.1.2 Standardized MedDRA Queries

No standardized MedDRA Queries (SMQs) will be performed.

6.2.1.3 Time to occurrence of any TEAE (by type of AE group)

TEAEs are also classified into a number of groups depending on the following factors:

- Serious TEAEs
- Death during the Safety Emergent Period
- Any Death (during the 24-month Period)
- Related Serious TEAEs
- TEAEs Leading to permanent Discontinuation of the study drug
- TEAEs NCI CTC Grade 3 or Higher
- MedDRA System Organ Class (SOC)

The time to occurrence for a subject with a TEAE for a given type (except any death) will be calculated (in years) as:

(First TEAE date of the given type during the Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25

With 'Analysis date of first dose intake' defined in Section 6.5.4. All adverse events collected during the Safety Emergent Period will be counted as TEAE, irrespective of use of rescue therapy.

Subjects who have not experienced a TEAE for that given type will be censored; the subject will be censored and the time to censoring for these subjects will be calculated (in years) as:

(Date of End of Safety Emergent Period – Analysis date of first dose intake +1) / 365.25

With date of end of Safety Emergent Period defined in Section 7.11.5

For any deaths during the 24-month period (including those occurring during the Post-Study Follow Up Period), time to occurrence for a subject who died during study / post study follow up) will be calculated (in years) as:

[End date for corresponding Fatal AE – Analysis date of first dose intake + 1] / 365.25

With 'End date for corresponding Fatal AE' which cover both study and post-study FU occurring up to Day 760 (i.e month 24 + one month follow up + 3 days window as per protocol).

Subjects who have not died; the subject will be censored and the time to censoring for these subjects will be calculated (in years) as:

Minimum between [Day 760 and Max (Date of last known date when subject alive (End of Post-study FU), Date of last contact (Post study FU visit/call), Date of last Study evaluation (EOS form)) – Analysis date of first dose intake +1) / 365.25.

Additional analyses censoring data post dialysis:

Time to occurrence with censoring all data at the initiation of permanent dialysis will be calculated as above where events occurring after initiation of dialysis will not be considered.

Subjects who initiated dialysis and who have not experienced an event prior to the dialysis will be censored and the time to censoring for these subjects will be calculated (in years) as:

(Date of first dialysis – Analysis date of first dose intake +1) / 365.25

Subjects who have not experienced an event and did not initiate dialysis during the Safety Emergent Period, the time to censoring will be calculated (in years) as

(Date of End of Safety Emergent Period – Analysis date of first dose intake +1) / 365.25.

For dialysis date imputation rules, refer to Section 7.11.2.

6.2.1.4 Definition of incidence rate

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

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

Where Total cumulative time at risk is the sum of individual time at risk defined as either time to occurrence of the event or time to censoring for subjects with no event. Time to occurrence of the event and time to censoring are defined in section 6.2.1.3

Number of subjects at risk is defined as the number of subjects with (censored or non-censored) times to the event of interest greater or equal to t.

6.2.1.5 Definitions of event rate

The event rate (per 100 patient year) during the Safety Emergent Period will be calculated as either:

$$\frac{\text{Number of events}}{\text{Total Years Safety period}} \times 100$$

Where Total Years Safety period is defined as [Sum of individual sa in days (Analysis date of last dose – Analysis date of first dose + 1)] / [(Duration of Safety Emergent Period in days / 365.25)].

6.2.1.6 AE up to plus 7 Days

Additional analyses restricted to treatment-emergent AEs that observed after starting administration of the test drug/comparative drug and up to the minimum between [(Analysis Date of last dose + 7 days + x), EOS visit date, Date of death], with x corresponding to additional days based on the last dosing frequency will be defined.

Time to occurrence of event will be calculated (in years) as:

(First event date occurring as defined above – Analysis date of first dose intake + 1) / 365.25

With 'Analysis date of first dose intake' defined in Section 6.5.4

Subjects who have not experienced an AE for that given type will be censored; the subject will be censored and the time to censoring for these subjects will be calculated (in years) as:

Minimum [(Analysis Date of Last Dose + 7 days + x), EOS visit date, Date of Death)] – Analysis date of first dose intake +1) / 365.25.

6.2.1.7 Adjudicated CV related Events

The variables are based on treatment-emergent events that have been confirmed using the centralized adjudication, regardless of the term reported by the investigator. The adjudication will be performed blinded to treatment assignment.

The following composite endpoints will be evaluated:

- Major Adjudicated Cardiovascular Event (death, non-fatal myocardial infarction and/or stroke) (MACE): time to first occurrence of death, myocardial infarction or stroke.
- MACE including, in addition, hospitalizations for either unstable angina and/or chronic heart failure (MACE+): time to first occurrence of death, myocardial infarction, stroke, hospitalization for either unstable angina and/or congestive heart failure.
- CV-MACE: time to first occurrence of CV death, myocardial infarction or stroke.
- Cardiovascular (CV)-MACE+: time to first occurrence of CV death, myocardial infarction, stroke, hospitalization for either unstable angina and/or congestive heart failure.

The following individual event endpoints will be evaluated:

- Death
- Cardiovascular (CV death)
- Myocardial infarction
- Stroke
- Unstable angina that requires hospitalization
- Congestive heart failure that requires hospitalization
- Deep vein thrombosis and/or pulmonary embolism
- Vascular access thrombosis
- Hypertensive emergency.

6.2.2 Vital Signs

The following endpoints will be assessed:

- Systolic blood pressure (SBP)
- Diastolic blood pressure (DBP)
- Pulse rate
- Respiratory rate
- Weight

Single measurements for blood pressure (BP) will be taken at three visits during the screening period. Measurements will be taken in triplicate with 2-minute intervals for all other visits. An average will be calculated from the three readings. In case of missing values

among the triplicates, the available readings will be used. Note that the average in the eCRF system will not be used.

Change from baseline will be calculated for all data collected during the study as the measurement taken at the specific visit minus the measurement at baseline visit. Baseline assessment is the assessment from day 1 visit. If day 1 assessment is missing, screening period assessment will be used in the analysis. For all vital parameters, the minimum and the maximum post-baseline value during the Safety Emergent Period will be derived. For this calculation, only values from day 2 up to the date of End of the Safety Emergent Period (see section 7.11.5) will be considered.

From the assessments collected in the 'Vital Signs - HD/HDF Subjects Only' form, only the pre-dialysis ones will be used for the by-visit summaries and the definition of potentially clinical vital signs criteria. The post-dialysis assessments will be listed only.

Vital signs values are potentially clinically significant (PCS) if they meet both the observed value criteria and the change from baseline criteria listed in Table 8

Table 8 Potentially Clinically Significant (PCS) Vital signs Criteria

Vital Sign Parameter	Flag	Criteria	
		Observed Values	Change from Baseline
Respiratory Rate (breaths per min)	High	≥ 20	Increase of ≥ 5
	Low	≤ 10	Decrease of ≥ 5
SBP (mmHg)	High	≥ 170	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20
DBP (mmHg)	High	≥ 110	Increase of ≥ 15
	Low	≤ 45	Decrease of ≥ 15
Pulse (beats per min)	High	≥ 120	Increase of ≥ 20
	Low	≤ 50	Decrease of ≥ 20
Weight (kg)	High	-	Increase of ≥ 10%
	Low	-	Decrease of ≥ 10%

Potentially Clinically Significant Vital Signs Criteria will be calculated at each study visit at any moment during the Safety Emergent Period using the worst value among all available measurements.

Time to occurrence of a PCS (for each Vital Signs PCS criteria)

For each potentially clinically significant vital sign criteria (i.e. combined criteria), the time to occurrence of a PCS at any moment during the Safety Emergent Period (see section 6.2) will be calculated (in years) as:

(First occurrence date – Analysis Date of First Dose intake + 1) / 365.25

With 'First occurrence date' defined as the first date when both criteria (i.e on observed and change from baseline) are met and 'Analysis Date of first dose' defined in section 6.5.4

Subjects without abnormality will be censored and time to censoring for these subjects will be calculated (in years) as:

(Date of last vital signs assessment where the parameter analyzed is non-missing during the Safety Emergent Period – Analysis date of first dose intake +1) / 365.25

6.2.3 Clinical laboratory variables

6.2.3.1 Potentially Clinically Significant (PCS) Laboratory Criteria

Laboratory test values are potentially clinically significant (PCS) if they meet either the low or high PCS criteria listed in Table 9 below.

Table 9 Potentially Clinically Significant (PCS) Laboratory Criteria

Potentially Clinically Significant Laboratory Criteria			
Laboratory Parameter	**Unit**	**Low PCS Criteria**	**High PCS Criteria**
Alanine Aminotransferase (ALT)	U/L	no lower limit	> 3X ULN *> 5X ULN#* *> 8X ULN#* *> 10X ULN#* *> 20X ULN#*
Aspartate Aminotransferase (AST)	U/L	no lower limit	> 3X ULN *> 5X ULN#* *> 8X ULN#* *> 10X ULN#* *> 20X ULN#*
Alkaline Phosphatase (ALP)	U/L	no lower limit	*> 1.5 X ULN#* > 3 X ULN
ALT or AST	U/L for ALT or AST	no lower limit	ALT or AST > 8X ULN
Total Bilirubin	µmol/L	no lower limit	> 1.5 X ULN *> 2 X ULN#*
Moderate Liver Abnormality**	U/L for ALT and AST, µmol/L for Total Bilirubin	no lower limit	ALT and/or AST > 3X ULN or Total Bilirubin > 2X ULN
Severe Liver Abnormality**	U/L for ALT and AST, µmol/L for Total Bilirubin	no lower limit	ALT and/or AST > 3X ULN and Total Bilirubin > 2X ULN
Gamma Glutamine Transaminase (GGT)	U/L	no lower limit	> 3X ULN
Calcium	mmo1/L	< 0.8 X LLN	>1.2 X ULN
Creatinine	µmo1/L		>1.5 X Baseline
Potassium	mmo1/L	< 0.75 X LLN	>1.2 X ULN
Sodium	mmo1/L	< 0.9 X LLN	>1.1 X ULN
Total Protein	g/L	< 0.9 X LLN	>1.1 X ULN
Blood Urea Nitrogen (BUN)	mmo1/L		>1.5 X Baseline
Neutrophils	10^6/L	≤1000	
Platelet Count	10^9/L	≤100	≥700
White Blood Cell Count	10^9/L	≤2.5	≥15
Lipase	U/L	no lower limit	> 3X ULN or > 2 X Baseline
LLN: Lower limit of normal, value provided by the laboratory ULN: Upper limit of normal, value provided by the laboratory			

Footnotes appear on next page

** a subject's ALT and Total Bilirubin laboratory draw date or AST and Total Bilirubin laboratory draw date must occur on the same blood sample in order to be counted.
# Additional criteria required for summary of Liver Function Tests only (see section 7.5.2.1)

Potentially Clinically Significant Laboratory Criteria will be calculated at each study visit and on-treatment (see section 6.2) using the worst value among all available measurements, except for Moderate and Severe Liver Abnormalities which will be calculated on-treatment only.

Time to occurrence of an abnormality (for selected criteria)

Time to occurrence of an abnormality, will be derived only for the following PCS criteria:

- Hemoglobin <6 g/dL
- Hemoglobin >14 g/dL
- Alanine Aminotransferase (ALT) > 3X ULN
- Aspartate Aminotransferase (AST) > 3X ULN
- Total Bilirubin > 1.5 X ULN

For each potentially clinically significant laboratory criteria, the time to occurrence of an abnormality for a subject with an abnormality at any moment during the Safety Emergent Period (see section 6.2) will be calculated (in years) as defined in section 6.2.2

Time to censoring will be defined (in years) as:

(Date of last laboratory assessment where parameter analyzed is non-missing during the Safety Emergent Period – Analysis date of first dose intake +1) / 365.25

6.2.3.2 Laboratory assessments

For all laboratory parameters, the minimum and the maximum values on treatment (i.e. during Safety Emergent Period, see section 6.2) will be defined.

In addition, each laboratory result will be classified as low (L), normal (N), or high (H) at each visit according to the laboratory supplied reference ranges.

Change from baseline will be calculated as the measurement taken at the specific visit minus the measurement at baseline visit.

Baseline assessment is the assessment from day 1 visit, except for Hb. If Day 1 assessment is missing, the screening or unscheduled assessment that is closest prior to Day 1 will be used.

For the lipid panel and glucose parameter, two baseline values will be defined based on fasting status: regardless of fasting and fasted.

6.2.4 Physical Examination

A comprehensive physical examination will be conducted during the screening period and at the EOT visit and recorded in the source documents. This examination will include general appearance and the following body regions and systems: head, eyes, ears, neck and throat (HEENT), lungs, heart, chest and back, abdomen, genitourinary, extremities, skin and any other, if deemed necessary.

A targeted examination (e.g. respiratory and cardiovascular) will be conducted and recorded in the source documents.

Only the date of the physical examination will be recorded in the eCRF. There will be no table or listing for the physical examination. Any clinically relevant adverse change will be recorded as an AE in the eCRF.

6.2.5 12-lead Electrocardiogram (ECG)

The 12-lead ECG measurements will be performed on all subjects at specific times. A single ECG measurement will be taken with the subject in the supine position, after the subject has been lying quietly for 5 minutes. Clinically significant abnormalities will be reported as an AE.

The visit, ECG date, Pulse, RR Interval, PR interval, QRS Interval, QT Interval, overall interpretation and relevant comments will be recorded in the eCRF.

Baseline assessment is the assessment from day 1 visit. If day 1 assessment is missing, screening period assessment will be used in the analysis.

QTc interval will be calculated using both

- Bazett (QTcB = QT/(RR Interval)$^{1/2}$) and
- Fridericia (QTcF = QT/(RR Interval)$^{1/3}$) corrections,

where QT is in msec and RR Interval in seconds; and if RR is not available, it will be replaced with 60/HR.

For all ECG parameters, the maximum post-baseline value on treatment will be defined (i.e. during Safety Emergent Period, see section 6.2).

ECG values are potentially clinically significant (PCS) if they meet or exceed the upper limit values listed in Table 10 below.

Table 10 ECG Parameters classification

ECG Parameter	Classification
QTc interval (msec)	> 450 msec, > 480 msec, > 500 msec;
QTc interval change (msec)	> 30 msec and > 60 msec
QRS (msec)	≥ 150 msec
PR (msec)	≥ 250 msec

Time to occurrence of PCS ECG:

For the two QTc criteria (QTc > 500 msec; change from baseline in QTc > 60 msec), the time to occurrence (in years) for a subject with occurrence of the PCS at any moment during Safety Emergent Period (defined in section 6.2) will be calculated (in years) as defined in section 6.2.2

Time to censoring will be defined as:

(Date of last ECG assessment where parameter analyzed is non-missing during the Safety Emergent Period – Analysis date of first dose intake +1) / 365.25

6.3 Pharmacokinetic Variables

All details of the population PK analysis will be described in a separate analysis plan.

6.4 Pharmacodynamic Variables

Not applicable.

6.5 Other Variables

6.5.1 Eligibility criteria

Eligibility at screening will be recorded as a yes/no variable for each criterion. The date of the informed consent for the subjects will also be documented.

6.5.2 Demographic and Baseline Characteristic Variables

Demographic characteristics will be recorded at screening (sex, the day, month and year of birth, age, race, height and weight).

Collection of date of birth depends on local regulations. Day of birth will be recorded in the eCRF as the first of the month when the day is not allowed to be collected. In cases where only year of birth is allowed to be collected, day and month will be recorded in the eCRF as the first of January. Age will be recalculated in SDTM and ADAM datasets.

If D_B is the Date of Birth and D_{First} is the Date of First Dose intake, Age is Age= $(D_{First}-D_B+1)/365.25$. If the Date of First Dose intake is not available, Date of Informed Consent will be used.

Based on recalculated age, three categories will be defined:

- <65 years,
- 65 – 74 years,
- ≥75 years

Each subject's body mass index (BMI) will be calculated as:

$$BMI\ (Kg/m^2) = Weight\ (Kg)/\ [Height\ (m)]^2,$$

in which the height will be converted from cm into m by dividing by 100.

Tobacco history and use will be recorded at screening.

The average maximum quantity of tobacco per week will be calculated using the average maximum quantity and frequency filled in the CRF. If the frequency is "Day", the average maximum quantity of tobacco per week will be determined as follow:

average maximum quantity per day x 7

If the frequency is "/Month", the average maximum quantity of tobacco per week will be determined as follows:

average maximum quantity per month / 4.3482

Screening Hb will be defined as the mean of all available central laboratory Hb values collected during screening period [i.e. 3 latest values for patients enrolled under protocol version 1.0 and 2.0, and 2 latest values for those under protocol version 3.0].

Baseline Hb is defined in section 6.1.1. Based on the mean screening Hb value, two categories will be defined:

- ≤ 8.0 g/dL
- > 8.0 g/dL

History of cardiovascular, cerebrovascular or thromboembolic diseases at baseline will be defined using selected MedDRA preferred terms based on the recorded medical history on the Cardiovascular Disease History and other medical history eCRF. CV History will be categorized as:

- Yes
- No

Countries and Regions

Subjects will be enrolled from the following 29 countries:

Region A (Western Europe and Israel)

- Denmark
- Finland
- France
- Germany
- Ireland
- Netherlands
- Portugal
- Spain
- Sweden
- UK
- Austria
- Israel

Region B (Central and Eastern Europe)

- Croatia
- Czech Republic
- Latvia
- Slovakia
- Slovenia
- Serbia
- Russia
- Poland
- Romania
- Hungary

- Bulgaria
- Ukraine
- Georgia
- Macedonia
- Montenegro
- Bosnia
- Belarus

Randomization will be stratified by region using two categories:

- Region A: Western Europe and Israel
- Region B: Central and Eastern Europe.

Time to Treatment Discontinuation

Time to Treatment Discontinuation in years is defined as:

Time to Treatment Discontinuation (years) = (Date of EOT Visit – Analysis date of first dose intake +1)/365.25

In case a subject completed the treatment period, time to censoring will be calculated as:

(Date of EOT Visit – Analysis date of first dose intake + 1) / 365.25

Time in years from Diagnosis of Anemia

Time from diagnosis of anemia in years is defined as:

Time from Diagnosis of Anemia (years) = (Analysis Date of First Dose intake – Date of Diagnosis)/365.25

In case of partial dates, imputation rules apply and are detailed in section 7.11.2

Time in years from Diagnosis of CKD

Time from diagnosis of CKD in years is defined as:

Time from Diagnosis of CKD (years) = (Analysis Date of First Dose intake – Date of Diagnosis of CKD)/365.25

In case of partial dates, imputation rules apply and are detailed in section 7.11.2

Time from Diagnosis of Targeted Medical History

Onset date and the start date of analysis for the study drug are collected and the time from diagnosis of targeted medical history in years is defined as

Time from Diagnosis of Targeted Medical History (years) =

(Analysis Date of First Dose intake – Onset Date)/365.25

In case of partial dates, imputation rules apply and are detailed in section 7.11.2

This will be calculated for each patient who was diagnosed with the targeted medical history: hypertension, diabetes mellitus type 1, type 2 and combined, dyslipidemia and vascular access.

eGFR

eGFR will be provided by the central laboratory only for the selected visits as described in the schedule of assessments. It will be calculated by the central lab using the following 4-variable Modification of Diet in Renal Disease (MDRD) equation:

eGFR (in mL/min per $1.73m^2$) = 175 x (SCr in mg/dL)$^{-1.154}$ x (Age in years)$^{-0.203}$ x (0.742 if female) x (1.21 if African American)

where SCr = serum creatinine concentration.

Since SCr will be collected for all the selected visits above plus additional ones, eGFR will be derived using the same formula and the derived eGFR will be used for the analysis.

Screening eGFR will be classified in the following categories: < 30 mL/min/1.73 m^2 versus ≥ 30 mL/min/$1.73m^2$, and <10, 10-15-<30, 30-<45, 45-<60, and ≥60 mL/min/$1.73m^2$

Iron status at screening

Subjects will be classified in one of the following four groups according to the TSAT and ferritin levels collected at Screening (prior to first drug intake):

- ferritin < 100 ng/mL and TSAT < 20%
- ferritin < 100 ng/mL and TSAT ≥ 20%
- ferritin ≥ 100 ng/mL and TSAT < 20%
- ferritin ≥ 100 ng/mL and TSAT ≥ 20%

Regarding Iron Repletion at screening, both TSAT and Ferritin should be coming from the same blood sample in cases that we have more than one record.

Iron status at Baseline

Subjects will be classified in one of the following four groups according to the TSAT and ferritin levels collected on Day 1:

- ferritin < 100 ng/mL and TSAT < 20%
- ferritin < 100 ng/mL and TSAT ≥ 20%
- ferritin ≥ 100 ng/mL and TSAT < 20%
- ferritin ≥ 100 ng/mL and TSAT ≥ 20%

Regarding Iron Repletion at baseline, both TSAT and Ferritin should be coming from the same blood sample in cases that we have more than one record on Day 1. If no Day 1 assessment is available, Iron Repletion at Screening will be used.

Use of ESAs during the last year prior to start of study treatment

Subjects will be classified as either having used ESAs or not during the last year. Previous use of ESAs is collected in the Treatment History for Anemia eCRF.

6.5.3 Previous and concomitant medication

Previous medication is defined as a medication with at least one dose taken before the date of first dose of study drug.

Concomitant medication is defined as a medication with at least one dose taken between the date of first dose (inclusive) and the date of the End of the Safety Emergent Period.

Previous and concomitant drug use will be recorded, including non-prescription medication, complementary and alternative medications. Handling of missing date information for prior or concomitant medications is given in section 7.11.2

If the medication start date and end date are both missing, the medication will be counted as both previous and concomitant.

If the medication start date is missing and the end date is prior the date of first drug administration, the medication will be counted as previous medication.

If the medication start date is missing and the end date is after the date of first drug administration, the medication will be counted as previous and concomitant medication.

6.5.4 Variables related to study drug

Randomization Arms

Table 11 below presents the groups to which subjects are randomized under protocol version 1.

Table 11 Randomization arms under protocol version 1

Randomization Arm Code (ARMCD)	Randomization Arm (ARM)
1A	Roxadustat - TIW correction - QW maintenance
2A	Roxadustat - TIW correction - BIW maintenance
3A	Roxadustat - TIW correction - TIW maintenance
B	Darbepoetin alfa

Table 12 below presents the groups to which subjects are randomized under protocol version 2 or 3.

Table 12 Randomization arms under protocol versions 2 and 3

Randomization Arm Code (ARMCD)	Randomization Arm (ARM)
A	Roxadustat TIW
B	Darbepoetin alfa

For the statistical analysis, treatments will be pooled across roxadustat under both protocol versions.

Analysis Date of First Dose Intake

Date of First Study Drug Dose Intake is collected in the Day 1 visit in the Randomization eCRF. In case of a missing/partial date, the earliest available date will be used. It will be on the same day than the randomization date and before the next dose date.

Analysis Date of Last Dose

Date of Last Study Drug Dose is collected at the End of Treatment visit in the End of Treatment eCRF. When this date is not known, all efforts should be made to obtain the date from the investigator. If it is still missing after all efforts have been made, then the visit date of the End of Treatment Visit will be used as the Analysis Date of Last Dose. If subject is lost to follow-up and none of these dates are available then the date of the last available assessment during the study will be used.

If analysis date of last dose is missing due to fact that date of last dose is a partial date with day unknown (month and year populated), then minimum between the date of death and end of month for the partial date will be used.

Duration of exposure in days

Treatment Period

For each subject, the Length of Time on treatment period will be calculated in days, using the following formula:

Analysis Date of Last Dose – Analysis Date first dose intake + 1

Safety Emergent Period

For each subject, the Length of Time during the Safety Emergent Period (see section 6.2) will be calculated in days, using the following formula:

End Date of Safety Emergent Period – Analysis Date first dose intake + 1

Patient-Exposure Years (PEY)

Treatment Period

For each subject, PEY on treatment period will be calculated in days as the duration of exposure divided by 365.25.

Safety Emergent Period

For each subject, PEY during the Safety Emergent Period will be calculated in days as the duration of exposure divided by 365.25.

Switch to maintenance (Yes/No)

If a subject has the variable Dose Decision ="Switch to Maintenance" in the roxadustat and darbepoetin alfa eCRF forms, *Dosing Decision* then the subject will be classified as Yes. Otherwise, the subject will be classified as never switched to maintenance.

Subjects who achieved maintenance

When a subject achieves Hb ≥11.0 g/dL and an Hb increase from BL Hb ≥1.0 g/dL as measured at two consecutive visits [dates] (with available data), separated by at least 5 days, the subject will be classified as a subject having achieved maintenance. In addition, subjects will be classified as having achieved maintenance using Hb central lab data.

Amount of Prescribed Medication

Roxadustat

The number of milligrams prescribed at each visit (including unscheduled visits) is captured in the *Changes in Dosing* eCRF. The investigator reported dose and frequency will be used when available. When the investigator reported dose and frequency are not available, the IRS reported dose and frequency will be used.

Prescribed weekly dose is defined as prescribed dose x 3 as the prescribed frequency is TIW for study drug for patients randomized following protocol amendment 2.0 or amendment 3.0. For patients randomized prior protocol amendment 2.0, other prescribed frequencies may be entered in the eCRF and should be used for the calculation of the prescribed weekly dose.

Each visit will have an associated start and end date as follows:

- Each visit (including unscheduled) has an associated date. This is the start date.
- Each visit (including unscheduled) will have an associated end date. This date will be the date of the next consecutive visit [including unscheduled visits and EOT] minus 1 day. This is the end date.

Time periods of interest are defined below (monthly is defined as a period of 4 weeks or 28 days) as follows:

Table 13 Time periods of interest

Time Period	Analysis Start Day	Analysis End Day*
Week 4 (Month 1)	Day 1	Day 28
Week 8 (Month 2)	Day 29	Day 56
Week 12 (Month 3)	Day 57	Day 84
Week 16 (Month 4)	Day 85	Day 112
Week 20 (Month 5)	Day 113	Day 140
Etc….		
Week 104 (Month 26)	Day 701	Day 728
Overall Treatment Period	Day 1	End Day of Treatment Period
Day 1 - Week 24	Day 1	End Day of Month 6 (day 168)
Day 1 – Week 36	Day 1	End Day of Month 9 (day 252)
Day 1 – Week 52	Day 1	End Day of Month 12 (day 364)

*or EOT whichever is first

This will allow to calculate the amount of prescribed medication in a time period as the sum of the daily prescribed amount within the time windows defined above.

In addition, amount of prescribed medication in mg/kg will be calculated. To convert a dose given in mg into a dose in mg/kg, the body weight recorded at day 1 will be used.

Darbepoetin alfa

The same methodology will be used for darbepoetin alfa. Note that prescribed frequency for darbepoetin alfa are: once weekly, every two weeks and monthly. The unit for darbepoetin alfa is microgram (µg).

In addition, amount of prescribed medication in µg/kg will be calculated. To convert a dose given in µg into a dose in µg/kg, the body weight recorded at day 1 will be used.

Amount of Consumed Medication

Roxadustat

The amount of medication taken will be estimated based on the amount of dispensed and returned medication. The dispensed and returned Investigational Product Medication is captured in the *Study Drug FG4592 Accountability* eCRF which includes the following kit information:

- Kit strength, kit treatment, kit dispensed, date of kit dispensed, kit strength total number of tablets dispensed
- Kit returned, returned date, total number of tablets returned

For each kit, the following will be calculated:

- Start Day of Exposure for each kit: study day kit dispensed
- End Day of Exposure for each kit: study day kit returned – 1 Day
- Amount dispensed for each kit: kit strength x number of tablets dispensed
- Amount returned for each kit: kit strength x number of tablets returned
- Amount consumed for each kit: amount dispensed – amount returned
- Daily consumed dose for each kit : amount consumed/(end day of exposure-start day of exposure +1)

Note: if nothing has been returned it is assumed that the medication has been taken. The same methodology as described for Amount Prescribed (planned) Medication will apply to calculate amount of consumed medication for each time period by summing up the different daily consumed amount of the different kits on a given day (subjects will be dispensed more than one kit on a given visit).

In addition, amount of consumed medication in mg/kg will be calculated. To convert a dose given in mg into a dose in mg/kg, the body weight recorded at day 1 will be used.

Darbepoetin alfa

The same methodology will be used for darbepoetin alfa. Note that instead of tablets dispensed/returned, there will be syringes dispensed/administered. Administration date is not captured, therefore the start and end day of exposure will be defined as follows:

- Start Day of Exposure for each kit: study day kit dispensed
- End Day of Exposure for each kit: for kits administered, the next consecutive date that darbepoetin alfa was dispensed – 1 Day

In addition, amount of consumed medication in µg/kg will be calculated. To convert a dose given in µg into a dose in µg/kg, the body weight recorded at day 1 will be used.

Compliance

Compliance will be calculated for the time periods defined in Amount of Prescribed (planned) Medication. Compliance in % will be calculated for each time period as:

$$\frac{\text{Amount consumed during time period}}{\text{Amount prescribed during time period}} \times 100$$

The following compliance categories will be defined:

- less than 50% (significant drug noncompliance)
- at least 50%, less than 75% (moderate drug noncompliance)
- at least 75%, less than 125% (acceptable compliance)
- greater or equal 125% (drug over compliance)
- unknown

Dose change, Dose decrease, dose increase and dose hold for roxadustat and darbepoetin alfa

Dosing changes (decrease, increase and dose hold) are derived from the study drug accountability data recorded in the eCRF for both roxadustat and darbepoetin alfa.

A dose change is the change in the number of milligrams or micrograms on the dose per intake (for example from 200 mg to 250 mg as dose increase).

For example, for roxadustat, a change from 200 TIW to 250 TIW is a change of 600 mg to 750 mg per week which is regarded as an increase in intake dose.

The same applies for darbepoetin alfa, but not weekly as it could be administered weekly, bi-weekly or monthly. A monthly-dose change will be calculated instead.

7 STATISTICAL METHODOLOGY

7.1 General Considerations

- All statistical comparisons will be made using two sided tests at the α=0.05 significance level unless specifically stated otherwise. Null hypotheses for superiority will be of no treatment difference and corresponding alternative hypothesis will be two-sided. Null hypothesis for non-inferiority testing will be of inferiority of roxadustat treatment and will be one-sided at the α=0.025.
- All data processing, summarization, and analyses will be performed using SAS® Version 9.3 (SAS Enterprise Guide 4.3) or higher. Specifications for tables, data listings and figures (TLFs) formats can be found in the TLF Specifications for this study.
- All data will be summarized by treatment arm (roxadustat and darbepoetin alfa) and for the total, unless specified otherwise.

- 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.
- All percentages will be rounded to one decimal place and lined up by the decimal place. The percentage will be suppressed when the count is zero.
- Any p-values will be rounded to four decimal places and will be presented as '< 0.0001' if they are less than 0.0001 after rounding.
- For continuous variables, descriptive statistics will include the number of subjects (n), mean, standard deviation, median, minimum and maximum. Frequencies and percentages will be displayed for categorical data. Percentages by categories will be based on the number of subjects with no missing data, i.e. will add up to 100%. Number of missing values will be shown in the frequencies tables.
- All data included in summary tables, inferential analyses or figures will also be listed.
- Listings will be done on all randomized subjects and all assessments (all collected data in the eCRF will be listed except the physical examination data).
- Model checking will be performed using graphical outputs provided by the SAS procedures and will be provided as part of the CSR section 13.1.9 (Raw SAS outputs).

For the definition of subgroups of interest please refer to Section 7.8.

7.2 Study Population

For this section, unless specified otherwise, PPS refers to the analysis set which excludes from the FAS all the subjects criteria listed in Table 5. See Section 5.3 for more details.

7.2.1 Disposition of Subjects

The following subject data will be summarized and presented:

- Number of subjects with informed consent, who discontinued before randomization and randomized (overall only);
- Number and percentage of subjects randomized in each analysis set, by treatment arm and overall;
- Number and percentage of subjects who completed and discontinued treatment, by primary reason for treatment discontinuation by treatment arm for All Randomized, SAF, FAS and PPS;
- Number and percentage of subjects who completed and discontinued the study, by primary reason for study discontinuation and by treatment arm for All Randomized and SAF,
- Number and percentage of subjects who completed and discontinued the post study follow-up period for All Randomized and SAF, and
- Number and percentage of subjects excluded from PPS by reason for exclusion defined in section 5.3, by treatment arm for the FAS.

The following data will be presented graphically by treatment arm for the SAF and the FAS:

- Treatment discontinuation by reason using bar chart;
- Treatment discontinuation by time interval and reason using bar chart; and

- Time to treatment discontinuation using a Kaplan-Meier plot.

In addition, the following graphs will be done by treatment arm, for the FAS and PPS:

- Treatment discontinuation for lack of efficacy using a cumulative incidence plot;
- Treatment discontinuation for adverse event using a cumulative incidence plot.
- Treatment discontinuation for withdrawal by subject using a cumulative incidence plot.

Time intervals will be analyzed using the following categories:

  - Less than 2 weeks
  - At least 2 weeks, less than 4 weeks
  - At least 4 weeks, less than 12 weeks
  - At least 12 weeks, less than 24 weeks
  - At least 24 weeks, less than 36 weeks
  - At least 36 weeks, less than 52 weeks
  - At least 52 weeks, less than 78 weeks
  - 78 weeks or more
  - Unknown.

In addition, the randomization stratification strata from both sources (CRF and IRT) will be reported by treatment arm. Discrepancy between stratification from CRF and IRT will be summarized and the total number of patients with discrepancy overall and for each stratification factor will be provided.

Screening failures will also be listed.

All data collected during the Post-Study Follow up period will be listed by visit (i.e. type of contact, subject status and occurrence of overnight hospitalizations).

7.2.2 Protocol Deviations

Protocol deviations, as defined in the study protocol (Section 8.1.6: Protocol Deviations) will be assessed for all randomized subjects. The number and percentage of subjects meeting any criteria will be summarized for each criterion and overall, by treatment arm and overall, as well as by study site. Subjects deviating from a criterion more than once will be counted once for the corresponding criterion. Subjects who have more than one protocol deviation will be counted once in the overall summary. A data listing will be provided by site and subject.

The protocol deviation criteria will be uniquely identified in the summary table and listing. The unique identifiers will be as follows:

- PD1 - Entered into the study even though they did not satisfy entry criteria,
- PD2 - Developed withdrawal criteria during the study and was not withdrawn,
- PD3 – Received wrong treatment or incorrect dose (incorrect active ingredient, that excludes rescue therapy (darbepoetin alfa) as per protocol guidelines),
  - PD3_1- Received wrong treatment kit
  - PD3_2- Received incorrect dose
- PD4 – Received prohibited concomitant treatment

7.2.3 Demographic and Other Baseline Characteristics

Demographic and other baseline/screening characteristics as per section 6.5.2 will be summarized by descriptive statistics and frequency tabulations.

Number and percentage of subjects randomized in each country and site will be presented for the All Randomized, SAF, FAS and PPS.

Descriptive statistics for age, weight, body mass index (BMI) and height at baseline will be presented. Frequency tabulations for sex and race will be presented. Descriptive statistics and frequency tabulations will also be presented for the subgroup variables presented in section 7.8. Additionally, demographic and other baseline characteristics will be presented for the following variables:

- Baseline and Screening Hb value as continuous and categorical (≤ 8.0 g/dL versus >8.0 g/dL),
- Baseline and Screening eGFR value as continuous and categorical (< 30 mL/min/1.73 m^2 versus ≥ 30 mL/min/1.73m^2 , and < 10, 10-<15, 15-<30, 30-<45, 45-<60, and ≥60 mL/min/1.73m^2).
- History of previous treatment with ESA: Yes vs. No.
- LDL Cholesterol : <LLN; LLN-ULN; > ULN
- CRP : ≤ ULN; > ULN

Demographic and baseline characteristics summaries above will be done for All Randomized, the SAF, FAS and PPS. This table will be repeated for subjects randomized after Amendment 1.0 and 2.0 are implemented (separately).

Baseline values of primary and key secondary variables will be presented for the same analysis sets and populations mentioned above:

- Hb (g/dL)
- LDL Cholesterol
- SF-36 PF and VT subscores
- MAP

All Medical History will be analyzed using the SAF, as described below:

Medical History other than anemia, CKD, cardiovascular disease and targeted medical history are coded in MedDRA, they will be summarized by System Organ Class and Preferred Term on the SAF.

Anemia history, chronic kidney disease history, targeted medical history, cardiovascular-vascular access thrombosis history, tobacco history and family history of cardiovascular disease will be summarized on the SAF.

The number and proportion of subjects with each typical symptom for CKD will be described, as well as the number and proportion of subjects with each typical symptom for anemia on the SAF.

Demographic and baseline data will also be listed.

7.2.4 Previous and Concomitant Medications

Previous medications are coded with WHO-DD, and will be summarized by therapeutic subgroup (ATC 2nd level), chemical subgroup (ATC 4th level) and preferred WHO name for the SAF.

As with previous medication, concomitant medication will be summarized similarly for the SAF. Subjects taking the same medication multiple times will be counted once per medication.

Treatment history for anemia will be summarized separately.

Missing dates imputation rules are detailed in section 7.11.2.

7.3 Study Drugs

Roxadustat (=FG-4592 Investigational Medicinal Product) is for oral administration and supplied as red coated oval tablets of 20, 50 and 100 mg.

Darbepoetin alfa (=Comparator Medicinal Product) is administered by SC or IV injection, and supplied as a solution for injection in a pre-filled syringe of 20, 30, 40, 60 and 100 µg.

7.3.1 Exposure

The following information on drug exposure will be presented by treatment arm and overall, for the SAF.

Exposure related variables are defined in 6.5.4

Descriptive statistics will be produced for:

- The amount of drug (roxadustat or darbepoetin alfa) the subject was exposed to during the treatment period (in mg or µg and in mg/kg or µg/kg), at study start, by month, and during the first 24, 36, 52 weeks and overall treatment period;

Duration of exposure will be summarized in two ways:

- Descriptive statistics and frequency tabulations will be presented on treatment period and on Safety Emergent Period;
- Exposure time will be categorized by treatment period and by Safety Emergent Period according to the following categories (and frequency tabulations):
    - Less than 2 weeks
    - At least 2 weeks, less than 4 weeks
    - At least 4 weeks, less than 12 weeks
    - At least 12 weeks, less than 24 weeks
    - At least 24 weeks, less than 36 weeks
    - At least 36 weeks, less than 52 weeks
    - At least 52 weeks, less than 78 weeks
    - 78 weeks or more

- Unknown.

In addition Patient-Exposure Year (PEY) will be calculated both on treatment period and on Safety Emergent Period.

For roxadustat, box-plots of average weekly prescribed dose (mg and mg/kg) by 4-weekly periods will be produced. A separate box-plot for darbepoetin alfa in another page will be produced, using a different unit (µg instead of mg).

Number and percentage of subjects with dose changes, increases, decreases or hold (any, one, two) will be summarized for both correction and the maintenance periods separately.

Study drug medication will also be listed showing for each subject and each visit the dispensed kit numbers and the actual medication in each kit. For instance, if a subject randomized to darbepoetin alfa, at one visit was mistakenly dispensed a kit containing roxadustat then the listing will show that roxadustat was given to the subject on the intended visit.

7.3.2 Treatment Compliance

Overall compliance with the dosing schedule will be examined for subjects in the SAF whose amount of dispensed and returned study drug and first and last days of treatment are known.

Percent overall compliance will be summarized in two ways, by month, and during the first 24, 36, 52 weeks and during the overall treatment period:

- Descriptive statistics will be presented,
- Percent compliance categories will be categorized according to the categories defined in section 6.5.4

Counts and percentages of subjects in each of these categories will be summarized.

Results will be displayed by treatment arm and overall.

7.4 Analysis of Efficacy

For all continuous efficacy variables, in addition to inferential analyses, descriptive statistics will be produced for the values and for the changes from baseline (BL) by visit.

Similarly, for all categorical efficacy variables, frequencies and proportions will be produced by analysis visit.

For this section, PPS refers to the analysis set which excludes from the FAS all the subjects criteria listed in Table 5. See Section 5.3 and Classification specifications for more details.

Analysis visits are detailed in section 7.11.4

Missing data imputation rules are detailed in section 7.11.1

7.4.1 Analysis of Primary Endpoint

7.4.1.1 Primary Analysis of the Primary Endpoint

The primary efficacy endpoint will be analyzed using the PPS.

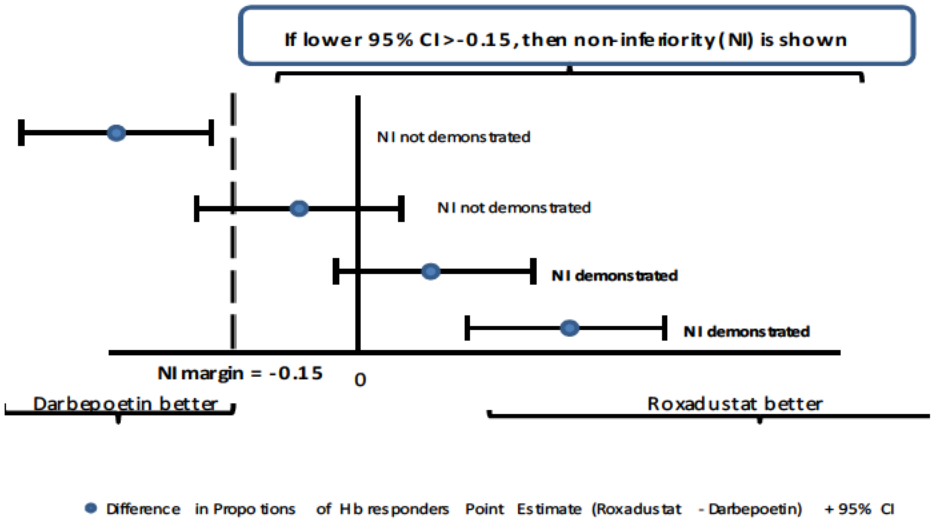
The proportion of responders in the primary efficacy variable will be compared using a Miettinen & Nurminen (MN) approach, adjusting for covariates and comparing roxadustat to darbepoetin alfa.

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

> H_0: Hb responder rate in the roxadustat group < Hb responder rate in the darbepoetin alfa group minus 15%
>
> versus
>
> H_1: Hb responder rate in the roxadustat group ≥ Hb responder rate in the darbepoetin alfa group minus 15%

The point estimates for the proportion of responders in the roxadustat group and the darbepoetin group and the difference in proportions between these rates will be calculated (roxadustat – darbepoetin). The Miettinen and Nurminen method will be used to calculate the two-sided 95% CI for the difference in rates, adjusting for stratification factors (covariates defined below). The null hypothesis stated above will be rejected if the difference in proportions of responders between the roxadustat group and darbepoetin alfa group lies entirely above -15%. If the resulting lower bound of the two-sided 95% CI between roxadustat and darbepoetin alfa is > -0.15, non-inferiority will be concluded (see figure 2 below).

Figure 2: Demonstrating Non-inferiority (NI)



The covariates will be:

- Region (Western Europe including Israel vs. Central and Eastern Europe)
- Baseline Hb values (≤8 g/dL vs. >8 g/dL)
- History of cardiovascular, cerebrovascular or thromboembolic diseases (Yes vs. No)
- Baseline eGFR (<30 mL/min/1.73 m^2 vs. ≥30 mL/min/1.73 m^2)

These covariates are also the randomization stratification factors. The analysis will use the

collected eCRF data to derive these covariates. Note that log-transformed eGFR might be used as covariate in the model instead of eGFR if it is demonstrated that eGFR is not normally distributed.

The SAS code will be similar to the one below:

```
proc freq;
  tables  region*CV_history* Hb_category*eGFR*arm*response/
riskdiff (cl=MN common);
run;
```

The stratified MN method has recently been implemented in SAS version 9.4. The current SAP states use of SAS version 9.3 or higher (see section 7.1). Therefore, depending on the availability of SAS version 9.4 at the time of analysis, the stratified MN statistics could be replaced by the alternative methodology based on the standard normal statistic proposed by Gart and Nam (see reference in section 9).

The SAS code will be similar to the one below:

```
proc genmod;
  class arm region CV_history;
  model response=eGFR_continuous region CV_history
Hb_continuous arm /dist=bin link=identity;
  lsmeans arm / diff cl;
run;
```

In addition, a 95% confidence interval for the proportion of roxadustat and darbepoetin alfa responders based on the exact method of Clopper-Pearson will be calculated and presented. Sensitivity analyses will be conducted to assess the consistency of the results before and after protocol amendments 1 and 2 (e.g., protocol versions 2.0 and 3.0 respectively).

Model checking:

Homogeneity among strata (covariates defined above) will be tested using the Gailsimon test. This test will assess the qualitative interaction between treatment effect and subjects' subsets in strata, i.e. assess if the direction of true treatment difference varies among subsets of subjects.

The SAS procedure will be similar to the following:

```
proc freq;

tables covariates*Treatment*response / gailsimon;

run;
```

A forest plot will be generated showing strata on the y-axis and differences in proportions and their 95% confidence interval on the x-axis. Stratas with 5 subjects or less may be combined.

7.4.1.2 Secondary Analyses (sensitivity) of the Primary Endpoint

The following analyses will be done as sensitivity analyses of the primary endpoint:

- The primary analysis [without rescue therapy] will be repeated on the FAS (S1);
- Sensitivity analysis [regardless of use of rescue therapy] will be performed on the PPS (S2);
- Sensitivity analysis [regardless of use of rescue therapy] will be performed on the FAS (S3);
- The primary endpoint will be analyzed using logistic regression including the covariates, on both PPS [without rescue therapy] (S4) and FAS [without rescue therapy] (S5). Baseline Hb (defined in 6.1.1) and baseline eGFR [or log(eGFR)] (defined in 6.1.3.17), will be included as continuous variables instead of categorical variables;
- Sensitivity analysis [regardless of use of rescue therapy] using logistic regression as mentioned above will be performed on the PPS (S6) and on the FAS (S7);
- The primary analysis [without rescue therapy] will be performed on the subgroup of subjects being randomized after the implementation of protocol v2.0 and v3.0 (S8). Analysis will be performed separately for each set of subjects.
- The primary analysis [without rescue therapy] on PPS will be performed using the modified definition of Hb response [i.e using only absolute value ≥ 11.g/dL] (S9).
- The primary analysis [without rescue therapy] on FAS will be performed using the modified definition of Hb response [i.e using only absolute value ≥ 11.g/dL] (S10).

For all sensitivity analyses except the primary analysis and primary analysis repeated excluding data from the site(s) with potential data quality issues, no hypothesis testing will be done, only confidence intervals presented. For the logistic regressions, the odds ratio (roxadustat vs. darbepoetin alfa) and their 95% confidence intervals will be produced, if convergence achieved. The non-inferiority margin of 15% will be transferred into an odds ratio margin as an indication only. Assuming a response rate 65% for roxadustat and 80% for darbepoetin with N=248 and N=208 respectively, this odds ratio will be 0.50.

The SAS procedure will be similar to the following:

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

Table 14 summarizes all sensitivity analyses to be performed with the primary endpoint. Table 14 also lists the sensitivity analysis to be performed excluding data from the site(s) with potential data quality issues. Data from this site(s) will not be excluded from the safety analyses.

Table 14 Primary and sensitivity analysis for the primary endpoint

Code	Set	Response Definition	Model	Covariates	Analysis is repeated excluding data from the site(s) with potential data quality issues
Primary	PPS	Without rescue therapy	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	No
S1	FAS	Without rescue therapy	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	Yes
S2	PPS	Regardless of rescue therapy,	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	No
S3	FAS	Regardless of rescue therapy	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	Yes
S4	PPS	Without rescue therapy	Logistic	Region, CV History, baseline eGFR, baseline Hb as continuous covariates	No
S5	FAS	Without rescue therapy	Logistic	Region, CV History, baseline eGFR, baseline Hb as continuous covariates	Yes
S6	PPS	Regardless of rescue therapy	Logistic	Region, CV History, baseline eGFR, baseline Hb as continuous covariates	No
S7	FAS	Regardless of rescue therapy	Logistic	Region, CV History, baseline eGFR, baseline Hb as continuous covariates	Yes
S8	PPS*	Without rescue therapy	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	No
S9	PPS	Modified definition, without use of rescue therapy	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	No
S10	FAS	Modified definition, without use of rescue therapy	MN CI	Region, History of CV, Baseline Hb, Baseline eGFR	Yes

*patients who were randomized after the implementation of protocol v2.0 and v3.0 (analysis will be performed separately for each set of subjects).

If a relevant baseline variable is identified for which a clinically important imbalance exists at baseline between treatment groups, additional sensitivity analyses of the primary endpoint

may be performed using a logistic model adjusting for this baseline variable. This will allow us to assess the impact of these imbalances on the treatment comparisons.

7.4.1.3 Additional Analyses of the Primary Endpoint

The analysis of the primary endpoint will be repeated by subgroup of interest. For definitions of subgroups of interest, see section 7.8

For the primary endpoint, subgroup analyses will be performed using the primary analysis on the PPS and the third sensitivity analyses (S3 in Table 14) on the FAS. If one of the subgroups is the same as a stratification factor, the factor will be omitted from the model.

Table 15 Additional Analyses of the Primary Endpoint

Code	Set	Response Definition	Endpoint*	Method	Covariates
A1	PPS	Without rescue therapy	Response by Subgroup	MN CI	Region, history of CV, Baseline Hb and Baseline eGFR[§]
A2	FAS	Regardless of rescue therapy	Response by Subgroup	MN CI	Region, history of CV, Baseline Hb and Baseline eGFR[§]

* Subgroup: (1) age, (2) sex, (3) region, (4) baseline hemoglobin category, (5) history of CV, (6) Baseline eGFR category, (7) Baseline CRP group (8) Iron repletion status.

[§] In case model does not converge for some subgroups, no adjustment on covariates will be done.

Subgroup analysis will be done by producing separate summaries similar to those produced for the primary analysis. In addition, forest plots will be generated per each subgroup showing subgroup factors on the y-axis and differences in proportions and their 95% confidence interval on the x-axis.

The potential existence of subgroup by treatment interaction will be visually inspected.

7.4.2 Analysis of Key Secondary Endpoints

The primary analysis set for the analysis of the key secondary endpoints will be the PPS for the non-inferiority tests and the FAS for the superiority tests.

All inferential analyses will evaluate treatment groups: roxadustat vs. darbepoetin alfa, unless specified otherwise.

Once the primary hypothesis has been rejected (section 7.4.1.1), the secondary variables will be tested using a fixed sequence testing procedure, as depicted in Table 16 in order to maintain the overall two-sided type I error rate at 0.05.

For non-inferiority at tests 1, 4, 5, 6 and 7, the test will be considered successful if:

- the lower bound of the two-sided 95% CI of the difference between two treatment arms is above a set margin for tests 1, 4, and 5,

- the upper bound of the two-sided 95% CI of the hazard ratio of treatment arms is below a set margin for tests 6 and 7.

For superiority at tests 2, 3, 8 and 9, the test will be considered successful if:

- the upper bound of the two-sided 95% confidence interval of the difference between treatment arms is below 0 for tests 2 and 8,
- the upper bound of the two-sided 95% confidence interval of the hazard ratio of the two treatment arms is below 1 for test 3,
- the upper bound of the two-sided 95% confidence interval of the hazard ratio of treatment arms is below 1 for test 9.

For each success, the test will be performed for the next comparison and so on. In case of test failure, the procedure will be stopped. Only the primary analysis of each secondary endpoint will be used to claim the significance.

Table 16 Key Secondary Endpoints fixed sequence testing procedure

Test	Analysis set	Endpoint	Comparison*
1	PPS	Hb change from BL to the average Hb of weeks 28 to 36 without having received rescue therapy (i.e. RBC transfusion for all subjects, or darbepoetin alfa for roxadustat subjects) within 6 weeks prior to and during this 8-week evaluation period	Non-inferiority of roxadustat versus darbepoetin alfa (the non-inferiority margin for the difference between groups is 0.75 g/dL).
2	FAS	LDL change from BL to the average of weeks 12 to 28	Superiority of roxadustat versus darbepoetin alfa
3	FAS	Time to first IV iron use during weeks 1 to 36	Superiority of roxadustat versus darbepoetin alfa
4	PPS	SF-36 PF sub-score change from BL to the average of weeks 12 to 28	Non-inferiority of roxadustat versus darbepoetin alfa (the non-inferiority margin is fixed as a difference of three points)
5	PPS	SF-36 vitality sub-score change from BL to the average of weeks 12 to 28	Non-inferiority of roxadustat versus darbepoetin alfa (the non-inferiority margin is fixed as a difference of three points)
Table continued on next page			

Test	Analysis set	Endpoint	Comparison*
6	PPS	MAP change from BL to the average MAP of weeks 20 to 28	Non-inferiority of roxadustat versus darbepoetin alfa (the non-inferiority margin for the difference between groups is 1 mmHg)
7	PPS	Time to first occurrence of hypertension during weeks 1 to 36, defined as: - systolic BP ≥170 mmHg AND systolic BP increase from BL ≥20 mmHg, or - diastolic BP ≥110 mmHg AND diastolic BP increase from BL ≥15 mmHg	Non-inferiority of roxadustat versus darbepoetin alfa (the non-inferiority margin is fixed as a hazard ratio of 1.3)
8	FAS	MAP change from BL to the average MAP of weeks 20 to 28	Superiority of roxadustat versus darbepoetin alfa
9	FAS	Time to first occurrence of hypertension during weeks 1 to 36, defined as : - systolic BP ≥170 mmHg AND systolic BP increase from BL ≥20 mmHg, or - diastolic BP ≥110 mmHg AND diastolic BP increase from BL ≥15 mmHg	Superiority of roxadustat versus darbepoetin alfa

*Subjects randomized to roxadustat QW, BIW and TIW under protocol v1 will be pooled together.

Details of the analysis for each of these secondary endpoints are given below.

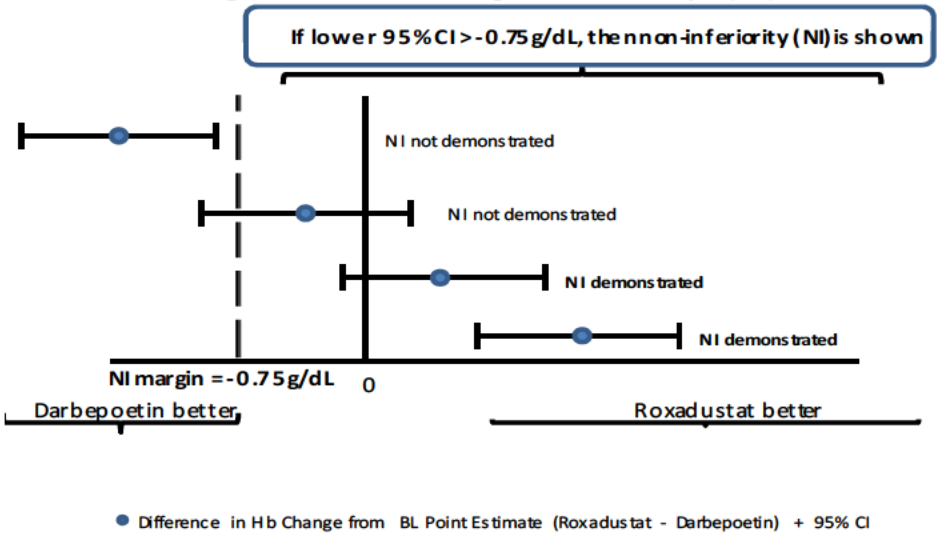
Descriptive statistics and frequency tabulations will be reported using all available data and imputed data, as defined in section 7.11.1.

7.4.2.1 Hb change from baseline to the average Hb in weeks 28-36

The primary analysis will be performed on the PPS [without rescue therapy see section 6.1.2.1]. Hb change from baseline to the average Hb value in weeks 28-36 will be analyzed using a Mixed Model of Repeated Measures (MMRM) with unstructured covariance matrix model (roxadustat versus darbepoetin alfa). The model will contain treatment arm, region, CV History, visits and visit by treatment as categorical variables and baseline Hb, baseline eGFR as continuous variables.

The non-inferiority margin for the difference between groups is 0.75 g/dL. Non Inferiority can be concluded if the lower bound of the two-sided 95% CI of the difference between the two treatments arms is above -0.75 g/dL. See figure 3 below:

Figure 3: Demonstrating Non-inferiority (NI)



MMRM model:

An MMRM model will be run for the purpose of implicit imputation of missing data by using all the available information from the observed data up to Week 36 via the within-patient correlation structure. The analysis will be based on the estimated difference between the two treatment arms overall mean effects throughout the evaluation period (weeks 28 to 36) based on this MMRM model.

The model will contain treatment arm, region, CV History, visits and visit by treatment as categorical variables and baseline Hb, baseline eGFR and baseline Hb by visit as continuous variables. Note that log-transformed eGFR might be used as covariate in the model instead of eGFR if it is demonstrated that eGFR is not normally distributed.

The unstructured covariance pattern model will be applied first. If the algorithm for unstructured covariance pattern does not converge, then heterogeneous Toeplitz structure will be used instead. If this second model also does not converge, then the (homogeneous) Toeplitz structure will be tried. Finally, if none of them converge, first order autoregressive (AR (1)) as a covariance structure will be used to achieve convergence.

A similar model as the general example below will be used:

$$c_{ikjn} = intercept + \beta_n M_{baseline,ikjn} + \tau_i + \alpha_n + (\alpha\tau)_{in} + \gamma_k + \varepsilon_{ikjn}$$

where

- c_{ikjn} is each analysis visit change from baseline of subject j in treatment arm i, and stratum k at time n,
- β_n is the slope of c_{ikjn} at visit n as a function of the baseline Hb,
- $M_{baseline,ikjn}$ is the baseline measurement of subject j in treatment arm i and stratum k at time n,
- τ_i , is the mean effect of treatment arm i,

- α_n is the mean effect at time *n,*
- $(\alpha\tau)_{in}$ is the interaction term between treatment arm *i* and time *n,*
- γ_k is the mean effect of stratum *k,*
- ε_{ikjn} is the residual at time *n* for subject *j* in treatment arm *i* and stratum *k.*

The SAS procedure will be similar to the following:

```
proc mixed;
 class subject_id treatment region cv_history visit;
 model change = treatment region cv_history  baseline_Hb
 baseline_eGFR visit treatment*visit baseline_Hb*visit;
 repeated visit /subject = subject_id type=un;
 lsmeans visit*treatment / cl alpha = 0.05;
 estimate 'Roxadustat at Weeks 28-36'
 int X covariates X / divisor=x;
 estimate 'Roxadustat v.s. Darbepoetin alfa at weeks 28-36'
     treatment 1 -1
     treatment*visit 0 0 0 0 .. 0.3 0.3 0.3 -0.3 -0.3 -0.3/ cl;
 where visit in
(,'Week1','Week2',...,'Week8','Week10',...,'Week36');
run;
```

One analysis Hb value for each visit will be used, as defined in analysis windows in Section 7.11.4

A forest plot will be generated showing strata on the y-axis and differences in estimated mean changes from BL to the average in weeks 28-36 (obtained from MMRM model) and their 95% confidence interval on the x-axis.

In addition, MMRM least square means and 95% confidence intervals will be calculated for each visit for the difference in treatment arms. MMRM least square means and their 95% confidence intervals will be plotted versus time.

Sensitivity analyses

Sensitivity analyses to be performed for this endpoint (Hb change from BL to the average Hb of weeks 28-36) are summarized below and detailed in Table 17.

- The primary analysis (MMRM, without rescue therapy) will be repeated on the FAS (S1);
- Sensitivity analyses will be performed (MMRM, regardless rescue therapy) on the PPS (S2) and on the FAS (MMRM, regardless recue therapy,) (S3);
- An Analysis of Covariance (ANCOVA) model with Last Observation Carried Forward (LOCF) will be fitted including region, history of CV, baseline Hb and baseline eGFR [or log(eGFR)] as continuous covariates, for the PPS (without rescue therapy) (S4) and the FAS (without rescue therapy) (S5);
- Another ANCOVA model with LOCF will be repeated on FAS population (regardless rescue therapy) (S6).

Table 17 Primary and sensitivity analyses for the Hb change from baseline

Code	Set	Sensitivity analysis	Endpoint	Method	Covariates
Primary	PPS	Without rescue therapy	Change to the Average Hb in weeks 28-36	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates.
S1	FAS	Without rescue therapy	Change to the Average Hb in weeks 28-36	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates.
S2	PPS	Regardless rescue therapy	Change to the Average Hb in weeks 28-36	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates.
S3	FAS	Regardless rescue therapy	Change to the Average Hb in weeks 28-36	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates.
S4	PPS	Without rescue therapy	Change to the Average Hb in weeks 28-36	ANCOVA with LOCF	Region, history of CV, baseline Hb and baseline eGFR as continuous covariates.
S5	FAS	Without rescue therapy	Change to the Average Hb in weeks 28-36	ANCOVA with LOCF	Region, history of CV, baseline Hb and baseline eGFR as continuous covariates.

Table continued on next page

Code	Set	Sensitivity analysis	Endpoint	Method	Covariates
S6	FAS	Regardless rescue therapy	Change to the Average Hb in weeks 28-36	ANCOVA with LOCF	Region, history of CV, baseline Hb and baseline eGFR as continuous covariates.

ANCOVA model

The following model will be used:

$$c_{ikj} = intercept + \beta \cdot M_{baseline,ikj} + \tau_i + \gamma_k + \varepsilon_{ikj}$$

where

- c_{ikj} is the change from baseline to the average Hb values in weeks 28-36 of subject j in arm i and stratum k
- β is the slope of c_{ikj} as a linear function of the baseline Hb
- $M_{baseline,ikj}$ is the baseline Hb of subject j in treatment arm i and stratum k
- τ_i is the mean effect of treatment arm i
- γ_k, is the mean effect of stratum k
- ε_{ikj} is the residual for subject j in treatment arm i and stratum k.

The SAS procedure will be similar to the following:

```
proc mixed;
class region history_CV treatment;
model change = treatment baseline_Hb baseline_eGFR region
history_CV;
run;
```

For the sensitivity analyses S4, S5 and S6, an additional covariate "Completer at Week 36" (Yes/No) will be added to the model above to account for LOCF.

For missing data imputation rules using an ANCOVA model with LOCF, refer to section 7.11.1

Descriptive analyses

In addition to the inferential analysis, central laboratory and HemoCue hemoglobin Hb values and their associated change from baseline, will be reported descriptively by visit. For central lab Hb values, the average of weeks 28-36 will also be reported.

The following data will be presented graphically by treatment arm:

- Hb results using mean values (+/- 95% CI) plot
- Hb change from baseline results using mean values (+/- 95% CI) plot.

A plot will be generated showing the change between the central laboratory and HemoCue hemoglobin values by visit (+/- 95% CI) during the Efficacy Emergent Period.

7.4.2.2 Change from baseline in LDL cholesterol to the average value of LDL cholesterol in weeks 12-28

Change from baseline in LDL cholesterol to the average value of LDL cholesterol in weeks 12-28 will be compared by treatment arms using a MMRM model as in Section 7.4.2.1 (with the addition of LDL at baseline as continuous covariate).

The analysis will be similar to the primary analysis provided in section 7.4.2.1, except that it will be a superiority test instead of a non-inferiority test. This superiority test will be considered successful if the upper bound of the two-sided 95% confidence interval of the difference between treatment arms is below 0.

The analysis will be done on the FAS. An additional analysis (not part of the sequence) will be done on the PPS, as detailed in Table 18

Table 18 Primary and sensitivity analysis for the LDL change from BL to the average LDL in weeks 12-28

Code	Set	Endpoint	Method	Covariates
Primary	FAS	Change from baseline to the Average LDL in weeks 12-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL LDL, BL Hb and BL eGFR as continuous covariates.
S1	PPS	Change from baseline to the Average LDL in weeks 12-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL LDL, BL Hb and BL eGFR as continuous covariates.

For missing LDL imputation rules for MMRM refer to Section 7.11.1.

This analysis will be done on all values (regardless the fasting status).

In addition to the inferential analysis, LDL cholesterol and LDL cholesterol change from baseline will be reported descriptively by visit. The average of weeks 12-28 will also be reported.

7.4.2.3 Time to first IV iron use during weeks 1 to 36 and Monthly IV iron use during Day 1 to Week 36

The average monthly IV iron use (mg) during Day 1 to Week 36 will be compared by treatment groups using an analysis of covariance (ANCOVA) model using region, history of CV, baseline Hb and baseline eGFR as covariates. Subjects with no medication record of IV Iron will be assumed that they received no IV Iron. For those subjects, monthly IV iron use will be set to zero mg. Subjects who received Intramuscular Iron are counted under IV Iron in the analyses.

The analysis will be done on the FAS. This superiority test will be considered successful if the upper bound of the two-sided 95% confidence interval of the difference between treatment arms is below 0. An additional analysis will be performed on the PPS.

Time to first use of IV Iron during the first 36 weeks will be analyzed similarly to time to hypertension (see section 7.4.2.7).

The analysis will be done on the FAS.

Superiority will be declared if the upper bound of the two-sided 95% confidence interval of the hazard ratio of the two treatment arms is below 1.

Table 19 Primary and sensitivity for Mean monthly IV iron use (mg) during Day 1 to Week 36 and Time to first IV iron use during weeks 1 to 36

Code	Set	Endpoint	Method	Covariates
Primary	FAS	Mean monthly IV iron use (mg) during Day 1 to Week 36	ANCOVA	Region, History of CV, Baseline Hb and Baseline eGFR
S1	PPS	Mean monthly IV iron use (mg) during Day 1 to Week 36	ANCOVA	Region, History of CV, Baseline Hb and Baseline eGFR
Primary	FAS	time to first IV iron use during weeks 1 to 36	Cox regression + Kaplan Meier	Stratified on Region and History of CV, and adjusted on BL Hb, BL eGFR as continuous covariates
S1	PPS	Time to first IV iron use during weeks 1 to 36	Cox regression + Kaplan Meier	Stratified on Region and History of CV, and adjusted on BL Hb, BL eGFR as continuous covariates

7.4.2.4 Change from baseline in SF-36 PF subscore to the average in weeks 12–28

Change from baseline in PF subscore of SF-36 to the average of weeks 12–28 will be compared by treatment arm, for the PPS and in the subsets of subjects with baseline PF subscore below 35 and equal or above 35. It will be done using a MMRM method adjusting for the region, history of CV, baseline SF-36 PF subscore, baseline Hb and baseline eGFR as covariates. Baseline SF-36 PF subscore, baseline Hb and baseline eGFR will be included as continuous variables. Non-inferiority can be concluded if the lower bound of the two-sided 95% CI of the difference between the two treatment groups is above -3 points (see figure 6 below).

The analysis for this variable will be similar to the one provided in section 7.4.2.1 (all the available information from the observed data up to analysis date of Week 28 will be used).

In addition, this model will be repeated for the FAS.

Figure 6: Demonstrating Non-inferiority (NI)

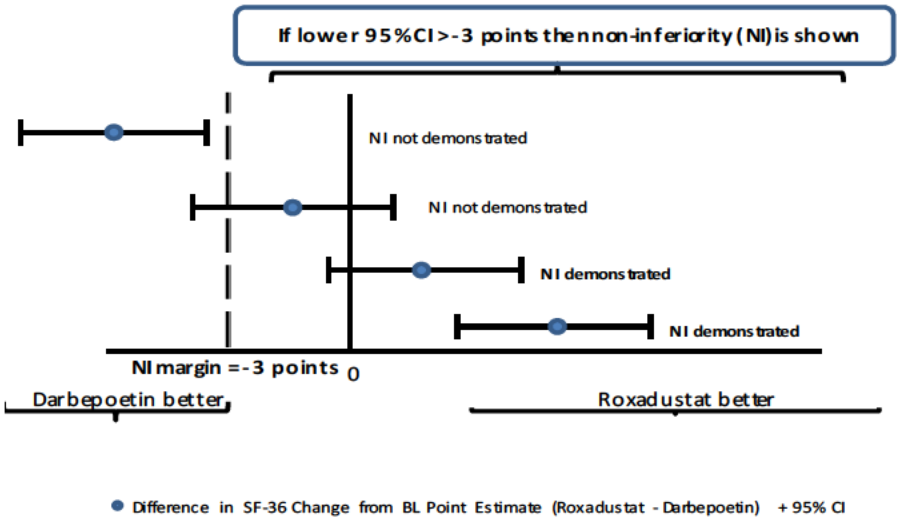


Table 20 Primary and sensitivity analysis for change from BL in SF-36 Physical Functioning (PF) sub-score to the average PF sub-score of weeks 12 to 28

Code	Set	Endpoint	Method	Covariates
Primary	PPS	SF36-PF Change from Baseline at Weeks 12-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb, BL SF-36 PF subscore and BL eGFR as continuous covariates.
S1	FAS	SF36-PF Change from Baseline at Weeks 12-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb, BL SF-36 PF subscore and and BL eGFR as continuous covariates.

In addition to the inferential analysis, SF-36 PF and SF-36 PF change from baseline will be reported descriptively by visit, using all available data. The average of weeks 12-28 will also be reported.

7.4.2.5 Change from baseline in SF-36 VT subscore to the average in weeks 12–28

Change from baseline in VT subscore of SF-36 to the average in weeks 12–28 will be compared by treatment arm, for PPS, and in the subsets of subjects with baseline VT subscore below 50 and equal or above 50. It will be done using a MMRM method adjusting for the region, history of CV, baseline SF-36 VT subscore, baseline Hb and baseline eGFR as covariates. Baseline SF-36 VT subscore, baseline Hb and baseline eGFR will be included as continuous variables. Non inferiority will be concluded as for previously (section 7.4.2.4).

The analysis for this variable will be similar to the one provided in section 7.4.2.1

In addition, this model will be repeated for the FAS.

Table 21 Primary and sensitivity analysis for change from BL in SF-36 VT sub-score to the average in weeks 12 to 28

Code	Set	Endpoint	Method	Covariates
Primary	PPS	SF36-VT Change from Baseline at Weeks 12-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb, BL SF-36 VT subscore and BL eGFR as continuous covariates.
S1	FAS	SF36-VT Change from Baseline at Weeks 12-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb, BL SF-36 VT subscore and BL eGFR as continuous covariates.

In addition to the inferential analysis, SF-36 VT and SF-36 VT change from baseline will be reported descriptively by visit. The average of weeks 12-28 will also be reported.

7.4.2.6 Change from baseline in mean arterial pressure (MAP) to the average MAP value of weeks 20–28

The blood pressure effect is assessed by the change from baseline in MAP.

Change from baseline in MAP to the average in weeks 20–28 will be compared by treatment arm using a MMRM model as in Section 7.4.2.1 (all the available information from the observed data up to Week 28 will be used)

For missing MAP imputation rules, refer to Section 7.11.1

Non-inferiority can be concluded if the upper bound of the two-sided 95% CI of the difference between the two treatment arms is below 1 mm Hg (see Figure 4 below). This will be calculated on the PPS.

As per Table 16, once the non-inferiority is concluded, superiority will be checked for this variable on the FAS. Superiority can be concluded if the upper bound of the two-sided 95% CI of the difference between the two treatment arms is below 0 mm Hg.

Figure 4: Demonstrating non-inferiority and superiority

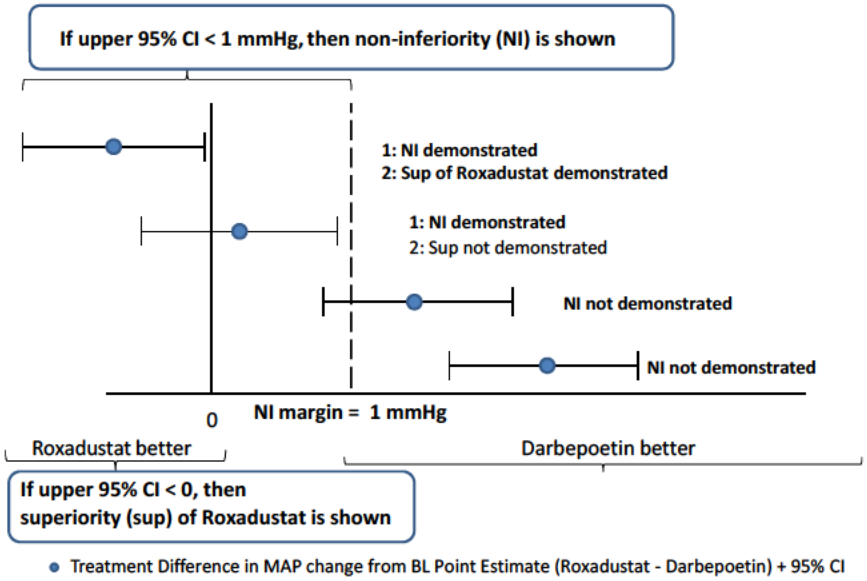


Table 22 Primary and sensitivity analysis for the MAP change from BL to the average MAP in weeks 20-28

Code	Set	Endpoint	Method	Covariates
Primary	PPS	Change from baseline to the Average MAP in weeks 20-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL MAP, BL Hb and BL eGFR as continuous covariates.
Primary	FAS	Change from baseline to the Average MAP in weeks 20-28	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL MAP, BL Hb and BL eGFR as continuous covariates.

In addition to the inferential analysis, MAP and MAP change from baseline will be reported descriptively by visit, using all available data. The average of weeks 20-28 will also be reported.

7.4.2.7 Time to first occurrence of hypertension during weeks 1 to 36

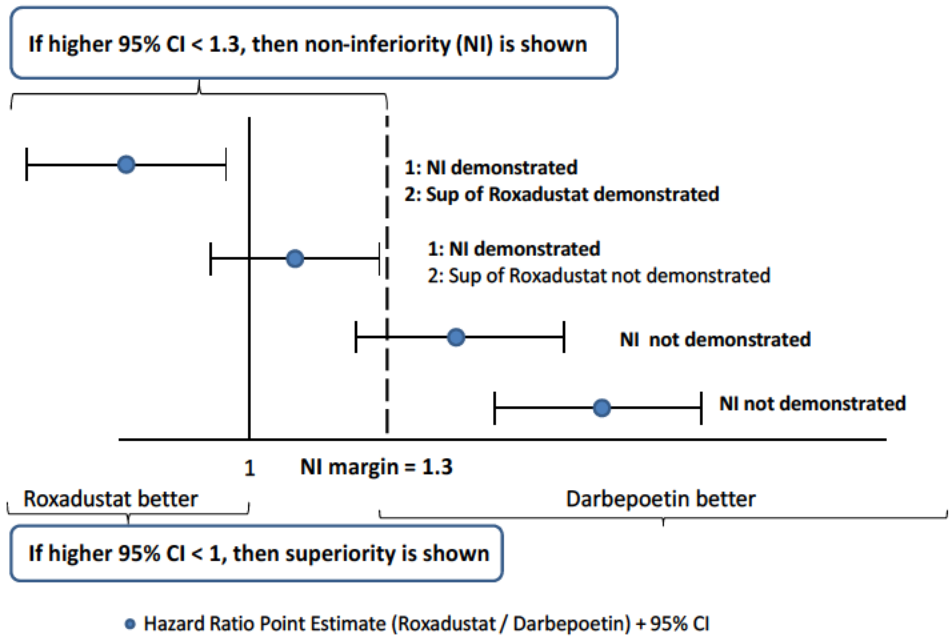
Time to first occurrence of an increase in blood pressure, including time to censoring, defined in Section 6.1.2.7 will be used in a Cox Proportional Hazards regression analysis, to compare treatment arms, stratified for region and history of CV and adjusted for baseline eGFR and baseline Hb (both continuous), and provide hazard ratio and their 95% confidence intervals.

Non-inferiority will be declared if the upper bound of the 2-sided 95% confidence interval, calculated on the PPS, is below 1.3 (see figure 5 below). As per Table 16 once the null

hypothesis is rejected, superiority will be checked for this variable on the FAS. Superiority will be concluded if the upper bound of the two-sided 95% CI of the hazard ratio of the two treatment arms is below 1.

An additional analysis (not part of the sequence) will be done on the All Randomized.

Figure 5: Demonstrating non-inferiority and superiority



The stratified Cox Model can be written:

$$\lambda_j(t;\underline{x}) = \lambda_{0j}(t)\exp(\alpha z)$$

where

j - indicator for stratum
z – treatment indicator – which is either roxadustat or darbepoetin alfa

The SAS procedure for the Cox regression will be similar to the following:

```
proc phreg;
   model time_to_event*cens_var(1) = treatment Hb_Bas eGFR_bas / rl;
   strata CVHist Region;
run;
```

Table 23 Primary and sensitivity analysis for the time to first occurrence of hypertension during weeks 1 to 36

Code	Set	Endpoint	Method	Stratas/Covariates
Primary	PPS	time to first occurrence of hypertension in week 1 to 36	Cox regression + Kaplan Meier	Stratified on Region and History of CV, and adjusted on BL Hb, BL eGFR as continuous covariates
Primary	FAS	time to first occurrence of hypertension during weeks 1 to 36	Cox regression + Kaplan Meier	Stratified on Region, History of CV and adjusted on BL Hb, BL eGFR as continuous covariates

In addition, the cumulative incidence curve of subjects with an increase in blood pressure from baseline during weeks 1 to 36 will be plotted by treatment arm.

The cumulative incidence will be calculated as one minus the Kaplan-Meier estimate of the survival function. Two types of analyses can be performed: modeling the cause specific hazard or modeling the hazard of the sub-distribution. The first approach has been chosen in this SAP, because competing risks are assumed not to exist, since the main interest is to study the treatments effect and this method provides results to be generalized across datasets with different competing risks. One minus the Kaplan-Meier estimate can be interpreted as the probability that an event of interest occurs to a subject by time t, in the absence of any competing risks.

In addition to the cumulative incidence plot, the cumulative incidence at Week 6, Week 12, Week 18, Week 24, Week 30 and Week 36 with the 95% confidence interval will be reported using Greenwood's formula.

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. This plot will be provided as part of the raw SAS outputs.

In addition, the number and percentage of subjects with occurrence of hypertension during weeks 1 to 36 will be reported by treatment arm on the PPS. For subjects who have experienced more than one hypertension, only their first event following study treatment will be used in the analysis.

In addition, the incidence rate (per 100 subject years at risk) will also be calculated as follows:

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

Where total cumulative time at risk is the sum of individual time at risk defined as either time to occurrence of the event or time to censoring for subjects with no event.

7.4.2.8 Additional Analyses of the Key Secondary Endpoints

Each of these key secondary endpoints will also be analyzed by the subgroups of interest defined in section 7.8, using only the primary analysis method, descriptively (no hypothesis testing).

Table 24 Additional Analyses of the Key Secondary Endpoints

Code	Set	Endpoint	Method	Covariates
A1	PPS	Change to the Average Hb in weeks 28-36 (without rescue therapy) by Subgroup	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates.
A2	FAS	Change from baseline to the Average LDL in weeks 12-28 by Subgroup	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL LDL, BL Hb and BL eGFR as continuous covariates.
A3	FAS	Mean monthly IV iron use (mg) during Day 1 to Week 36 by Subgroup	ANCOVA	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates.
A4	PPS	SF36-PF Change from Baseline at Weeks 12-28 by Subgroup	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb, BL SF-36-PF subscore and BL eGFR as continuous covariates.
A5	PPS	SF36-VT Change from Baseline at Week 28 by Subgroup	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL Hb ,BL SF-36-VT subscore and BL eGFR as continuous covariates.
A6	PPS	Change from baseline to the Average MAP in weeks 20-28 by Subgroup	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL MAP, BL Hb and BL eGFR as continuous covariates.
A7	FAS	Change from baseline to the Average MAP in weeks 20-28 by Subgroup	MMRM	Region, History of CV, visits and visits by treatment as categorical variables. BL MAP, BL Hb and BL eGFR as continuous covariates.
A8	PPS	time to first occurrence of hypertension in week 1 to 36 by Subgroup	Cox regression + Kaplan Meier	Stratified on Region and History of CV, and adjusted on BL Hb, BL eGFR as continuous covariates
A9	FAS	time to first occurrence of hypertension in week 1 to 36 by Subgroup	Cox regression + Kaplan Meier	Stratified on Region, History of CV and adjusted on BL Hb, BL eGFR as continuous covariates

* Subgroup: (1) age, (2) sex, (3) region, (4) baseline hemoglobin category, (5) history of CV (6) Baseline eGFR category, (7) Baseline CRP group (8) Iron repletion status.

Forest plots will be produced where all subgroup factors will appear on the y-axis and the appropriate statistic comparing roxadustat versus darbepoetin alfa and the 95% confidence interval will appear on the x-axis.

7.4.3 Analysis of the Additional Secondary Endpoints

All the analyses below will be superiority tests. These inferential analyses will be performed on the FAS and presented treatment effect as Roxadustat versus Darbepoetin alfa using 95% confidence intervals.

Descriptive statistics will be presented by treatment arm.

7.4.3.1 Hb change from BL to the average Hb value of weeks 28 to 52 regardless of the use of rescue therapy

The analysis will be done similarly as in section 7.4.2.1, on the FAS except that it will be for weeks 28-52 and regardless the use of rescue therapy.

7.4.3.2 Time to achieve the first Hb response (without rescue therapy and regardless of rescue therapy) during the Efficacy Emergent Period

Time to achieve the first Hb response, without rescue therapy, will be analyzed using the same methods described in section 7.4.2.7. In addition, the analysis will be repeated regardless of rescue therapy.

Superiority will be declared if the lower bound of the two-sided 95% confidence interval of the hazard ratio is above 1.

Cumulative incidence will be provided at 4, 8, 16, 24, 52, 72 and 96 weeks with their associated 95% confidence interval reported using Greenwood's formula.

7.4.3.3 Hb averaged over weeks 28 to 36, 44 to 52, 72 to 80, 96 to 104 without having received rescue therapy within 6 weeks prior to and during this 8-week period

Averaged Hb over weeks 28-36, 44 to 52, 72 to 80 and 96 to 104 will be described by treatment arm on the FAS.

The SAS procedure will be similar to the one provided in section 7.4.2.1

In addition, the number and proportion of subjects with average Hb over weeks 28-36 and over weeks 44-52, 72-80 and 96-104 and by visit within the <10 g/dL, 10-12 g/dL, ≥10 g/dL and >12 g/dL categories will be shown.

7.4.3.4 Hb value and Hb change from BL Hb to each post-dosing time point (regardless of rescue therapy)

Hb value and change from BL Hb to each post dosing time point will be described by treatment arm on the FAS.

The analysis and SAS procedure will be the same as in section 7.4.2.1, except that it is regardless of rescue therapy.

7.4.3.5 Hb change from BL Hb to the average Hb value of weeks 28 to 36, 44 to 52, 72 to 80, 96 to 104 regardless of the use of rescue therapy

This analysis will be done similarly as in section 7.4.2.1 on the FAS except that it will be for weeks 28-36, 44-52, 72-80 and 96-104 regardless use of rescue therapy.

7.4.3.6 Categorical analysis of Hb values

Proportion of Hb values:

The proportion of Hb values within 10-12 g/dL or ≥10 g/dL is a quantitative variable in the 0 – 100 range for each subject. Descriptive statistics for this variable will be presented by treatment arm. This variable will be presented in weeks 28-36, in weeks 44-52, in weeks 72-80 and in weeks 96 -104, on the FAS.

Percentage of time during the Efficacy Emergent Period:

Descriptive statistics for the percentage of time with Hb values falling in each interval interval (<10.0 g/dL, within 10.0-12.0 g/dL, ≥10 g/dL, > 12.0 g/dL, > 13.0 g/dL and > 14.0 g/dL) between the first and last Hb assessment during the Efficacy Emergent Period will be presented by treatment arm.

7.4.3.7 Occurrence and time to first potential Excessive Hematopoiesis (EH)

The number and percentage of subjects, and the number of events, will be reported based on Hb values from the central lab with Hb increase by >2.0 g/dL between any two visits within 4 weeks of treatment.

Time to first occurrence of potential EH will also be analyzed similarly as in Section 7.4.2.7 (except that all data during the Efficacy Emergent Period will be taken into account for the analysis). The results will be presented by period (First 6 weeks, Week 7 - Week 27, Week 28 - Week 52 and > Week 52)

7.4.3.8 Occurrence (number) of hospitalizations, number of days of hospitalization per PEY and time to first hospitalization

The number and percentage of subjects with hospitalization will be reported. Descriptive statistics and frequency tabulations by treatment arm of the total duration of hospitalization (days), the average duration of each hospitalization (days), the number of hospitalizations, number of days of hospitalization per PEY and reason for hospitalization will also be reported.

Time to first hospitalization will also be analyzed on the FAS similarly as in Section 7.4.2.7 (except that all data during the Efficacy Emergent Period will be taken into account for the analysis).

7.4.3.9 Occurrence and time to first use of RBC transfusions, number of RBC packs per subject, volume of RBC transfused per subject

The number and percentage of subjects with RBC transfusion for any reason will be reported. Descriptive statistics by treatment arm for number of RBC packs and volume of blood

transfused will be reported as well as number of RBC packs and volume per PEY. Time to first use of RBC transfusion will be analyzed during the Efficacy Emergent Period similarly as in Section 7.4.2.7 using the FAS.

7.4.3.10 Occurrence and time to first use of rescue therapy [composite of RBC transfusions [all subjects] and darbepoetin alfa use [roxadustat treated subjects only])

The number and percentage of subjects with rescue therapy during the treatment period will be reported by treatment arm.

Time to first use of rescue therapy will also be analyzed similarly as in Section 7.4.2.7 (except that all data during the Efficacy Emergent Period will be taken into account for the analysis).

The analysis will be done on the FAS.

Superiority will be declared if the upper bound of the two-sided 95% confidence interval of the hazard ratio of the two treatment arms is below 1.

7.4.3.11 Occurrence of iron supplementation. Mean monthly IV iron (mg) per subject during weeks 37-52 and weeks 53-104 (monthly defined as a period of 4 weeks). Time to first IV Iron during the efficacy emergent period

Occurrence of Iron use will be summarized by time period (during the efficacy emergent period, during the first 36 weeks, during week 37 – 52 and week 53- 104) and by type of Iron (IV Iron only, Oral Iron only, IV and Oral iron and No iron).

Mean monthly IV Iron use during weeks 37-52 and weeks 53-104 will be analyzed same way as for the first 36 weeks (see section 7.4.2.3).

Time to first use of IV Iron will be analyzed similarly to use of rescue therapy (see section 7.4.3.10).

The analysis will be done on the FAS.

Superiority will be declared if the upper bound of the two-sided 95% confidence interval of the hazard ratio of the two treatment arms is below 1.

7.4.3.12 Change from BL to each post-dosing study visit in Total cholesterol, LDL/High-density Lipoprotein (HDL) ratio, Non-HDL cholesterol, Apolipoproteins A1 and B, ApoB/ApoA1 ratio

Descriptive statistics (value, change from baseline) by visit, treatment arm will be presented regardless of fasting status. Descriptive statistics will also be reported with fasted values.

7.4.3.13 Occurrence of mean LDL cholesterol <100 mg/dL (2.59 mmol/L), calculated over weeks 12 to 28, and weeks 36 to 52 of treatment

The number and percentage of subjects with mean LDL cholesterol less than 100 mg/dL (2.59 mmol/L)(regardless fasting status) on average in weeks 12-28 and 36-52 will be reported by treatment arm and baseline category (< 2.59 mmol/L , ≥ 2.59 mmol/L and

overall). Descriptive statistics and frequency tabulations will also be reported regardless of fasting status.

7.4.3.14 Occurrence of achieved antihypertensive treatment goal (SBP <130 mmHg systolic and DBP<80 mmHg) based on the mean SBP and mean DBP calculated over weeks 12 to 28 and 36 to 52

The number and percentage of subjects with achieved antihypertensive treatment goal over an evaluation period defined as the average of available values in weeks 12-28 and 36-52 will be reported by treatment arm.

7.4.3.15 Health related Quality of Life Questionnaires Change from BL to the average value of weeks 12 to 28 and 36-52

The following questionnaires will be analyzed:

- SF-36 (Physical Component Score (PCS));
- FACT-An (Anemia Subscale ("Additional Concerns"), Total FACT-An Score);
- EQ-5D 5L (VAS Score);
- Work Productivity and Activity Impairment Questionnaire: Anemic Symptoms (WPAI:ANS).

SF-36 and FACT-An

Descriptive statistics and frequency tabulations (value, change from baseline) will be presented for SF-36 (subscale and component scores) and FACT-An (total and subscale scores) by visit and treatment arm. Mean values will also be plotted over time and by treatment arm.

In addition, for the average value in weeks 12-28 and 36-52, an inferential analysis, similar to the one defined in section 7.4.2.1 will be performed for the following endpoints:

1. Physical Component Scores of SF-36 (SF-36 PCS)
2. Anemia Subscale ("Additional Concerns") of FACT-An Scores
3. Total FACT-An Scores

For SF-36, at each visit, the frequency and proportion of subjects with an increase from baseline <3/≥ 3 points and <5/≥5 points will be reported for Physical Functioning, Vitality and Physical Component scores.

For FACT-An, at each visit, the frequency and proportion of subjects with a increase from baseline <3/≥3 points in the FACT-An Anemia and a change from baseline <7/≥7 points in the Total FACT-An score will be reported.

EQ-5D 5L

For the EQ-5D 5L VAS score, change from baseline to each visit and to the average of weeks 12-28 and 36-52 will be described by treatment arm. Mean values will be plotted over time and by treatment arm. For the VAS score average value in weeks 12-28, an inferential analysis, similar to the one defined in section 7.4.2.1 will be performed.

For the 5 EQ-5D 5L qualitative domains, the number and percentage of subjects in each response level value will be reported by visit and treatment arm.

Work Productivity and Activity Impairment (WPAI)

The number and proportion of employed subjects will be reported by visit and treatment arm.

WPAI calculated variables will be summarized by descriptive statistics and frequency tabulations, by visit and treatment arm.

Furthermore, change from baseline to each visit and to the average value in weeks 12-28 and 36-52 will be described by treatment arm.

7.4.3.16 Patients' Global Impression of Change (PGIC)

Patients' Global Impression of Change will be summarized descriptively by visit and treatment arm.

7.4.3.17 Iron, HbA1c, and CKD progression parameters

Changes from baseline to each study visit will be calculated for these parameters:

- Serum ferritin
- TSAT
- Serum Iron
- HbA1c level
- Fasting blood glucose
- eGFR (including slope over time)
- Serum creatinine (log transformed)
- Albumin/creatinine ratio in urine (log transformed)

Descriptive statistics and frequency tabulations will be presented for these parameters and for the change from baseline by visit and treatment arm. Mean values will also be plotted versus visit by treatment arm.

For HbA1c and Fasting blood glucose, summaries will be presented by Baseline Diabetes Mellitus status (regardless of type).

Geometric Mean (GM) and coefficient of variation (CV) will be displayed for serum creatinine, eGFR and albumin/creatinine ration in urine. In addition, GM Ratio from baseline and 95% CI will be presented for these parameters by transforming the change from baseline of the log-transformed back to the original scale. This will be presented by treatment arm.

CKD Progression parameters:

- ***Annualized eGFR slope over time:***

The annualized eGFR (expressed in ml/min per 1.73 m² per year) will be determined using the SAS code below:

```
proc mixed data=egfr2;
class usubjid strata_except_egfr_hb TRTPN TRTP;
```

```
model aval= strata_except_egfr_hb HGBBL atptn
HGBBL*atptn  egfrbl*atptn atptn*TRTPN*TRTP
atptn*strata_except_egfr_Hb     / solution ddfm=kr
outpredm=pred cl;
random intercept atptn / subject=usubjid type=un ;
estimate 'Annnualized slope Roxadustat'
    intercept 0
    strata_except_egfr_hb 0 0 0 0
     HGBBL 0
     atptn 1
    HGBBL*atptn &meanHbBase
    egfrbl*atptn &meanegfrbl
    atptn*TRTPN*TRTP 1 0
    atptn*strata_except_egfr_Hb &prop_strata
  /cl;
estimate 'Annnualized slope Darbepoetin'
    intercept 0
    strata_except_egfr_hb 0 0 0 0
     HGBBL 0
     atptn 1
    HGBBL*atptn &meanHbBase
    egfrbl*atptn &mean egfrbl
     atptn*TRTPN*TRTP 0 1
     atptn*strata_except_egfr_Hb &prop_strata
  /cl;
ods output lsmeans=lsmeans_un  estimates=est_un;
ods output FitStatistics=FitStatistics SolutionF=SolutionF;
run;
```

Annualized eGFR slope over time is estimated by a random slopes and intercepts model using all available eGFR values (one baseline and all post-treatment values up to End of Treatment Period or start of dialysis) adjusted on Baseline Hb, Region, CV history at Baseline and the interaction terms (Baseline eGFR by timepoint and Baseline Hb by timepoint). All assessments collected after initiation of acute or chronic dialysis will be excluded from the analysis.

All time to event endpoints as defined in Section 6.1.3.17 will be analyzed using the same methods described in Section 7.4.2.7 on the FAS.

The number and percentage of subjects who have initiated dialysis or kidney transplant at any moment during the safety emergent period will be reported. A listing of dialysis data will also be provided.

7.4.4 Analysis of Exploratory Variables: hs-CRP (High Sensitivity C-Reactive Protein) and sTFR (Soluble Transferrin Receptor)

For each visit, descriptive statistics with the absolute values and change from baseline for hs-CRP and sTFR will be displayed.

7.5 Analysis of Safety

Safety analyses will be performed using the Safety Analysis Set (SAF).

Missing dates' imputation rules for AE onset and stop date are detailed in section 7.11.2

For each safety parameter, the last non-missing assessment prior to the first dose of study drug will be used as the baseline for all analyses, unless specified otherwise.

7.5.1 Adverse Events

All adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.0. They will be summarized by System Organ Class (SOC) and Preferred Term (PT).

7.5.1.1 Overview

An overview table will include the following details, by treatment arm:

- Number and percentage of subjects with TEAEs,
- Number and percentage of subjects with drug related TEAEs,
- Number and percentage of subjects with serious TEAEs,
- Number and percentage of subjects with serious drug related TEAEs,
- Number and percentage of subjects with TEAEs leading to permanent discontinuation of study drug,
- Number and percentage of subjects with drug related TEAEs leading to permanent discontinuation of study drug,
- Number and percentage of subjects with drug related TEAEs leading to death
- Number and percentage of subjects with NCI CTC Grade 3 or higher TEAEs,
- Number and percentage of subjects with TEAEs leading to death and
- Number of deaths occurring during the Safety Emergent Period and overall.

7.5.1.2 Proportion of subjects with TEAEs by SOC/PT

The number and percentage of subjects with TEAEs, as classified by SOC and PT will be summarized by treatment arm. Summaries will be provided for:

- TEAEs
- TEAEs (PT only)
- Drug related TEAEs,
- Serious TEAEs,
- Drug related serious TEAEs,
- TEAEs with NCI CTC Grade 3 or higher
- TEAEs by relationship
- TEAEs by severity
- Drug related TEAEs by severity
- TEAEs leading to permanent discontinuation of study drug,
- Drug related TEAEs leading to permanent discontinuation of study drug,
- TEAEs leading to death,
- Drug-related TEAEs leading to death,
- TEAEs excluding serious adverse events, that equal to or exceed a threshold of 5.0% in any treatment arm,

- Common TEAEs that equal to or exceed a threshold of 5.0% in any treatment arm
- TEAEs by relationship to the study drug,
- AEs occurring during the post study follow up period (i.e after Analysis date of Last Dose + 28 days) and
- Serious AEs occurring during the post study follow up period

For the summaries by severity or relationship to the study drug in the subject count, if a subject has multiple TEAEs with the same SOC or PT, but with different severity or relationship, then the subject will be counted only once with the worst severity and highest degree of relationship, however, if any of the severity or relationship values are missing then the subject will be counted only once with missing severity or relationship. In the adverse event count, the adverse events will be presented in each category they were classified to. Drug related TEAEs will be presented in a similar way by severity only.

Maximum severity or relationship will be defined as the worst severity and highest degree of relationship on-treatment.

No queries will be done for SMQs.

7.5.1.3 Event-rates per 100 patient-years

The number of events and event rate (per 100 patient years) during the Safety Emergent Period with TEAEs, as classified by SOC and PT will be calculated by treatment arm. Summaries will be provided for:

- TEAEs
- TEAE NCI CTC Grades 3 or higher
- Serious TEAEs,
- TEAEs leading to death.

7.5.1.4 Incidence rates and cumulative incidence

Since the percentage of adverse events might be different between arms due to the difference in the discontinuation rates, in addition to the frequency tables above, the incidence rate (per 100 subject years at risk) and the cumulative incidence at 6 months, 12, 18 months & 24 months with the 95% confidence interval, using Greenwood's formula, will be reported by treatment arm.

It will be done for each of the following event types:

- Serious TEAEs,
- Deaths occurring during the Safety Emergent Period,
- Any death (including post-study follow up period),
- Related serious TEAEs,
- AEs leading to discontinuation of study drug, and
- TEAE NCI CTC Grades 3 or higher.

Hazard ratios of each of the TEAE categories of special interest above will be computed by using Cox Proportional Hazards regression stratified on region and history of cardiovascular,

cerebrovascular or thromboembolic diseases and adjusted on baseline Hb and baseline eGFR as continuous covariates. Hazard ratios (roxadustat as relative to darbepoetin alfa) and their 95% CI will be calculated for each TEAE category.

A dot-and-forest plot will be produced showing each of the above TEAE categories in the y-axis. The incidence rates by treatment arm, the stratified Cox hazard ratio (roxadustat as relative to darbepoetin alfa) and its 95% CI will be shown on the x-axis.

In addition, cumulative incidence plots for subjects experiencing each of the TEAE categories above will be produced by treatment arm.

Incidence rate (per 100 subject years at risk) in each treatment arm by SOC and PT and the cumulative incidences by SOC will also be produced for the most common TEAEs (≥ 5% in total SAF).

In addition, cumulative incidence plots of subjects experiencing at least one most common TEAEs in a SOC will be produced. A dot-and-forest plot will be produced showing each SOC in the y-axis and the incidence rates, the Cox hazard ratio and its 95% CI in the x-axis. The SOC will be sorted by the hazard ratio.

7.5.1.5 Sensitivity/Subgroup analyses

- *Subgroups of interest:*

The number and percentage of subjects reporting TEAEs in each treatment arm will be tabulated by SOC and PT for the subgroups of interest defined in Section 7.8

- *Baseline eGFR category (< 15 vs ≥ 15) only:*

The following summaries will be provided by comparing Baseline eGFR categories (< 15 vs ≥ 15):

- Event rate per 100 patient-years for TEAEs by SOC and PT
- Incidence rate and cumulative incidence for serious TEAEs
- Incidence rate and cumulative incidence for deaths during the safety emergent period
- Incidence rate and cumulative incidence for drug-related serious TEAEs
- Incidence rate and cumulative incidence for AEs leading to discontinuation of study drug
- Incidence rate and cumulative incidence for TEAE NCI CTC Grades 3 or higher

- *Incidence rate and cumulative incidence censoring events occurring on or after the start of chronic dialysis:*

The same analysis in incidence as described in section 7.5.1.4 will be performed (i.e incidence rate, hazard ratios, dot-and-forest plots and cumulative incidence plots) by censoring all data occurring at initiation of chronic dialysis (see section 6.2.1.3 for more details regarding the definition of time to event and time to censoring) :

It will be done for:

- Serious TEAEs
- Serious TEAEs by baseline eGFR category (< 15 vs ≥ 15)

- Deaths occurring during the Safety Emergent Period,
- Deaths occurring during the Safety Emergent Period by eGFR category (< 15 vs ≥ 15)
- Drug-related serious TEAEs,
- Drug-related serious TEAEs by eGFR category (< 15 vs ≥ 15)
- AEs leading to discontinuation of study drug,
- AEs leading to discontinuation of study drug by eGFR category (< 15 vs ≥ 15)
- TEAE NCI CTC Grades 3 or higher
- TEAE NCI CTC Grades 3 or higher by eGFR category (< 15 vs ≥ 15)
- Incidence rate for TEAE by SOC and most common PTs
- Cumulative incidence by SOC for most common PTs

7.5.1.6 AEs up to 7 days

Additional summaries or analyses with a subset of the treatment–emergent events irestricted up to Analysis Date of Last Dose adjusted for dosing frequency + 7 days:

- Overview summary table of subjects with AEs
- Number and percentage of subjects with AEs, as classified by SOC and PT
- Number and percentage of subjects with AEs leading to death, as classified by PT
- Incidence rate for serious AEs (i.e incidence rate, hazard ratios, and cumulative incidence plots as described in section 7.5.1.4 will be provided)
- Incidence rate for AEs NCI CTC Grades 3 or higher (i.e incidence rate, hazard ratios, and cumulative incidence plots as described in section 7.5.1.4 will be provided)

7.5.1.7 Adjudicated events

The pre-specified adjudicated cardiovascular and thrombo-embolic events will be descriptively presented by treatment group. Time to the different type of events will be presented using Kaplan-Meier estimates. Comparison between roxadustat and darbepoetin will be presented using hazard ratios and 95% confidence intervals from the Cox regression model used for all other time to event endpoints. If appropriate, censoring at start of dialysis can be applied.

All data will also be listed.

7.5.2 Clinical Laboratory Evaluation

Descriptive statistics for laboratory values (in SI units) and changes from baseline at each assessment time point and for the maximum and minimum on-treatment (i.e during Safety Emergent Period) value will be presented by treatment arm for the quantitative laboratory parameters.

Maximum and minimum on-treatment values will be determined using all the original values and not the derived windows.

Shift tables and number and percentage of subjects with shift to low and shift to high will be reported by treatment arm for the quantitative laboratory parameters.

Box plots of quantitative laboratory values (in SI units) versus visit will be produced by treatment arm (two arms in one page).

A plot for each parameter of mean (+/- 95% CI) versus visit will be produced by treatment arm (two arms in one page).

Summary by visit for qualitative laboratory parameters will be provided by treatment arm.

All clinical laboratory data will also be listed.

Potentially clinically significant (PCS) laboratory abnormalities

For each potentially clinically significant (PCS) criterion defined in Table 9, the percentage of subjects with abnormalities by visit and at any moment during the Safety Emergent Period and who did not meet the criteria at baseline will be reported by treatment arm.

For the selected PCS criteria (Hb <6 g/dL, Hb >14 g/dL, ALT > 3X ULN, AST > 3X ULN, Total Bilirubin > 1.5X ULN), incidence rate (per 100 subject years at risk) and the cumulative incidence at 6 months, 12, 18 months & 24 months with the 95% confidence interval using Greenwood's formula will also be reported, using only subjects who did not meet the criteria at baseline, by treatment arm. Risk of PCS abnormalities will be compared using the same Cox model as used in section 7.5.1. Hazard ratio and its 95% will be calculated for the frequency of roxadustat as relative to darbepoetin alfa. A dot-and-forest plot will be produced showing the PCS abnormalities above on the y-axis and the incidence rates, the Cox hazard ratio and their 95% CI on the x-axis.

In addition, cumulative incidence plots for subjects experiencing each PCS abnormality will be produced by treatment arm (two arms in one page).

7.5.2.1 Liver function tests

Descriptive summary of PCS values in Liver Enzymes and Total Bilirubin will be provided as per Astellas standard TLB_005 and FFG_008.

In addition, a matrix scatter plot of Liver Enzymes and Bilirubin (as in Astellas standard FFG_009) will be plotted showing the maximum ALT, AST, ALP and total bilirubin during the Safety Emergent Period crossed against each other. Different dots will be used for roxadustat and darbepoetin alfa.

Individual displays of Liver Enzymes and Bilirubin parameters, listed in 6.2.3.1, will be reported for all subjects with either ALT or AST > 3 x ULN or total bilirubin > 2 x ULN during Safety Emergent Period (as in Astellas standard FFG_011) .

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

7.5.3 Vital Signs

Descriptive and changes from baseline for vital signs (systolic blood pressure, diastolic blood pressure, respiratory rate, weight and pulse) at each assessment time point during the study

and for the maximum and minimum on-treatment (i.e during Safety Emergent Period) value will be presented by treatment arm.

Maximum on-treatment value will be determined using all the original values and not the derived windows.

A plot for each parameter of mean (+/- 95% CI) versus visit will be produced by treatment arm (two arms in one page).

Potentially Clinically Significant (PCS) Vital signs Criteria (10 Combined) will be analyzed in the same way as explained in section 7.5.2 for PCS laboratory abnormalities.

All vital signs data will also be listed.

7.5.4 Electrocardiograms (ECGs)

Descriptive and changes from baseline for ECG parameters (Pulse, PR interval, RR interval, QRS interval, QT interval, and QTc interval) at each assessment time point and for the maximum and minimum on-treatment (i.e during Safety Emergent Period) value will be presented treatment arm.

Maximum on-treatment value will be determined using all the original values and not the derived windows.

The number and percentage of subjects with post-baseline PCS values (see Table 10) will be tabulated by treatment arm. The percentages are to be calculated relative to the number of subjects with available baseline and at least one post-baseline assessment. The numerator will be total number of subjects with at least one post-baseline PCS ECG value. Shift tables may be presented.

The following PCS QTc Criteria (both QTcB and QTcF):

- QTc > 500 msec
- Change from baseline in QTc > 60 msec

will be analyzed separately in the same way as explained in Section 7.5.2 for PCS laboratory abnormalities.

A plot for each parameter of mean (+/- 95% CI) versus visit will be produced by treatment arm (two arms in one page).

In addition, ECG parameters will be reported according to Astellas standards TEG_003 and TEG_004.

All ECG data will also be listed.

7.5.5 Pregnancies

A listing of pregnancy test results will be provided.

7.6 Analysis of PK

The statistical methods for PK data will be described in a separate analysis plan. Results of the population PK analysis will not be reported in the Clinical Study Report but in a separate population PK report.

Plasma concentration data of roxadustat will be listed.

7.7 Analysis of PD

Not Applicable.

7.8 Subgroups of Interest

Selected efficacy and safety endpoints will be summarized for the subgroups defined on the basis of the categorized variables listed below in Table 25

Table 25 Subgroups of interest

Grouping variables	Subgroups
Age group	< 65 years 65 - 74 years ≥ 75 years
Sex	Female Male
Region	Western Europe and Israel Central and Eastern Europe
Baseline Hb	≤8 g/dL >8 g/dL
History of cardiovascular, cerebrovascular or thromboembolic diseases	Yes No
Baseline eGFR	<30 mL/min/1.73 m^2 ≥ 30 mL/min/1.73m^2
Baseline eGFR	<15 mL/min/1.73 m^2 ≥ 15 mL/min/1.73m^2
Baseline CRP group	CRP ≤ULN vs. CRP > ULN
Baseline Iron repletion status	Baseline TSAT ≥20% and Baseline ferritin ≥100 ng/mL Yes vs. No

For a group variable, subgroup analysis will not be done in case the size of a at least one of the subgroup is too small (i.e. less than 5% of the patient population).

7.9 Other Analyses

No other analyses are planned.

7.10 Interim Analysis (and Early Discontinuation of the Clinical Study)

An interim analysis will be performed when all subjects have completed 36 weeks of treatment. Since all primary and key secondary analyses have been defined over this period of time, no multiplicity adjustment is required as the information fraction at the interim analysis is 100%. Results of the endpoints measured over a period longer than 36 weeks will

be mainly described to provide full disclosure, but no interferential analyses will be done. For safety, the interim analysis will include all available data at the time of the last subject reaching Week 36. Adjudicated data will not be part of the interim analysis. Once the study is completed, efficacy and safety will be analyzed again and reported including the complete study treatment of 104 weeks. More details about the data cut-off process can be found in the Data Cut-Off Specifications document.

Safety data and dosing decisions will be monitored on an ongoing basis. Ongoing review of safety data will be completed by an independent Data and Safety Monitoring Board (DSMB).

7.11 Handling of Missing Data, Outliers, Visit Windows, and Other Information

7.11.1 Missing Data

For relevant analyses without rescue therapy, for subjects who used rescue therapy, the reported Hb values after the initiation of rescue therapy will be set to missing (instead of the reported values) for 6 weeks from the start date of rescue therapy (or the end in case the duration of rescue therapy > 1 day).

The following imputations will be performed for continuous endpoints, unless specified otherwise:

- An MMRM model will be run for the purpose of implicit imputation of missing data by using all the available information from the observed data via the within-patient correlation structure for continuous endpoints with inferential analysis.

For all secondary variables analyzed using MMRM, an additional sensitivity analysis will be performed using an ANCOVA model with LOCF method using BL value and the stratification factors as covariates:

- Hb (continuous): Available data in time period under evaluation will be used. If no Hb is available during the time period under evaluation, the average of the previous four post-baseline Hb values will be carried forward. For analyses without rescue therapy, for subjects who used rescue therapy, the last Hb value prior to the initiation of rescue therapy will be carried forward for imputation for 6 weeks from the end date of rescue therapy (in place of the collected values). If at least one non-missing Hb assessment is available during the period of evaluation, then no imputation will be performed. When the start date is different than the end date of rescue therapy, the last observation will be carried forward during the period as well.

No imputation will be done for endpoints with no inferential analysis.

7.11.2 Missing Dates

As a general rule, the worst case scenario imputation rule is usually used. A start date is generally imputed to the first possible day, unless the available information in the partly missing date is equal to the one in the reference date. In this case, the substituted date is set to the reference date. An end date is generally imputed to the last possible day.

Completely missing dates will not be imputed.

Diagnosis of anemia, CKD and Targeted Medical History

The following rules will be applied to impute partially missing dates of diagnosis of Anemia, CKD and targeted medical history, as defined in Table 26 below.

Table 26 Definitions of the Analysis Date of Diagnosis of Anemia, CKD and Targeted Medical History

Reported Date (from the eCRF)	Analysis Date (Derived)
--/MM/YYYY	01/MM/YYYY
--/--/YYYY	01/01/YYYY
DD/--/----, or --/MM/----, or --/--/----	No imputation

Previous or Concomitant medication

For previous or concomitant medications, including rescue medications as well as for chronic dialysis, partially missing start dates and/or stop dates will be imputed as defined in Table 27 and Table 28 below:

Table 27 Definitions of the Previous or Concomitant Medication Analysis Start Date

Reported Date (from the eCRF)	Analysis Date (Derived)
--/MM/YYYY	01/MM/YYYY
--/--/YYYY	01/01/YYYY
DD/--/----, or --/MM/----, or --/--/----	No imputation

If the imputed start date is after the stop date, then the imputed start date will be one day prior to the stop date.

Table 28 Definitions of the Previous or Concomitant Medication Analysis Stop Date

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

Adverse Event

For adverse events, partially missing start dates and/or stop dates will be imputed as defined in Table 29 and Table 30 below:

Table 29 Definitions of the Analysis Adverse Event Onset Date

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

* If the month and year is the same as the month and year of first drug intake, use date of the first drug intake.

Table 30 Definitions of the Analysis Adverse Event End Date

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

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

7.11.3 Outliers

As a general rule all values, including outliers will be analyzed.

7.11.4 Visits Windows

The study protocol gives the overall study schedule and the permissible intervals for visits expressed as the number of days relative to the first study medication date (Day 1).

For all study assessments reported by visit, the value which assessment day is the latest collected within the corresponding analysis visit window will be used. If more than one value is collected that day, then the latest value will be used in the analysis.

Analysis Visit windows, as depicted in Table 31 below, will be used for the following study assessments reported by visit:

- Central laboratory parameters (except Lipid Panel),
- Vital Signs,

- ECG parameters.
- Exposure listings

Table 31 Analysis Visit Windows

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Screening		Day -42 to Day -1	Screening
Day 1	Day 1	Day 1	Baseline
Week 1	Day 7 * (Week #) + 1	Day 2 to Day 11	Week 1
Week 2	Day 7 * (Week #) + 1	Day 12 to Day 21	Week 2
Week 4 – 22	Day 7 * (Week #) + 1	[Target Day – 7, Target Day +6]	Week 4 - 22
Week 24	Day 7 * (Week #) + 1	[Target Day – 7, Target Day + 13]	Week 24
Week 28 – 100	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 13]	Week 28 - 100
Week 104/EOT (for completers)	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 28]	Week 104
EOT Visit (for premature discontinuations) and unscheduled (on-treatment)	NA	NA	Analysis Visit corresponding to the actual visit window
EOT + 2 Weeks Visit	14 Days after EOT visit day	NA	EOT + 2 weeks
EOS Visit	28 Days after EOT visit day	NA	EOS
Unscheduled	NA	[Day of EOT + 1, Day of EOT + 20]	EOT + 2 weeks
Unscheduled	NA	[Day of EOT + 21, Day of EOT +31]	EOS

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

Analysis Visit windows, as depicted in Table 32 below, will be used for the quality of life efficacy study assessments:

Table 32 Analysis Visit Windows for QoL

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Day 1	Day 1	Day 1	Baseline
Week 8	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 13]	Week 8
Week 12	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 27]	Week 12
Week 28	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 27]	Week 28

Table continued on next page

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Week 36	Day 7 * (Week #) + 1	[Target Day – 28, Target Day +27]	Week 36
Week 52	Day 7 * (Week #) + 1	[Target Day – 56, Target Day + 83]	Week 52
Week 76	Day 7 * (Week #) + 1	[Target Day – 84, Target Day + 97]	Week 76
Week 104/EOT (for completers)	Day 7 * (Week #) + 1	[Target Day – 98, Target Day + 28]	Week 104/EOT
EOT Visit (for premature discontinuations)	NA	NA	Analysis Visit corresponding to the actual visit window
Unscheduled	NA	NA	Analysis Visit corresponding to the actual visit window
Any	NA	All the assessments not falling within any of the analysis visit windows defined above	Not Windowed

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

Analysis Visit windows, as depicted in Table 33 below, will be used for the Lipid Panel, including LDL cholesterol efficacy study assessment:

Table 33 Analysis Visit Windows for Lipid Panel

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Day 1	Day 1	Day 1	Baseline
Week 4	Day 7 * (Week #) + 1	Day 2 to Day 42	Week 4
Week 8	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 13]	Week 8
Week 12	Day 7 * (Week #) + 1	[Target Day – 14, Target Day + 27]	Week 12
Week 20	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 27]	Week 20
Week 28	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 27]	Week 28
Week 36	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 27]	Week 36
Week 44	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 27]	Week 44
Week 52	Day 7 * (Week #) + 1	[Target Day – 28, Target Day + 55]	Week 52
Week 68	Day 7 * (Week #) + 1	[Target Day – 56, Target Day + 55]	Week 68

Table continued on next page

CRF Visit	Target Day[a]	Analysis Visit Windows Actual Assessment Day	Analysis Visit
Week 84	Day 7 * (Week #) + 1	[Target Day – 56, Target Day + 69]	Week 84
Week 104/EOT (for completers)	Day 7 * (Week #) + 1	[Target Day – 70, Target Day + 28]	Week 104
EOT Visit (for premature discontinuations)	NA	NA	Analysis Visit corresponding to the actual visit window
EOS Visit	NA	Last assessment between Day 2 and study termination day	EOS
Unscheduled	NA	NA	Analysis Visit corresponding to the actual visit window

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

For the MMRM analyses, which requires one value per visit, one analysis Hb value for each planned visit will be used.

For the ANCOVA analyses, which use the average, all available values in the analysis windows will be used for the calculation.

7.11.5 End of Safety Emergent Period

The end of Safety Emergent Period will be defined as Minimum [(Analysis Date of Last Dose + x +28 days), EOS visit date, date of death] with x the number of days calculated from the last frequency corresponding to the last non missing dosing record collected by the site (i.e dose holds will therefore not be considered) .

X will be derived as follows (standardized on 3 times per week, the standard frequency for roxadustat) :

if frequency = “3 TIMES PER WEEK” then x=0
if frequency = “2 TIMES PER WEEK” then x=1
if frequency = “1 TIME PER WEEK” then x=5
if frequency = “EVERY 2 WEEKS” then x=12
if frequency = “QM” then x =26

Additional frequencies might be recorded. For those cases, X will be derived using the same rule as above.

With Analysis Date of Last Dose defined in section 6.5.4

7.11.6 End of Efficacy Emergent Period

For all subjects, the end of Efficacy Emergent Period will be defined as EOT visit. In case a subject died during the treatment period, the end of Efficacy Emergent Period will be set to the last non-missing Hb assessment.

8 DOCUMENT REVISION HISTORY

Version	Date	Changes	Comment/rationale for change
1.0	26 May 2014	NA	Document finalized
2.0	22 August 2016	1) Reduction of the number of sensitivity analyses for the secondary endpoints. 2) New ordering for the key secondary endpoints. 3) Implementation of the Time to event approach for the PPS and adjustment of the relevant sections. 4) An additional analysis set has been defined 'All Randomized' as per Astellas standard. 5) Use of the Safety Emergent Period (i.e last dose + 28 days) as evaluation period by default for the safety endpoints. 6) Use of the Efficacy Emergent Period (i.e last dose + 7 days) as evaluation period by default for most of the efficacy time to event endpoints. 7) Use of "Time to censoring" instead of"time at risk" for patient with no event for more clarity. 8) Clarification of time to censoring for events evaluated during the Safety Emergent Period/Efficacy Emergent Period. 9) Time to PCS ECG focused to QTc instead of all ECG assessments. 10) Update regarding derivation of MAP values: average of the 3 measurements instead of 2 (now in line with FG SAP). 11) Implementation of additional secondary endpoint: Hb change	1) Due to harmonization with Fibrogen 060 study SAP and to limit the additional analyses. 2) Due to the importance of the QoL endpoints and the harmonization with Fibrogen 060 study SAP. 3) PPS definition has been revised in order to limit the exclusion of data by using a time to event approach rather than creating one PPS set for each period of interest. 4) This additional analysis set has been added in protocol amendment 1.0. 5) Use of the Safety Emergent Period (i.e last dose + 28 days) as evaluation period by default for the safety endpoints. 6) Decision agreed with the study team to use a consistent approach for time to event efficacy endpoints by extending the evaluated period up to 7 days after last dose. 7) Clarification of the wording. Time to censoring is more appropriate. 8) Based on the new definitions of Efficacy/Safety Emergent period, time to censoring for all time to events have been updated accordingly. 9) Time to event for PCS ECG other than QTcF was considered not of interest. 10) Using the average of all 3 measurements is more efficient and is consistent with Fibrogen 060 study SAP. 11) This additional endpoint has been added in protocol

		from BL to the average Hb value of weeks 28-52 regardless of the use of rescue therapy for alignment with global program.	amendment 1.0.
		12) Analysis for key secondary endpoints has been updated: Mixed Model of Repeated Measures (MMRM) replacing Analysis of Covariance (ANCOVA) with Last Observation Carried Forward (LOCF) and Removal of Dunnett's multiple comparison procedure.	12) This has been added in protocol amendment 1.0.
		13) Additional secondary endpoint has been deleted: Iron use - Monthly oral iron (mg) use per subject during weeks 1 to 36, and weeks 36 to 52 (monthly defined as a period of 4 weeks).	13) This additional endpoint has been deleted in protocol amendment 2.0.
		14) Implementation of time to first hospitalization, time to first occurrence of serum creatinine doubled, time to first potential EH (2 different criteria separately), time to dialysis.	14) Due to harmonization with Fibrogen 060 study SAP and the expected difference in treatment durations.
		15) Update regarding the analysis on cholesterol data: use of all data as main analysis and fasted only as additional instead of fasted only as main.	15) Due to harmonization with Fibrogen 060 study SAP.
		16) Reduction in the number of analyses of subgroup of interests.	16) Due to harmonization with Fibrogen 060 study SAP.
		17) Addition of the analysis of percentage of time when Hb > 12, 13 or 14.	17) The concept of percentage of time with high Hb values was implemented in order to account for differences in individual treatment durations.
		18) MMRM model added for number of RBC and volume of RBC during the treatment period.	18) Inferential analysis for RBC was missing in the v1.0.
		19) Removal of the descriptive analysis on change from day of first dialysis to week 4 and week 12 of dialysis for SF-36, FACT-AN scores, EQ-5D, WPAI.	19) Exploratory summary statistics of SF-36, FACT-AN scores, EQ-5D, WPAI for patients on dialysis not required since this is a small subset of patients and not based on the randomized population.
		20) Additional tables run on group of subjects enrolled after protocol amendment implementation.	20) Added in order to assess the impact of protocol amendment 1.0 and 2.0.
		21) Definition of hypertension updated	21) Due to harmonization with

		using ≥ instead of > for the SBP and DBP thresholds 22) Addition of subgroup analyses on Baseline VT and PF subscores for SF-36 23) Change in the choice of covariance structure for MMRM models. 24) Implementation of de facto sensitivity analyses on the FAS for the primary endpoint and secondary endpoints related to Hb	Fibrogen 060 SAP. 22) Due to harmonization with 608 SAP for HEOR purpose. 23) Due to harmonization with Fibrogen 060 SAP. 24) Added to perform sensitivity analyses accounting for all Hb assessments collected during the study (real practice analysis on ITT)
2.0	07 December 2016	NA	Document finalized
3.0	19 September 2018	1) The alternative definitions of the Primary Efficacy Endpoint Hb response have been clarified by: a) Removing 'without discontinuation' and 'regardless of discontinuation'. b) Including an additional definition dropping the criterion of at least 1.0 g/dL increase. 2) For the key secondary endpoint Hb change from BL to the average Hb in weeks 28-36 have been clarified by removing 'without discontinuation' and 'regardless of discontinuation' and created one definition for regardless rescue therapy. 3) The key secondary endpoint Mean monthly IV iron use (mg) per subject during weeks 1 to 36 (monthly defined as a period of 4 weeks) has been revised as Time to first IV iron use during weeks 1 to 36.	1a), 2) All Hb values up to the End of Treatment assessment will be considered independent if subject discontinued yes/no. 1b) Since the latest protocol amendment 2.0 allows subjects with screening Hb between 10.0 and 10.5 to be enrolled, this additional definition of Hb response has been added (i.e. Hb ≥ 11.0 g/dL only). 3) Small number of subjects were treated with IV iron and therefore the mean monthly IV iron dose could not be considered as a continuous endpoint. Instead we will look at it as a binary endpoint with a time to event analysis. Results for both endpoints will be described.
3.0	19 September 2018	4) Derivation details has been added or further clarifications have been included where needed: a. Occurrence and time to first potential EH: b. Iron Use: Derivation of iron based on the ATC 3th class has been clarified.	4) Details added to support the statistical programming of the derived variables.

		c. Number of days of hospitalization per patient-exposure-year (PEY). d. Concomitant medication definition. e. FACT-An total score. f. Time to Treatment Discontinuation: time to censoring derivation. g. Time to occurrence of any TEAE (by type of AE group): Bullet point 'TEAEs leading to Death' is updated as 'Death during the Safety Emergent Period'. h. For HRQoL endpoint PGIC, only the last assessment will be used. i. For FACT-An, the missing score 'FACT-An Fatigue Score' is added. j. TEAE definition for AE both start and stop dates are completely missing and the adverse event occurs on Day 1 check box is missing is added. k. PCS Laboratory Criteria: Hemoglobin >120 g/L is updated as Hemoglobin >14 g/dL. l. For the statistical analysis, treatments will be pooled across roxadustat under both protocol versions. m. Demographic and Other Baseline Characteristics: eGFR categories 60-<90 and ≥90 mL/min/1.73m2 is combined into one category. n. Volume transfused when the number of RBC units is given o. Occurrence of rescue therapy as a binary variable.	
3.0	19 September 2018	5) More extensive analyses of adverse events are included: a. For adverse events incidence rate, number of subjects at risk. b. Time to occurrence of any death (during the 24-month Period including follow-up) c, Definitions of event rate d. AE up to plus 7 Days	5) Additional AE analysis have been included to harmonize with the other placebo-controlled Phase 3 studies with roxadustat. The adjudicated events for this study will not be part of the pooled CV analysis as this study is still ongoing. In addition, this

		e. Adjudicated CV related Events	study is the only active-controlled study in non-dialysis patients.
3.0	19 September 2018	6) Iron, HbA1c and CKD progression parameters: a. clarified that these are exploratory variables. b. Clarified that any assessment for creatinine and eGFR occurring after the initiation of acute or chronic dialysis will be excluded for the summaries and further analysis. c. Updated the list of exploratory endpoints related to CKD progression. The composite endpoint as well as the components will be included.	6) Clarification added and to harmonize the CKD progression endpoints in line with the other placebo-controlled Phase 3 studies.
3.0	19 September 2018	7) For the primary aanalysis of the primary endpoint, SAS code using version 9.3 is added 8) Section 7.4.1.2 Secondary Analyses (sensitivity) of the Primary Endpoint In the Table a column is added for the analyses repeated excluding data from the site(s) with potential data quality issues.	7) Details of the SAS code has been added knowing that SAS 9.4 is not implemented. 8) As a sensitivity analysis of efficacy data.
3.0	19 September 2018	9) Section 7.4.1.3 Additional Analyses of the Primary Endpoint and section 7.4.2.8 Additional Analyses of the Key Secondary Endpoints a) analysis for subgroups Baseline CRP group and Iron repletion status are added. b) The following footnote is added for the table in section 7.4.1.3 Additional Analyses of the Primary Endpoint: In case model does not converge for some subgroups, no adjustment on covariates will be done. c) . ≥10.0 g/dL is added for all categorical endpoints using the Hb values.	9 a) To provide analysis for these subgroups of interest. b) To clarify about the model convergence issue. c) to harmonize across the development program
3.0	19 September 2018	10) Interpretation of fixed sequencing test is updated for tests 3, 6, 8 and 9.	10) For test 3, this is due to the change in endpoint as mentioned in Section

		11) for EQ-5D 5L, an inferential analysis for the VAS score average value in weeks 12-28 is added. 12) Additional subgroups are added for Baseline eGFR (< 15 vs ≥ 15), Baseline CRP group (≤ULN vs > ULN) and Baseline Iron repletion status (TSAT ≥20% and ferritin ≥100 ng/mL Yes vs No).	6.1.2 Key Secondary Efficacy Endpoints. For tests 2, 6 and 8, typographical error. 11) To provide inferential analysis for EQ-5D VAS score. 12) Additional clinically important subgroups have been identified and included in the analysis.
3.0	19 September 2018	13) Section 7.10 Interim Analysis (and Early Discontinuation of the Clinical Study) Clarified that results of the endpoints measured over a period longer than 36 weeks will be mainly described to provide full disclosure, but no interferential analyses will be done. Also clarified that Adjudicated data will not be part of the interim analysis. 14) Section 7.11.1 Missing Data For other endpoints, paragraph related to LOCF method imputation is removed. 15) End of Safety Emergent Period and End of Efficacy Period are revised to take into account the different dosing frequency of darbepoetin treatment.	13) More details on the interim analysis is provided, as there is no separate interim analysis plan compiled. 14) LOCF method will not be applied due to comments from the FDA. 15) Darbepoetin is administered in a different (less frequent) dosing regimen compared to roxadustat. The definition of the periods have been revised to align both treatment regimens with respect to the end date.

9 REFERENCES

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

ICH Harmonized Tripartite Guideline E 9. Statistical Principles for Clinical Trials, February 1998. (www.ich.org; Guidelines; "Efficacy" Topics)

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

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

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

Miettinen, O. and Nurminen, M., Comparative Analysis of Two Rates, Statistics in Medicine, 1985;4:213-226,

Gail, M. and R. Simon. Testing for Qualitative Interactions between Treatment Effects and Patient Subsets. Biometrics. Vol. 41 No. 2 (June 1985):361-372.

10 APPENDICES

10.1 Appendix 1: SF-36 v2

Your Health and Well-Being

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

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

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

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

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

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

SF-36 v2

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

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

SF-36 v2

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

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

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

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

SF-36 v2

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

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

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

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

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

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

SF-36 v2

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

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

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

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

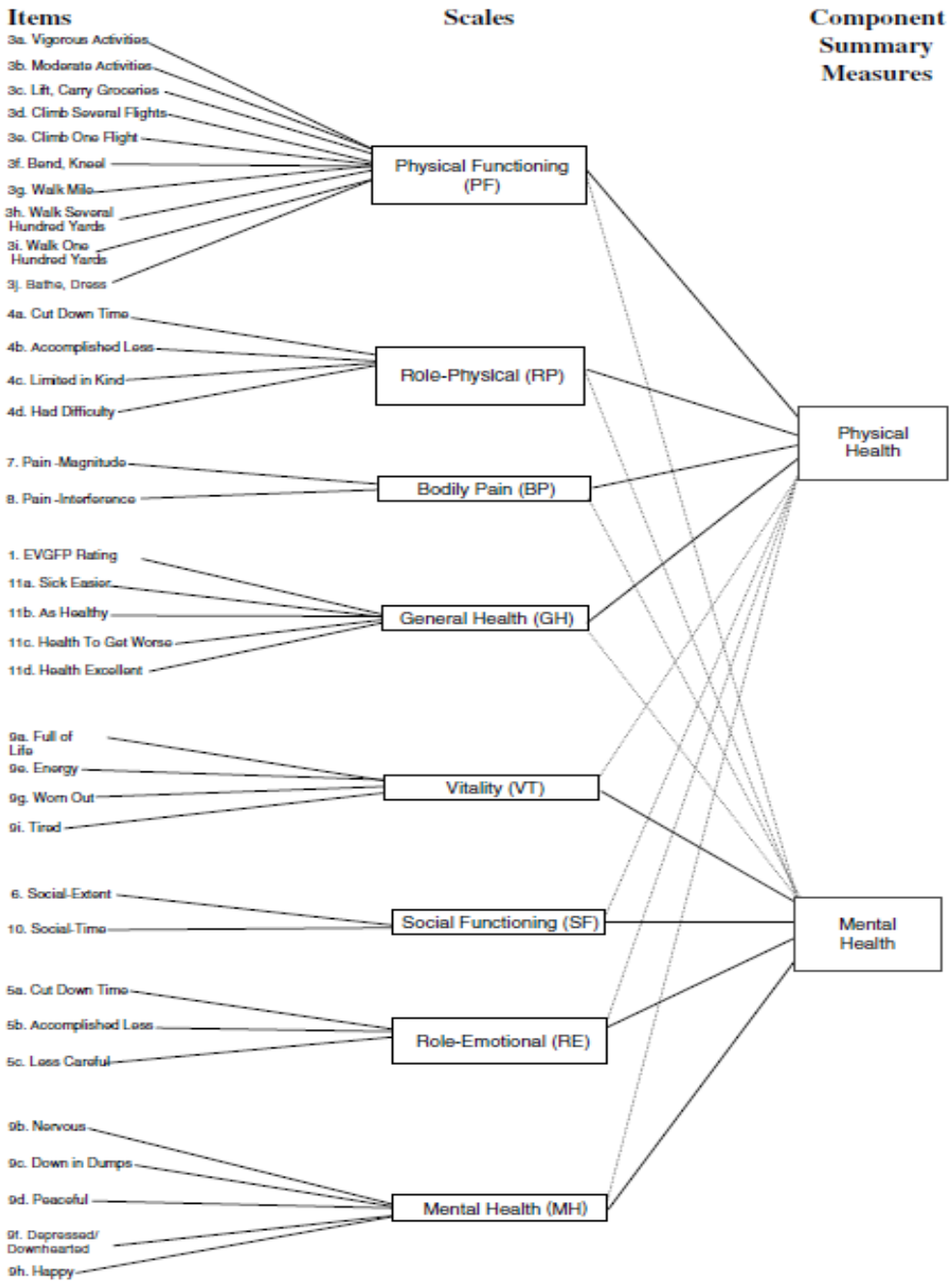
SF-36 v2

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

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

Thank you for completing these questions!

SF-36 Model


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

10.2 Appendix 2: FACT-An (Version 4)

10.2.1 FACT-An Questionnaire

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

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

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

FACT-An (Version 4)

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

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

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

FACT-An (Version 4)

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

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

10.2.2 FACT-An Scoring Guidelines

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

Subscale	Item Code	Reverse item?		Item response	Item Score
PHYSICAL WELL-BEING (PWB)	GP1	4	-	_____	=_____
	GP2	4	-	_____	=_____
	GP3	4	-	_____	=_____
	GP4	4	-	_____	=_____
Score range: 0-28	GP5	4	-	_____	=_____
	GP6	4	-	_____	=_____
	GP7	4	-	_____	=_____
				Sum individual item scores: _____	
				Multiply by 7: _____	
				Divide by number of items answered: _____	
				=PWB subscale score	
SOCIAL/FAMILY WELL-BEING (SWB)	GS1	0	+	_____	=_____
	GS2	0	+	_____	=_____
	GS3	0	+	_____	=_____
	GS4	0	+	_____	=_____
Score range: 0-28	GS5	0	+	_____	=_____
	GS6	0	+	_____	=_____
	GS7	0	+	_____	=_____
				Sum individual item scores: _____	
				Multiply by 7: _____	
				Divide by number of items answered: _____	
				=SWB subscale score	
EMOTIONAL WELL-BEING (EWB)	GE1	4	-	_____	=_____
	GE2	0	+	_____	=_____
	GE3	4	-	_____	=_____
	GE4	4	-	_____	=_____
Score range: 0-24	GE5	4	-	_____	=_____
	GE6	4	-	_____	=_____
				Sum individual item scores: _____	
				Multiply by 6: _____	
				Divide by number of items answered: _____	
				=EWB subscale score	

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

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

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

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

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

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

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

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

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

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

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

FACIT-Fatigue Subscale Scoring Guidelines

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

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

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

10.3 Appendix 3: 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 ☐

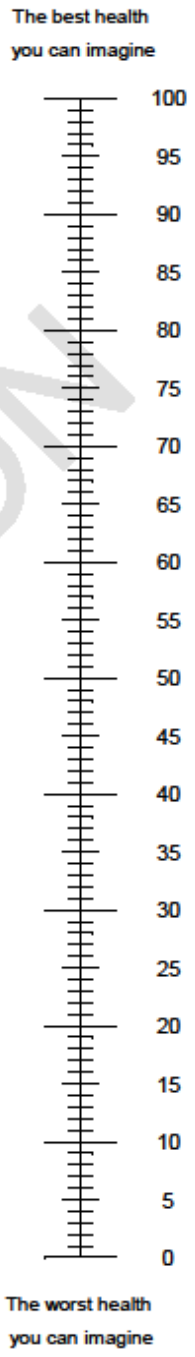
ANXIETY / DEPRESSION

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

10.4 Appendix 4: WPAI:ANS V2.0

The following questions ask about the effect of your anaemic symptoms on your ability to work and perform normal daily activities. *Please fill in the blanks or circle a number, as indicated.*

1. Are you currently employed (working for pay)? _____NO _____YES
   If NO, tick "NO" and skip to question 6.

The next questions refer to the **past seven days**, not including today.

2. During the past seven days, how many hours did you miss from work because of problems associated with your anaemic symptoms? *Include hours you missed on sick days, times you went in late, left early, etc., because of your anaemic symptoms .*
Do not include time you missed to participate in this study.

   _____HOURS

3. During the past seven days, how many hours did you miss from work because of any other reason, such as annual leave, holidays, time off to participate in this study?

   _____HOURS

4. During the past seven days, how many hours did you actually work?

   _____HOURS *(If "0", skip to question 6)*

5. During the past seven days, how much did your anaemic symptoms affect your productivity while you were working?

Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual.
If anaemic symptoms affected your work only a little, choose a low number. Choose a high number if anaemic symptoms affected your work a great deal.

WPAI:ANS

Consider only how much anaemic symptoms affected productivity while you were working.

Anaemic symptoms had no effect on my work	0	1	2	3	4	5	6	7	8	9	10	Anaemic symptoms completely prevented me from working

CIRCLE A NUMBER

6. During the past seven days, how much did your anaemic symptoms affect your ability to perform your normal daily activities, other than work at a job?

By normal activities, we mean the usual activities you perform, such as working around the house, shopping, childcare, exercising, studying, etc. Think about times you were limited in the amount or kind of activities you could perform and times you accomplished less than you would like. If anaemic symptoms affected your activities only a little, choose a low number. Choose a high number if anaemic symptoms affected your activities a great deal.

Consider only how much anaemic symptoms affected your ability to do your normal daily activities, other than work at a job.

Anaemic symptoms had no effect on my daily activities	0	1	2	3	4	5	6	7	8	9	10	Anaemic symptoms completely prevented me from doing my daily activities

CIRCLE A NUMBER

10.5 Appendix 5: Patient Overall Impression of Change

Since the start of the study, my general state of health is:

tick one box only

- ☐ [1] Very Much Improved
- ☐ [2] Much Improved
- ☐ [3] Minimally Improved
- ☐ [4] No Change
- ☐ [5] Minimally Worse
- ☐ [6] Much Worse
- ☐ [7] Very Much Worse

10.6 Appendix 6: Medication WHO Drug Dictionary Codes

Name	Code
ESA except darbepoetin alfa	ATC level 4 = B03XA [WHODD drug code = '00909301001', '00928301001', '07973701001', '01703101001']
Darbepoetin alfa	ATC level 4 = B03XA [WHODD drug code = '02198701001']
IV Iron	CM3ATCL="IRON Preparations" AND route = intravenous
Oral Iron	CM3ATCL="IRON Preparations" AND route = oral
RBC transfusion	ATC level 4 = B05AX [WHODD drug code = '01186901001']
Any investigational drug	WHODD drug code = 99999701001
Hypoxia-inducible factor HIF-PHI	ATC level 4 = B03XA
Iron-chelating agents	ATC code = V03AC01, V03AC02 and V03AC03
Androgens	ATC level 3 = G03B and G03E
Dapsone	ATC level 4 = J04BA
Acetaminophen/paracetamol	ATC level 4 = N02BE, N02AA, R05X

10.7 Appendix 7: Signatures

List of Key Contributors and Approvers

Key Contributors

The following contributed to or reviewed this Statistical Analysis Plan as relevant to their indicated discipline or role.

Primary author

PPD	IQVIA

Contributors and Reviewers

PPD	Data Science
PPD	Data Science
PPD	Clinical Science
PPD	IQVIA
PPD	Medical Science, Chiltern

Approval Signatories

(E-signatures are attached at end of document)

PPD was the study statistician for this study.

PPD was the Global Statistician Leader and biostatistics peer reviewer of this Statistical Analysis Plan.

I approve the content of this Statistical Analysis Plan.

PPD

Medical Science – Nephrology
Astellas Pharma Europe B.V.