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

**ISN/Protocol 1517-CL-0608**

## ClinicalTrials.gov Identifier: NCT01887600

## Date of Statistical Analysis Plan: Final Version 5.0, dated 02 Aug 2018

**Sponsor: Astellas Pharma Europe B.V. (APEB)**

Sylviusweg 62
2333 BE Leiden
The Netherlands

# STATISTICAL ANALYSIS PLAN

Final Version 5.0, dated 02 August 2018

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

ISN: 1517-CL-0608
EudraCT number: 2012-005180-27

Sponsor name
Astellas Pharma Europe B.V. (APEB)
Sylviusweg 62,
2333 BE Leiden
The Netherlands

## Table of Contents

## List of In-Text Tables

# I. LIST of ABBREVIATIONS AND KEY TERMS

## List of Abbreviations

| Abbreviations | Description of abbreviations |
|---|---|
| AE | Adverse Event |
| ALP | Alkaline Phosphatase |
| ALT | Alanine Aminotransferase (GPT) |
| ANCOVA | Analysis of Covariance |
| AnS | Anemia Subscale |
| anti-HCV Ab | Anti-hepatitis C Virus Antibody |
| APEB | Astellas Pharma Europe B.V. |
| Apo | Apolipoproteins |
| ASC | Analysis Set Classifications |
| ASP1517 | FG-4592 (codename of investigational product) or roxadustat (international nonproprietary name) |
| AST | Aspartate Aminotransferase (GOT) |
| AT | Aminotransferase |
| ATC | Anatomical Therapeutic Chemical |
| BL | Baseline |
| BL Hb | Baseline Hemoglobin (please refer to key definitions for infoFrmation) |
| BMI | Body Mass Index |
| BP | Blood Pressure |
| BUN | Blood Urea Nitrogen |
| CBC | Complete Blood Count |
| CHr | Reticulocyte Hemoglobin Content |
| CI | Confidence Interval |
| CKD | Chronic Kidney Disease |
| CMH | Cochran-Mantel-Haenszel |
| CRF | Case Report Form |
| CRO | Contract Research Organization |
| CRP | C Reactive Protein |
| CS | Classification Specifications |
| CSE | Composite Safety Endpoint |
| CSR | Clinical Study Report |
| CV | Coefficient of Variation |
| DBP | Diastolic Blood Pressure |
| dL | Deciliter |
| DSMB | Data Safety Monitoring Board |
| ECG | Electrocardiogram |

| Abbreviations | Description of abbreviations |
|---|---|
| eCRF | Electronic CRF |
| EDC | Electronic Data Capture |
| eGFR | Estimated Glomerular Filtration Rate |
| EH | Excessive Hematopoiesis |
| EOS | End of Study |
| EOT | End of Treatment |
| 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) |
| FSI | First Subject In |
| FWB | Functional Well-being |
| g | gram |
| GDS | Global Data Science |
| GGT | Gamma Glutamyl Transferase |
| GM | Geometric Mean |
| 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 |
| HIV | Human Immunodeficiency Virus |
| HR | Heart Rate |
| HRQoL | Health-Related Quality of Life |
| hs-CRP | High Sensitivity C-reactive protein |

| Abbreviations | Description of abbreviations |
|---|---|
| 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 |
| INN | International Nonproprietary Name |
| INR | International Normalized Ratio |
| IPCW | Inverse Probability of Censoring Weighting |
| IRT | Interactive Response Technology |
| ISN | International Study Number |
| IV | Intravenous(ly) |
| Kg | Kilograms |
| LDL | Low-density Lipoprotein |
| LA-CRF | Liver Abnormality Case Report Form |
| LFT | Liver Function Tests |
| LLN | Lower Limit of Normal |
| LOCF | Last Observation Carried Forward |
| LSO | Last Subject Out |
| MACE | Major cardiovascular adverse events: myocardial infarction, stroke, death from all causes |
| MACE+ | Myocardial infarction, stroke, death from all causes, chronic heart failure requiring hospitalization, unstable angina requiring hospitalization |
| MAP | Mean Arterial Pressure |
| MDRD | Modification of Diet in Renal Disease |
| MedDRA | Medical Dictionary for Regulatory Activities |
| mg | Milligram |
| MI | Myocardial Infarction |
| mL | Milliliters |
| 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 |
| O | Optional |
| PCS | Physical Component Score |
| PD | Protocol Deviation |
| PDAS | Pharmacodynamic Analysis Set |

| Abbreviations | Description of abbreviations |
|---|---|
| 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 |
| 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 |
| SD | Standard Deviation |
| SI | International System of Units |
| SF-36 | Short Form 36 |
| 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 |
| SMQ | Standardized MedDRA Query |
| SOC | System Organ Class |
| SOP | Standard Operating Procedure |
| SPA | Special Protocol Assessment |
| SQ | Subcutaneous |
| sTfR | Soluble Transferrin Receptor |
| SWB | Social Well-being |

| Abbreviations | Description of abbreviations |
|---|---|
| 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-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 (AE) is as any untoward medical occurrence in a subject administered the study drug, roxadustat or placebo, or who has undergone study procedures and which does not necessarily have a causal relationship with this treatment. AE collection starts after obtaining signed informed consent and continues until the End of Study visit. AEs will not be collected during the period between first screen where subject has failed screening and first rescreening visit. |
| Baseline | 1) Observed values/findings which are regarded as calibrated zero status in the present study; 2) Time when 'Baseline' is observed. |
| Baseline Hemoglobin (Hb) value | Baseline Hb is defined as the mean of four central laboratory Hb values: four latest Hb values prior or on the same date as first study drug intake (pre-dose). |
| Discontinuation | The act of concluding participation in either the study treatment or the study, 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. |
| 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 or is contemporaneous with randomization. |
| Extended treatment period | Period of time that patient is treated from end of primary treatment period (52 weeks) up to 104 weeks |
| Hb Response | • Hb ≥11.0 g/dL and a Hb increase from baseline (BL) by ≥1.0 g/dL in any subject with BL Hb>8.0 g/dL, OR<br>• an increase from BL by ≥2.0 g/dL in any subject with BL Hb ≤8.0 g/dL<br>at two consecutive visits [dates] (with available data) separated by at least 5 days during the first 24 weeks of treatment without rescue therapy (i.e., red blood cell [RBC] transfusion, ESA, or intravenous [IV] iron) prior to Hb response. |
| 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. |

| Terms | Definition of terms |
|---|---|
| Investigational period | Period of time where major interests of protocol objectives are observed, and where the test drug or placebo (sometimes without randomization) is usually given to a subject, and continues until the last assessment after completing administration of the test drug or placebo. |
| Post study follow-up | Period of time from EOS visit to projected week 108 or until the last subject randomized reaches EOS, whichever comes first. This period is only applicable to subjects who discontinued treatment. These subjects will be followed up on a 6-monthly frequency for vital status and hospitalizations. |
| Primary treatment period | Period of time that subject is treated from first treatment up to 52 weeks |
| 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 test drug or comparative drug is given to a subject |
| Randomization | Action to allocate a subject to the treatment group or treatment cohort. Subjects will be randomized to roxadustat or placebo at day 1. |
| Roxadustat | International Nonproprietary Name (INN) of ASP1517/FG-4592 investigational product |
| Rescreening | Process of repeating screening. If a subject fails screening they may be re-screened once if deemed appropriate; all screening procedures will be repeated. Renal ultrasound only to be repeated if not within 12 weeks prior to randomization. |
| Rescreening failure | Subject who is rescreened, but did not fulfill protocol inclusion and/or exclusion criteria for a second time and failed to receive randomized treatment, or decided not to participate anymore (withdrew consent) prior to the treatment period. |
| Safety Emergent Period | Defined as the evaluation period from the Analysis date of first drug intake up to 28 days after the Analysis Last Dose |
| Screening | 1) Process for retrieving candidates for the study. 2) Process for checking the eligibility of subjects usually done during the “pre-investigational period” |
| Screening failure | Screened subject, but did not fulfill protocol inclusion and/or exclusion criteria and failed to receive randomized treatment, or decided not to participate anymore (withdrew consent) prior to the treatment period. |
| Screening Hb value | Mean of subject’s three last Hb valued collected during the screening period and prior to the day of randomization. |
| 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. |

| Terms | Definition of terms |
|---|---|
| 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, either as a recipient of the investigational product(s) or as a control. |
| Time to event | Time from a defined starting point (analysis date of first dose intake) to the 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 | Period of time from first study drug intake until last study drug intake. Minimum 52 weeks to a maximum of 104 weeks or until the last subject randomized to treatment has completed 40 weeks of treatment (or at the forecasted week 40 date if this subject discontinues treatment early) |
| 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 endpoints and other data.

The SAP is finalized and signed prior to any of the following: study unblinding, database hard lock, interim analysis, or accumulation of substantial amount of data in an open-label study to ensure lack of bias. For operational efficiency an earlier time is usually targeted and wherever possible, the SAP should be developed in parallel with protocol finalization. If the expected interval between First Subject In (FSI) and soft-lock is less than 12 weeks, then the SAP should be approved by 12 weeks prior to the planned date of soft-lock. If needed, revisions to the approved SAP may be made prior to database hard lock. Revisions will be version controlled.

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

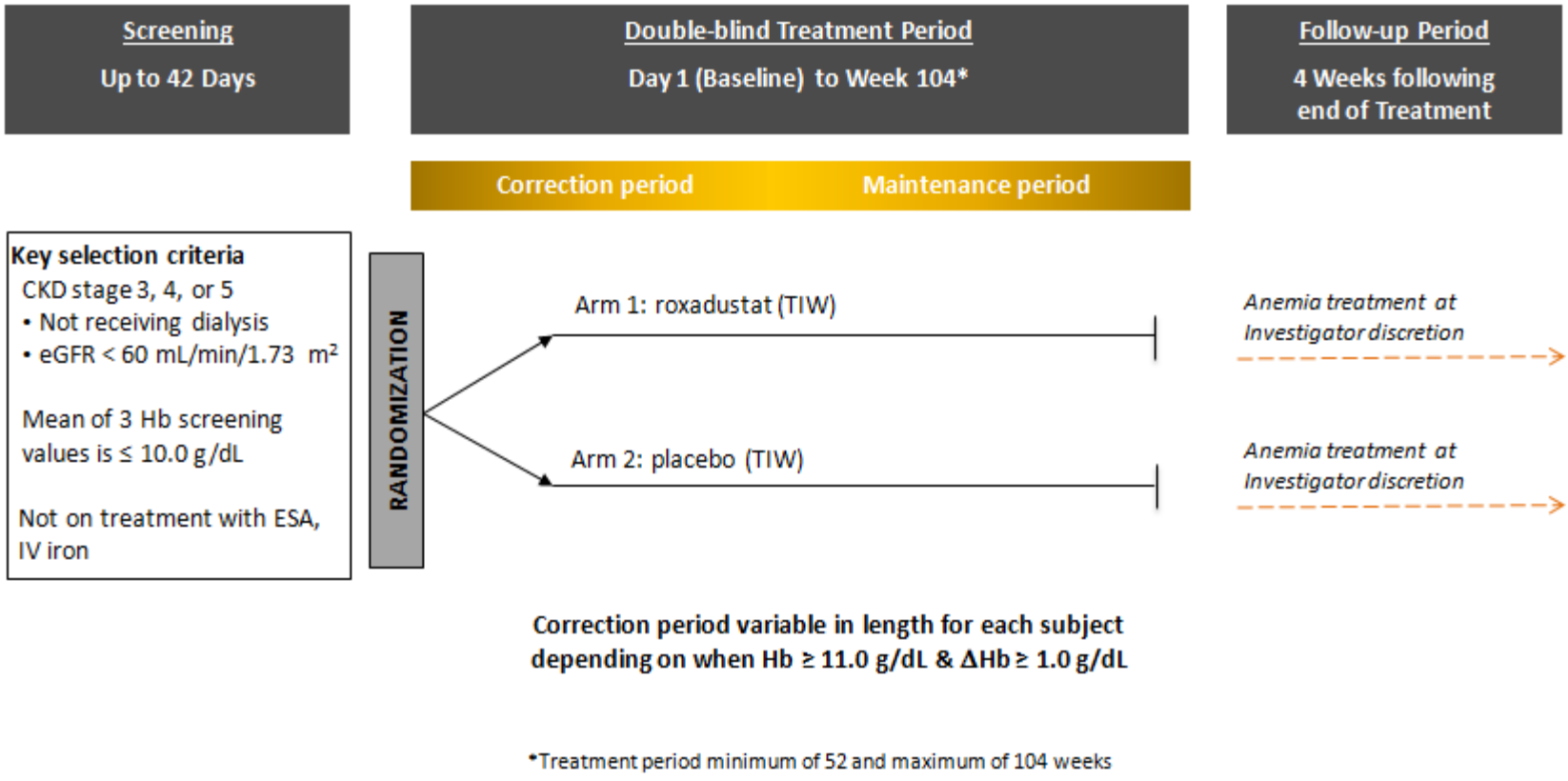
All details of the Pharmacokinetics Analysis Set (PKAS)_will be described in a separate analysis plan, and a separate PKAS modeling report will be written.

This SAP is based on protocol version 2.0, dated 17 December 2014 and on Case Report Form (CRF) version 17.0, dated 26 April 2017.

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

## Flow Chart



## Table 1 Schedule of Assessments

| **Study Period:** | **Screening** | | | **Treatment**[a] | | | | **Follow-up** | | | | **Post study Follow- up** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Up to 6 Weeks** | | | | | | | | | | | |
| **Visit / Week:** | **S1** | **S2** | **S3** | **Day 1**[b] | **Weekly** (wks 1 to 2) ± 2 days | **Every 2 Weeks** (wks 4 to 24) ± 2 days | **Every 4 Weeks** (wks 28 to 100) ± 3 days | **EOT** (wk 104) ± 3 days | **EOT + 2 wks** ± 3 days | **EOS** (EOT **+ 4 wks**) ± 3 days | **Unscheduled Visits** | Every 6 months until projected wk 108 |
| Written informed consent | X | | | | | | | | | | | |
| Randomization | | | | X | | | | | | | | |
| Eligibility criteria | X | | | X | | | | | | | | |
| Demographics and medical history including tobacco use | X | | | | | | | | | | | |
| Weight | X | | | X | | wks 12, 24 | wks 36, 52, 76 | X | | X | O[d] | |
| Physical examination | X | | | X | | wks 12[c] 24[c] | wks 36[c], 52[c], 76[c] | X | | X[c] | O[c, d] | |
| Blood pressure[e], heart rate[e], respiratory rate[g] | X | X | X | X | X | X | X | X | | X | X | |
| CBC with WBC differential, red cell indices and platelet count | X | | | X | X | wks 4, 8, 12, 20 | wks | X | | X | O[d] | |
| Reticulocyte count, Hemoglobin content of reticulocytes (CHr) | X | | | X | X | wks 4, 6, 8, 12, 16, 20 | wk 28 and every following 8 wks | X | | X | O[d] | |
| Hemoglobin[h] | | X | X | | | X | X | | X | | X | |
| HemoCue® assessment [i] | | | | X | X | X | X | | | | X | |
| Serum chemistry (incl LFT) | X | | | X | | wks 4, 8, 12, 20 | wk 28 and every following 8 wks | X | X | X | O[d] | |
| LFTs [j] | | | | | wk 2 | wks 6, 16 | | | | | O[d] | |
| Serum Lipid panel (fasting whenever possible) | X | | | X | | wks 4, 8, 12, 20 | wks 28, 36, 44, 52, 68, 84 | X | | X | O[d] | |
| Serum iron, ferritin, TIBC, TSAT | X | | | X | | wks 4, 8, 12, 20 | wk 28 and every following 8 wks | X | | X | O[d] | |
| HbA1c | X | | | X | | wk 12 | wks 28, 36, 44, 60, 84 | X | | X | O[d] | |
| Vitamin $B_{12}$, folate | X | | | | | | | | | | | |
| HIV Immunoassay, HBsAg, anti-HCV antibody | X | | | | | | | | | | | |

*Table continued on next page*

| **Study Period:** | **Screening** | | | **Treatment**[a] | | | | **Follow-up** | | | | **Post study Follow- up** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Up to 6 Weeks** | | | | | | | | | | | |
| **Visit / Week:** | **S1** | **S2** | **S3** | **Day 1**[b] | **Weekly** (wks 1 to 2) ± 2 days | **Every 2 Weeks** (wks 4 to 24) ± 2 days | **Every 4 Weeks** (wks 28 to 100) ± 3 days | **EOT** (wk 104) ± 3 days | **EOT + 2 wks** ± 3 days | **EOS** (EOT **+ 4 wks**) ± 3 days | **Unscheduled Visits** | Every 6 months until projected wk 108 |
| Serum Pregnancy test[k] | X | | | | | wks 12, 24 | wks 36, 48, 60, 72, 84, 96 | X | | | O[d] | |
| eGFR (Cr Clear Modified Diet Abbreviated) [l] | X | | | X | | wk 20 | wks 36, 52, 68, 84 | X | | X | O[d] | |
| Special laboratory analytes (hepcidin, sTfR, hs-CRP) | | | | X | | wks 4, 12, 20 | wks 36, 52 | X | | X | | |
| Archival serum/plasma samples for biomarkers | | | | X | | wks 4, 12, 20 | wks 52, 76 | X | | X | | |
| Blood sample for population PK | | | | | wks 2 to 8[m] | | | | | | | |
| Genotyping[n] | | | | | X | | | | | | | |
| Urinary testing[o] | | | | X | | wks 12, 24 | wks 36, 52, 64, 76, 88 | X | | | O[d] | |
| Quality of Life Questionnaires [p] | | | | X | | wks 8, 12, 28 | wks 36, 52, 76 | X | | | O[d] | |
| 12-lead ECG | X | | | X | | wks 12, 24 | wks 36, 52, 76 | X | | | O[d] | |
| Renal ultrasound[q] | X | | | | | | | | | | O[d] | |
| Dose adjustment review[r] | | | | | | X | X | | | | O[d] | |
| Hospitalization recording[s] | X | X | X | X | X | X | X | X | X | X | X | X |
| Adverse event recording | X | X | X | X | X | X | X | X | X | X | X | |
| Concomitant medication recording | X | X | X | X | X | X | X | X | X | X | X | |
| Procedure and non-drug therapy recording | X | X | X | X | X | X | X | X | X | X | X | |
| Study drug dispensing[t] | | | | X[u] | X | X | X | | | | O[d] | |
| Vital Status, SAEs, cardiovascular and thromboembolic AEs | | | | | | | | | | | | X |

S1/S2/S3 = screening visit 1, 2 and 3; EOT = End of Treatment; EOS = End of Study; wk(s) = week(s); X = mandatory test/assessment; O = optional test/assessment; see below for footnotes)

**Note:** see Appendix 3 from Protocol: Instructions for Subjects Moving from Protocol v1.0 to Protocol v2.0

**Note**: see Appendix 4 from Protocol: Instructions for Subjects Requiring Dialysis

*Table footnotes continued on next page*

[a] In case of premature discontinuation or withdrawal during the treatment period, the subject should complete the EOT and EOS visits. Thereafter, this subject will continue to be followed up at a 6-monthly frequency for vital status and hospitalizations until their projected date of completion (i.e., projected week 108 date) or, if earlier, until the last subject randomized reaches EOS, or until consent withdrawn.

[b] All study assessments to be performed prior to first study drug administration

[c] Targeted physical examination only (e.g., respiratory and cardiovascular).

[d] The study drug dosing is to be reviewed and if needed new or additional study drug is to be dispensed.

[e] Blood pressure (BP) measured singly during the screening period, and in triplicate at all other visits. It is recommended during the treatment period, blood pressure measurement 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 will be recorded prior to, and after dialysis (hemodialysis [HD]/hemodiafiltration [HDF] subjects only).

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

[g] Respiratory rate measured singly during all visit. It is recommended during the treatment period, respiratory rate 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] Hemoglobin (Hb) should be collected at all the visits where complete Blood Count (CBC) is not collected

[i] Hb will be assessed by HemoCue on the blood sample, collected for Central Laboratory hemoglobin assessment

[j] Liver Function Tests (LFTs) to be collected at visits where full Serum Chemistry is not collected

[k] Collect from female subjects of child bearing potential only.

[l] Calculated by the Central Laboratory.

[m] Sampling roxadustat will be done at 6 time points over 1 to 3 visits. See Section 5.6 from Protocol. At each pharmacokinetic visit, an additional sample will be collected for albumin and alpha-acid glycoprotein determination.

[n] Optional assessment. A separate informed consent form must be signed before genotyping sample is collected. Sample collection can be done at any timepoint thought the treatment period of the study.

[o] 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. At day 1, weeks 24, 52 and 76 and EOT a urine sample will be archived for potential future biomarker analysis.

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

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

[r] Dose adjustment review from week 4 onward, and every 4 weeks thereafter until EOT (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.

[s] Telephone or in-person follow-up call with subject

[t] For subjects requiring dialysis, it is recommended for HD/HDF subjects that study drug is administered any time after completion of dialysis (if dosing is scheduled on a dialysis day).

[u] Intake of initial study drug on day of randomization.

# 3 STUDY OBJECTIVES AND DESIGN

## 3.1 Study Objectives

### 3.1.1 Primary Objective

The primary objective of this study is to evaluate the efficacy of roxadustat in the treatment of anemia in non-dialysis CKD subjects.

### 3.1.2 Secondary Objectives

The secondary objectives in this study are to:

- Evaluate the safety of roxadustat in the treatment of anemia in non-dialysis CKD subjects.
- Evaluate HRQoL benefit of roxadustat treatment in subjects with non-dialysis CKD anemia.
- Evaluate the need for anemia rescue therapy with roxadustat in subjects with non-dialysis CKD anemia: RBC transfusion, ESA, or IV iron.

## 3.2 Study Design

### 3.2.1 General

This is a phase 3, multi-center, randomized, double-blind, placebo controlled study in anemic subjects with Stage 3, 4 or 5 CKD who are not on dialysis. This study is planned to recruit 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 anemia associated with CKD. Study FGCL-4592-060 with similar design is conducted by FibroGen Inc, in study centers across North America, Latin America and Asia Pacific.

### 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 ≤ 10.0 g/dL upon repeated screening measurements. Subjects do not need to be iron replete at BL; inclusion is permitted if ferritin ≥30 ng/mL (≥ 67.4 pmol/L) and Transferrin Saturation (TSAT) ≥ 5%. Anemia of non-renal origin is to be excluded. Washout periods of at least 12 weeks for any prior ESA or IV iron treatment or 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 will take roxadustat or placebo orally as a combination of tablets of different strengths. All tablets for subjects receiving roxadustat will contain active ingredient whereas all tablets for subjects receiving placebo will contain just placebo. The study will consist of three study periods as follows:

- Screening period: up to 6 weeks.
- Treatment period: minimum 52 weeks (primary treatment period) up to a maximum of 104 weeks (extended treatment period) or until the last subject randomized to treatment has completed 40 weeks of treatment (or at the forecasted week 40 date if this subject discontinues treatment early). Treatment period for last patient randomized is 52 weeks.
- Post-treatment follow-up period: 4 weeks.
- Study termination and post-study follow-up period

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

Subjects that have discontinued treatment prior to their projected week 104 will continue to be followed for vital status and hospitalizations in post study follow up.

**Screening Period**

During the screening period, subjects' eligibility for study participation will be assessed.

**Treatment Period**

The initial protocol Version 1.0 included a random allocation to 6 treatment arms in a 2:2:2:1:1:1 ratio. These 6 arms included placebo or roxadustat as double-blind treatment in a 2:1 ratio and each of these had 3 different dosing frequencies for the maintenance period. (TIW, BIW or QW). Following FDA's recommendation, the protocol was amended with the removal of the dosing frequencies BIW and QW and keeping only the TIW dosing frequency. This SAP is based on protocol version 2.0 using the pooled placebo and pooled roxadustat arms as treatment groups for comparison.

After subjects have been confirmed eligible for study participation, they will be randomized to receive 1 of 2 treatment arms. The randomization will result in a 2:1 ratio of subjects receiving roxadustat or placebo, respectively.

The initial study drug dose (per dose amount) is based on a tiered, weight-based dosing scheme shown in Table 2

**Table 2 Initial Study Drug (Roxadustat/ Placebo) Dosing**

| **Study Drug** (Dose Frequency) | **Weight** (≥ 45 to ≤ 70 kg) | **Weight** (> 70 to ≤ 160 kg) |
|---|---|---|
| Roxadustat/Placebo (TIW) | 70 mg | 100 mg |

Study drug will be dosed initially for Hb correction, until subjects achieve central Hb value of ≥ 11.0 g/dL and Hb increase from BL of ≥ 1.0 g/dL at two consecutive study visits separated by at least 5 days (correction period).

Once Hb correction is reached the subject will enter the maintenance period. The aim of the maintenance period is to treat to a Hb level of 11.0 g/dL by maintaining the Hb levels between 10.0 g/dL and 12.0 g/dL. 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 and thereafter every four weeks until the end of treatment. Subjects will be treated with roxadustat or placebo for at least 52 weeks and will continue taking the double-blind treatment as they were assigned until a maximum of 104 weeks. Depending on the rate of recruitment, the maximum treatment period will be 104 weeks for subjects who were randomized early into the study.

- The last subject randomized will stop treatment at the minimum of 52 weeks.
- When the last subject randomized reaches 40 weeks of treatment (or the forecasted week 40 date, if the last subject randomized discontinues treatment early)
    - Subjects beyond 52 weeks treatment will stop treatment at this point.
    - Subjects that have not yet reached 52 weeks treatment will continue until they reach 52 weeks treatment.

For details on study drug dosing, see Protocol, Section 5.1.

**Follow-up Period**

After the Treatment period, subjects proceed to the 4-week post-treatment follow-up period.

**Post Study Follow-up** *(only for subjects prematurely discontinued from treatment)*

Subjects that have stopped treatment prior to their projected week 104 will complete the EOT visit and EOS visit. Thereafter, these subjects will continue to be followed up on a 6-monthly frequency for vital status and hospitalizations until their projected date of completion (i.e., projected week 108 date) or, if earlier, until the last subject randomized reaches EOS, or until consent withdrawn.

### 3.2.4 Comparator

Placebo has been chosen as comparator to adequately assess the efficacy, safety and benefit of achieving Hb correction and maintenance in anemic subjects treated with roxadustat. Scientifically, efficacy and benefit of a new investigational medicinal product is most convincingly established by demonstrating superiority in a placebo-controlled trial.

## 3.3 Randomization

A randomized double-blind 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.

Randomization and treatment assignments will be performed via Interactive Response Technology (IRT) prepared on behalf of the Sponsor (under the responsibility of the Data Science (DS) Department of APEB). Specific procedures for randomization through the IRS are contained in the study procedures manual.

A total of 450-600 planned subjects will be randomized to receive one of the 2 treatment arms in a 2:1 ratio as follows:

- Roxadustat (planned 300-400 subjects)
- Placebo (planned 150-200 subjects)

Randomization will be stratified by the following four factors:

- Region (region A versus region B)*

  * assignment to region (see Section 6.5.2) will be determined based on health care system comparability.
- Screening Hb values ( ≤ 8.0 g/dL versus > 8.0 g/dL)
- History of cardiovascular, cerebrovascular or thromboembolic diseases (Yes versus No).
- eGFR (< 30 mL/min/1.73 m2 versus ≥ 30 mL/min/1.73 m2).

Subjects randomized to roxadustat QW, BIW or TIW prior to implementation of protocol v2.0 will be pooled together. Subjects randomized to placebo will also be pooled together.

# 4 SAMPLE SIZE

A minimum of 450 and up to 600 subjects are planned to be randomized to receive roxadustat or placebo (2:1 with approximately 300 roxadustat versus 150 placebo) in a double-blind manner in order to support the primary endpoint(s) of the study.

**EU (EMA)**

Three hundred subjects for the roxadustat treatment group and 150 subjects for the placebo treatment group are needed to achieve at least power of 95% to demonstrate a statistically significant difference with a 5% two-sided significance level between roxadustat and placebo in the primary endpoint assuming that the proportion of subjects with response in the roxadustat group is at least 65% and in the placebo group is at most 25%.

**USA (FDA)**

A sample size of 450 will allow the study to have at least power of 99% to detect a 1.0 g/dL difference in mean Hb values between the two treatment groups, assuming that the common standard deviation is 1.2 g/dL using an analysis of variance (ANOVA) test with a 5% two-sided significance level.

Most importantly, this sample size is required for a meta-analysis of composite safety endpoint by pooling studies. In case the minimum size of four hundred and fifty subjects is not large enough to achieve the required number of MACE/MACE+ events, an increase in the sample size up to 600 is regarded.

# 5 ANALYSIS SETS

In accordance with International Conference on Harmonization (ICH) recommendations in guidelines E3 and E9, the analysis sets below 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 subjects.

Criterion for exclusion from All Randomized is defined as follows:

- Not randomized

The selection of subjects for the All Randomized will be confirmed in the Analysis Set Classification (ASC) meeting.

The All Randomized will be used for summaries, selected primary and secondary analyses of efficacy endpoints, as well as selected demographic and baseline characteristics.

The All Randomized Set will exclude the 3 subjects from site 70051, which has been terminated prematurely due to GCP violations including data integrity issues. Consequently, these subjects will not be included in any of the analyses.

## 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 non-missing post-dose Hb assessment. 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

The selection of subjects for the FAS will be confirmed in the Analysis Set Classification (ASC) meeting.

The FAS will be used for summaries, selected primary and secondary analyses of efficacy endpoints, as well as selected demographic and baseline characteristics.

## 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 3. This PPS will be used for all disposition, demography and baseline characteristics.

**Table 3 Criteria for excluding a 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 reference period or until EOT, whatever comes first. |
| 5 | Subjects where breaking of the randomization code occurs during the reference period or until EOT, whatever comes first. |
| 6 | Administration of wrong randomization study drug for more than one week during the reference period or until EOT, whatever comes first |
| 7 | 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. |
| 8 | 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 includes the subjects from the SAF population who meet the following criteria:

- Received at least one dose of roxadustat
- At least one quantifiable plasma concentration of roxadustat was obtained; dosing and sampling history has been recorded.

PKAS will be defined separately and all analyses will be reported in a separate report.

## 5.6 Pharmacodynamic Analysis Set (PDAS)

Not applicable in this study.

# 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 7 days after the Analysis date of Last Dose (defined in Section 6.5.4) or EOT Visit, whichever occurs first. This period will be used as reference

period for the time to event analyses related to efficacy endpoints, unless specified otherwise. 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

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

#### 6.1.1.1 Primary Efficacy Endpoint for EU (EMA)

The primary efficacy endpoint is a binary variable, Hb response (Yes/No), where Yes is defined as:

- 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 at least 5 days during the first 24 weeks of treatment without having received rescue therapy (RBC transfusion, ESA, or IV iron) prior to Hb response.

Both 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 Table 31, Section 7.11.4).

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

Baseline Hb is defined as the mean of four central laboratory Hb values: four latest Hb values prior or on the same date as first study drug intake (pre-dose).

Definition of rescue therapy is provided in Section 6.1.2.3

#### 6.1.1.2 Primary Efficacy Endpoint for US (FDA)

Hb change from baseline (BL) to the average Hb of weeks 28 to 52 regardless of rescue therapy.

The central laboratory reported Hb values will be used for this analysis.

All available Hb values obtained from the central laboratory will be used (i.e., both scheduled and unscheduled Hb values). Hb values in analysis visit windows at weeks 28, 32, 36, 40, 44, 48 and 52 will be used for the calculation of the average of weeks 28 to 52 (see Table 31 and Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

In case a subject does not have any available Hb value within this evaluation period refer to Section 7.11.1 for imputation rules.

Baseline Hb is defined as the mean of four central laboratory Hb values: four latest Hb values prior or on the same date as first study drug intake.

### 6.1.2 Key Secondary Efficacy Endpoints

The key secondary efficacy endpoints in this study are listed in Table 4

**Table 4 Key Secondary Efficacy Endpoints**

| Number | Endpoint |
|---|---|
| 1 | Hb change from BL to the average Hb in weeks 28-36, without having received rescue therapy 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 to 28. |
| 3 | Occurrence and time to first use of rescue therapy [composite of RBC transfusions, IV iron supplementation and rescue ESA] |
| 4 | Change from BL in SF-36 Vitality (VT) sub-score to the average VT sub-score of weeks 12 to 28. |
| 5 | Change from BL in SF-36 Physical Functioning (PF) sub-score to the average PF sub-score of weeks 12 to 28. |
| 6* | Change from BL in mean arterial pressure (MAP) to the average MAP value of weeks 20 to 28. |
| 7* | Occurrence and time to first occurrence of hypertension (defined as either SBP ≥ 170 mmHg AND an increase from BL ≥ 20 mmHg or as DBP ≥ 110 mmHg, AND an increase from BL of ≥ 15 mmHg |
| 8* | Rate of progression of CKD measured by annualized eGFR slope over time |

*: These key secondary endpoints will not be included in the hierarchical testing procedure.

#### 6.1.2.1 Hb change from BL to the average Hb in weeks 28-36, without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period

All available Hb values obtained from the central laboratory will be used (i.e., both scheduled and unscheduled Hb values). Hb values in analysis visit windows at weeks 28, 32 and 36 will be used for the calculation of the average of weeks 28 to 36 (see Table 31 and Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

In case a subject does not have any available Hb value within this evaluation period, or in case a subject requires rescue therapy within 6 weeks prior to and during this 8-week evaluation period, refer to Section 7.11.1 for imputation rules.

Baseline Hb is defined in Section 6.1.1.1

#### 6.1.2.2 Change from BL in Low Density Lipoprotein (LDL) Cholesterol to the Average LDL Cholesterol of Weeks 12 to 28

The analysis will be done on all values (fasted and non fasted) of Day 1 and weeks 12 to 28.

All available LDL values will be used (regardless the fasting status), i.e., both scheduled and unscheduled LDL values. 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 Table 33 and 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 Use and time to first use of rescue therapy (composite of RBC transfusions, ESA use, and IV iron)

RBC transfusion is collected in the Blood Transfusions form of the eCRF. The use of ESAs and IV iron is collected in the Concomitant Medication form of the eCRF (entries where "Given as Anemia Therapy?" is ticked). These medications will be coded into the ATC and WHO-DRL dictionaries.

The following WHO-DRL codes will be classified as ESA: '00909301001', '00928301001', '02198701001', '07973701001', '01703101001'. The following WHO-DRL code where route is INTRAVENOUS will be classified as IV IRON: '00023501001' and '90135401001'. The following WHO-DRL code will be classified as RBC transfusion: '01186901001'.

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 started at End of Treatment visit will not be considered rescue for patients completed the treatment. 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

#### 6.1.2.4 Change from BL in SF-36 Vitality (VT) Sub-score to the Average VT Sub-score of Weeks 12 to 28

For details on the calculation of the SF-36 VT scale subscore, see Section 6.1.3.14.1.

All available SF-36 VT values will be used i.e., both scheduled and unscheduled SF-36 VT values. SF-36 PF values in analysis visit windows at weeks 12 and 28 will be selected for the calculation of the average VT sub-score of weeks 12 to 28 (see Table 32 and Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

For missing SF-36 VT, refer to Section 7.11.1 for the imputation rules. Baseline assessment is the assessment from Day 1 visit.

#### 6.1.2.5 Change from BL in SF-36 Physical Functioning (PF) Sub-score to the Average PF Sub-score of Weeks 12 to 28

For details on the SF-36 PF subscore, refer to Section 6.1.3.14.1. Similar rules as for Section 6.1.2.4 will be used for the calculation of the SF-36 VT scale sub-score.

#### 6.1.2.6 Change from BL in Mean Arterial Pressure (MAP) to the Average MAP Value of Weeks 20 to 28

Blood pressure will be measured singly for the three 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 will be calculated for each visit using the three readings. If less than three readings are available, all 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 during the Safety Emergent Period will be used, i.e. both scheduled and unscheduled MAP values. MAP 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 Table 31 and Section 7.11.4 for the analysis windows definition and the differentiation between the MMRM and ANCOVA analyses).

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 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

Occurrence of an increase in blood pressure is a binary variable (Yes/No), defined as:

- systolic blood pressure (SBP) increase from BL ≥20 mmHg AND SBP ≥170 mmHg, or
- diastolic blood pressure (DBP) increase from BL ≥15 mmHg AND DBP ≥110 mmHg.

The date of occurrence of hypertension is defined as the first date where SBP criterion or DBP criterion is met, whichever occurs first.

At each visit, SBP and DBP are calculated as the average from the 3 readings.. If less than three readings are available, the non-missing readings will be used in the calculation of the average.

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.

Only events starting during the Safety Emergent Period (defined in Section 6.2) will be taken into account.

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) / 365.25Where 'Analysis date of first dose intake' is defined in Section 6.5.4, and where 'First event date' is the first date of occurrence of hypertension.The time to censoring for a subject without the event of interest is calculated as: (Date of last vital signs assessment during the Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25Refer to Sections 6.2 and 7.11.5 for more details regarding the Safety Emergent Period.

#### 6.1.2.8 Rate of progression of CKD measured by annualized eGFR slope over time

Any eGFR values obtained from start of dialysis treatment (acute or chronic) will be excluded for the summaries and statistical analyses.

### 6.1.3 Additional Secondary Efficacy Endpoints

The additional secondary efficacy endpoints are listed in Table 5

**Table 5 Additional Secondary Efficacy Endpoints**

| Number | Endpoint |
|---|---|
| | **Hb correction and maintenance** |
| 1 | Hb level averaged over weeks 28 to 36, 44 to 52, and 96 to 104 without use of rescue therapy within 6 weeks prior to and during this evaluation period. |
| 2 | Time to achieve the first Hb response as defined by primary endpoint. |
| 3 | Hb change from BL to each post-dosing time point. |
| 4 | Hb change from BL to the average Hb value of weeks 28 to 36, 44 to 52, and 96 to 104 regardless of the use of rescue therapy. |
| 5 | 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, and 96 to 104 without use of rescue therapy within 6 weeks prior to and during these 8-week evaluation periods. |
| | **Hospitalizations** |
| 6 | Occurrence (number) of hospitalizations,number of days of hospitalization per patient-year exposure and time to first hospitalization. |
| | **Rescue Therapy Use** |
| 7 | Time to first use of rescue therapy (composite of RBC transfusions, ESA use, and IV iron) in the first 24 weeks of treatment. |
| 8 | Occurrence and time to first use of RBC transfusions, number of RBC packs per month subject, volume of RBC transfused per month. |
| 9 | Occurrence and time to first use of ESA. Number of ESA-Weeks per year |
| 10 | Occurrence and time to first use of IV iron supplementation. Mean monthly IV iron (mg) per subject during day 1 to week 36, weeks 37-52 and weeks 53-104 (monthly defined as a period of 4 weeks). |
| | **Change in Cholesterol Levels, Apolipoproteins** |
| 11 | 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. |
| 12 | Occurrence of mean LDL cholesterol <100 mg/dL calculated over weeks 12 to 28. |
| | **Blood Pressure Effect** |
| 13 | Occurrence of achieved antihypertensive treatment goal in CKD subjects (SBP< 130 mmHg and DBP< 80 mmHg) based on the mean SBP and mean DBP calculated over weeks 12 to 28. |

*Table continued on next page*

| Number | Endpoint |
|---|---|
| | **HRQoL** |
| 14 | Change from BL to the average value of weeks 12 to 28 (SF-36 Physical Component Score (PCS), Anemia Subscale ("Additional Concerns") of Functional Assessment of Cancer Therapy (FACT-An) Score, Total FACT-An Score, EQ-5D 5L VAS Score and Work Productivity and Activity Impairment (WPAI:ANS). |
| 15 | Patient Global Impression of Change (PGIC). |
| | **Hepcidin, Iron status, HbA1c, and CKD progression** |
| 16 | Changes from BL to each study visit (when measured) in Serum hepcidin, Serum ferritin, TSAT, HbA1c level, Fasting blood glucose, eGFR, Urine albumin/creatinine ratio, Time to (and proportion of subjects) Serum Creatinine having doubled during the study and Proportion of subjects with ESRD. |

#### 6.1.3.1 Hb level averaged over weeks 28 to 36, 44 to 52, and 96 to 104 without use of rescue therapy within 6 weeks prior to and during this evaluation period.

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 and 96-104 will be categorized into the following categories:

- <10.0 g/dL,
- 10.0-12.0 g/dL,
- >12.0 g/dL,
- ≥ 10.0 g/dL.

In case a subject does not have any available Hb value within this evaluation period, or in case a subject requires rescue therapy within 6 weeks prior to and during this 8-week evaluation period, refer to Section 7.11.1 for imputation rules

#### 6.1.3.2 Time to achieve the first Hb response as defined by 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 6.1.1.1

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 'First date of both values that meet the criteria for response' and 'Analysis date of first dose intake' is defined in Section 6.5.4

For a subject without Hb response or with rescue therapy before 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

#### 6.1.3.3 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 window, as defined in Table 31 and Section 7.11.4

Baseline Hb is defined in Section 6.1.1.1

At each visit, Hb will also be categorized into the following categories: <10 g/dL, 10-12 g/dL and >12 g/dL.

#### 6.1.3.4 Hb change from BL to the average Hb value of weeks 28 to 36, 44 to 52, and 96 to 104 regardless of the use of rescue therapy

The same rules as defined in Section 6.1.2.1, but regardless use of rescue therapy.

#### 6.1.3.5 Categorical analysis of Hb values

The following endpoints will be analyzed : proportion of Hb values within 10.0-12.0 g/dL and ≥10.0 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, ≥ 10.0 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 and (≥10.0 g/dL)/Total number of Hb values*100

in weeks 28 to 36, 44 to 52 and 96 to 104 without use of rescue therapy within 6 weeks prior to and during this 8 week evaluation period. All scheduled and unscheduled hemoglobin values that belong to each period will be taken into account using the analysis windows defined in Table 31, 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, ≥10.0 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:

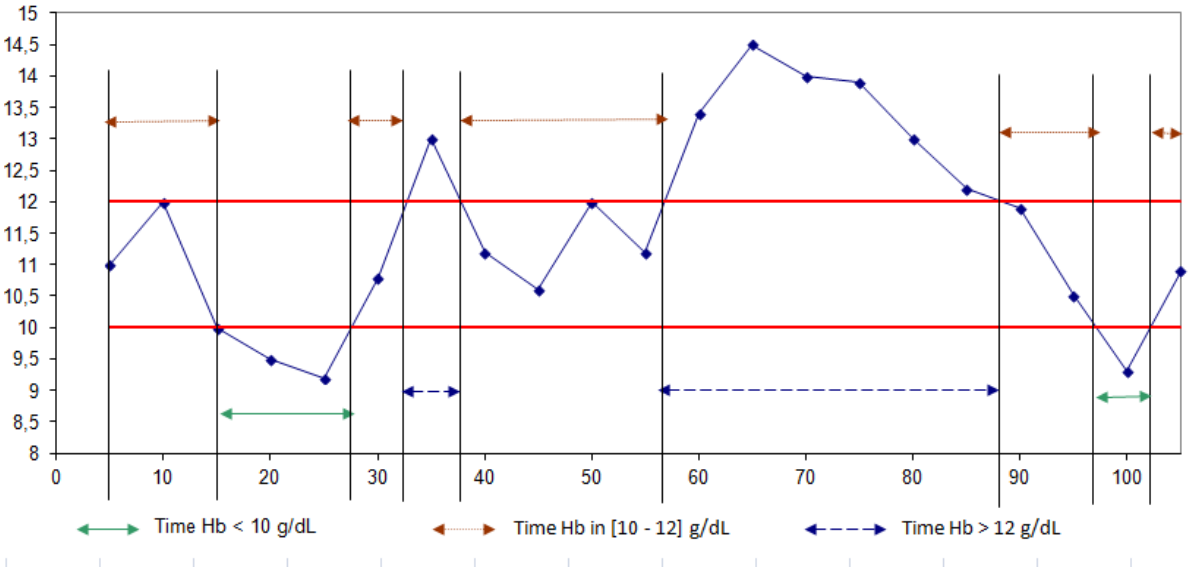
$$x = x_0 + (y - y_0)\frac{(x_1 - x_0)}{(y_1 - y_0)}$$

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

**Figure 1 Example of Linear Extrapolation**




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

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

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

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.

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 occurence 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:

(Date of last hemoglobin assessment during the 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.

#### 6.1.3.6 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 (Minimum ((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 Emergen 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 . Duration of treatment exposure is defined in Section 6.5.4

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

#### 6.1.3.7 Occurrence and time to first use of rescue therapy [composite of RBC transfusions, IV iron supplementation and ESA treatment] during the first 24 weeks

Rescue medication is defined in Section 6.1.2.3. Start of rescue medication is only counted as event if it was started within the first 24 weeks of treatment. Subjects without event were censored at the date of the Week 24 visit.

#### 6.1.3.8 Occurrence and time to first use of RBC transfusions, number of RBC packs per month, volume of RBC transfused per month

The blood transfusion form of the eCRF in the cumulative visit will be used to derive the number of RBC packs.

Monthly volume of blood transfused and the monthly total number of RBC units/packs (for each subject, the sum of blood volume and units transfused during the Efficacy Emergent Period / divided total number of days multiplied by 28 days) will be derived.

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

When transfused volume is not given but the number of RBC units is given, the volume will be estimated as number of RBC units times 250 mL (for transfusion of packed cell units) or number of 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

#### 6.1.3.9 Occurrence and time to first use of IV iron supplementation. Mean monthly IV iron (mg) per subject during day 1 to week 36, weeks 37-52 and weeks 53-104 (monthly defined as a period of 4 weeks)

The use of IV iron is collected in the *Concomitant Medication* form of the eCRF. The route of administration (Intravenous) is also captured in the eCRF. All medications are coded with WHO-DD. Records selected will be those coded as IRON PREPARATIONS, ATC 3$^{rd}$ level code: B03A and where route is INTRAVENOUS.

Having received IV Iron is a binary variable (Yes/No), where "Yes" is defined as having at least one record selected during the Efficacy Emergent Period. The Efficacy Emergent Period will be divided in periods of 28 days and for each of these periods, the monthly mean of IV iron will be calculated.

Only use of IV Iron that was ongoing or started during the Efficacy Emergent Period will be taken into account.

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

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

Monthly is defined as a period of 4 weeks.

Subjects without a relevant concomitant medication record will be assumed that they used no IV iron, thus set to 0 mg.

The same method will be used to calculate monthly IV iron use for week 37-52 and 53-104. Analysis visits will be used as indicated in Section 7.11.4.

For subjects with use of IV iron, the time to first use of IV iron is calculated as:

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

With 'First event date' defined as 'Date of first IV iron' during the Efficacy Emergent Period and 'Analysis date of first dose intake' defined in Section 6.5.4

For a subject without use of IV iron, the time to censoringis 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 ESA. Number of ESA-Weeks per year

The concomitant medication form of the eCRF in the cumulative visit will be used to detect rescue medication with ESAs as defined in Section 6.1.2.3

The number of ESA-weeks per year are defined as the duration of ESA exposure divided by the total efficacy period in years. Each period when ESA [ATC code = B03XA] was taken will be summed by subjects as follows:

[sum (each period (min(End date of efficacy emergent period, End date of ESA therapy) + X days – Start date of ESA therapy +1)/7 )]/total period in years;

X will be defined as the duration of the effect of ESA following the last ESA administration, based on the following rules:

If Frequency = 1 PER WEEK then X=7
If Frequency = 2 PER WEEK then X=3
If Frequency = 3 PER WEEK then X=2
If Frequency = 1 PER MONTH then X=28
If Frequency = ONCE then X=0
If Frequency = BIM (Bi-monthly) then X=14
If Frequency = QD (daily) then X=1

If Frequency = QM (monthly) then X=28
If Frequency =QOD (every other day) then X=2
If Frequency = TID (three times daily) then X=1
If Frequency = 4 TIMES PER WEEK then X=2
If Frequency = EVERY 3 WEEKS then X=21

Additional frequency may be considered depending on the data.

For subjects with use of ESA, the time to first use of ESA is calculated as:

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

With 'First event date' defined as 'Date of first ESA' and 'Analysis date of first dose intake' defined in Section 6.5.4

For a subject without use of ESA, 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

#### 6.1.3.11 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

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:

- Total cholesterol
- LDL cholesterol
- HDL cholesterol
- Low-density lipoprotein (LDL) / high-density lipoprotein (HDL) ratio
- Non-HDL cholesterol
- Triglyceride
- Apolipoproteins A1 and B (ApoA1 and ApoB)
- 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 Table 33 and 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.

#### 6.1.3.12 Occurrence of mean LDL cholesterol <100 mg/dL calculated over weeks 12 to 28

The evaluation period is defined as the average of all available LDL cholesterol values in weeks 12-28 (visit at 12, 20 and 28 weeks, as defined in Table 33 and Section 7.11.4). The occurrence of mean LDL cholesterol <100 mg/dL over weeks 12 to 28 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.

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

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

#### 6.1.3.13 Occurrence of achieved antihypertensive treatment goal in CKD subjects (SBP < 130 mmHg and DBP < 80 mmHg) based on the mean SBP and mean DBP calculated over weeks 12 to 28

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 to 28 during the Safety Emergent Period, similarly as in Section 6.1.2.6 (analysis windows defined in Table 31, 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.

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

#### 6.1.3.14 Change from BL to the average value of weeks 12 to 28 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 will be provided by an external vendor (QualityMetric).

##### 6.1.3.14.1 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 10.1). 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 10.1). 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 QualyMetric using QualityMetric Health Outcomes(tm) Scoring Software 4.5.

Change from baseline to the average value in weeks 12-28 will be calculated for the Physical Component Scores of SF-36 (SF-36 PCS), following the same rules as defined in Section 6.1.2.5

For missing SF-36 PCS values, refer to Section 7.11.1 imputation rules. 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 each visit 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.

#### 6.1.3.14.2 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. A final higher score indicates better QoL (see Appendix 10.2).

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.

If there are missing items, subscale scores can be standardized. 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]

When there are missing data, standardizing by subscale in this way is acceptable (Webster 2003) as long as more than 50% of the items were answered (e.g., a minimum of 4 of 7 items, 4 of 6 items, etc). The total score is then calculated as the sum of the un-weighted subscale scores. The FACT scale is considered to be an acceptable indicator of a subject's quality of life as long as overall item response rate is greater than 80% (e.g., at least 22 of 27 FACT-G items completed). This is not to be confused with individual subscale item response rate, which allows a subscale score to be prorated for missing items if greater than 50% of items are answered. In addition, a 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 10.2. The following 9 scores will be calculated:

- PWB subscale score
- SWB subscale score
- EWB subscale score
- FWB subscale score
- AnS subscale score
- FACT-An TOI score
- FACT-G total score
- FACT-An total score
- Fatigue subscale score

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

Change from baseline to the average value in weeks 12-28 will be calculated for the FACT-An subscore, following the same rules as defined in Section 6.1.2.4

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.

#### 6.1.3.14.3 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 10.3).

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 10.3

Change from baseline to the average value in weeks 12-28 will be calculated for the EQ 5D 5L VAS, following the same rules as defined in Section 6.1.2.4

Frequency distributions will be described for each visit for:

- EQ-5D 5L Mobility Score
- EQ-5D 5L Self-Care Score
- EQ-5D 5L Usual Activities Score
- EQ-5D 5L Pain/Discomfort Score
- EQ-5D 5L Anxiety/Depression Score

The evaluation period for the VAS score is defined as the average of available EQ-5D 5L VAS scores of weeks 12-28 (visit at W12 and W28).

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.

#### 6.1.3.14.4 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 10.4).

For subjects who are currently employed, the following four items 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

Change from baseline to the average value in weeks 12-28 and in weeks 36-52 will be calculated for these items, following the same rules as defined in Section 6.1.2.4. The evaluation period score is defined as the average of available WPAI: ANS subscore of weeks 12-28 (visit at W12 and W28) and weeks 36-52 (visit at W36 and W52).

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 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 10.5).

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.

#### 6.1.3.16 Hepcidin and Iron, HbA1c and CKD progression parameters

Changes from baseline to each study visit (see analysis windows in Table 31, Section 7.11.4) will be calculated for these parameters:

1. Serum hepcidin
2. Serum ferritin
3. Serum Iron
4. TSAT
5. HbA1c level
6. Fasting blood glucose
7. Serum creatinine (log transformed)
8. Albumin/creatinine ratio in urine (log transformed)

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

For all variables 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.

CKD Progression variables:

- Time to CKD progression (composite of doubling serum creatinine, chronic dialysis or renal transplant, and 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 renal 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

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

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

(End of 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

- Time to chronic dialysis or renal transplant or occurrence for a subject who died

Occurrence of chronic dialysis or renal 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

With 'First event date' defined as 'first occurrence of chronic dialysis or renal transplant, occurrence of subject who died (whichever occurring first)' and 'Analysis date of first dose intake' defined in Section 6.5.4.

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

(End of 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

- Time to doubling of serum creatinine

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

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

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

(End of 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

- Time to doubling of serum creatinine or chronic dialysis or renal 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 renal transplant during the Safety Emergent Period will be derived.

The endpoint is defined as time to doubling serum creatinine or chronic dialysis or renal transplant what ever comes first:

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

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

(End of 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

- Time to chronic dialysis or renal transplant:

  Occurrence of chronic dialysis or renal transplant (whichever occurring first) during the Safety Emergent Period (see Section 6.2) will be calculated and time to event in years will be defined as :

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

With 'First event date' defined as 'Date of dialysis or Date of renal transplant (whichever occurring first) and 'Analysis date of first dose intake' defined in Section 6.5.4.

For a subject without dialysis or transplant, the time to censoring will be calculated 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

- Occurrence of End Stage Renal Disease is taken directly from the eCRF with the following categories: f 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
- ESRD-free survival is defined as time form first dose alive and not progressed to ESRD. The time to event will be derived where event is death or ESRD whatever is first.

  The time to event will be derived (in years) as:

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

  where 'date of event' is defined as date of occurrence of End Stage Renal Disease (recorded End Stage Renal Disease as AE) or date of death, whichever comes first, during the Safety Emergent Period.

  For a subject wthout event, the time to censoring will be calculated (in years) as:

  (End of Safety Emergent Period – Analysis date of first dose intake + 1) / 365.25
- Proportion and Time to at least a 40% eGFR decrease from baseline, based on all eGFR values before start of acute or chronic dialysis

Time to at least 40% decrease in eGFR from baseline, chronic dialysis or renal transplant:

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

Occurrence of chronic dialysis or renal transplant (whichever occurring first) 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

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

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

(End of 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

- 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

With 'First event date' defined as 'First occurrence of 40% decrease in eGFR from baseline, and 'Analysis date of first dose intake' defined in Section 6.5.4

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

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

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

### 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 Table 31, 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, respiratory rate and weight),
- Clinical laboratory variables (hematology, biochemistry including liver enzymes and total Bilirubin, and urinalysis),
- Physical examination,
- 12-lead electrocardiogram (ECG).
- Vascular Access Thrombosis

The **Safety Emergent Period** will be defined as the evaluation period from the Analysis date of first drug intake up to 28 days after the Analysis Last Dose date (defined in Section 6.5.4). Refer to Section 7.11.5 for more details on the derivation of the date of End of the Safety Emergent Period. 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 .

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 a 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 deaths (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 deaths) 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 occuring 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 lastcontact (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.

#### 6.2.1.4 AE within 7 days

Additional analyses restricted to AEs that observed after starting administration of the test drug/comparative drug and up to Analysis Date of Last Dose + 7 days wil be defined. If the AE 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.

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

> (First event date of the given type occurring from Day 1 (post dose) up to Analysis date of minimum of (Last Dose + 7 days, end of safety period)) – Analysis date of first dose intake + 1) / 365.25

With 'Analysis date of first dose intake' defined in Section 6.5.4 and Analysis Date of Last Dose defined in Section 6.5.3

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, ***Max*** (EOS, Date of Death)] – Analysis date of first dose intake +1) / 365.25

#### 6.2.1.5 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.

#### 6.2.1.6 Definitions of event rate

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

$$\frac{\text{Number of events}}{\text{Patient Exposure Years}} \times 100$$

Or

$$\frac{\text{Number of subjects experienci ng an event}}{\text{Patient Exposure Years}} \times 100$$

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

The definition used will be indicated in the description of the corresponding analysis in Section 7

### 6.2.2 Vital Signs

The following endpoints will be assessed:

- Systolic blood pressure (SBP)
- Diastolic blood pressure (DBP)
- Pulse

In addition, the following assessments will be done:

- 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, the average in the eCRF system will not be used.

In case of missing values within a visit, the available readings will be used.

The position and date for the assessment will also be recorded. Change from baseline will be calculated as the measurement taken at the specific visit minus the measurement at baseline visit.

For all vital parameters, the minimum and the maximum post-baseline value during the Safety Emergent Period will be defined. For this calculation, only values from day 2 up to the date of End of the Safety Emergent Period (see Section 7.11.5).

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 missing visits, the last observation will be carried forward.

From the assessments collected in the 'Vital Signs - HD/HDF Subjects Only' form, only the pre-dialysis ones will be used for the classification of analysis visit (see Table 31, Section 7.11.4) and the definition of potentially clinical vital signs criteria.

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 6 (10 combined criteria).

**Table 6 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 |
| Systolic Blood Pressure (mmHg) | High | ≥ 170 | Increase of ≥ 20 |
| | Low | ≤ 90 | Decrease of ≥ 20 |
| Diastolic Blood Pressure (mmHg) | High | ≥ 110 | Increase of ≥ 15 |
| | Low | ≤ 45 | Decrease of ≥ 15 |
| Pulse (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 and on-treatment (at any moment during the Safety Emergent Period) using the worst value among all available measurements.

**Time to occurrence of PCS Vital signs:**

For each potentially clinically significant vital signs criteria (i.e., 10 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 intake' 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

Post-dialysis Vital Signs data will be listed and PCS criteria (observed value) only will be evaluated for them and flagged in the listing. No summary tables or Time to Event will be performed for them due to insufficient data.

### 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 7 below.

**Table 7 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<br>*> 5X ULN#*<br>*> 8X ULN#*<br>*> 10X ULN#*<br>*> 20X ULN#* |
| Aspartate Aminotransferase (AST) | U/L | no lower limit | > 3X ULN<br>*> 5X ULN#*<br>*> 8X ULN#*<br>*> 10X ULN#*<br>*> 20X ULN#* |
| Alkaline Phosphatase (ALP) | U/L | no lower limit | *> 1.5 X ULN#*<br>> 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<br>*> 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 LNL | >1.1 X ULN |
| Total Protein | g/L | < 0.9 X LNL | >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 |
| Triacylglycerol/Lipase | U/L | no lower limit | > 3X ULN or > 2 X Baseline |
| LLN: Lower limit of normal, value provided by the laboratory<br>ULN: Upper limit of normal, value provided by the laboratory | | | |

*Hemoglobin in SI unit (for conventional unit g/dL, divide by 10)

** 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 visitand 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 (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 the day 1 visit, except for Hb (see Section 6.1.1).

If Day 1 is missing, the screening or unscheduled assessment that is closest prior to Day 1 will be used.

Screening is defined as the screening or unscheduled assessment that is closest to Day 1.

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 Interval is not available, it will be replaced with 60/HR.

For all ECG parameters, the maximum post-baseline value on treatment will be defined (see Section 6.2).

ECG values are potentially clinically significant (PCS) if they meet or exceed the upper limit values listed in Table 8 below.

**Table 8 ECG Parameters classification**

| ECG Parameter | Classsification |
|---|---|
| QTc interval (msec) | > 450 msec, > 480 msec, > 500 msec; |
| QTc interval change (msec) | > 30 msec and > 60 msec<br>(increase from baseline) |
| 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

### 6.2.6 Vascular Access Thrombosis (VAT)

For each AE listed as VAT, additional information regarding the event will be recorded.

## 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 follows:

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 the three latest central laboratory Hb values prior to the day of randomization.

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 determined by performing a medical review of preferred terms recorded on any of the medical history forms. . History will be categorized as:

- Yes
- No

**Countries and Regions**

Subjects will be enrolled from the following countries:

Region A (Western Europe)

- Spain
- UK
- Italy
- Belgium

Region B (Rest of the World)

- Bulgaria
- Hungary
- Russia
- Romania
- Poland
- Turkey
- South Africa
- Serbia
- Ukraine
- Georgia

- Estonia
- Greece
- Panama
- Peru
- Guatemala
- Columbia
- Dominican Republic
- Belarus

Region B includes additional countries which were not part of the initial list of countries from Central and Eastern Europe as per IRT due to extended recruitment in Central and South Americas as well as Africa.

Randomization will be stratified by region using two categories:

- Region A: Western Europe
- Region B: Rest of the World

A limited amount of subjects is expected to be recruited out of Europe and will be randomized under the Central and Eastern Europe level of the Interactive Response Technology .

**Time to Treatment Discontinuation**

Time to Treatment Discontinuation in years is defined as:

Time to Treatment Discontinuation (years) = (‘Date of Treatment Discontinuation - (‘Analysis date of first dose intake +1)/365.25

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

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

With Analysis data of last dose defined in Section 6.5.4

**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, type2 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:

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

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. Checks will be performed by programming in order to ensure that derived eGFR values will match with eGFR provided by ICON for the selected visits.

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

**Iron Repletion 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 < 30 ng/mL or TSAT < 5%
- 30 ≤ ferritin < 100 ng/mL and 5% ≤ TSAT < 20%
- 30 ≤ ferritin < 100 ng/mL and TSAT ≥ 20%
- ferritin ≥ 100 ng/mL and 5% ≤ 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 Repletion 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 < 30 ng/mL or TSAT < 5%
- 30 ≤ ferritin < 100 ng/mL and 5% ≤ TSAT < 20%
- 30 ≤ ferritin < 100 ng/mL and TSAT ≥ 20%
- ferritin ≥ 100 ng/mL and 5% ≤ 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 or not ESAs during the last year. Previous use of ESAs is collected in the Treatment History for Anemia Form.

### 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 concomitant medication.

### 6.5.4 Variables related to study drugs

**Randomization/Treatment Arms**

Table 9 below presents the groups to which subjects are randomized under the initial protocol version 1.0.

**Table 9 Randomization arms under Protocol version 1.0**

| Randomization Arm Code (OARMCD) | Correction Period | Maintenance period | Randomization Arm (OARM) |
|---|---|---|---|
| 1A | TIW | QW | Roxadustat TIW/QW |
| 1P | TIW | QW | Placebo TIW/QW |
| 2A | TIW | BIW | Roxadustat TIW/BIW |
| 2P | TIW | BIW | Placebo TIW/BIW |
| 3A | TIW | TIW | Roxadustat TIW/TIW |
| 3P | TIW | TIW | Placebo TIW/TIW |

Table 10 below presents the groups to which subjects are randomized under protocol version 2.0 and later.

**Table 10 Treatment arms under protocol version 2.0**

| Randomization Arm Code (ARMCD) | Randomization Arm (ARM) |
|---|---|
| A | Roxadustat |
| B | Placebo |

For the statistical analysis, treatments will be pooled across placebo and 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 Extended Treatment visit in the End of Extended 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 Extended 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 (overall)**

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

('Analysis Date of Last Dose' - 'Analysis Date first dose intake') + 1

**Amount of Prescribed (planned) Medication**

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.

The prescribed daily and weekly dose at each visit (including unscheduled visits) will be calculated as follows:

Daily Prescribed dose in mg is Prescribed weekly Dose divided by 7

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. 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 11 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 Extended 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

**Extended treatment period only applies for subjects who do not discontinue prior to week 52.

This will allow us 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.

**Amount of Consumed Medication**

The consumed Investigational Product Medication is captured in the *Study Drug 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)

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.

**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 (not reported cumulatively) 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 Changes**

Dosing changes are collected in the Study Drug - Dosing Decisions eCRF.

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

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

For each subject the total number of dose changes will be calculated.

# 7 STATISTICAL METHODOLOGY

## 7.1 General Considerations

- All statistical comparisons will be made using two sided tests at the $\alpha$=0.05 significance level unless specifically stated otherwise. Null hypotheses for superiority testing will be of no treatment difference and corresponding alternative hypothesis will be two-sided. Null hypotheses for non-inferiority testing will be of inferiority of roxadustat treatment and will be one-sided at the $\alpha$=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 placebo) 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.

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 3. See Section 5.3 for more details.

### 7.2.1 Disposition of Subjects

The following subject data will be summarized and presented:

- Number and percentage 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 in each period (primary treatment period and extended treatment period), by primary reason for treatment discontinuation and 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 FAS and the SAF:

- 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 daweeks, 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

Subjects who screened failure 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, 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
    - PD3_1- Received wrong treatment kit
    - PD3_2- Received incorrect dose
- PD4 - Received excluded concomitant treatment

### 7.2.3 Demographic and Other Baseline Characteristics

Demographic and other baseline/screening characteristics 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 and SAF. 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/1.73$m^2$, and <10, 10-<15, 15-<30, 30-<45, 45-<60, and >=60 mL/min/1.73$m^2$
- History of previous treatment with ESA: Yes vs. No.

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

Selected demographics collected at screening will be done on Screen Failure subjects.

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.

Anemia history, CKD history, targeted medical history, cardiovascular history, tobacco history and family history of cardiovascular disease will be summarized.

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.

Demographic and baseline data will also be listed.

### 7.2.4 Previous and Concomitant Medications

Previous and concomitant 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.

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 (=Investigational Product Medicine) is for oral administration and supplied as red coated oval tablets of 20, 50 and 100 mg. Roxadustat and placebo tablets are identical in size, shape and color to preserve the blinding method.

### 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 Section 6.5.4

Descriptive statistics will be produced for:

- The amount of drug (roxadustat and placebo) the subject was exposed to during the treatment period (in mg and in mg/kg), by month, during the first 24, 36, 52 weeks and overall;
- Number and percentage of subjects with dose increases, decreases or interruptions during the treatment period.

Duration of exposure will be summarized in two ways:

- Descriptive statistics will be presented overall, during the first 24 weeks, 36 weeks and 52 weeks;
- Exposure time will be categorized overall, and by treatment 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.

Box-plots of monthly dose and monthly dose/kg by month will be produced by treatment arm (roxadustat and placebo).

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 placebo, 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 total study drug count and first and last days of treatment are known.

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

- 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 actual 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.

Quantitative endpoints with repeated measures over time will be analyzed using MMRM and ANCOVA with Multiple Imputations (MI) as the preferred methods for imputation of missing data, as per FDA recommendation at the time of SPA assessment of FGCL-4592-060.

For this section, PPS refers to the analysis set which excludes from the FAS all the subjects criteria listed in Table 3. See Section 5.3 and and Classification specifications for more details.

Analysis visits windows 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(s)

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

#### 7.4.1.1 EU (EMA) Primary Endpoint

##### 7.4.1.1.1 Primary Analysis of the EU (EMA) Primary Endpoint

The EU (EMA) primary efficacy endpoint will be analyzed using the FAS. See Section 7.11.1 for imputation rules, in case of missing values in the evaluation period or in case of rescue therapy prior to Hb response.

The proportion of responders in the primary efficacy variable will be compared using a Cochran–Mantel–Haenszel (CMH) test adjusting for the region, history of CV, baseline Hb and baseline eGFR, comparing roxadustat to placebo.

The EU (EMA) primary hypothesis to be tested for the primary efficacy analysis is:

- $H_0$: Hb responder rate in the roxadustat group = Hb responder rate in Placebo group

  versus

- $H_1$: Hb responder rate in the roxadustat group $\neq$ Hb responder rate in Placebo

The CMH adjusted odds ratio (roxadustat versus placebo) and its 95% confidence interval will be provided. Superiority of roxadustat versus placebo will be declared if the lower bound of the two-sided 95% confidence interval of the CMH odds ratio is higher than 1.

The SAS procedure will be similar to the following:

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

The covariates will be:

- Region [West Europe, Rest of the World))
- 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 $m^2$ )

The analysis will use the collected eCRF data to derive these covariates. Note that the choice was made to use Baseline Hb and eGFR as covariates rather than Screening Hb and eGFR which were stratification factors since Baseline values are the latest pre-dose assessements (and no difference between Screening and Baseline is expected).

In addition, a 95% confidence interval for the proportion of each roxadustat and Placebo based on the exact method of Clopper-Pearson will be calculated and presented.

The SAS procedure will be similar to the following:

```
proc freq;
tables Treatment / binomial(exact) alpha=0.05;
run;
```

**Model checking**

The assumption that the odds ratios are homogeneous across strata will be tested using the Breslow-Day test for stratified 2x2 tables. The model will be checked using covariates above defined by raw data.

In addition, this test will be repeated replacing the covariates by country. For this analysis, countries with less than 20 subjects in the FAS will be pooled in a grouped level per region.

This model checking information will be provided as part of the raw SAS outputs.

#### 7.4.1.1.2 Secondary Analyses (sensitivity) of the EU (EMA) Primary Endpoint

Table 12 summarizes all sensitivity analyses to be performed with the EU (EMA) primary endpoint.

**Table 12 Primary and sensitivity analyses for the EU (EMA) primary endpoint**

| Code | Set | Rescue Therapy | Endpoint | Method | Covariates |
|---|---|---|---|---|---|
| Primary | FAS | Without rescue therapy | Response | CMH | Region, history of CV, Baseline Hb and Baseline eGFR |
| S1 | All Randomized | Without rescue therapy | Response | CMH | Region, history of CV, Baseline Hb and Baseline eGFR |
| S2 | PPS | Without rescue therapy | Response | CMH | Region, history of CV, Baseline Hb and Baseline eGFR |
| S3 | FAS | Regardless rescue therapy | Response | CMH | Region, history of CV, Baseline Hb and Baseline eGFR |
| S4 | FAS | Without rescue therapy | Response | Logistic | Region, history of CV, Baseline Hb and Baseline eGFR as continuous covariates |
| S5 | FAS* | Without rescue therapy | Response | CMH | Region, history of CV, Baseline Hb and Baseline eGFR |

*patients who were randomized after the implementation of protocol v2.0

For the logistic regressions, the odds ratio (roxadustat vs. placebo) and their 95% confidence intervals will be produced, if convergence achieved.

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;
```

For all sensitivity analyses, no hypothesis testing will be done, only confidence intervals presented and graphically represented in a forest plot.

If a relevant baseline variable is identified for which a clinically important imbalance exists at baseline between treatment groups, additional sensitivity analysis on the primary endpoint may be performed using a logistic model adjusting for this baseline variable. This will allow the assessment of the impact of these imbalances on the treatment comparisons.

#### 7.4.1.1.3 Additional Analyses of the EU (EMA) Primary Endpoint

The analysis of the EU (EMA) primary endpoint will be repeated by subgroup of interest. For definitions of subgroups of interest, see Section 7.8

For the EU (EMA) primary endpoint, subgroup analyses will be performed using the primary analysis in Table 12 If one of the subgroups is the same as a stratification factor, the factor will be omitted from the model.

**Table 13 Additional Analyses of the EU (EMA) Primary Endpoint**

| Code | Set | Rescue Therapy | Endpoint* | Method | Covariates |
|---|---|---|---|---|---|
| A1 | FAS | Without rescue therapy | Response by Subgroup | CMH | 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 category and (8) Baseline Iron Repletion Status category

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 CMH adjusted odds ratios and their 95% confidence interval on the x-axis.

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

### 7.4.1.2 US (FDA) Primary Endpoint

#### 7.4.1.2.1 Primary Analysis of the US (FDA) Primary Endpoint

The US (FDA) primary efficacy endpoint will be analyzed using the All Randomized Set

The change from baseline to the average Hb of weeks 28 to 52 will be computed from an analysis of covariance (ANCOVA) model with multiple imputations (MI), adjusting for covariates (covariates defined below), comparing roxadustat to placebo.

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

> $H_0$: Hb mean change from baseline to the average level from Week 28 to Week 52 in the roxadustat group = Hb mean change from baseline in the placebo group
>
> versus:
>
> $H_1$: Hb mean change from baseline to the average level of Week 28 to Week 52 in the roxadustat group $\neq$ Hb mean change from baseline in the placebo group.

Difference of least square means (roxadustat minus placebo) and its 100x(1-α/2)% confidence interval will be estimated for the change from baseline to the average of weeks 28 to 52. Superiority of roxadustat versus placebo will be considered successful if the lower bound of the two-sided 95% confidence interval of the difference between treatment arms (roxadustat minus placebo) is higher than 0.

The covariates will be:

- Region (West Europe, Ex-Central East Europe)
- Baseline Hb values ( continuous)
- History of cardiovascular, cerebrovascular or thromboembolic diseases (Yes vs. No)
- Baseline eGFR (continous)

The analysis will not use the randomization strata as stratification errors during the randomization process may occur. The analysis will use the collected eCRF data to derive these covariates.

**ANCOVA with MI model:**

The MI ANCOVA model will be used to compare the roxadustat and placebo groups in a fixed sequence procedure:

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

1. Generate 1000 datasets, using seed 654289, where intermittent missing hemoglobin data will be imputed for each treatment relying on non-missing data from all subjects within each treatment group using the Monte Carlo Markov Chain MCMC imputation model with treatment, baseline hemoglobin, baseline eGFR, region, history of cardiovascular, cerebrovascular or thromboembolic diseases and the available non missing hemoglobin for each scheduled Week.
   The MCMC statement in the SAS PROC MI procedure with monotone option will be used. As a result, each dataset will only have missing ending data, or a monotone missing data pattern.
2. For each dataset from step 1, missing ending data (hemoglobin up through end of evaluation period) will be imputed using seed 472794. As a result, 1000 imputed complete datasets will be generated.
   - Missing data at Week 1 will be imputed using the regression imputation model with baseline stratification factor, baseline and hemoglobin from Week 1, using the SAS PROC MI procedure with the REGRESSION option in the MONOTONE statement.
   - The SAS PROC MI procedure will use data separately from each treatment subjects to impute the missing data for a specific Week (i.e. only those that need the imputation for the Week). Since subjects from the different treatment groups for that Week are excluded from the step, they will not contribute to the imputation for the Week.
   - Repeat for all other scheduled Weeks sequentially (Week 2 to the end of evaluation period). Subjects whose missing data were imputed for previous Weeks will contribute to the imputation for the current Week.

   The regression imputation model includes an intercept and the slopes of the hemoglobin from previous Weeks and the stratification factors.
3. Analyze each imputed dataset using the ANCOVA using the mean of all observed or imputed Hb values within the evaluation period. The model will contain terms for baseline Hb measurement as a covariate and treatment arm and the other randomization stratification factors except screening Hb (≤8 vs >8) as fixed effects.

4. Combine estimates from the results of each of the 1000 ANCOVA runs using SAS PROC MIANALYZE.

Report the results of the least-squares mean estimates of the change from baseline in hemoglobin during the evaluation period, the estimates of treatment effect (e.g., least-squares mean change form baseline in hemoglobin for the treatment group minus the least-squares mean change from baseline in hemoglobin for the placebo group) and the corresponding p-values and 95% CIs during the evaluation period.

All available Hb values will be used for the calculation of the average in weeks 28 to 52, as defined in analysis windows in Section 7.11.4.

Hb imputation rules using MI are detailed in Section 7.11.1.

A forest plot will be generated showing strata on the y-axis and differences in mean changes from BL to the average in week 28-52 and their 95% confidence interval on the x-axis.

**Model checking:**

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline Hb in the x-axis.

Different dot styles will be used for the two treatment arms. Solid black symbols (square) will be used for roxadustat and non-solid red symbols (X) will be used for placebo.

Residual plots will be done for the ANCOVA with MI analyses (Primary, S1 and S2 in Table 14).

In addition, an empirical cumulative distribution function of the residuals will be plotted for the same ANCOVA analysis.

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.

**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 28-52 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.

#### 7.4.1.2.2 Sensitivity Analyses of the US (FDA) Primary Endpoint

Table 14 summarizes all sensitivity analyses to be performed with the US (FDA) primary endpoint.

**Table 14 Primary and sensitivity analysis for the US (FDA) primary endpoint**

| Code | Set | Rescue Therapy | Endpoint | Method | Covariates |
|---|---|---|---|---|---|
| Primary | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | ANCOVA with MI | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S1 | FAS | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | ANCOVA with MI | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S2 | PPS | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | ANCOVA with MI | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S3 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | MMRM | Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates. |
| S4 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | PMM (Last Mean Carried Forward) | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S5 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | PMM (Baseline Carried Forward, Roxadustat only) | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S6 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | PMM (Baseline Carried Forward, both groups) | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S7 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 | PMM (Jump to control) | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S8 | All Randomized | Without rescue therapy | Change to the Average Hb in weeks 28-52 | ANCOVA with MI | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| S9 | All Randomized | Without rescue therapy | Change to the Average Hb in weeks 28-52 | MMRM | Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates. |

For all sensitivity analyses, no hypothesis testing will be done, only confidence intervals will be presented and graphically represented in a forest plot.

**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 52 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 52) 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.

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 $n$,
- $\beta_n$ 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 $n$,
- $\tau_i$ , is the mean effect of treatment arm $i$,
- $\alpha_n$ is the mean effect at time $n$,
- $(\alpha\tau)_{in}$ is the interaction term between treatment arm $i$ and time $n$,
- $\gamma_k$ is the mean effect of stratum $k$,
- $\varepsilon_{ikjn}$ 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 v.s. Placebo at weeks 28-52'
     treatment 1 -1
     treatment*visit 0 0 0 0 .. 0.2 0.2 0.2 0.2 0.2 -0.2 -0.2 -0.2
 -0.2  / cl;
 where visit in
('Day1','Week1','Week2',...,'Week8','Week10',...,'Week52');
run;
```

One analysis Hb value for each visit will be used, as defined in analysis windows in Section 7.11.4 and Table 31.

A forest plot will be generated showing strata on the y-axis and differences in mean changes from BL to the average in weeks 28-52 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.

**Model checking:**

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline Hb in the x-axis.

Different dot styles will be used for the two treatment arms. Solid black symbols (square) will be used for roxadustat and non-solid red symbols (X) will be used for placebo.

Residual plots will be done for the MMRM analysis. The plot will be repeated by visit at Weeks 28, 32, 36, 40, 44, 48 and 52.

In addition, an empirical cumulative distribution function of the residuals will be plotted for the MMRM analysis.

The model will also be checked using the stratification factors used for the randomization [IRS].

**Pattern Mixture Model (PMM)**

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

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

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

**A. Timing and Extent of Missing Data**

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

- Full data cases are defined as subjects with non-missing hemoglobin for all scheduled weeks of the Treatment period.
- Missing data cases are defined as subjects with a missing hemoglobin on at least one scheduled Week of the treatment period. The missing data cases are further grouped into intermittent missing and monotone missing cases.
    - Intermittent missing hemoglobin cases are defined as subjects with a missing hemoglobin for at least one scheduled week of but not on consecutive scheduled weeks up to end of the evaluation period.
    - Monotone missing hemoglobin cases are defined as subjects who have consecutive scheduled Weeks with missing hemoglobin up to the end of

evaluation period. A subject who is a Monotone missing case could have intermittent missing hemoglobin prior to the ending Week.

Subjects will be grouped as follows:

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

Should the incidence of Monotone missing data cases and intermittent missing data cases be relatively small, then those cases will be combined so that the groups are full data cases and missing data cases. The summary of missing pattern in first 52 scheduled visits will be presented in a table/graph.

**B. Assumptions on Missing Data Mechanism**

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

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

**C. Last Mean Carried Forward**

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

**D. PMM –Baseline Carried Forward (Roxadustat only and both groups)**

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

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

**E. PMM – Jump to Control**

A pattern-mixture model using a Prior Visit Mean carried forward jump to control multiple imputation method (Carpenter et al, 2013) will also be used as another sensitivity analysis similar to the PMM- Prior Visit Mean carried forward except for step 3, where the joint distribution of the patient's observed and missing data are considered to be multivariate normal with mean and covariance matrix from the control (placebo) treatment arm.

#### 7.4.1.2.3 Additional Analyses of the US (FDA) Primary Endpoint

The analysis of the US (FDA) primary endpoint will be repeated by subgroup of interest. For definitions of subgroups of interest, see Section 7.8

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

In addition, a sensitivity analysis will be performed by adding to the MMRM model all Hb values that will be collected even after the end of efficacy emergent period up to EOT (EOT + 2 weeks assessments should not be included), See analysis A3 in Table 15 below.

**Table 15 Additional Analyses of the US (FDA) Primary Endpoint**

| Code | Set | Rescue Therapy | Endpoint* | Method | Covariates |
|---|---|---|---|---|---|
| A1 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 by Subgroup | ANCOVA with MI | Region, History of CV, BL Hb, BL eGFR as continuous covariates |
| A2 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 by Subgroup | MMRM | Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and BL eGFR as continuous covariates. |
| A3 | All Randomized | Regardless rescue therapy | Change to the Average Hb in weeks 28-52 including all Hb values | MMRM | Region, History of CV, visits and visits by treatment as categorical variables. BL Hb and 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 category [only for A1] and (8) Baseline Iron Repletion Status category [only for A1]

Subgroup analyses 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 change from baseline 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 the difference between the treatment arms: roxadustat versus placebo.

Once the primary hypothesis has been rejected for the EU (EMA) primary endpoint, the key secondary endpoints 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. If the null hypothesis is rejected for a test, the claim of superiority (or non-inferiority) will be

considered successful and the test will progress to the next comparison in sequence as follows:

**Table 16 Key Secondary Endpoints fixed sequence testing procedure**

| Test | Analysis set | Endpoint | Comparison |
|---|---|---|---|
| 1 | FAS | Hb change from BL to the average Hb in weeks 28-36, without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period. | Superiority of roxadustat versus placebo |
| 2 | FAS | Change from BL in Low Density Lipoprotein (LDL) cholesterol to the average LDL cholesterol of weeks 12 to 28. | Superiority of roxadustat versus placebo |
| 3 | FAS | Use and time to first use of rescue therapy during the treatment period | Superiority of roxadustat versus placebo |
| 4 | FAS | Change from BL in SF-36 Vitality (VT) sub-score to the average VT sub-score of weeks 12 to 28 for all subjects | Superiority of roxadustat versus placebo |
| 5 | FAS | Change from BL in SF-36 Physical Functioning (PF) sub-score to the average PF sub-score of weeks 12 to 28 for all subjects | Superiority of roxadustat versus placebo |
| 6 * | PPS | Change from BL in mean arterial pressure (MAP) to the average MAP value of weeks 20 to 28. | *: Non-inferiority of roxadustat versus placebo (the non-inferiority margin for the difference between groups is fixed as 2 mmHg) |
| 7 * | PPS | Time to first occurrence of hypertension (defined as an increase from BL of ≥ 20 mmHg systolic blood pressure (SBP) and SBP ≥ 170 mmHg or an increase from baseline of ≥15 mmHg diastolic blood pressure (DBP) and DBP ≥ 110 mmHg | *: Non-inferiority of roxadustat versus placebo (the non-inferiority margin for the difference between groups is fixed as a hazard ratio of 1.3) |
| 8 * | FAS | Rate of progression of CKD measured by annualized eGFR slope over time | *: Superiority of roxadustat versus placebo |

*: These key secondary endpoints will not be included in the hierarchical testing procedure.

No multiplicity adjustment will be done for the last three endpoints. Details of the analysis for each of these secondary endpoints are given below.

#### 7.4.2.1 Hb change from baseline to the average Hb in weeks 28-36, without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period

Hb change from baseline to the average Hb value in weeks 28-36 will be compared by treatment arms using a MMRM model as in Section 7.4.1.2

The analysis will be similar to the analysis provided in Section 7.4.1.2, except that evaluation period will be from baseline to weeks 28-36, without having received rescue therapy within 6 weeks prior to and during this 8-week evaluation period.

The analysis will be done on the FAS.

This superiority test will be considered successful if the lower bound of the two-sided 95% confidence interval of the difference between treatment arms (roxadustat minus placebo) is higher than 0.

An additional analysis (not part of the sequence) will be done for this endpoint on the FAS regardless the use of rescue therapy,as detailed in Table 17:

**Table 17 Primary and sensitivity analysis for the Hb change form BL to the average Hb in weeks 28-36**

| Code | Set | Rescue Therapy | Endpoint | Method | Covariates |
|---|---|---|---|---|---|
| Primary | 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. |
| S1 | 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. |

#### 7.4.2.2 Change from BL in Low Density Lipoprotein (LDL) cholesterol to the average LDL cholesterol of weeks 12 to 28

LDL change from baseline to the average LDL value in weeks 12-28 will be compared by treatment arms using a MMRM model as in Section 7.4.1.2 (with the addition of LDL at baseline as continuous covariate). The analysis will be done on the FAS.

The analysis will be similar to the analysis provided in Section 7.4.1.2. This superiority test will be considered successful if the upper bound of the two-sided 95% confidence interval of the difference between treatment arms (roxadustat minus placebo) is below 0.

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

Theses analyses 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 for fasted values. The average of weeks 12-28 will also be reported.

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline LDL-cholesterol in the x-axis.

#### 7.4.2.3 Use and time to first use rescue therapy during the treatment period [composite of RBC transfusions, IV iron supplementation and rescue ESA]

Time to first use of rescue therapy will also be analyzed similarly as in Section 7.4.2.7 except that only superiority will be tested and 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.

In addition, the number and percentage of subjects with rescue therapy during the Efficacy Emergent Period will be reported by treatment group.

**Table 19 Primary and sensitivity analysis for use and time to first use rescue therapy**

| Code | Set | Endpoint | Method | Covariates |
|---|---|---|---|---|
| Primary | FAS | Time to first use rescue therapy | Cox regression + Kaplan Meier | Stratified on Region, History of CV and adjusted on BL Hb, BL eGFR as continuous covariates |
| S1 | All Randomized | Time to first use rescue therapy | Cox regression + Kaplan Meier | Stratified on Region, History of CV and adjusted on BL Hb, BL eGFR as continuous covariates |
| S2 | FAS | time to first use rescue therapy | Cox regression + survival curve (Breslow estimator) using IPCW method | Stratified on Region, History of CV and adjusted on BL Hb, BL eGFR as continuous covariates |

For the IPCW method, refer to Appendix 10.7 for details. If the conclusions of the analysis using stratified cox model and IPCW method are not concordant, additional analysis using IPCW method will performed for each of the endpoints separately.

- Time to first use of RBC transfusion
- Time to first use of IV iron supplementation
- Time to first use of ESA

Details of these endpoints are provided in Sections 7.4.3.8, 7.4.3.9 and 7.4.3.10

#### 7.4.2.4 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 of weeks 12–28 will be compared by treatment arm for all subjects (primary analysis) and in the subsets of subjects with baseline vitality subscore below 50 and equal or above to 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.

The analysis will be similar to the analysis provided in Section 7.4.1.2. Superiority will be declared if the lower bound of the two-sided 95% confidence interval of the difference between roxadustat and placebo is above 0.

The analysis will be done on the FAS.

**Table 20 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 | FAS | SF36-VT Change from Baseline at Weeks 12-28 for all subjects | 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 for subjects with BL VT subscore below 50 | 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. |
| S2 | FAS | SF36-VT Change from Baseline at Weeks 12-28 for subjects with BL VT subscore equal or above 50 | 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. |

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline SF-36 VT subscore in the x-axis.

In addition to the inferential analyses, SF-36 VT and SF-36 VT 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 BL in SF-36 Physical Functioning (PF) sub-score to the average PF sub-score of weeks 12 to 28

Change from baseline in PF subscore of SF-36 to the average of weeks 12–28 will be compared by treatment arm for all subjects (primary analysis) 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.

The analysis will be similar to the analysis provided in Section 7.4.1.2. Superiority will be declared if the lower bound of the two-sided 95% confidence interval of the difference between roxadustat and placebo is above 0.

The analysis will be done on the FAS.

**Table 21 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 for all subjects**

| Code | Set | Endpoint | Method | Covariates |
|---|---|---|---|---|
| Primary | FAS | SF36-PF Change from Baseline at Weeks 12-28 for all subjects | 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 for subjects with BL PF subscore below 35 | 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. |
| S2 | FAS | SF36-PF Change from Baseline at Weeks 12-28 for subjects with BL PF subscore equal or above 35 | 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. |

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline SF-36 PF subscore in the x-axis.

In addition to the inferential analyses, 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.

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

The blood pressure effect is shown by use of the change from baseline in MAP.

MAP change from baseline to the average MAP value in weeks 20-28 will be analyzed using a MMRM model as in Section 7.4.1.2.1 (with the addition of MAP at baseline as continous covariate) . The analysis will be done on the PPS.

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 roxadustat and placebo (roxadustat minus placebo), calculated on the PPS is below 2 mm Hg.

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

**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. |
| S1 | All Randomized | 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 analyses, 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.

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline MAP in the x-axis.

#### 7.4.2.7 Time to first occurrence of hypertension (defined as either SBP ≥170 mmHg AND an increase from BL ≥ 20 mmHg or as DBP ≥ 110 mmHg, AND an increase from BL ≥ 15 mmHg)

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 of the hazard ratio (roxadustat as relative to placebo), calculated on the PPS, is below 1.3. As per Table 16, once the null hypothesis is rejected, superiority will be checked for this variable on the FAS, as part of the fixed sequence testing procedure. Superiority will be concluded if the upper bound of the two-sided 95% CI of the hazard ratio of the two treatment arms (roxadustat as relative to placebo) is below 1.

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 placebo

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;
```

The SAS procedure for the Cox regression using IPCW method uses above similar code with weight statement. Refer to Appendix 10.7 for details.

**Table 23 Primary and sensitivity analysis for the time to first occurrence of hypertension**

| Code | Set | Endpoint | Method | Stratas/Covariates |
|---|---|---|---|---|
| Primary | PPS | time to first occurrence of hypertension | Cox regression + Kaplan Meier | Stratified on Region and History of CV, and adjusted on BL Hb, BL eGFR as continuous covariates |
| S1 | All Randomized | time to first occurrence of hypertension | Cox regression + Kaplan Meier | Stratified on Region, History of CV and adjusted on BL Hb, BL eGFR as continuous covariates |
| S2 | FAS | time to first occurrence of hypertension | 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 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 risk.

In addition to the cumulative incidence plot, the cumulative incidence at 3, 6 and 9 months 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 an increase from baseline in blood pressure during the Safety Emergen Period will be reported by treatment arm (roxadustat and placebo) on the PPS. For subjects who have experienced more than one increase in blood pressure, 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 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.

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.

#### 7.4.2.8 Rate of progression of CKD measured by annualized eGFR slope over time

Geometric Mean (GM) and coefficient of variation (CV) will be displayed for eGFR. In addition, change from baseline and 95% CI will be presented.

The annualized eGFR slope over time (expressed in ml/min per 1.73 m² or % 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 &mean egfrbl
          atptn*TRTPN*TRTP 1 0
          atptn*strata_except_egfr_Hb &prop_strata
        /cl;
      estimate 'Annnualized slope Placebo'
          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.

Superiority will be declared if the lower bound of the two-sided 95% confidence interval of the difference in Least Square Means between roxadustat and placebo is above 0.

**Model checking:**

Residual plots will be produced showing the following:

- Model residuals in the y-axis and model predictions in the x-axis,
- Model residuals in the y-axis and baseline Hb in the x-axis.

Different dot styles will be used for the two treatment arms. Solid black symbols (square) will be used for roxadustat and non-solid red symbols (X) will be used for placebo.

If the residual diagnostics show heteroscedasticity, additional analysis will be performed where annualized eGFR slope over time (% per year) is estimated by a random slopes and intercepts model using all available (log-transformed) 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 log(eGFR) by timepoint and Baseline Hb by timepoint).

Least Square Means will be transformed back to the original scale and expressed as annual percent changes.

#### 7.4.2.9 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 | FAS | 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 | SF36-PF Change from Baseline at Weeks 12-28 for all subjects 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. |

*Table continued on next page*

| Code | Set | Endpoint | Method | Covariates |
|---|---|---|---|---|
| A4 | FAS | SF36-VT Change from Baseline at Week 28 for all subjects 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. |
| A5 | 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. |
| A6 | PPS | time to first occurrence of hypertension 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 category [only for A1] and (8) Baseline Iron Repletion Status category [only for A1]

Forest plots will be produced where all subgroup factors will appear in the y-axis and the appropriate statistic comparing roxadustat to placebo and the 95% confidence interval will appear in the x-axis.

### 7.4.3 Analysis of Additional Secondary Efficacy 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 placebo.

Descriptive statistics will be presented by treatment arm.

#### 7.4.3.1 Hb level averaged over weeks 28 to 36, 44 to 52, and 96 to 104 without use of rescue therapy within 6 weeks prior to and during these 8-week evaluation periods

Averaged Hb over weeks 28-36, 44 to 52 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 and 96-104 and by visit within the <10 g/dL, 10-12 g/dL and >12 g/dL categories will be reported.

#### 7.4.3.2 Time to achieve the first Hb response as defined by primary endpoint for EU (EMA)

Time to achieve the first Hb response, without rescue therapy, will be analyzed using the same methods described in Section 7.4.2.7. on the FAS.

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, and 24 weeks with their associated 95% confidence interval reported using Greenwood's formula.

#### 7.4.3.3 Hb change from BL to each post-dosing time point

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

#### 7.4.3.4 Hb change from BL to the average Hb value of weeks 28 to 36, 44 to 52, and 96 to 104 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 28-36, 44-52 and 96-104.

#### 7.4.3.5 Categorical analysis for 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 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, > 12.0 g/dL, ≥10.0 g/dL, > 13.0 g/dL and > 14.0 g/dL) between the first and last Hb assessement during the Efficacy Emergent Period will be presented by treatment arm.

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)

#### 7.4.3.6 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 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).

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. Both descriptive summary and time to first hospitalization will be reported by eGFR category (<15 vs ≥15).

As a sensitivity analysis, the above analysis will be repeated using IPCW method (Appendix 10.7).

The analysis will be done on the FAS.

#### 7.4.3.7 Occurrence and time to first use of rescue therapy during the First 24 Weeks [composite of RBC transfusions, IV iron supplementation and rescue ESA]

The number and percentage of subjects with rescue therapy during the first 24 weeks 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.

#### 7.4.3.8 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 will be reported. Descriptive statistics by treatment arm for number of RBC packs and volume of blood transfused will be reported. Time to first use of RBC transfusion will be analyzed during the Efficacy Emergent Period similarly as in Section 7.4.2.7

Mean monthly number of RBC packs and volume of RBC transfused during the Efficacy Emergent Period will be compared by treatment arm using a ANCOVA model as in Section 7.4.1.2.1 (except that no multiple imputation will be performed). Subjects with no medication record of RBC will be assumed that they received no RBC and therefore number of packs and volume will be set to 0. The number and percentage of subjects with number of RBC Packs/Volume >0 will be presented.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.

#### 7.4.3.9 Occurrence and time to first use of IV iron supplementation. Mean monthly IV iron (mg) per subject during day 1 to week 36, weeks 37-52 and weeks 53-104 (monthly defined as a period of 4 weeks)

The number and percentage of subjects who received IV Iron and the total amount of IV Iron used during the Efficacy Emergent Period will be reported descriptively by treatment arm.

The incidence and cumulative incidence of subjects receiving IV Iron will be calculated similarly as in Section 7.4.2.7

Mean monthly IV iron (mg) per subject during day 1 to week 36, weeks 37-52 and weeks 53-104 will be summarized by treatment arm. Subjects with no medication record of IV Iron will be assumed that they received no IV Iron (for patients under treatment during that month).

The analysis will be done on the FAS.

Furthermore, the total amount of IV Iron used during the efficacy emergent period as well as for Day 1 to week 36, weeks 37 to 52 and weeks 53 to 104.

Time to first use of IV iron will also be analyzed similarly as in Section 7.4.2.7

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.

#### 7.4.3.10 Occurrence and time to first use of ESA. Number of ESA-Week per year

The number and percentage of subjects with ESA use as rescue therapy, and the number of ESA-weeks/year will be reported descriptively by treatment arm.

Time to first use of ESA 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 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 and treatment arm. Descriptive statistics will also be reported regardless of fasting status.

#### 7.4.3.12 Occurrence of mean LDL cholesterol <100 mg/dL calculated over weeks 12 to 28

The number and percentage of subjects with mean LDL cholesterol less than 100 mg/dL (fasting values and regardless fasting status) on average in weeks 12-28 will be reported by treatment arm and by baseline value.

#### 7.4.3.13 Occurrence of achieved antihypertensive treatment goal in CKD subjects (SBP< 130 mmHg and DBP< 80 mmHg) based on the mean SBP and mean DBP calculated over weeks 12 to 28.

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 will be reported by treatment arm.

#### 7.4.3.14 Health related Quality of Life Questionnaires Change from BL to the average value of weeks 12 to 28

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 (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, an inferential analysis, similar to the one defined in Section 7.4.2.1 will be performed for the following endpoints:

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

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

**EQ-5D 5L**

For the EQ-5D 5L VAS score, change from baseline to each visit and to the average of weeks 12-28 will be described by treatment arm.

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 descriptively by visit and treatment arm.

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

#### 7.4.3.15 Patients' Global Impression of Change (PGIC)

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

#### 7.4.3.16 Hepcidin and Iron, HbA1c and CKD progression parameters

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

- Serum hepcidin
- Serum ferritin
- TSAT
- HbA1c level
- Fasting blood glucose
- 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.

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

All time to event endpoints as defined in Section 6.1.3.16 will be will be analyzed using the same methods described in Section 7.4.2.7 on the FAS. Superiority will be declared if the lower bound of the two-sided 95% confidence interval of the hazard ratio is above 1.

As a sensitivity analysis, the above stratified cox model will be repeated using IPCW method (Appendix 10.7) for the following endpoints:

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

  If the conclusions of the analysis using stratified cox model and IPCW method are not concordant, additional analysis using IPCW method will performed for each of the following endpoints separately.

  - Time to doubling of serum creatinine
  - Time to chronic dialysis or renal transplant
  - Time to occurrence for a subject who died or chronic dialysis or renal transplant

- Time to at least 40% decrease in eGFR from baseline, chronic dialysis or renal transplant

  If the conclusions of the analysis using stratified cox model and IPCW method are not concordant, additional analysis using IPCW method will performed for each of the following endpoint separately.

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

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 by treatment arm.

Genotyping will be shipped to a delegated CRO and analyzed under the responsibility of Bioanalysis-Europe of Astellas Pharma Europe B.V. A separate report will be provided.

## 7.5 Analysis of Safety

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

Missing dates' imputation rules for AE onset date 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 causally 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 causally drug related TEAEs leading to permanent discontinuation of study drug,
- Number and percentage of subjects with NCI CTC Grade 3 or higher TEAEs,
- Number and percentage of subjects with TEAEs leading to death
- Number and percentage of subjects with drug related 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 (by PT only)
- Drug related TEAEs,
- TEAEs with NCI CTC Grade 3 or higher
- TEAEs by severity

- Drug related TEAEs by severity
- Serious TEAEs,
- Drug related serious TEAEs,
- TEAEs leading to permanent discontinuation of study drug,
- Drug related TEAEs leading to permanent discontinuation of study drug,
- TEAEs excluding serious adverse events that equal to or exceed a threshold of 5.0% in any treatment arm (roxadustat or placebo),
- TEAEs leading to death,
- Drug-related TEAEs leading to death,
- Common TEAEs that equal to or exceed a threshold of 5.0% in any treatment arm (roxadustat or placebo)
- TEAEs by relationship to the study drug,
- AEs occurring during the post study follow up period
- Serious AEs during the post-stiudy follow up period
- Onset of common (>=5% in Any treatment arm(roxadustat or placebo)) TEAEs by treatment duration: <3 months, ≥3 to ≤6 months, >6 to ≤12 months, >12 months

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 (see Section 6.2).

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
- TEAEs censored at start of chronic dialysis
- 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 treatment 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 of special interest:

- Serious TEAEs,
- TEAE for each SOC (>10% for a SOC in total SAF)
- 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 CV and adjusted on baseline Hb and baseline eGFR as continuous covariates. Hazard ratios (roxadustat as relative to placebo) and their 95% CI will be calculated for each TEAE category.

As a sensitivity analysis, the above analysis will be repeated using IPCW method (Appendix 10.7) for the following event types of special interest:

- Serious TEAEs,
- TEAE for each SOC (>10% for a SOC in total SAF)
- Deaths occurring during the Safety Emergent Period, and
- TEAE NCI CTC Grades 3 or higher.

A dot-and-forest plot will be produced showing each of the above TEAE categories on the y-axis. The incidence rates by treatment arm, the stratified Cox hazard ratios (roxadustat as relative to placebo) and their 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 within 7 days

Additional summaries or analyses with events to be considered restricted to any adverse event starting up to Analysis Date of Last Dose + 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, dot-and-forest plots 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, dot-and-forest plots and cumulative incidence plots as described in Section 7.5.1.4 will be provided)

Pre-specified adjudicated cardiovascular and thrombo-embolic events will be analyzed in meta-analyses across multiple phase 3 studies and compared between treatment groups (roxadustat versus control). The statistical method for this analysis pooling studies will be detailed in a Pooled Statistical Analysis Plan (pSAP). The results will be presented in a separate report.

All data will also be listed. All Adverse Events collected from site 70051 will be listed separately.

### 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 and treatment 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 Section 6.2.3.1, 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.

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 of PCS abnormalities 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 placebo. A dot-and-forest plot will be produced showing the PCS abnormalities above in the y-axis and the incidence rates, the Cox hazard ratio and their 95% CI in 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 placebo.

Individual displays of Liver Enzymes and Bilirubin parameters, listed in Section 6.2.3.1, will be reported for all subjects with ALT and / or AST > 3 x ULN or total bilirubin > 2 x ULN during Safety Emergent Period.

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 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).

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 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.

The number and percentage of subjects with post-baseline PCS values (see Table 8) will be tabulated by treatment group. 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<br>65 - 74 years<br>≥ 75 years |
| Sex | Female<br>Male |
| Region | Western Europe<br>Rest of the World |
| Baseline Hb | ≤ 8 g/dL<br>> 8 g/dL |
| History of cardiovascular, cerebrovascular or thromboembolic diseases | Yes<br>No |
| *Table continued on next page* | |

| Grouping variables | Subgroups |
|---|---|
| Baseline eGFR | < 30 mL/min/1.73m$^2$<br>≥ 30 mL/min/1.73m$^2$ |
| Baseline eGFR[1] | < 15 mL/min/1.73m$^2$<br>≥ 15 mL/min/1.73m$^2$ |
| Baseline CRP [2] | ≤ ULN<br>> ULN |
| Baseline Iron Repletion Status [2] | (TSAT >= 20% and ferritin >= 100 ng/mL) vs (TSAT < 20% or ferritin < 100 ng/mL) |
| [1]: *only for the two primary efficacy endpoints and also for selected AE summaries.* | |
| [2]: *only for the two primary efficacy endpoints and the key secondary endpoint on Hb (week 28- week 36)* | |

## 7.9 Other Analyses

The current analysis plan is based on version 2.0 of the protocol. An exploratory analysis on Hb maintenance will be conducted in the sub-set of patients treated for at least 24 weeks under protocol version 1.0. The sub-set includes patients that discontinued treatment prior to Week 24 under protocol version 1.0. The pooled roxadustat arm will be seperated into 3 groups according to their randomized treatment: (TIW, BIW, QW). Tables with summary statistics by visit up to Week 24 will be derived for the Hb values and the average weekly prescribed dose by the 4 treatment groups (placebo, roxadustat TIW, roxadustat BIW, roxadustat QW). In addition, the change from baseline to the Hb value average over Weeks 12-24 (excluding rescue medication) will be analysed using the MMRM model using Hb and eGFR at baseline as co-variates. In addition, average weekly dose will be summarized by treatment arm by 4-weekly period until subject switched to protocol version 2.0.

## 7.10 Interim Analysis (and Early Discontinuation of the Clinical Study)

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

- An ANCOVA model with multiple imputations (MI) will also be run.

No imputations 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<br>--/MM/----, or<br>--/--/---- | No imputation |

**Previous or Concomitant medication:**

For previous or concomitant medications, including rescue medications and chornic 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<br>--/MM/----, or<br>--/--/---- | 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<br>30/MM/YYYY, or<br>29/MM/YYYY, or<br>28/MM/YYYY |
| --/--/YYYY | 31/12/YYYY |
| DD/--/----, or<br>--/MM/----, or<br>--/--/---- | No imputation |

**AE Onset date**

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<br>--/02/2008<br>--/02/2008<br>--/02/2008 | DD/MM/YYYY<br>14/02/2008<br>14/02/2007<br>14/02/2009 | <br>14/02/2008*<br>01/02/2008<br>01/02/2008 |
| --/--/YYYY<br>--/--/2008<br>--/--/2008<br>--/--/2008 | DD/MM/YYYY<br>14/02/2008<br>14/02/2007<br>14/02/2009 | <br>14/02/2008<br>01/01/2008<br>01/01/2008 |
| DD/--/----<br>--/MM/----<br>--/--/---- | | 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<br>30/MM/YYYY or<br>29/MM/YYYY or<br>28/MM/YYYY |
| --/--/YYYY | 31/12/YYYY |
| DD/--/----, or<br>--/MM/----, or<br>--/--/---- | 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.

**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) or unscheduled (on-tretament) | 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 |
| 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 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 |

[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 |
| 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 |

*Table continued on next page*

| CRF Visit | Target Day[a] | Analysis Visit Windows Actual Assessment Day | Analysis Visit |
|---|---|---|---|
| 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 + 28 days , max(EOS Visit, Date of death))*** in case Analysis date of Last Dose = Date of Last Dose
- ***Minimum [(Analysis date of Last Dose + 28 days , max(EOS Visit, Date of death))*** in case Analysis date of Last Dose = Date of the End of Extended Treatment Visit (due to missing date of Last Dose) and the subject is treatment completer or discontinued due to any reason apart from Lost to follow up.
- ***Analysis date of Last dose*** in case Analysis date of Last Dose = Date of the End of Extended Treatment Visit (due to missing date of Last Dose) and the subject discontinued due to Lost to follow up.
- ***Analysis date of Last dose*** in case Analysis date of Last Dose = Date of last available assessment (due to missing date of last dose and date of the End of Extended Treatment Visit). In case a subject died, then ***Minimum [***Date of last available assessment ***+ 28 days , Date of death)]***.

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 :

- ***Minimum (Analysis date of Last Dose + 7 days , EOT Visit)***

# 8 DOCUMENT REVISION HISTORY

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| 1.0 | 28-May-2014 | NA | Document finalized |
| 2.0 | 21-January-2016 | 1) Reduction of the number of sensitivity analyses for the secondary endpoints.<br>2) New ordering for the key secondary endpoints.<br>3) Implementation of the Time to event approach for the PPS and adjustment of the relevant sections.<br>4) Change from ITT to All Randomized as per Astellas standard. Definition remains unchanged.<br>5) Use of the Safety Emergent Period (i.e last dose + 28 days) as evaluation period by default for the safety endpoints.<br>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.<br>7) Use of "Time to censoring" instead of"time at risk" for patient with no event for more clarity.<br>8) Clarification of time to censoring for events evaluated during the Safety Emergent Period/Efficacy Emergent | 1) Due to harmonization with Fibrogen 060 study SAP and to limit the additional analyses .<br>2) Due to the importance of the QoL endpoints and the harmonization with Fibrogen 060 study SAP.<br>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.<br>4) Astellas standards requirement.<br>5) Decision agreed by the study team to use a consistent approach for all the safety endpoints by extending the evaluated period up to 28 days after last dose, which matched the AE analyses.<br>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.<br>7) Clarification of the wording. Time to censoring is more appropriate.<br>8) Based on the new definitions of Efficacy/Safety Emergent period, time to censoring for all time to |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | Period. | events have been updated accordingly. |
| | | 9) Time to PCS vital signs added. | 9) Time to PCS were already planned for lab and ECG but not for Vital signs. Decision was taken to provide them as well. |
| | | 10) Time to PCS ECG focused to QTc instead of all ECG assessments. | 10) Time to event for PCS ECG other than QTcF was considered not of interest. |
| | | 11) Additional censoring rules for the post-dialysis data (used for some time to event related to AEs). | 11) In order to assess the impact of dialysis on incidence rates of AEs, sensitivity analyses will be conducted by censoring patients at the date of dialysis. |
| | | 12) Update regarding derivation of MAP values : average of the 3 measurements instead of 2 (now in line with FG SAP). | 12) Using the average of all 3 measurments is more efficient and is consistent with Fibrogen 060 study SAP. |
| | | 13) Implementation of primary efficacy endpoint for FDA (change in Hb from baseline to the average level between week 28 and week 52) and description of statistical analyses including sensitivity analyses for missing data | 13) This additional endpoint has been added in protocol amendment 1.0.Sensitivity analyses for missing data were added following FDA feedback on Fibrogen 060 study. Both Fiborgen and Astellas now in line regarding this endpoint. |
| | | 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) Reduction in the number of analyses of subgroup of interests. | 15) Due to harmonization with Fibrogen 060 study SAP. |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | 16) Addition of the analysis of percentage of time when Hb > 12, 13 or 14. | 16) Due to harmonization with Fibrogen 060 study SAP. |
| | | 17) MMRM model added for number of RBC and volume of RBC during the treatment period. | 17) The concept of percentage of time with high Hb values was implemented in order to account for differences in individual treatment durations. |
| | | 18) Removal of descriptive analysis on the categorical change form baseline in Total FACT-An Score. | 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) Descriptive statistics for categorical change not necessary since analysis on absolute change already planned and considered sufficient. |
| | | 20) Updated rule for complete missing start date and end date for concomitant medications. In that case, the medication will be considered as both previous and concomitant instead of only concomitant.For the analysis of monthly IV Iron use, MMRM has been replaced by ANCOVA. | 20) 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. |
| | | 21) Additional tables run on group of subjects enrolled after protocol amendment implementation | 21) Previous rule considered too conservative since a missing year of start is usually for previous medications than concomitant. Decision was taken to consider such medications as both previous and concomitant.<br>22) MMRM model for monthly IV Iron was not adapted due the data distribution. Since montly average will be |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | | calculated, ANCOVA is more appropriate.<br>23) Added in order to assess the impact of protocol amendment 1.0. |
| 3.0 | 06-November-2017 | 1) Removal of the reporting for date informed consent for genotyping and genotyping sample.<br><br>2) Number and percentage of subjects with drug related TEAEs leading to death added.<br><br>3) Use of SBP ≥170 mmHg AND an increase from BL ≥ 20 mmHg or as DBP ≥ 110 mmHg, AND an increase from BL ≥ 15 mmHg for the definition of hypertension.<br><br>4) Use of the Safety Emergent Period as evaluation period to be considered a concomitant medication.<br><br>5) ATC codes for Iron-chelating agents were updated.<br><br>6) Belarus was added in the list of countries for Region B.<br><br>7) Use of the Efficacy Emergent Period as evaluation period to be considered for the definition of rescue therapy.<br><br>8) Removal of the criterion "at least one Hb value >13 g/dL during the Efficacy Emergent Period" for the definition of Potential Excessive Hematopoiesis (EH). | 1) Decision agreed by the study team.<br><br>2) Astellas standards requirement<br><br>3) Due to harmonization with Fibrogen 060 study SAP.<br><br>4) Decision agreed by the study team.<br><br>5) The ATC codes in previous version were not correct as they actually did not cover the Iron supplementation agents<br><br>6) Protocol requirement.<br><br>7) Decision agreed with the study team to use a consistent approach for rescue therapy by extending the evaluated period up to 7 days after last dose.<br>8) Due to harmonization with Fibrogen 060 study SAP. |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | 9) For the analysis of number of RBC packs and volume, MMRM has been replaced by ANCOVA | 9) MMRM model for RBC packs and volume was not adapted due the data distribution. Since monthly average will be calculated, ANCOVA is more appropriate. |
| | | 10) Removal of the calculation of the amount of IV iron in monthly intervals during the Efficacy Emergent Period. | 10) Decision agreed by the study team not to report Mean IV Iron for each month but only for the period of interest. |
| | | 11) Removal of the supplementary analysis for each post-dosing timepoint for subjects with mean LDL cholesterol less than 100 mg/dL. Average W12-W28 should be done for both regardless and fasting. | 11) Decision agreed by the study team. |
| | | 12) Implementation of the derivation of the hospitalization duration. When hospitalization is ongoing, the date of end of the Efficacy Emergent Period will be used. | 12) Decision agreed by the study team to be more accurate. |
| | | 13) Removal of other baseline factors including: Diabetes mellitus, age, gender, BMI, baseline eGFR, baseline CRP, ferritin, TSAT, CHr, Hypertension from ANCOVA with MI. | 13) Due to harmonization with Fibrogen 060 study SAP. |
| | | 14) Update the calculation of the number of days of hospitalization per PEY as: Minimum ((Date of discharge, End of Efficacy Emergent Period) – Date of admission + 1)] / [(Duration of Efficacy Emergent Period. | 14) Decision agreed by the study team to be consistent with the definition. |
| | | 15) No comparison for IV Iron between treatment arms. | 15) Due to the very small number of subjects with IV Iron. |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | 16) Figure added with the comparison between central laboratory hemoglobin (g/dL) and HemoCue hemoglobin (g/dL). | 16) Decision agreed by the study team to check the differences between the two methods. |
| | | 17) Removal of the analysis for subjects who have reached Hb >= 11.0 g/dL prior to week 28. | 17) Decision agreed by the study team |
| | | 18) For time to first occurrence of serum creatinine doubled compared with baseline, a subject will be counted with an event either for doubled serum creatinine or ESRD, whichever comes first. Otherwise, we censor at date of last non-missing serum creatinine assessment during the Safety Emergent Period. | 18) Decision agreed by the study team to be consistent with the definition for time to first occurrence of serum creatinine doubled compared with baseline. |
| | | 19) For WPAI scale, the number and proportion of employed subjects was summarized by visit and treatment arm. | 19) Decision agreed by the study team to check the number of employed subjects by visit and treatment arm. |
| | | 20) Fatigue subscale score added for FACT-An questionnaire. | 20) Descriptive statistics fatigue were added due to a request from Health Economics and Outputs Research Department. |
| | | 21) Time to All Deaths including Post study FU added. | 21) Decision agreed by the study team. |
| | | 22) Time to AEs by eGFR categories (<15 vs ≥ 15 mL/min/1.73m$^2$) added. | 22) Decision agreed by the study team. |
| | | 23) Number of events and event rate (per 100 patient years) during the Safety Emergent Period with TEAEs added. | 23) Decision agreed by the study team to compare treatments in a fair way due to potential difference in exposure between the two arms. |
| | | 24) New sensitivity analysis (MMRM) for change to the | 24) Decision agreed by the study team. |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | average Hb in weeks 28-52 including all Hb values up to end of treatment visit.<br><br>25) Interaction term baseline*visit added to the main MMRM model for Hb analysis | <br><br>25) Based on internal statistical guidance and check on model fit. |
| 4.0 | 08-March-2018 | 1) Two additional subgroup analyses [Baseline CRP category (≤ULN vs. > ULN)] and Baseline Iron Repletion Status category [(TSAT >= 20% and ferritin >= 100 ng/mL) vs. others] added for the primary efficacy endpoints and the key secondary efficacy endpoint in Hb (week 28 – week 36) | 1) Baseline CRP group (CRP ≤ULN vs. CRP > ULN) is considered as a marker for inflammation which may affect treatment response. A subgroup analysis is added to explore the consistency of the effect of roxadustat on the Hb related primary efficacy parameters. Iron repletion status may affect treatment response and is added for consistency with Fibrogen 060 study SAP. |
| 5.0 | 11-Jul-2018 | 1) Time to Rescue Therapy during the first 24 weeks was replaced by Time to Rescue Therapy during the treatment period as Key Secondary Efficacy endpoint. Endpoint at 24 weeks was therefore considered as Other Secondary endpoint<br><br>2) VT and PF SF-36 scores have been switched in the hierarchical order of the key secondary endpoints<br>3) eGFR slope over time was added at the bottom of the list of key secondary endpoint<br>4) Details were added in All Randomized section regarding exclusion of subjects from site 70051<br><br>5) Additional analyses on adverse | 1) Due to harmonization with Fibrogen 060 study SAP.<br><br>2) Due to harmonization with Fibrogen 060 study SAP.<br>3) Due to harmonization with Fibrogen 060 study SAP.<br>4) Site 70051 was terminated due to GCP violations and it is now clearly mentioned in the population section that those subjects are excluded from all analysis set.<br>5) Due to harmonization |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | event were added considering events ococuring up to last dose +_7 days (i.e OT-7)<br>6) Details regarding derivation of CV history were added in relevant section<br>7) Time to first dialysis was changed to Time to chronic dialysis or renal transplant | with Fibrogen 060 study SAP.<br>6) To provide additional details about derivation.<br>7) Due to harmonization with Fibrogen 060 study SAP. |
| 5.0 | 18-Jul-2018 | **Time to event endpoints:**<br>For the following endpoints (new or existing), IPCW analysis is added. Additional endpoints of interest are added as indicated by new.<br>a. Rescue therapy:<br>Time to first use of rescue therapy [composite of RBC transfusions, IV iron supplementation and rescue ESA]<br>Time to first use of RBC transfusion<br>Time to first use of IV iron supplementation<br>Time to first use of ESA<br>b. Hospitalization:<br>Time to first hospitalization<br>c. CKD progression:<br>Time to CKD progression (composite of doubling serum creatinine, chronic dialysis or renal transplant, and Death) (new)<br>Time to doubling serum creatinine (new)<br>Time to chronic dialysis or renal transplant<br>Time to chronic dialysis or renal transplant or occurrence for a subject who died (new)<br>Time to at least 40% decrease in eGFR from baseline, chronic dialysis or renal transplant (new)<br>Time to at least 40% decrease in eGFR from baseline (new)<br>d. Adverse Events:<br>Time to serious TEAE<br>Time to death<br>Time to TEAE with grade 3 or higher<br>Time to TEAE by SOC (>10% for a SOC in total SAF) | Due to harmonization with Fibrogen 060 study SAP. |
| 5.0 | 18-Jul-2018 | Appendix 8 is added about inverse probability censoring weighted model | To provide details about inverse probability censoring weighted |

| Version | Date | Changes | Comment/rationale for change |
|---|---|---|---|
| | | | model |
| 5.0 | 18-Jul-2018 | Rate of progression of CKD measured by annualized eGFR slope over time: Analysis is updated to use non-log transformed eGFR values. | To perform log transformed analysis only if residual analysis indicates heteroscedasticity |
| 5.0 | 18-Jul-2018 | A summary is added for onset of common (>=5% in Any Treatment Group) TEAEs by Treatment Duration: <3 months, ≥3 to ≤6 months, >6 to ≤12 months, >12 months | To provide summary of common TEAEs by treatment duration. |
| 5.0 | 02-Aug-2018 | End point time to doubling of serum creatinine or ESRD is updated as Time to doubling of serum creatinine or chronic dialysis or renal transplant | To be consistent with Time to CKD progression Endpoint. |

# 9 REFERENCES

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

Gart J, Nam J. Approximate Interval Estimation of the Difference in Binomial Parameters: Correction for Skewness and Extension to Multiple Tables. Biometrics. 1990, 46: 637-643

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

Hedeker D, Gibbons R. Application of Random-Effectes Pattern-Mixture Models for Missing Data in Longitudinal Studies. Psychological Methods 1997, Vol 2, No 1 64-78.

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)

Ratitch B, O'Kelly M. Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures, PharmaSUG2011.

Spiessens B, Debois M. Adjusted significance levels for subgroup analyses in clinical trials. Contemporary Clinical Trials, Volume 31, Issue 6; November 2010.

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.

# 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 |

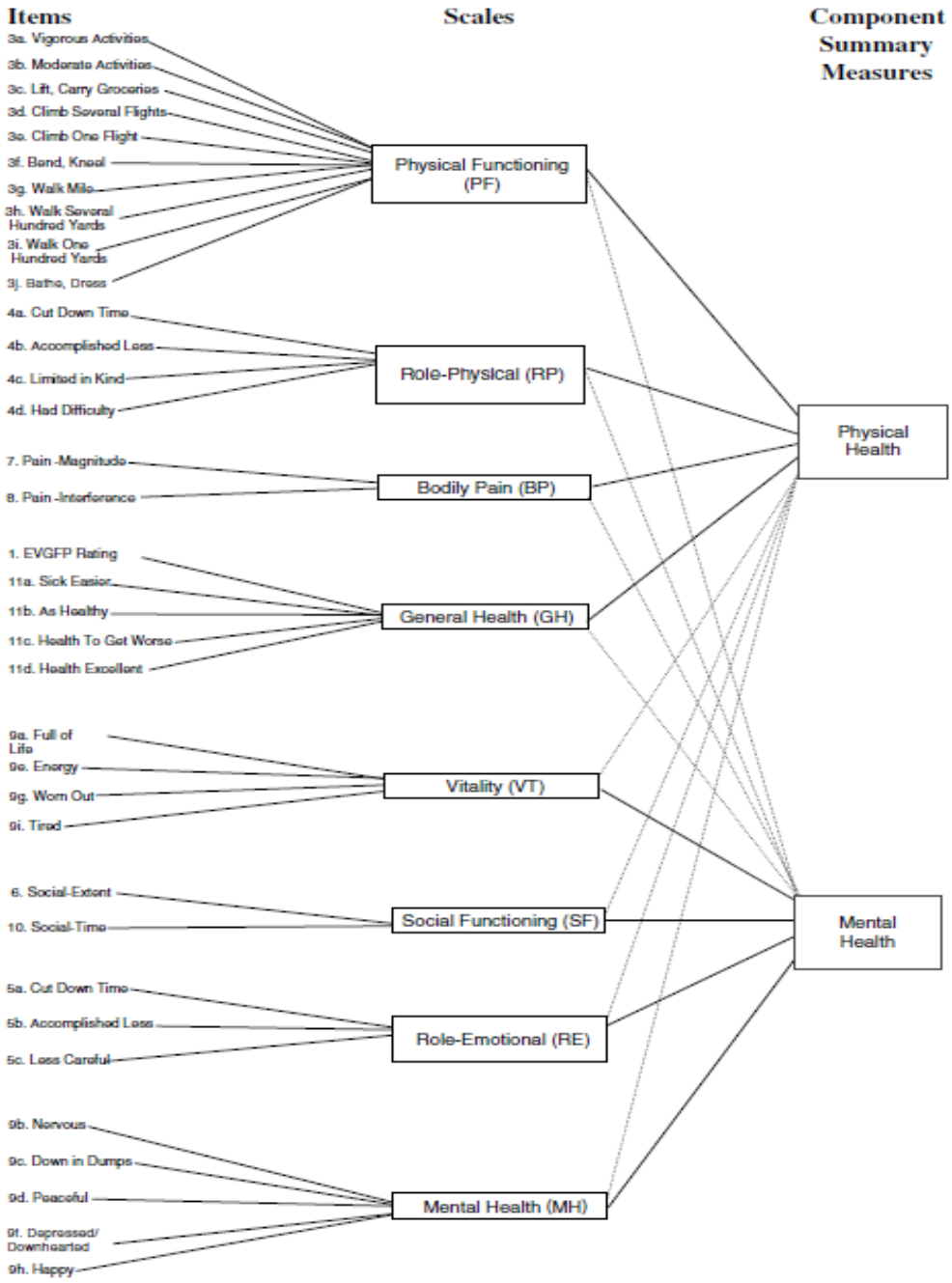
## SF-36 v2

**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 | + | _____ | =_____ |
| | GF5 | 0 | + | _____ | =_____ |
| *Score range:* 0-28 | 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 | - | _____ | =_____ |
| | An3 | 4 | - | _____ | =_____ |
| *Score range:* 0-80 | An4 | 4 | - | _____ | =_____ |
| | An5 | 0 | + | _____ | =_____ |
| | An6 | 4 | - | _____ | =_____ |
| | An7 | 0 | + | _____ | =_____ |
| | An8 | 4 | - | _____ | =_____ |
| | An9 | 4 | - | _____ | =_____ |
| | An10 | 4 | - | _____ | =_____ |
| | B1 | 4 | - | _____ | =_____ |
| | An11 | 4 | - | _____ | =_____ |
| | An12 | 4 | - | _____ | =_____ |
| | BL4 | 0 | + | _____ | =_____ |
| | An13 | 0 | + | _____ | =_____ |
| | An14 | 4 | - | _____ | =_____ |
| | An15 | 4 | - | _____ | =_____ |
| | An16 | 4 | - | _____ | =_____ |

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

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

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

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

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

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

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

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

## FACIT-Fatigue Subscale Scoring Guidelines

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

| Subscale 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 ❑

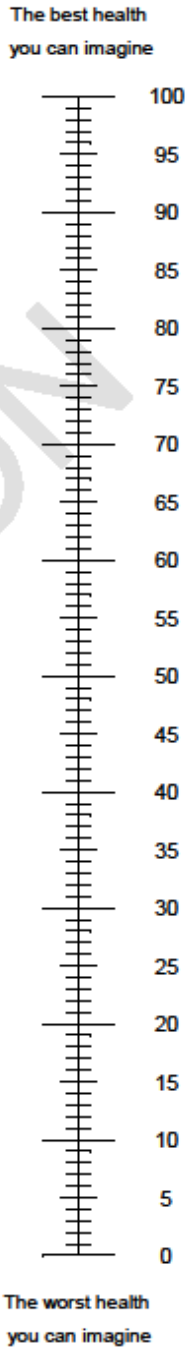
**PAIN / DISCOMFORT**
- I have no pain or discomfort ❑
- I have slight pain or discomfort ❑
- I have moderate pain or discomfort ❑
- I have severe pain or discomfort ❑
- I have extreme pain or discomfort ❑

**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 = ☐



## 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 | ATC level 4 = B03AC [WHODD drug code = '00023501001', '90135401001'] |
| 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: Inverse probability censoring weighted model

Background

The possible presence of differential SDD rates between the roxadustat arm and the placebo arm, which may indicate informative censoring, will have a strong impact on the use of conventional survival methodology using Cox regression or other non-parametric survival analysis methods, which may lead to heavy bias in the estimation of the hazard ratio (HR). Patients who dropped out may have dependent or nonrandom censoring. The IPCW approach presented below is often used under these situations.

Step 1: Model the censoring mechanism.

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

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

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

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

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

## 10.8 Appendix 8: 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

[redacted]

### Contributors and Reviewers

[redacted]

# Approver Signatories

(E-signatures are attached at end of document)

[redacted] was the study statistician for this study.

[redacted] was the Global Statistician Leader and biostatistics peer reviewer of this Statistical Analysis Plan

I approve the contents of this Statistical Analysis Plan:

[redacted]
Astellas Pharma Europe B.V.

# ELECTRONIC SIGNATURE PAGE

**Document Type** : Statistical Analysis Plan

**Document Control Number** : mgc1400722

**Amendment Number** : N/A

**International Study Number** : 1517-CL-0608

**Departmental Study Number** : N/A

**Actual Version Number** : Version 5.0

**Document Version** : 6.0

**Nonclinical Initial SD Approved Date (UTC)** : N/A

| Date (UTC) | Signed by | Sign Off Meaning |
|---|---|---|
| 08/02/2018 20:35:49 | [redacted] | Study Statistician Approval |
| **Full Name / Legal Name** | [redacted] | |
| 08/02/2018 20:58:05 | [redacted] | Statistics Lead |
| **Full Name / Legal Name** | [redacted] | |
| 08/03/2018 07:23:46 | [redacted] | Scientific Lead Approval |
| **Full Name / Legal Name** | [redacted] | |
| | | |
| **Full Name / Legal Name** | | |
| | | |
| **Full Name / Legal Name** | | |
| | | |
| **Full Name / Legal Name** | | |
| | | |
| **Full Name / Legal Name** | | |
| | | |
| **Full Name / Legal Name** | | |
| | | |
| **Full Name / Legal Name** | | |
| | | |
| **Full Name / Legal Name** | | |

*UTC: Coordinated Universal Time