# STATISTICAL ANALYSIS PLAN

| | |
|---|---|
| **Protocol:** | A Randomized, Open-label, Active-controlled Study Evaluating the Efficacy and Safety of Dose Conversion from a Long-acting Erythropoiesis-stimulating Agent (Mircera®) to Three Times Weekly Oral Vadadustat for the Maintenance Treatment of Anemia in Hemodialysis Subjects |
| **Compound:** | Vadadustat (AKB-6548) |
| **Protocol Number:** | AKB-6548-CI-0039 |
| **US IND Number:** | 102,465 |
| **Phase:** | Phase 3b |
| **Sponsor:** | Akebia Therapeutics, Inc.<br>245 First Street<br>Cambridge, MA 02142<br>United States of America |
| **Author:** | [redacted]<br>Firma Clinical Research<br>221 Schilling Circle, Suite 1888<br>Hunt Valley, MD 21031 |
| **Initial SAP Date:** | 21 July 2022/Version 1.0 |
| **Amendment #1:** | 31 October 31, 2022/Version 1.1 |





NCT Number: NCT04707768
This NCT number has been applied to the document for purposes of posting on Clinicaltrials.gov

# Signature Page

[redacted] MD, PhD
**Akebia Therapeutics, Inc**
**Senior Medical Director, Clinical Development**

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 28-Nov-2022 | 3:50 PM EST
83CA1E151FFC49F08FF35398211C33F4

**Signature** 28-Nov-2022 | 3:51 PM EST **Date**

[redacted] PhD
**Akebia Therapeutics, Inc**
**Vice President, Biometrics**

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 28-Nov-2022 | 3:49 PM EST
2C57E53FDEFA4E1FAB9B91D0904135D4

**Signature** 28-Nov-2022 | 3:49 PM EST **Date**

[redacted] PhD
**Lead Statistician**
**Firma Clinical Research, LLC**

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I am the author of this document
Signing Time: 28-Nov-2022 | 7:29 PM EST
42F26B12FD2045F9AEEB2F4A6E6F2D81

**Signature** 28-Nov-2022 | 7:29 PM EST **Date**

# REVISION HISTORY

| Version | Version Date | Author | Summary of Changes Made |
|---|---|---|---|
| 1.0 | 21Jul2022 | [redacted] | Initial |
| 1.1 | 31Oct2022 | [redacted] | Sections 6.11.2 and 6.11.3<br>Removed the hierarchical testing. The treatment comparison is based on the combined vadadustat dosing group vs. Mircera for both primary and secondary efficacy endpoints. The comparisons between vadadustat 900 mg TIW vs. Mircera and vadadustat 600 mg TIW vs. Mircera are the sensitivity analyses. |

# Table of Contents

## List of Abbreviations and Definitions

The following abbreviations and specialist terms are used in this document.

| Abbreviation | Definition |
|---|---|
| AE | adverse event |
| AESI | adverse event of special interest |
| ALP | alkaline phosphatase |
| ALT | alanine aminotransferase |
| ARCRS | American Renal Clinical Research Services |
| AST | aspartate aminotransferase |
| ANCOVA | analysis of covariance |
| BMI | body mass index |
| BUN | blood urea nitrogen |
| CBC | complete blood count |
| CI | confidence interval |
| CKD | chronic kidney disease |
| CPK | creatine phosphokinase |
| CRO | contract research organization |
| CRP | C-reactive protein |
| CV | cardiovascular |
| DBP | diastolic blood pressure |
| DILI | drug-induced liver injury |
| eCRF | electronic case report form |
| EOT | end of treatment |
| EPO | erythropoietin |
| ESA | erythropoiesis-stimulating agent |
| ET | early termination |
| [redacted] | [redacted] |
| FAS | full analysis population |
| FCS | fully conditional specification |
| Hb | hemoglobin |
| HDL | high density lipoprotein |

| Abbreviation | Definition |
|---|---|
| HF | heart failure |
| INR | international normalized ratio |
| IV | intravenous |
| [redacted] | [redacted] |
| MAR | missing at random |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MMRM | mixed model repeated measures |
| MNAR | missing not at random |
| LDH | lactate dehydrogenase |
| LDL | low density lipoprotein |
| LLD | lower limits of detection |
| MCH | mean corpuscular hemoglobin |
| MCHC | mean corpuscular hemoglobin concentration |
| MCS | mental health composite score |
| MCV | mean corpuscular volume |
| PCS | physical health composite score |
| PEP | primary efficacy period |
| [redacted] | [redacted] |
| [redacted] | [redacted] |
| PP | per protocol |
| PT | preferred term |
| RBC | red blood cell |
| RDW | red blood cell distribution width |
| SAE | serious adverse event |
| SAP | statistical analysis plan |
| SBP | systolic blood pressure |
| SD | standard deviation |
| SE | standard error |
| SEP | secondary efficacy period |

| Abbreviation | Definition |
|---|---|
| SF-12 | 12-Item Short Form Survey |
| SGPT | serum glutamic pyruvic transaminase |
| SOC | system organ class |
| TEAE | treatment-emergent adverse event |
| TFL | table, listing, figure |
| TIBC | total iron binding capacity |
| TIW | three times weekly |
| TSAT | transferrin saturation |
| ULN | upper limit of normal |
| US | United States |
| USRC | U.S. Renal Care |
| [redacted] | [redacted] |
| WBC | white blood cell |

# 1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical analysis methods that will be used to analyze, and report results for Akebia Protocol AKB-6548-CI-0039.

This SAP is based on the final study protocol, version 1.0 (dated 16 September 2020). Details of study design, conduct and data collection are described in the study protocol, the electronic case report form (eCRF), and general eCRF completion guidelines. The specifications of tables, figures, and data listings are included in a separate document (i.e., TFL shells).

If the analyses described in the protocol differs from those in this SAP, the analysis methods described in the SAP prevail.

# 2. STUDY OBJECTIVES

## 2.1 PRIMARY OBJECTIVE

To demonstrate the efficacy and safety of vadadustat administered three times weekly (TIW) compared to long-acting ESA (Mircera) for the maintenance treatment of anemia in hemodialysis subjects.

# 3. STUDY DESIGN

## 3.1 OVERVIEW

This is a multi-center, randomized, open-label, active-controlled study of the efficacy and safety of conversion from long-acting ESA (Mircera) to vadadustat TIW for the maintenance treatment of anemia in hemodialysis patients. Following a Screening period of up to 8 weeks (56 days), subjects who meet all inclusion and none of the exclusion criteria will be randomized 1:1:1 to vadadustat 600 mg TIW, vadadustat 900 mg TIW, or to remain on Mircera according to the dialysis center's protocol. The planned number of subjects are 450 at approximately 60 investigative sites in the United States (US) at outpatient hemodialysis centers.

Randomization will be stratified across dialysis organization: USRC (U.S. Renal Care), ARCRS (American Renal Clinical Research Services), and Frenova (Frenova Renal Research).

Following randomization there will be 2 periods during the study:

- **Conversion and Maintenance Period (Weeks 0 to 52):** conversion to vadadustat TIW or to remain on Mircera (Weeks 0 to 20). There will be a primary efficacy

evaluation period (Weeks 20 to 26) and a secondary efficacy evaluation period (Weeks 46 to 52).

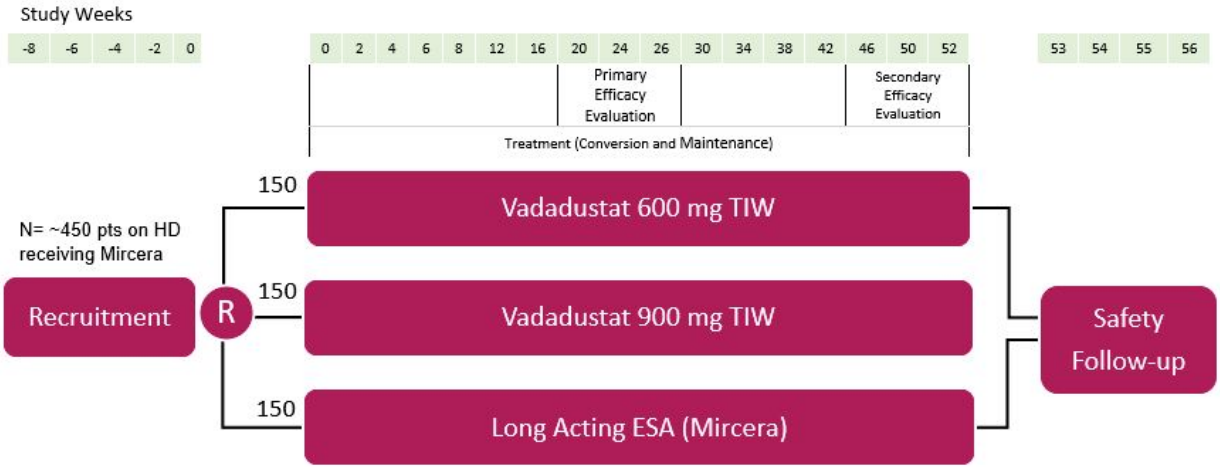
- **Safety Follow-up Period (Early Termination [ET] and Follow-Up):** post-treatment safety follow-up visit (ET/End of Treatment [EOT] +4 weeks) either in person or via telephone.

End of Study:

- **Study Completion**
  The study will be considered completed after all randomized subjects have completed their final study visit (Week 52/or ET and Safety Follow-Up at Week 56).
- **Subject Completion**
  Completion of treatment: A subject will be considered as having completed the study treatment after completion of their study treatment through Week 52.

  Completion of study: A subject will be considered as having completed the study after completion of their study treatment through Week 52 and the safety follow-up visit.

**Figure 1. Study Schema**




ESA: erythropoiesis-stimulating agent; HD: hemodialysis; R: randomized; TIW: three times weekly

Dosing will be initiated at Baseline and the first dose of vadadustat will be administered at the hemodialysis unit after other Baseline procedures have been completed. Subjects randomized to vadadustat may take vadadustat with or without food and will be instructed to swallow the tablet(s) whole. Subjects are to take vadadustat at roughly the same time each day.

Mircera assigned subjects will be administered IV Mircera at the hemodialysis unit according to the dialysis center's protocol.

Dose adjustments will be guided by central laboratory Hb concentrations throughout the study to determine if vadadustat or Mircera doses will be adjusted, interrupted, or maintained. Mircera dosing will be based on standard of care laboratory measures according to the dialysis center's protocol. Vadadustat dose adjustment will be based on the guidelines in the protocol Section 7.1.2.

The minimum dose of vadadustat will be 300 mg TIW (1 tablet TIW) and the maximum dose will be 1200 mg TIW. Subjects whose dose of vadadustat is interrupted due to elevated Hb will continue in the study. Unless contraindicated, treatment will be resumed whenever possible and assessed at every study visit following study drug interruption.

Clinical evaluations will be conducted during the course of the study based on the schedule of activities in the protocol Table 1. If the evaluations will occur on a hemodialysis day, the clinical evaluations should be completed before dialysis, if applicable.

## 3.2 Study Population

The study population will consist of subjects greater than or equal to 18 years of age who are receiving chronic, outpatient in-center hemodialysis three times weekly (TIW), requiring erythropoiesis-stimulating agent (ESA) treatment and are on maintenance treatment with Mircera. Study subjects will also be required to have a mean Screening hemoglobin (Hb), based on 2 results, between 8.5 and 11.0 g/dL (inclusive).

## 3.3 Sample Size Determination

For the primary efficacy analysis, it is assumed that the difference in mean change from Baseline in Hb for vadadustat will be the same as the active control, Mircera, and the common standard deviation for the mean change from Baseline will be assumed to be 1.2 g/dL. The noninferiority margin of -0.75 g/dL will be used (for vadadustat minus Mircera). With the 1:1:1 randomization ratio of vadadustat 600 mg TIW, vadadustat 900 mg TIW, and Mircera, and approximately 150 subjects in each arm, the noninferiority test will have >90% power with consideration of a 30% drop out rate.

## 3.4 Study Medication Stopping Rules

Vadadustat must be permanently discontinued if a subject meets 1 of the following criteria:

- ALT or AST >3x ULN and total bilirubin >2x ULN
- ALT or AST >3x ULN and international normalized ratio (INR) >1.5
- ALT or AST >8x ULN
- ALT or AST remains >5x ULN over 2 weeks (re-challenge generally should be avoided with ALT or AST >5x ULN unless there are no other good therapeutic options)

- ALT or AST >3x ULN with symptoms (e.g., fatigue, nausea, vomiting, right upper quadrant pain, fever, rash) or eosinophilia
- undergo a solid organ (including kidney), hematopoietic stem cell, or bone marrow transplantation

### 3.5 MEASURES TO MINIMIZE BIAS: RANDOMIZATION AND BLINDING

This will be an open-label study. Treatment assignment will be done through a validated electronic system by block randomization with a block size of 6. Treatments will be administered in an open-label fashion. The sponsor study team will be blinded to 'by treatment' aggregated analyses except for the unblinded statistician. This will be detailed in a separate document of "Blinding Plan".

In addition, the study involves the use of an Independent Data Monitoring Committee (IDMC), and an identical schedule of visits, procedures, and assessments for all treatment groups in order to reduce the potential bias. An external independent group of unblinded biostatisticians and programmers will provide the support to IDMC.

## 4. ANALYSIS POPULATION

The following analysis populations will be used in this study:

- Randomized population: defined as all randomized subjects. This population will be analyzed based upon the randomized treatment. This is the Intention to Treat (ITT) population.
- Full analysis population (FAS): defined as randomized subjects receiving 1 or more doses of study medication and had at least one post-dose Hb assessment. This population will be analyzed based upon the randomized treatment.
- Per protocol (PP) population: defined as all randomized subjects who received study medication during the primary evaluation period, had at least 1 Hb assessments during the primary evaluation period, and have no important protocol deviations affecting the primary endpoint analyses, i.e., up to Week 26. This population will be analyzed based upon the actual treatment received. Protocol deviations leading to exclusion from the PP population will be specified prior to database lock on a blinded basis and recorded in a separate document.
- Safety population: defined as all subjects who received at least 1 dose of study medication. This population will be analyzed based upon the actual treatment received. Subjects who received in error some vadadustat and some Mircera (excluding rescue therapy) will be classified by the more frequently received drug.

Efficacy analyses will utilize the randomized/ITT, full analysis (FAS), and PP populations while safety analyses will utilize the safety population.

# 5. ENDPOINTS

## 5.1 PRIMARY EFFICACY ENDPOINT

Mean change in Hb between Baseline (average of the last 2 available pretreatment Hb) and the primary evaluation period (average Hb from Weeks 20 to 26, inclusive).

## 5.2 SECONDARY EFFICACY ENDPOINT

Mean change in Hb between Baseline (average of the last 2 available pretreatment Hb) and the secondary evaluation period (average Hb from Weeks 46 to 52, inclusive).

## 5.3 OTHER EFFICACY ENDPOINTS

Other endpoints include the following:

- Proportion of subjects having a Hb value within the target range (10.0 to 11.0 g/dL) during the primary evaluation period (Weeks 20 to 26).
- Proportion of subjects having a Hb value within the target range (10.0 to 11.0 g/dL) during the secondary evaluation period (Weeks 46 to 52).
- Proportion of subjects having average Hb value within the target range (10.0 to 11.0 g/dL) during the primary evaluation period (Weeks 20 to 26)
- Proportion of subjects having average Hb value within the target range (10.0 to 11.0 g/dL) during the secondary evaluation period (Weeks 46 to 52)
- Mean change from baseline and percentage change from baseline in the primary efficacy period (PEP) and secondary efficacy period (SEP) for iron-related parameters
- Proportion of subject receiving rescue by analysis periods.
- Proportion of subjects receiving RBC transfusions by analysis periods.
- Proportion of subjects receiving ESA medication by analysis periods.
- Time to first rescue by narrow and broad (include any rescue, ESA, and RBC respectively)
- Proportion of subjects who received any IV iron therapy by analysis period)
- Mean weekly IV dose elemental iron by analysis period and route of administration
- Mean change in serum glucose and lipid parameters between Baseline and the PEP (Weeks 20 to 26) and SEP (Weeks 46 to 52)
- Mean change in iron metabolism (including iron, hepcidin, ferritin, transferrin saturation [TSAT], total iron binding capacity [TIBC]) between Baseline and the PEP (Weeks 20 to 26) and SEP (Weeks 46 to 52)

### 5.4 PRIMARY SAFETY ENDPOINTS

Treatment-emergent adverse events (TEAEs) and Treatment-emergent Serious Adverse Events (SAEs).

### 5.5 SECONDARY SAFETY ENDPOINTS

Secondary endpoints include the following:

- Proportion of subjects with Hb >11.0, Hb >12.0, >13.0, or >14.0 g/dL.
- Proportion of subjects with Hb <7.0, <8.0, <9.0, or <10.0 g/dL.
- Proportion of subjects with Hb increase >1.0 g/dL within any 2-week interval or >2.0 g/dL within any 4-week interval.

### 5.6 OTHER SAFETY ENDPOINTS

Safety will also be evaluated by the following assessments:

- Laboratory test results
- Vital signs

[redacted]

## 6. STATISTICAL METHODOLOGY AND ANALYSES

### 6.1 GENERAL CONSIDERATIONS

Data collected throughout the study will be summarized using descriptive statistics and listed in by-subject listings. Continuous variables will be summarized using number of subjects with data, mean, standard deviation (SD), median, minimum, and maximum. For categorical variables, the number and percentage of subjects in each category will be tabulated. Summaries will be provided by treatment group and for vadadustat treatment groups combined within appropriate analysis populations, and by time point/time period, as appropriate.

For Hb, Baseline will be calculated as the average of the last 2 central laboratory Hb measurements of samples taken at or prior to the first dose. For blood pressure, baseline will be calculated as the average of the blood pressure within same visit, and then select

the last available results on or prior to the first dose. For other parameters, unless otherwise specified, Baseline will be defined as the last available value prior to the first dose of study medication.

The primary efficacy endpoint as well as the secondary and all other efficacy endpoints will be summarized using descriptive statistics by treatment group, and by analysis visit and/or analysis period as appropriate. Mean values of Hb and selected other efficacy parameters will be plotted across analysis visits/periods by treatment group.

Hb values will be assessed through the central laboratory for dose adjustments, efficacy, and safety evaluations.

If the central laboratory uses assays that have lower limits of detection (LLD), all laboratory results below the LLD will be imputed with the LLD. Results reported as greater than a value (i.e., "> value") will be imputed as 1.5 x that value.

Mircera will be converted to IV epoetin equivalent unit per kilogram per week (U/kg/week) as 1:220. The conversion has been derived from published literature (Paganini, 1995; Kaufman1998Levin, 2008Wright2015) and input from clinical experts.

Study days are defined as follows:

Study Day = [Event date – First dosing date + 1], if on or after first dosing date

[Event date – First dosing date], if before first dosing date.

Note that with the definition above, days of "0" will not be used. As such, the study day can be interpreted as the number of days before or after first day of dosing. Event date refers to the date associated with the result being summarized. In some cases, this is the date of an assessment or measurement; in other cases, this is the onset date of an adverse or outcome event.

If the subject numbers, combined over two treatment groups, in one of the strata are less than 30, the strata will be combined with the next large strata for efficacy and safety analyses.

### 6.2 Visit and Analysis Time Period Classification

All assessments will be mapped to the corresponding analysis visit (i.e., target week) based on study day, including unscheduled visits. Table 1 is a general window for each visit based on the protocol. Table 2 will be applied for less frequent assessments such as [redacted] C-Reactive Protein (CRP), Complete Blood Count (CBC) with differential, Serum Chemistry, Lipid Panel, Reticulocyte count, Biomarkers, Erythropoietin and Dialysis Adequacy. All other assessments will apply to the analysis window from Table 1.

If more than one Hb assessment is available in a post-baseline analysis period (i.e., PEP or SEP), then all results will be used to calculate the average of Hb. But if more than one record is available in same day, the last assessment (i.e., at most one record per day) will

be used to calculate the average. For other assessments, if more than one assessment is available in each window, the assessment which is closest to the target day will be used for analysis. If the assessments are equally close, the last assessment will be used.

**Table 1. Classification of Assessments at Every Visit (General)**

| Analysis Time Period | Visit | Target Day | Analysis Window |
|---|---|---|---|
| Baseline | Screening Visit 1 | - | - |
| | Screening Visit 2 | - | - |
| | Baseline | Day 1 | <=Day 1 (prior to first dose)[a] |
| Week 2-6 | Week 2 | Day 14 | Day 1 (post dose) – 21 |
| | Week 4 | Day 28 | Day 22 – 35 |
| | Week 6 | Day 42 | Day 36 – 49 |
| Week 8-16 | Week 8 | Day 56 | Day 50 – 70 |
| | Week 12 | Day 84 | Day 71 – 98 |
| | Week 16 | Day 112 | Day 99 – 126 |
| Primary Efficacy Period (PEP) (Weeks 20-26) | Week 20 | Day 140 | Day 127 – 154 |
| | Week 24 | Day 168 | Day 155 – 175 |
| | Week 26 | Day 182 | Day 176 – 196 |
| Week 30-42 | Week 30 | Day 210 | Day 197 – 224 |
| | Week 34 | Day 238 | Day 225 – 252 |
| | Week 38 | Day 266 | Day 253 – 280 |
| | Week 42 | Day 294 | Day 281 – 308 |
| Secondary Efficacy Period (SEP) (Weeks 46-52) | Week 46 | Day 322 | Day 309 – 336 |
| | Week 50 | Day 350 | Day 337 – 357 |
| | Week 52 | Day 364 | Day 358 – 378 or ET/EOT assessment date + 15 days |
| Safety Follow-up[b] | Follow-up | ET/EOT assessment date + 28 days | >= ET/EOT assessment date + 15 days |

[a] If patient has no first dose date, day 1 will be the randomization date.
[b] for Vital Sign, Dialysis Access Type, and Therapeutic Phlebotomy

**Table 2. Classification of Assessments for Selected Assessment**

| Analysis Time Period | Visit | Target Day | Analysis Window |
|---|---|---|---|
| Baseline | Screening Visit 1 | - | - |
| | Screening Visit 2 | - | - |
| | Baseline | Day 1 | <=Day 1 (prior to first dose)[a] |
| Week 2-6 | Week 4 | Day 28 | Day 1 (post dose) – 49 |
| Week 8-16 | Week 12 | Day 84 | Day 50 – 126 |
| Primary Efficacy Period (PEP) (Weeks 20-26) | Week 26 | Day 182 | Day 127 – 196 |
| Secondary Efficacy Period (SEP) (Weeks 46-52) | Week 52 | Day 364 | Day 309 – 378 or ET/EOT assessment date + 15 days |

[a] If patient has no first dose date, day 1 will be the randomization date.

### 6.3 Handling of Missing Values

Missing data will be handled using a procedure specific to each variable and particular analysis as described in the sections relevant to each endpoint. If no method for missing data is discussed, descriptive analyses will be based upon observed data without imputation. Appendix B describes the general approaches to be used for missing data.

The algorithms for imputation of partial dates are applied based on following manner.

*Adverse Event Onset*

- If date is completely missing, date is set to date of first dose.
- If year is present and month and day are missing or year and day are present, and month is missing:
  - If year = year of first dose, then set month and day to month and day of first dose.
  - If year < year of first dose, then set month and day to December 31.
  - If year > year of first dose, then set month and day to January 1.
- If month and year are present and day is missing:
  - If year = year of first dose:
    - month = month of first dose, then set day to day of first dose date.
    - month < month of first dose, then set day to last day of month.
    - month > month of first dose, then set day to first day of month.
  - If year < year of first dose, then set day to last day of month.
  - If year > year of first dose, then set day to first day of month.
- For all other cases, set date to date of first dose.

*Adverse Event End Date*

- If year is present and month and day are missing or year and day are present and month is missing, set end month and day to December 31.
- If month and year are present and day is missing, set the day to last day of the month.
- If fatal event, date is set to minimum of imputed end date and death date.
- For all other cases, set date to missing.

*Concomitant Medication Start Date*

- If start date is completely missing, start date is the date of informed consent except for the records that medication end date precedes date of informed consent.
- If start year is present and month and day are missing or year and day are present and month is missing, set start month and day to January 1.
- If start year and month are present and day is missing, set start day to 1st day of month.

*Concomitant Medication End Date*

- If end date is completely missing, end date is the last available date in the study.
- If end year is present and month and day are missing or year and day are present and month is missing, set end month and day to December 31.
- If end year and month are present and day is missing, set end day to last day of the month.

The imputed dates must be logical, ensuring that no end date is after database lock or death or before the start date.

### 6.4 WEEK 26 ANALYSIS

The protocol mentioned that a formal week 26 analysis for the primary efficacy endpoint (the 2 comparisons between the vadadustat groups and Micera with hierarchical testing) may be conducted after the last patient completes the primary efficacy period (Week 26).Because of the complete response letter received in March 2022 from FDA regarding the Vadadustat NDA 215192 , Akebia decided not to conduct this week 26 analysis.

### 6.5 SUBJECT DISPOSITION

The number and percentage of subjects randomized and included in each analysis population will be summarized by treatment group and overall. Reasons for excluding subjects from the analysis populations will be presented in a by-subject listing.

The number of randomized subjects who completed study medication treatment, discontinued from study medication early, and completed or discontinued from the study and reasons for discontinuation will be summarized by treatment group and overall.

Time to study treatment discontinuation will be estimated using the Kaplan-Meier product-limit method and log rank test as described in Appendix A. Subject completed

study treatment will be censored at date of study completion. The hazard ratio and corresponding 95% confidence interval for the comparison of vadadustat to and Mircera control group will be estimated using a Cox proportional hazards model, with treatment group, randomization stratification factor (dialysis organization), and baseline Hb as covariates. Kaplan-Meier plot of time to study treatment discontinuation will be provided.

### 6.6 PROTOCOL DEVIATIONS

Definitions of important and non-important protocol deviations will be described in a master Protocol Deviation log file, including those protocol deviations leading to exclusion from the PP populations. These will be specified prior to database lock on a blinded basis.

Protocol deviations will be summarized in the randomized population as the number and percentage of subjects with a protocol deviation by treatment group and overall. A by subject listing of protocol deviations, indicating those exclusionary from the PP population, will be provided.

### 6.7 DEMOGRAPHIC AND BASELINE CHARACTERISTICS

The following item will be summarized by treatment group and overall based on randomized population:

Demographic variables: age, age group 1 (<65 years, >=65 years), age group 2 (>=18-64 years, >=65-84 years, >=85 years), sex, ethnicity, race, height, pre- and post-dialysis weight, body mass index (BMI)

BMI ($kg/m^2$) = weight (kg)/(height (m))at Screening V2

Baseline characteristics: stratification factor (Dialysis organization), Mircera monthly dose amount, Baseline IV/Oral iron dose use (mg/week), Kt/V result, history of diabetes mellitus, cardiovascular disease, retinal disorder, statin used, systolic blood pressure (SBP), diastolic blood pressure (DBP) and heart rate.

Baseline Laboratory Measurements: Hb, hepcidin, iron indices (ferritin, TSAT, TIBC, serum iron), serum glucose, lipid profile (total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides) and CRP.

Continuous variables will be summarized as descriptive statistics. Categorical variables will be summarized by using number and percentage.

### 6.8 MEDICAL HISTORY

Medical history terms will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), and the number and percent of subjects with medical history will be summarized by system organ class (SOC) and preferred terms (PT) for each treatment group and overall based on the safety population.

Etiology of CKD and dialysis history will be summarized by treatment group and overall based on safety population.

### 6.9 Prior and Concomitant Medications

Prior and concomitant medications will be coded using World Health Organization Drug dictionary (WHO DD). Data collected from eCRF page "Prior and Concomitant Medications" and "ESA Administration" will be included and summarized for each treatment group and overall based on the safety population. Anti-Hypertensive Medications will be summarized for each analysis period by treatment group and overall based on the safety population.

Prior medications will be defined as any medications that were taken 30 days prior to SV1 to before the first dose of study medication. Concomitant medications will be defined as any medications taken on or after first dose of study medication.

### 6.10 Study Drug Dosing and Compliance

Due to differences in data collection between the two treatment groups, the calculated compliance rate for a given time period will be derived differently for each treatment group:

Compliance rate for vadadustat group (%) = (treatment duration in the given time period (+ 7/3, if the treatment duration in the give time period is not multiple of 7 days)– number of missed dose during the given time period*7/3)/(treatment duration in the given time period (+ 7/3, if the treatment duration in the give time period is not multiple of 7 days))*100.
Any tablet is taken in a day will be considered compliant for that day.

Compliance rate for Mircera group (%) = Sum of minimum (treatment duration of the dosing record + the dosing frequency for the exposure, treatment duration up to the start date -1 of the next dosing record) during the give time period/(treatment duration in the given time period + the dosing frequency of the last dosing record in the given time period for the exposure)*100.

Total duration (week) of dose exposure and interruption will be provided by treatment group for safety population.

The following summary will be provided by treatment group for both safety population and FAS as well:

- Study Treatment Compliance Rates by analysis period
- Average weekly dose by analysis period
- Number of subjects who took vadadustat by dose level and by analysis visit
- Maximum increase in dose amount of Mircera by analysis period
- Study treatment modifications and discontinuations

## 6.11 EFFICACY ANALYSES

The general approach for analysis of continuous outcomes will be analysis of covariance (ANCOVA) with multiple imputation for missing data or mixed models for repeated measurements (MMRM) on observed data.

For binary variables, the general approach will be Mantel-Haenszel estimation of risk difference stratified by dialysis organization with or (without) multiple imputation for handling missing data [Mantel, 1959].

### 6.11.1 Primary Estimand

The primary estimand defining the treatment effect of interest in this trial uses the treatment policy strategy specified in the ICH E9 (R1) Addendum [FDA 2021]. In this trial, the research question is to compare the assignment of Mircera to vadadustat, regardless of intercurrent events. The primary analysis will be performed in the randomized population in accordance with intent-to-treat. This means the analysis will include individuals who never received treatment, discontinued treatment, were rescued with any therapy, and withdrew consent. The estimate of the treatment effect of interest will be the one regardless of whether the intercurrent event occurred.

The primary estimand for this trial is defined by the following components:

- Treatment: vadadustat with rescue therapy as needed versus Mircera with rescue therapy as needed
- Population: randomized population.
- Variable: change in Hb from Baseline (average pretreatment Hb) to the primary evaluation period (average Hb from Weeks 20 to 26, inclusive).
- Intercurrent event: never received treatment, discontinued treatment, rescued with any therapy or withdrew consent.
- Population-level summary: difference in mean change in Hb from baseline to the primary evaluation period between the vadadustat groups and Mircera control group.

### 6.11.2 Primary Efficacy Analyses

The primary efficacy endpoint is defined as the mean change in Hb between Baseline (average pretreatment Hb) and the primary evaluation period (average Hb from Weeks 20 to 26, inclusive).

The primary analysis of the primary endpoint will use the randomized population with imputation of data. All intercurrent events will be handled by treatment policy strategy (Section 6.11.1). Unobserved data after intercurrent events, including Hb outcomes and any covariates to be used in the models, will be imputed with the randomized treatment group to which the subject was randomized using multiple imputation under fully conditional specification (FSC) method as described in Appendix B. Data up to

scheduled visit of week 26 will be included in the primary efficacy analyses. Analysis of covariance (ANCOVA) as described in Appendix A will be used to calculate the LS mean (SE) by treatment group and LS mean difference (SE), 95% confidence intervals (CI) of the difference in mean change in Hb from baseline to the primary evaluation period between the vadadustat groups and Mircera control group, with the randomization stratification factor (Dialysis organization) and Baseline Hb as covariates.

Treatment comparisons will be performed between the vadadustat and Mircera groups. Subjects in the vadadustat starting dose groups of 600 mg or 900 mg TIW will have their doses adjusted to maintain Hb levels of 10.0 to 11.0 g/dL following the protocol's dosing algorithm.

- Noninferiority of vadadustat will be established if the lower limit of 2-sided CI is greater or equal to -0.75 g/dL for combined vadadustat dosing groups minus Mircera.

If the lower limit of the 2-sided 95% confidence interval for the difference between the mean Hb change from baseline to Weeks 20 to 26 in the vadadustat group and the mean Hb change from baseline in the Mircera group is above zero, superiority will have been established and the finding will be interpreted as providing evidence of a greater change from baseline in Hb for vadadustat relative to the control arm. Because the same 95% confidence interval is used to the superiority test, no adjust of multiplicity is needed.

#### 6.11.2.1 Sensitivity Analyses of the Primary Efficacy Endpoint

To assess the robustness of the findings from the primary analysis, the following sensitivity analyses will be conducted:

- Primary analysis will be repeated using the FAS population.
- Primary analysis will be repeated using the PP population with the actual treatment received.
- Primary analysis will be repeated with imputation of data which may have been affected by the intercurrent events. All per-visit Hb values within four weeks of administration of rescue therapy (narrow or broad-on-treatment defined in Section 6.11.4.4) or after EOT visit will be set as missing prior to imputation.
- Primary analysis will be repeated with method of mixed model for repeated measures (MMRM) based on observed data only. The MMRM model as described in Appendix A will include baseline Hb, stratification factor (dialysis organization), treatment group, analysis period (throughout 26-weeks), treatment-by-analysis time period interaction, baseline-by-analysis time period interaction as covariate.
- Tipping point analyses will be performed to assess the effect of the missing data.

- Primary analysis will be repeated for the vadadustat 900 mg TIW starting dose group vs. Mircera and the vadadustat 600 mg starting dose group vs. Mircera as exploratory analyses.

### 6.11.3 Secondary Efficacy Analyses

The study has one secondary efficacy endpoint which will be analyzed formally only if the primary analysis meets the noninferiority margin.

Mean change in Hb between Baseline (average pretreatment Hb) and the secondary evaluation period (average Hb from Weeks 46 to 52, inclusive) will be analyzed using the same methodology as specified for the primary efficacy endpoint. Evaluation of the average change in Hb will employ the approach described for the primary endpoint assessing Weeks 46 to 52 instead of Weeks 20 to 26.

Sensitivity analyses analogous to those performed for the primary efficacy endpoint will be repeated to assess mean change in Hb from Baseline to Weeks 46 to 52. For ANCOVA models, the analysis period of Weeks 46 to 52 will replace the analysis period of Weeks 20 to 26, and for MMRM, the models will be extended to include analysis period up to week 52. The analysis will be repeated for the vadadustat 900 mg TIW starting dose group vs. Mircera and vadadustat 600 mg starting dose group vs. Mircera as exploratory analyses.

After the non-inferiority for both primary and key secondary endpoints have been established, If the lower limit of the 2-sided 95% confidence interval for the difference between the mean Hb change from baseline to Weeks 20 to 26 in the vadadustat group and the mean Hb change from baseline in the Mircera group is above zero, superiority will have been established and the finding will be interpreted as providing evidence of a greater change from baseline in Hb for vadadustat relative to the control arm. Because the same 95% confidence interval is used to the superiority test, no adjustment for multiplicity is needed.

### 6.11.4 Other Efficacy Analyses

Except for specifying the analysis population, other efficacy endpoint analyses will be performed using the randomized analysis population.

#### 6.11.4.1 Hb value within the target range

All subjects will be defined as either being in target range (10.0 to 11.0 g/dL) in Weeks 20 through 26 (“yes”) or not (“no”), based on at least one Hb value meeting the target range during the visit windows in Weeks 20 through 26. This analysis includes the observed Hb values only.

The same criterion will be used based on the average Hb value during the visit windows in Weeks 20 through 26. To ensure uniformity across all analyses, the missing binary outcomes will be computed from the imputed complete dataset in the primary analysis.

The proportion of subjects within target range will be calculated, 95% CI for the proportions will be reported along with the difference in proportions as well as the odds ratio; a Mantel-Haenszel method stratified by dialysis organization will be used to calculate the 95% CI for the statistics of treatment comparison.

Noninferiority will be evaluated using a 2-sided CI for the difference in proportions. Noninferiority will have been established if the lower limit of the CI is above -15%.

If the lower limit of the 2-sided CI for the difference in proportions is above zero, superiority will have been established and the finding will be interpreted as providing evidence of a higher proportion of subjects being within the target range for vadadustat relative to the control arm.

Secondary evaluation period (Weeks 46 through 52) will be analyzed using the same method.

#### 6.11.4.2 Endpoints Related to Iron

Mean change from baseline and percentage change from baseline are presented for hepcidin, ferritin, TIBC, serum iron, and TSAT, ANCOVA model will be used for treatment comparisons with respect to mean change from baseline in the PEP and SEP.

The proportion of subjects who received any iron therapy and the mean weekly dose of elemental iron will be summarized by route of administration, analysis periods, and treatment groups. The calculation of mean IV (or Oral) weekly dose will include subjects with iron administration and will count only weeks in which a subject was still being followed in the denominator. For proportion of subjects who received each (or both) of the relevant routes of elemental iron, 95% CI for the proportions will be reported along with the difference in proportions as well as the odds ratio; a Mantel-Haenszel method stratified by dialysis organization will be used to calculate the 95% CI for the statistics of treatment comparison. For mean weekly dose of elemental iron, ANCOVA model will be used for treatment comparisons.

#### 6.11.4.3 Other Laboratory Parameters

Mean changes from baseline in PEP and SEP will be summarized by treatment group for serum glucose and lipid parameters – including total cholesterol, LDL, HDL, and triglycerides. ANCOVA model will be used for treatment comparisons.

#### 6.11.4.4 Definition of Rescue

The eCRF collects reasons for ESA medication and RBC transfusion. An Event is considered as rescue only when the event occurred between first dose date to permanent study treatment discontinuation. The series of rescue definitions are defined as follows:

- Narrow:
  - For RBC transfusion, when "Low Hemoglobin" is the reason indicated on the eCRF.
  - For ESA medication:
    - Vadadustat arm: the reason of ESA usage is the subject's Hb less than 9.0 g/dL.
    - Mircera arm: two scenarios are defined as below:
      i. the reason of other ESA or Mircera usage is the subject's Hb less than 9.0 g/dL or Mircera dose increase at least 50%.
      ii. the reason of other ESA or Mircera usage is the subject's Hb less than 9.0 g/dL or Mircera dose increase at least 100%.
- Broad:
  - Any RBC transfusion.
  - For ESA medication:
    - Vadadustat arm: any ESA usage.
    - Mircera arm: two scenarios are defined as below:
      i. Any other ESA usage, or Mircera usage due to the subject's Hb less than 9.0 g/dL or Mircera dose increase at least 50%.
      ii. Any other ESA usage, or Mircera usage due to the subject's Hb less than 9.0 g/dL or Mircera dose increase at least 100%.

#### 6.11.4.4.1 Analysis of Rescue

The number and percentage of subjects receiving rescue by rescue type, treatment group and analysis period will be summarized. Confidence interval will be 95% CI for these proportions along with the difference in proportions as well as the odds ratio. A Mantel-Haenszel test stratified by the dialysis organization will be used to calculate the 95% CI for the statistics of treatment comparison.

Time to first rescue therapy administration by narrow and broad definitions separately will be estimated using the Kaplan-Meier product-limit method and log rank test. Subjects without events will be censored at last dose date. The hazard ratio and corresponding 95% CI for the comparison of vadadustat to Mircera will be estimated using a Cox proportional hazards model, with treatment group, randomization stratification factor (dialysis organization), and baseline Hb as covariates. Kaplan-Meier plots of time to first rescue will be provided.

#### 6.11.4.4.2 RBC Transfusion

The same analyses of any rescue will be performed for RBC transfusion by narrow and broad separately.

#### 6.11.4.4.3 ESA Rescue

The same analyses of any rescue will be performed for RBC transfusion by narrow and broad separately. In addition, the increases in dose relative to last dose will be characterized in three categories by percentage increase: <50%, ≥50% and <100%, and ≥100%. Subjects will be presented by their maximum category of increase within each analysis periods.

## 6.12 SAFETY ANALYSES

All analyses of the safety data will be summarized by treatment group and overall based on the safety population unless otherwise specified.

### 6.12.1 Adverse Events

A treatment-emergent adverse event (TEAE) is one that begins or worsens after treatment initiation. TEAEs will be summarized using the number and percentage of subjects with TEAEs for all subjects in the safety population.

All AEs will be coded by MedDRA. The MedDRA version used for reporting this study will be described in a footnote on AE related outputs. TEAEs will be summarized by SOC and PT for each treatment group. All TEAE summaries will provide the number of subjects reporting at least 1 TEAE. TEAE summaries will be ordered by decreasing pooled percentage for SOC, and PT within SOC.

An overall summary table will show number of subjects with at least 1 of each of the following:

- TEAE
- Drug-related TEAE
- Severe TEAE
- Treatment-emergent SAE
- Drug-related treatment-emergent SAE
- TEAE leading to withdrawal of study drug
- Drug-related TEAE leading to withdrawal of study drug
- Fatal TEAE
- All Deaths: Any deaths reported during the study no matter whether deaths are caused by TEAEs

Summaries by SOC and PT will be provided for the following types of TEAEs:

- TEAE
- TEAE by severity
- Drug-related TEAEs (including all categories for relationship to study medication other than “Unrelated”, as determined by the investigator)
- Treatment-emergent serious adverse event (SAE)

- Drug-related treatment-emergent SAE
- TEAEs leading to withdrawal of study drug
- Drug-related TEAE leading to withdrawal of study drug
- Fatal TEAE
- Adverse events of special interest (AESI) listed on Appendix E

A summary by treatment group and PT will be presented for TEAEs reported by at least 5% of subjects in either treatment group.

TEAEs will be summarized by 1) worst severity, and 2) worst causality by SOC and PT. For each subject and each PT, the worst severity recorded will be used in the by-severity summaries. Similarly, the worst causality (most related to treatment) attributed will be used in the by-causality summaries. If severity or causality is missing, data will be imputed to the worst category.

Treatment-emergent SAEs and deaths will be summarized by treatment group and overall, including the number and percentage for subject with Treatment-emergent SAEs, number of deaths (all causes, resulting from adverse events). Summary by SOC and PT will include subjects affected by Treatment-emergent SAEs, occurrence (all, causally related to treatment), number of death (all causes), and deaths causally related to treatment.

A by-subject listing of all TEAEs will be provided. This listing will be presented by treatment group and will include center, subject identifier, age, sex, race, AE (SOC, PT, and verbatim term), study day of onset, study day of resolution, duration, severity, seriousness, relationship to the study medication, action taken, outcome and causality.

A by-subject listing of all Aes happening before first dose date will be provided.

### 6.12.2 Laboratory Data

#### 6.12.2.1 Hb-related Safety Endpoints

Hb-related safety endpoints will be defined using data from the central laboratory. No imputation will be performed for missing data. For each endpoint in Table 3, the proportion of the study group who satisfy the definition and 95% CI for these proportions will be reported along with the difference in proportions as well as the odds ratio; a Mantel-Haenszel method stratified by the dialysis organization will be used to calculate the 95% CI for the statistics of treatment comparison.

**Table 3. Hb-related Safety Endpoints from First Dose Date to Last Visit**

| Variable | Criterion for "Yes" |
|---|---|
| Hb >11.0 g/dL | Any Hb >11.0 g/dL at any time after Day 1 |
| Hb >12.0 g/dL | Any Hb >12.0 g/dL at any time after Day 1 |
| Hb >13.0 g/dL | Any Hb >13.0 g/dL at any time after Day 1 |
| Hb >14.0 g/dL | Any Hb >14.0 g/dL at any time after Day 1 |
| Hb<7.0 g/dL | Any Hb <7.0 g/dL at any time after Day 1 |
| Hb<8.0 g/dL | Any Hb <8.0 g/dL at any time after Day 1 |
| Hb<9.0 g/dL | Any Hb <9.0 g/dL at any time after Day 1 |
| Hb<10.0 g/dL | Any Hb <10.0 g/dL at any time after Day 1 |
| Hb increase >1.0 g/dL within any 2-week interval | Difference between 2 Hb values within any 2 weeks >1.0 g/dL after Day 1 |
| Hb increase >2.0 g/dL within any 4-week interval | Difference between 2 Hb values within any 4 weeks >2.0 g/dL after Day 1 |

Note: Subjects without any post-baseline data will be excluded from this analysis.

#### 6.12.2.2 Liver Function Abnormality

A summary of liver function abnormalities by analysis period will be provided. Any subject with at least 1 of the following liver function abnormalities will be summarized:

- Alanine aminotransferase (ALT) >2x and <=3x upper limit of normal (ULN)
- ALT >3x and <=5x ULN;
- ALT >5x and <=10x ULN;
- ALT >10x ULN
- Aspartate aminotransferase (AST) >2x and <=3x ULN;
- AST >3x and <=5x ULN;
- AST >5x and <=10x ULN;
- AST >10x ULN
- Total bilirubin >2x and <=3x ULN;
- Total bilirubin >3x ULN

In addition, a table will summarize the occurrence of events that satisfy the following criteria of liver injury:

- (ALT or AST >3x ULN and <=5x ULN) and total bilirubin >2x ULN;
- (ALT or AST >5x ULN and <=10x ULN) and total bilirubin >2x ULN;
- ALT or AST >10x ULN and total bilirubin >2x ULN

#### 6.12.2.3 Other Laboratory Safety Endpoints

Descriptive statistics for laboratory values (in US conventional and SI units) and changes from baseline at each analysis visit will be summarized for the following laboratory parameters collected in the study including but not limited to the following:

- Complete blood count (CBC): Hb, hematocrit, red blood cell (RBC), MCV, MCH, MCHC, platelets, RDW, white blood cell (WBC) with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils), reticulocyte count
- Serum chemistry: sodium, potassium, bicarbonate, chloride, calcium, magnesium, phosphorus, glucose, creatinine, BUN, CPK, uric acid, albumin, total protein
- Liver function tests: Total bilirubin, ALP, ALT/SGPT, AST/SGOT, LDH
- Iron indices: Iron, ferritin, TSAT, TIBC
- Lipid Profile: Total cholesterol, LDL, HDL, Triglycerides
- CRP
- EPO

A summary of the number and percentage of subjects experiencing clinically significant values for the following parameter definitions will be presented by analysis period:

- Potassium >6.0 mmol/L

### 6.12.3 Physical Examination

With only the Date of Assessment collected (if performed) on the physical examination in the eCRF, no output is planned for it. If there are clinically significant abnormal findings, details can be found in the listings of medical history or adverse event.

### 6.12.4 Vital Signs

For blood pressure, a total of 2 measurements at intervals of at least 2 minutes will be performed. For change from baseline related summaries, the average of the blood pressure within the same visit will be calculated as the results for the visit. For the individual value criteria, an individual result will be used instead of the average.

Descriptive statistics for vital signs (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, temperature, weight (pre-dialysis and post-dialysis)) and their change from baseline will be provided by each analysis visit.

The number and percentage of subjects experiencing the following findings by analysis periods will be provided:

- SBP >=180 mmHg
- SBP >=160 mmHg
- DBP >=110 mmHg
- SBP or DBP: Increase from baseline >=20 mmHg

- Meet Both/Either Criteria
  - SBP >=160 mmHg and increase from baseline >=20 mmHg
  - DBP >=110 mmHg and increase from baseline >=20 mmHg

[redacted]

[redacted]

[redacted]

[redacted]

### 6.14 SUBGROUP ANALYSES

The same analyses of the primary efficacy endpoint and secondary efficacy endpoint will also be performed for subgroups based on the following. Subgroup analyses will be performed only if the total number of outcomes in a stratum, combined over the two treatment groups, is at least 30.

- Dialysis organization [USRC (U.S. Renal Care), ARCRS (American Renal Clinical Research Services), Frenova (Frenova Renal Research)]
- Age (<65 years, >=65 years)
- Sex (Male, Female)
- Race (White, Black, All others)
- Baseline Hb (<10.0 g/dL, >=10.0 g/dL)

# 7. CHANGES TO PROTOCOL-PLANNED ANALYSES

Removed the hierarchical testing scheme for both primary and secondary efficacy endpoints. Treatment comparisons will be performed between vadadustat (combining the 2 randomized groups with starting dose of 600mg and 900mg) and Mircera groups. Subjects in the vadadustat starting dose groups of 600 mg or 900 mg TIW will have their doses adjusted to maintain Hb levels of 10.0 to 11.0 g/dL following the protocol's dosing algorithm. The subgroup analysis with different starting dose with vadadustat vs. Mircera will be considered as exploratory analyses.

# 8. REFERENCES

Clopper, C. J., and Pearson, E. S. (1934). "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial." Biometrika 26:404–413.

FACIT Administration and Scoring Guidelines: http://www.ser.es/wp-content/uploads/2015/03/FACIT-F_INDICE.pdf

FDA Guidance for Industry: 'Non-inferiority Clinical Trials' March 2010. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials

Food and Drug Administration E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials Guidance for Industry. 2021. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9r1-statistical-principles-clinical-trials-addendum-estimands-and-sensitivity-analysis-clinical

Functional Assessment of Cancer Therapy – Anemia (FACT-An): https://www.facit.org/measures/FACT-An

Levin NW, Fishbane S, Cañedo FV, et al. Intravenous methoxy polyethylene glycol-epoetin beta for haemoglobin control in patients with chronic kidney disease who are on dialysis: a randomised non-inferiority trial (MAXIMA). *Lancet* 2007; 370:1415-21. Erratum in: *Lancet* 2008; 371:386.

Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Development of the kidney disease quality of life (KDQOL)instrument. Qual Life Res 1994, Oct;3(5):329-38

Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Kidney Disease Quality of Life Short Form (KDQOL-SFTM). A Manual for Use and Scoring. Ver. 1.3. Santa Monica: RAND; 1997. p. 7994

Mantel N, Haenszel W. Statistical aspects of analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; 22:719–48

Paganini EP, Eschbach JW, Lazarus JM, et al. Intravenous versus subcutaneous dosing of epoetin alfa in hemodialysis patients. Am J Kidney Dis 1995; 6:331-40.

Ware JE Jr, Kosinski M, Keller SD. How to Score the SF-12 Physical and Mental Health Summary Scales (3rd ed. Boston: The Health Institute, New England Medical Center, 1998.

[redacted]

[redacted]

zardratio 'Cox proportional hazard' TRT;
**RUN;**

[redacted]

[redacted]

For efficacy endpoints of having average Hb value in the target range in Weeks 20-26 and Weeks 46-52, these endpoints are binary; they will be analysed according to the Mantel-Haenszel estimate of stratified risk differences. These binary outcome measures will be computed from the continuous imputed data which will serve as the input data set for the computation of the estimate of risk difference. The code for this analysis is provided as below.

[redacted]

## B.3. Tipping Point Sensitivity Analysis for Missing Data

This multiple imputation model for the primary efficacy model described above assumes that data in both treatment arms are missing at random (MAR) and would therefore follow the trend of observed data. The tipping point analysis assesses the effect of potential deviations from this assumption and explores the consequences of assuming that data in the vadadustat arm are missing not at random (MNAR) (i.e., subjects in the vadadustat arm with missing outcome are assumed to have a lower Hb values than subjects in the Mircera arm). The specific steps are as follows:

1. The missing information on Hb in the vadadustat arm will be multiply imputed using a shift parameter S applied to lower the mean Hb values in that arm. Missing data in the Mircera arm will continue to be imputed assuming MAR, using the same code as for the primary efficacy analysis subset to the Mircera arm.
2. The multiply imputed data will be analysed using PROC MIANALYZE to obtain an estimated treatment effect and its associated 95% CI.
3. Steps 1 and 2 will then be repeated. The shift parameter S starts from 0, which corresponds to the primary efficacy analysis with no shift effect and decreased by 0.1 in each step until the analysis reaches the "tipping point" (i.e., lower limit of CI < -0.75), the point at which the effect of vadadustat is no longer noninferior to that of Mircera. The more the tipping point diverges from the observed data, the more robust the conclusion based on primary efficacy analysis.

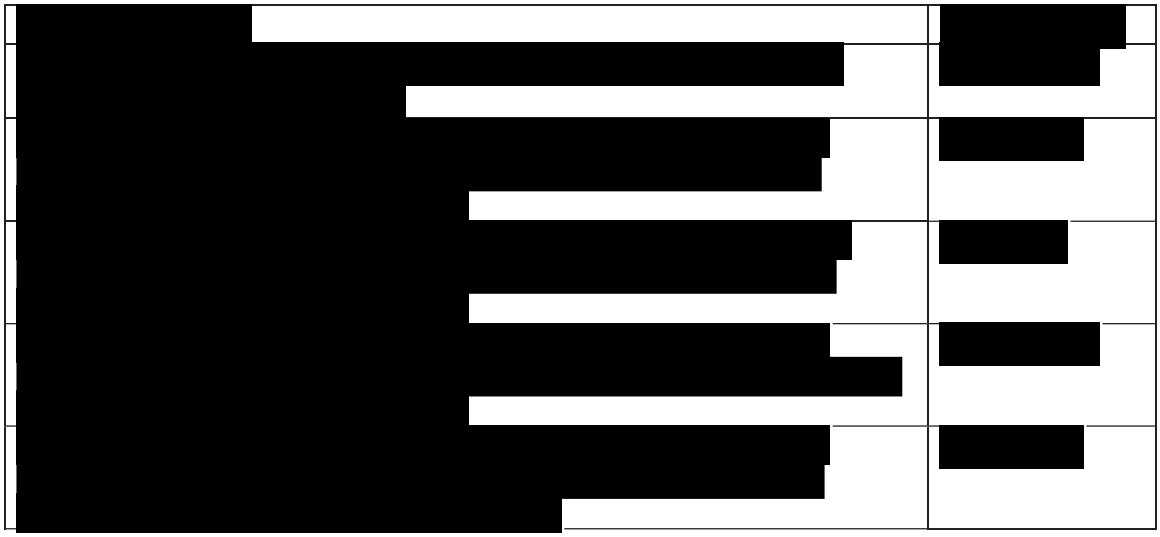
[redacted]

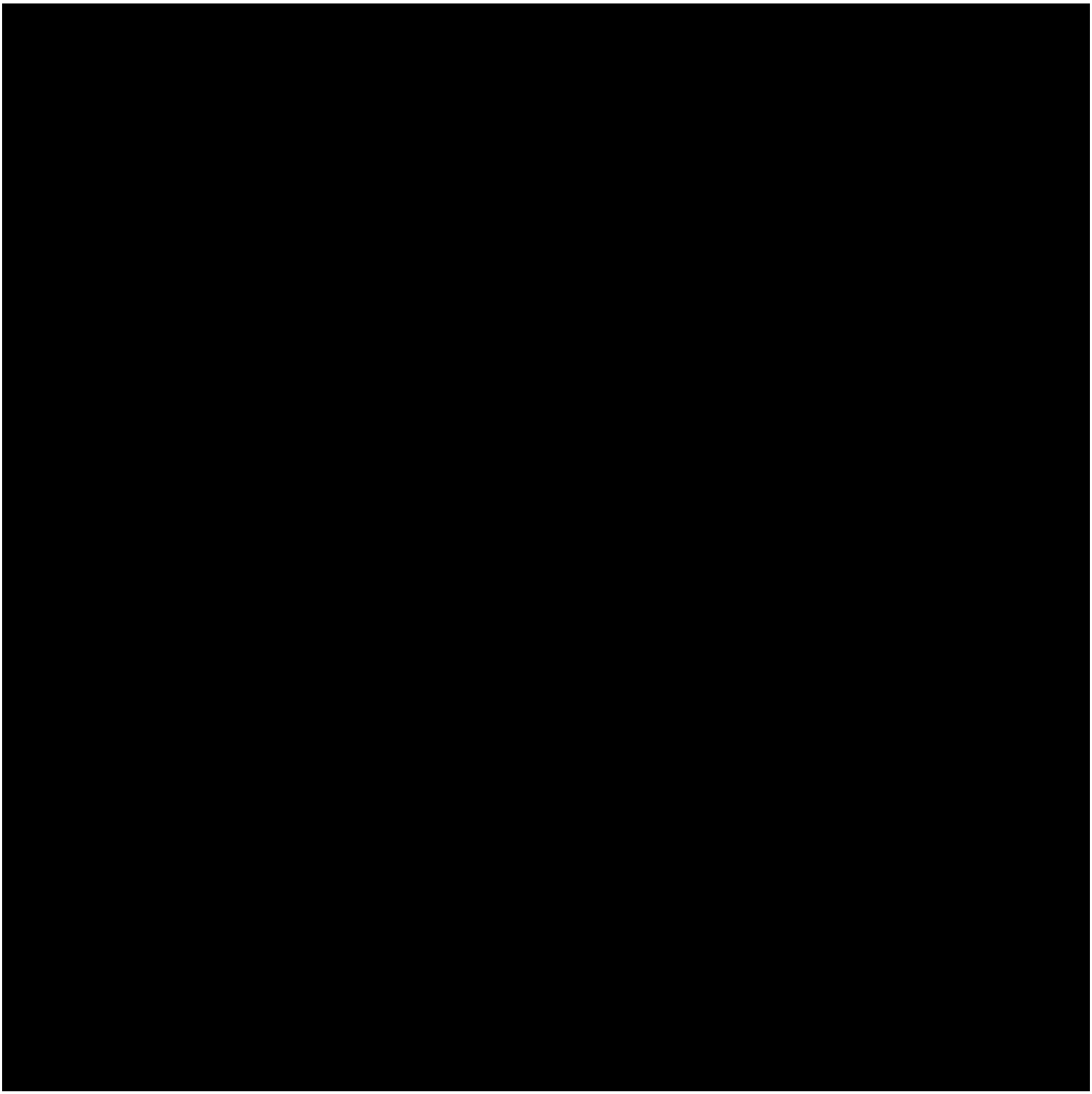
## B.4. Random Seed Specification

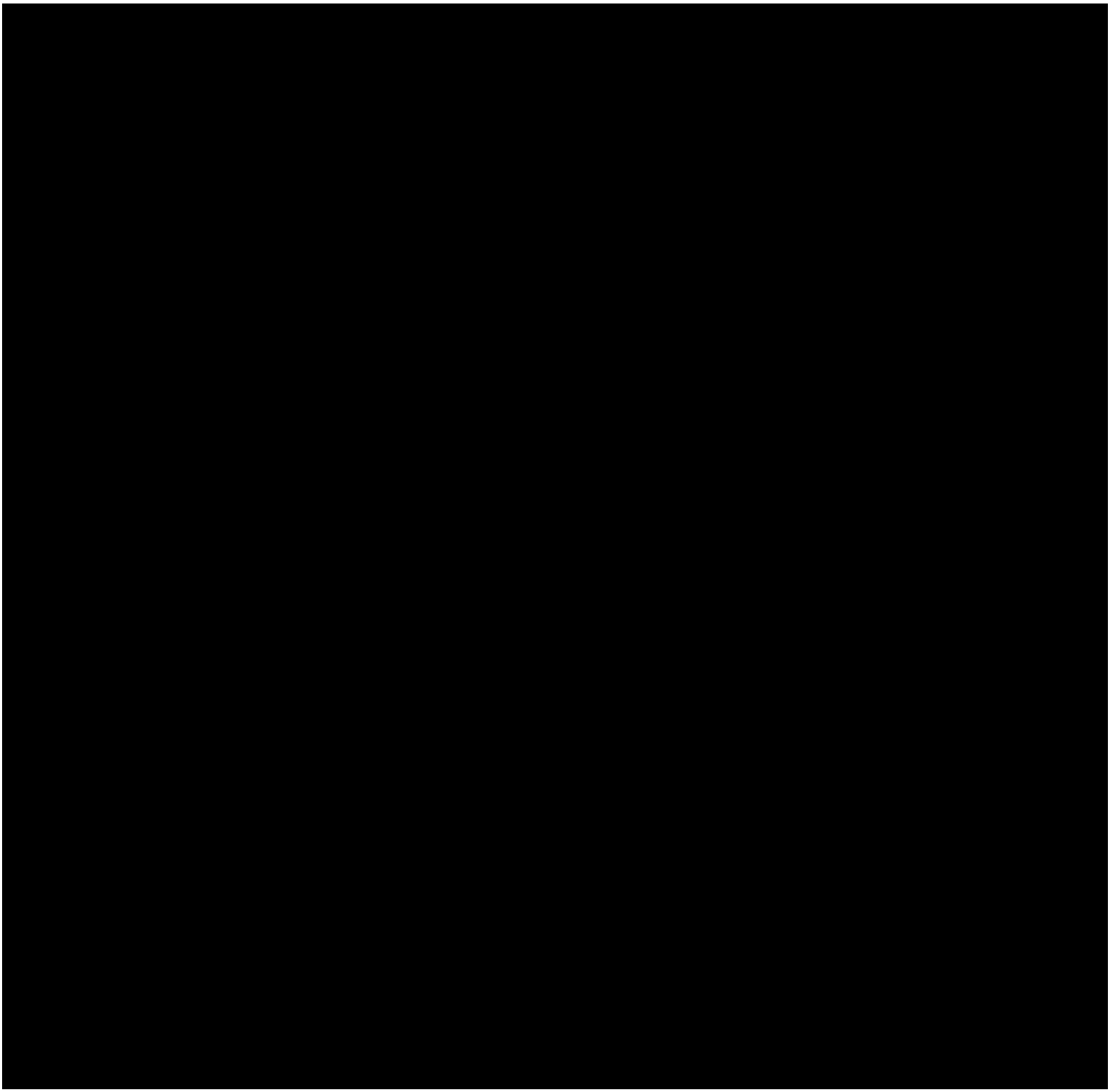
Assigned seeds in proc mi for multiple imputations will be generated randomly from a master seed 65480039. The SAS code and the generated seed numbers are listed below.

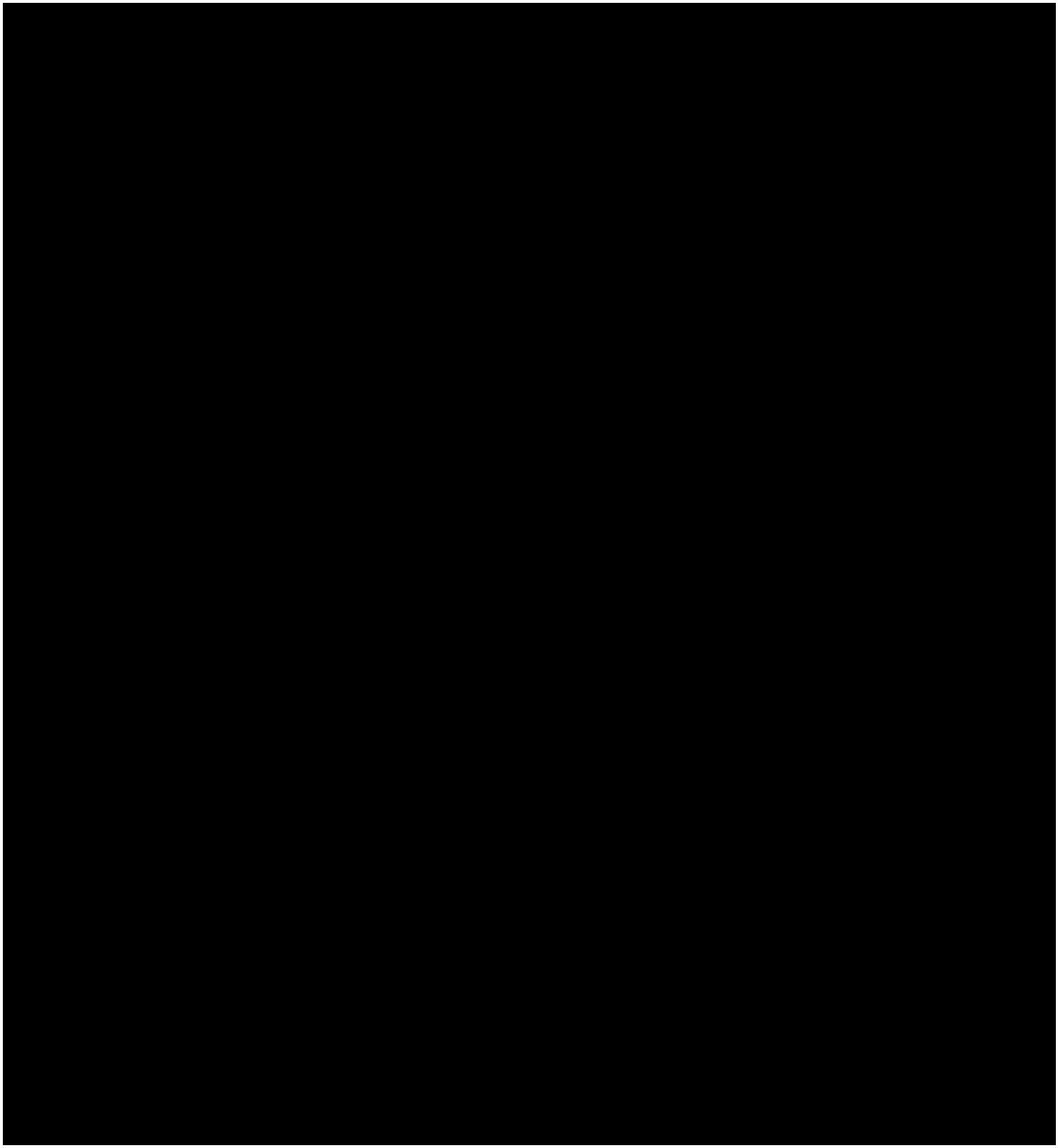
[redacted]

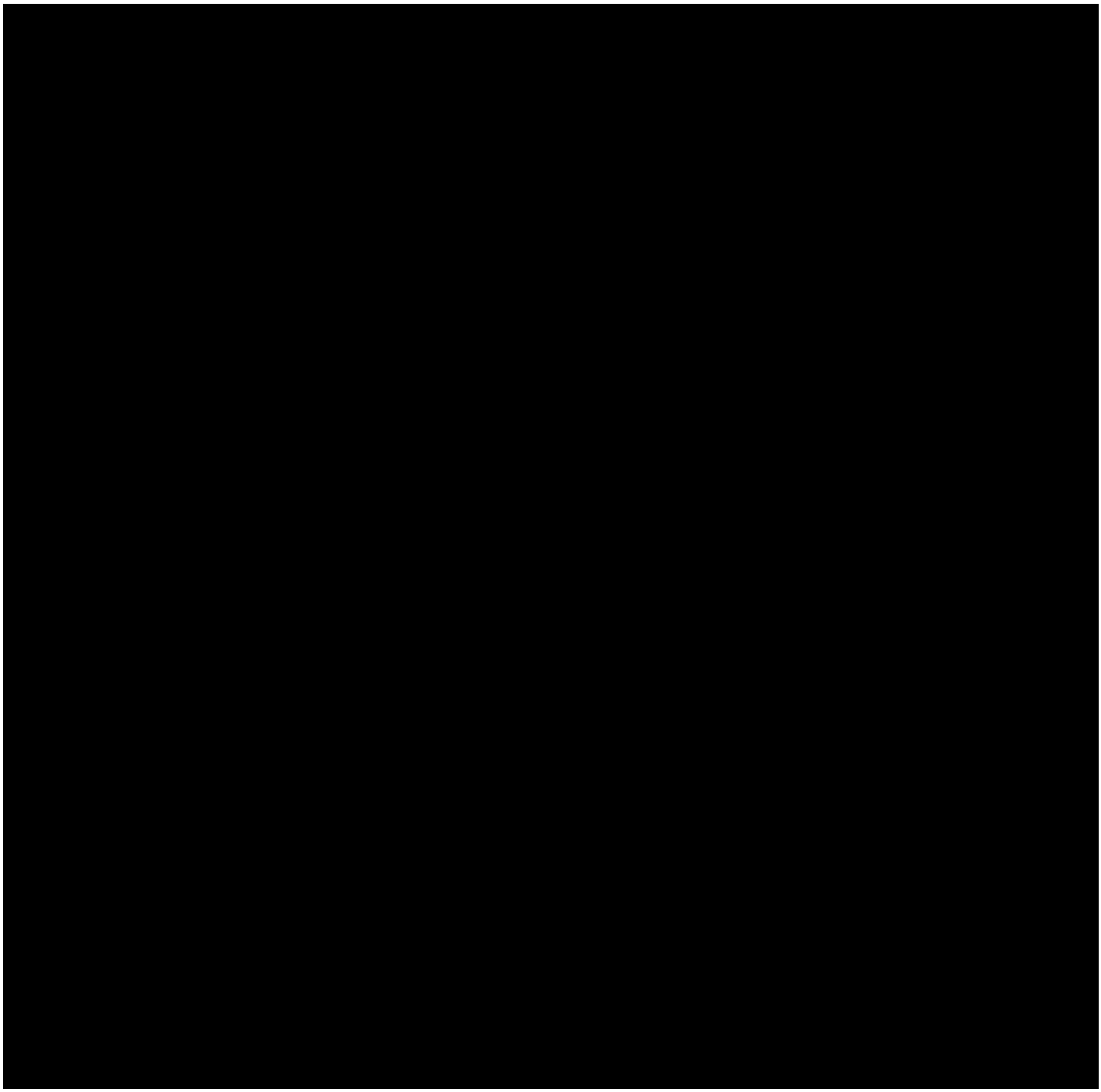
[redacted]

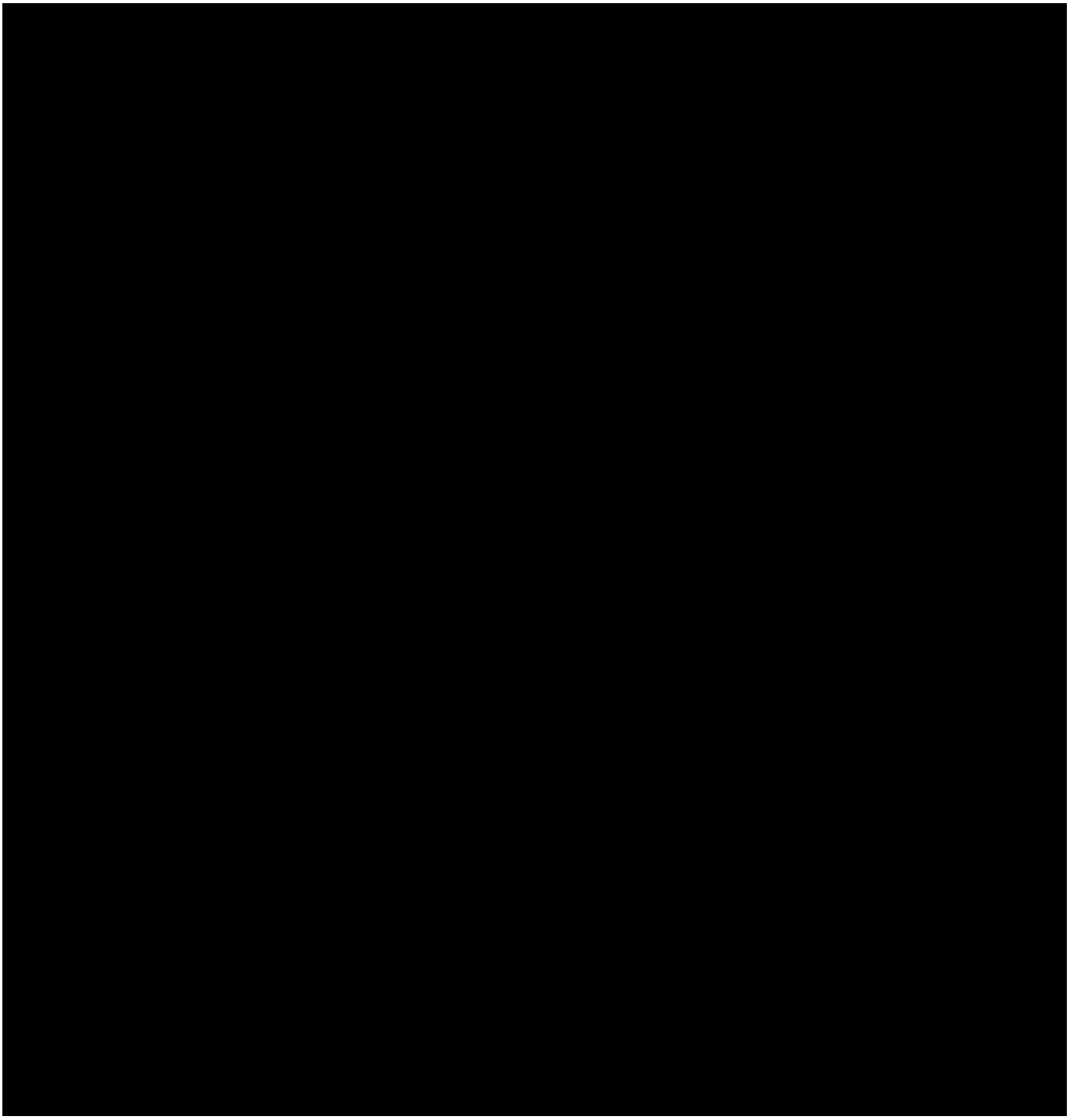


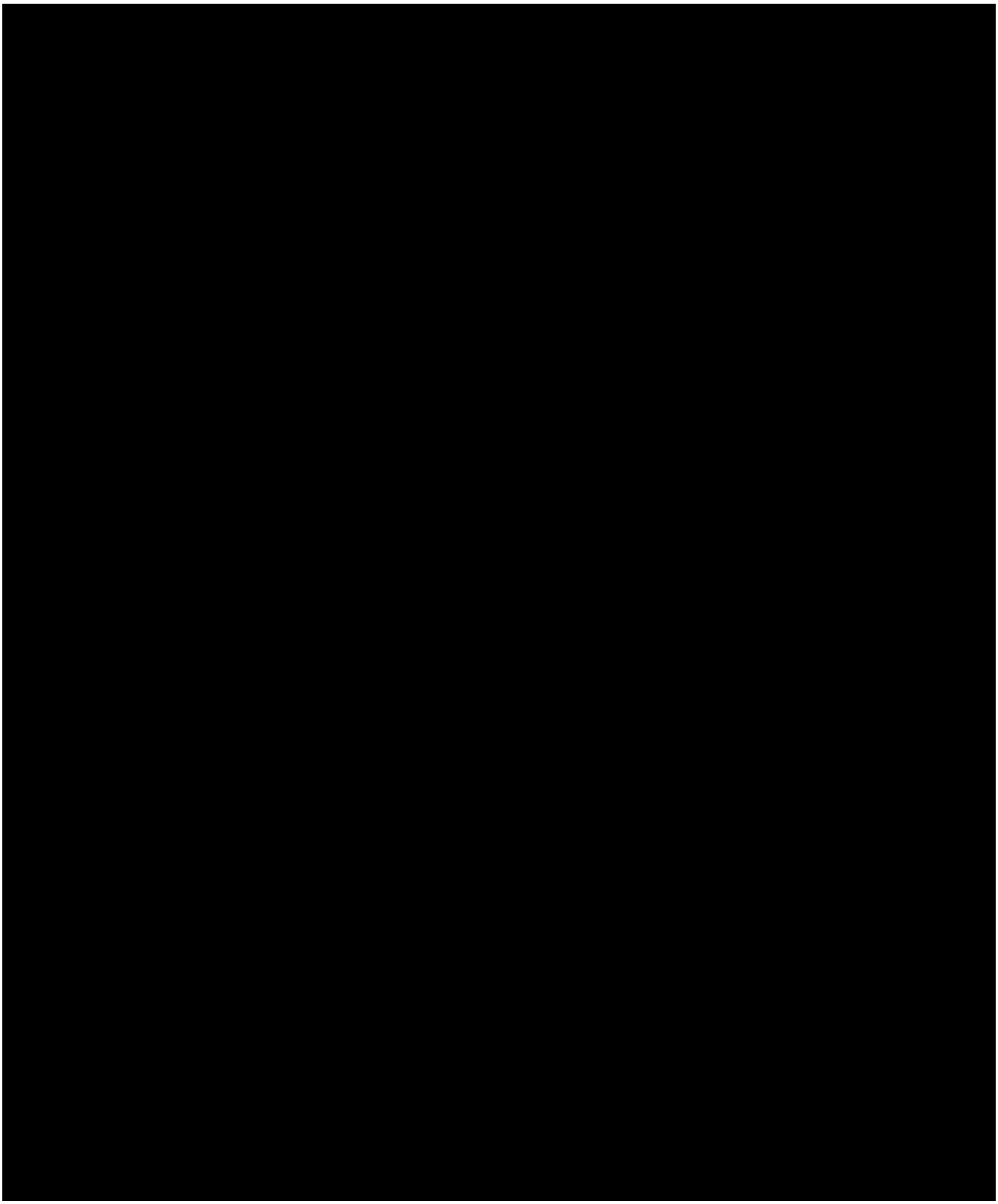












## Appendix E. Adverse Events of Special Interest (AESI)

The following items will be presented as AESI.

- Worsening of Hypertension: Hypertension MedDRA SMQ Narrow
- Hepatotoxicity: Drug related hepatic disorders Comprehensive SMQ broad
- Pulmonary Hypertension: Pulmonary Hypertension SMQ narrow
- Adrenal disorder: high-level group term (HLGT) Adrenal gland disorders not elsewhere classified (NEC), high-level term (HLT) Adrenal cortex tests
- Malignancies: Malignant or unspecified tumors MedDRA SMQ broad
- Congestive heart failure: Cardiac failure MedDRA SMQ Narrow
- Adverse event preferred terms used in key queries from FDA briefing document for Roxadustat are listed below:

| Medical Topic | PT Term |
|---|---|
| Thrombosis | Cerebral infarction |
| | Embolic cerebral infarction |
| | Ischaemic stroke |
| | Cerebellar infarction |
| | Lacunar stroke |
| | Embolic stroke |
| | Brain stem stroke |
| | Lacunar infarction |
| | Thrombosis in device |
| | Arteriovenous fistula thrombosis |
| | Arteriovenous graft thrombosis |
| | Vascular access site thrombosis |
| | Vascular graft thrombosis |
| | Graft thrombosis |
| | Shunt thrombosis |
| | Acute myocardial infarction |
| | Myocardial infarction |
| | Deep vein thrombosis |
| | Thrombosis |
| | Atrial thrombosis |
| | Peripheral artery thrombosis |
| | Subclavian vein thrombosis |
| | Brachiocephalic vein thrombosis |

| Medical Topic | PT Term |
|---|---|
| | Subclavian artery thrombosis |
| | Vena cava thrombosis |
| | Thrombophlebitis superficial |
| | Arterial thrombosis |
| | Thrombophlebitis |
| | Jugular vein thrombosis |
| | Venous thrombosis |
| | Pelvic venous thrombosis |
| | Venous thrombosis limb |
| | Cardiac ventricular thrombosis |
| | Intracardiac thrombus |
| Device/shunt thrombosis/occlusion/malfunction/Stenosis | Thrombosis in device |
| | Arteriovenous fistula thrombosis |
| | Arteriovenous graft thrombosis |
| | Vascular access site thrombosis |
| | Vascular graft thrombosis |
| | Medical device site thrombosis |
| | Device occlusion |
| | Arteriovenous fistula occlusion |
| | Vascular access site occlusion |
| | Vascular access complication |
| | Vascular access malfunction |
| | Arteriovenous graft site stenosis |
| | Shunt occlusion |
| | Shunt malfunction |
| | Vascular graft stenosis |
| | Anastomotic stenosis |
| | Vascular access site complication |
| | Vascular graft occlusion |
| Device/shunt thrombosis | Thrombosis in device |
| | Arteriovenous fistula thrombosis |
| | Arteriovenous graft thrombosis |
| | Vascular access site thrombosis |
| | Vascular graft thrombosis |

| Medical Topic | PT Term |
|---|---|
| | Graft thrombosis |
| | Shunt thrombosis |
| | Medical device site thrombosis |
| | Device related thrombosis |
| | Injection site thrombosis |
| Seizure FDA | Epilepsy |
| | Epileptic encephalopathy |
| | Seizure |
| | Generalised tonic-clonic seizure |
| | Idiopathic partial epilepsy |
| | Partial seizures |
| | Tonic convulsion |
| Stroke | Cerebral infarction |
| | Embolic cerebral infarction |
| | Ischaemic stroke |
| | Cerebellar infarction |
| | Lacunar stroke |
| | Embolic stroke |
| | Brain stem stroke |
| | Lacunar infarction |
| | Cerebrovascular accident |
| | Haemorrhagic stroke |
| | Brain stem haemorrhage |
| Sepsis/septic shock | Device related sepsis |
| | Enterococcal sepsis |
| | Sepsis |
| | Urosepsis |
| | Streptococcal sepsis |
| | Pseudomonal sepsis |
| | Staphylococcal sepsis |
| | Septic shock |
| | Sepsis syndrome |
| | Biliary sepsis |
| | Bacterial sepsis |

| Medical Topic | PT Term |
|---|---|
| | Fungal sepsis |
| | Citrobacter sepsis |
| | Listeria sepsis |
| | Abdominal sepsis |
| | Septic encephalopathy |
| | Escherichia sepsis |