Official Title	A Phase 2, Open Label, Randomized, Dose Ranging, Safety, Efficacy, Pharmacokinetic and Pharmacodynamic Study of AG-348 in Adult Patients With Pyruvate Kinase Deficiency
NCT ID	NCT02476916
Document Version and Date	Statistical Analysis Plan, Version 2.0, 21 August 2020

STATISTICAL ANALYSIS PLAN

A Phase 2, Open Label, Randomized, Dose Ranging, Safety, Efficacy, Pharmacokinetic and Pharmacodynamic Study of AG-348 in Adult Patients with Pyruvate Kinase Deficiency

AG348-C-003

Version: v2.0

Date: 21-Aug-2020

CCI

Prepared by:

PPD

Study Statistician

CCI

CCI

Name and Title
(Printed)

Date
(DD MMM YYYY)

Approved by:

PPD

VP, Clinical Development

CCI

CCI

Name and Title
(Printed)

Date
(DD MMM YYYY)

PPD

Head of Biometrics

CCI

CCI

Name and Title
(Printed)

Date
(DD MMM YYYY)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Definition
AE	Adverse event
AESI	Adverse event of special interest
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Class
BID	Twice daily
BMI	Body mass index
CI	Confidence Interval
CTCAE	Common Terminology Criteria for Adverse Events
DRT	Data Review Team
DXA	Dual-energy x-ray absorptiometry
ECG	Electrocardiogram
EOS	End of Study
EPO	Erythropoietin
EOT	End of Treatment
eCRF	Electronic case report form
Hb	Hemoglobin
HCT	Hematocrit
HDL-C	High-density lipoprotein-cholesterol
Hp	Haptoglobin
HLT	MedDRA High Level Term
LDH	Lactate dehydrogenase
LFT	Liver function test
LLN	Lower limit of normal
Max	Maximum value
MedDRA	Medical Dictionary for Regulatory Activities
Min	Minimum value
MRI	Magnetic resonance imaging
PK	Pyruvate kinase

Abbreviation	Definition
PKR	Pyruvate kinase isoform R
PT	Preferred Term
QTcB	Heart rate-corrected QT interval using the Bazett's formula
QTcF	Heart rate-corrected QT interval using the Fridericia's formula
RBC	Red blood cell
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SOC	System organ class
TEAE	Treatment-emergent adverse event
TIBC	Total iron binding capacity
ULN	Upper limit of normal
WHO	World Health Organization

1. VERSION HISTORY

This statistical analysis plan (SAP) describes the analysis for protocol AG348-C-003 Version 8.0 (dated 04-Sep-2019).

CCI

[redacted]

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in both Core Period and Extension Period of study AG348-C-003 except for

pharmacokinetic and pharmacodynamic (PD) data, which will be described in a separate SAP. This document may modify the plans outlined in the protocol.

The final clinical study report (CSR) will include all data from the start of the Core Period of the study up to the End of Study (EOS) for all subjects, which is defined as the time at which all subjects have completed the study, or have been lost to follow-up.

Additional analyses of the data may be performed for publication or regulatory reporting purposes. In the following sections, references to "data cutoff date/EOS date" are meant to indicate that the data cutoff date will be used for analyses to be reported in a CSR or for other regulatory reporting purposes before the end of the study, and the EOS date will be used for analyses to be reported in the final CSR.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Objectives

3.1.1. Primary Objective

The primary objective of the Core Period of the study is to:

- Evaluate the safety and tolerability of up to 24 weeks of AG-348 administration in subjects with pyruvate kinase deficiency (PK deficiency).

The primary objective of the Extension Period of the study is to:

- Evaluate the long-term safety and tolerability of AG-348 administration in subjects with PK deficiency.

3.1.2. Secondary Objectives

The secondary objectives of the Core Period of the study are to:

- Evaluate the pharmacokinetics of AG-348 and the metabolite AGI-8702.
- Evaluate the pharmacodynamic (PD) response of adenosine triphosphate (ATP) and 2,3-diphosphoglycerate (2,3-DPG) after administration of AG-348.
- Evaluate indicators of clinical activity of AG-348 in subjects with PK deficiency, including changes in hemoglobin (Hb), hematocrit (HCT), reticulocyte count, haptoglobin (Hp), carboxyhemoglobin (COHb), lactate dehydrogenase (LDH), total and indirect bilirubin, erythropoietin (EPO), hepcidin, ferritin, and transferrin saturation (serum iron/iron-binding capacity).

The secondary objectives of the Extension Period are to:

- Evaluate indicators of clinical activity of AG-348 in subjects with PK deficiency, including changes in Hb, HCT, reticulocyte count, Hp, COHb, LDH, total and indirect bilirubin, EPO, hepcidin, ferritin, and transferrin saturation (serum iron/iron-binding capacity).

- Evaluate the optimal maintenance dose of AG-348 for each individual subject during the Extension Period.

CCI

[redacted]

3.2. Endpoints

3.2.1. Safety Measures and Endpoints

Safety will be evaluated by all of the following:

- Monitoring of adverse events (AEs), including serious adverse events (SAEs), adverse events of special interest (AESIs), and AEs leading to discontinuation. All AEs will be graded using CTCAE, Version 4.03.
- Safety laboratory parameters (hematology, chemistry, urinalysis, coagulation).
- Physical examination findings (including neurological examination).
- Vital signs.
- 12-lead electrocardiograms (ECGs).
- DXA scans.
- Serum sex hormone levels (testosterone [total and free], estrone, and estradiol), bone turnover markers (serum osteocalcin-N-mid and serum C-terminal telopeptide [CTX]), 25-hydroxy vitamin D2 and D3, total cholesterol, high-density lipoprotein-cholesterol (HDL-C), and triglycerides will be monitored for evidence of potential inhibition of aromatase by AG-348.
- Menstruating female subjects will also keep a paper-based menstrual cycle diary throughout the Core and Extension Periods.

3.2.2. Clinical Activity Measures and Endpoints

- Monitoring of potential indicators of clinical activity will include evaluating changes in Hb, HCT, reticulocyte count, Hp, COHb, LDH, total and indirect bilirubin, EPO, hepcidin, ferritin, and transferrin saturation.

3.2.3. Pharmacokinetic and Pharmacodynamic Measures and Endpoints

The pharmacokinetic and PD profile of AG-348 will be evaluated by all of the following:

- Approximately the first 10 subjects treated during the Core Period, contingent on clinical site feasibility, will undergo extensive pharmacokinetic sampling. The remainder of treated subjects will undergo limited pharmacokinetic sampling.
- During the Core Period, serial blood sampling for determination of concentration-time profiles of AG-348 and its metabolite AGI-8702 will be conducted following the first dose and the morning Day 15 dose, and additional trough levels of AG-348 and AGI-8702 will be obtained.
- Pharmacodynamic assessments during the Core Period will include 2,3-DPG, ATP (secondary objectives), CCI [redacted] will only be conducted in clinical sites able to perform these assessments during the Core Period.

  Approximately the first 10 subjects treated during the Core Period will undergo extensive PD sampling. The remainder of treated subjects will undergo limited PD sampling.
- During the Core Period, serial blood sampling for determination of levels of ATP and 2,3-DPG will be conducted following the first dose and the morning Day 15 dose, and additional trough levels of ATP and 2,3-DPG will be obtained.

CCI [redacted]

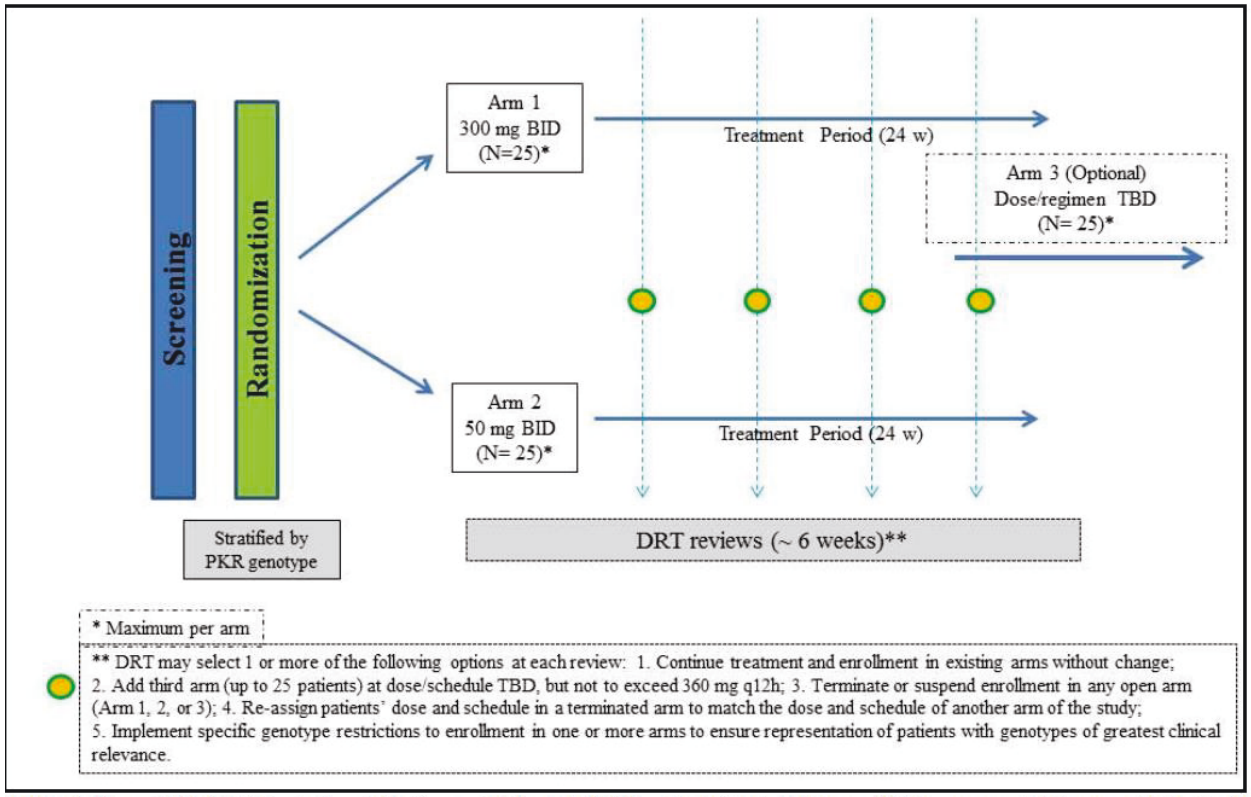
4. STUDY DESIGN

Study AG348-C-003 is a Phase 2, open-label, 2-arm, multicenter, randomized, dose-ranging study in adult subjects with PK deficiency. The study will be divided into a Core Period and an Extension Period.

- During the Core Period, subjects will receive multiple doses of AG-348 (hereinafter referred to as mitapivat) for up to 24 weeks.
- Subjects who are eligible can enter the Extension Period to receive mitapivat for up to 8 years following the end of the Core Period.

In the Core Period, subjects will be initially randomized to one of 2 BID doses of mitapivat as shown in Figure 1.

Figure 1: Study Schema: Core Period



Abbreviations: BID (q12h) = twice daily (every 12 hours); DRT = Data Review Team; PKR = pyruvate kinase red blood cell isoform; TBD = to be determined; w = week(s).

Because PK deficiency is a rare disease with a limited eligible subject population and because the underlying pathophysiology and clinical phenotype of affected subjects is heterogeneous due to the wide variety of mutations in PKR that cause the disease, it is deemed important to focus closely on dose findings in this first-in-patient study. Therefore, in addition to initiating this

study with 2 different doses of mitapivat administered BID, a DRT will review ongoing study data and adapt the study design, dose and schedule of mitapivat if indicated.

The DRT may exercise 1 or more of the following options during the Core Period:

- Continue treatment and enrollment in existing arms without change
- Terminate or suspend enrollment to allow further review of clinical data in Arm 1 and/or Arm 2 (and/or potential Arm 3).
- Re-assign subject's doses and schedule in a terminated arm to match the dose and schedule of another arm of the study. In this case, the subjects in the terminated arm will remain in their original arm, ie, they will not count towards the enrollment quota of the arm whose dose and schedule are being adopted
- Implement specific genotype restrictions to enrollment in 1 or more arms to ensure representation of subjects with genotypes of greatest clinical relevance

The DRT may exercise 1 or more of the following options during the Extension Period:

- Continue treatment without change.
- Re-assign subjects' doses and schedule to an existing dose and schedule that has been determined to be safer, and/or have a better PD response, or produce signs of clinical activity.
- Terminate or suspend treatment to allow further review of clinical data (eg, for unacceptable safety/tolerability, poor PD response, or lack of signs of clinical activity).

If Arm 3 is implemented, the dose of mitapivat selected will not exceed 360 mg BID. As of Protocol v8.0 there are no plans to add Arm 3 because the Core Period is complete.

5. ANALYSIS DATA SETS

Only subjects who signed informed consent will be included in the analysis set below.

- The Full Analysis Set (FAS) will include all subjects who are randomized. Subjects will be classified according to the treatment arm assigned at randomization.
- The Efficacy Analysis Set is a subset of the FAS and will include all subjects who received study drug for at least 3 weeks.
- The Safety Analysis Set will include all subjects who received at least 1 dose of study treatment. Subjects will be classified according to the treatment arm assigned at randomization if received at least once, or to the first treatment received if the assigned treatment at randomization is never received.

Table 2 summarizes the use of the analysis sets.

Table 2: Analysis Sets for Each Endpoint

Endpoints	Full Analysis Set	Efficacy Analysis Set	Safety Analysis Set
Demographic and other baseline characteristics	✓		
Disposition	✓		
Major protocol deviations	✓		
Exposure and concomitant therapies			✓
Efficacy	✓	✓	
Safety			✓

6. GENERAL STATISTICAL CONSIDERATIONS

6.1. Randomization, Blinding, Unblinding, and Crossover

Subjects will initially be randomized in a 1:1 ratio to one of the two treatment arms:

- Arm 1: Mitapivat 300 mg BID
- Arm 2: Mitapivat 50 mg BID

The protocol allowed for the possibility that an additional Arm 3 could be implemented with a dose of mitapivat not to exceed 360 mg BID; however, as noted in Section 4, as of protocol v8.0 Arm 3 will not be implemented.

The randomization will be stratified by PKR mutation (R510Q, R486W, R479H, other). Mutation status is defined by the presence of at least 1 of the indicated mutations; subjects with more than 1 stratified mutation will be assigned based on Sponsor's discretion.

Since this is an open-label study of mitapivat only, randomization will not be blinded.

The study allows subjects to receive a dose different from the dose assigned at randomization based on observed Hb concentrations.

6.2. Sample Size Determination and Decision Rules

6.2.1. Sample Size Determination

Due to the rare disease setting, the sample size in each treatment arm is determined by feasibility Up to 25 subjects may be randomized in each treatment arm. The actual number of subjects randomized will depend on the safety reviews and decisions made by the DRT.

Table 3 provides the probability within a dose arm of detecting 1 or more AE with varying sample size and the true underlying AE rates.

Table 3: Probability of Observing at Least 1 Specific AE Given Different Underlying AE Incidence Rates

Sample Size	True Underlying AE Rate		
	15%	**10%**	**5%**
5	56%	41%	23%
10	80%	65%	40%
15	91%	79%	54%
25	98%	93%	72%
50	> 99%	99%	92%

6.2.2. Decision Rules

There are no formal statistical decision rules in this study.

6.3. Definitions

6.3.1. Study Drug and Study Treatment

Both study drug and study treatment are defined as mitapivat.

There are 2 treatment arms in this study, as shown in Section 6.1:

- Arm 1: mitapivat 300 mg BID
- Arm 2: mitapivat 50 mg BID

6.3.2. Start and End Dates of Study Treatment

The start of study treatment is the earliest date/time of administration of a non-zero dose of the study treatment.

The end of study treatment is the latest date/time of administration of a non-zero dose of the study treatment on or before the data cutoff date/EOS date.

6.3.3. Study Day

The study day for assessments or events occurring on or after the start of study treatment (eg, AE onset, laboratory assessment) will be calculated as:

Study day = Date of the assessment or event − start of study treatment + 1.

The study day for assessments or events occurring before the start of study treatment (eg, lab assessment during the Screening Period, medical history) will be negative and calculated as:

Study day = Date of the assessment or event − start of study treatment.

There is no study day 0. The study day will be displayed in data listings.

6.3.4. Baseline

Efficacy Evaluations

For efficacy CCI laboratory parameters [Hb, HCT, reticulocyte count, Hp, COHb, LDH, total and indirect bilirubin, EPO, hepcidin, ferritin, and transferrin saturation (serum iron/iron-binding capacity)], baseline is defined as the average of all screening assessments within 45 (42+3) days before the date of randomization (including assessments on the date of randomization) for subjects randomized and not dosed or before the start of study treatment for subjects randomized and dosed. Assessments collected within 61 days after a transfusion will be excluded from the baseline derivation.

Baseline for efficacy laboratory parameters will be derived based on central laboratory data; if no central laboratory data are available before the start of study treatment, then local laboratory data will be used to derive the baseline.

Baseline Characteristics

For summaries of baseline characteristics based on the FAS, baseline will be defined as follows:

- For subjects randomized and not dosed: the last assessment on or before the date of randomization
- For subjects randomized and dosed: the last assessment on or before the start of study treatment

Safety Evaluations

For alanine aminotransferase (ALT) and aspartate aminotransferase (AST), baseline is defined as the average of all screening assessments collected within 45 (42+3) days before the start of study treatment. Baseline will be derived based on central laboratory data; if no central laboratory data are available before the start of study treatment, then local laboratory data will be used to derive the baseline.

For other laboratory assessments:

- Prior to deriving the baseline,
  - If there are multiple records with the same assessment day and time from the same laboratory, the average value will be used
  - If there are multiple records with the same assessment day and time from different laboratories, the value from the central laboratory will be used
- The baseline will then be the last value on or before the start of study treatment.

For all other safety parameters, the last assessment on or before the start of study treatment will be used as the baseline.

If, per protocol, an assessment (efficacy, baseline characteristic, or safety) is to be performed on study day 1, before the first dose of study treatment, and the assessment time, time of first dose of study treatment, or both, is missing (or not collected), it will be assumed that the assessment is performed before study treatment administration. Unscheduled assessments will be used in the

determination of baseline; however, an unscheduled assessment on study day 1 will be considered to have been obtained after study treatment administration.

If no assessment meets the definition of baseline for an evaluation (efficacy, baseline characteristic, or safety), the baseline will be set to missing.

6.3.5. On-Treatment Period

The on-treatment period starts on the date of the start of study treatment and ends 30 days after the end of study treatment.

Within the on-treatment period the following dosing periods are defined:

- **Core Period** starts on the date of start of study treatment and
  - Ends 30 days after the end of study treatment, if the subject does not enter the Extension Period
  - Ends on the end of treatment date of the Core Period, recorded in the Core Period EOT eCRF, if the subject enters the Extension Period
- **Extension Period** starts 1 day after end of treatment date of the Core Period and ends 30 days after the end of study treatment**.**

On-treatment period is the Cumulative Period with Core Period and Extension Period combined.

6.4. General Methods

6.4.1. Data Handling After Cutoff Date

For analyses of the data prior to EOS, the data after the cutoff date may not undergo the cleaning process and will not be displayed in any listings or used for summary statistics, statistical analyses, or imputations.

6.4.2. Standard Derivations and Reporting Conventions

The following conversion factors will be used to convert days into weeks, months or years: 1 week=7 days, 1 month=30.4375 days, and 1 year=365.25 days.

The following derivations will be implemented.

- Age (years):
  - (date of informed consent − date of birth+1)/365.25
  - If only day of birth is missing: Age (years): (year/month of informed consent−year/month of birth)
  - If day and month of birth are missing: Age (years): (year of informed consent−year of birth)

  The integer part of the calculated age will be used for reporting purposes.
- Body mass index (BMI; kg/m^2)=weight (kg)/height $(m)^2$

- Duration (in days) from a reference date (eg, randomization date, start date of study treatment)
  - date of event − reference date + 1, if the date of the event is on or after the reference date
  - date of event − reference date, if the date of the event is before the reference date

Reporting conventions will be as follows:

- Mean and median will be displayed to one more decimal place than the raw data.
- Standard deviation (SD) will be displayed to two more decimal places than the raw data.
- Percentages will be displayed to 1 decimal place (however, percentages corresponding to 0 counts will be reported as 0 rather than 0.0 and 100 percent will be reported as 100 rather than 100.0).
- Unless otherwise specified, rounding will be performed to the closest integer/first decimal using the common mid-point between the two consecutive values, eg, 5.11 to 5.14 will be rounded to 5.1, and 5.15 to 5.19 will be rounded to 5.2.
  - Non-zero percentages that are <0.1 before rounding will be displayed as "<0.1", eg, 0.09 will be reported as <0.1 rather than as 0.1.

6.4.3. Pooling of Data Across Sites

In order to provide overall estimates of treatment effects, data will be pooled across sites. The "site" factor will not be considered in statistical models or subgroup analyses given the high number of participating sites in contrast to the anticipated small number of subjects randomized at each site.

6.4.4. Continuous and Categorical Variables

Continuous variables will be summarized using descriptive statistics, ie, number of non-missing values, mean, SD, median, quartiles, minimum, and maximum. Categorical variables will be summarized by frequency distributions (number and percentage of subjects within a given category in the analysis data set). Unless otherwise specified, the calculation of percentages will include the "missing" category. Therefore, counts of missing observations will be included in the denominator and presented as a separate category. For summaries by visit, percentages will be based on the number of subjects with data available for that visit, unless otherwise specified.

6.4.5. Unscheduled Visits

Generally, data collected at unscheduled visits will be included and summarized for both safety and efficacy analyses in the same manner as the data collected at scheduled visits. Data collected at unscheduled visits will be included in by-subject listings together with the data collected at scheduled visits.

Summaries of outliers (eg, worst value, worst change from baseline, worst Common Terminology Criteria for Adverse Events [CTCAE] grade) during the on-treatment period for safety endpoints such as laboratory measurements and ECG parameters will include data from both scheduled and unscheduled visits.

Individual longitudinal plots for laboratory measurements during the on-treatment period will include data from both scheduled and unscheduled visits.

Descriptive statistics (mean, SD, median, quartiles, minimum, maximum) by nominal visit will only be provided for DXA scans results. Data collected at unscheduled and scheduled postbaseline visits will be mapped to scheduled visits using analysis visit windows, and then values at scheduled postbaseline visits will be derived based on the rules described below.

For efficacy CCI endpoints [Hb, HCT, reticulocyte count, reticulocyte percentage, Hp, COHb, LDH, total and indirect bilirubin, EPO, hepcidin, ferritin, and transferrin saturation (serum iron/iron-binding capacity)], data collected at unscheduled and scheduled postbaseline visits will be mapped to scheduled visits using analysis visit windows, and then values at scheduled postbaseline visits will be derived based on the rules described below. Descriptive statistics by nominal visit and longitudinal plots during the on-treatment period for efficacy endpoints such as Hb concentration will be provided using the derived values at scheduled visits.

Analysis Visit Windows

For the evaluation of Hb, hemolysis markers (indirect bilirubin, LDH, and Hp), iron markers [serum iron, serum ferritin, total iron-binding capacity (TIBC), transferrin saturation, and hepcidin], erythropoietic markers (reticulocyte count, reticulocyte percentage, and EPO), HCT, and COHb, the analysis visit windows will be derived based on the target study day for the scheduled visits as follows. Note that based on the scheduled of assessments, a Week 12 Visit, for example, will have a target study day of 1+(12×7)=85.

- Visit windows will be implemented for scheduled visits after Day 1.
- For analysis visit(n):
  - Start day of visit window =1+end day of window for visit(n−1). If n=1, start day of the visit window is study day 2.
  - End day of visit window = [(target day for analysis visit(n) + target day for analysis visit(n+1))/2]–1 except for the last scheduled visit. The end day of the last visit window on or before the data cutoff date is the min(cutoff date/EOS date, end of on-treatment period).

For DXA scan results the analysis visit window for the first scheduled assessment at Week 24 Visit will start on study day 86. The derivation for the end day of Week 24 Visit and the remaining visit windows will follow the same rule of deriving “analysis visit(n)” specified above.

Derivation of Values at Scheduled Postbaseline Visits Based on Analysis Visit Windows

For efficacy laboratory parameters [Hb, hemolysis markers (indirect bilirubin, LDH, and Hp), iron markers (serum iron, serum ferritin, TIBC, transferrin saturation, and hepcidin)],

erythropoietic markers (reticulocyte count, reticulocyte percentage, and EPO)], HCT, and COHb, any assessments obtained within 61 days after a transfusion will be excluded. In addition:

- Central laboratory assessment(s) (scheduled or unscheduled) within the visit windows will be used
- If no central laboratory value is within the visit window, local laboratory assessment(s) within the visit window will be used

If multiple assessments are identified within a visit window for a parameter, the following rules will be applied:

- The assessment measured closest to the target study day of the scheduled visit will be used
- If there are multiple assessments equidistant to the target study day, the average value will be used for efficacy laboratory parameters

6.5. Methods for Handling Missing Data

6.5.1. Adverse Event and Concomitant Medication Start Dates

If the end date is non-missing and the imputed start date is after the end date, the end date will be used as the start date.

(1) Missing day only

- If the month and year are the same as the month and year of the date of the start of study treatment, the date of the start of study treatment will be used.
- If the month and year are before the month and year of the date of the start of study treatment, the last day of the month will be used.
- If the month and year are after the month and year of the date of the start of study treatment, the first day of the month will be used.

(2) Missing day and month

- If the year is the same as the year of the date of the start of study treatment, the date of the start of study treatment will be used.
- If the year is before the year of the date of the start of study treatment, 31 December will be used.
- If the year is after the year of the date of the start of study treatment, 01 January will be used.

(3) Missing day, month, and year

- The date of the start of study treatment will be used.

6.5.2. Adverse Event and Concomitant Medication End Dates

If the start date is non-missing and the imputed end date is before the start date, the start date will be used as the end date. If an imputation for an AE end date results in an AE end date that is after the data cutoff date/EOS date, the AE will be considered as ongoing at the data cutoff date/EOS date.

(1) Missing day only

- The last day of the month will be used.

(2) Missing day and month

- 31 December will be used.

(3) Missing day, month, and year

- The event will be regarded as ongoing.

6.5.3. Exposure

No imputation will be done for the date of the first dose of study drug.

If the date of the last dose of study drug is missing or partially missing, it will be imputed as follows (separately for each study drug):

- If the last date of study drug is completely missing and there is no End of Treatment Disposition eCRF page for the study drug AND there is no death date, the subject should be considered to be ongoing and the data cutoff date/EOS date for the analysis will be used as the last dosing date.
- If the last date of study drug is completely or partially missing and there is EITHER an End of Treatment Disposition eCRF page for the study drug OR a death date (on or before the data cutoff date/EOS date), then the imputed last dose date is:

  =Last day of the year, if only the year is available and Year <Year of min(EOT date, death date)

  =Last day of the month, if both the year and month are available and Year=Year of min(EOT date, death date) and Month <Month of min(EOT date, death date)

  =min(EOT date, death date), for all other cases

7. STATISTICAL ANALYSES

All summaries will be tabulated by treatment arm (Section 6.3.1) and overall (both treatment arms combined), unless otherwise specified.

7.1. Subject Disposition

For all subjects screened in the study, the following will be summarized:

- Number of subjects screened in the study

- Frequency (number and percentage) of subjects who discontinued the study before randomization, overall and by reason for discontinuation. Percentages will be calculated based on the number of subjects screened in the study.

In addition, the frequency of subjects in each of the analysis sets described in Section 5 will be summarized by treatment arm. Percentages will be calculated only for analysis sets that are a subset of the FAS or a subset of the safety analysis set.

The following summaries will be presented by treatment arm and for both treatment arms combined based on the FAS:

- Frequency of subjects in each randomization strata
- Frequency of subjects randomized in each geographic region, country, and site
- Frequency of subjects randomized and not treated, overall and by reason for discontinuation
- For each of the Core and Extension Periods, separately
  - Frequency of subjects with study drug ongoing
  - Frequency of subjects who discontinued study drug, overall and by the reason for discontinuation of study drug
- Frequency of subjects who completed the study
- Frequency of subjects ongoing in the study
- Frequency of subjects who discontinued the study, overall and by the reason for study discontinuation

The frequency of subjects with disposition reason, in each epoch, due to reasons associated with COVID-19 will further be summarized under the main reason for discontinuation.

Disposition for all screened subjects and randomization data will be provided in by-subject listings.

7.2. Protocol Deviations

All major protocol deviations that impact the safety of the subjects, the conduct of the study or the evaluation of the study results will be reported based on the FAS. These will include

- Subjects randomized despite not satisfying the eligibility criteria
- Subjects who develop withdrawal criteria while on the study but are not withdrawn
- Subjects who are randomized under the wrong stratification factor(s)
- Subjects who receive an excluded concomitant medication

In addition, for each category of major protocol deviations, those related to COVID-19 will also be summarized.

Major protocol deviations will be provided in a by-subject listing.

7.3. Demographic and Other Baseline Characteristics

The following summaries will be presented based on the FAS, unless otherwise specified.

7.3.1. Demographics and Physical Measurements

Demographic characteristics and physical measurements at baseline will be summarized as follows:

- Demographic characteristics
  - Sex: Male, Female (child bearing potential status will be summarized for female subjects)
  - Race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, other, unknown
  - Ethnic origin: Hispanic or Latino, Not Hispanic or Latino, not reported
  - Age (years): summary statistics
  - Age categories:
    - <65, ≥65 years
    - <35, ≥35 years
  - Physical measurements
  - Height (cm)
  - Weight (kg)
  - BMI (kg/m2)

Demographic data for all screened subjects will be provided in a by-subject listing.

7.3.2. Disease Characteristics

The following baseline characteristics of the underlying disease will be summarized based on the data entered in the eCRF:

- Baseline Hb concentration [both continuously and by categories (<8.5 g/dL, ≥8.5 g/dL)]
- Mutation class (missense/missense, missense/non-missense, non-missense/non-missense)
- Baseline ferritin
- DXA scan results by location (femoral total and adjusted spine): Bone mineral density (BMD) and their corresponding T-scores and Z-scores. Frequency of subjects with T-scores in 3 categories (≤-2.5, >-2.5-<-1, ≥-1.0)
- Prior splenectomy status (Yes, No; if Yes, age of splenectomy)

- Prior cholecystectomy status (Yes, No; if Yes, age of cholecystectomy)
- Prior chelation status (Yes, No); the status is "Yes" if a subject has received chelation therapy within 52 weeks (364 days) before the first dose of study treatment.

Data on disease characteristics will be provided in by-subject listings.

7.3.3. Medical History

Medical history will be summarized in frequency tabulations according to the latest version of the Medical Dictionary for Regulatory Activities (MedDRA) by System Organ Class (SOC) and Preferred Term (PT).

Medical history and surgical history of splenectomy will be provided in by-subject listings.

In addition, the medical condition collected (from Targeted Medical History eCRF) will be summarized.

7.3.4. Prior Therapies

The following summaries will be presented based on the safety analysis set.

Prior medications are defined as medications (from the Prior and Concomitant Medications eCRF) that are started before the start of study treatment.

All non-study medications will be coded according to the ATC classification and PT using the latest version of the World Health Organization (WHO) Drug Dictionary Enhanced. All prior medications will be summarized in frequency tabulations according to WHO ATC third level and PT.

Prior medications will be provided in a by-subject listing.

Prior transfusions are collected from the Transfusion History eCRF.

Transfusion history will be provided in a by-subject listing.

7.4. Exposure to Study Drug and Compliance

The following summaries will be presented based on the safety analysis set, separately for the Core Period and for the Cumulative Period, unless otherwise specified.

7.4.1. Treatment Duration and Exposure

Exposure will be summarized as dose received [cumulative dose, and as dose received relative to planned dose (RDI)]. Duration of exposure for mitapivat will be summarized as a continuous variable as well as in categories (>0 - 4, >4 - 8, >8 - 12, >12 - 20, >20 – 24 weeks, and in 12-week intervals beyond 24 weeks).

- Duration of exposure (days)=last dose date−first dose date+1
- Actual cumulative dose (mg)=sum of actual total daily doses
- Planned cumulative dose (mg)=sum of planned total daily doses

  - Actual total daily dose is considered
    - correct if the "Explanation" on the "Study Medication Compliance" eCRF is not "Patient Missed Dose(s)" or "Incorrect Dose"
    - incorrect, otherwise
  - Planned total daily dose=
    - actual total daily dose, if correct total daily dose is taken
    - most recent previous correct actual total daily dose, if incorrect total daily dose is taken
- RDI (%) = 100×actual cumulative dose (mg)/planned cumulative dose (mg)

7.4.2. Dose Modifications

The summary of dose modifications will include:

- The frequency of subjects with at least 1 dose modification, separately for dose increase, dose reduced, dose interrupted
- Summary of reasons for dose modification, separately for dose increase, dose reduced, dose interrupted

Dose modifications and reasons are collected on the Study Medication Compliance eCRF.

7.4.3. Optimal Maintenance Dose

Subjects who have had dose taper to identify their optimal maintenance dose will be summarized by their starting dose before the dose taper and their first dose that they were on for more than 3 weeks after dose taper.

7.5. Concomitant Therapies

The following summaries will be presented based on the safety analysis set.

Concomitant medications are defined as non-study medications (from the Prior and Concomitant Medications eCRF) that are started during the on-treatment period or are started before the start of the study treatment and end or remain ongoing during the on-treatment period.

All non-study medications will be coded according to ATC code and PT using the latest version of the WHO Drug Dictionary. All concomitant medications will be summarized in frequency tabulations according to WHO ATC third level and PT.

Concomitant procedures are defined as procedures (from the Prior and Concomitant Procedures eCRF) that are started during the on-treatment period or are started before the start of the study treatment and end or remain ongoing during the on-treatment period.

The concomitant procedures will be coded by the latest version of MedDRA by SOC and PT and will be summarized in frequency tabulations by SOC and PT.

Concomitant transfusions are collected in the "On Study Transfusions" eCRF page.

Concomitant transfusions will be provided in a by-subject listing.

7.6. Efficacy Analyses

The following analyses will be based on the FAS and the EAS, unless otherwise specified.

7.6.1. Hb Response in the Core Period

Hb response is defined as postbaseline change from baseline in Hb ≥1.5 g/dL at >50% assessments in the Core Period, excluding those within 2 months (61 days) of transfusion.

This definition ensures that neither single increases in Hb nor oscillating increases in the 1.5 g/dL range will be counted as Hb responses. Because there are 9 scheduled postbaseline assessments in the Core Period (ie, Weeks 1, 2, 3, 6, 9, 12, 16, 20, and 24), subjects must have Hb values ≥1.5 g/dL above baseline for at least 5 scheduled assessments (assuming total compliance with protocol-specified visits, without any unscheduled visits) to be considered Hb responders (ie, achieving an Hb response).

The frequency of subjects with an Hb response will be summarized along with the 2-sided 95% exact CI using the Clopper-Pearson method (exact CI for a binomial proportion as computed by default by the FREQ procedure using the EXACT option).

7.6.2. Other Indicators of Clinical Activity

Hb, hemolysis markers (indirect bilirubin, LDH, and Hp), iron markers (serum iron, serum ferritin, TIBC, transferrin saturation and hepcidin), and erythropoietic markers (reticulocyte count, reticulocyte percentage, and EPO), HCT, and COHb will be summarized by visit.

By-subject longitudinal plots will be presented with Hb at baseline and postbaseline, and actual dose over time. The plots will further include age, sex, race, PKR mutation (R510Q, R486W, R479H, other), mutation class (missense/missense, missense/non-missense, non-missense/non-missense), prior splenectomy status, baseline chelation status, and postbaseline chelation status.

Similar plots will be presented separately for reticulocyte percentage, markers of hemolysis (indirect bilirubin, LDH, and Hp), iron markers (serum iron, serum ferritin, TIBC, transferrin saturation and hepcidin), and erythropoietic markers (reticulocyte counts and EPO).

7.7. Safety Analysis

Summaries of safety data will be presented based on the safety analysis set, separately for the Core Period and for the on-treatment (Cumulative) Period, unless otherwise specified.

7.7.1. Adverse Events

Treatment-emergent adverse events (TEAEs) are AEs with a first onset date during the on-treatment period or worsening from baseline. All summaries described below will be based on TEAEs, if not otherwise specified.

All AEs will be listed by subject and AEs with onset outside of the on-treatment period will be flagged in the listings. Unless otherwise specified, TEAEs will be summarized according to the

latest version of MedDRA by SOC and/or PT, severity (based on CTCAE v4.03 grading), seriousness, and relation to study treatment in decreasing frequency based on the frequencies observed overall.

Each subject will be counted only once within each SOC or PT. If a subject experiences multiple TEAEs under the same PT within a SOC for the same summary period, only the TEAE assessed as related or with the worst severity, as applicable, will be included in the summaries of relationship and severity. If a subject has TEAEs with missing and non-missing grades, the maximum of the non-missing grades will be displayed. No imputation of missing grades will be performed.

The following will be summarized:

- TEAEs by SOC and PT
- TEAEs by SOC, PT, and worst grade
- Most common TEAEs and Grade ≥3 TEAEs by PT; these will include TEAEs (any grade) reported in ≥10% of subjects in either treatment arm or Grade ≥3 TEAEs reported in ≥5% of subjects in either treatment arm. These thresholds may be changed based on the observed data without an amendment to this SAP.
- Treatment-related TEAEs by SOC and PT
- Treatment-related TEAEs by SOC, PT, and worst grade
- Grade ≥3 TEAEs, by SOC and PT
- Treatment-related Grade ≥3 TEAEs, by SOC and PT
- Serious TEAEs, by SOC and PT
- Treatment-related Serious TEAEs, by SOC and PT
- TEAEs leading to discontinuation of study drug, by SOC and PT
- TEAEs leading to interruption of study drug, by SOC and PT
- TEAEs leading to dose reduction, by SOC and PT
- TEAEs leading to death, by SOC and PT
- Treatment-related TEAEs leading to death, by SOC and PT

In addition, the following will be summarized by actual dose at TEAE onset.

- First occurrence of TEAEs by PT
- First occurrence of serious TEAEs by PT
- TEAEs by PT
- Serious TEAEs by PT

7.7.1.1. Adverse Events of Special Interest

Transaminase increase is an AESI for mitapivat and will be reported by the investigator in the AESI eCRF page if there is a transaminase increase of >2.5×baseline or an increase in AST or ALT to Grade ≥2 in severity, whichever is lower.

Additional TEAEs of interest for mitapivat are as follows:

- AEs of endocrinological interest (identified based on the criteria outlined in the mitapivat program specified Safety Search Criteria)
- Insomnia (PTs under HLT of "Disturbances in Initiating and Maintaining Sleep" or identified based on the criteria outlined in the mitapivat program specified Safety Search Criteria)

The following will be summarized for AESIs and the additional TEAEs of interest:

- AESIs/TEAEs of interest by PT
- AESIs/TEAEs of interest by PT and worst grade
- Grade ≥3 AESIs/TEAEs of interest by PT
- AESIs/TEAEs of interest leading to discontinuation of study drug by PT
- Serious AESIs/TEAEs of interest by PT
- AESIs/TEAEs of interest leading to death by PT

In addition, the following will be summarized by actual dose at TEAE onset for the additional AE of interest "Insomnia":

- First occurrence of TEAEs by PT
- First occurrence of serious TEAEs by PT
- TEAEs by PT
- Serious TEAEs by PT

7.7.1.2. Adverse Events Associated with COVID-19

The selection of AEs associated with COVID-19 will be based on the MedDRA MSSO list of PTs. The following will be summarized:

- TEAEs associated with COVID-19, by SOC and PT
- Grade ≥3 TEAEs associated with COVID-19, by SOC and PT
- Serious TEAEs associated with COVID-19, by SOC and PT
- TEAEs associated with COVID-19 leading to discontinuation of study drug, by SOC and PT
- TEAEs associated with COVID-19 leading to interruption of study drug, by SOC and PT

- TEAEs associated with COVID-19 leading to dose reduction, by SOC and PT
- TEAEs associated with COVID-19 leading to death, by SOC and PT

7.7.2. Death

The frequency of subjects in the safety analysis set who died will be tabulated based on information from the EOS eCRF. Deaths will be summarized for the following categories:

- On-treatment death: Deaths within 30 days after the last dose of study treatment (ie, deaths during the on-treatment period)
- Post-treatment death: Deaths more than 30 days after the last dose of study treatment (ie, deaths after the end of the on-treatment period)
- Overall: All deaths

In addition, deaths related to COVID-19 will be summarized.

Deaths for all screened subjects will be provided in a by-subject listing.

7.7.3. Clinical Laboratory Data

Clinical laboratory test results will be expressed in SI units. Preferred unit (g/dL) will also be used for Hb in efficacy analysis.

For each laboratory test (chemistry, hematology, coagulation) performed in the study, a by-subject listing of laboratory test results will be presented with the corresponding CTCAE grades (if applicable), laboratory normal ranges, and flags for values below lower limit of normal (LLN) or above upper limit of normal (ULN).

Parameters with CTCAE grades available:

Clinical laboratory test results will be graded according to CTCAE v4.03 as applicable. Grading will be derived based on the numerical thresholds defined by the CTCAE criteria. Non-numerical qualifiers will not be taken into consideration in the derivation of CTCAE grading.

Laboratory test results classified according to CTCAE will be described using the worst grade. For parameters graded with 2 separate toxicity criteria, such as potassium (hypokalemia/hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (eg, hypokalemia) grades at baseline and postbaseline will be set to 0 when the variables are derived for summarizing high direction toxicity (eg, hyperkalemia), and vice versa.

The frequency of subjects with laboratory toxicities during the on-treatment period will be tabulated as follows. The denominator used to calculate percentages for each laboratory test is the number of subjects evaluable for CTCAE grading for that parameter (ie, those subjects for whom a Grade of 0, 1, 2, 3 or 4 can be derived).

- The summary of laboratory parameters by CTCAE grade will include the number and percentage of subjects with Grade 1, 2, 3, 4; Grade 3-4; and Any Grade (Grades 1-4) during the on-treatment period. The highest CTCAE grade during the on-treatment period is considered the worst grade

- The shift table will summarize baseline CTCAE grade versus worst CTCAE grade during the on-treatment period. The highest CTCAE grade during the on-treatment period is considered the worst grade
- Newly occurring or worsening laboratory abnormalities (Any Grade, Grade 3-4) during the on-treatment period will also be summarized

Parameters with CTCAE grades not available:

Results of laboratory tests that are not part of CTCAE will be presented according to the following categories: below the LLN, within normal limits, and above the ULN according to the laboratory normal ranges.

Shift tables will display the frequency of subjects with shifts from baseline missing, <LLN, normal, or >ULN to each of <LLN, normal or >ULN during the on-treatment period.

7.7.3.1. Hematology

For **WBC differential counts** [total neutrophil, lymphocyte, monocyte, eosinophil, and basophil counts], the absolute value will be used when reported. When only percentages are available (relevant primarily for neutrophils and lymphocytes, because the CTCAE grading is based on the absolute counts), the absolute value is derived as follows:

Derived differential absolute count=(WBC count)×(Differential %value/100)

If the range for the differential absolute count is not available (ie, the range is only available for the percentage) then Grade 1 will be attributed as follows:

- Lymphocyte count decreased:
  - Derived absolute count does not meet Grade 2-4 criteria, and
  - % value <% LLN value, and
  - Derived absolute count ≥800/mm^3
- Neutrophil count decreased:
  - Derived absolute count does not meet Grade 2-4 criteria, and
  - % value<% LLN value, and
  - Derived absolute count ≥1,500/mm^3

7.7.3.2. Chemistry

Liver function tests: Alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and total bilirubin are used to assess possible drug-induced liver toxicity. The ratios of test result to ULN will be calculated and categorized for these parameters during the on-treatment period.

The summary of liver function tests will include the following categories. The frequency of subjects with each of the following during the on-treatment period will be summarized:

- ALT >3×ULN, ALT >5×ULN, ALT >10×ULN, ALT >20×ULN
- AST >3×ULN, AST >5×ULN, AST >10×ULN, AST >20×ULN
- (ALT or AST) >3×ULN, (ALT or AST) >5×ULN, (ALT or AST) >10×ULN, (ALT or AST) >20×ULN
- total bilirubin >2×ULN
- Concurrent ALT >3×ULN and total bilirubin >2×ULN
- Concurrent AST >3×ULN and total bilirubin >2×ULN
- Concurrent (ALT or AST) >3×ULN and total bilirubin >2×ULN
- Concurrent (ALT or AST) >3×ULN and total bilirubin >2×ULN and ALP ≥2×ULN
- Concurrent (ALT or AST) >3×ULN and total bilirubin >2×ULN and (ALP <2×ULN or missing)

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, ie, a subject with an AST >10×ULN will also appear in the categories >5×ULN and >3×ULN. Liver function test elevation and possible Hy's Law cases will be summarized using frequency counts and percentages.

An evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot will be created, with different symbols for different treatment arms, by graphically displaying:

- Peak serum ALT (/ULN) vs peak total bilirubin (/ULN) including reference lines at ALT=3×ULN and total bilirubin=2×ULN
- Peak serum AST (/ULN) vs peak total bilirubin (/ULN) including reference lines at AST=3×ULN and total bilirubin=2×ULN

In addition, the following individual longitudinal plots and by-subject listings will be provided:

- Individual longitudinal plot of ALT including subjects with at least one ALT during the on-treatment period >2.5×baseline or worsening to CTCAE Grade ≥2 during the on-treatment period
- Individual longitudinal plot of AST including subjects with at least one AST during the on-treatment period >2.5×baseline or worsening to CTCAE Grade ≥2 during the on-treatment period
- Listing of all total bilirubin, ALT, AST, and ALP values for subjects with a postbaseline total bilirubin >2×ULN, ALT >3×ULN, or AST >3×ULN
- Listing of all total bilirubin, indirect bilirubin, ALT, AST and ALP values for subjects with a postbaseline ALT >ULN or AST >ULN

In addition, a shift table from baseline to the worst CTCAE grade of ALT and AST during the on-treatment period will be provided. For each subject:

- If the worst CTCAE grade of ALT is worse than that of AST during the on-treatment period, the baseline CTCAE grade of ALT will be used
- If the worst CTCAE grade of AST is worse than that of ALT during the on-treatment period, the baseline CTCAE grade of AST will be used
- If AST and ALT have the same worst CTCAE grade during the on-treatment period, the lower baseline CTCAE grade of ALT and AST will be used

For **calcium**, CTCAE grading is based on corrected calcium and ionized calcium. Corrected Calcium is calculated from albumin and calcium as follows:

> Corrected calcium (mmol/L)=measured total calcium (mmol/L)+0.02×[40−serum albumin (g/L)]

7.7.3.3. Sex Steroid Test

For sex steroid test results, shift tables will display the frequency of subjects with shifts from baseline missing, < LLN, normal, > ULN to each of < LLN, normal or > ULN during the on-treatment period.

In addition, individual longitudinal plots will be provided for each sex hormone by sex.

7.7.3.4. Pregnancy Test

Pregnancy test results will be presented in a by-subject listing.

7.7.4. Vital Signs and Physical Measurements

All physical measurements and vital sign assessments (height, weight, BMI, systolic blood pressure, diastolic blood pressure, pulse rate, temperature) will be presented in a by-subject listing.

7.7.5. Electrocardiograms

ECG summaries will include all ECG assessments from the on-treatment period. QTcB and QTcF interval will be derived based on RR and QT interval (see below), if not collected in the eCRF.

Selecting Primary QT Interval Correction for Heart Rate

The analysis of QT interval data is complicated by the fact that the QT interval is highly correlated with heart rate. Because of this correlation, formulas are routinely used to obtain a corrected QT interval, denoted QTc, which is independent of heart rate. This QTc is intended to represent the QT interval at a standardized heart rate. Several correction formulas have been proposed in the literature. For this analysis several of those methods of correction will be used, as described below. The QT interval corrected for heart rate by the Bazett's formula, QTcB, is defined as

$$QTcB = \frac{QT}{\sqrt{RR}}$$

and the QT interval corrected for heart rate by the Fridericia's formula, QTcF, is defined as

$$QTcF = \frac{QT}{\sqrt[3]{RR}},$$

where RR represents the RR interval of the ECG, in seconds and can be derived as RR (sec)=60/heart rate (bpm).

Although Bazett's correction is the historical standard, it does not perform well when heart rate fluctuates. Fridericia's formula may perform better under these conditions.

ECG Summaries

The following analyses will be performed for each applicable ECG parameter (RR, PR, QRS, QT, and QTc) during the on-treatment period. The denominator to calculate percentages for each category is the number of subjects evaluable for the category.

- Pearson correlation between QT and RR interval, QTc (QTcF, QTcB) and RR interval using baseline assessments
- Frequency of subjects with notable ECG values, defined as those in the following categories:
  - QT/QTc interval increase from baseline >30 ms, >60 ms
  - QT/QTc interval > 450 ms, > 480 ms, > 500 ms
  - PR interval >200 ms
  - QRS duration >120 ms

All ECG assessments and qualitative ECG abnormalities will be presented in by-subject listings.

7.7.6. DXA Scans

DXA scan results including bone mineral density (BMD), T-scores, Z-scores during the on-treatment period will be summarized by location (total femur and adjusted spine) and visit. For T-scores, shift from baseline to the worst post-baseline result during on-treatment period (≤-2.5, >-2.5 - <-1.0, ≥-1.0) will be provided.

All DXA scan results will be presented in a by-subject listing.

7.7.7. Menstrual Cycle Diary

Menstrual cycle diary data collected from women of childbearing potential during the on-treatment period will be summarized by regular contraceptive status (oral contraceptives or depot injection). The following summaries will be included:

- Total number of menstrual cycles reported
- Total number of abnormal menstrual cycles in the following categories: heavier, lighter, longer, shorter, sooner and later than usual.

Menstrual cycle diary data will be presented in a by-subject listing with regular contraceptive status flagged.

CCI

[redacted]

7.9. Interim Analyses

No formal interim analysis is planned for this study.

8. REFERENCES

Clopper, C., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika, 26*(4), 404-413.