Janssen Research & Development *

Statistical Analysis Plan

A Phase 2 Proof-of-Concept Study to Separately Evaluate the Activity of Talacotuzumab (JNJ-56022473) or Daratumumab in Transfusion-Dependent Subjects with Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) who are Relapsed or Refractory to Erythropoiesis-Stimulating Agent (ESA) Treatment

Protocol 56022473MDS2002; Phase 2

JNJ-56022473 (Talacotuzumab)
JNJ-54767414 (Daratumumab)

Status:	Approved
Date:	21 September 2018
Prepared by:	Janssen Research & Development, LLC
Document No.:	

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for study JNJ-56022473-MDS-2002.

2. TRIAL OBJECTIVES

2.1. PRIMARY OBJECTIVE

The primary objective of the study is to evaluate the efficacy (transfusion independence [TI]) of talacotuzumab (JNJ-56022473) or daratumumab in transfusion-dependent subjects with low or intermediate-1 risk MDS whose disease has relapsed during or is refractory to ESAs.

2.2. SECONDARY OBJECTIVES

The secondary objectives of this study are:

- To evaluate the safety of talacotuzumab or daratumumab in the study population.
- To evaluate the clinical benefit of talacotuzumab and daratumumab in this study population through:
    - Time to TI and duration of TI.
    - Rate of HI, CR, and PR.
    - Overall survival (OS).
    - Progression to AML.
    - Rate and amount of supportive care, including transfusions and myeloid growth factors.
- To characterize the PK of talacotuzumab and daratumumab in the study population.
- To evaluate the immunogenicity of talacotuzumab and daratumumab in subjects with MDS.

2.3. EXPLORATORY OBJECTIVES

The tertiary/exploratory objectives of this study are:

- To evaluate pharmacodynamic biomarkers for talacotuzumab activity.
- To determine CD123 and CD38 expression and to explore the association with clinical outcomes.

3. TRIAL DESIGN

3.1. OVERVIEW OF STUDY DESIGN

This is a Phase 2, multicenter, open-label, randomized study to evaluate the safety and efficacy of single-agent talacotuzumab or single-agent daratumumab in subjects with low or intermediate-1 risk MDS who are transfusion-dependent and whose disease has relapsed during or is refractory

to ESA treatment. Approximately 60 subjects (30 to receive talacotuzumab and 30 to receive daratumumab) will be enrolled in this study.

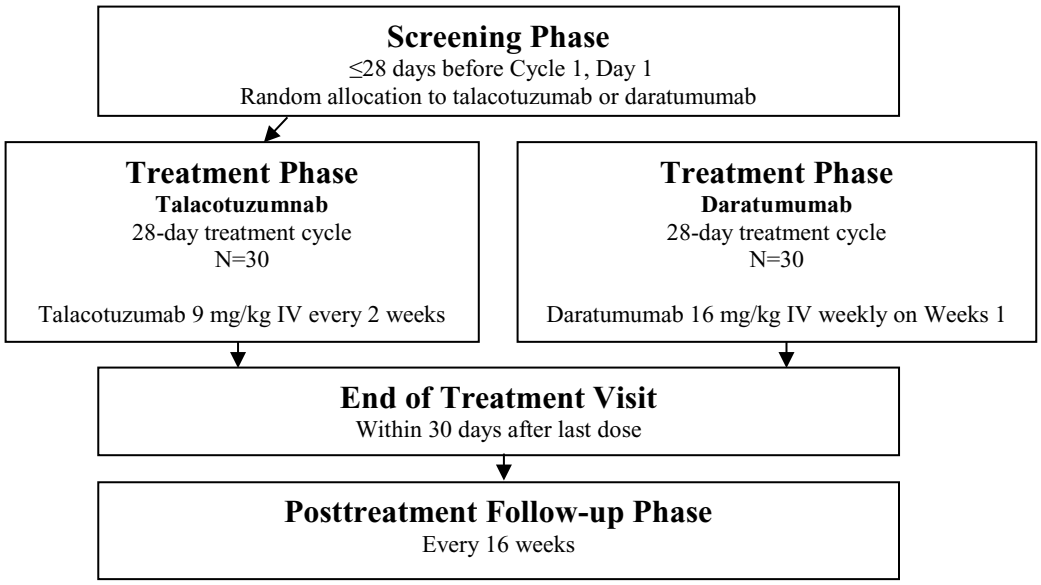
The purpose of this study is to separately evaluate 2 agents with different mechanisms of action that may be effective in subjects with low-risk MDS. The design of this study will allow a preliminary assessment of efficacy and safety of both agents, to determine if either or both warrant further evaluation in phase 3 studies. Subjects will be randomized in a 1:1 ratio to receive talacotuzumab or daratumumab. Randomization will be stratified based on number of RBC units transfused burden (4 or >4) during the 8 weeks prior to randomization.

Talacotuzumab will be administered at 9 mg/kg intravenously (IV) every 2 weeks. Daratumumab will be administered at 16 mg/kg IV weekly on Weeks 1 to 8, every 2 weeks for Weeks 9 to 24, and every 4 weeks thereafter. Cycle length is 28 days for both agents. Study drugs will continue to be administered until disease progression, lack of response, unacceptable toxicity, withdrawal of consent, or study end.

The clinical cutoff for the purpose of the primary endpoint analysis will be 6 months after randomization of the last subject. The end of the study is defined as 1 year after the last subject has been randomized or anytime the sponsor terminates the study.

There will be no statistical hypothesis testing, and each arm will be assessed separately. Statistical methods focus on the evaluation of the likelihood that either treatment regimen will result in a true TI rate of 30% or greater against a minimum acceptable value of 15%. A diagram of the study design is provided below in Figure 1.

Figure 1: Schematic Overview of the Study



3.2. STATISTICAL HYPOTHESIS

The primary hypothesis of this study is that the treatment with talacotuzumab or daratumumab separately will produce a TI rate 30% or greater against a minimal acceptable value of 15%. The statistical evaluation of the hypothesis will use a Bayesian approach to assess the likelihood of whether a true TI rate ≤15% can be ruled out.

3.3. SAMPLE SIZE JUSTIFICATION

Approximately 60 (30 to receive talacotuzumab and 30 to receive daratumumab) subjects will be enrolled in the study to ensure at least 30 evaluable (see Section 4.3 for definition) subjects in each treatment arm.

On the basis of historical data, the TI rate with treatment of best supportive care is expected to be approximately 7.5% in subjects with low or intermediate-1 risk MDS [17, 22, 23]. The primary objective is to evaluate the likelihood that treatment with talacotuzumab or daratumumab can result in a true TI rate ≥30%. A true TI rate ≤15% (twice the TI rate of best supportive care) will be considered as of no clinical interest in the context of this Phase 2 study. The number of subjects to be randomized into each arm will be 30, which will ensure at least 90% probability that the true TI rate is greater than the minimum acceptable value (MAV) 15% if at least 8 (26.7%) subjects reach transfusion independence. This sample size determination utilizes a Bayesian approach with *Beta(0.5, 0.5)* as the prior distribution and *Beta(0.5+**m**, 0.5+30-**m**)* as the posterior distribution for the true TI rate, where ***m*** is the observed number of subjects who achieve TI.

3.4. STATISTICAL METHODS

The objective of the study is to explore if treatment with talacotuzumab or daratumumab will result in an 8-week RBC TI rate ≥30%.

There is no statistical hypothesis testing for the primary objective, and there is no formal statistical comparison between the 2 arms. Instead, statistical criteria, based on observed number of subjects achieving TI (number of TIs), are provided to determine if the outcome of each treatment arm is "meeting", "not meeting" the primary objective, or statistically "inconclusive". This statistical assessment will be used to guide the decision making for further development of either or both treatment therapies. The evaluation will be performed against the target value (TV) 30% and the minimum acceptable value (MAV) 15% for the TI rate. TV of 30% is defined according to study primary objective and is deemed clinically meaningful. A TI rate below 15%, the MAV, will be of no clinical interest. TI rates between MAV and TV will be considered modest, which will define a statistically inconclusive outcome. The statistical criteria is determined to control false positive and false negative error rates each at 10% level by evaluating its operating characteristics using a Bayesian approach with ***Beta(0.5, 0.5)*** as the prior distribution and ***Beta(0.5+m, 0.5+N-m)*** as the posterior distribution of the true TI rate, where ***N*** is the total number of subjects and ***m*** is the observed number of TIs. According to study design, ***N***=30 for each treatment arm.

The TI rate point-estimate and the corresponding 2-sided 80% credible intervals (CI) will be calculated for each treatment arm. The 80% CI will be symmetric, i.e., 10% each tail.

Given ***N***=30 for each treatment arm, as shown in Figure 2 and Table 1 , the lower bound of the 2-sided 80% CI will be >15% (the MAV) and the upper bound >30% (the TV) if the observed number of TI is ≥8, meaning statistically 15% TI rate has been ruled out and 30% ruled in, i.e., the study primary objective is "met". If number of TIs is ≥13, then the outcome will be considered highly desirable as the lower bound of the CI will be >30% (the TV). Number of TIs ≤5 will be an undesirable outcome as the upper bound of the CI is <30% and the lower bound is <15%, implying the primary objective is "not met". When number of TI=6 or 7, the outcomes are statistically inconclusive since the lower bound of the CI is <15% and the upper bound is >30%; and in such a case, results of additional sensitivity analyses and other endpoints (secondary, safety, PK, PD/biomarker) will play a more important role in the decision making.

Table 1: Observed TI, 2-sided Credible Interval, and Statistical Decision Criteria (N=30)

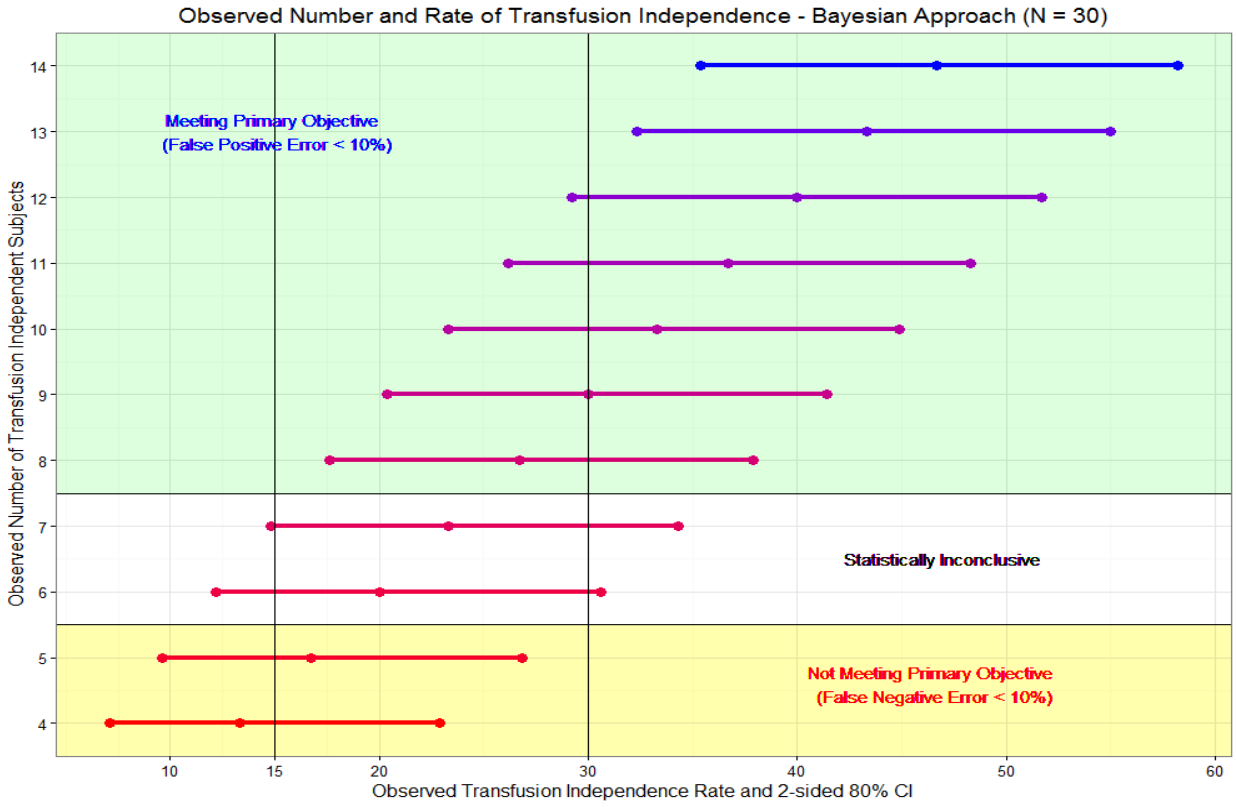
Number (%) of Observed TI	80% C.I.[1]	"Meeting" Primary Objective
4 (13.3)	(7.1, 22.9)	No[2]
5 (16.7)	(9.6, 26.8)	
6 (20.0)	(12.2, 30.6)	Inconclusive
7 (23.3)	(14.8, 34.3)	
8 (26.7)	(17.6, 37.9)	Yes[3]
9 (30.0)	(20.4, 41.4)	
10 (33.3)	(23.3, 44.9)	
11 (36.7)	(26.2, 48.3)	
12 (40.0)	(29.2, 51.7)	
13 (43.3)	(32.3, 55.0)	
14 (46.7)	(35.4, 58.2)	

[1] Using ***B(0.5, 0.5)*** as the prior distribution for true IT rate

[2] With false negative error rate <10%; i.e., when observing # of TI ≤5, there is a <10% chance that the true TI rate is ≥30%

[3] With false positive error rate <10%; i.e., when observing # of TI ≥8, there is a <10% chance that the true TI rate is ≤15%

Figure 2: Plot Showing the Decision Rule for Each Treatment Arm



The statistical method described in this section is also applicable if the number of subjects in each arm is not 30, in which case the posterior distribution corresponding to the actual sample size will be used.

3.5. RANDOMIZATION AND BLINDING

Central randomization will be implemented in this study using an IWRS. Randomization will be stratified based on transfusion burden (4 or >4 units) prior to randomization, and then assigned randomly on a 1:1 basis to receive either talacotuzumab or daratumumab. Transfusion burden is defined as the maximum number of RBC units transfused over any 8 consecutive weeks during the 16 weeks prior to randomization. However, due to safety concerns, enrollment to the talacotuzumab arm was closed (refer to Section 3.1 Overview of Study Design for details).

3.6. CHANGE IN STUDY CONDUCT

Enrollment to the talacotuzumab arm was closed based upon the occurrence of a Grade 4 infusion-related reaction in the first subject who received talacotuzumab. All analyses and summaries described in this document will include data only in subjects who received daratumumab. Data listings will be provided separately for the single subject who received talacotuzumab.

4. GENERAL DESCRIPTIONS AND DEFINITIONS

4.1. BASELINE AND STUDY PHASES

Baseline Value: Unless specified otherwise, the baseline value is defined as the last non-missing value collected on or before the first dose.

Screening Phase: Up to 28 days prior to the first dose of study drug, during which subject eligibility will be reviewed and approved by the sponsor.

Treatment Phase: Between the date of first dose until study drug discontinuation. The Treatment Phase will be divided by cycles, based on the nominal treatment cycles as recorded on the CRF. Each cycle is planned to be 28 days.

Post-treatment Follow-up Phase: After the End of Treatment Phase until the end of study participation or end of study, which is defined as 1 year after randomization of the last subject or anytime the Sponsor terminates the study, whichever comes first.

4.2. STUDY DAY AND CONVENTION OF TIME COMPUTATIONS

Study Day 1 is defined as the day of first dose.

Study Day = assessment date – Study Day 1 + 1 for assessments performed on or after Study Day 1; Study Day = assessment date – Study Day 1 for assessments performed before Study Day 1.

Cycle Day = assessment date – date of the first dose for the cycle + 1.

By convention, one month equals to 30.4375 days and one year equals to 365.25 days. Duration in days between 2 timepoints will be calculated as (end date – start date + 1).

4.3. ANALYSIS SETS

Intent-to-Treat (ITT) analysis set: includes all subjects who are assigned to receive daratumumab. This analysis set will be used for all summaries of disposition, demographic, baseline disease characteristics, and efficacy analyses.

Safety analysis set: includes all subjects who receive at least one dose of daratumumab. This analysis set will be used for all safety analyses and analyses of exposure.

5. SUBJECT INFORMATION

5.1. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

All demographic and baseline characteristic variables will be summarized for the ITT analysis set.

The stratification factors used during randomization (prior RBC transfusion burden) will be tabulated.

Two panels of baseline clinical laboratory tests specified will be summarized: hematology and serum chemistry. Frequencies of National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) grade at baseline will be reported.

Disease characteristics to be summarized will include WHO classification, IPSS risk category (low, intermediate-1), prior ESA treatment (Yes, No), serum EPO level (≤ 500 mU/mL, > 500 mU/ml), and time since initial diagnosis (months).

Hepatitis serology tests at baseline will be summarized by test type and results.

Transfusion history summaries will include total transfusion units, and prior RBC transfusion burden (4, >4 units). Prior RBC transfusion burden is defined as the maximum number of RBC units transfused over an 8-week period during the 16 weeks prior to Study Entry.

Medical history will be summarized including abnormality in physical examination.

In addition, summaries of bone marrow findings at baseline will be provided.

5.2. DISPOSITION INFORMATION

Subject enrollment will be summarized by country and investigator.

Disposition information will be summarized for the ITT analysis set at the end of treatment and at the end of study, respectively.

5.3. EXTENT OF EXPOSURE

All exposure summaries are to be presented for the Safety analysis set.

The number of treatment cycles received will be summarized along with the duration of treatment exposure; dose intensity, which is equal to the sum of the total dose (mg/kg) received in all cycles divided by the number of cycles received; relative dose intensity, which is defined as intensity of total actual dose / intensity of total planned dose; and incidence of dose modifications (aborted infusion, infusion interruption, reduced dose, infusion delay within the cycle, and cycle delay), as well as reasons of dose modifications. Skipped doses are not captured in the database but will be derived and reviewed by study team physicians and provided in a data listing.

5.4. PROTOCOL DEVIATIONS

Subjects with major protocol deviations will be listed. Protocol deviations will be based on clinical review mainly on the following aspects (but not limited to): (1) eligibility criteria, (2) treatment compliance, and (3) prohibited concomitant medication. Protocol deviations will be closely

monitored during the execution of the study and the final set of protocol deviation criteria will be finalized before database lock.

5.5. PRIOR AND CONCOMITANT MEDICATIONS

Medications administered prior to the first dose of study drug will be considered prior medications. Concomitant therapies include those taken on or after first dose date through 30 days after last dose of study drug. Using this definition, a medication can be classified as both prior and concomitant.

Use of prior and concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) class and drug generic term. Prior anticancer therapy will be summarized by type (e.g., radiotherapy, surgery, chemotherapy). Best response to prior ESA therapy will also be summarized.

Concomitant therapies of special interest will be summarized, including: myeloid growth factors [e.g., granulocyte-colony stimulating factor (G-CSF)], anticoagulants and antiplatelet agents. Other concomitant therapies, including anti-infectives (antibacterials, antivirals, and antimycotics), corticosteroids, anti-arrhythmics and other cardiac supportive therapy, anti-histamines, anti-emetics, anti-diarrheals will be summarized.

6. EFFICACY

6.1. PRIMARY EFFICACY ENDPOINT

6.1.1. DEFINITION

The primary efficacy endpoint is the rate of RBC TI lasting for at least 8 weeks. The 8-week RBC TI rate is defined as the proportion of subjects without any RBC transfusion during any consecutive 8 weeks (56 days) starting from Study Day 1 until the end of study. The starting date of 8-week RBC TI must be between Study Day 1 and End of treatment Phase ending date (date of last dose of study medication + 30 days).

6.1.2. HANDLING OF MISSING TRANSFUSION ASSESSMENTS

In order to be considered transfusion independent, a subject must have completed continuous transfusion assessments throughout the qualifying observation period where his/her transfusion independence is determined. An apparent transfusion free period with one or more missing transfusion assessments will not be considered a qualifying 8-week or 24-week TI period. Subjects with no qualifying observation periods will be considered non-responders. In addition, the number (%) of patients who have missed one or more visits will be tabulated.

6.1.3. ANALYSIS METHODS

Number and percent of subjects achieving TI and the corresponding 2-sided 80% confidence

interval (CI) will be summarized.

6.2. MAJOR SECONDARY ENDPOINTS

6.2.1. DEFINITION

24-Week RBC TI Rate

24-week RBC TI rate is defined as the proportion of subjects without any RBC transfusion during any consecutive 24 weeks (168 days) starting from Study Day 1 until the end of study. The starting date of 24-week RBC TI must be between Study Day 1 and End of treatment Phase ending date (date of last dose of study medication + 30 days).

Time to 8-Week (24-Week) RBC TI

Pertaining only to subjects who achieved TI, time to the 8-week (24-week) RBC TI is defined as the interval from Study Day 1 to the first day of the first 8-week (24-week) RBC TI period.

Duration of RBC TI

Pertaining only to subjects who achieved TI for at least 8 weeks, duration of RBC TI is defined as the maximal number of consecutive transfusion-free days that are ≥56 days (8 weeks). For subjects maintaining transfusion-free by the time of the cutoff date for the analysis or who withdraw from the study, duration will be censored at the date of the last transfusion evaluation. Duration of 24-week TI is defined similarly.

Rate of Hematologic Improvement per International Working Group Response Criteria 2006 (IWG 2006)

The IWG criteria for HI define specific responses of cytopenias in the three hematopoietic lineages: erythroid (HI-E), platelet (HI-P), and neutrophil (HI-N). Rate of hematologic improvement of erythroid lineage is the focus of this study , which is defined as the proportion of subjects who achieve HI-E.

HI-E is defined as a hemoglobin (Hb) rise of at least 1.5 g/dL above pretreatment and lasting at least 8 weeks or reduction of at least 4 units of RBC transfusions/8 weeks compared with the prior RBC transfusion burden. Pretreatment Hb level is defined as the average of all the Hb values in the 8 weeks prior to study entry, including the value on Cycle 1 Day 1 and excluding values that were within 14 days after transfusion (thus considered to be influenced by transfusion). If there were no Hb values that met this definition of not being influenced by transfusions, then the baseline value is used.

Rate of Complete Remission (CR) or Partial Remission (PR)

Defined as the proportion of subjects who achieve CR, or PR per IWG 2006 [4].

Overall Survival (OS)

OS is defined as the interval from Study Day 1 to death from any cause. Survival time of living subjects will be censored on the last date a subject is known to be alive or lost to follow-up.

Time to Progression to AML

Time to progression to AML is defined as the interval from Study Day 1 to the date of AML progression (bone marrow or peripheral blood blasts ≥20%). Subjects who have not progressed to AML at the cutoff date for the analysis or who withdraw from the study will be censored at the date of the last disease evaluation.

Amount and Relative Change in RBC Transfusions

Amount of RBC transfusions is defined as the total number of RBC transfusion units in a given post-baseline 8-week (any consecutive 56 days period) interval during study. The starting date of a post-baseline 8-week interval must be between Study Day 1 and End of treatment Phase ending date (date of last dose of study medication + 30 days). Relative change in RBC transfusions is defined as:

$$\text{Relative change in RBC Transfusion} = \frac{\text{amount of RBC transfusions} - \text{prior RBC transfusion burden}}{\text{prior RBC transfusion burden}} * 100\ \%$$

Note that prior RBC transfusion burden is the maximum number of RBC units transfused over an 8-week period during the 16 weeks prior to Study Entry.

The best RBC transfusion reduction is defined as the "best" or the "minimum" number of RBC transfusions in any post-baseline 8-week interval. The time to the best transfusion reduction in an 8-week interval (in weeks), the minimum number of RBC transfusions received during this interval, and the relative change from baseline will be summarized.

Rate and Duration of Myeloid Growth Factor Usage

Rate of myeloid growth factor usage is defined as the proportion of subjects receiving any myeloid growth factors starting from Study Day 1; duration of myeloid growth factor administered starting from Study Day 1.

6.2.2. ANALYSIS METHODS

Rate of 24-week TI, CR, PR, and hematologic improvement (HI), and other binary endpoints, will be summarized with percentage along with its 95% 2-sided exact Clopper-Pearson confidence interval.

The time to 8-week RBC TI and time to 24-week RBC TI will be summarized descriptively.

The Kaplan-Meier method will be used to estimate the distribution of duration of RBC TI for 8-week and 24-week TI responders, respectively. The distribution of OS and time to progression to AML will be summarized using similar Kaplan-Meier methods.

Amount and relative change in RBC transfusions will be summarized descriptively by time point.

7. SAFETY

Safety analyses will be performed using the safety analysis set. The safety parameters to be evaluated are the incidence, intensity, and type of adverse events, clinically significant changes in the subject's physical examination findings, vital signs measurements, clinical laboratory results (hematology and chemistry), and deaths. Exposure to study treatment and reasons for discontinuation will be tabulated. Descriptive statistics will be reported for all safety data.

7.1. ADVERSE EVENTS

Treatment-emergent adverse events are adverse events that (1) occur after the first dose of study drug through the 30 days following the last dose of study drug or until subsequent anti-cancer therapy if earlier; (2) any adverse event considered study drug-related regardless of the start date of the event; or (3) any event present at baseline but worsens in severity or is subsequently considered drug-related by the investigator. Treatment-emergent adverse events will be summarized by System Organ Class and preferred term, by intensity (National Cancer Institute (NCI) common terminology criteria for adverse events (CTCAE), Version 4.03), and by study drug relationship.

For each treatment-emergent adverse event, the percentage of subjects who experience at least 1 occurrence of the given event will be summarized. Tables will be sorted by frequency in incidence (the highest to lowest incidence). The same summary will be provided for serious treatment-emergent AEs, and drug-related serious treatment-emergent AEs, as well as treatment-emergent AEs leading to treatment discontinuation, death, dose modifications or delays.

The verbatim terms used in the electronic case report form (eCRF) by investigators to identify AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

7.3. CLINICAL LABORATORY TESTS

Laboratory data of hematology and clinical chemistry from baseline up to 30 days after last dose will be reported in SI units.

Summary statistics (N, mean, SD, median, and range) will be calculated for the raw data and for their changes from baseline by scheduled visits, as well as for last value and for the changes from baseline to the last value. Within a cycle, summary statistics from the worst values will be provided. Individual values outside the normal ranges will be identified (by "H" for high and "L" for low) in the data listings displaying the absolute values for each subject.

Displays of over-time summaries will be presented for the following key laboratory parameters: hemoglobin, platelet count, white blood cell count, manual absolute peripheral blast count, absolute neutrophil count (ANC), creatinine, total bilirubin (with fractionation if abnormal), ALT, AST, ALP, lactate acid dehydrogenase, and albumin.

Shift tables for each cycle will be produced for selected laboratory parameters, including hemoglobin, platelet count, white blood cell count, peripheral blast count, ANC, total bilirubin, ALT, AST, and ALP. These tables will summarize by cycle the number of subjects with each baseline CTCAE grade and changes to the maximum CTCAE grade in the cycle. For laboratory parameters without CTCAE grade, shifts from baseline to highest per cycle (Low, Normal, High) categories will be used.

Shift tables from baseline to worst value on treatment (from treatment start to 30 days after last dose or the end of treatment visit date, whichever is later) will also be provided. The worst toxicity grade during the treatment will be tabulated.

In addition, incidence of persistent (≥2 or 4 weeks) and severe (≥3 or 4 grade) cytopenia (thrombocytopenia/neutropenia) will be evaluated. Maximum post-baseline CTCAE V 4.03 grade for anemia, leukopenia, neutropenia, and thrombocytopenia will be tabulated.

7.4. VITAL SIGNS AND PHYSICAL EXAMINATION FINDINGS

Descriptive statistics (N, mean, SD, median and range) of vital signs (temperature, blood pressure (systolic and diastolic)) values and changes from the baseline will be summarized at each scheduled time point. The percentage of subjects with values beyond the clinically important limits will be summarized.

Screening physical examination should include body weight, height, and the evaluation of head, eye, ear, nose, and throat (HEENT), cardiovascular, dermatological, musculoskeletal, respiratory, gastrointestinal, and neurological systems. Only a limited symptom-directed physical examination and weight assessment is required on Day 1 of all cycles after baseline. New or worsened abnormalities should be recorded as adverse events if appropriate.