STATISTICAL ANALYSIS PLAN

Syros Pharmaceuticals, Inc.

SY-1425-301

Protocol Title:	A Randomized, Double-blind, Placebo-controlled Phase 3 Study of Tamibarotene Plus Azacitidine Versus Placebo Plus Azacitidine in Newly Diagnosed Adult Patients Selected for RARA-positive Higher-risk Myelodysplastic Syndrome
NCT Number:	NCT04797780
Protocol Version and Date:	Version 4.0; 23 March 2023
Sponsor:	Syros Pharmaceuticals, Inc. 35 CambridgePark Drive, 4th Floor Cambridge, MA 02140 USA
Document Version and Date:	Version 4.0; 09 February 2024

1 STATISTICAL ANALYSIS PLAN APPROVAL

Sponsor:	Syros Pharmaceuticals, Inc.
Clinical Protocol Number:	SY-1425-301
Protocol Title:	A Randomized, Double-blind, Placebo-controlled Phase 3 Study of Tamibarotene Plus Azacitidine Versus Placebo Plus Azacitidine in Newly Diagnosed Adult Patients Selected for RARA-positive Higher-risk Myelodysplastic Syndrome
Document File Name:	SY-1425-301 SAP V4.0 Final 09Feb2024
Document Version and Effective Date:	Version 4.0; 09 February 2024
Approved By:	

PPD **Date**

Syros Pharmaceuticals, Inc.

PPD **Date**

Syros Pharmaceuticals, Inc.

Authored by PPD

Syros Pharmaceuticals, Inc.

Date

2 TABLE OF CONTENTS

1 Statistical Analysis Plan Approval 2
2 Table of Contents 4
3 List of Abbreviations 8
4 Introduction 11
5 Study Objectives 12
5.1 Primary Study Objective 12
5.2 Key Secondary Study Objective 12
5.3 Secondary Study Objectives 12
5.4 Exploratory Study Objectives 13
6 Investigational Plan 13
6.1 Overall Study Design 13
6.2 Schedule of Activities 15
6.3 Treatments 15
6.3.1 Treatments Administered 15
6.3.2 Method of Assigning Participants to Treatment Groups and Blinding 16
6.4 Efficacy and Safety Endpoints 17
6.4.1 Efficacy Endpoints 17
6.4.2 Safety Endpoints 32
6.4.3 Pharmacokinetic Endpoints 34
7 Statistical Methods 35
7.1 General Methodology 35
7.1.1 Reporting Conventions 35
7.1.2 Summarization by Visit 36
7.1.3 Data Handling Rules 36
7.1.4 Standard Calculations 40
7.2 Analysis Sets 41
7.2.1 Modified Intent-to-Treat Analysis Set 41
7.2.2 Intent-to-Treat Analysis Set 41
7.2.3 Safety Analysis Set 1 41
7.2.4 Safety Analysis Set 2 41
7.2.5 Per-Protocol Analysis Set 42
7.2.6 Pharmacokinetic Analysis Set 42
7.3 Study Participants 42
7.3.1 Disposition of Participants 42
7.3.2 Protocol Deviations 43
7.3.3 Demographic and Baseline Characteristics 43
7.3.4 MDS Disease History 44
7.3.5 Medical History 45
7.3.6 Prior and Post-treatment Cancer Therapies 45
7.3.7 Prior and Concomitant Medications 45
7.3.8 Prior and Concomitant Blood Transfusions 46

7.3.9 Other Therapies, Treatments, and Procedures....47
7.3.10 Extent of Study Drug Exposure....47
7.4 Efficacy Evaluation....49
7.4.1 Datasets Analyzed....49
7.4.2 Strata Pooling Strategy....49
7.4.3 Primary Efficacy Endpoint Analysis Methods....50
7.4.4 Key Secondary Efficacy Endpoint Analysis Methods....51
7.4.5 Secondary Efficacy Endpoints Analysis Methods....54
7.4.6 Exploratory Efficacy Endpoint Analysis Methods....64
7.4.7 Statistical/Analytical Issues....68
7.5 Pharmacokinetic Evaluation....72
7.6 Safety Evaluation....72
7.6.1 Adverse Events....72
7.6.2 Clinical Laboratory Evaluation....74
7.6.3 Vital Signs, Physical Findings, and Other Observations Related to Safety....76
7.7 Determination of Sample Size....78
8 Changes in the Conduct of the Study or Planned Analyses....79
8.1 Summary of Changes in Version 4.0....79
8.1.1 Duration of CR and Duration of Overall Response....79
8.1.2 Missing Data and Imputation Rules....79
8.1.3 Safety Analysis Sets....79
8.1.4 Central Mutation Data....79
8.1.5 Transformation to AML....80
8.1.6 Exploratory Comparison of Response Categories and TI....80
8.1.7 Subgroup Analyses....80
8.1.8 Proposed Revised IWG Response Criteria for Exploratory Analyses....80
8.1.9 Other Changes....80
8.2 Summary of Changes in Version 3.0....80
8.2.1 Modified Intent-to-Treat Analysis Set....80
8.2.2 Primary Estimand....81
8.2.3 Secondary Estimand....81
8.2.4 Supplementary Primary and Secondary Estimands....81
8.2.5 Other Changes from the Protocol....81
8.3 Summary of Changes in Version 2.0....81
9 Reference List....82
Appendix 1. Search Strategy for Selected Adverse Event Groups and Their Associated Standardized MedDRA Query (SMQ) or High Level Term (HLT) or Custom MedDRA Query (CMQ)....83
Appendix 2. Proposed Revised IWG Response Criteria....84

TABLE OF TABLES

Table 1 List of Abbreviations....8
Table 2 Study Treatments....16
Table 3 Primary Estimand....18
Table 4 Key Secondary Estimand....20
Table 5 Secondary Estimands....23
Table 6 Supplementary Primary and Secondary Estimands....27
Table 7 IWG Response Criteria for MDS....29
Table 8 Response Criteria for Hematologic Improvement for MDS....31
Table 9 Protocol-Required Safety Laboratory Tests....32
Table 10 Date Imputation Rules for AEs....37
Table 11 Date Imputation Rules for Medical History/Prior and Concomitant Medications/Prior Cancer Therapy....38
Table 12 Assignments for Event and Censoring Dates for Overall Survival Analysis....52
Table 13 Assignments for Event and Censoring Dates for DOCR Analysis....55
Table 14 Assignments for Event and Censoring Dates for Duration of Overall Response Analysis....57
Table 15 Assignments for Primary and Alternative Event and Censoring Dates for EFS Analysis....61
Table 16 Planned Stopping Boundaries for CR Futility Analysis and FA....69
Table 17 Planned Stopping Boundaries for OS Futility Analysis and FA....69
Table 18 Vital Signs ACC/AHA Blood Pressure Categories....77
Table 19 CTCAE Grading Criteria for Temperature....77

TABLE OF FIGURES

Figure 1 Study Design....15

3 LIST OF ABBREVIATIONS

Table 1 List of Abbreviations

Abbreviation	Definition
ACC	American College of Cardiology
AE	Adverse event
AHA	American Heart Association
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
ANC	Absolute neutrophil count
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BID	Twice per day
BMI	Body mass index
CMH	Cochran-Mantel-Haenszel
CMQ	Custom MedDRA Query
CO_2	Bicarbonate
CR	Complete remission
CR_h	Complete remission with partial hematologic recovery
CRF	Case report form
CRO	Contract research organization
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DOCR	Duration of complete remission

Abbreviation	Definition
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDC	Electronic data capture
EFS	Event-free survival
EORTC QLQ-30	European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire 30
EoT	End of Treatment
EQ-5D-5L	EuroQol 5 dimensions
FAB	French-American-British
FDA	Food and Drug Administration
H_A	Alternate hypothesis
HI	Hematologic improvement
HLT	High Level Term
HR-MDS	Higher-risk myelodysplastic syndrome
HRQOL	Health-related quality of life
HSCT	Hematopoietic stem cell transplantation
HSD	Hwang-Shih-DeCani
H_0	Null hypothesis
ICH	International Council for Harmonisation
IDMC	Independent Data Monitoring Committee
INR	International Normalized Ratio
IPCW	Inverse probability censored weighting
IPSS	International Prognostic Scoring System

Abbreviation	Definition
IPSS-M	Molecular International Prognostic Scoring System
IPSS-R	Revised International Prognostic Scoring System
ITT	Intent-to-Treat
IWG	International Working Group
LSM	Least square means
LOCF	Last observation carried forward
mCR	Marrow CR
MDS	Myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MMRM	Mixed models for repeated measures
NCI	National Cancer Institute
NGS	Next-Generation Sequencing
ORR	Overall response rate
OS	Overall survival
PCS	Potentially clinically significant
PK	Pharmacokinetics
PP	Per-Protocol
PR	Partial remission
PT	Preferred term
QTcF	QTc interval corrected for heart rate using Fridericia formula
Q1	First quartile
Q3	Third quartile
RARA	Retinoic acid receptor alpha

Abbreviation	Definition
RBC	Red blood cell
RMST	Restricted mean survival time
ROC	Receiver operating characteristic
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SI	Système International
SMQ	Standardized MedDRA Query
SOC	System organ class
TB	Total bilirubin
TEAE	Treatment-emergent adverse event
TI	Transfusion independence
ULN	Upper limit of normal
vs	Versus
WBC	White blood cell
WHO	World Health Organization

4 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide comprehensive and detailed descriptions of the methods and presentation of data analyses proposed for Syros Pharmaceuticals, Inc.'s Protocol SY-1425-301 (A Randomized, Double-blind, Placebo-controlled Phase 3 Study of Tamibarotene Plus Azacitidine Versus Placebo Plus Azacitidine in Newly Diagnosed Adult Patients Selected for RARA-positive Higher-risk Myelodysplastic Syndrome). Descriptions of planned analyses are provided in order to avoid post hoc decisions that may affect the interpretation of the statistical analysis. The statistical methods applied in the design and planned analyses of this study are consistent with the International Council for Harmonisation (ICH) guideline *Statistical Principles for Clinical Trials* (E9) (1998).

This SAP will be finalized prior to data analysis and futility interim database lock to provide full details to be presented in the clinical study report (CSR). Any changes between the statistical methods provided in the clinical study protocol and this SAP will be explained herein; any changes or deviations from this SAP relative to the final analysis will be fully documented in the CSR. Minor changes or deviations from the templates for tables, figures, and listings need not be documented in the CSR.

5 STUDY OBJECTIVES

5.1 Primary Study Objective

The primary objective of this study is to characterize and compare the complete remission (CR) rate of tamibarotene plus azacitidine versus (vs) placebo plus azacitidine.

5.2 Key Secondary Study Objective

The key secondary objective of this study is to:

- characterize and compare the overall survival (OS) of tamibarotene plus azacitidine vs placebo plus azacitidine

5.3 Secondary Study Objectives

The secondary objectives of this study are to:

- characterize and compare the transfusion independence (TI) rate of tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize and compare the overall response rate (ORR) of tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize the duration of complete remission (DOCR) and duration of overall response of tamibarotene plus azacitidine or placebo plus azacitidine,
- characterize the time to CR and time to initial response of tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize and compare the event-free survival (EFS) rate of tamibarotene plus azacitidine vs placebo plus azacitidine,
- compare changes in health-related quality of life (HRQOL) of tamibarotene plus azacitidine vs placebo plus azacitidine, and
- characterize the safety of tamibarotene plus azacitidine vs placebo plus azacitidine.

5.4 Exploratory Study Objectives

The exploratory objectives of this study are to:

- characterize and compare the cytogenetic CR rate of tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize and compare the molecular response rate of tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize and compare the change in transfusion requirements of tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize and compare the rate of consolidation treatment with allogeneic hematopoietic stem cell transplantation (HSCT) following treatment with tamibarotene plus azacitidine vs placebo plus azacitidine,
- characterize the pharmacokinetics (PK) of tamibarotene, and
- characterize myelodysplastic syndrome (MDS) molecular features associated with response and with loss of response for tamibarotene plus azacitidine vs placebo plus azacitidine.

6 INVESTIGATIONAL PLAN

6.1 Overall Study Design

This is a Phase 3, multi-center, randomized, double-blind, placebo-controlled study comparing the activity of tamibarotene plus azacitidine to placebo plus azacitidine as first-line therapy in retinoic acid receptor alpha (RARA)-positive participants with newly diagnosed higher-risk MDS (HR-MDS). Enrollment in countries in North America, Europe, and Israel is planned.

At the Pre-screening Visit, blood samples will be collected and sent to the Almac Diagnostics Laboratory for assessment of RARA biomarker to determine study eligibility. Pre-screening assessments for the RARA biomarker will be performed within 45 days of Cycle 1 Day 1, and screening assessments for all other eligibility criteria will be performed within 30 days of Cycle 1 Day 1.

Approximately CCI HR-MDS participants will be screened, with approximately CCI participants randomized in the study. Participants will be randomized 2:1 to receive either tamibarotene plus azacitidine or placebo plus azacitidine. Randomization will be stratified by the Revised International Prognostic Scoring System (IPSS-R) risk group (Intermediate, High, and Very High Risk) and by geographical region (North America, Western Europe, plus Israel, and Eastern Europe).

Response will be assessed by the investigator per the modified International Working Group (IWG) MDS criteria (Cheson 2006). Bone marrow aspirates will be collected to measure response on Day 1 of Cycles 2 and 4, followed by every third cycle (7, 10, 13, etc.), with bone marrow aspirates collected at other times as clinically determined based upon clinical findings or changes in peripheral blood counts. Following the response assessment on Day 1 of Cycle 7, those in CR may have the frequency of bone marrow aspirates for response assessment reduced to every 6 cycles.

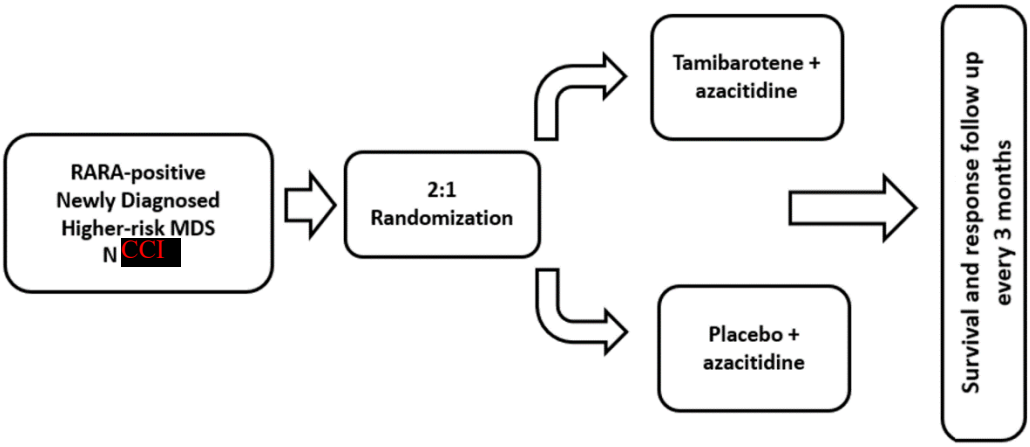
Participants will undergo response and safety evaluations throughout their study participation as detailed in the protocol Schedule of Activities (Section 1.3 in the study protocol). Participants may continue to receive study drug until experiencing an unacceptable toxicity, disease progression (including transformation into acute myeloid leukemia [AML]), relapse, decision to pursue post-remission therapy (such as HSCT) or an alternative anticancer therapy, participant withdraws consent, or the investigator determines it is in the best interest of the participant to discontinue study drug.

An End of Treatment (EoT) Visit should occur within 3 days of the last dose of study drug (or within 3 days of decision to permanently stop drug treatment, if this decision directly follows a period when study drug has been on hold) and before the start of any subsequent anticancer therapy. A Safety Follow-up Visit should occur 30 days (±3 days) after the EoT Visit and if possible before the start of any subsequent anticancer therapy. After the EoT Visit, participants who have not progressed/relapsed will enter the Disease Follow-up period, during which response assessments will be performed for up to 5 years after discontinuation of study drug (irrespective of initiation of HSCT or other subsequent anticancer therapy) or until disease progression/relapse, death, or the occurrence of the expected number of OS events to support the final OS analysis, whichever occurs first. It is estimated that a final analysis of OS will occur approximately 70 months following the enrollment of the first participant.

Participants who progress/relapse will enter Survival Follow-up and will be followed to document the start of subsequent anticancer therapy, the date of transformation to AML, the date of death, and OS status until the expected number of OS events to support the final OS analysis is observed.

An Independent Data Monitoring Committee (IDMC) will be established and will review the data at the 2 planned interim futility analyses and the safety interims, as specified in the IDMC Charter.

Figure 1 Study Design



Abbreviations: MDS = myelodysplastic syndrome; RARA = retinoic acid receptor alpha

6.2 Schedule of Activities

Refer to the protocol (Section 1.3) for the complete Schedule of Activities.

6.3 Treatments

6.3.1 Treatments Administered

Study drug is defined as any investigational intervention(s), marketed product(s), or placebo intended to be administered to a participant according to the study protocol beginning on Cycle 1 Day 1.

Table 2 Study Treatments

Intervention Name	Tamibarotene	Tamibarotene-matched Placebo	Azacitidine
Type	Drug	Drug	Drug
Dose Formulation	Tablet	Tamibarotene-matched tablet	Reconstituted as solution or suspension
Unit Dose Strength(s)	2-mg tablets	Not applicable	100-mg single-use vials
Dosage Level and Regimen	6 mg BID; Days 8 through 28 of each 28-day treatment cycle.	3 tablets that match tamibarotene (BID); Days 8 through 28 of each 28-day treatment cycle	75 mg/m^2 once per day; Days 1 through 7 of each 28-day treatment cycle. If dosing on Days 6 and 7 is not possible due to logistical limitations, these doses may be delayed to Days 8 and 9.
Route of Administration	Oral	Oral	Intravenous or subcutaneous
Use	Experimental	Placebo comparator	Background intervention

Abbreviations: BID = twice per day

6.3.2 Method of Assigning Participants to Treatment Groups and Blinding

The study is double blinded, and randomization will be performed on Cycle 1 Day 1 using an interactive response system. Participants can be randomized earlier than Cycle 1 Day 1, but within 72 hours of the first dose of study drug, to accommodate operational needs.

Approximately CCI participants with HR-MDS will be randomly assigned to study drug. Participants will be randomized 2:1 to receive either tamibarotene plus azacitidine or placebo plus azacitidine. Randomization will be stratified by the IPSS-R risk group (Intermediate, High, and Very High Risk). In addition, stratification of analyses by geographical region (North America, Western Europe, plus Israel, and Eastern Europe) will be performed, as there may be differences in available supportive care treatments and in access to subspecialty consultants and specialized hospital care among these regions.

All participants, investigators, pharmacist(s) performing drug preparation, and site staff will be blinded to treatment assignment.

Personnel responsible for the bioanalysis of the tamibarotene concentrations in plasma will be provided with the randomization code in order to identify samples for analysis. The results will not be shared with the sponsor until the study is unblinded.

In addition, at the time of CR futility analysis, preliminary PK analysis will be performed by designated contract research organization (CRO), and the limited personnel from the CRO will have access to the unblinded PK data to build the population PK model.

To maintain study integrity, OS data will be kept blinded for Syros at the time of CR final analysis and beyond. Only 1 independent SAS programmer at Syros, who does not directly support this study, will have access to the OS data; the OS data will be stored in a separate study folder and can only be accessed by the independent SAS programmer. Further details will be outlined in the Blinding Maintenance Plan.

6.4 Efficacy and Safety Endpoints

Unless otherwise specified, all efficacy and safety endpoints will use participants' all available data up to the data cut-off.

6.4.1 Efficacy Endpoints

6.4.1.1 Primary Efficacy Endpoint and Estimand

The primary efficacy endpoint is CR defined as CR achieved by Cycle 7 Day 1, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006). Peripheral blood counts to support the achievement of CR must be in the absence of transfusions or growth factors within the previous 7 days. The CR should be confirmed at least 4 weeks from the date of marrow sampling. Instructions for the assessment of IWG response including the provision of a 2-week window after the response assessment bone marrow aspiration for unscheduled hematology evaluations to meet the IWG criteria of CR are provided in the Study SY-1425-301 protocol, Section 8.1.1. The calculation of the rate of CR for each treatment group is:

CR rate (%) = number of CR responders achieved by Cycle 7 Day 1 / number of participants in the modified Intent-to-Treat (mITT) analysis set × 100. The denominator is the mITT population in each treatment group.

Non-CR responders (mITT – CR responders) are participants who do not achieve CR or meet any of the following criteria:

1. Do not have any post-baseline response assessments.
2. All response assessment results are missing or not done.
3. Discontinue prior to having a post-baseline assessment.
4. Withdraw consent before achieving a response.
5. Die before achieving a response.
6. Start a new anticancer therapy before achieving a response.
7. Experience prolonged treatment interruption requiring permanent withdrawal of study drug before achieving a response.

Table 3 Primary Estimand

	Estimand Category				
Objective	**Treatment**	**Population (Based on mITT)**	**Variable/ Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
CR: To demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in CR in RARA-positive newly diagnosed HRMDS patients	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk	CR achieved by Cycle 7 Day 1, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006)	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	• CR rate by treatment arm • Difference in proportion of CR rate between 2 treatment arms

Abbreviations: CR = complete remission; HR-MDS = higher-risk myelodysplastic syndrome; IPSS-R = Revised International Prognostic Scoring System; IWG = International Working Group; MDS = myelodysplastic syndrome; mITT = Modified Intent-to-Treat; RARA = retinoic acid receptor alpha; WHO = World Health Organization

6.4.1.2 Key Secondary Efficacy Endpoint and Estimand

The key secondary efficacy endpoint is:

- OS, defined as the duration from the date of randomization to the date of death due to any cause. OS will be calculated in months as:

  OS (months) = (date of death – date of randomization + 1) / 30.4375.

Table 4 Key Secondary Estimand

Objective	Estimand Category				
	Treatment	**Population (Based on ITT)**	**Variable/ Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
OS: To demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in OS in RARA-positive newly diagnosed HR-MDS patients	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive patients with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk	OS, defined as the time from the date of randomization to date of death from any cause	Treatment discontinuation; Treatment policy Initiation of HSCT or other subsequent anticancer therapy; Treatment policy	Kaplan-Meier estimate of OS and hazard ratio

Abbreviations: HR-MDS = higher-risk myelodysplastic syndrome; HSCT = hematopoietic stem cell transplantation; IPSS-R = Revised International Prognostic Scoring System; ITT = Intent-to-Treat; OS = overall survival; RARA = retinoic acid receptor alpha; WHO = World Health Organization

6.4.1.3 Secondary Efficacy Endpoints and Estimands

The secondary efficacy endpoints are:

- TI, defined as a period of at least 56 days with no red blood cell (RBC) or platelet transfusion since the date of randomization to the last dose of study drug + 30 days, the initiation of post-treatment therapy, or death, whichever occurs first. TI rate is calculated as:

TI rate (%) = number of participants who achieve TI / number of participants in the mITT analysis set × 100. The denominator is the mITT population in each treatment group.

- Overall response, defined as achieving CR, partial remission (PR), marrow CR (mCR), or subcategories of hematologic improvement (HI), as determined by the investigator per the modified IWG MDS criteria (Cheson 2006). ORR will be calculated as:

  ORR (%) = number of overall responders / number of participants in the mITT analysis set × 100. The denominator is the mITT population in each treatment group.

  Overall response may also be analyzed in line with the updated IWG response criteria (Zeidan 2023).

- DOCR, defined as the duration from the date of first documented evidence of CR to the date of documented relapse of disease or disease progression, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first. Among CR responders, DOCR will be calculated in months as:

  DOCR (months) = (first date of documented relapse of disease, disease progression, or death due to any cause - date of first documented evidence of CR + 1) / 30.4375.

- Duration of overall response, defined as the duration from the date of first documented evidence of CR, PR, mCR, or HI, to the date of documented disease progression or relapse of disease, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first. Among overall responders, duration of overall response will be calculated in months as:

  Duration of overall response (months) = (first date of documented disease progression, relapse of disease, or death due to any cause - date of first documented evidence of CR, PR, mCR, or HI + 1) / 30.4375.

- Time to CR, defined as the duration from the date of randomization to the date of the first documented evidence of CR as determined by the investigator per the modified IWG MDS criteria (Cheson 2006). Among CR responders, this endpoint will be calculated as:

  Time to CR (months) = (date of the first documented evidence of CR – date of randomization + 1) / 30.4375.

- Time to initial response, defined as the duration from the date of randomization to the date of the first documented evidence of CR, PR, mCR, or HI as determined by the

investigator per the modified IWG MDS criteria (Cheson 2006). Among overall responders, this endpoint will be calculated as:

Time to Initial Response (months) = (date of the first documented evidence of CR/PR/mCR/HI – date of randomization + 1) / 30.4375.

- EFS, defined as the time from the date of randomization to the date of transformation to AML or death due to any cause, whichever occurs first. EFS will be calculated in months as:

  EFS (months) = (date of transformation to AML or death due to any cause - date of randomization + 1) / 30.4375.

- Change in HRQOL as measured by the European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire 30 (EORTC QLQ-30) and EuroQol 5 dimensions (EQ-5D-5L).

Table 5 Secondary Estimands

	Estimand Category				
Secondary Objectives	**Treatment**	**Population**	**Variable/Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
TI: To demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in TI	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk in mITT	TI, defined as a period of at least 56 days with no RBC or platelet transfusion since the date of randomization to the last dose of study drug +30 days, the initiation of post-treatment therapy, or death, whichever occurs first	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	• TI rate by treatment arm
Overall Response: To demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in overall response	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk and who are responders (CR PR, mCR, or HI) in mITT	Overall response, defined as achieving CR, PR, mCR, or HI as determined by the investigator per the modified IWG MDS criteria (Cheson 2006)	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	• ORR by treatment arm

Secondary Objectives	Estimand Category				
	Treatment	**Population**	**Variable/Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
DOCR (CR achieved by Cycle 7 Day 1): To characterize the DOCR of tamibarotene plus azacitidine or placebo plus azacitidine	Tamibarotene plus azacitidine and placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk and those who achieved CR by Cycle 7 Day 1 in mITT	DOCR, defined as the duration from the date of first documented evidence of CR to the date of documented relapse of disease or disease progression, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first	Treatment discontinuation and HSCT; Treatment policy Initiation of subsequent anticancer therapy other than HSCT prior to relapse; Hypothetical strategy	Kaplan-Meier estimate of DOCR
Duration of Overall Response: To characterize the duration of overall response (DOR) of tamibarotene plus azacitidine or placebo plus azacitidine	Tamibarotene plus azacitidine and placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk and who are responders (CR PR, mCR, or HI) in mITT	Duration of overall response defined as the duration from the date of first documented evidence of CR, PR, mCR, or HI, to the date of documented disease progression or relapse of disease as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first	Treatment discontinuation and HSCT; Treatment policy Initiation of subsequent anticancer therapy other than HSCT prior to relapse; Hypothetical strategy	Kaplan-Meier estimate of duration of overall response

Secondary Objectives	Estimand Category				
	Treatment	**Population**	**Variable/Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
Time to CR (CR achieved by Cycle 7 Day 1): Characterize the time to CR of tamibarotene plus azacitidine and placebo plus azacitidine	Tamibarotene plus azacitidine and placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk and who achieved CR by Cycle 7 Day 1 in mITT	Time to CR, defined as the duration from the date of randomization to the date of the first documented evidence of CR as determined by the investigator per the modified IWG MDS criteria (Cheson 2006)	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	Summary statistics (n, mean, median, min, max)
Time to Initial Response Characterize the time to initial response of tamibarotene plus azacitidine and placebo plus azacitidine	Tamibarotene plus azacitidine and placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk who are responders (CR PR, mCR, or HI in mITT	Time to initial response, defined as the duration from the date of randomization to the date of the first documented evidence of CR, PR, mCR, or HI as determined by the investigator per the modified IWG MDS criteria (Cheson 2006)	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	Summary statistics (n, mean, median, min, max)
EFS: To demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in EFS	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk in ITT	EFS, defined as the time from the date of randomization to the date of transformation to AML or death due to any cause, whichever occurs first.	Treatment discontinuation; Treatment policy Initiation of HSCT or other subsequent anticancer therapy; Treatment policy	Kaplan-Meier estimate of EFS

	Estimand Category				
Secondary Objectives	**Treatment**	**Population**	**Variable/Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
Change in HRQOL: Compare changes in HRQOL of participants treated with tamibarotene plus azacitidine vs placebo plus azacitidine	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk in mITT	Observed and change from baseline for EORTC QLQ-30 and EQ-5D-5L parameters	Treatment discontinuation; Treatment policy Patients who die, withdraw consent, or start a new anticancer therapy; Composite strategy	Summary of change from baseline for EORTC QLQ-30 and EQ-5D-5L parameters from Linear mixed model

Abbreviations: AML = acute myeloid leukemia; CR = complete remission; DOCR = duration of complete remission; EFS = event-free survival; EORTC QLQ-30 = European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire 30; EQ-5D-5L = EuroQol 5 dimensions; HSCT = hematopoietic stem cell transplantation; HI = hematologic improvement; HR-MDS = higher-risk myelodysplastic syndrome; HRQOL = health-related quality of life; HSCT = hematopoietic stem cell transplantation; IPSS-R = Revised International Prognostic Scoring System; IWG = International Working Group; mCR = marrow CR; MDS = myelodysplastic syndrome; mITT = Modified Intent-to-Treat; ORR = overall response rate; PR = partial remission; RARA = retinoic acid receptor alpha; RBC = red blood cell; TI = transfusion independence; WHO = World Health Organization.
Note: DOCR and time to CR analysis will be based on the analysis set for those who achieved CR by Cycle 7 Day 1 in mITT.

Table 6 Supplementary Primary and Secondary Estimands

	Estimand Category				
Secondary Objectives	**Treatment**	**Population**	**Variable/Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
CR (achieved at any time): To demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine	Tamibarotene plus azacitidine versus placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk and who achieved CR at any time in mITT	CR as determined by the investigator per the modified IWG MDS criteria (Cheson 2006)	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	• CR rate by treatment arm • Difference in proportion of CR rate between 2 treatment arms
DOCR (CR achieved at any time): To characterize the DOCR of tamibarotene plus azacitidine or placebo plus azacitidine	Tamibarotene plus azacitidine and placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk and those who achieved CR at any time in mITT	DOCR, defined as the duration from the date of first documented evidence of CR to the date of documented relapse of disease or disease progression, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first	Treatment discontinuation and HSCT; Treatment policy Initiation of subsequent anticancer therapy other than HSCT prior to relapse; Hypothetical strategy	• Kaplan-Meier estimate of DOCR

	Estimand Category				
Secondary Objectives	**Treatment**	**Population**	**Variable/Endpoint**	**Intercurrent Event; Strategy**	**Population Level Summary**
Time to CR (CR achieved at any time): To characterize the time to CR of tamibarotene plus azacitidine and placebo plus azacitidine	Tamibarotene plus azacitidine and placebo plus azacitidine	RARA-positive participants with newly diagnosed HR-MDS according to the WHO classification, classified by the IPSS-R risk category as intermediate, high, or very high risk and who achieved CR at any time in mITT	Time to CR, defined as the duration from the date of randomization to the date of the first documented evidence of CR as determined by the investigator per the modified IWG MDS criteria (Cheson 2006)	Treatment discontinuation; Treatment policy Participants who die, withdraw consent, or start a new anticancer therapy before achieving a response; Composite strategy	Summary statistics (n, mean, median, min, max)

Abbreviations: CR = complete remission; DOCR = duration of complete remission; HR-MDS = higher-risk myelodysplastic syndrome; HSCT = hematopoietic stem cell transplantation; IPSS-R = Revised International Prognostic Scoring System; IWG = International Working Group; MDS = myelodysplastic syndrome; mITT = Modified Intent-to-Treat; RARA = retinoic acid receptor alpha; WHO = World Health Organization.
Note: CR, DOCR, and time to CR analysis will be based on the analysis set for those who achieved CR at any time in mITT.

6.4.1.4 Exploratory Efficacy Endpoints

The exploratory efficacy endpoints are:

- Cytogenetic CR, defined as participants achieving cytogenetic CR (Cheson 2006), as determined by the investigator. The cytogenetic CR rate will be calculated as:

  Cytogenetic CR rate (%) = number of cytogenetic CR responders / mITT participants with baseline cytogenetics abnormality × 100.

- Molecular response, defined as participants achieving a reduction in variant allele frequency by next-generation sequencing.

- Transfusion requirement, defined as the frequency and amount of blood products (RBC or platelet) received.

- The rate of consolidation with allogeneic HSCT is defined as the proportion of participants who are treated with allogeneic HSCT. This endpoint is calculated as:

  Consolidation with allogeneic HSCT rate (%) = number of participants treated with allogeneic HSCT / number of participants in the mITT analysis set × 100. The denominator is the mITT population in each treatment group.

- Clinical response and genetic mutations and/or gene expression markers at baseline, at the time of response, and loss of response.

Response types (i.e., CR, PR, mCR, etc.) included in the efficacy endpoints are defined in Table 7 and Table 8. A sample of the EORTC QLQ-30 can be found in Appendix 9, and a sample of the EQ-5D-5L can be found in Appendix 10 of the Study SY-1425-301 protocol.

Table 7 IWG Response Criteria for MDS

Response Criteria[a]	***Peripheral Blood***				***BMB (%)***	***Other***
	Hemoglobin (g/dL)	**Neutrophils per (L)**	**Platelets per (L)**	**Blasts (%)**		
CR	≥11	≥1.0 × 10^9	≥100 × 10^9	0	≤5	Normal maturation of all cell lines, persistent dysplasia will be noted

Response Criteria[a]	***Peripheral Blood***				***BMB (%)***	***Other***
	Hemoglobin (g/dL)	**Neutrophils per (L)**	**Platelets per (L)**	**Blasts (%)**		
PR					Decreased by ≥50% from baseline, but >5%	All CR criteria if abnormal before treatment except BMB
mCR	If HI response, note in addition to mCR				Decreased by ≥50% from baseline, and ≤5%	
Stable disease						Failure to achieve PR and no evidence of progression for >8 weeks
Failure						Death during treatment, or disease progression: worsening cytopenia, increase in % BMBs, progression to a more advanced MDS FAB subtype
Relapse after CR or PR						At least one of the following: Return to pretreatment BMB %. Decrement of ≥50% from maximum remission/response levels in granulocytes or platelets. Reduction in Hgb ≥1.5g/dL or transfusion dependence.
Cytogenetic Response			**Evaluation**			
Complete			Disappearance of chromosomal abnormality with no appearance of new ones.			
Partial			≥50% reduction of chromosomal abnormality			

Response Criteria[a]	*Peripheral Blood*				*BMB (%)*	*Other*
	Hemoglobin (g/dL)	**Neutrophils per (L)**	**Platelets per (L)**	**Blasts (%)**		
Disease Progression, participants with:	**Evaluation Criteria:**					
<5% bone marrow blasts	≥50% increase to >5% bone marrow blasts					
5%-<10% bone marrow blasts	≥50% increase to >10% bone marrow blasts					
10%-<20% bone marrow blasts	≥50% increase to ≥20% bone marrow blasts					
For all categories	At least 50% decrease from maximum remission/response in granulocytes or platelets or reduction in Hgb by ≥ 2 g/dL or transfusion dependence					

Abbreviations: BMB = bone marrow blasts; CR = complete remission; FAB = French-American-British; HI = hematologic improvement; Hgb = hemoglobin; IWG = International Working Group; mCR = marrow CR; MDS = myelodysplastic syndrome; PR = partial remission
[a] Response must last at least 4 weeks.

The IWG criteria for HI define specific responses of cytopenias in the 3 hematopoietic lineages: erythroid (HI-E), platelet (HI-P), and neutrophil (HI-N). The HIs are measured in participants with pretreatment abnormal values: hemoglobin level less than 110 g/L (11 g/dL) or RBC-transfusion dependence, platelet count less than 100×10^9/L or platelet-transfusion dependence, and absolute neutrophil count (ANC) less than 1.0×10^9/L.

Table 8 Response Criteria for Hematologic Improvement for MDS

Hematologic Improvement	Response Criteria (Response Lasting 8 Weeks)
Erythroid response (pretreatment <11 g/dL)	Hgb increase by ≥1.5 g/dL Relevant reduction of units of RBC transfusions by an absolute number of at least 4 RBC transfusions/8 weeks compared with the pretreatment transfusion number in the previous 8 weeks. Only RBC transfusions given for a Hgb of ≤9.0 g/dL pretreatment will count in the RBC transfusion response evaluation.
Platelet response (pretreatment $<100 \times 10^9$/L)	Absolute increase of $\geq 30 \times 10^9$/L if starting with $>20 \times 10^9$/L platelets Increase from $<20 \times 10^9$/L to $>20 \times 10^9$/L and by at least 100%
Neutrophil response (pretreatment $<1 \times 10^9$/L)	At least a 100% increase and an absolute increase $>0.5 \times 10^9$/L
Progression or relapse after HI in the absence of another explanation	At least one of the following: At least 50% decrease from maximum response levels in granulocytes or platelets; Reduction in Hgb by ≥1.5 g/dL; Transfusion dependence.

Abbreviations: Hgb = hemoglobin; HI = hematologic improvement; MDS = myelodysplastic syndrome; RBC = red blood cell
Adapted from: Cheson 2006.

6.4.2 Safety Endpoints

The safety endpoints for this study are incidence of adverse events (AEs) and changes in clinical laboratory values, electrocardiogram (ECGs), vital sign, and Eastern Cooperative Oncology Group (ECOG) measurements.

6.4.2.1 Adverse Events

An AE is any untoward medical occurrence in a clinical study participant, temporally associated with the use of study drug, whether or not considered related to the study drug. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of study drug whether or not related to the study drug. An AE can arise from any use of the study drug and from any route of administration, formulation, or dose, including an overdose.

AEs will be captured from the time of the first study drug administration through the Safety Follow-up Visit.

6.4.2.2 Laboratory Parameters

All protocol-required local laboratory tests, as defined in Table 9, will be conducted in accordance with local laboratory practices.

Table 9 Protocol-Required Safety Laboratory Tests

Laboratory Tests	**Parameters**
Hematology	Platelet count Hemoglobin WBC (leukocyte count including differential): Neutrophils (ANC, calculated from the leukocyte count and WBC differential count) Lymphocytes Monocytes Eosinophils Basophils % Blasts

Laboratory Tests	**Parameters**
Clinical chemistry	Blood urea nitrogen/urea Creatinine CO_2 Uric acid Albumin Sodium Phosphorus Triglycerides Total cholesterol Magnesium Calcium Potassium Chloride Glucose Amylase Lipase Total protein ALP Lactate dehydrogenase Total and direct bilirubin AST/serum glutamic-oxaloacetic transaminase ALT/serum glutamic-pyruvic transaminase
Coagulation	Prothrombin time or international normalized ratio Activated partial thromboplastin time/ partial thromboplastin time

Laboratory Tests	Parameters
Routine urinalysis	Specific gravity pH Protein RBC WBC Leukocyte esterase Ketones Nitrite
Pregnancy testing	Highly sensitive serum or urine human chorionic gonadotropin pregnancy test (as needed for women of childbearing potential)[a]

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; CO_2 = bicarbonate; RBC = red blood cell; WBC = white blood cell

[a] Local high-sensitivity urine pregnancy testing will be standard for the protocol unless serum testing is required by local regulation or Institutional Review Board/Independent Ethics Committee.

6.4.2.3 Other Safety Endpoints

6.4.2.3.1 Vital Signs

Vital signs will be measured in a semi-supine position after 5 minutes of rest and will include temperature, systolic and diastolic blood pressure, pulse rate, and respiratory rate.

6.4.2.3.2 Electrocardiograms

Triplicate 12-lead ECGs will be obtained using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QTc intervals (utilizing Fridericia's correction for QTc).

6.4.3 Pharmacokinetic Endpoints

The PK endpoints will be provided in a separate report and is outside the scope of this SAP.

7 STATISTICAL METHODS

7.1 General Methodology

Data will be analyzed by Syros Pharmaceuticals, Inc. biostatistics personnel or designee. Statistical analyses will be reported with tables, figures, and listings, using recommended ICH numbering. Output specifications for all tables, figures, and listings will be in conformance with the guidelines specified by the ICH in Appendix 7 of the *Electronic Common Technical Document Specification* (Apr 2003).

7.1.1 Reporting Conventions

Efficacy tables, safety tables (except overall AE tables), and figures will be summarized by treatment group (tamibarotene + azacitidine and placebo + azacitidine). Other tables will also include a column for all participants combined. In general, all data collected, and any derived data will be presented in participant data listings for the relevant populations. Listings will be ordered by treatment group, participant number, and assessment or event date. The treatment group presented in listings will be based on the planned assignment, unless otherwise noted.

In general, continuous variables will be summarized to indicate the number of participants with available data (n), mean, standard deviation (SD), median, minimum, and maximum values. Categorical variables, unless otherwise specified, will be summarized by the number of participants with available data (n) and number and percentage of participants in each category. If there is a subcategory under a categorical variable, this subcategory will be summarized by the number and percentage of participants based on the number of participants under this categorical variable. A category for 'Missing' will be added and populated to make the total percentage of each category equal to 100%. Select ordinal data may be summarized using both descriptive statistics and counts and percentages of participants in each category, as appropriate.

Non-zero percentages will be rounded to 2 decimal places. Rounding conventions for presentation of summary statistics will be based on the precision of the variable of summarization, as it is collected in its rawest form (i.e., on the electronic case report form [eCRF] or as provided within an external file) and are outlined as follows:

- The mean and median will be rounded to 1 more decimal place than the precision of the variable of summarization, and up to 3 decimal places.
- Measures of variability (e.g., SD, SE) will be rounded to 2 more decimal places than the precision of the variable of summarization, and up to 4 decimal places.
- Minimum and maximum values will be presented using the same precision as the variable of summarization, and up to 2 decimal places.

Other statistics (e.g., CIs) will be presented using the same general rules outlined above or assessed for the most appropriate presentation based on the underlying data.

The primary endpoint will be tested at a full alpha level of 2.5%. To control the overall Type I error rate of the primary and key secondary endpoints, hierarchical testing will be performed in the following order: (1) primary endpoint; (2) key secondary endpoint. If hypothesis testing of the primary endpoint is rejected, the key secondary endpoint will be tested at the full alpha level of 2.5%; otherwise, the testing will be stopped. Secondary and exploratory efficacy endpoints will be tested using a two-sided, 5% significance level, and two-sided 95% CIs will be reported when appropriate, unless otherwise specified. P-values will be rounded to 4 decimal places. P-values less than 0.0001 will be displayed as "<0.0001"; p-values greater than 0.9999 will be displayed as ">0.9999".

When applicable, analyses will be stratified by IPSS-R risk group (Intermediate, High, and Very High Risk) and by geographical region (North America, Western Europe, plus Israel, and Eastern Europe).

7.1.2 Summarization by Visit

Scheduled analysis visits are visits at scheduled timepoints as specified in the protocol. Data summarized by study visit will be based on the nominal, scheduled visit label as reported on the eCRF. For safety analysis, in the case of multiple observations at a specific visit, the latest observation will be used. If more than 1 observation is made on the same day, an average value if continuous or the worst value if categorical will be included in the analysis.

Data collected at unscheduled visits will not be included in the by-visit summaries but will be considered when endpoint derivations potentially include multiple visits (e.g., determination of baseline value, determination of worst post-baseline value, etc.). All data will be included in participant listings.

7.1.3 Data Handling Rules

7.1.3.1 Threshold Values

Unless otherwise noted, values that include a threshold sign (i.e., '<' or '>') will be imputed for summary tables in the following manner:

- A value that is 1 unit less than the limit of quantitation will be used for the calculation of descriptive statistics if the datum is reported in the form of "<x" (where x is considered the limit of quantitation). For example, if the values are reported as <50 and <5.0, values of 49 and 4.9, respectively, will be used for calculation of summary statistics. An exception to this rule is any value reported <1. For the values reported as <1 or <0.1, a value of 0.9 or 0.09 will be used for calculation of summary statistics.

- A value that is 1 unit above the limit of quantitation will be used for the calculation of descriptive statistics if the datum is reported in the form of ">x" (where x is considered the limit of quantitation). Values with decimal points will follow the same logic as above.

- The limit of quantitation will be used for the calculation of descriptive statistics if the datum is reported in the form of "≤x" or "≥x" (where x is considered the limit of quantitation).

The original data including the threshold signs will be displayed in the listings.

7.1.3.2 Baseline Definition

For all endpoints, unless otherwise specified, the baseline value will be the latest pre-dose assessment with a non-missing value, including those from unscheduled visits. If time is not collected, Day 1 assessments are assumed to be taken prior to first dose and used as baseline. When scheduled assessments and unscheduled assessments occur on the same day of first dose and time of the assessment or time of first dose is not available, the following convention to determine baseline will be used:

- If both a scheduled and an unscheduled visit are available on the day of the first dose, and time is missing, the scheduled assessment will be used as baseline.

- If all scheduled assessments are missing on the day of the first dose and an unscheduled assessment is available, the unscheduled assessment will be used as baseline.

For participants who did not receive study treatment during the study, baseline will be defined as the latest, non-missing collected value. Unless otherwise stated, if baseline data are missing, no derivation will be performed, and baseline will be set to missing.

For ECG analyses, if the latest, non-missing pre-dose value is from triplicate, the participant level baseline is defined as the mean of the triplicate baseline assessments.

7.1.3.3 Imputation of Partial or Missing AE Dates for Analysis

Rules for imputing partial or missing AE dates for analysis are provided in Table 10.

Table 10 Date Imputation Rules for AEs

Adverse Events	**Partial/missing AE dates recorded in the CRF will be imputed using the following conventions:**

Missing start day	The first of the month will be used unless it is before the start date of study treatment; in this case, the study treatment start date will be used and hence the event is considered treatment emergent.
Missing start day and month	No imputation.
Missing end day	Last day of the month will be used.
Missing end day and month	No imputation.
Completely missing start/end day	No imputation.

Abbreviations: AE = adverse event; CRF = case report form

7.1.3.4 Imputation of Partial or Missing Medical History/Concomitant Medication Dates for Analysis

Rules for imputing partial or missing medical history or concomitant dates for analysis are provided in Table 11.

Table 11 Date Imputation Rules for Medical History/Prior and Concomitant Medications/Prior Cancer Therapy

Medical History/Prior and Concomitant Medications/Prior Cancer Therapy	**Partial/Missing Medical History/Prior and Concomitant Medication/Prior Cancer Therapy Dates Recorded in the CRF Will Be Imputed Using the Following Conventions:**
Missing start day	'01' will be used for the day.
Missing start day and month	'01' will be used for the day and 'Jan' will be used for the month.
Missing end day	'28/29/30/31' will be used for the day (dependent on the month and year) and the earlier of last day of the month and the end of study date will be used as the imputed value except in the following two cases: 1) For prior cancer therapy 2) If it is not ongoing and the start date is prior to the study treatment start date For these two cases, the end date will be imputed as (study treatment start date -1).

Missing end day and month	'31' will be used for the day and 'Dec' will be used for the month, and the earlier of 31 Dec of the given year and the end of study date will be used as the imputed value except in the following two cases: 1) For prior cancer therapy 2) If it is not ongoing and the start date is prior to the study treatment start date For these two cases, the end date will be imputed as (study treatment start date -1).
Completely missing start/end day	No imputation.

7.1.3.5 Imputation of Partial or Missing New Anticancer Therapy Dates for Efficacy Evaluation

Start dates for follow-up anticancer therapy will be temporarily imputed in order to define event and censoring rules for EFS, response rate, time to event, duration of response, or time to response (e.g., start date for new anticancer therapy). The imputed dates will be stored in the anticancer therapy-derived dataset.

- If missing start day, month, and year, then no imputation.
- If missing start day and month, then no imputation.
- If missing start day only, then do the following:
    - If partial date falls in the same month as the last dose of study treatment, then assign to earlier of (date of last dose of study treatment + 1, last day of month).
    - If partial date falls in the same month as the participant's last assessment and the participant's last assessment is transformation to AML, then assign to earlier of (date of transformation to AML + 1, last day of month).
    - If both rules above apply, then assign to latest of the 2 dates.
    - Otherwise, impute missing day to the first of the month.
- If missing end date, then no imputation should be done.

7.1.3.6 Imputation of Partial or Missing Date of Birth and the Calculation of Age

Age in years will be calculated as (date of main informed consent – date of birth + 1) / 365.25 and will be stored in the adsl dataset.

- If missing day, month, and year of birth, then no imputation.
- If missing day and month of birth, then impute date of birth to June 30.
- If missing day only, then impute date of birth to 15th of the month.

7.1.3.7 Imputation of Partial or Missing Date of Death

Date of death will be temporarily imputed in order to define event and censoring rules for time to event endpoints. The imputed dates will be stored in the adsl dataset.

- If missing day, month, and year, then no imputation.
- If missing day and month, then impute to later of January 1st or (last known alive date + 1).
- If missing day only, then impute to later of first date of the month or (last known alive date + 1).

7.1.3.8 Calculation of ECG Data

Three reads are taken during ECG at each visit. The average of the available reads will be used as the ECG data at the corresponding visit. If none of the three reads are available, then the observation will be considered as missing, and no imputation will be performed.

7.1.3.9 Post-Baseline Missing Values

- If post-baseline Common Terminology Criteria for Adverse Events (CTCAE) grade for laboratory parameters is missing, there will be no imputation for missing values, and in the shift from baseline to worst post-baseline CTCAE grade by parameter tables, a missing category will be added in order to have the total number of post-baseline participants equal to the total number at baseline.
- Missing category for post-baseline values will be added to ECG shift table, similar to the bullet above. The same rules will be applied for vital sign shift tables and ECOG shift tables.
- If AE grade is missing, the previous worst grade under the same preferred term (PT) will be assigned if it is available, otherwise a grade of 4 will be assigned for classification purpose in the tables, figures, and listings.
- If the transfusion eCRF page is missing or if the blood product received information is missing, then there will be no imputation, and the participants will be classified as having missing transfusion dependence status. If the date of the transfusion is missing, then the transfusion dependence status at baseline and post-baseline will both be missing.

7.1.4 Standard Calculations

Where appropriate, the calculated study day of each assessment or event will be presented with the assessment or event date on participant data listings, where study day will be determined as:

- The assessment/event date minus the date of first dose of study drug, if the assessment/event date is prior to the date of first dose; and

- The assessment/event date minus the date of first dose of study drug, plus 1, if the assessment/event date is on or after the date of first dose.

Other variables requiring calculations will be derived using the following formulas:

- **Days:** A duration between 2 dates expressed in days will be calculated as the later date – the earlier date + 1.
- **Months:** A duration expressed in months will be calculated by dividing the duration in days by (30.4375).
- **Years:** A duration expressed in years will be calculated by dividing the duration in days by 365.25.
- **Change from Baseline:** Change from baseline will be calculated as the post-baseline value minus the baseline value.
- **Percentage Change from Baseline:** Percentage change from baseline will be calculated as the change from baseline divided by the baseline value, multiplied by 100.

7.2 Analysis Sets

7.2.1 Modified Intent-to-Treat Analysis Set

The mITT Analysis Set includes the first approximately CCI randomized participants. Treatment groups and stratification for this analysis set will be determined according to the treatment and stratification assignments at the time of randomization.

7.2.2 Intent-to-Treat Analysis Set

The Intent-to-Treat (ITT) Analysis Set includes all participants who are randomized. Treatment groups and stratification for this analysis set will be determined according to the treatment and stratification assignments at the time of randomization.

7.2.3 Safety Analysis Set 1

The Safety Analysis Set 1 includes all participants in the ITT Analysis Set who have received any amount of study drug (tamibarotene, placebo, or azacitidine). Treatment groups for this analysis set will be determined according to the actual treatment the participants received. Participants randomized to the tamibarotene plus azacitidine arm who discontinue treatment prior to receiving any doses of tamibarotene will be included in the placebo plus azacitidine arm in Safety Analysis Set 1.

7.2.4 Safety Analysis Set 2

The Safety Analysis Set 2 includes all participants in the mITT Analysis Set who have received any amount of study drug (tamibarotene, placebo, or azacitidine). Treatment groups for this analysis set will be determined according to the actual treatment the

participants received. Participants randomized to the tamibarotene plus azacitidine arm who discontinue treatment prior to receiving any doses of tamibarotene will be included in the placebo plus azacitidine arm in Safety Analysis Set 2.

7.2.5 Per-Protocol Analysis Set

The Per-Protocol (PP) Analysis Set includes all participants in the mITT analysis set who are considered to be sufficiently compliant with the protocol as determined by the sponsor before database lock. Treatment groups for this analysis set will be determined according to the treatment assignment at the time of randomization.

7.2.6 Pharmacokinetic Analysis Set

The PK Analysis Set includes all participants who are randomized and have received at least 1 dose of tamibarotene and have at least 1 whole blood sample collected and assayed for measurement of tamibarotene plasma concentration and the time of sampling and the time of dosing on the day of sampling is known.

7.3 Study Participants

7.3.1 Disposition of Participants

Participant disposition will be summarized for all participants and the mITT analysis set by treatment group (placebo + azacitidine or tamibarotene + azacitidine) and for all participants combined (total). Summaries will include the following:

- the number and percentage of participants in each stratum (IPSS-R Risk Group and Geographical Region)
- the number and percentage of participants in each analysis set
- the number and percentage of participants by primary reason for exclusion from the PP analysis set
- the number and percentage of participants completing and discontinuing each study treatment (tamibarotene/placebo and azacitidine)
- the number and percentage of participants by primary reason for discontinuation of each study treatment (tamibarotene/placebo and azacitidine)
- the number and percentage of participants by primary reason for study termination

The primary reasons for exclusion from the PP analysis set will be identified by the Sponsor prior to database lock and unblinding of the study.

A listing of participant disposition by treatment group will be provided. Additionally, listings of randomization schemes and codes (including strata), discontinuation from the study, discontinuation from study treatment, and study eligibility will be provided. Screen failures, which are defined as participants who had signed the main informed consent to

participate in the clinical study but are not subsequently randomly assigned to study treatment, will be included in the study eligibility listing but will not be included in any summary tables for analysis. Participants who had signed a prescreening informed consent and did not have *RARA*-positive test results will be considered prescreening failures; these participants will not be included in the screening failure listings.

7.3.2 Protocol Deviations

Major protocol deviations will be summarized by treatment group and over all participants combined for the ITT analysis set. Major protocol deviations are protocol deviations, including non-compliance with the protocol, captured on-study that are deemed by the Sponsor to potentially impact the efficacy or safety conclusions of the study.

Protocol deviation review will be conducted and agreed upon by Sponsor and CRO stakeholders regularly throughout the study. All major protocol deviations will be determined and appropriately categorized (as specified in the study Protocol Deviation Assessment Plan) prior to database lock and unblinding of the treatment groups. The number and percentage of participants with protocol deviations that impact the efficacy or safety conclusions, as well as the number and percentage of participants with deviations within all categories, will be presented.

The reasons for protocol deviations will be summarized for the PP analysis set.

By-participant listings of all protocol deviations, including those related to COVID-19, will be provided.

7.3.3 Demographic and Baseline Characteristics

Demographic variables and baseline characteristics will be summarized using descriptive statistics (mean, standard deviation, and median, as well as the minimum and maximum values) by treatment group and overall with the number of non-missing observations for the continuous variables;

Categorical variables will be summarized with the number and percentage of participants in each category by treatment group and overall. The number of participants with missing information will also be summarized.

The mITT and ITT analysis sets will be used for demographic and baseline characteristics analyses.

The demographic variables are:

- Age (years)
- Age categories (18 to <65, ≥65 to <75, ≥75)
- Sex (male, female)

- Race
- Ethnicity
- Height (cm)
- Weight (kg)
- Body mass index (BMI)

BMI will be calculated as: weight (kg) / [height (cm) / 100]^2.

The baseline variables are:

- ECOG performance status (0, 1, 2, 3)

The demographic and baseline data listed above, as well as female childbearing potential, will be presented in listings.

7.3.4 MDS Disease History

MDS disease history characteristics are:

- Time since higher-risk diagnosis of MDS (months)
- WHO classification at study entry (MDS with Excess Blasts – MDS-EB-1, MDS with Excess Blast – MDS-EB2)
- MDS type (primary or not treatment-related MDS, secondary or treatment-related MDS)
- IPSS-R category (intermediate, high, very high)
- Geographical region (North America, Western Europe, plus Israel, and Eastern Europe)
- Bone marrow blast (%)
- Bone marrow blast category (<5%, ≥5% to <10%, ≥10%)
- Cytogenetic risk status (very good, good, intermediate, poor, very poor)
- Molecular abnormalities for genes commonly associated with MDS
- Low hemoglobin (<11 g/dL)
- Low platelets (<100 × 10^9/L)
- Low ANC (<1.0 × 10^9/L)

The above MDS disease characteristics will be summarized by treatment group and overall for the mITT and ITT analysis sets. If available, central Next-Generation Sequencing (NGS) data will be used to characterize the molecular features of participants in the treatment groups, otherwise local laboratory results will be used.

Time since higher-risk diagnosis in months will be calculated as: (Date of randomization - Date of Higher-Risk Diagnosis of MDS + 1) / 30.4375. The continuous variables will be summarized using descriptive statistics with the number of non-missing observations, and the categorical variables will be summarized with the number and percentage of participants in each category, and the number of participants with missing information will also be summarized.

A listing of MDS disease history will be provided.

7.3.5 Medical History

Medical history is collected prior to dosing and will be summarized by treatment group and overall for the mITT and ITT analysis sets. Verbatim terms on case report forms (CRFs) will be mapped to PTs and system organ classes (SOCs) using the Medical Dictionary for Regulatory Activities (MedDRA), version 23.1 or higher.

Summaries will be ordered by descending incidence of SOC and descending incidence of PT within each SOC for all participants. A listing of medical history will be provided.

7.3.6 Prior and Post-treatment Cancer Therapies

All prior cancer therapy data, including regimen number, indication, drug/agent name, and start/end dates, will be collected prior to dosing. Any anticancer therapy that is started after the end of study treatment will also be recorded. Medications will be coded using the WHO Drug Dictionary Global B3 Sep 1, 2020 or a more recent version and will be mapped to Anatomical Therapeutic Chemical (ATC) drug class (level 4) and drug name. Prior and post-treatment cancer therapies will be summarized separately by treatment group and overall for the mITT and ITT analysis sets. Summaries will be ordered by descending incidence of ATC class and descending incidence of drug name within ATC class for all participants.

Separate by-participant listings of prior and post-treatment cancer therapies will be provided.

7.3.7 Prior and Concomitant Medications

Medications will be coded using the WHO Drug Dictionary Global B3 Sept. 1, 2020 or a more recent version. Medications entered on the eCRF will be mapped to ATC drug class (level 4) and drug name.

Prior and concomitant medications will be summarized separately, and the study phase of each medication will be determined programmatically based on medication start and end dates. A prior medication is defined as any medication taken with an end date prior to the first dose of study drug. A concomitant medication is defined as any medication that (1) started before the first dose of study drug and was continuing at the time of the first dose of study drug, or (2) started on or after the date of the first dose of study drug through 30 days after the participant's last dose of study drug.

For both prior and concomitant medication summaries, the number and percentage of participants receiving any medication will be summarized by treatment group and overall, as will the number and percentage of participants receiving any medication by ATC drug class and generic drug name. Participants reporting use of more than 1 medication at each level of summarization (any medication received, ATC class, and generic drug name) will be counted only once. ATC class terms will be displayed in descending order of incidence for all participants, as will generic drug names within each ATC class. The study phase during which each medication was received (e.g., prior or concomitant) will be presented on the listing of prior and concomitant medications.

mITT and ITT analysis sets will be used for prior and concomitant medication analysis.

7.3.8 Prior and Concomitant Blood Transfusions

Any transfusions from 16 weeks prior to randomization through the Disease Follow-up period will be recorded. Transfusions will be classified as prior or concomitant based on the start date. A prior transfusion is defined as any transfusion that started up to 16 weeks prior to the date of randomization. A concomitant transfusion is defined as any transfusion that started on or after the date of randomization through the Disease Follow-up period. Prior and concomitant blood transfusions will be summarized separately.

For both prior and concomitant transfusion summaries, the number and percentage of participants receiving any RBC or platelet transfusion will be summarized by treatment group and overall, as will the number and percentage by transfusion type (RBCs, platelets, or other blood product). Participants with more than 1 transfusion at each level of summarization (any transfusion received, transfusion type) will be counted only once. Additionally, transfusion burden (units/56 days) will be summarized using descriptive statistics as part of the prior transfusion and concomitant blood transfusion summary separately. Transfusion burden is defined as the total number of transfusion units received from 16 weeks prior to randomization to randomization for prior transfusion, and the total number of transfusion units received from randomization through Disease Follow-up period for concomitant blood transfusion. Transfusion burden will be analyzed by each 56 days blocks for both prior transfusion and concomitant blood transfusion. The study phase during which each transfusion was received (e.g., prior or concomitant) will be presented on the listing of blood transfusions.

mITT and ITT analysis sets will be used for prior and concomitant blood transfusions analysis.

7.3.9 Other Therapies, Treatments, and Procedures

Any other therapies, treatments, or procedures that occur during the study will be provided in a by-participant listing.

7.3.10 Extent of Study Drug Exposure

The total number of cycles (maximum cycle reached) will be summarized using descriptive statistics for Safety Analysis Sets 1 and 2 by study drug (tamibarotene/placebo and azacitidine) and treatment group. Additionally, the number and percentage of participants will be summarized by the total number of cycles in which they were treated, according to the following categories: 1 to 2, 3 to 4, 5 to 6, and 7+ cycles.

For tamibarotene/placebo, actual and planned total dose, duration of exposure, and actual, planned, and relative dose intensity will be summarized using descriptive statistics for Safety Analysis Sets 1 and 2 by study drug.

- The duration of exposure in months will be calculated as

duration of exposure (months) = (date of last dose of tamibarotene/placebo – date of first dose of tamibarotene/placebo + 1) / 30.4375.

- Actual dose intensity in dose units (mg for tamibarotene/placebo) per month will be calculated as

actual dose intensity (mg/month) = total dose received (mg) / duration of exposure (months).

For the above actual dose intensity formula, the total dose received will be calculated in 2 steps:

1. Sum of the doses in each cycle (Cycle 1 (mg) = X1 dose, Cycle 2 (mg)= X2 dose, ... Cycle n (mg)= Xn dose);

2. Sum of non-zero dosing in all cycles (Cycle 1 (mg) + Cycle 2 (mg) + ... + Cycle n (mg)).

- Planned dose intensity in dose units (mg for tamibarotene/placebo) per month will be calculated as

planned dose intensity (mg/month) = planned total dose (mg) / duration of exposure (months).

The planned total dose of tamibarotene and placebo will be calculated based on the total planned dose collected in the CRF up to EOT. For the above planned dose intensity formula, the planned total dose is the sum of the planned total doses in all cycles.

- Relative dose intensity for tamibarotene/placebo will be calculated as

relative dose intensity = actual dose intensity (mg/month) / planned dose intensity (mg/month).

For azacitidine, duration of exposure, total number of cycles of azacitidine, total number of doses, and average number of doses per cycle will be summarized using descriptive statistics for Safety Analysis Sets 1 and 2 by treatment group.

- *Duration of treatment with azacitidine (months)= (date of last dose of azacitidine – date of first dose of azacitidine + 1) / 30.4375.*
- *Total number of cycles will be calculated based on the maximum cycle reached.*
- *Total number of doses of azacitidine will be calculated in 2 steps:*

  1. Count the number of doses in each cycle.

  2. Sum the counts in step 1.
- *Average number of azacitidine doses per cycle = total number of doses of azacitidine / total number of cycles of azacitidine treatment.*

Summary of duration of exposure to study treatment will include categorical summaries. Duration of exposure will be categorized into time intervals (i.e., <1 month, 1 to <2 months, 2 to <3 months, 3 to <4 months, 4 to <6 months, 6 to <8 months, and ≥8 months); frequency counts and percentages will be presented for the number and percentage of participants in each interval.

Additionally, the number and percentage of participants who have interrupted/held, missed/incorrect, discontinued, increased, and decreased tamibarotene/placebo doses, and the number and percentage of participants who have delayed, held, missed, discontinued, increased, and decreased azacitidine doses will be presented for Safety Analysis Sets 1 and 2 for each treatment group.

The summary of exposure will also be provided for the mITT Analysis Set who have received at least 1 dose of study treatment (tamibarotene, placebo, or azacitidine). The total number of cycles according to the following categories: 1 to 2, 3 to 4, 5 to 6, and 7+ cycles, total dose received, duration of exposure, dose intensity, will be summarized for by treatment group for each study drug. The number of azacitidine doses and the average number of azacitidine dose per cycle will be summarized by treatment group.

All study treatment administration and accountability data will be listed for each study drug.

7.4 Efficacy Evaluation

7.4.1 Datasets Analyzed

All efficacy summaries, except for OS and EFS, will be based on the mITT analysis set. An additional summary of the primary endpoint CR will be based on the PP analysis set. The OS and EFS analyses will be based on the ITT analysis set. The ITT analysis set will be used for sensitivity analyses of efficacy endpoints, including CR, ORR, DOCR, duration of overall response, time to CR, and time to initial response. Efficacy endpoints will be based on all disease assessments during the study, unless otherwise specified.

Listings of MDS disease response assessments, hematology laboratory results, bone marrow morphology and cytogenetics results, and post-treatment survival follow-up will be provided to support the efficacy analyses.

7.4.2 Strata Pooling Strategy

There are a total of 6 strata as follows:

- Very high & North America, Western Europe, and Israel
- High & North America, Western Europe, and Israel
- Intermediate & North America, Western Europe, and Israel
- Very high & Eastern Europe
- High & Eastern Europe
- Intermediate & Eastern Europe

For stratified analysis of any efficacy endpoint, the following steps will be used to pool strata if there is insufficient information in any one stratum.

Step 1: First look at the region strata, if any region stratum <6 participants, this stratum will be pooled to the opposite region stratum with a higher number (e.g., if Very high & Eastern Europe = 2 participants and Very high & North America, Western Europe, and Israel = 3 participants, then 2 participants will be pooled with 3 participants). If there is a tie, then Eastern Europe will be pooled to North America, Western Europe, and Israel. (e.g., if Very high &Eastern Europe = 2 participants and Very high &North America, Western Europe and Israel = 2 participants, Eastern Europe will be pooled to Very high &North America, Western Europe, and Israel).

Step 2: If any stratum is <6 participants after Step 1, then the following strategy of pooling will be taken in order of:

Very High pool to High from the same region (e.g., if Very high & North America, Western Europe, and Israel is <6 participants, then it will be pooled to High & North America, Western Europe, and Israel)

High pool to Very High from the same region

Intermediate pool to High from the same region.

Step 3: If any stratum is <6 participants after Step 2, then repeat Step 1 and Step 2.

7.4.3 Primary Efficacy Endpoint Analysis Methods

7.4.3.1 Primary Estimand Analysis for CR

The final analysis of CR is planned for the time when the first approximately CCI randomized participants out of the planned CCI participants have completed the Cycle 7 Day 1 response assessment or discontinued treatment, whichever comes first.

The primary endpoint is CR achieved by Cycle 7 Day 1, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006). There is expected to be a high correlation between CR achieved by Cycle 7 Day 1 and CR achieved at any time (Silverman 2006; VIDAZA UPSI), and therefore, the same effect size and other design assumptions are justified in the hypothesis testing below. The objective is to demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in CR in RARA-positive newly diagnosed HR-MDS participants.

The null hypothesis (H_0) to be tested is:

In RARA-positive participants with newly diagnosed HR-MDS, CR rate for participants in the tamibarotene plus azacitidine treatment group is the same as the CR rate for participants in the placebo plus azacitidine treatment group.

The alternate hypothesis (H_A) to be tested is:

In RARA-positive participants with newly diagnosed HR-MDS, CR rate for participants in the tamibarotene plus azacitidine treatment group is higher than the CR rate for participants in the placebo plus azacitidine treatment group.

The following analyses will be conducted separately:

[1] Primary analysis of primary estimand (investigator response in mITT analysis set)

[2] Sensitivity analysis of primary estimand (investigator response in ITT analysis set)

[3] Sensitivity analysis of primary estimand (investigator response in PP analysis set)

The CR rate and 95% exact binomial CIs based on the Clopper-Pearson exact method will be calculated by treatment group.

The CR rate will be analyzed using the Cochran-Mantel-Haenszel (CMH) test to compare the 2 treatment groups with a 1-sided 2.5% level of significance. The strata for the test will be those used for stratified randomization, with potential pooling of strata if there is insufficient information in any stratum; the pooling strategy is specified in Section 7.4.2. The point estimate of the difference of the proportions for CR rate between the 2 treatment groups will be provided along with its 95% CI.

For participants who did not receive the correct randomization strata assignment (e.g., the International Prognostic Scoring System (IPSS) risk score category entered in the electronic data capture (EDC) is different from the one assigned in the final randomization schedule), the strata assignment entered in the EDC will be used for the sensitivity analyses. If there is insufficient information in any stratum, the pooling strategy specified in Section 7.4.2 will be used.

A by-participant listing of the derived efficacy data used for analysis of the CR rate will be provided.

7.4.3.2 Supplementary Estimand Analysis for CR

The supplementary endpoint of CR is CR achieved at any time, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006). The analysis will be performed using the same statistical methodologies as applied to the primary estimand for participants in the mITT Analysis Set as described in Section 7.4.3.1.

7.4.4 Key Secondary Efficacy Endpoint Analysis Methods

The key secondary endpoint is the OS, defined as the time from the date of randomization to the date of death from any cause. The objective is to demonstrate the superiority of tamibarotene plus azacitidine compared to placebo plus azacitidine in OS in RARA-positive newly diagnosed HR-MDS participants.

The final analysis of OS will occur when a total of CCI death events is reached, which is approximately 70 months following the enrollment of the first participant.

The H_0 to be tested is:

In RARA-positive participants with newly diagnosed HR-MDS, OS for participants in the tamibarotene plus azacitidine treatment group is the same as the OS for participants in the placebo plus azacitidine treatment group.

The H_A to be tested is:

In RARA-positive participants with newly diagnosed HR-MDS, OS for participants in the tamibarotene plus azacitidine treatment group is superior to the OS for participants in the placebo plus azacitidine treatment group.

The primary analyses of OS will be based on the ITT analysis set.

OS is defined as the time from the date of randomization to the date of death due to any cause. Participants who do not die at the time of a data cutoff date (interim and/or final) will be censored at the last date the participant is known to be alive or at the analysis cutoff date, whichever is earlier.

The distribution of OS will be estimated for each treatment group using Kaplan-Meier methodology and compared between the 2 treatment groups using the log-rank test stratified by IPSS-R risk group (Intermediate, High, and Very High Risk) and by geographical region (North America, Western Europe, plus Israel, and Eastern Europe) at 1-sided 2.5% significance level.

The hazard ratio and corresponding 95% CI between the 2 treatment groups will be estimated using the stratified Cox proportional hazards model with the same stratification factors that are used for the stratified log-rank test. For both the stratified log-rank test and the stratified Cox proportional hazards model, the strata will be those used for stratified randomization, with potential pooling of strata specified in Section 7.4.2.

The number and percentage of participants with an event or who were censored will be summarized by treatment group. OS will be displayed by Kaplan-Meier curves. The 25th percentile, median, and 75th percentile OS (along with corresponding 95% CIs) will also be presented. Landmark analyses (OS rates, 95% CI, and number of participants at risk) will occur at 12 and 24 months post first dose and will be analyzed using the Kaplan-Meier method. The CIs will be calculated using the method of Brookmeyer and Crowley (Brookmeyer 1982).

A by-participant listing of the derived efficacy data used for determining OS will be provided.

Event and censoring rules for OS are provided in Table 12.

Table 12 Assignments for Event and Censoring Dates for Overall Survival Analysis

Situation Number	Situation	Date of Death or Censoring	Outcome
1	Death during the study	Date of death	Event
2	Participant still alive at data cutoff / death occurred after data cutoff	Date of cutoff	Censored

3	Participant lost to follow-up before data cutoff	Date of last known to be alive	Censored

7.4.4.1 Sensitivity Analyses of the Key Secondary Endpoint

Sensitivity Analysis 1: OS analysis based on unstratified log-rank test.

The distribution of OS will be estimated for each treatment group using Kaplan-Meier methodology and compared between the 2 treatment groups using the unstratified logrank test at 1-sided 2.5% significance level.

The hazard ratio and corresponding 95% CI between the 2 treatment groups will be estimated using the Cox proportional hazards model with treatment only as covariate.

The number and percentage of participants with an event or who were censored will be summarized by treatment group. OS will be displayed by Kaplan-Meier curves. The 25th percentile, median, and 75th percentile OS (along with corresponding 95% CIs) will also be presented. Landmark analyses (OS rates, 95% CI, and number of participants at risk) will occur at 12 and 24 months post first dose and will be analyzed using the Kaplan-Meier method. The CIs will be calculated using the method of Brookmeyer and Crowley (Brookmeyer 1982).

Sensitivity Analysis 2: Violation of proportional hazard assumption.

If the proportional hazard assumption does not hold for the OS analysis, the restricted mean survival time (RMST) may be substituted as appropriate. RMST is the participant's life expectancy until time t and can be estimated nonparametrically by the area under the Kaplan-Meier curve up to time t (Tian 2020). Time t will be chosen at the 90th percentile of observed follow-up times (Tian 2020).

The estimated RMST difference in OS between treatment groups will be reported with 95% CIs. The SAS procedure RMSTREG will be used with a linear link function.

Sensitivity Analysis 3: Adjusting for post-study therapy.

A sensitivity analysis of OS will be performed using an inverse probability censored weighting (IPCW) method. Logistic regression will be used to identify the baseline clinical covariates, such as IPSS-R, age, baseline bone marrow blast, etc., that could potentially be confounding factors. A weighted cox regression model will be used by applying the resulting weights from the logistic regression (Watkins 2013).

To obtain a less biased estimate of treatment effect, an IPCW method will be performed for OS, adjusting for post-study therapy. To compensate for the impact of switching to another therapy, participants with similar characteristics to those participants that switched treatment will be identified and a higher weight will be assigned to the

participants that stayed on treatment and a lower weight will be assigned to the participants receiving post study therapy.

To calculate these weights, the likelihood of remaining uncensored will be estimated by logistic regression. Specifically, 2 logistic regression models, one using only baseline covariates and the other using both baseline and time-dependent covariates, will be performed. The coefficient between these 2 estimated probabilities of switching will be used for the assigned weights. Subsequently, participants who switched will have a lower weight than participants that did not switch.

7.4.5 Secondary Efficacy Endpoints Analysis Methods

7.4.5.1 Transfusion Independence

TI is defined as a period of at least 56 days with no RBC or platelet transfusion since the date of randomization to the last dose of study drug + 30 days, the initiation of posttreatment therapy, or death, whichever occurs first. The overall TI rate (based on receiving neither RBC nor platelet transfusion for post-baseline period of any at least 56 days) with respective 95% exact binomial CIs based on the Clopper-Pearson exact method will be calculated by treatment group. The stratified CMH test will be applied to compare the TI rates between the 2 treatment groups. The RBC TI rate and the platelet TI rate, along with 95% exact binomial CIs, will also be calculated and summarized by treatment group. Analyses of the TI rates will be based on the mITT analysis set.

A by-participant listing of the derived efficacy data used for determining the TI rate will be provided.

7.4.5.2 Overall Response

The ORR is defined as achievement of CR, PR, mCR, or HI as determined by the investigator per the modified IWG MDS criteria (Cheson 2006).

The following analyses will be conducted separately:

[1] Primary analysis of ORR estimand (investigator response in mITT analysis set)

[2] Sensitivity analysis of ORR estimand (investigator response in ITT analysis set)

The ORR and 95% exact binomial CIs based on the Clopper-Pearson exact method will be calculated and summarized by treatment group. The stratified CMH test will be applied to compare the ORRs between the 2 treatment groups.

The number and percentage of participants and 95% exact binomial CIs based on the Clopper-Pearson exact method for the individual rates of CR, PR, mCR, and HI; combined rates of CR or PR; combined rates of CR or mCR; combined rates of CR, PR, or mCR; combined rates of CR, mCR, PR, and HI (for participants who failed to achieve

CR, mCR, or PR); and categories of HI (for participants who failed to achieve CR, mCR, or PR) will also be summarized.

For participants who had mCR, the subcategories of mCR with HI and mCR without HI will be summarized. mCR with HI is defined as achieving both mCR and HI as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), whereas mCR without HI is defined as achieving mCR but not HI (based on the same IWG MDS criteria).

CR and overall response may also be analyzed in line with the proposed revised IWG response criteria for HR-MDS (Zeidan 2023). A by-participant listing of the derived efficacy data used for determining the ORR will be provided.

7.4.5.3 Duration of Complete Remission

The duration of complete remission (DOCR) is defined as the duration from the date of first documented evidence of CR to the date of documented relapse of disease or disease progression, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first.

DOCR will be calculated for participants with CR achieved by Cycle 7 Day 1.

The following analyses will be conducted separately:

[1] Primary analysis of DOCR estimand (investigator response in mITT analysis set)

[2] Sensitivity analysis of DOCR estimand (investigator response in ITT analysis set)

DOCR will be estimated using Kaplan-Meier product-limit estimates. The median estimate of DOCR and corresponding 95% CIs based on the method of Brookmeyer and Crowley (Brookmeyer 1982) will be provided for each treatment group. The median follow-up time of DOCR and its corresponding 95% CI will be calculated using the reverse Kaplan-Meier approach (Sathish 2019). The analysis will only be for participants who are CR responders. Event and censoring rules for DOCR are provided in Table 13. A sensitivity analysis of the DOCR will also be performed using the alternative censoring rule identified in situation number 6 by using both mITT and ITT analysis sets.

Table 13 Assignments for Event and Censoring Dates for DOCR Analysis

Situation Number	Situation (Occurring on or After First Date of CR Response)	Date of Event or Censoring	Outcome
1	No post-baseline response assessments after the first CR and the participant has	First date of CR response	Censored

Situation Number	Situation (Occurring on or After First Date of CR Response)	Date of Event or Censoring	Outcome
	not died or developed disease progression/relapse (if the participant has died or developed disease progression /relapse then follow the rules indicated in situations 7 and 8)		
2	Disease progression/relapse documented at scheduled visits within extended loss-to-follow-up time [1]	Date of assessment of disease progression/relapse	Event
3	Disease progression/relapse documented between scheduled (unscheduled) visits within extended loss-to-follow-up time [1]	Date of assessment of disease progression/relapse	Event
4	No disease progression/relapse (or death)	Date of last adequate response assessment [2]	Censored
5	New anticancer therapy other than HSCT started prior to documented disease progression/relapse or death	Date of last adequate response assessment prior to start of new anticancer therapy other than HSCT	Censored
6	HSCT started prior to documented disease progression/relapse or death	Date of last adequate response assessment occurring prior to documented disease progression/relapse or death, irrespective of HSCT date [2]	Censored
		Date of assessment of disease progression/relapse or death [SA]	*Event*
7	Death without an extended loss-to-follow-up time [1]	Date of death	Event
8	Death or developed disease progression/relapse after an extended loss-to-follow-up time (≥2 missed consecutive assessment) [1]	Date of last adequate response assessment [2] prior to the ≥2 missed consecutive assessments	Censored

Abbreviations: CR = complete remission; HSCT = hematopoietic stem cell transplantation

[SA] Alternative rule for handling of events for DOCR.

[1] Extended loss-to-follow-up time is defined as the last non-missing adequate response assessment and event (death or disease progression/relapse). Extended loss-to-follow-up time is 208 days. This cutoff is calculated based on the protocol scheduled procedure: the 28-day treatment cycle has ±3 days window, response assessment is expected to be conducted every 3 months (with ±14 days window). Thus, 2 consecutive scheduled assessments window is 2 × [(30 × 3 + 14] = 208 days. Within extended loss-to-follow up time is defined as ≤208 days; after an extended loss-to-follow-up time is defined as >208 days.

[2] An adequate assessment is defined as an assessment where the Investigator performed all the procedures required for assessment and documented the outcome.

7.4.5.4 Duration of Overall Response

The duration of overall response is defined as the duration from the date of first documented evidence of CR, PR, mCR, or HI, to the date of documented disease progression or relapse of disease as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first.

The following analyses will be conducted separately:

[1] Primary analysis of duration of overall response estimand (investigator response in mITT analysis set)

[2] Sensitivity analysis of duration of overall response estimand (investigator response in ITT analysis set)

The duration of overall response will be estimated using Kaplan-Meier product-limit estimates. The median estimate of duration of overall response and corresponding 95% CIs based on the method of Brookmeyer and Crowley (Brookmeyer 1982) will be provided for each treatment group. The median follow-up time of duration of overall response and corresponding 95% CI will be calculated using reverse Kaplan-Meier approach (Sathish 2019). The analysis will only be for participants who are overall responders. Event and censoring rules for duration of overall response are provided in Table 14. A sensitivity analysis of the duration of overall response will also be performed using the alternative censoring rule identified in situation number 6 in both mITT and ITT analysis sets.

Table 14 Assignments for Event and Censoring Dates for Duration of Overall Response Analysis

Situation Number	Situation (Occurring on or After First Date of Response)	Date of Event or Censoring	Outcome
1	No post-baseline response assessments after the first response and the participant has	First date of overall response	Censored

Situation Number	Situation (Occurring on or After First Date of Response)	Date of Event or Censoring	Outcome
	not died or developed disease progression/relapse (if the participant has died or developed disease progression/relapse then follow the rules indicated in situations 7 and 8)		
2	Disease progression/relapse documented at scheduled visits within extended loss-to-follow-up time [1]	Date of assessment of disease progression/relapse	Event
3	Disease progression/relapse documented between scheduled (unscheduled) visits within extended loss-to-follow-up time [1]	Date of assessment of disease progression/relapse	Event
4	No disease progression/relapse (or death)	Date of last adequate response assessment [2]	Censored
5	New anticancer therapy other than HSCT started prior to documented disease progression/relapse or death	Date of last adequate response assessment prior to start of new anticancer therapy	Censored
6	HSCT (prior to documented disease progression/relapse or death)	Date of last adequate response assessment occurring prior to documented disease progression/relapse or death, irrespective of HSCT date [2]	Censored
		Date of assessment of disease progression/relapse or death [SA]	*Event*
7	Death without extended loss-to-follow-up time [1]	Date of death	Event
8	Death or disease progression/relapse after an extended loss-to-follow-up time (≥2 missed consecutive assessment) [1]	Date of last adequate response assessment [2] prior to the ≥2 missed consecutive assessments	Censored

Abbreviations: CR = complete remission; HSCT = hematopoietic stem cell transplantation

[SA] Alternative rule for handling of events for duration of Overall Response.

[1] Extended loss-to-follow-up time is defined as the last non-missing adequate response assessment and event (death or disease progression/relapse). Extended loss-to-follow-up time is 208 days. This cutoff is calculated based on the protocol scheduled procedure: the 28-day treatment cycle has ± 3 days window, response assessment is expected to be conducted every 3 months (with ± 14 days window). Thus, 2 consecutive scheduled assessments window is 2 × [(30) × 3 + 14] = 208 days. Within extended loss-to-follow up time is defined as ≤208 days; after an extended loss-to-follow-up time is defined as >208 days.

[2] An adequate assessment is defined as an assessment where the Investigator performed all the procedures required for assessment and documented the outcome.

The duration of CR/PR will also be summarized. It is defined as the duration from the date of first documented evidence of CR or PR to the date of documented relapse of disease or disease progression, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006), or death due to any cause, whichever occurs first. Duration of CR/PR will be estimated using Kaplan-Meier product-limit estimates. The median estimate of duration of CR/PR and corresponding 95% CIs based on the method of Brookmeyer and Crowley (Brookmeyer 1982) will be provided for each treatment group. The analysis will only be for participants who are CR/PR responders. Event and censoring rules for duration of CR/PR are provided in Table 13.

7.4.5.5 Time to CR

Time to CR is defined as the duration from the date of randomization to the date of the first documented evidence of CR as determined by the investigator per the modified IWG MDS criteria (Cheson 2006).

Time to CR will be calculated for participants with CR achieved by Cycle 7 Day 1.

The following analyses will be conducted separately:

[1] Primary analysis of time to CR estimand (investigator response in mITT analysis set)

[2] Sensitivity analysis of time to CR estimand (investigator response in ITT analysis set)

Time to CR will be summarized using descriptive statistics by treatment group and only for participants who are responders from the mITT analysis set.

Time to CR will be summarized using descriptive statistics by treatment group and only for participants who are responders from the ITT analysis set for the sensitivity analysis.

7.4.5.6 Time to Initial Response

Time to initial response is defined as the duration from the date of randomization to the date of the first documented evidence of CR, PR, mCR, or HI as determined by the investigator per the modified IWG MDS criteria (Cheson 2006).

The following analyses will be conducted separately:

[1] Primary analysis of time to initial response estimand (investigator response in mITT analysis set)

[2] Sensitivity analysis of time to initial response estimand (investigator response in ITT analysis set)

Time to initial response will be summarized using descriptive statistics by treatment group and only for participants who are responders from the mITT analysis set.

Time to initial response will be summarized using descriptive statistics by treatment group and only for participants who are responders from the ITT analysis set for the sensitivity analysis.

7.4.5.7 Event-Free Survival

EFS is defined as the time from the date of randomization to the date of transformation to AML or death due to any cause, whichever occurs first.

EFS analyses will be based on the ITT analysis set at the time of OS analysis.

EFS will be estimated using Kaplan-Meier product-limit estimates and will be presented for each treatment group and only for participants who have an event or who are censored in the ITT analysis set. Event and censoring rules for EFS are provided in Table 15.

A log-rank test stratified by IPSS-R risk group (Intermediate, High, and Very High Risk) and by geographical region (North America, Western Europe, plus Israel, and Eastern Europe) will be used to analyze the difference in EFS between treatment groups at 2-sided 5% significance level.

The number and percentage of participants with an event or who were censored will be summarized by treatment group. The 25^{th} percentile, median, and 75^{th} percentile EFS (along with corresponding 95% CIs) will also be presented. Landmark analyses (EFS rates, 95% CI, and number of participants at risk) will occur at 6 and 12 months post first dose and will be analyzed using the Kaplan-Meier method. The CIs will be calculated using the method of Brookmeyer and Crowley (Brookmeyer 1982).

Additionally, a separate summary of reasons for censoring, plots of the Kaplan-Meier estimates, a by-participant listing of the derived efficacy data used for the EFS analysis, and a listing of reasons for censoring will be provided.

Sensitivity Analysis 1: EFS analysis using alternative censoring rules.

A sensitivity analysis of EFS will also be performed using the alternative censoring rules identified in situation numbers 4, 7, and 9 in Table 15.

Sensitivity Analysis 2: Violation of proportional hazard assumption.

If the proportional hazard assumption does not hold for the EFS analysis, the RMST may be substituted as appropriate. RMST is the participant's life expectancy until time t and can be estimated nonparametrically by the area under the Kaplan-Meier curve up to time t (Tian 2020). Time t will be chosen at the 90th percentile of observed follow-up times (Tian 2020).

The estimated RMST difference in EFS between treatment groups will be reported with 95% CIs. The SAS procedure RMSTREG will be used with a linear link function.

Table 15 Assignments for Primary and Alternative Event and Censoring Dates for EFS Analysis

Situation Number	Situation	Date of Event or Censoring	Outcome
1	No (or inadequate) baseline disease assessments [1] and the participant has not died (if the participant has died follow the rules for death indicated in situations 7 to 9).	Date of randomization	Censored
2	No post-baseline response assessments and the participant has not died (if the participant has died follow the rules for death indicated in situations 7 to 9)	Date of randomization	Censored
3	Transformation to AML documented at scheduled visits	Date of assessment of transformation to AML	Event

	without extended loss-to-follow-up time [2]		
4	Transformation to AML documented between scheduled visits without extended loss-to-follow-up time [2]	Date of assessment of transformation to AML	Event
		Date of next scheduled response assessment [SA]	*Event [SA]*
5	No transformation to AML (or death).	Date of last adequate response assessment [3]	Censored
6	New anticancer therapy and/or HSCT started (prior to documented transformation to AML or death).	Date of last adequate response assessment occurring prior to documented transformation to AML or death, irrespective of new systemic therapy and/or HSCT date [3]+	Censored
		Date of death or transformation to AML [SA]	*Event [SA]*
7	Death without extended loss-to-follow-up time [2]	Date of death	Event
8	Death or transformation to AML after an extended loss-to-follow-up time [2]	Date of last adequate response assessment [3] prior to documented transformation to AML (prior to missed assessments)	Censored
		Date of death or transformation to AML [SA]	*Event [SA]*
9	Death or transformation to AML after an extended loss-to-follow-up time [2] from randomization	Date of randomization	Censored

Abbreviations: AML = acute myeloid leukemia; EFS = event-free survival; HSCT = hematopoietic stem cell transplantation
[SA] Alternative rule for handling of events for supplementary estimand of EFS.

[1] Adequate baseline assessments are defined as at least one non-missing bone marrow sample or peripheral blood (hemoglobin, neutrophils, platelets, and blasts) result at baseline.

[2] Extended loss-to-follow-up time is defined as the last non-missing adequate response assessment and event (death or disease progression/failure/relapse). Extended loss-to-follow-up time is 208 days. This cutoff is calculated based on the protocol scheduled procedure: the 28-day treatment cycle has ± 3 days window, response assessment is expected to be conducted every 3 months (with ± 14 days window). Thus, 2 consecutive scheduled assessments window is 2 × [30 × 3 + 14] = 208 days. Within extended loss-to-follow up time is defined as ≤208 days; after an extended loss-to-follow-up time is defined as >208 days.

[3] An adequate assessment is defined as an assessment where the Investigator performed all the procedures required for assessment and documented the outcome.

Transformation to AML

Since EFS will be analyzed at the time of the final OS analysis, transformation to AML in the mITT Analysis Set will be analyzed at the time of the primary CR analysis. The number of participants who had disease transformation to AML will be summarized by treatment group, overall, and separately by two groups: those who achieved CR by Cycle 7 Day 1 and those who did not.

7.4.5.8 Change in HRQOL as Measured by the EORTC QLQ-30 and EQ-5D-5L

The EORTC QLQ-30 and EQ-5D-5L test results will be summarized separately by treatment group, parameter, and visit based on the mITT analysis set. For EORTC QLQ-30, results for global health status, each of the 5 functional scales (physical, role, emotional, cognitive, and social functioning), and each of the 8 symptom scales (fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, and diarrhea) will be summarized. The scores for global health status, functional scales, and symptom scales will be calculated and standardized using the methods described in *The EORTC QLQ-C30 Scoring Manual* (3rd Edition) (Fayers 2001). For EQ-5D-5L, results for overall health, each of the 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), and the visual analogue scale will be summarized.

Descriptive statistics will be presented for observed values and changes from baseline by parameter at each visit and endpoint (last observation carried forward [LOCF] up to the 30-day Safety Follow-up Visit) where the questionnaires were scheduled to be collected per the clinical study protocol. Change from baseline in each parameter will be analyzed using a mixed model for repeated measures (MMRM). The model will include treatment, visit (restricted up to the 30-day Safety Follow-up Visit), and treatment-by-visit interaction as fixed effects and baseline as covariate. An unstructured covariance structure will be used to model the between- and within-participant errors.

The Kenward-Roger method will be used to estimate the degrees of freedom. Treatment least square means (LSMs) will be estimated within the framework of the MMRM using type 3 sums of squares. Differences in LSM between treatment groups (and associated standard errors and 95% CI) will be used for statistical inference. The LSM difference, standard error, and 95% CI will be reported for each parameter and visit.

A by-participant listing of all EORTC QLQ-30 and EQ-5D-5L results will be provided.

7.4.5.9 Supplementary Estimand Analyses for Secondary Efficacy Endpoints

The supplementary estimand of secondary efficacy endpoints are DOCR and time to CR with CR achieved at any time, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006).

The analysis will be performed using the same statistical methodologies as applied to the analysis for DOCR and time to CR with CR achieved by Cycle 7 Day 1 for participants in the mITT Analysis Set as described in Section 7.4.5.3 and Section 7.4.5.5.

7.4.6 Exploratory Efficacy Endpoint Analysis Methods

7.4.6.1 Cytogenetic Complete Remission

The cytogenetic CR rate as determined by the investigator per the modified IWG MDS criteria (Cheson 2006) will be analyzed using the same statistical methodologies as applied to the primary efficacy endpoint for participants in the mITT Analysis Set with a chromosomal abnormality at baseline, as described in Section 7.4.3.1. A by-participant listing of the derived efficacy data used for determining cytogenetic CR rate will be provided.

7.4.6.2 Molecular Response

Analysis of molecular response will be detailed in a separate document and is outside the scope of this SAP.

7.4.6.3 Transfusion Requirement

Transfusion requirement will be listed and summarized as described in Section 7.3.8.

7.4.6.4 Consolidation with Allogeneic HSCT Rate

The number and percentage of participants treated and not treated with allogeneic HSCT, along with 95% exact binomial CIs based on Clopper-Pearson exact method, will be summarized by treatment group for the mITT Analysis Set.

7.4.6.5 Complete Remission and Genetic Mutations and/or Gene Expression Markers

Any additional analyses of CR and genetic mutations and/or gene expression markers at baseline, at time of response, and loss of response will be detailed in a separate document and are outside of the scope of this SAP. All biomarker data that are collected in the clinical database will be listed.

7.4.6.6 Response Rate per the Proposed Revised IWG Response Criteria

The following response rate analyses will be performed based on the mITT analysis set:

- CR/CR with partial hematologic recovery (CR_h) Rate

CR/CR_h is defined as achievement of CR or CR_h as determined by the proposed revised IWG MDS criteria (Zeidan 2023, Appendix 2).

The number and percentage of participants achieving CR or CR_h and 95% exact binomial CIs based on the Clopper-Pearson exact method will be calculated and summarized by treatment group.

- $CR/PR/CR_h$ Rate

$CR/PR/CR_h$ is defined as achievement of CR, PR, or CR_h as determined by the proposed revised IWG MDS criteria (Zeidan 2023, Appendix 2). The number and percentage of participants achieving CR, PR, or CR_h and 95% exact binomial CIs based on the Clopper-Pearson exact method will be calculated and summarized by treatment group.

7.4.6.7 Transfusion Rate

Participants who have been on treatment for at least 56 days are considered evaluable for this analysis.

Definitions:

- Baseline Transfusion Dependence:

Transfusion of ≥1 unit of RBCs and/or platelets during the 56 days prior to date of randomization.

- Baseline TI:

No RBC or platelet transfusions during the 56 days prior to date of randomization.

- Post-baseline TI:

Post-baseline TI is defined as a period of at least 56 days with no RBC or platelet transfusion since the date of randomization to the last dose of study drug + 30 days, the initiation of post-treatment therapy, or death, whichever occurs first.

Analyses:

The mITT Analysis Set will be used for the following transfusion rate analyses.

Overall transfusion status (RBC or platelets), RBC transfusion status, and platelet transfusion status will be analyzed separately as below:

- Rate of conversion of baseline transfusion dependence to post-baseline TI in the baseline transfusion-dependent mITT participants.

The number and percentage of mITT participants who are transfusion dependent at baseline and the number and percentage of baseline transfusion-dependent mITT participants who achieved post-baseline TI, along with respective 95% exact binomial CIs, will be summarized by treatment group.

- Rate of maintenance of baseline TI in the baseline transfusion-independent mITT patients.

The number and percentage of mITT participants who are transfusion independent at baseline and the number and percentage of baseline transfusion-independent mITT participants who maintained TI post-baseline will be summarized by treatment group.

In addition, the median number of units of RBC and platelets received in 56 days blocks (e.g., Day 1 to Day 56, Day 57 to Day 112, and etc.) starting from Cycle 1 Day 1 will be reported for each treatment group.

7.4.6.8 Sensitivity Analysis for Selecting RARA Overexpression

A sensitivity analysis will be performed to explore the cutoff values of delta Cq used in the RARA biomarker assay for selecting for responses for participants in the mITT Analysis Set. A logistic regression model with CR (responder vs non-responder) as the binary dependent variable and treatment group, delta Cq, and interaction between treatment group and delta Cq as covariates will be used to find the delta Cq cutoff value that optimally selects for the response from the receiver operating characteristic (ROC) curve. First, test if there is interaction between the treatment group and delta Cq by using the Wald Chi-Square test. If no interaction is found, the interaction term (treatment

group × delta Cq) will be removed from the model, and only treatment group and delta Cq will be kept in the model as covariates. If there is an interaction, further investigation must be performed to qualitatively characterize the differences in RARA overexpression.

The steps to finding the best delta Cq cutoff value are as follows:

Step 1: Use PROC LOGISTIC in SAS with the CTABLE option in the model statement to display a table with statistics for each range of cut points, including true positive percentage and true negative percentage.

```
Proc logistic data=xx;
  Class treatment;
  Model response (Event = '1') = delta_Cq treatment/ctable;
Run;
```

Step 2: Calculate Youden's J statistic as

J = true positive percentage + true negative percentage – 100.

Step 3: Obtain the Max (J) and find the corresponding probability level.

Step 4: Use the fitted logit equation from SAS output:

Logit (probability level) = Intercept + B × delta Cq + C × treatment,

where B is the coefficient of delta Cq, C is the coefficient of treatment, treatment is either 0 or 1 for placebo plus azacitidine or tamibarotene plus azacitidine.

Step 5: Using the probability level from Step 3 and the formula in Step 4, the optimal cutoff value of delta Cq can be calculated for each treatment arm.

Results from the logistic regression model, as well as a plot of the ROC curve, will be presented by treatment group for the mITT analysis set.

7.4.6.9 Exploratory EFS Analysis

Based on the proposed revised IWG response criteria (Zeidan 2023, Appendix 2), EFS is defined as the time from the date of study randomization to the date of the first of the following events:

- PD (based on the criteria in IWG 2023)

- Failure to achieve CR, PR, CR_L, CR_h, or HI within 6 months of study entry
- Relapse from CR, PR, CR_L, CR_h, or HI
- Death from any cause

EFS (months) = (date of event – date of study randomization + 1) / 30.4375.

EFS analysis will be performed as specified in Section 7.4.5.7 for the ITT Analysis Set.

7.4.6.10 Exploratory Comparison of Response Categories and TI

The TI status (achieved TI versus did not achieve TI) of participants in the mITT Analysis Set will be summarized by treatment group for the following response categories: best overall response of CR by Cycle 7 Day 1, CR or PR, mCR with HI, mCR without HI, HI (in participants not achieving CR, PR, or mCR), and participants without a response.

7.4.7 Statistical/Analytical Issues

7.4.7.1 Interim Analysis

There will be 2 planned interim futility analyses, 1 for the CR and 1 for OS, respectively; these will be non-binding and will be conducted by the IDMC. The IDMC may recommend terminating the study for unfavorable results at an interim analysis.

Interim Futility Analysis for CR

An interim futility analysis of the CR rate is planned when the 95th randomized participant has completed the Cycle 7 Day 1 assessment or discontinued treatment, whichever comes first. The study may be terminated at CR interim futility analysis, if the 1-sided p-value from the stratified CMH test comparing the CR rates (tamibarotene plus azacitidine treatment group versus placebo plus azacitidine treatment group) is CCI . The non-binding futility bound is derived using a Hwang-Shih-DeCani (HSD) spending function with gamma = -4 (Hwang 1990). Refer to Section 7.4.3 for a description of the analysis of the CR rate.

Interim Futility Analysis for OS

If the CR rate is statistically significant at the time of the primary efficacy analysis of the CR rates between the 2 treatment groups, an interim OS futility analysis will be conducted at the final CR analysis, which is when approximately CCI information fraction (CCI of the total CCI events) has been observed, and the boundary for OS will be updated based on actual observed information fractions at the interim. The study may be terminated at OS interim futility analysis if the 1-sided p-value from the stratified logrank test comparing the distribution of OS (tamibarotene plus azacitidine treatment group

versus placebo plus azacitidine treatment group) is CCI. The non-binding futility bound is derived using Lan-DeMets spending function. Refer to Section 7.4.4 for a description of the analysis of OS.

Table 16 Planned Stopping Boundaries for CR Futility Analysis and FA

Endpoint	Look	IF	Number of Participants	Boundary (1-sided p-value)	Minimum Critical Difference of CR Rate at the Boundary
CR rate	Futility	0.5	CCI	CCI	0.0045
	FA	1	CCI	CCI	0.1349

Abbreviations: CR = complete remission; FA = final analysis; IF = information fraction.

Table 17 Planned Stopping Boundaries for OS Futility Analysis and FA

Endpoint	Look	IF	Number of Events	Boundaries (1-sided p-value)	Approx. Observed Hazard Ratios at the Boundary
OS	Futility	0.249	CCI	CCI	1.261
	FA	1	CCI	CCI	0.78

Abbreviations: FA = final analysis; IF = information fraction; OS = overall survival

7.4.7.2 Multiplicity Adjustments

The study will be claimed positive when the primary endpoint (CR rate) is statistically significant. To maintain the overall type I error among the primary endpoint and the key secondary endpoint, the CR rate will be tested at 1-sided alpha of CCI, and if it is significant, OS will be tested at 1-sided alpha of CCI.

7.4.7.3 Examination of Subgroups

The subgroup analyses will be conducted using the mITT analysis set for CR-related analyses.

To determine whether the treatment effect is consistent across various subgroups, subgroup analyses of the CR rate, based on whether participants achieved CR by Cycle 7 Day 1, will be performed for the participant subgroup categories listed below. For the geographical region subgroup, the 3 regions (North America, Western Europe, or Israel) were pooled together as 1 category in the subgroup analysis because these 3 regions and Eastern Europe may have differences in available supportive care treatments, access to subspecialty consultants, and highly specialized hospital care.

- IPSS-R Risk Status:
  - Intermediate
  - High
  - Very High

- Geographical Region:
  - North America, Western Europe, or Israel (3 pooled regions)
  - Eastern Europe

- Age Group:
  - ≥18 to <65
  - ≥65 to <75
  - ≥75

- Gender:
  - Female
  - Male

- Race:
  - Asian
  - Black or African American
  - White
  - Other/Multiple

- Bone Marrow Blasts at Baseline:
  - ≤10%
  - >10% to <15%
  - ≥15%

- RARA Overexpression
  - 1st quartile
  - 2nd quartile
  - 3rd quartile
  - 4th quartile

- ECOG
  - 0 or 1
  - 2 or 3

- Disease Type
  - Primary or not-treatment related MDS
  - Secondary or treatment related MDS

- TP53
  - Mutated

  - Wild type

- Cytogenetic Risk Status
  - Very Good or Good
  - Intermediate
  - Poor
  - Very Poor

- ASXL1/TET2
  - ASXL1 mutated
  - ASXL1 wild type/TET2 mutated
  - ASXL1 wild type/TET2 wild type

- RUNX1
  - Mutated
  - Wild type

- DNMT3A
  - Mutated
  - Wild type

- TP53 Multihit
  - Yes
  - No

- Number of Mutated Genes
  - 0
  - 1
  - 2
  - ≥3

In addition, a subgroup analysis for genetic risk, potentially utilizing IPSS-Molecular (IPSS-M) risk category, may be performed. Detailed analysis will be performed separately, and this analysis is outside the scope of this SAP. Subgroup analyses of the different mutations will be based on central NGS data, if available, otherwise local laboratory results will be used.

A forest plot of the estimated treatment effect along with 95% CIs of CR rate will be provided by the subgroup factors listed above. A subgroup analysis may not be performed if the number of participants in the subgroup in each treatment group is not sufficiently large (e.g., <10%). Additional subgroup analyses may be performed post-hoc, as appropriate.

Subgroup analysis of OS will be performed based on the ITT Analysis Set for IPSS-R risk status, geographical region, age, race, gender and any other factors that are deemed appropriate.

7.5 Pharmacokinetic Evaluation

Concentration-time data for tamibarotene will be included in the CSR as listings. Other PK analysis will be provided in a separate report and is outside the scope of this SAP.

7.6 Safety Evaluation

All safety analyses will be conducted for Safety Analysis Set 1, unless otherwise specified. For safety analysis presented by study visit, the baseline value will be defined as the last non-missing value reported prior to the first study drug administration.

7.6.1 Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as those AEs with onset after the first dose of study drug or existing events that worsened after the first dose during the study up until the last dose of study drug + 30 days. TEAEs will be summarized by treatment group. Selected TEAEs will also be summarized. The selected TEAEs include rash, pain, hypertriglyceridemia, venous thrombosis, anemia/RBC count decreased, thrombocytopenia/platelet count decreased, neutropenia/neutrophil count decreased, and leukopenia/white blood cell (WBC) count decreased. A detailed search strategy for the selected TEAE groups is provided in Appendix 1.

AEs with missing start and/or stop dates will be considered TEAEs. Imputation rules for partial and missing dates are described in Table 10.

Verbatim terms on CRFs will be mapped to PTs and SOCs using MedDRA, version 23.1 or higher. AEs will be graded using the National Cancer Institute (NCI) CTCAE version 5. If an AE grade is missing, the previous worst grade under the same PT will be assigned if it is available, otherwise a grade of 4 will be assigned for classification purposes in the TLFs.

The CTCAE Grade will be summarized by grouping Grade 1 and Grade 2 together, then Grade 3 and Grade 4 together, as well as Any Grade for all relevant TEAE tables.

Summaries of the following types will be presented by treatment group:

- Overall summary of the number of TEAEs, TEAEs by relationship to tamibarotene/placebo and azacitidine, TEAEs by CTCAE grade, serious AEs (SAEs), SAEs related to tamibarotene/placebo and azacitidine, selected TEAEs, TEAEs requiring concomitant medications, TEAEs leading to dose interruption of tamibarotene/placebo and azacitidine, TEAEs leading to dose reduction of tamibarotene/placebo and azacitidine, TEAEs leading to discontinuation of

tamibarotene/placebo and azacitidine, TEAEs leading to death (overall summary will also be analyzed based on mITT who received at least 1 study dose)
- TEAEs by SOC and PT
- TEAEs by PT
- TEAEs occurring in ≥10% of participants in any treatment group by PT
- TEAEs related to tamibarotene/placebo by PT
- TEAEs related to azacitidine by PT
- TEAEs by CTCAE grade, SOC and PT
- TEAEs by CTCAE grade and PT
- TEAEs related to tamibarotene/placebo by CTCAE grade and PT
- TEAEs related to azacitidine by CTCAE grade and PT
- SAEs by SOC and PT
- SAEs by PT
- SAEs occurring in ≥5% of participants in any treatment group by PT
- SAEs related to tamibarotene/placebo by PT
- SAEs related to azacitidine by PT
- Selected TEAEs by PT
- Selected TEAEs related to tamibarotene/placebo by PT
- Selected TEAEs related to azacitidine by PT
- Selected TEAEs by PT and CTCAE grade
- Selected TEAEs related to tamibarotene/placebo by PT and CTCAE grade
- Selected TEAEs related to azacitidine by PT and CTCAE grade
- TEAEs requiring concomitant medications by PT
- TEAEs related to tamibarotene/placebo requiring concomitant medications by PT
- TEAEs related to azacitidine requiring concomitant medications by PT
- TEAEs leading to dose interruption of tamibarotene/placebo by PT
- TEAEs leading to dose interruption of azacitidine by PT
- TEAEs leading to dose reduction of tamibarotene/placebo by PT
- TEAEs leading to dose reduction of azacitidine by PT
- TEAEs leading to discontinuation of tamibarotene/placebo by PT
- TEAEs leading to discontinuation of azacitidine by PT

The overall summary of TEAEs will also be created for Safety Analysis Set 2, while all other safety outputs will be created only for Safety Analysis Set 1.

All deaths will be summarized based on the number and percentage of participants. This summary will classify participants by time of death relative to the last dose of treatment (>30 days or ≤30 days) and analyze the primary cause of death. Off-treatment deaths (deaths occurring >30 days after the last dose of treatment) will only be summarized at the prespecified IA and the final OS analysis or based on a regulatory request.

Summaries by SOC and PT will be ordered by descending incidence of SOC and PT within each SOC based on the tamibarotene plus azacitidine group. Summaries by PT will be ordered by descending incidence of PT for the tamibarotene plus azacitidine group. At each level of summarization (e.g., any AE, SOC, and PT), participants experiencing more than 1 TEAE will be counted only once. In the summary of TEAEs by severity grade, participants will be counted once at the highest severity reported at each level of summarization; in the summary of TEAEs by relationship, participants will be counted once at the closest relationship to the indicated study drug (tamibarotene/placebo or azacitidine).

AE data will be presented in data listings. SAEs, selected AEs, AEs requiring concomitant medications, AEs leading to dose interruption/dose reductions/discontinuation of study drug (tamibarotene or azacitidine), AEs leading to death, and deaths will be presented in separate data listings.

7.6.2 Clinical Laboratory Evaluation

All descriptive summaries of laboratory results will be based on data analyzed by the local laboratory and presented in Système International (SI) units. All clinical laboratory data (hematology, serum chemistry, coagulation, urinalysis, and pregnancy test) will be listed. Laboratory measurements identified as abnormal (i.e., outside the normal range) will be flagged in the listings.

Clinical laboratory measurements, including hematology, serum chemistry, and coagulation, will be summarized by treatment group and overall. Descriptive statistics will be presented for observed values and changes from baseline at each visit where parameters were scheduled to be collected per the clinical study protocol.

Urinalysis parameters will be summarized using the following categories: normal, abnormal, not clinically significant or abnormal, clinically significant. A 3-by-3 contingency table will be presented to summarize the shift from the baseline category to the worst post-baseline category. Summary results will include the count and percentage of participants within each shift category and treatment group.

Where applicable, hematology, chemistry, and coagulation results for select parameters will be assigned a toxicity grade based on the NCI CTCAE, version 5.0. If the quantitative criteria for grading are equivalent for 2 grades and the differentiation is described by clinical interventions, the clinical intervention component will not be considered and the highest CTCAE grade will be assigned. Similarly, death related to AE (i.e., Grade 5) cannot be determined with available laboratory-based data collection and, thus, will not be summarized as a category. Laboratory parameters that include multiple sets of criteria for each direction (e.g., separate criteria for potassium measures to assess hyperkalemia and hypokalemia) will be summarized separately to reflect each set of criteria.

Six-by-Six contingency tables will be presented for laboratory tests where toxicity grading can be applied to summarize the shift from the baseline grade to the worst post-baseline grade. Grades will be presented as none (Grade 0; i.e., measurements did not meet any CTCAE criteria for Grades 1 through 4), Grade 1, Grade 2, Grade 3, Grade 4, or missing grade. Summary results will include the count and percentage of participants within each shift category.

In addition, non-CTCAE-graded hematology, chemistry, and coagulation parameters will be summarized using the following categories: low, normal, high, and missing. A 4-by-4 contingency table will be presented to summarize the shift from the baseline category to the worst post-baseline category. Summary results will include the count and percentage of participants within each shift category and treatment group.

Prothrombin Time will not be summarized and will only be available in a listing. International Normalized Ratio (INR) will be summarized.

7.6.2.1 Liver Safety Assessment

The following potentially clinically significant (PCS) criteria in liver function tests for alkaline phosphatase (ALP), alanine transaminase (ALT), total bilirubin (TB), aspartate transaminase (AST), INR, and their combination are defined as follows. The participant's worst post baseline value (highest for all parameters and lowest for ALP, for combination of values where ALP <2 × ULN) will be used.

- ALT or AST >3× upper limit of normal (ULN)
- ALT or AST >5×ULN
- ALT or AST >8×ULN
- ALT or AST >10×ULN
- ALT or AST >20×ULN
- TB >2×ULN
- ALP >1.5×ULN
- INR >1.5
- ALT or AST >3×ULN and TB >2×ULN (*)
- ALT or AST >3×ULN and INR >1.5 (*)
- ALT or AST >3×ULN and ALP <2×ULN and TB >2×ULN (*)

(*) Combination of values measured within same day or within 1 day apart

The number and percentage of participants with PCS values in liver enzymes and TB will be presented by treatment group. A listing of participants who met the above PCS criteria will be provided.

7.6.3 Vital Signs, Physical Findings, and Other Observations Related to Safety

7.6.3.1 Vital Signs

Vital sign measurements will be summarized by treatment group and overall. Descriptive statistics will be presented for results and change from baseline at each time point where vital signs were scheduled to be collected.

A shift from baseline to the worst post-baseline blood pressure category (normal to stage 2 hypertension) summary will be presented by treatment group and overall according to each of the American College of Cardiology (ACC) and American Heart Association (AHA) 2017 guidelines as outlined in the Table 18.

Table 18 Vital Signs ACC/AHA Blood Pressure Categories

Blood Pressure Category	Systolic Blood Pressure		Diastolic Blood Pressure
Normal	<120 mmHg	and	<80 mmHg
Elevated	120 - 129 mmHg	and	<80 mmHg
Stage 1 Hypertension	130 - 139 mmHg	or	80 - 89 mmHg
Stage 2 Hypertension	≥140 mmHg	or	≥90 mmHg

Abbreviations: ACC = American College of Cardiology; AHA = American Heart Association

Additionally, temperature will be summarized by treatment group and overall for each visit according to CTCAE grading criteria guidelines as outlined in Table 19.

Table 19 CTCAE Grading Criteria for Temperature

CTCAE Grade	Criteria
Grade 0	<38.0 degrees C (<100.4 degrees F)
Grade 1	38.0 – 39.0 degrees C (100.4 - 102.2 degrees F)
Grade 2	>39.0 – 40.0 degrees C (102.3 - 104.0 degrees F)
Grade 3	>40.0 degrees C (>104.0 degrees F) for ≤24 hrs
Grade 4	>40.0 degrees C (>104.0 degrees F) for >24 hrs

Abbreviations: CTCAE = Common Terminology Criteria for Adverse Events

All vital signs data including scheduled and unscheduled time points will be presented in the participant data listing.

7.6.3.2 12-Lead Electrocardiogram

Twelve-lead ECG continuous parameters (heart rate, PR interval, QT interval, QRS duration, QTc interval corrected for heart rate using Fridericia formula [QTcF] interval) will be summarized by treatment group and overall. Descriptive statistics will be presented for observed values and changes from baseline at each visit where parameters were scheduled to be collected.

ECG recordings will be classified by the investigator as “normal,” “abnormal, not clinically significant,” or “abnormal, clinically significant.” A 3-by-3 contingency table will be presented to summarize the shift from the baseline category to the worst post-baseline value. Summary results will include the count and percentage of participants within each shift category and treatment group.

Prolonged QTcF intervals will be summarized as QTcF measurements (msec) that are <450, ≥450 to 480, >480 to 500, and >500 msec at each visit where ECG is routinely

collected per the clinical study protocol. Change from baseline categories will also be summarized for measurements that represent a change 0 to <30, ≥30 to 60, or >60 msec relative to the baseline value. Summary results will include the percentage of participants within each category and treatment group.

All ECG data including scheduled and unscheduled time points will be listed. Abnormalities occurring at any point in the study post-baseline will also be included in the ECG listing.

7.6.3.3 Physical Examination

Any abnormal clinically significant finding identified during the physical examination will be recorded as medical history or AE depending on when it started or worsened. A by-participant listing of the physical examination dates will be provided.

7.6.3.4 ECOG Performance Status

ECOG performance scores will be summarized by treatment group and total. The number and percentage of participants will be presented by ECOG performance score (ranging from 0 to 5) at each time point where ECOG performance scoring was scheduled to be collected. The change from baseline by visit will be summarized by descriptive statistics. Additionally, a 6-by-6 contingency table will be presented to summarize the shift from the baseline category to the worst post-baseline score. Summary results will include the count and percentage of participants within each shift category and treatment group. By-participant ECOG performance status results will be provided in a listing.

7.7 Determination of Sample Size

There is expected to be a high correlation between CR achieved by Cycle 7 Day 1 and CR achieved at any time (Silverman 2006; VIDAZA UPSI), and therefore, same effect size and other design assumptions are justified in the sample size determination.

CCI participants will provide CCI power to detect the difference in CR rates between the tamibarotene plus azacitidine treatment group and the placebo plus azacitidine treatment group, with assumed CR rates of CCI versus CCI in the 2 treatment groups, respectively, a 2:1 randomization, and 1-sided alpha of CCI.

A total of CCI death events will provide CCI power to detect the difference in OS between tamibarotene plus azacitidine treatment group and the placebo plus azacitidine treatment group, with assumed median survival of CCI months versus CCI months (hazard ratio = CCI) in the 2 treatment groups, respectively, a 2:1 randomization, and 1-sided alpha of CCI.

A total of approximately CCI participants (CCI in the tamibarotene plus azacitidine treatment group and CCI in the placebo plus azacitidine treatment group) are expected to

be randomized in this study to obtain the CCI death events and final analysis time at approximately 70 months.

8 CHANGES IN THE CONDUCT OF THE STUDY OR PLANNED ANALYSES

The planned analyses outlined in this SAP include changes from the protocol as outlined below:

8.1 Summary of Changes in Version 4.0

8.1.1 Duration of CR and Duration of Overall Response

In Sections 6.4.1.3, 7.4.5.3, and 7.4.5.4, "failure" has been removed as one of the possible events for DOCR and duration of overall response calculation because it is expected that worsening of response after CR/PR/mCR/HI will be captured by disease progression or relapse of disease. This change has been applied throughout the document.

8.1.2 Missing Data and Imputation Rules

In Section 7.1.3.4, Table 11 has been updated to ensure that proper imputation methods are used for both events that should end prior to the treatment start date (e.g., prior cancer therapy) and events that should end after the treatment start date (e.g., concomitant medications). Section 7.1.3.6 was added to detail the imputation rules for missing date of birth and the calculation of age. Sections 7.1.3.7, 7.1.3.8, and 7.1.3.9 were added to detail the imputation rules for missing date of death, calculation of ECG data, and post-baseline missing values, respectively.

8.1.3 Safety Analysis Sets

In Section 7.2.3, clarification was added for the Safety Analysis Set (defined as Safety Analysis Set 1, based on all randomized patients), and Section 7.2.4 was added to define Safety Analysis Set 2 (based on the first approximately CCI randomized patients). All safety analyses will be based on Safety Analysis Set 1, while additional summaries of exposure and overall TEAEs will be created based on Safety Analysis Set 2 to determine the benefit risk compared to the primary analysis. The safety sections have been updated throughout the document to reflect this change.

8.1.4 Central Mutation Data

In Sections 7.3.4 and 7.4.7.3, clarification has been added that the mutation status of genes at baseline will be based on central NGS data, if available. Otherwise, local laboratory data will be used.

8.1.5 Transformation to AML

In Section 7.4.5.7, it was clarified that EFS will only be analyzed at the time of the final OS analysis. Participants with disease transformation to AML will be summarized at the time of the primary analysis.

8.1.6 Exploratory Comparison of Response Categories and TI

Section 7.4.6.10 was added to detail an additional exploratory analysis examining the TI status of participants based on CR in order to determine association between the two.

8.1.7 Subgroup Analyses

In Section 7.4.7.3, new subgroup analyses of CR rate based on disease type (primary versus secondary) and mutation status have been added to find possible predictors.

8.1.8 Proposed Revised IWG Response Criteria for Exploratory Analyses

Appendix 2 has been added with a table summarizing the proposed revised response criteria based on Zeidan 2023 used to calculate response categories. Exploratory analyses for response rate (CR/CRh and CR/PR/CRh) in Section 7.4.6 and EFS in Section 7.4.6.9 have been clarified that these analyses are based on the proposed revised IWG response criteria.

8.1.9 Other Changes

- In Section 7.4.5.4, details about the duration of CR/PR were included. These details were moved from a separate exploratory section to be incorporated in the duration of overall response section for consistency of presentation.

- In Section 7.4.5.8, clarification was added that the visual analogue scale will also be summarized in the analyses of HRQOL since these data are collected.

- In order to allow minor flexibility in enrollment, it was clarified in Sections 7.2.1 and 7.4.3.1 that the primary analysis will be conducted when approximately CCI participants have completed their Cycle 7 Day 1 assessment or discontinued treatment, whichever comes first.

8.2 Summary of Changes in Version 3.0

8.2.1 Modified Intent-to-Treat Analysis Set

The mITT analysis set was added to the SAP to ensure the Food and Drug Administration (FDA)'s feedback of minimum duration of treatment is incorporated. The mITT analysis set includes the first CCI randomized participants; at the time of the CR analysis, this will

ensure at least 6 months of follow-up data for CCI participants, which is the designed sample size for the CR analysis.

8.2.2 Primary Estimand

FDA's Clinical and Statistics Information Request was received on 14 August 2023 regarding the Statistical Analysis Plan (SAP) version 2.0. Per FDA clinical information request 1b, the primary efficacy endpoint is CR defined as CR achieved by Cycle 7 Day 1, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006).

8.2.3 Secondary Estimand

- DOCR was changed to DOCR for CR achieved by Cycle 7 Day 1.
- Time to CR was changed to time to CR for CR achieved by Cycle 7 Day 1.

8.2.4 Supplementary Primary and Secondary Estimands

The supplementary primary and secondary estimands table (Table 6) was added to the SAP to retain the planned CR, DOCR, and time to CR with CR achieved at any time, as determined by the investigator per the modified IWG MDS criteria (Cheson 2006). The corresponding supplementary analyses were added in Section 7.4.3.2 and Section 7.4.5.9.

8.2.5 Other Changes from the Protocol

- In the overall study design of protocol, there is redundant information for participants who progress/relapse. The original wording is: "Patients who progress/relapse will enter Survival Follow-up and will be followed to document the start of subsequent anticancer therapy, **the date of disease progression/relapse**, the date of transformation to AML, the date of death, and OS status until the expected number of OS events to support the final OS analysis is observed." The bolded text was removed from this SAP.
- An exploratory EFS analysis was added to characterize EFS according to the new IWG 2023 response criteria (Zeidan 2023).
- The objectives of DOCR and duration of overall response in the estimand table from the protocol (Table 4) were not correctly specified, and corrections were made in the SAP in Table 5.

8.3 Summary of Changes in Version 2.0

The SAP was updated to Version 2.0 to align with the updated Protocol Version 4.0. The full list of changes is provided in SY-1425-301 SAP Amendment Summary Version 2.0.

9 REFERENCE LIST

Brookmeyer R, Crowley J. A CI for the median survival time. Biometrics. 1982; 38:29-41.

Cheson BD, Greenberg PL, Bennett JM, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. Blood. 2006;108(2):419-25.

Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.

Hwang IK, Shih WJ, DeCani JS. Group sequential designs using a family of type I error probability spending functions. Stat Med. 1990; 9(12):1439-45.

National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 5, November 27, 2017. Available from: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf.

Sathish N, Wu C. Let's Flip: An Approach to Understand Median Follow-up by the Reverse Kaplan-Meier Estimator from a Statistical Programmer's Perspective. PharmaSUG 2019-Paper ST-81.

Silverman LR, McKenzie DR, Peterson BL, et al. Further analysis of trials with azacitidine in patients with myelodysplastic syndrome: studies 8421, 8921, and 9221 by the Cancer and Leukemia Group B. J Clin Oncol. 2006;24(24):3895-903.

Tian L, Jin H, Uno H, et al. On the empirical choice of the time window for restricted mean survival time. Biometrics. 2020;76(4):1157-1166.

VIDAZA™ US Prescribing Information. *Celgene Corporation*. 2022.

Watkins C, Huang X, Latimer N, et al. Adjusting overall survival for treatment switches: commonly used methods and practical application. Pharm Stat. 2013;12(6):348-357.

Zeidan AM, Platzbecker U, Bewersdorf JP, et al. Consensus proposal for revised International Working Group response criteria for higher risk myelodysplastic syndromes. Blood. 2023. Epub ahead of print.

APPENDIX 1. SEARCH STRATEGY FOR SELECTED ADVERSE EVENT GROUPS AND THEIR ASSOCIATED STANDARDIZED MEDDRA QUERY (SMQ) OR HIGH LEVEL TERM (HLT) OR CUSTOM MEDDRA QUERY (CMQ)

Selected TEAE Group	**Search Strategy**
Rash	HLT *Rashes, eruptions and exanthems NEC*
	Hypersensitivity (SMQ) broad search
Pain	HLT *Pain and discomfort NEC*
	HLT *Bone Related Signs and Symptoms*
	HLT *Musculoskeletal and Connective Tissue Pain & Discomfort*
Hypertriglyceridemia	HLT *Elevated triglycerides*
	PT *Blood triglycerides increased*
	PT *Blood triglycerides abnormal*
Venous thrombosis	*Embolic and thrombotic events, venous (SMQ)* narrow search
Anemia/Red blood cell count decreased	*Hematopoietic erythropenia (SMQ)* narrow search
Thrombocytopenia/Platelet count decreased	*Hematopoietic thrombocytopenia (SMQ)* narrow search
Leukopenia/White blood cell count decreased	*Hematopoietic leukopenia (SMQ)* narrow search
Neutropenia/Neutrophil count decreased	*Hematopoietic leukopenia (SMQ)* narrow search filtered for any PT with “neutro”

Abbreviations: CMQ = Custom MedDRA Query; HLT = high level term; NEC = not elsewhere classified; PT = preferred term; SMQ = Standardized MedDRA Query

Note: This list is based on MedDRA version 23.1 and will be updated based on the dictionary or relevant SMQ/HLT updates at the time of the analysis.

APPENDIX 2. PROPOSED REVISED IWG RESPONSE CRITERIA

The table below highlights the criteria used to determine some of the response categories based on the proposed revised IWG response criteria (Zeidan 2023).

Response	**Proposed Revised Response Criteria (2023)**
Complete Remission (CR)	Bone marrow: - ≤5% myeloblasts Peripheral blood: - Hemoglobin (Hgb) ≥10 g/dL - Platelets ≥100 × 10^9/L - Neutrophils ≥1.0 × 10^9/L - Blasts 0%
Marrow CR (mCR)	Eliminated as a response category
Partial Remission (PR)	All CR criteria except: - Bone marrow blasts decreased by ≥50% over pretreatment but still ≥5%
Stable Disease (SD)	Eliminated as a response category
Complete Remission with Limited Count Recovery (CR_L) - CR unilineage (CR_{uni}) or CR bilineage (CR_{bi})	Bone marrow: - <5% myeloblasts Peripheral blood: - Blasts 0% Not meeting CR, but either one (CR_{uni}) or two (CR_{bi}) of the following criteria: - Hgb ≥10 g/dL - Platelets ≥100 × 10^9/L - Neutrophils ≥1.0 × 10^9/L
Complete Remission with Partial Hematologic Recovery (CR_h)	Bone marrow: - ≤5% myeloblasts Peripheral blood: - Platelets ≥50 × 10^9/L - Neutrophils ≥0.5 × 10^9/L - Blasts 0%
Progressive Disease	Fulfilling either one of the criteria below: - ≥50% increase in blasts AND absolute increase of blast percentage by at least 5% from pre-treatment prior to current line of therapy - Repeated (more than once) need for red blood cell OR platelet transfusions within 8 weeks (e.g., separated by

	at least 7 days) not related to acute intercurrent illness (e.g., sepsis, gastrointestinal [GI] bleed) or treatment effect in the absence of hematologic improvement of at least one other blood lineage as defined above. - ≥50% increase in blasts from baseline assessment to ≥ 20% blasts
Disease Relapse	Fulfilling any one of the criteria below: - Bone marrow blasts by at least 5% and ≥50% increase from prior assessment; or reappearance of blasts in the blood; or development of extramedullary disease (myeloid sarcoma) - Decrement of blood counts ≥50% from maximum remission/response levels in Hgb, platelets AND absolute reduction of Hgb by 1.5 g/dl and to Hgb <10 g/dL or platelets $<100 \times 10^9$/L or neutrophils $<1.0 \times 10^9$/L - Repeated (more than once) need for red blood cell or platelet transfusions (e.g., separated by 7 days) not related to acute intercurrent illness (e.g., sepsis, GI bleed) or treatment effect in the absence of hematologic improvement of at least one other blood lineage as defined above.