Statistical Analysis Plan

NCT Number: NCT03268954

Title: A Phase 3, Randomized, Controlled, Open-label, Clinical Study of Pevonedistat Plus Azacitidine Versus Single-Agent Azacitidine as First-Line Treatment for Patients With Higher-Risk Myelodysplastic Syndromes, Chronic Myelomonocytic Leukemia, or Low-Blast Acute Myelogenous Leukemia

Study Number: Pevonedistat-3001

Document Version and Date: Amendment 4, 16 Jun 2021

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

STATISTICAL ANALYSIS PLAN

STUDY NUMBER: Pevonedistat-3001

A Phase 3, Randomized, Controlled, Open-label, Clinical Study of Pevonedistat Plus Azacitidine Versus Single-Agent Azacitidine as First-Line Treatment for Patients with Higher-Risk Myelodysplastic Syndromes, Chronic Myelomonocytic Leukemia, or Low-Blast Acute Myelogenous Leukemia

PHASE 3

Version**:** Amendment 4

Date**:** 16 June 2021

Prepared by:

[redacted], PhD, [redacted], Global Statistics

Based on:

Protocol Version: Amendment 13

Protocol Date: 26 May 2021

1.1 Approval Signatures

Study Title: A Phase 3, Randomized, Controlled, Open-label, Clinical Study of Pevonedistat Plus Azacitidine Versus Single-Agent Azacitidine as First-Line Treatment for Patients with Higher-Risk Myelodysplastic Syndromes, Chronic Myelomonocytic Leukemia, or Low-Blast Acute Myelogenous Leukemia

Approvals:

[redacted], MD, PhD
[redacted], Global Statistics

Date

2.0 TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

3.0 LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
ALT	alanine aminotransferase
AML	acute myelogenous leukemia
ANC	absolute neutrophil count
AST	aspartate aminotransferase
BSA	body surface area
CMH	Cochran-Mantel-Haenszel
CMML	chronic myelomonocytic leukemias
CR	complete remission
CRi	complete remission with incomplete blood count recovery
CHW	Cui-Hung-Wang (test)
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EFS	event-free survival
EQ-5D-5L	EuroQoL 5 dimensions 5 levels (a quality of life questionnaire of the "EuroQoL Group Association" that was expanded to a 5-level instrument)
EORTC	European Organization for the Research and Treatment of Cancer
EORTC QLQ-C30	European Organization for the Research and Treatment of Cancer Core Quality of Life Questionnaire
EOT	end of treatment
EU	European Union
ex-US	all countries and regions excluding the United States
FA	final analysis
FAB	French-American-British
HI	hematologic improvement
HMA	hypomethylating agent
HR MDS	higher-risk myelodysplastic syndromes
HRQOL	health-related quality of life
IA	interim analysis
IDMC	independent data monitoring committee
IRC	independent review committee
IPSS-R	Revised International Prognostic Scoring System
ITT	intent-to-treat
ITD	internal tandem duplication
IV	intravenous(ly)
IWG	International Working Group
IWRS	interactive web response system
K-M	Kaplan-Meier
LFT	liver function test

Abbreviation	Term
MDS	myelodysplastic syndromes
MedDRA	Medical Dictionary for Regulatory Activities
NAE	NEDD8-activating enzyme
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
ORR	overall response rate
ORR 2	overall response rate 2
OS	overall survival
PD	progressive disease; disease progression
PK	pharmacokinetic(s)
PP	per protocol
PR	partial remission
PRO	patient-reported outcome
PS	[Eastern Cooperative Oncology Group] performance status
PT	prefer term
RBC	red blood cell
SAE	serious adverse event
SAP	statistical analysis plan
SC	Subcutaneous
SMQ	Standard MedDRA Query
SOC	system organ class
SOE	Schedule of Events
TEAE	treatment-emergent adverse event
US	United States
WBC	white blood cell
WHO	World Health Organization

4.0 OBJECTIVES

4.1 Primary Objectives

The primary objectives are:

- To determine whether the combination of pevonedistat and azacitidine improves EFS when compared with single-agent azacitidine. (An event is defined as death or transformation to AML in patients with MDS or CMML, whichever occurs first, and is defined as death in patients with low-blast AML.)

4.2 Key Secondary Objective

The key secondary objective is:

- To determine whether the combination of pevonedistat and azacitidine improves OS when compared with single-agent azacitidine.

4.3 Secondary Objectives

Other secondary objectives are:

- To determine whether the combination of pevonedistat and azacitidine improves 6-month and 1-year survival rates when compared with single-agent azacitidine.
- To determine whether the combination of pevonedistat and azacitidine improves 30- and 60-day mortality rates when compared with single-agent azacitidine.
- To determine in patients with HR MDS, patients with HR CMML, and patients with combined HR MDS/CMML whether the combination of pevonedistat and azacitidine delays time to AML transformation when compared with single-agent azacitidine.
- To determine whether the combination of pevonedistat and azacitidine, when compared with single-agent azacitidine, improves the rate of CR (CR in patients with HR MDS or CMML, or low-blast AML), CR+CRi in patients with low-blast AML, CR+marrow CR (in patients with HR MDS or CMML), CR+PR+hematologic improvement (HI) (in patients with HR MDS or CMML), CR+marrow CR+PR (in patients with HR MDS or CMML), CR+marrow CR+PR+HI (in patients with HR MDS or CMML), overall response, overall response by Cycle 6, and overall response 2. Overall response in patients with HR MDS or CMML is defined as CR+PR; overall response in patients with low-blast AML is defined as CR+CRi+PR. Overall response 2 in patients with HR MDS or CMML is defined as CR+PR+HI; overall response 2 in patients with low-blast AML is defined as CR+CRi+PR.
- To determine whether the combination of pevonedistat and azacitidine, when compared with single-agent azacitidine, improves duration of CR (CR for HR MDS or CMML or low-blast AML), CR+CRi for low-blast AML, overall response (CR+PR for HR MDS or

CMML, CR+CRi+PR for low-blast AML), and overall response 2 (CR+PR+HI for HR MDS or CMML, CR+CRi+PR for low-blast AML).

- To determine whether the combination of pevonedistat and azacitidine improves rate of transfusion independence when compared with single-agent azacitidine. RBC and platelet transfusion independence requires that the patient receive no RBC or platelet transfusions for a period of at least 8 weeks during the time period from the first dose of study drug administration through 30 days after the last dose of any study drug.
- To determine whether the combination of pevonedistat and azacitidine increases the duration of RBC transfusion independence, platelet transfusion independence, or platelet and RBC transfusion independence, when compared with single-agent azacitidine.
- To determine whether the combination of pevonedistat and azacitidine improves time to first CR or PR or CRi (in patients with low blast AML) when compared with single-agent azacitidine.
- To determine in patients with HR MDS, patients with HR MDS/CMML, and patients with HR CMML whether the combination of pevonedistat and azacitidine improves rates of HI when compared with single-agent azacitidine.
- To determine whether the combination of pevonedistat and azacitidine does not increase inpatient hospital admission related to HR MDS, CMML, or low-blast AML when compared with single-agent azacitidine.
- To determine whether the combination of pevonedistat and azacitidine delays time to PD, relapse after CR (low-blast AML), relapse after CR or PR (HR MDS/CMML), or death when compared with single-agent azacitidine.
- To determine whether the combination of pevonedistat and azacitidine maintains overall health status/quality of life and fatigue domain scores as measured by the PRO instrument EORTC QLQ-C30 when compared with single-agent azacitidine.
- To collect plasma concentration-time data for pevonedistat to contribute to future population PK analyses of pevonedistat.
- To compare ORR, EFS, and OS in patients who have TP53 mutations and 17p deletions and/or are determined to be in an adverse cytogenetic risk group at Baseline, across treatment arms.
- To determine in patients with HR MDS or CMML, and low-blast AML whether the combination of pevonedistat and azacitidine delays time to subsequent therapy when compared with single-agent azacitidine. Subsequent therapy is defined as agent(s) with antileukemic/anti-MDS activity (eg, cytarabine, anthracyclines, purine analogues, and hypomethylating agents other than azacitidine). Patients who discontinue study treatment to receive single-agent azacitidine off study would not be counted as receiving subsequent therapy.

4.4 Safety Objective

The safety objective is:

- To evaluate the safety of the combination of pevonedistat and azacitidine when compared with single-agent azacitidine.

4.5 Exploratory Objectives

The exploratory objectives are:

- To explore the relationship between response with EFS and OS, including CR with EFS and OS, ORR by Cycle 6 and ORR with EFS and OS, and overall response rate 2 (ORR 2) by Cycle 6 and ORR 2 with EFS and OS.
- To explore the relationship between molecular characteristics in patients identified at Screening (such as somatic mutations, cytogenetic abnormalities, epigenetic modifications, gene expression patterns) with ORR, EFS, OS, and other clinical endpoints of interest, across treatment arms.
- To determine impact of therapy on the depth and durability of response, immune repertoire and treatment emergent resistance by following genomic, epigenomic, and/or protein-based markers in bone marrow or blood and correlation with ORR, EFS, OS, and other clinical efficacy endpoints of interest, across treatment arms.
- To explore the relationship between molecular markers associated with poor prognosis in HR MDS or CMML, or low-blast AML, including genes associated with transformation to AML, identified in bone marrow aspirates at Screening with ORR, EFS, OS, and other clinical endpoints of interest, across treatment arms.
- To assess the time to improvement in functioning and symptoms as measured by the PRO instrument EORTC QLQ-C30, including the following EORTC QLQ-C30 scores: fatigue, physical functioning, role functioning, and dyspnea and using the supplemental items of the EORTC item bank: fatigue, shortness of breath, and physical functioning.
- To evaluate health utilization and calculate utility values using the preference-based PRO instrument EQ-5D-5L.
- To determine in patients with low-blast AML, the percentage of CR and CRi that are cytogenetic remissions.
- To determine measurable residual disease (MRD) status in patients who achieve CR in Cycle 4 or Cycle 7 and determine its relationship to EFS.

4.6 Study Design

This study is a multicenter, global, randomized, controlled, open-label, phase 3 clinical study of the combination of pevonedistat and azacitidine versus single-agent azacitidine administered in patients with HR MDS, CMML, and low-blast AML (see Section 7.1 in the protocol, inclusion criteria 2 and 3 for definitions) who have not previously received chemotherapy or other

antineoplastic agents including HMAs such as decitabine or azacitidine. Patients with nonproliferative CMML (ie, WBC <13,000/μL) are included because these patients were also included in both randomized studies of azacitidine conducted in the US and EU and were shown to have response rates similar to MDS patients.

General eligibility may be assessed before the formal Screening period if it is part of standard clinical practice. However, per the Schedule of Events (SOE) (Appendix C in the protocol), formal screening will occur during the Screening period, which may last up to 28 days before randomization. The sponsor's project clinician (or designee) will confirm patient eligibility before randomization by the investigator.

It is expected that approximately 450 patients, including at least 350 patients with HR MDS or CMML and at least 100 patients with low-blast AML, will be enrolled in this study. At enrollment, patients with HR MDS or CMML, or low-blast AML will be randomized at a 1:1 ratio to receive study drug (either single-agent azacitidine or the combination of pevonedistat and azacitidine) in 28-day treatment cycles. All patients will be stratified into 4 categories: low-blast AML, IPSS-R risk groups of very high, high, or intermediate for MDS/CMML [1]. Note that patients with HR MDS or CMML with indeterminate cytogenetics findings at Screening should be assigned a cytogenetics prognostic variable of 2 points, ie, intermediate, for determining overall Prognostic Risk Category/Score; see Section 9.4.4 in the protocol. All patients will receive azacitidine (75 mg/m^2 [IV or SC]) on Days 1 through 5, Day 8, and Day 9. Patients randomized to the combination arm will also receive pevonedistat (20 mg/m^2 via 60 ([±10]-minute infusion) on Days 1, 3, and 5. Modifications to the dose and schedule may be allowed as detailed in the SOE (Appendix C in the protocol) and Section 8.1 in the protocol.

Patients, including those who achieve a CR, may receive study treatment until they experience unacceptable toxicity, relapse, transformation to AML, or PD as defined in Section 9.4.20 in the protocol. Patients may be allowed to continue study treatment (either treatment arm) if they meet the criteria for PD based only on bone marrow blast count (without AML transformation in patients with HR MDS or CMML) if, in the clinical judgment of the investigator, the patient is still receiving clinical benefit from this treatment and the continuation is endorsed by the sponsor's project clinician (or designee). Patients who meet the criteria for PD and continue on study under these conditions must be reconsented before continuing study treatment. Patients may choose to discontinue therapy at any time.

Patients will attend the End-of-Treatment (EOT) visit 30 days (+10 days) after the last dose of study drug or before the start of subsequent antineoplastic therapy if that occurs sooner.

Following the EOT visit, patients with HR MDS or CMML will enter EFS follow-up, if their disease has not transformed to AML. Patients will have monthly assessments to include physical exam, clinical blood tests, HRQOL assessments, hospitalization assessment, and disease assessment as outlined in the SOE (Appendix C in the protocol). Patients who discontinue study treatment without evidence of progression (ie, PD or transformation to AML) will have a bone marrow aspirate and hematology tests (see Table 9.d in the protocol) sent to the central laboratory at time of suspected progression (see SOE [Appendix C in the protocol] and Table A in the protocol). Patients will continue monthly EFS follow-up study visits until their disease

transforms to AML or they start subsequent therapy (see Study Diagram in Appendix B in the protocol). Patients who have started subsequent therapy will have EFS follow-up but will not be required to have monthly visits; at the time of suspected transformation to AML they will have a bone marrow aspirate and hematology tests (specimens sent to central laboratory).

Following the EOT visit, patients with low-blast AML will enter response follow-up, if they have no evidence of PD and they have not started subsequent therapy. Patients will have monthly assessments to include physical exam, clinical blood tests, HRQOL assessments, hospitalization assessment, and disease assessment as outlined in the SOE (Appendix C in the protocol). Patients who discontinue study treatment while not in CR and without evidence of PD will have a bone marrow aspirate and hematology tests sent to the central laboratory at the time of suspected PD. Patients who discontinue treatment while in CR will also have a bone marrow aspirate and hematology tests performed at the time of suspected relapse sent to the central laboratory (see Table A in the protocol). Patients will continue monthly response follow-up visits, until they relapse from CR or meet the criteria for PD (see Study Diagram in Appendix B in the protocol).

Following the EFS and response follow-up visits, or the EOT visit (for patients with HR MDS or CMML who discontinue study treatment because of transformation to AML, or patients with low-blast AML who discontinue study treatment because of PD), patients will enter OS follow-up and will be contacted every 3 months until death to document subsequent therapies and survival status (see Study Diagram in Appendix B in the protocol).

Disease response assessments for all HR MDS and CMML patients will be based on the Modified International Working Group (IWG) response criteria for MDS [2] as detailed in Section 9.4.20 in the protocol. Disease response assessments for patients with low-blast AML will be based on the Revised Recommendations of the IWG for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia as detailed in Section 9.4.20 in the protocol [3]. Formal disease assessments for study endpoints will be determined based on bone marrow aspirate blast counts analyzed at a central laboratory, clinical laboratory evaluations performed at a central laboratory (local laboratory results may be used for time-sensitive clinical decisions), and local transfusion data.

Inpatient hospital admissions related to HR MDS or CMML, or low-blast AML, as well as transfusion independence, will be monitored as secondary efficacy endpoints. RBC and platelet transfusion independence requires that the patient receive no RBC or platelet transfusions, respectively, for a period of at least 8 weeks. Treatment-emergent resistance will also be monitored.

A bone marrow aspirate and biopsy will be collected at Screening, and bone marrow aspirates will be collected during treatment and follow-up for blast count evaluation (to inform disease burden assessment).

Bone marrow aspirates will be obtained at Screening and at additional time points described in Table A and in Section 9.4.24 in the protocol for assessing response and/or for translational research purposes. Samples will be collected and analyzed from patients in both treatment arms

and sent to a central laboratory. Bone marrow aspirates collected at Screening will be used to analyze tumor cytogenetics, baseline somatic mutations, and other molecular characteristics such as gene expression profile and epigenetic status. Bone marrow aspirates collected at the specified time points during treatment and/or at relapse will be used to evaluate depth and duration of response by following parameters such as residual tumor cells, residual mutation load, and changes in epigenetic modifications. Such analysis will also be used to identify treatment-emergent mutations. Developing potential biomarkers of pevonedistat-mediated activity may require analysis of the data from this study in combination with data from other clinical studies of pevonedistat.

Sparse sampling for the determination of pevonedistat plasma concentrations and, if appropriate, its metabolites will be collected from each patient in the Combination Pevonedistat Plus Azacitidine Arm as described in the SOE (Appendix C in the protocol) to contribute to a population PK analysis of pevonedistat co-administered with azacitidine.

Adverse events and ECOG PS will be assessed, and ECGs, clinical laboratory values, and vital signs will be obtained, to evaluate the safety and tolerability of the study drug treatments. Toxicity will be evaluated according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), version 4.03, effective date 14 June 2010. Dose modification guidelines are presented in Section 8.3 in the protocol.

PROs will be evaluated using the EORTC QLQ-C30, version 3.0, and EQ-5D-5L questionnaires.

5.0 ANALYSIS ENDPOINTS

5.1 Primary Endpoints

The primary endpoint is:

- EFS: time from randomization to the date of an EFS event (defined as death or transformation to AML in patients with MDS or CMML, whichever occurs first, and defined as death in patients with low-blast AML).

5.2 Key Secondary Endpoint

The key secondary endpoint is:

- OS.

5.3 Other Secondary Endpoints

Other secondary endpoints are:

- Six-month and 1-year survival rates.
- Thirty-day and 60-day mortality rates.
- Time to AML transformation in patients with HR MDS, patients with HR CMML, and patients with combined HR MDS/CMML.

- Rate of CR (CR in patients with HR MDS or CMML, or low-blast AML), CR+CRi in patients with low-blast AML, CR+marrow CR (in patients with HR MDS or CMML), CR+PR+HI (in patients with HR MDS or CMML), CR+marrow CR+PR (in patients with HR MDS or CMML), CR+marrow CR+PR+HI (in patients with HR MDS or CMML), overall response, overall response by Cycle 6, and overall response 2. Overall response in patients with HR MDS or CMML is defined as CR+PR; overall response in patients with low-blast AML is defined as CR+CRi+PR. Overall response 2 in patients with HR MDS or CMML is defined as CR+PR+HI; overall response 2 in patients with low-blast AML is defined as CR+CRi+PR.
- Duration of CR (CR for HR MDS or CMML, or low-blast AML), CR+CRi for low-blast AML, overall response (CR+PR for HR MDS or CMML, CR+CRi+PR for low-blast AML), and overall response 2 (CR+PR+HI for HR MDS or CMML, CR+CRi+PR for low-blast AML).
- Rates of RBC and platelet transfusion independence.
- Duration of RBC transfusion independence, platelet transfusion independence, and platelet and RBC transfusion independence.
- Time to first CR or PR or CRi (for patients with low-blast AML).
- Rates of HI in patients with HR MDS, patients with HR CMML and patients with HR MDS/CMML.
- Patients who have inpatient hospital admission(s) related to HR MDS or CMML (collected through transformation to AML or until initiation of subsequent therapy, whichever occurs first) or low-blast AML (collected through initiation of subsequent therapy).
- Time to PD, relapse after CR (low-blast AML), relapse after CR or PR (HR MDS/CMML), or death.
- HRQOL assessed using the EORTC QLQ-C30
- Plasma concentration-time data for pevonedistat.
- ORR, EFS, and OS in patients who have TP53 mutations, 17p deletions, and/or are determined to be in an adverse cytogenetic risk group in both treatment arms.
- Time to Subsequent Therapy.

5.4 Safety Endpoints

The safety endpoints are:

- AEs.
- SAEs.
- Results of clinical laboratory tests.

- ECOG performance status (PS)

5.5 Exploratory Endpoints

The exploratory endpoints are:

- Correlation of response with EFS and OS in ITT and the disease subpopulations (patients with HR MDS/CMML and patients with low-blast AML):
  - CR with EFS and OS.
  - ORR by Cycle 6 and ORR with EFS and OS.
  - ORR 2 by Cycle 6 and ORR 2 with EFS and OS.
- Identification of molecular abnormalities such as somatic mutations, cytogenetic abnormalities, as well as methylation and gene expression patterns in patients at Screening in both treatment arms and correlation with clinical efficacy.
- Evaluation of changes in markers such as somatic mutations, cytogenetic abnormalities, methylation and proteomic signatures as well as immune repertoire in bone marrow aspirates or blood at specified times during therapy and at relapse, in both treatment arms, and correlation with clinical efficacy.
- Relationship with molecular markers associated with poor prognosis in HR MDS or CMML, or low-blast AML such as FLT3 ITD, RUNX-1, EZH2, ASXL1, N-RAS, K-RAS, in both treatment arms and correlation with clinical efficacy.
- Time to improvement in functioning and symptoms as measured by the PRO instrument EORTC QLQ-C30, including the following EORTC QLQ-C30 scores: fatigue, physical functioning, role functioning, and dyspnea and using the supplemental items of the EORTC item bank: fatigue, shortness of breath, and physical functioning.
- Evaluation of health utilization and calculate utility values using a preference-based PRO instrument EQ-5D-5L.
- Rates of overall cytogenetic response by treatment arms will be analyzed using IRC assessments in ITT population, patients with HR MDS/CMML and patients with low-blast AML, separately.
- Rate of cytogenetic complete remissions among patients with best overall response of CR or CRi.
- MRD status in patients who achieve CR in Cycle 4 or Cycle 7 and determine its relationship to EFS.

6.0 DETERMINATION OF SAMPLE SIZE

There is 1 primary endpoint, of EFS, and 1 key secondary endpoint of OS. The total number of patients was calculated on the basis of maintaining 83% power to test the key secondary endpoint OS in patients with HR MDS at a 1-sided alpha level of 0.025, as well as sufficient

representation of patients with low-blast AML. The study is also adequately powered to test the primary endpoint, EFS, in patients with HR MDS and in the ITT population for full approval in the US submission and the ex-US submission, respectively, as well as to test OS in the ITT population.

A total of approximately 450 patients, including at least 350 patients with HR MDS or CMML and at least 100 patients with low-blast AML, will be enrolled, with 320 patients with HR MDS required to obtain 202 OS events in patients with HR MDS. Patients will be randomized in a 1:1 ratio to the 2 treatment arms, stratified by low-blast AML, IPSS-R risk group of very high, high, or intermediate for MDS/CMML [1]. Assuming an enrollment rate that increases from an initial 1 patient per month to 22 patients per month after the first 10 months, with a dropout rate of approximately 10%, the study will take approximately 27 months for patient accrual and an additional approximately 38 months for follow-up from the enrollment of the last patient to obtain the target 202 OS events in HR MDS population.

The primary endpoint of EFS will be analyzed based on the patients with HR MDS for US submission and the ITT population for ex-US submission at the second interim analysis (IA2), which will be performed when the adaptive EFS event size (from a minimum of 147 to a maximum of 249) has occurred in patients with HR MDS. The minimal planned event size of 147 EFS events from patients with HR MDS is based on an optimistic assumption of a hazard ratio of 0.585 (median EFS of 17.09 months in the combination arm versus 10 months in the azacitidine alone arm) with approximately 90% power at a 1-sided alpha of 0.025. The maximal planned event size of 249 EFS events from patients with HR MDS is based on a relatively conservative assumption of a hazard ratio of 0.663 (median EFS of 15.08 months in the combination arm versus 10 months in the azacitidine alone arm) with approximately 90% power at a 1-sided alpha of 0.025. It is projected that the EFS event size from the ITT population ranges approximately from 158 (approximately 92% power, 1-sided alpha=0.025, HR=0.585, median EFS of 22.22 months in the combination arm versus 13 months in the azacitidine alone arm) to 305 (approximately 95% power, 1-sided alpha=0.025, HR=0.663, median EFS of 19.61 months in the combination arm versus 13 months in the azacitidine alone arm).

The key secondary endpoint of OS will be analyzed at the final analysis, which will be performed when approximately 202 OS events have occurred in patients with HR MDS. With the assumption of a hazard ratio of 0.663 for OS in patients in HR MDS (median OS of 36.95 months in the combination arm versus 24.5 months in the azacitidine alone arm), 202 OS events provide approximately 83% power at a 1-sided alpha of 0.025. It is projected that approximately 280 OS events will occur to the ITT population, providing approximately 93% power to test OS in the ITT population (1-sided alpha=0.025, HR=0.663, median OS of 36.95 months in the combination arm versus 24.5 months in the azacitidine alone arm).

7.0 METHODS OF ANALYSIS AND PRESENTATION

7.1 General Principles

All statistical analyses will be conducted using SAS® Version 9.2, or higher. Where appropriate, variables will be summarized descriptively by study visit. For the categorical variables, the count

and proportions of each possible value will be tabulated by treatment group. The denominator for the proportion will be based on the number of subjects who provided non-missing responses to the categorical variable. For continuous variables, the number of subjects with non-missing values, mean, median, SD, minimum, and maximum values will be presented by treatment. The Kaplan-Meier survival curves and 25th, 50th (median), and 75th percentiles will be provided along with their 95% CIs for time-to-event data.

Means and medians will be presented to 1 more decimal place than the recorded data. The standard deviations (SDs) will be presented to 2 more decimal places than the recorded data. Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate.

P-values will be rounded to 3 decimal places prior to assessment of statistical significance.

The analyses will be performed and summarized using Tables and/or figures for the corresponding analysis sets of overall population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML, and patients with low-blast AML), wherever it is appropriate. The by-patient listings will also be provided for the above specified five populations, wherever it is appropriate.

7.1.1 Randomization and Stratification

The randomization scheme will be generated by an independent vendor. Before dosing, a randomization number will be assigned to each patient. The randomization assignment will be implemented by an interactive web-based response system (IWRS).

Patients will be randomized to receive the combination of pevonedistat and azacitidine or azacitidine alone in a 1:1 ratio, stratified into 4 categories: low-blast AML, IPSS-R risk groups of very high, high, or intermediate for HR MDS/CMML [1].

Randomization error may occur during study conduct; if this happens in the study, corrected stratification factors will be used in the primary analyses. Both original stratification factors at randomization and corrected stratification factor will be summarized.

7.1.2 Blinding and Unblinding

This is an open-label study; investigators and patients will know their individual treatment assignment. However, aggregate results by treatment group will be blinded to the study team, investigators, patients, and IRC throughout the study conduct. Individual treatment assignment will be blinded to the study team and IRC throughout the study conduct. The independent statistical center (ISC) and IDMC will have access to un-blinded individual patient level data. The periodic safety analyses will be generated for the IDMC by an ISC.

7.1.3 Definition of Study Days

Study Day 1 is defined as the date on which a subject is administered their first dose of the medication. Other study days are defined relative to the Study Day 1 with Day 1 being Study Day 1 and Day -1 being the day prior to Study Day 1.

7.1.4 Definition of Study Visit Windows

All data will be categorized on the basis of the scheduled visit at which they are collected. These visit designators are predefined values that appear as part of the visit tab in the electronic case report form (eCRF). The analysis of PK data will be based on the actual elapsed time post dose.

7.1.5 Conventions for Missing Adverse Event Dates

Every effort will be made to avoid missing/partial dates in on-study data.

Adverse events with stop dates that are completely or partially missing will be imputed as follows:

- If the stop date has a month and year but the day is missing, the last day of the month will be imputed.
- If the stop date has a year but the day and month are missing, the 31st of December will be imputed.

After the imputation, the imputed dates will be compared against the date of death, if available. If the date is later than the date of death, the date of death will be used as the imputed date instead.

Adverse events with start dates that are completely or partially missing will be imputed as follows:

- If the start date has a month and year but the day is missing, the first day of the month will be imputed.
- If this date is earlier than the first dose date, then the first dose date will be used instead.
- If this date is later than the stop date (possibly imputed), then the stop date will be used instead.
- If the start date has a year, but the day and month are missing, the 15th of June will be imputed.
- If this date is earlier than the first dose date, then the first dose date will be used instead.
- If this date is later than the stop date (possibly imputed), then the stop date will be used instead.

If the start date of an event is completely missing, then it is imputed with the first dose date.

7.1.6 Conventions for Missing Concomitant Medication/ Subsequent Therapies Dates

Concomitant therapies with start dates that are completely or partially missing will be analyzed as follows:

- If the start date has a month and year but the day is missing, the therapy will be included in the summary table if the month and year of the start date of the event are:

- On or after the month and year of the date of the first dose of study drug

  and

- On or before the month and year of the date of the last dose of study drug plus 30 days
- If the start date has the year but the day and month are missing, the therapy will be included in the summary table if the year of the start date of the event is:
- On or after the year of the date of the first dose of study drug

  and

- On or before the year of the date of the last dose of study drug plus 30 days

If the start date of an event is completely missing, then the therapy will be included in the summary table.

Subsequent therapies with start dates that are completely or partially missing will be analyzed as follows:

- When the month and year are present but the day is missing:
- If the onset month and year are the same as the month and year of the last dose of study drug, the day of the last dose + 1 will be imputed.
- If the onset month and year are not the same as the month and year of the last dose of study drug, the first day of the month is imputed.
- When only a year is present:
  - If the onset year is the same as the year of the last dose of study drug, the date of last dose + 1 will be imputed.
  - If the onset year is not the same as the year of the last dose of study drug, the first day of the year is imputed.
- If no components of the onset date are present, the date of the last dose of study drug + 1 will be imputed.

7.1.7 Missing/Partial Dates in Screening Visit

The following rules apply to dates recorded in the Screening visits:

- If only the day component is missing, the first day of the month will be used if the year and the month are the same as those for the first dose of study drug; otherwise, the 15th will be used.
- If only a year is present, and it is the same as the year of the first dose of study drug, the 15th of January will be used unless it is later than the first dose, in which case the date of the first of January will be used, unless other data indicate that the date is earlier.

If only a year is present, and it is not the same as the year of the first dose of study drug, the 15th of June will be used, unless other data indicate that the date is earlier.

7.1.8 Definition of Baseline Values

Unless otherwise specified, for each safety parameter, the baseline value is defined as the value collected at the time closest to, but prior to, the start of study drug administration. For analysis of ECG data, the baseline value is the screening value.

7.2 Analysis Sets

7.2.1 Intent-to-Treat Population

The Intent-to-Treat (ITT) population is defined as all patients who are randomized. Patients in this population will be analyzed according to the treatment they were randomized to receive, regardless of any dosing errors.

The ITT population will be used for the primary, secondary efficacy analyses, and resource utilization and patient reported outcome analysis.

7.2.2 Safety Population

The safety population is defined as all patients who receive at least 1 dose of pevonedistat plus azacitidine or azacitidine alone. Patients will be analyzed according to the actual treatment they received. Patients who received any dose of pevonedistat will be included in the combination pevonedistat plus azacitidine arm, and patients who did not receive any dose of pevonedistat will be included in the single-agent azacitidine arm, regardless of their randomized treatment.

Safety population will be used for all safety related analyses such as adverse events (AE), concomitant medication, laboratory tests, and vital signs.

7.2.3 Per-protocol (PP) Population

The PP population is a subgroup of the ITT population, consisting of all patients who receive at least one dose of study drug and who do not have major protocol deviations, as determined by the study clinician, who is blinded to study drug assignment.

All decisions to exclude patients from the PP population will be made by the Takeda Project Clinician or designee prior to unblinding the study for IAs or FA purpose.

The PP population will be used as a sensitivity analysis of the ITT population for the primary efficacy endpoint EFS.

All patients in the PP population will be analyzed according to the actual treatment received.

7.2.4 Response-evaluable Population

The response-evaluable population is defined as patients who receive at least 1 dose of study drug, have a disease assessment at Screening (baseline evaluation), and at least 1 postbaseline disease assessment.

7.3 Disposition of Subjects

Patient disposition includes the number and percentage of patients for the following categories: patients randomized, patients in each of the study populations, patients discontinued from treatment, primary reason for discontinuation from treatment, patients discontinued from the study, primary reason for discontinuation from the study, and completion of study. All percentages will be based on the number of patients in the ITT population or the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML).

Patients will be considered to have completed the study if they are followed until death or until the sponsor terminates the study.

A listing will present data concerning patient disposition.

The stratification strata at randomization is based on the disease diagnosis at screening. For the patients whose disease diagnosis changed after screening and before randomization, a listing will be provided.

7.4 Demographic and Other Baseline Characteristics

7.4.1 Demographics

Baseline demographics will be summarized for all patients in the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately. Baseline demographic data to be evaluated will include age at date of informed consent, sex, ethnicity, race, height, weight, body surface area (BSA), and other parameters as appropriate.

Patient enrollment by region and country will also be summarized by treatment arms.

BSA is calculated using the following formula based on the patient's height and weight collected at screening. If a weight at screening is not available, the weight at Cycle 1 Day 1 pre-dose can be used.

$$\text{BSA} = \sqrt{\frac{Ht(cm) \times Wt(kg)}{3600}}$$

No inferential statistics will be generated.

Demographic data will also be presented in a by-patient listing.

7.4.2 Inclusion/Exclusion Criteria

All inclusion/exclusion information on enrolled patients will be included in a by-patient listing. The listing will include whether all criteria were satisfied. For patients who did not satisfy the criteria, the criteria number will be listed with the deviation.

7.4.3 Baseline Disease Status

Analyses on baseline disease characteristics will be performed for the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

Baseline disease characteristics (HR MDS/CMML: disease type of de novo or secondary, disease subtype if secondary, French-American-British (FAB) category, WHO classification of tumors, Modified Charlson Comorbidity Index, and IPSS-R category; low-blast AML: disease type of de novo or secondary, disease subtype if secondary, revised WHO classification of AML, evidence of extramedullary disease, cytogenetics and mutation based risk stratification; HRMDS/CMML/AML: months from initial diagnosis) and ECOG performance status will be summarized for all patients by treatment arm. Separate by-patient listings will also be presented for baseline disease characteristics and ECOG performance status.

HRMDS/CMML patients will be classified into 3 categories based on the IPSS-R score: intermediate, high, very high.

Separate tables for HR MDS/CMML patients, HR MDS, HR CMML and low-blast AML patients will summarize the numbers and percentages of patients who had prior therapy, prior radiation (including the total lifetime dose of radiation received and months from last prior radiation to first dose of pevonedistat), prior surgery (including months from last prior surgery to first dose of pevonedistat) and prior transplants for all patients in the safety population. Separate by-patient listings will also be presented for prior therapies, prior radiation, prior surgery, and prior transplants.

Months from diagnosis to the randomization date for each treatment is calculated by:

$$\frac{\text{randomization date - date of diagnosis}}{365.25/12}$$

Distribution of stratification factor also will be summarized.

A separate table will summarize the characteristics of the bone marrow aspirate samples taken at screening. This will include myeloid/erythroid ratio, myeloblast percentage, cytogenetic results, cell maturing status, and presence of Auer rods. The table also will include a summary of the percentage of patients with baseline mutations in genes known to be prognostic indicators in HR MDS/CMML or AML (e.g. TP53, SF3B1, ASXL1, NPM1, FLT3, CEBPA etc.). Baseline bone marrow aspirate data and mutation status of genes also will be presented in by-patient listings.

A separate table will summarize the results of the bone marrow biopsy samples taken at screening. This will include myeloblast percentage. Bone marrow biopsy data and sample collection also will be presented in by patient listings.

HR MDS/CMML patients will be classified into five IPSS-R cytogenetic risk groups: very good, good, intermediate, poor, and very poor, using their cytogenetic results collected at screening. The screening IPSS-R cytogenetic risk classification results will be summarized by treatment arm and will also be included in a by-patient listing. Combining cytogenetic risks and baseline genetic mutations, low-blast AML patients will also be classified into the European

LeukemiaNet (ELN) risks groups: Favorable, Intermediate, and Adverse. The ELN risk classification results will be summarized by treatment arm and will also be included in a by-patient listing. The ELN risk classification criteria uses the version published by Dohner H et al. 2017 [5].

A listing will be generated for patients who receive hydroxycarbamide (hydroxyurea) at enrollment, which includes screening WBC and screening bone marrow aspirate myeloblasts.

7.5 Medical History and Concurrent Medical Conditions

General medical history and prior medications will be listed for the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

7.6 Medication History and Concomitant Medications

Concomitant medications will be coded by Preferred Term using the World Health Organization (WHO) Drug Dictionary. The number and percentage of patients taking concomitant medications from the first dose through the end of the on-treatment period will be tabulated by Anatomical Therapeutic Chemical (ATC) classification pharmacological subgroup and WHO drug generic term for each treatment arm in the Safety population. A by-patient listing will also be presented for concomitant medications.

Concomitant procedures will not be coded but will be presented in a data listing for the Safety population.

7.7 Study Drug Exposure and Compliance

7.7.1 Extent of Exposure

Pevonedistat

An overall summary of drug exposure for pevonedistat will be presented, including the number of cycles, the mean number of doses per cycle, and the distribution of the number of cycles (numbers and percentages of patients who are treated for at least 1 cycle, 2 cycles, 3 cycles, …), for the treatment arm of pevonedistat plus azacitidine in the safety population, the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

Patients will be considered to have been treated for a cycle if they receive at least one dose of pevonedistat during the 28 days of that cycle.

The mean number of doses per cycle will be calculated for each patient and then summarized for the treatment arm of pevonedistat plus azacitidine in the safety population.

Dosing intensity will be summarized for the treatment arm of pevonedistat plus azacitidine in the safety population. Percent Dosing Intensity will be calculated using the following equations for Daily Expected Dose (mg), Daily Prepared Dose (mg), and Daily Dose Received (mg):

$$\text{Daily Expected Dose} = \text{Dose Level Assigned at Study Entry } (\text{mg/m}^2) * \text{Body Surface Area } (\text{m}^2)$$

$$\text{Daily Prepared Dose} = \text{Scheduled Dose Level } (\text{mg/m}^2) * \text{Body Surface Area } (\text{m}^2)$$

$$\text{Daily Dose Received} = \text{Daily Prepared Dose} * \left(\frac{\text{Volume of IV bag actually infused (mL)}}{\text{Prepared Volume}}\right)$$

Daily Expected Dose and Daily Prepared Dose may differ if there are dose decreases. The scheduled dose level will be collected on the electronic case report form (eCRF) for each dosing day. Daily Expected Dose and Daily Prepared Dose will be calculated on the BSA measured at baseline unless the patient experiences a >5% change in body weight from the weight used for the most recent BSA calculation. If the patient experiences a >5% change in body weight from the weight used for the most recent BSA calculation, body surface area (BSA) will be calculated on Cycle 1, Day 1, and at subsequent visits and will be used for the calculation of Daily Expected Dose and Daily Prepared Dose at the corresponding visits.

Total Dose Received, Total Dose Expected, and Dosing Intensity for MLN4924 will be based on the following formulas:

$$\text{Total Dose Received} = \text{Sum of Daily Dose Received across all days that MLN4924 was administered}$$

$$\text{Total Dose Expected} = \text{Daily Expected Dose} * 3 \text{ doses per cycle} * \text{number of treated cycles}$$

$$\text{Percent Dosing Intensity} = \frac{\text{Total Dose Received}}{\text{Total Dose Expected}} * 100$$

If there are dose increases, the Dosing Intensity may exceed 100%. The number of patients with >100%, 100% intensity, 80% - <100%, 50 - <80, and <50% intensity will be summarized for the treatment arm of pevonedistat plus azacitidine.

Azacitidine

For azacitidine dosing, the percentage of all doses that were administered IV or SC will be summarized by treatment arm. The extent of exposure will be summarized by treatment arm in a similar manner as pevonedistat.

Daily Expected Dose, Daily Prepared Dose for Aza IV, Daily Dose Received for Aza IV, Daily Dose Received for Aza SC, Total Dose Received, Total Dose Expected, and Dosing Intensity for azacitidine will be based on the following formulas:

$$\text{Daily Expected Dose} = 75 \text{ mg/m}^2 * \text{BSA}$$

$$\text{Daily Prepared Dose (Aza IV)} = \text{Scheduled Dose Level } (\text{mg/m}^2) * \text{Body Surface Area } (\text{m}^2)$$

$$\text{Daily Dose Received (Aza IV)} = \text{Daily Prepared Dose} * \left(\frac{\text{Volume of IV bag actually infused (mL)}}{\text{Prepared Volume}}\right)$$

$$\text{Daily Dose Received (Aza SC)} = \text{Daily Dose Received}$$

$$\text{Total Dose Received} = \text{Sum of Actual Dose across all days of dosing}$$

$$\text{Total Dose Expected} = \text{Sum of "Daily Expected Dose * 7 doses per cycle" across all treated cycles}$$

$$\text{Percent Dosing Intensity} = \frac{\text{Total Dose Received}}{\text{Total Dose Expected}} * 100$$

Dosing intensity for azacitidine will be summarized by treatment arm in a similar manner to pevonedistat dosing intensity.

Dosing data will also be presented in by-patient listings.

7.7.2 Treatment Compliance and Modifications

The actions on study drugs will be summarized by treatment arm in the safety population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately. Data will be summarized for Cycle 1 only as well as all cycles. A patient will count only once for each type of action.

7.7.3 Duration of Follow-up

The duration of follow-up is defined as time from randomization to the date of death or last known visit. If a subject dies, the duration will equal the date of death minus the date of study start + 1, with a censor variable = 1 (censored for follow-up). If a subject is alive, the duration will equal the date when the subject was last known to be alive minus the date of study start + 1, with a censor variable = 0 (event for follow-up).

7.8 Efficacy Analysis

All available efficacy data will be included in data listings and tabulations. Data that are potentially spurious or erroneous will be examined according to standard data management operating procedures. In general, missing data will be treated as missing, and no data imputation will be applied, unless otherwise specified.

All efficacy evaluations will be conducted using the ITT population, the HR MDS population, and other disease subpopulations (patients with HR MDS/CMML, patients with HR CMML and patients with low-blast AML), separately, unless otherwise specified. Analyses based on other subsets of patients will be specified when needed.

In the following analyses, the stratified analyses will be conducted for the ITT population, the HR MDS population and other disease subpopulations (patients with HR MDS/CMML, and patients with HR CMML), except for the patients with low-blast AML. Unless specified otherwise, the stratification factor includes the randomization strata of low-blast AML, IPSS-R risk groups of very high, high, or intermediate (for HR MDS/CMML) for the ITT population, and IPSS-R risk groups of very high, high, or intermediate for the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS and patients with HR CMML). For the subpopulation of patients with low-blast AML, the unstratified analyses will be conducted since there is no stratification within this subpopulation at randomization.

7.8.1 Primary Efficacy Endpoint

There is one primary endpoint, EFS, with an event defined as death or transformation to AML for HR MDS/CMML patients or defined as death for low-blast AML patients. EFS is defined as the time from the date of randomization to the date of the occurrence of an event.

Transformation to AML for HR MDS/CMML patients is defined according to WHO Classification [6].

For HR MDS/CMML patients, patients without documentation of transformation to AML or death will be censored at the date of the last response assessment. Stem cell transplantation (SCT) is not considered as an alternate antineoplastic therapy.

For low-blast AML patients, patients without documentation of death will be censored at the date that the patient was last known to be alive.

The details regarding the handling of missing assessments and censoring for HR MDS/CMML patients for the EFS analysis are based on the FDA rules and presented in Table 7.a.

Table 7.a Handling of Missing Assessments and Censoring for HR MDS/CMML Patients for EFS Primary Analysis Based on the FDA Rules

Situation	Date of Event or Censoring	Outcome
No baseline and/or no post baseline assessment, no subsequent anticancer therapy after study treatment, no death	Date of Randomization	Censored
Transformation to AML documented between scheduled visits	Date of documented transformation to AML	Event
No documented EFS event	Date of last adequate assessment[a]	Censored
Lost to follow-up, withdraw consent before any documented EFS event	Date of last adequate assessment[a]	Censored
Alternate antineoplastic therapy started prior to transformation to AML or death (excluding stem cell transplantation)	Date of last adequate assessment prior to starting alternate antineoplastic therapy	Censored
Death before first assessment	Date of death	Event
Death between adequate assessment visits	Date of death	Event

a Adequate disease assessment is defined as there is sufficient data to evaluate a patient's disease status.

The primary endpoint, EFS, will be analyzed based on the HR MDS population for US submission and the ITT population for ex-US submission at IA2.The analyses of EFS will also be performed for other disease subpopulations (patients with HR MDS/CMML and patients with HR CMML) at IA2. In addition, at the first interim analysis (IA1), EFS will be analyzed based on the ITT population, the HR MDS population, and the HR MDS/CMML population to assess the futility and re-estimated the EFS event size. The analysis of the primary endpoint EFS will use IRC assessment.

IA1

The first IA will be performed when approximately 74 EFS events in patients with HR MDS (around 50% information relative to the minimal planned event size of 147 EFS events for the EFS final analysis in patients with HR MDS at IA2) have occurred, targeting to evaluate EFS for futility and re-assess EFS event size for IA2. If EFS hazard ratio >1.0 in all 3 of the following populations: patients with HR MDS, patients with HR MDS/CMML and the ITT population, the

study will stop; otherwise, EFS event-size re-estimation will be performed in patients with HR MDS using the conditional power for the EFS FA that will conducted for HR MDS patients at IA2 and the study will continue to IA2. The conditional power for EFS in the patients with HR MDS will be calculated by ISC and the detailed calculations are in Section 7.12. The IDMC will compare the conditional power with the pre-specified EFS event size adaptation rule in IDMC charter Appendix 5 and inform the SEC of the target EFS event size for IA2. The prespecified EFS event size re-estimation adaptation rule is a step function which is calculated based on Liu and Hu [4]. EFS will be analyzed separately for ITT population, the patients with HR MDS, and the patients with HR MDS/CMML. The unadjusted stratified Cox model will be used to estimate the hazard ratio and its 2-sided 95% CIs for the treatment effect. The Kaplan Meier (K-M) survival curves and K-M medians (if estimable), along with their 2-sided 95% CIs, will also be provided for each treatment arm.

IA2

IA2 will be performed when approximately the adaptive total EFS event size (from a minimum of 147 to a maximum of 249) in HR MDS informed by the IDMC has occurred. EFS will be analyzed based on the ITT population, the HR MDS population, the HR MDS/CMML population, and the HR CMML population. Initially, separate testing will be performed for EFS in patients with HR MDS for US submission and in the ITT population for ex-US submission at the full alpha of 2.5% at IA2. The analyses of EFS in the patients with HR CMML and in the patients with HR MDS/CMML are not included in the hierarchical testing procedure and will be used for exploratory purposes only.

The null and alternative hypotheses for EFS are:

H_0: EFS in combo Arm = EFS in aza Arm

H_a: EFS in combo Arm > EFS in aza Arm

The 1-sided Cui-Hung-Wang (CHW) [7] weighted log-rank test will be used to compare EFS between the 2 treatment arms. For patients with HR MDS, the weights are prespecified based on the observed number of EFS events in patients with HR MDS at IA1 and the minimum 147 EFS events in patients with HR MDS planned for IA2. For other populations (ITT, HR MDS/CMML, HR CMML), the weights are prespecified based on the observed number of EFS events in the specific population of interest at IA1 and the estimated number of EFS events in the specific population of interest at IA2 (matching to the minimum number of 147 EFS events in patients with HR MDS planned for IA2) (see Appendix in Section 9.0 for the estimated number of EFS events).

If the prespecified number of approximately 202 OS events for final analysis (FA) are expected (based on blinded study data) to be available close to the IA2, the IA2 (ie, EFS FA) and the FA (ie, OS FA) will be performed as a single analysis, when approximately 202 OS events and the adaptive EFS event size have occurred in HR MDS patients. The weights of the CHW test statistics for testing EFS at this single analysis will be the same as used at IA2.

When constructing the CHW weighted log-rank test statistics for EFS at IA2, the final cleaned IA1 data will be used.

The unadjusted stratified Cox model will be used to estimate the hazard ratio and its 2-sided 95% CIs for the treatment effect. The Kaplan Meier (K-M) survival curves and K-M medians (if estimable), along with their 2-sided 95% CIs, will also be provided for each treatment arm.

In addition, for the ITT population, the patients with HR MDS/CMML, the patients with HR MDS and the patients with HR CMML, a stratified Cox regression model will be used to further evaluate the treatment effects on EFS after adjusting for some prognostic factors. Besides treatment, the following prognostic factors will be included in the model simultaneously: age (<65, 65-74, >=75), de novo versus secondary, region (North America vs ex North America), baseline ECOG score (0-1 vs 2), baseline peripheral WBC (<15,000 per µL vs >=15,000 per µL), baseline platelet (<100,000 vs >=100,000). Additional exploratory analyses may be performed if deemed necessary.

For patients with EFS events, the reasons leading to the determination of EFS will be tabulated. For patients without EFS events, the main reason for censoring will also be tabulated.

Sensitivity Analysis

EFS assessed by IRC using different censoring mechanisms for patients with HR MDS or CMML will be analyzed in the patients with HR MDS for US-submission, and the ITT population for ex-US submission, and other disease subpopulations (patients with HR MDS/CMML, and patients with HR CMML); for example, not censoring for HR MDS/CMML patients who discontinue treatment and go on alternative antineoplastic therapy based on the EMA rules. The details of the handling of missing assessments and censoring for HR MDS/CMML patients for sensitivity analyses are presented in Table 7.b.

Table 7.b Handling of Missing Assessments and Censoring for HR MDS/CMML Patients for EFS Sensitivity Analysis Based on the EMA Rules

Situation	Date of Event or Censoring	Outcome
Alternate antineoplastic therapy started prior to transformation to AML or death	Date of documented transformation to AML or death	Event

Additional sensitivity analyses for EFS for the ITT population, the patients with HR MDS, the patients with HR MDS/CMML, and the patients with HR CMML will include:

- EFS assessed by investigator.
- EFS assessed by IRC in the PP population
- SCT is considered as alternate antineoplastic therapy for patients with HR MDS/CMML, and patients with HR MDS/CMML who have SCT prior to transformation to AML or death will be censored at the date of last adequate assessment prior to starting SCT.
- EFS event after missing more than one visit is censored. Two censoring windows of 60 and 90 days will be considered. An EFS event will be censored at the last adequate disease

assessment before the EFS event, if it occurs beyond 60 (90) days after the last disease assessment visit.

In case of any updates on the IA1 data after IA1 is performed, instead of using the final cleaned IA1 data, the original IA1 data will be used to construct the CHW weighted log-rank test statistics for EFS analysis at IA2.

All the sensitivity analyses will be performed by using the same methods as used for the primary analysis.

Subgroup Analysis

Subgroup analyses will be performed for EFS relative to baseline stratification factors, indication (HR MDS, HR CMML, low-blast AML or HR MDS/CMML), demographic data such as sex, race, and age, disease characteristics such as de novo or secondary, for the ITT population, the HR MDS population, the patients with HR MDS/CMML, and the patients with HR CMML. The statistical model will be adjusted accordingly to fit the subgroup analyses. The unstratified log rank test will be used for comparing the treatment arms in the exploratory manner. The details on subgroups are presented in Table 7.c:

Table 7.c List of Subgroups

Subgroup	Definition of Group
Age	< 65 years, ≥65 and ≤ 74 years, ≥ 75 years
Sex	male vs female
Race	White vs non-White
Region	North America (NA) vs ex NA
IPSS-R risk category & low-blast AML	very high, high, intermediate, or low-blast AML
Indication	HR MDS/CMML, HR MDS, HR CMML, or low-blast AML
Low-blast AML ELN2017 risk	adverse, intermediate and favorable
Baseline ECOG performance status	(0 or 1) vs 2
Disease type	de novo, secondary
Baseline peripheral WBC	<15,000 per µL vs >= 15,000 per µL
Baseline platelet	<100,000 vs >=100,000

7.8.2 Key Secondary Efficacy Endpoint

OS is the key secondary efficacy endpoint. The analysis of OS for US submission will be based on patients with HR MDS; the analysis for ex-US submission will be based on the ITT population in the overall population (which includes patients with HR MDS or CMML, or low-blast AML). The analysis will also be performed for patients with low-blast AML, the patients with HR MDS/CMML and the patients with HR CMML. OS is defined as the time from the date of randomization to the date of death due to any cause. Patients without documented death at the time of the analysis will be censored at the date that the patient was last known to be alive.

OS will be tested in different disease populations at IA2 and at finial analysis (FA) following the multiple hierarchical testing procedures as described in Section 7.8.3. The FA will be performed when approximately 202 OS events have occurred in patients with HR MDS. The 1-sided alpha of 0.0001 will be used at IA2, and the remaining 1-sided alpha of 0.0249 at will be used at FA for testing OS in the ITT population, patients with HR MDS, patients with low-blast AML and patients with HR MDS/CMML. Separate multiple hierarchical testing procedures will be performed for US submission and ex-US submission, as described in Section 7.8.3. OS will also be tested in the patients with HR CMML, but the test is for the exploratory purpose only and not included in the hierarchical testing procedures.

The null and alternative hypotheses for OS are:

H_0: OS in combo Arm = OS in aza Arm

H_a: OS in combo Arm > OS in aza Arm

The 1-sided CHW weighted log-rank test will be used to compare OS between the 2 treatment arms and calculate the p-values. The weights are pre-specified based on the observed number of OS events in the specific population of interest at IA1, the estimated numbers of OS events in the specific population of interest at IA2 (matching to the minimum number of 147 EFS events in patients with HR MDS planned for IA2) and at FA (matching to the targeted 202 OS events in patients with HR MDS planned at FA) (see Appendix in Section 9.0 for the estimated number of OS events).

If the prespecified number of approximately 202 OS events for FA are expected (based on blinded study data) to be available close to the IA2, the IA2 (ie, EFS FA) and the FA (ie, OS FA) will be performed as a single analysis, when approximately 202 OS events and the adaptive EFS event size have occurred in HR MDS patients. The weights of the CHW test statistics for testing OS are pre-specified based on the observed number of OS events at IA1, and the targeted 202 OS events in patients with HR MDS and the estimated number of OS events in other specific population of interest matching to the targeted 202 OS events in patients with HR MDS at FA (see Appendix in Section 9.0 for the estimated number of OS events). The1-sided alpha of 0.025 will be used for testing OS at this single analysis following the hierarchical testing procedures described in Section 7.8.3.

When constructing the CHW weighted log-rank test statistics for OS at IA2/FA, the final cleaned IA1 data will be used.

Hazard ratios, along with the 2-sided 95% CIs will be estimated using the unadjusted stratified Cox model for the ITT population, patients with HR MDS/CMML, patients with HR MDS and patients with HR CMML and using the unadjusted Cox model for the patients with low-blast AML. Kaplan-Meier curves, Kaplan-Meier medians (if estimable), and survival probability at 30 days, 60 days, 6 months and 1 year, together with the 95% CIs, will be calculated for each treatment group.

In addition, for the ITT population, the patients with HR MDS/CMML, the patients with HR MDS and the patients with HR CMML, a stratified Cox regression model will be used to further evaluate the treatment effects on OS after adjusting for some prognostic factors. Besides

treatment, the following prognostic factors will be included in the model simultaneously: age (<65, 65-74, >=75), de novo versus secondary, region (North America vs ex North America), baseline ECOG score (0 or1 vs 2), baseline peripheral WBC (<15,000 per µL vs >= 15,000 per µL), baseline platelet (<100,000 vs >=100,000). For the patients with low-blast AML, the similar analysis will be conducted using a unstratified Cox regression model, which includes treatment, the above prognostic factors and ELN 2017 cytogenetic risk categories for low-blast AML (adverse, intermediate, favorable) in the model. Additional exploratory analyses may be performed if deemed necessary.

Subgroup analyses will be performed using subgroups defined for EFS analyses. The unstratified log-rank test will be used to compare the treatment arms within the subgroups.

To adjust for the potential confounding effects of post-treatment transplantation after patients discontinued study treatment, the sensitivity analysis will be performed for OS with censoring at the start of post-treatment transplantation for the ITT population, the patients with HR MDS/CMML, the patients with HR MDS, the patients with HR CMML, and the patients with Low-blast AML.

To adjust for the impact by COVID-19 pandemic, another sensitivity analysis for OS will be performed as appropriate for ITT and HR MDS populations and all other disease subpopulations by excluding the deaths caused by COVID-19 from the OS events and censoring patients who die of COVID-19 at the dates of their death. The same statistical methods as used for the primary analysis of OS will be used for both sensitivity analyses.

To adjust for the potential confounding effects of subsequent therapies after patients discontinue study treatment, the following 2 methods will be used whenever appropriate:

- Marginal Structural Models (MSMs) by Robins et al., 2000 [9]
- Inverse Probability of Censoring Weighted (IPCW) method by Robins and Finkelstein, 2000 [10]

In the MSM and IPCW analyses, in order to derive weights adjusting for the time-fixed and time-varying confounding effects due to taking alternative therapies, the covariates that affect disease progression and post-progression treatment, and the OS endpoint will be used. Baseline covariates include region (North America (NA) vs ex NA), age (< 65 years, >=65 and <= 74 years, >= 75 years), race (white, non-white), ECOG score (0 or 1 vs 2), disease type (de novo or secondary, baseline peripheral WBC, baseline platelets, disease risk group (for AML, ELN 2017, adverse, intermediate, favorable, for HR MDS/CMML, IPSS-R, very high, high, intermediate). Time-varying covariates include duration of treatment exposure, disease progression status at each study visit, RBC Transfusion at each study visit, platelet transfusion at each study visit, indication of subsequent therapy (yes or no), additional therapy with transplantation. The final criteria for selected covariates would need to be statistically have a p-value of less than or equal to 0.1 in the multivariate Cox regression models for weight calculations, and in the final Cox model for OS. If there are more than 5% missing in the covariate, then this covariate will be dropped from the weighting calculation and final OS model. For both MSM and IPCW analyses, logistic regression models on repeated measurements will be used to approximate the Cox

models in the weight derivations from which stabilized weights will be derived per subject per observation. Adjusted K-M curves will also be presented along with hazard ratios (HRs), 95% confidence intervals for HRs, and adjusted p-values based on MSM and IPCW approaches. SAS proc PHREG procedure with counting process type of data input, which takes multiple observations per subject, will be used as the final Cox model for OS for both MSM and IPCW approaches.

Additional analyses of OS may be performed if deemed necessary.

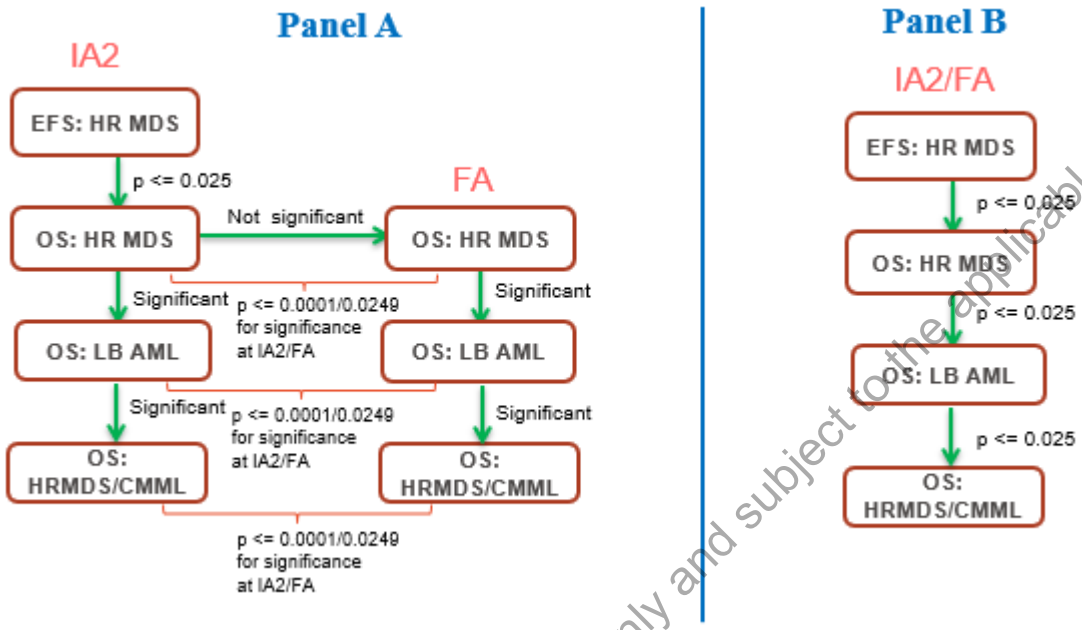
7.8.3 Multiple Hierarchical Testing Procedure for Testing Primary and Key Secondary Efficacy Endpoints at IA2 and FA

IA2 will be performed when the approximate adaptive EFS event size (from a minimum of 147 to a maximum of 249), informed by the IDMC, has occurred in patients with HR MDS. The FA will be performed when approximately 202 OS events have occurred in patients with HR MDS.

Separate multiple hierarchical testing procedures for the US submission and the ex-US submission will be adopted to test the primary endpoint of EFS and secondary endpoint of OS in the HR MDS population (US submission), the ITT population (ex-US submission), and other disease populations at IA2 and FA, with each procedure having a total 1-sided alpha of 0.025, as depicted on Panel A of Figure Figure 7.a (US submission testing procedure) and on Panel A of Figure 7.b (ex-US submission testing procedure).

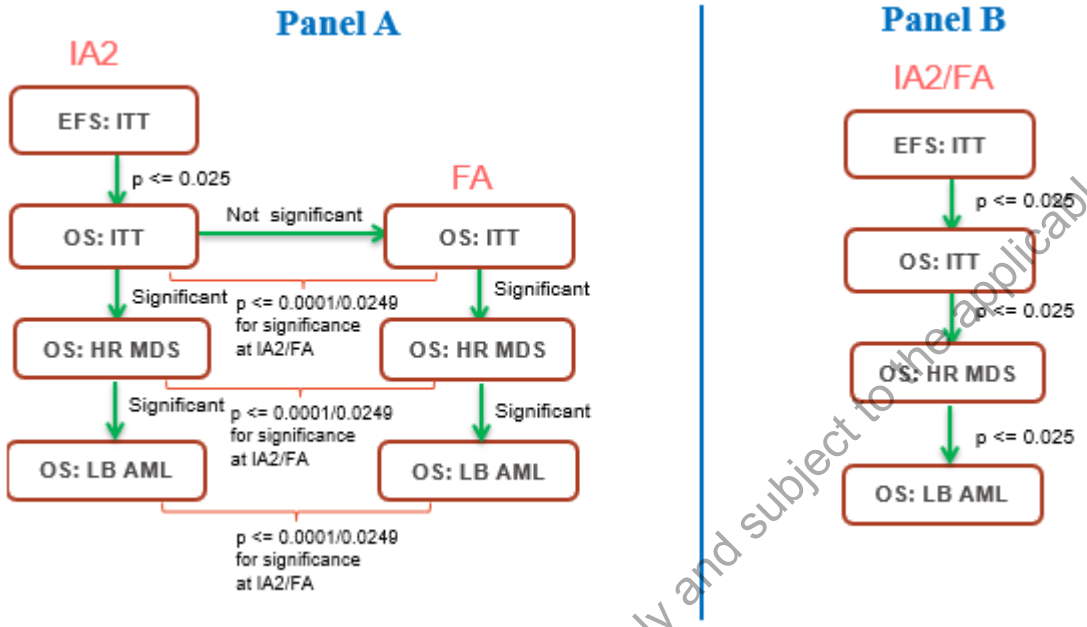
If the prespecified number of approximately 202 OS events for FA are expected (based on blinded study data) to be available close to the IA2, the IA2 (ie, EFS FA) and the FA (ie, OS FA) will be performed as a single analysis, when approximately 202 OS events and the adaptive EFS event size have occurred in HR MDS patients. As originally planned, separate multiple hierarchical testing procedures for the US submission and the ex-US submission will be used to test the primary endpoint of EFS and the key secondary endpoint of OS in the HR MDS population (US submission), the ITT population (ex-US submission), and other disease populations at this single analysis, with a total 1-sided alpha of 0.025 for each procedure. The 2 hierarchical testing procedures originally specified at IA2 and FA will be performed as a single testing procedure with a 1-sided alpha of 0.025 used for each test, as depicted on Panel B of Figure 7.a (US submission testing procedure) and on Panel B of Figure 7.b (ex-US submission testing procedure).

Figure 7.a Multiple Hirarchical Testing Procedure at IA2 and FA for the US Submission



CMML=chronic myelomonocytic leukemia, EFS=event-free survival, FA=final analysis, HR MDS=higher-risk myelodysplastic syndromes, IA2=second interim analysis, IA2/FA=IA2 and FA are performed as a single analysis, LB AML=low-blast acute myelogenous leukemia, OS=overall survival, US=United States.

Figure 7.b Multiple Hierarchical Testing Procedure at IA2 and FA for the ex US Submission



EFS=event-free survival, ex-US=all countries and regions excluding the United States, FA=final analysis, HR MDS=higher-risk myelodysplastic syndromes, IA2=second interim analysis, IA2/FA=IA2 and FA are performed as a single analysis, LB AML=low-blast acute myelogenous leukemia, OS=overall survival.

At IA2:

- For the US submission, EFS in patients with HR MDS will be tested first, at a 1-sided alpha of 0.025. The CHW weighted log-rank test statistic will be used. If the test of EFS in patients with HR MDS is statistically significant, OS in patients with HR MDS will be subsequently tested using the CHW weighted test statistic. If the test of OS in patients with HR MDS is statistically significant, then subsequent statistical testing using the CHW weighted test statistic will be performed in the following order: OS in patients with low-blast AML, and OS in patients with HR MDS/CMML.
- For the ex-US submission, EFS in the ITT population will be tested first, at a 1-sided alpha of 0.025. The CHW weighted test statistic will be used. If the test of EFS in the ITT population is statistically significant, OS in the ITT population will be subsequently tested using the CHW weighted test statistic. If the test of OS in patients in the ITT population is statistically significant, then subsequent statistical testing using the CHW weighted test statistic will be performed, first for OS in patients with HR MDS and then for OS in patients with low-blast AML.

If EFS is not statistically significant in both patients with HR MDS and the ITT population at IA2, the study will stop for futility. If OS is statistically significant in both patients with HR

MDS and the ITT population at IA2, the study will stop for efficacy. Otherwise, the study will continue to the FA.

At the FA:

- For the US submission, if the test of OS in the HR MDS population fails to achieve statistical significance at IA2, OS in patients with HR MDS will be tested first using the CHW weighted test statistic. If statistical significance is achieved, subsequent statistical testing will be performed in the following order: OS in patients with low-blast AML, and OS in patients with HR MDS/CMML. The CHW weighted test statistic will be used for all both tests.
- For the ex-US submission, if the test of OS in the ITT population fails to achieve statistical significance at IA2, OS in the ITT population will be tested first using the CHW weighted test statistic. If statistical significance is achieved, subsequent statistical testing will be performed, first for OS in patients with HR MDS and then for OS in patients with low-blast AML. The CHW weighted test statistic will be used for both tests.

EFS in the HR MDS population (US submission) and in the ITT population (ex-US submission) will be tested only at IA2 using a 1-sided alpha of 0.025 for each test. All other tests including OS in the HR MDS population, OS in the ITT population, OS in the low-blast AML population, and OS in the HR MDS/CMML population, will use a 1-sided alpha of 0.0001 at IA2 and the remaining 1-sided alpha of 0.0249 at FA.

The testing procedures for the US and ex-US submissions will be performed independently, with a total one-sided alpha of 0.025 for each testing procedure. Under the sequential procedure, the testing will be stopped once 1 test fails to achieve statistical significance. No benefit will be claimed for untested hypotheses.

When the EFS event size for the HR MDS population for IA2 is re-estimated at IA1, the numbers of OS and EFS events in other populations of interest for IA2 will be changed accordingly. The use of CHW weighted test statistics [7] can preserve a type-I error rate when the event sizes are changed. In addition, multiple testing procedures (multiple hypotheses and multiple looks) that apply the sequential procedure and the pre-specified alpha spending approach at multiple looks also maintain control of the familywise type I error rate, as discussed in Glimm et al [11].

7.8.4 Secondary Efficacy Endpoint(s)

Other secondary efficacy parameters are: 6-month and 1-year survival rates; time to AML transformation; rate of CR (includes CR in patients with HR MDS or CMML, or low-blast AML), CR+CRi in patients with low-blast AML, CR+ marrow CR in patients with HR MDS or CMML, CR+PR+HI in patients with HR MDS or CMML, CR+ marrow CR+PR in patients with HR MDS or CMML, CR+ marrow CR+PR+HI in patients with HR MDS or CMML, overall response, overall response by Cycle 6 and overall response 2; duration of CR (for patients with HR MDS or CML, or low-blast AML), CR+CRi (in patients with low-blast AML), overall response, and overall response 2; RBC and/or platelet transfusion independence; duration of

RBC and/or platelet transfusion independence; time to first CR or PR or CRi (for patients with low-blast AML); rates of HI in patients with HR MDS or CMML; percent of patients who have at least 1 inpatient hospital admission(s) related to HR MDS or CMML, or low-blast AML; time to PD, relapse after CR or PR, or death; and ORR, EFS and OS in patients that have TP53 mutations, 17p deletions and/or are determined to be in an adverse cytogenetic risk group in both treatment arms.

Disease response-related endpoints will be analyzed using both IRC and investigator's assessments, unless otherwise specified.

The analyses of secondary efficacy endpoints will be performed at IA2 or at the single analysis of IA2/FA if the IA2 and the FA are performed as a single analysis.

6-month and 1-year Survival Rates

K-M estimates and the 95% CIs of 6-month and 1-year survival rates will be provided by treatment arm and overall based on the ITT population and disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately. Survival rate differences will be calculated, along with its 95% CI.

Time to AML Transformation

Time to AML transformation is defined as time from randomization to documented AML transformation. This definition only applies to HR MDS and CMML patients, so this analysis will only be carried out in the patients with HR MDS, the patients with HR MDS/CMML, and the patients with HR CMML. Patients who died before progression to AML will be censored at the date of death. Patients without documented AML transformation at the time of the analysis will be censored at the date of the last assessment.

A stratified 1-sided log-rank test will be used to compare time to AML transformation between the 2 treatment arms. Hazard ratios, along with the 2-sided 95% CIs will be estimated using the unadjusted stratified Cox model. Kaplan-Meier curves, Kaplan-Meier medians (if estimable), and survival probability at 30 days, 60 days, 6 months and 1 year, together with the 95% CIs, will be calculated for each treatment group.

Rates of CR (in patients with HR MDS or CMML, or low-blast AML), CR+CRi (in patients with low-blast AML), CR+ marrow CR (in patients with HR MDS or CMML), CR+PR+HI (in patients with HR MDS or CMML), CR+ marrow CR+PR (in patients with HR MDS or CMML), CR+ marrow CR+PR+HI (in patients with HR MDS or CMML), overall response (CR+PR in patients with HR MDS or CMML; CR+CRi+PR in patients with low-blast AML), overall response by Cycle 6 (CR+PR in patients with HR MDS or CMML; CR+CRi+PR in patients with low-blast AML by Cycle 6) and overall response 2 (CR+PR+HI in patients with HR MDS or CMML; CR+CRi+PR in patients with low-blast AML)

Rates of CR (in patients with HR MDS or CMML, or low-blast AML), CR+CRi in patients with low-blast AML, CR+ marrow CR in patients with HR MDS or CMML, CR +PR+HI in patients with HR MDS or CMML, CR+ marrow CR+PR in patients with HR MDS or CMML, CR+

marrow CR+PR+HI in patients with HR MDS or CMML, overall response (CR+PR in patients with HR MDS or CMML; CR+CRi+PR in patients with low-blast AML), overall response by Cycle 6 (CR+PR in patients with HR MDS or CMML; CR+CRi+PR in patients with low-blast AML by Cycle 6) and overall response 2 (CR+PR+HI in patients with HR MDS or CMML; CR+CRi+PR in patients with low-blast AML**)** are disease response-related endpoints, which will be analyzed using both IRC and investigator's assessments for the Response-Evaluable population. They are respectively defined as the proportion of patients who achieve the corresponding response in the corresponding group of patients.

Rates of CR, overall response, overall response by Cycle 6 and overall response 2 will be summarized in the overall response-evaluable population and in the response-evaluable subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

The number and percentage of patients for each definition of response will be summarized by treatment group. Stratified Cochran-Mantel-Haenszel (CMH) chi-square test will be used to compare the 2 treatment arms for the ITT population, the patients with HR MDS/CMML, the patients with HR MDS and the patients with HR CMML. And for the patients with low-blast AML, the unstratified Cochran-Mantel-Haenszel (CMH) chi-square test will be used. The CMH chi-square test p-value, the relative risk with its 2-sided 95% CIs will be calculated. The absolute rate difference will be provided with its 95% CIs using the asymptotic method.

Patients who are not response-evaluable will be treated as non-responders.

Duration of CR, CR+CRi (in patients with low-blast AML), Overall Response, and Overall Response 2

Duration of CR/CR+CRi/overall response/overall response 2, is defined as the time from the date of first documentation of a CR/CRi or CR/PR or better/ overall response 2, to the date of first documentation of PD or relapse from CR (in patients with low-blast AML) or relapse after CR or PR (in patients with HR MDS/CMML) for responders of CR/CRi or CR/PR or better/ overall response 2, respectively. Responders without documentation of PD or relapse from CR (in patients with low-blast AML) or relapse from CR or PR (in patients with HR MDS/CMML) will be censored at the date of their last response assessment that is SD or better. Duration of CR/CR+CRi/overall response/overall response 2 will be summarized descriptively using the K-M method. Kaplan Meier (K-M) survival curves and K-M medians (if estimable) will be provided for each treatment arm. Duration of CR+CRi is summarized only in patients with low-blast AML. Duration of CR/overall response will be summarized in the overall response-evaluable population and in the response-evaluable subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately. Duration of overall response 2 will be summarized in the overall response-evaluable population, in the response-evaluable subpopulation (patients with HR MDS/CMML, patients with HR MDS and patients with HR CMML), separately.

Rate of RBCs and/or Platelet Transfusion Independence

Analysis of RBC and/or platelet transfusion independence will be based on the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

A patient is defined as RBC and/or platelet-transfusion independent if he/she receives no RBC and/or platelet transfusions for a period of at least 8 weeks [9,10] during the time period from the first dose of study drug administration through 30 days after the last dose of any study drug. Rate of RBC and/or platelet-transfusion independence is defined as number of patients who become RBC and/or platelet transfusion independent divided by the number of patients who are RBC and/or platelet transfusion dependent at Baseline (i.e., patients received RBC and/or platelet-transfusion within 8 weeks before the first dose of study drug administration).

The number of patients who are RBC and/or platelet transfusion dependent/independent at Baseline and post Baseline, as well as rate of RBC and/or platelet transfusion independence, will be summarized by treatment group. Rate of RBC and/or platelet transfusion independence will be compared using stratified CMH test for the ITT population, the patients with HR MDS/CMML, the patients with HR MDS and the patients with HR CMML and using unstratified CMH test for the patients with low-blast AML. P values and 2-sided 95% CIs of the relative risk will be provided. The absolute rate difference will be provided with its 95% CIs using the asymptotic method.

Duration of RBC and/or Platelet Transfusion Independence

Analysis of duration of RBC and/or platelet transfusion independence will be based on patients who are RBC and/or platelet transfusion independent during the trial. This analysis will be performed for ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

Duration of transfusion independence is defined as the longest time between the last RBC and/or platelet transfusion before the start of the RBC and/or Platelet transfusion-independent period and the first RBC and/or Platelet transfusion after the start of the transfusion-independent period, which occurs ≥ 8 weeks later. The patients who remain transfusion independent from the initial transfusion or the first dose of study drug administration till the time of the analysis will be censored. If the time of analysis is before 30 days after the last dose study, the patients will be censored at the time of analysis or death date, whichever occurs first. If the time of analysis exceeds 30 days after the last dose study medication, the patients will be censored at 30 days after the last dose of study medication or death date, whichever occurs first. Duration of RBC and/or platelet transfusion independence will be summarized descriptively using the K-M method. Kaplan Meier (K-M) survival curves and K-M medians (if estimable) will be provided for each treatment arm.

Time to First CR or PR or CRi (for patients with low-blast AML)

Analysis of time to first CR or PR or CRi (for patients with low-blast AML) will be based on the response-evaluable population.

Time to first CR or PR or CRi is defined as time from randomization to first documented CR or PR or CRi, whichever occurs first. Time to first CR or PR or CRi will be analyzed similarly as time to AML transformation.

For the responders of CR, PR or CRi, the number of responders with CR, PR or CRi, mean, SD, median, minimum and maximum of time to first CR, PR or CRi in months will be presented by treatment in a separate table.

Rates of HI

Rates of HI will be analyzed using IRC assessments, investigator's assessments and also programmed HI for the response-evaluable population of patients with HR MDS, patients with HR CMML and patients with HR MDS/CMML. Stratified CMH test will be used to compare the 2 treatment arms. The difference in rates and the associated 95% CIs will be presented using the asymptotic method.

Percent of Patients Who Have at Least One Inpatient Hospital Admission Related to HR MDS, HR CMML, or low-blast AML

Analysis of percent of patients who have at least one inpatient hospital admission related to HR MDS, CMML, or low-blast AML will be based on the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately. Inpatient hospital admission data will be collected through transformation to AML (HR MDS/CMML patients) or until initiation of subsequent therapy (all patients), whichever occurs first.

The number and percentage of patients who have at least one inpatient hospital admission(s) related to HR MDS, CMML, or low-blast AML will be summarized by treatment group. The rate difference and the associated 95% CIs will be provided using the asymptotic method.

Time To PD, Relapse after CR (low blast AML) or Relapse after CR or PR (HR MDS/CMML), or Death

Analysis of time to PD, relapse after CR (low blast AML), relapse after CR or PR (HR MDS/CMML), or death will be based on the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

Time to PD, relapse after CR (low blast AML), relapse after CR or PR (HR MDS/CMML), or death is defined as the time from the date of randomization until the date of the first documentation of disease progression, relapse after CR (low blast AML), relapse after CR or PR (HR MDS/CMML), or death due to any cause, whichever occurs first. Time to PD, relapse after CR or PR, or death will be analyzed similarly as time to AML transformation.

ORR, EFS and OS in Patients who Have TP53 Mutations, or 17p Deletions and/or are Determined to be in Adverse Cytogenetic Risk Group

ORR, EFS and OS in patients who have TP53 mutations, or 17p deletions and/or are determined to be in an adverse cytogenetic risk group will be analyzed descriptively for ITT population and

the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately, using the similar method as ORR, EFS and OS.

Time to Subsequent Therapy

Time to subsequent therapy is defined as the time from the date of randomization to the date of the first documented subsequent therapy (excluding stem cell transplantation). Subsequent therapy is defined as agent(s) with antileukemic/anti-MDS activity (eg, cytarabine, anthracyclines, purine analogues, and hypomethylating agents other than azacitidine). Patients who discontinue study treatment to receive single-agent azacitidine off study would not be counted as receiving subsequent therapy. Patients who do not receive subsequent therapy at the time of the analysis will be censored at the date of death or last contact. Time to subsequent therapy is a time-to-event variable, which will be analyzed based on the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately, using the similar method as time to transformation to AML.

7.8.5 Additional Efficacy Endpoint(s)

The analyses of additional efficacy endpoints below will be performed at IA2 or at the single analysis of IA2/FA if the IA2 and the FA are performed as a single analysis.

Allogeneic Stem Cell Transplantation

Analysis of rate of Allogeneic Stem Cell Transplantation (ASCT) will be based on the ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

The number and percentage of patients who have received ASCT will be summarized by treatment group. The rate difference and the associated 95% CIs will be provided using the asymptotic method.

Duration of HI

Analysis of duration of HI will be using IRC assessments, investigator's assessments and also programmed HI for the response-evaluable population of patients with HR MDS or CMML.

Duration of HI is defined as the longest time from an established hematological improvement to the subsequent first non-HI. Patients remain HI or have no information about the status of HI from an established hematological improvement until the time of the analysis will be censored at the date of the last disease assessment. Duration of HI will be summarized descriptively using the K-M method. Kaplan Meier (K-M) survival curves and K-M medians (if estimable) will be provided for each treatment arm.

Rate of RBC and/or Platelet Transfusion

Analysis of RBC and/or platelet transfusion will be performed for ITT population and the disease subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

The rate of RBCs and/or platelet transfusion is defined as the total number of RBCs and/or platelet transfusion from the first dose through 30 days after the last dose of any study drug divided by the total number of patient-years in each group. Patient-year is calculated from the first dose to 30 days after the last dose of any study drug. The rate and the associated 95% CIs will be provided by treatment group.

7.9 Pharmacokinetic/Pharmacodynamic Analysis

7.9.1 Pharmacokinetic Analysis

Pevonedistat plasma concentration-time data will be presented in listings. Pharmacokinetic data will be used to perform future population PK analyses of pevonedistat using a nonlinear mixed effects modeling approach and to assess the effect of various covariates on pevonedistat PK. These analyses may additionally include data collected in other pevonedistat clinical studies. The analysis plan for the population PK analysis will be separately defined, and the results of these analyses will be reported separately.

7.9.2 Pharmacodynamic Analysis

Not applicable.

7.10 Other Outcomes

Not applicable.

7.10.1 Patient-Reported Outcomes (PROs)

Patient-reported outcome assessments using the EORTC QLQ-C30 and 10 supplemental EORTC items will be analyzed using patients with patient-reported outcome measurements at baseline and at least one post baseline measurement in the ITT population.

Fifteen subscale scores can be calculated from the 30 EORTC QLQ-C30 items (Table 7.d) [12]. The summary score of the EORTC QLQ-C30 will be calculated from the mean of 13 of the 15 QLQ-C30 scales (the Global Health Status/QoL scale and the Financial Difficulties scale are not included). Emphasis will be placed on the fatigue, physical functioning, role functioning, and dyspnea subscales.

Table 7.d Description of EORTC QLQ-C30 Domain Scores

Scale scores	Subscale scores	Included items
Functional	Physical functioning (PF)	1-5
	Role functioning (RF)	6, 7
	Emotional functioning (EF)	21-24
	Cognitive functioning (CF)	20, 25
	Social functioning (SF)	26, 27
Global Health/ QoL	Global health status / QoL (QL)	29, 30
Symptom	Fatigue (FA)	10, 12, 18
	Nausea and vomiting (NV)	14, 15
	Pain (PA)	9, 19
	Dyspnea (DY)	8
	Insomnia (SL)	11
	Appetite loss (AP)	13
	Constipation (CO)	16
	Diarrhea (DI)	17
	Financial difficulties (FI)	28

Higher scores represent better health states for the functional scales, whereas lower scores represent better health states for the symptom scales/items [13].

Descriptive Statistics

The descriptive statistics of actual values and changes from baseline of the subscale scores for the EORTC QLQ-C30 will be summarized by treatment arm over time in a table and accompanying set of figures. The estimated mean in the changes from baseline in the 2 treatment arms and the 95% CIs will be provided at each time point. Also, the mean differences between treatment groups along with 95% CIs and p values will be presented at each time point.

The scores for the 10 supplemental EORTC items will be summarized by treatment arm over time in a table.

The descriptive analyses will be conducted in the overall population, by indication (HR MDS/CMML, HR MDS, HR CMML and low blast AML), and by clinical response (responders and nonresponders). Responders are HR MDS/CMML patients who achieve HI, PR, or CR, or low-blast AML patients who achieve PR, CRi or CR.

Change from baseline scores using cumulative distribution function (CDF) figures

The change in EORTC QLQ-C30 scale scores from baseline to various time points (e.g., cycles 3, 6, and 12) will be plotted using cumulative distribution functions (CDFs) for each treatment arm. The x-axis represents the changes in score (range: -100 to 100) and the y-axis represents the cumulative percentage of patients with a given change in score. The CDF figures will be

presented in the overall population, by indication (HR MDS/CMML, HR MDS, HR CMML, and low blast AML), and by clinical response (responders and nonresponders).

Analysis based on meaningful change

The number and percentage of patients with a change in EORTC QLQ C30 subscale scores from baseline achieving a clinically meaningful change from baseline will be summarized by treatment group over time in a table and figure. Specifically, patients with a change in score from baseline for a given threshold in a direction reflecting deteriorating functioning or increased symptoms at a given time point will be classified as "worsened", whereas those with a change in score from baseline in a direction reflecting improved QOL/functioning or decreased symptoms at a given time point will be classified as "improved". Those with no change in score from study entry or a change in score within the threshold for a clinically meaningful change will be classified as "stable".

The meaningful change analyses will be conducted in the overall population, by indication (HR MDS/CMML, HR MDS, HR CMML and low blast AML), and by clinical response (responders and nonresponders).

Change from baseline scores using cumulative distribution function (CDF) figures

The change in EORTC QLQ-C30 scale scores from baseline to various time points (e.g., last visit prior to EOT) will be plotted using cumulative distribution functions (CDFs) for each treatment arm. The x-axis represents the changes in score (range: -100 to 100) and the y-axis represents the cumulative percentage of patients with a given change in score. The CDF figures will be presented in the overall population, by indication (HR MDS/CMML, HR MDS, HR CMML, and low blast AML), and by clinical response (responders and nonresponders).

Time to HRQoL improvement

The time to HRQOL improvement will be compared across treatment arms. Improvement will be defined as an increase from study entry based on a clinically meaningful change for the functional scales and global health status/QoL. Improvement will be defined as a decrease from study entry based on a clinically meaningful change for the symptom scales.

Time to HRQoL improvement will be defined as the time from the date of randomization to the date of the first score that meets the definition for improvement. Patients without any HRQoL improvement will be censored at the date of last HRQoL measurement.

A 2-sided stratified log-rank test will be used to compare the treatment groups with respect to time to HRQoL improvement. In addition, an unadjusted, stratified Cox model will be used to estimate the hazard ratio and its 95% CI. The Kaplan-Meier survival curves will also be provided for each treatment group.

The time to HRQoL improvement analyses will be conducted in the overall population and by indication (HR MDS/CMML, HR MDS, HR CMML and low blast AML).

Analysis based on linear mixed effects models

The change from baseline in the EORTC QLQ-C30 subscale scores will be analyzed using repeated-measures linear mixed-effects (random-intercept only) models by incorporating the measurements across different time points, including treatment arm, time (a discrete variable), the interaction between treatment arm and time, baseline score, and stratification factors (i.e. low-blast AML, IPSS-R risk groups of very high, high, or intermediate for HR MDS/CMML) as covariates. The estimated mean in the changes from baseline in the 2 treatments and the 95% CIs will be provided at each time point and overall in a table as well as in a figure. In addition, the mean differences between treatment groups along with 95% CIs and p values will be presented at each time point in a table as well as in a figure. The analyses using repeated-measures linear mixed-effects models will be conducted for all subscales in the overall population and by indication (HR MDS/CMML, HR MDS, HR CMML, and low blast AML).

Missing Data

Details of scoring and initial handling of missing data are included in the EORTC QLQ-C30 scoring guidelines. Sensitivity analyses may be conducted to study the impact of missing data.

Compliance

Compliance proportion for EORTC QLQ-C30 based on the full ITT population will be summarized over time by treatment arm and overall. At each visit, the percentage of patients who completed at least one item of the EORTC QLQ-C30 will be calculated. Two denominators will be used for the calculation:

- The total number of patients initially randomized in the treatment arm
- The number of patients in the treatment arm who are still in the study at that visit

Compliance proportion for the 10 supplemental EORTC items will also be summarized over time by treatment arm and overall. At each visit, the percentage of patients who completed at least one of the 10 supplemental EORTC items will be calculated using the same two denominators as above.

Further exploratory analyses of PRO will be described in a specific standalone PRO analysis plan.

7.10.2 Health Economics (Health Care Resource Use)

EQ-5D-5L quality of life questionnaire scores (including 5-dimension descriptive system, VAS score and utilities) will be summarized in descriptive statistics for treatment arms over time. In addition, the change from baseline of VAS score and utilities (time tradeoff) will also be summarized in descriptive statistics for treatment arms over time. Specifically, utilities for UK, France, Germany, Italy, and Spain are of interest. The directly elicited value sets will be used for these country-specific utilities calculations, if available; otherwise, the crosswalk value sets will be used.

Compliance proportion for EQ-5D-5L will be summarized over time by treatment arm and overall. At each visit, the percentage of patients who completed at least one item of the EQ-5D-5L will be calculated. Two denominators will be used for the calculation:

1. The total number of patients initially randomized in the treatment arm
2. The number of patients in the treatment arm who are still in the study at that visit

Healthcare utilization data will be summarized in descriptive statistics of medical encounters (numbers and rates of encounters, reasons for encounters, and length of stay). Categories of interest include hospitalizations, emergency department visits, and outpatient visits.

Transfusion data will be summarized in descriptive statistics. Transfusions will be categorized as platelet (pooled platelet concentrate or plateletpheresis (single donor)) or red blood cell. For each type of transfusion, we will present the number of patients receiving a transfusion, the number of transfusions per patient, the location of the transfusion (in hospital vs. outpatient), and the number of units per transfusion by treatment arm.

Further modeling will be performed separately at post hoc analyses.

7.10.3 Biomarker

Biomarker analyses will be summarized using the ITT population.

- Baseline cytogenetic and mutation data will be summarized descriptively by treatment arm.
- EFS, OS, transformation to AML and ORR will be evaluated in patient subgroups defined by presence of: (a) mutations in genes that have been shown to have prognostic value in HR MDS/CMML and/or AML, (b) mutations in genes or groups of genes shown to be frequently mutated in MDS, AML and CMML (c) mutations that have shown correlation with response to treatment with hypomethylating agents, (d) mutations in selected pathways such as apoptosis, cell cycle, proliferation, DNA damage, immune response and development (e) high risk cytogenetic markers and (g) mutations in methylation pathway genes or changes in methylation patterns.
- Detailed statistical methods for identifying predictive markers for response to the pevonedistat plus azacitidine combination in HR MDS/CMML and/or AML will be specified in a separate biomarker analysis plan.
- Changes in allelic mutation burden in the bone marrow during treatment will be analyzed by treatment arm and correlated ORR, EFS, OS and time to transformation to AML.
- For the subgroup of patients who relapse following initial response to treatment, the mutational status taken at study entry and at relapse will be compared.
- Further exploratory analyses may be carried out based on emerging scientific knowledge.

7.11 Safety Analysis

Safety evaluations will be based on the incidence, severity, type of AEs, clinically significant changes or abnormalities in the patient's physical examination, vital signs, ECG, and clinical laboratory results.

These analyses will be performed using the safety population. All analyses will be performed by treatment arm for the safety population and the safety subpopulations (patients with HR MDS/CMML, patients with HR MDS, patients with HR CMML and patients with low-blast AML), separately.

All safety analyses will be performed at IA2. As appropriate, selected safety analyses will also be performed at FA. If IA2 and FA are performed as a single analysis, all safety analyses will be performed at this single analysis.

7.11.1 Adverse Events

7.11.1.1 Adverse Events

AEs will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA). Treatment-emergent is defined as any AE that occurs after administration of the first dose of study treatment and through 30 days after the last dose of any study drug.

AEs will be tabulated by system organ class (SOC), high level term (HLT), and preferred term (PT) by treatment arm. Summary tabulations include the following categories:

- Treatment-emergent AEs
- Drug-related treatment-emergent AEs
- Treatment-emergent Grade 3, 4 and 5 AEs (presented by grade and overall)
- Treatment-emergent drug-related Grade 3, 4 and 5 AEs (presented by grade and overall)
- Treatment-emergent AEs resulting in study drug discontinuation
- Treatment-emergent AEs that require dose modification of pevonedistat or azacitidine.
- Treatment-emergent SAEs
- Treatment-emergent drug-related SAEs
- Nonserious treatment-emergent AEs (≥5% in any arm)

Patients with the same AE more than once will have that event counted only once within each body system, once within each High Level Term, and once within each Preferred Term.

Treatment-emergent AEs will be tabulated by SOC, HLT, PT, and highest intensity. Most commonly reported (at least 10% of all patients) treatment-emergent AEs will be presented by preferred term. Most commonly reported (at least 10% of all patients) treatment-emergent AEs by preferred term will also be summarized by treatment cycles (Cycle 1, Cycle 2-3, Cycle 4-5,

Cycle 6-7, Cycle 8-9, Cycle 10-11, Cycle 12-13, Cycle 14+). All adverse events will also be reported in by-patient listings. By patient line listing of TEAEs will be generated.

Adverse events with start dates that are completely or partially missing will be analyzed as follows:

- If the start date has month and year but day is missing, the event will be considered treatment emergent if both the month and year of the start date of the event are on or after the month and year of the date of the first dose of study drug and on or before the month and year of the date of the last dose of study drug plus 30 days.
- If the start date has year, but day and month are missing, the event will be considered treatment emergent if the year of the start date of the event is on or after the year of the date of the first dose of study drug and on or before the year of the date of the last dose of study drug plus 30 days. If the start date of an event is completely missing, then the event is assumed to be treatment emergent.

7.11.1.2 Serious Adverse Events

The number and percentage of patients experiencing at least one treatment emergent serious AE (SAE) will be summarized by MedDRA SOC, HLT, and PT. Similar summary will be generated for treatment emergent drug-related SAEs.

A by-patient listing of the SAEs will be presented (the patient listing will contain all SAEs regardless of treatment emergent AE status).

An additional listing of treatment emergent C1D1 grade 2 or higher SAEs will also be generated.

7.11.1.3 Deaths

A by-subject listing of the deaths will be presented. All deaths occurring on-study and during follow-up will be displayed (regardless of treatment emergent AE status).

All deaths will be summarized by treatment arm, including deaths occurring on-study and death during follow-up separately.

On-study death is defined as the death that occurs between the first dose of study drug and 30 days after the last dose of study drug.

Thirty-/60-day mortality rate is defined as the proportion of patients who survive at most 30/60 days from the first dose of study drug, which will be summarized by treatment arm in a table.

7.11.1.4 Adverse Events Resulting in Discontinuation of Study Drug

The number and percentage of patients experiencing at least one adverse event resulting in discontinuation of study drug will be summarized by MedDRA SOC, HLT, and PT.

A by-patient listing of treatment emergent AEs resulting in discontinuation of study drug will be presented.

The separate summary table of the TEAEs coded to one of the following PTs or HLTs in Sections 7.11.1.5 - 7.11.1.9 will be generated by treatment arm. This table will be organized by TEAE category and then by HLT or PT specified under the TEAE category.

7.11.1.5 Dyspnoea

- PT Bendopnoea
- PT Dyspnoea
- PT Dyspnoea at rest
- PT Dyspnoea paroxysmal nocturnal
- PT Laryngeal dyspnoea
- PT Nocturnal dyspnoea

7.11.1.6 Anaemia of chronic disease and Anaemia of malignant disease

- PT Anaemia of chronic disease
- PT Anaemia of malignant disease

7.11.1.7 Infusion site reactions

- HLT Infusion site reactions
- PT Infusion site abscess
- PT Infusion site cellulitis
- PT Infusion site infection
- PT Infusion site joint infection

7.11.1.8 Injection site reactions

- HLT Injection site reactions
- PT Injection site abscess
- PT Injection site cellulitis
- PT Injection site infection
- PT Injection site joint infection

7.11.1.9 Gastrointestinal disorders

- PT Abdominal pain
- PT Constipation

- PT Diarrhoea
- PT Nausea
- PT Vomiting

The separate summary table of TEAEs will be generated for the following SMQs in Sections 7.11.1.10 -7.11.1.21 by arm. This table will be organized by SMQ and then by PT falling under each SMQ.

7.11.1.10 Pneumonia

- Infective pneumonia (SMQ broad and narrow)
- Eosinophilic pneumonia (SMQ broad and narrow)

7.11.1.11 Haemorrhages SMQ (broad and narrow)

Patients with events coded to a PT subsumed by Haemorrhages SMQ will be summarized respectively for occurrences of thrombocytopenia, platelet count decreased, and lab platelet toxicity grade of at least 2.

By-patient line listing for patients with any grade TEAE of Haemorrhages (defined by SMQ Haemorrhages) and concurrent within 5 days Thrombocytopenia/Platelet Count Decreased will be provided. If a patient has a treatment emergent haemorrhage and also either treatment emergent thrombocytopenia or post-baseline lab of platelet count decreased within a 5 day window of the haemorrhage, this patient is considered concurrent.

A summary table of patients with concurrent Haemorrhage (SMQ) within 5 days of thrombocytopenia or platelet count decreased will also be generated.

7.11.1.12 Acute Renal Failure (SMQ broad and narrow)

7.11.1.13 Cardiac failure (SMQ broad and narrow)

7.11.1.14 Cardiomyopathy (SMQ broad and narrow)

7.11.1.15 Cardiac arrhythmia (SMQ broad and narrow)

7.11.1.16 Drug related hepatic disorders – comprehensive search (SMQ broad and narrow)

7.11.1.17 Vascular disorders

- Embolic and thrombotic events (SMQ broad and narrow)
- Hypertension (SMQ broad and narrow)

7.11.1.18 Haemodynamic oedema, effusions and fluid overload (SMQ narrow terms only)

7.11.1.19 Haematopoietic cytopenias (SMQ narrow terms only) and sub-SMQs

The sub-SMQs are:

- Haematopoetic cytopenias affecting more than one type of blood cell
- Haematopoetic erythropenia
- Haematopoetic leukopenia
- Haematopoetic thrombocytopenia

7.11.1.20 Dehydration (SMQ broad and narrow)

7.11.1.21 Hypersensitivity (SMQ broad and narrow)

7.11.1.22 Infection

By- patient line listing for Patients with ANY GRADE TEAE of Infections (defined by SOC Infections and Infestations. or HLGT Respiratory Tract Infections, or HLT Lower Respiratory Tract Inflammatory and Immunologic Conditions) and concurrent within 5 days of Febrile neutropenia (defined by PT Febrile neutropenia) will be generated. The listing will include: subject ID, Treatment group, Febrile neutropenia (reported term and PT, Start date/End date, Days from first dose/Days from last dose, Seriousness), Infections (reported term and PT, Start date/End date; Days from first dose/Days from last dose; Seriousness).

By- patient line listing for Patients with ANY GRADE TEAE of Infections (defined by SOC Infections and Infestations. or HLGT Respiratory Tract Infections, or HLT Lower Respiratory Tract Inflammatory and Immunologic Conditions) and concurrent within 5 days of Neutropenia (defined by PT Neutropenia, PT Neutrophil Count Decreased, PT White Cell Count Decreased) will be generated. The listing will include: subject ID, Treatment group, Neutropenia (reported term and PT, Start date/End date, Days from first dose/Days from last dose, Seriousness), Infections (reported term and PT, Start date/End date; Days from first dose/Days from last dose; Seriousness).

A summary table of patients with concurrent infections and infestation TEAE (defined by SOC Infections and Infestations. or HLGT Respiratory Tract Infections, or HLT Lower Respiratory Tract Inflammatory and Immunologic Conditions) within 5 days of febrile neutropenia or neutropenia will also be generated.

7.11.1.23 Change in Transfusion

A listing will be generated for patients who take platelets and/or red blood cells as concomitant medications during study to display transfusion trend over time (Week 1-4, Week 5-8, Week 9-12, and Week 13+).

7.11.1.24 Hydration

A listing of patient that required hydration as a concomitant medication during study to display hydration trend over time (Week 1-4, Week 5-8, Week 9-12, and Week 13+) by treatment arm.

7.11.1.25 Dose Modifications due to LFT Abnormalities

A listing of patients that required dose modification of due to LFT abnormalities defined by TEAEs coded to the following HLTs and PTsduring study to display trend over time (Week 1-4, Week 5-8, Week 9-12, and Week 13+) for pevonedistat in combination arm and for azacitidine by treatment arm.

- Acute hepatic failure (PT)
- Hyperbilirubinemia (PT)
- Blood alkaline phosphatase (PT)
- Hepatic function abnormal (PT)
- Blood alkaline phosphatase abnormal (PT)
- Liver function analyses (HLT)
- Blood alkaline phosphatase increased (PT)

7.11.1.26 Dose Modifications due to Renal Abnormalities

A listing of patients that required dose modification due to renal abnormalities [as defined by the TEAEs listed in section 7.11.1.12 (Acute renal failure SMQ)] during study to display trend over time (Week 1-4, Week 5-8, Week 9-12, and Week 13+) for pevonedistat in combination arm and for azacitidine by treatment arm.

7.11.1.27 Dose Modifications due to Myelosuppression

A listing of patient that required dose modification due to myelosuppression [defined by the TEAEs listed in below], plus 2 additional PTs of Thrombocytopenia and Platelet count decreased)] during study to display trend over time (Week 1-4, Week 5-8, Week 9-12, and Week 13+) for pevonedistat in combination arm and for azacitidine by treatment arm.

- Anaemia
  - Anaemia of chronic disease
  - Haemoglobin decreased
  - Anaemia of malignant disease
  - Mean cell haemoglobin decreased
  - Anaemia
  - Haematocrit decreased

  - Red blood cell count decreased
  - Neutropenia
  - Agranulocytosis
- Neutropenia
  - Granulocyte count decreased
  - Neutropenic infection
  - Band neutrophil count decreased
  - Neutropenic sepsis
  - Band neutrophil percentage decreased
  - Neutrophil count abnormal
  - Febrile neutropenia
  - Neutrophil count decreased
  - Idiopathic neutropenia
  - Neutrophil percentage abnormal
  - Leukopenia
  - Neutrophil percentage decreased

7.11.1.28 Overall Summary

The number and percentage of patients who experience any of the following groups will be summarized by treatment arm and azacitidine route:

- Any treatment emergent adverse event (including separate summaries of maximum toxicity grade experienced (Grade 1 to Grade 5))
- Drug-related treatment emergent adverse event (including separate summaries of maximum toxicity grade experienced (Grade 1 to Grade 5))
- Serious treatment emergent adverse event
- Drug related serious treatment emergent adverse event
- Treatment emergent adverse events resulting in study drug discontinuation
- Treatment emergent adverse events that required dose modification
- On-study deaths

The summary tables will also be provided for overall TEAE summary by age (<65, >=65 and <75, and >=75), including on-study deaths, Grade 3 or higher TEAEs, drug related adverse

events, serious adverse events, drug related serious adverse events, adverse events leading to treatment discontinuations.

7.11.2 Clinical Laboratory Evaluations

For the purposes of summarization in both the tables and listings, all laboratory values will be converted to standardized units. If a lab value is reported using a non-numeric qualifier (eg, less than (<) a certain value, or greater than (>) a certain value), the given numeric value will be used in the summary statistics, ignoring the non-numeric qualifier.

Laboratory test results from the central laboratory will be used when they are available. Laboratory test results from local laboratory will be used only when no central laboratory test result exists at the same scheduled sample collection time point.

If a patient has repeated laboratory values for a given time point, the value from the last evaluation will be used.

Laboratory test results will be summarized according to the scheduled sample collection time point. Change from baseline will also be presented. Unscheduled laboratory test results will be listed and included in laboratory shift tables.

Shift tables of the change in NCI CTC from baseline to the post baseline worst CTC grade will be generated for relevant measurements. Summary tables will be generated to display the actual values and percent changes from baselines for selected labs. Graphical displays will be used to show changes in laboratory measures over time for patients:

- Box graphs and line plots of individual tests over time by treatment arm and azacitidine route.
- Scatter plots of baseline versus worst post-baseline values for all patients. Separate plotting characters will be used for each combination of treatment arm by azacitidine route. These will be generated for only selected labs in Table 7.e.

Table 7.e Selected Labs

Panel	Test	CTCAE Shift Table	Box Graphs/Line Plots	Scatter Plots	Summary Tables
Chemistry	Albumin	X	X		
	ALT	X	X		X
	AST	X	X		X
	Alkaline Phosphatase	X	X		
	Carbon Dioxide		X		X
	Direct Bilirubin		X		
	Total Bilirubin	X	X		X
	Blood Urea Nitrogen		X	X	
	Corrected Calcium	X	X		
	Chloride		X	X	
	Creatinine	X	X		
	Creatinine Clearance		X	X	X
	Glucose	X	X		
	Lactate Dehydrogenase (LDH)		X	X	
	Magnesium	X	X		
	Phosphate	X	X		X
	Potassium	X	X		X
	Sodium	X	X		
	Urate	X	X		
Hematology	Platelets	X	X		X
	Hemoglobin	X	X		
	Leukocytes	X	X		
	Neutrophils (ANC)	X	X		X
	Monocytes		X		
Additional	Reticulocyte		X		X
	Ferritin		X		X

For patients with neutrophil lab results reported as segmented neutrophils and neutrophil bands, ANC will be calculated as:

If neutrophil values is available, ANC= neutrophils. Otherwise, ANC=total leukocyte count × total percentage of neutrophils (segmented neutrophils + band neutrophils)

Example:

If total leukocyte count = 4.3 x 10^3; segmented neutrophils = 48%; band neutrophils = 2% Then: 4300 × (0.48 + 0.02) = 4300 × 0.5 = ANC of 2150

Creatinine clearance will be derived using one of the Cockcroft-Gault and CKD-epi formulas as follows:

Cockcroft-Gault equation:

For males:

$$\text{Creatinine Clearance (mL/min)} = \frac{(140 - \text{age[years]}) \times \text{weight [kg]}}{0.81 \times (\text{serum creatinine } [\mu\text{mol/L}])}$$

OR

$$\text{Creatinine Clearance (mL/min)} = \frac{(140 - \text{age[years]}) \times \text{weight [kg]}}{72 \times (\text{serum creatinine [mg/dL]})}$$

For females:

$$\text{Creatinine Clearance (mL/min)} = \frac{0.85 \times (140 - \text{age[years]}) \times \text{weight [kg]}}{0.81 \times (\text{serum creatinine } [\mu\text{mol/L}])}$$

OR

$$\text{Creatinine Clearance (mL/min)} = \frac{0.85 \times (140 - \text{age[years]}) \times \text{weight [kg]}}{72 \times (\text{serum creatinine [mg/dL]})}$$

A cap value of 125 will be set to creatinine clearance (calculated from Cockcroft-Gault equation) higher than 125.

CKD-EPI equation (http://nephron.com/epi_equation):

For males:

$$\text{GFR (mL/min/1.73 m}^2) = 141 \times \min(\text{Scr}/0.9, 1)^{-0.411} \times \max(\text{Scr}/0.9, 1)^{-1.209} \times 0.993^{\text{Age}},$$

where Scr = serum creatinine (mg/dL).

For black males:

$$\text{GFR(mL/min/1.73 m}^2) = 141 \times \min(\text{Scr}/0.9, 1)^{-0.411} \times \max(\text{Scr}/0.9, 1)^{-1.209} \times 0.993^{\text{Age}} \times 1.159,$$

where Scr = serum creatinine (mg/dL).

For females:

$$GFR\ (mL/min/1.73\ m^2) = 141 \times \min(Scr/0.7, 1)^{-0.329} \times \max(Scr/0.7, 1)^{-1.209} \times 0.993^{Age} \times 1.018$$

,

where Scr = serum creatinine (mg/dL).

For black females:

$$GFR\,(mL/min/1.73\,m^2) = 141 \times \min(Scr/0.7, 1)^{-0.329} \times \max(Scr/0.7, 1)^{-1.209} \times 0.993^{Age} \times 1.018 \times 1.159$$

,

where Scr = serum creatinine (mg/dL). All chemistry and hematology lab data will also be presented in by-patient listings.

The percentage of marrow progenitor cells in peripheral blood will be presented in by-patient listings, including leukemic blasts, myeloblasts, promyelocytes, myelocytes, metamyelocytes, and uncharacterized blasts.

In addition, the urinalysis parameters will be presented in by-patient listings. These include turbidity and color, pH, specific gravity, protein, ketones, bilirubin, occult blood, nitrite, urobilinogen, glucose, erythrocytes, leukocyte esterase, and leukocytes.

Events that potentially met the biochemical criteria for Hy's law (eg, patients with any elevated aminotransferase of >3x ULN and alkaline phosphatase <2x ULN, in association with an increase in bilirubin ≥2x ULN) will also be provided for overall and by cycles. Incidences of the following will be provided:

- >3x-, >5x-, >10x-, and >20x ULN elevations of AST and/or ALT;
- Any elevations of bilirubin: elevated total bilirubin to >2x ULN;
- Any elevations of alkaline phosphatase >1.5x ULN;
- Elevation of aminotransferase (>3x ULN) accompanied by elevated bilirubin (>1.5x ULN, >2x ULN); and
- Potential Hy's law cases. The Sponsor qualifies these as "potential" cases, since a bona fide case definition requires that no other cause nor other drug has been shown to be causative than the test article. In some advanced cases with more cholestasis, the alkaline phosphatase may be >2x ULN.

7.11.3 Vital Signs

Boxplots over time for temperature, DBP, SBP, and heart rate during Cycle 1 will be generated. Vital sign data will also be presented in a by-patient listing.

Summary table of weight and percent change from baseline in weight over time will be provided.

7.11.4 12-Lead ECGs

The number and percent of patients experiencing abnormal ECG results will be summarized for each time point by treatment arm and azacitidine route.

QTcF and QTcB will be derived using the following formulas.

$$QTcF = \frac{QT_{uncorrected}}{\left(\frac{60}{Ventricular\ Rate}\right)^{1/3}} \qquad QTcB = \frac{QT_{uncorrected}}{\sqrt{\frac{60}{Ventricular\ Rate}}}$$

ECG findings will also be presented in by-patient listings.

7.11.5 Other Observations Related to Safety

Eastern Cooperative Oncology Group performance status and change from baseline will be summarized. Shifts from baseline to the worst postbaseline score will be tabulated by treatment arm.

7.12 Interim Analysis

There is 1 primary endpoint, EFS, 1 key secondary endpoint of OS, and adaptive EFS event-size re-estimation.

There are two interim analyses and one final analysis. The first IA is planned to evaluate EFS for futility and re-assess the EFS events size for patients with HR MDS for IA2.The second IA will be EFS FA in the patients with HR MDS for US submission and in the ITT population for ex-US submission. The FA will evaluate OS.

The IAs will be conducted by the independent statistical center (ISC) and presented to the IDMC for review.

The first IA will be performed when approximately 74 EFS events in patients with HR MDS (around 50% information relative to the minimal planned event size of 147 EFS events for patients with HR MDS for the EFS final analysis) have occurred, targeting to evaluate EFS for futility and re-assess EFS event size for IA2 of the EFS FA. If EFS hazard ratio >1.0 in all three of the following populations: patients with HR MDS, patients with HR MDS/CMML and the ITT population, the study will stop; otherwise, EFS event-size re-estimation will be performed in patients with HR MDS using the conditional power. The EFS event-size adaptation rule is a prespecified stepwise function based on the conditional power, which is calculated on the basis of Liu and Hu [4], to avoid the back calculation problem. The adaptation rules will be outlined in a separate document and will not be accessible to the sponsor's study team until completion of the study. The rules will be available only to the IDMC, and the sponsor's independent design statistician, the sponsor's head of biostatistics, and the statistics representative in the sponsor's executive committee (if different from the sponsor's head of biostatistics), who are not involved in the study conduct.

The conditional power is calculated assuming the observed effect size in patients with HR MDS at the first IA is the true effect size in patients with HR MDS. The calculation is based on B-value [14]. Let D_1 denote the number of EFS events observed at the first IA and D_2 denote the minimal planned number of EFS event at the EFS final analysis in patients with HR MDS, then the information fraction for the conditional power calculation is

$$t = \frac{D_1}{D_2}.$$

Let Z_t denote the negative value of the logrank test statistic in patients with HR MDS at the first IA. The B-value at the first IA is calculated as

$$B(t) = Z_t\sqrt{t}.$$

When data are monitored at t, $B(t)$ is observed, and the conditional power assuming the trend indicated by the interim data is calculated as

$$CP_t = 1 - \Phi\left\{\frac{z_{0.975} - \frac{B(t)}{t}}{\sqrt{1-t}}\right\},$$

where $\Phi(\cdot)$ is the standard normal cumulative density function.

The second IA is to evaluate the primary endpoint EFS (for US submission: in patients with HR MDS; for ex-US submission: in the ITT population) for efficacy. It will be performed when approximately the adaptive EFS event size (from a minimum of 147 to a maximum of 249), informed by the independent data monitoring committee, has occurred in patients with HR MDS. If EFS is not statistically significant in both the HR MDS population and the ITT population at IA2, the study will stop for futility. If statistical significance is achieved in the HR MDS population, then OS will be tested in the HR MDS population for US submission. If statistical significance is achieved in the ITT population, then OS will be tested in the ITT population for ex-US submission. If OS is statistically significant in both the HR MDS population and the ITT population at IA2, the study will stop for efficacy. Otherwise, the study will continue to the FA. The FA will be conducted to evaluate the key secondary endpoint OS in the HR MDS population (for US submission) and in the ITT population (for ex-US submission) when approximately 202 OS events have occurred in patients with HR MDS. Separate multiple hierarchical testing procedures for the US submission and the ex-US submission will be adopted to test the primary endpoint of EFS and secondary endpoint of OS in the HR MDS population (US submission), the ITT population (ex-US submission), and other disease populations at IA2 and FA, with each procedure having a total 1-sided alpha of 0.025, as described in Section 7.8.3.

7.13 Changes in the Statistical Analysis Plan

Reference materials for this statistical plan include Clinical Study Protocol Pevonedistat-3001 Amendment 13 (Protocol Amendment 13 dated 26 May 2021).

The major changes to the previous SAP include:

Clarifying that the methods and testing procedures for analyzing EFS and OS when the prespecified number of approximately 202 OS events for FA are expected (based on blinded study data) to be available close to the IA2, the IA2 (ie, EFS FA) and the FA (ie, OS FA) will be performed as a single analysis. In this situation, separate multiple hierarchical testing procedures for the US submission and the ex-US submission will be used to test the primary endpoint of EFS and the key secondary endpoint of OS in the HR MDS population (US submission), the ITT population (ex-US submission), and other disease populations at this single analysis, with a total 1-sided alpha of 0.025 for each procedure. The 2 hierarchical testing procedures originally specified at IA2 and FA will be performed as a single testing procedure with a 1-sided alpha of 0.025 used for each test as depicted on Panel B of Figure 7.a (US submission testing procedure) and on Panel B of Figure 7.b (ex-US submission testing procedure). The weights of the CHW test statistics for testing EFS are the same as those used for testing EFS. The weights of the CHW test statistics for testing OS are prespecified based on the observed number of OS events at IA, and the targeted 202 OS events in patients with HR MDS and the estimated number of OS events in the other specific population of interest matching to the targeted 202 OS events in patients with HR MDS for FA (see Appendix in Section 9.0 for the estimated number of OS events).

8.0 REFERENCES

1. Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Sole F, et al. Revised international prognostic scoring system for myelodysplastic syndromes. Blood 2012;120(12):2454-65.
2. Cheson BD, Greenberg PL, Bennett JM, Lowenberg B, Wijermans PW, Nimer SD, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. Blood 2006;108(2):419-25.
3. Cheson BD, Bennett JM, Kopecky KJ, Buchner T, Willman CL, Estey EH, Revised recommendations of the International Working Group for diagnosis, standardization of response criteria, treatment outcomes, and reporting standards for therapeutic trials in acute myeloid leukemia. Journal of Clinical Oncology 2003;21(24):4642-9.
4. Liu Y, Hu MX. Testing multiple primary endpoints in clinical trials with sample size adaptation. Pharmaceutical Statistics 2016; 15(1):37-45.
5. Dohner H, Estey E. et al. Diagnosis and management of AML in adults. Blood 2017; 129(4):424-427.
6. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood 2002;100(7):2292-302.
7. Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics 1999; 55(3):853-7.
8. Lan KK, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika. 1983; 70:659– 663
9. Robins JM, Hernan M, Brumback B. Marginal strucural models and causal inference in epidemiology. Epidemiology 2000; 11(5): 550-560.
10. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 2000;56(3):779-88.
11. Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine 2010; 29 (): 219-228.
12. Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels: European Organisation for Research and Treatment of Cancer; 2001.
13. Silverman LR, McKenzie DR, Peterson BL, Holland JF, Backstrom JT, Beach CL, 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. Journal of Clinical Oncology 2006; 24(24):3895-903.
14. K. K. Gordon Lan, & Wittes, J. The B-Value: A Tool for Monitoring Data. Biometrics 1988; 44(2): 579-585.

9.0 APPENDIX

9.1 Estimated Number of EFS and OS Events at IA2 and FA

Table 9.a and Table 9.b below provide the estimated number of EFS and OS events in different disease populations at IA2 (matching to the minimum number of 147 EFS events in patients with HR MDS planned for IA2) and FA (matching to the targeted 202 OS events in patients with HR MDS planned at FA), respectively. The estimations were based on the actual study enrollment. For the estimations of EFS events numbers, the assumptions of hazard ratios and median EFS time defined in the protocol were used. The estimations of OS event numbers used the assumption of hazard ratios defined in the protocol and the median OS times based on the emerging data from the final analysis of the phase 2 study, which had the same study design and assessed pevonedistat combination with azacitidine among the same disease populations as is done in this study.

Table 9.a Estimated Number of EFS and OS Events at IA2

Endpoint	Population	IA2 (assuming 147 EFS events in patients with HR MDS)
EFS	ITT	210
	HR MDS	147
	HR CMML	12
	HR MDS/CMML	159
OS	ITT	182
	HR MDS	120
	HR CMML	11
	HR MDS/CMML	131
	Low-blast AML	51

Table 9.b Estimated Number of OS Events at FA

Endpoint	Population	FA (202 OS events in patients with HR MDS)
OS	ITT	294
	HR MDS	202
	HR CMML	18
	HR MDS/CMML	220
	Low-blast AML	74

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
[redacted]	Biostatistics Approval	18-Jun-2021 15:38 UTC