1.0 Title Page

Statistical Analysis Plan

Study M16-043

A Randomized, Double-Blind, Placebo Controlled Study of Venetoclax Co-Administered with Low Dose Cytarabine Versus Low Dose Cytarabine in Treatment Naïve Patients with Acute Myeloid Leukemia Who Are Ineligible for Intensive Chemotherapy

Date: 30 Nov 2018

Version 3.0

2.0 Table of Contents

List of Tables

3.0 Introduction

This statistical analysis plan (SAP) describes the full efficacy and safety statistical analyses for venetoclax Study M16-043 (Viale-c) Protocol Amendment 4 dated 29 November 2018. It will provide details of statistical methods and describe analysis conventions to guide the statistical programming work.

Unless noted otherwise, all analyses will be performed using SAS® version 9.3 or later (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

4.0 Study Objectives, Design and Procedures

4.1 Objectives

The primary objective of this study is to evaluate if venetoclax when co-administered with low dose cytarabine (LDAC) improves overall survival (OS) versus LDAC and placebo, in treatment naïve subjects with acute myeloid leukemia (AML).

The secondary objectives of the study are

- To evaluate if venetoclax when co-administered with LDAC improves composite complete remission rate (complete remission + complete remission with incomplete blood count recovery, CR + CRi).
- To evaluate if venetoclax when co-administered with LDAC improves complete remission and complete remission with partial hematologic recovery rate (CR + CRh).
- To evaluate if venetoclax when co-administered with LDAC improves the proportion of subjects achieving a composite CR (CR + CRi) by the initiation of Cycle 2.
- To evaluate if venetoclax when co-administered with LDAC improves complete remission rate (CR).
- To evaluate if venetoclax when co-administered with LDAC reduces fatigue based on patient reported outcome (PRO) assessment of the Patient Reported

Outcomes Measurement Information System (PROMIS), and Fatigue Short Form (SF) 7a.

- To evaluate if venetoclax when co-administered with LDAC improves subjects Global Health Status/Quality of Life (GHS/QoL) based on the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core (EORTC QLQ-C30).
- To evaluate if venetoclax when co-administered with LDAC improves event-free survival (EFS).
- To evaluate if venetoclax when co-administered with LDAC improves the transfusion independence rates.
- To evaluate if venetoclax when co-administered with LDAC improves the proportion of subjects achieving complete remission and complete remission with partial hematologic recovery (CR + CRh) by the initiation of Cycle 2.
- To evaluate if venetoclax when co-administered with LDAC improves the MRD response rate.
- To evaluate if venetoclax when co-administered with LDAC improves the response rates and overall survival in molecular subgroups (e.g., IDH1/2, FLT3).

The exploratory objectives of this study are

- Exploration of biomarkers predictive of venetoclax activity and duration of response may be performed. These analyses may be part of a multi-study assessment to compare responses to the therapies and/or disease state. Potential analyses may include, but will not be limited to:
  - To evaluate BCL-2 expression and outcome measures of overall survival and complete remission rate.
- To evaluate the impact of venetoclax based on the remaining subscales/items from the EORTC QLQ-C30 and EQ-5D-5L.

4.2 Study Design and Plan

This is a Phase 3, randomized, double-blind, placebo controlled, multicenter study of venetoclax co-administered with LDAC versus LDAC in treatment naïve patients with acute myeloid leukemia who are ineligible for intensive induction chemotherapy sponsored by AbbVie. Subjects will be randomized to one of the two treatment arms in a 2:1 ratio, both of which will have treatment cycles of 28 days.

- Arm A: Venetoclax 600 mg orally QD on Days 1 – 28 plus LDAC 20 mg/m^2 SC QD on Days 1 – 10
- Arm B: Placebo for venetoclax 600 mg orally QD on Days 1 – 28 plus LDAC 20 mg/m^2 SC QD on Days 1 – 10

Approximately 210 subjects will be randomized to meet scientific and regulatory objectives without enrolling an undue number of subjects in alignment with ethical considerations.

4.3 Sample Size

The primary endpoint of the study is overall survival and sample size calculation is based on the following assumptions:

- Median OS of 6 months for placebo plus LDAC arm
- Median OS of 11 months for venetoclax plus LDAC arm (hazard ratio of 0.545)
- Interim analysis of OS at 75% of death events with O'Brien-Fleming boundary
- 2:1 randomization ratio to venetoclax plus LDAC, and placebo plus LDAC arm

With the above assumptions, a total of 133 death events will provide 90% power to detect statistically significant difference between treatment arms at two-sided alpha level of 0.05. A total of approximately 210 subjects (140 in venetoclax plus LDAC arm and 70 in

placebo plus LDAC arm) will be randomized into the study to obtain the 133 death events.

4.4 Interim Analysis

To ensure subject safety, an Independent Data Monitoring Committee (IDMC) will review unblinded safety data at the following proposed time points:

- The first meeting to review the unblinded safety data will occur approximately 3 months after the 20th patients has been randomized and dosed
- The subsequent reviews of unblinded safety data will occur approximately every 3 months after the first review of unblinded safety data.

Additional ad hoc meetings may occur as needed for the IDMC to fulfill its responsibilities.

Further, one interim analysis (IA) of OS will be performed once 100 death events (75% of the total 133 events) are observed. The Lan-DeMets alpha spending function with O'Brien-Fleming boundary will be used to ensure that the one-sided false positive rate will be 0.025 or less for overall survival. The IDMC will make a recommendation based on overall survival to either to stop for success at IA or proceed to the final analysis using pre-specified stopping rule. The planned stopping boundaries are further described in Table 1 below for IA and final analyses (FA) of OS endpoints. The actual stopping boundaries at IA and FA of OS endpoint will be derived using Lan-DeMets alpha spending function, based on the observed number of death events in the extracted database. The actual stopping boundaries will be presented in the IDMC charter addendum. Details of the IDMC review will be presented in the IDMC charter.

For the IA, if the IDMC inform the sponsor that OS data is statistically significant in favor of Arm A, upon notification from the Sponsor Steering Committee after consultation with regulatory authorities as required, the sponsor will then unblind the study to prepare regulatory submissions globally. Otherwise, study will remain blinded and continue to follow protocol specified procedures until the FA.

Table 1. Planned Stopping Boundaries at the IA and FA of OS Endpoint

Look	# of Events	Efficacy Stopping Boundaries (One-Sided p-Value)
IA	100	0.010
FA	133	0.022

Interim statistical analyses and summaries for presentation to IDMC will be prepared by Axios. AbbVie personnel will not have access to the interim analyses prepared for the IDMC.

5.0 Analysis Sets

5.1 Definition for Analysis Sets

Two study populations will be analyzed, defined as follows:

- The full analysis set consists of all subjects randomized by IVRS/IWRS. The full analysis set will be used for efficacy analyses. The data from full analysis set will be analyzed by the treatment arm assignment given at the time of randomization, even if the subject takes the incorrect drugs that do not match the assigned treatment, or does not receive any treatment, or does not follow the protocol until completion.
- The safety analysis set consists of all subjects who take at least one dose of venetoclax/placebo or LDAC. The safety analysis set will be used for safety analyses. The data from safety analysis set will be analyzed by the actual treatment that subject received.

5.2 Variables Used for Stratification of Randomization

The subject randomization will be stratified by AML status (de novo, secondary), age ($18 - < 75, \geq 75$), and region (US, EU, China, Japan (JP), Rest of World (ROW)).

6.0 Analysis Conventions

General Considerations

Unless otherwise noted, for all statistical analyses, statistical significance will be determined by a two-sided *P* value ≤ 0.05. The date of randomization is defined as the date that the IVRS/IWRS issues a randomization number.

Definition of Study Drug

Unless otherwise specified, the study drug in this document refers to venetoclax/placebo and LDAC. The first dose date of study drug is defined as the date the 1st dose of venetoclax/placebo or LDAC, whichever occurs first, is administrated. The last dose date of study drug is defined as the date the last dose of venetoclax/placebo or LDAC, whichever occurs later, is administrated.

Definition of Baseline

All baseline summary statistics and analyses will be based on characteristics prior to the initiation of any component of study drug (or randomization for non-treated subjects). Unless otherwise stated, baseline for a given variable will be defined as the last value for that variable obtained prior to the first dose of any component of study drug or randomization for non-treated subjects.

Definition of Final Observation

The final observation in Table 2 and Table 3 is defined as the last non-missing observation collected after the first dose of study drug to the last dose of study drug + 30 days.

Definition of Data Cutoff Date

The data cutoff date for interim analysis (IA) will be the projected date of the 100th death occurred in the full analysis set. The data cutoff date for final analysis (FA) will be the projected date of the 133th death occurred in the full analysis set.

The same data cutoff date will be implemented for both efficacy and safety analyses.

Stratification Factor for Efficacy Analyses

AML status (secondary, de novo) and age (18 – < 75, ≥ 75) will be used in all stratified analyses of the efficacy endpoints. The stratification factor value under which the subject is randomized by the IVRS/IWRS will be used in the efficacy analyses.

Definition of Cycle Rx Days in Each Cycle

Cycle Rx Days for each cycle are calculated for each time point relative to first dose of study drug in each cycle. The day of the first dose of study drug in each cycle is defined as Cycle Rx Day 1, while the day prior to the first dose of study drug is defined as Cycle Rx Day –1 (there is no Cycle Rx Day 0).

Definition of Analysis Windows

All time points and corresponding time windows are defined for each cycle are based on Cycle Rx Day 1 of each cycle.

For visit-wise clinical laboratory analyses, vital signs analyses, and quality of life analyses, the time windows specified in Table 2 and Table 3 describe how efficacy and safety data are assigned to protocol specified visits respectively. Analysis time windows are constructed using the following algorithm:

- Determine the nominal Cycle Rx day for each scheduled visit.
- Determine the window around a specific nominal Cycle Rx day as in Table 2 and Table 3.
- If more than one assessment are included in a time window of the assessment, the one closest to the nominal day should be used. If there are two observations with equal distance to the nominal day, the later one will be used in analyses. If multiple values are collected on the same day, the arithmetic average will be calculated and used as the value of that day for the mean change analyses.

- For analysis of tumor lysis syndrome (TLS) laboratory variables meeting the Howard criteria where multiple values were collected over the course of a day (within 4 hours prior to dosing and 6 – 8 hours post dose), all values will be used and no averages will be taken.
- For analysis of liver enzyme laboratory variables meeting criteria for potential drug-induced liver injury, all values will be used and no averages will be taken.
- For laboratory shift tables, the Common Terminology Criteria (CTC) grade will be assigned based on all observed laboratory values and no averages will be taken. If multiple grades are collected on the same day, the worst grade will be used as the value of that day for the shift table analyses.

Table 2. Time Windows for Analysis of Hematology, Chemistry, and Vital Signs Parameters

Scheduled Visit	Nominal Cycle Rx Day	Time Window (Cycle Rx Day Range)
Baseline	≤ 1	See the baseline definition (Section 6.0)
Cycle 1 Day 2	2	[2]
Cycle 1 Day 3	3	[3]
Cycle 1 Day 4	4	[4]
Cycle 1 Day 5	5	[5]
Cycle 1 Day 8	8	[6, 11]
Cycle 1 Day 15	15	[12, 18]
Cycle 1 Day 21	21	[19, 25]
Cycle 2 Day 1	1	[–3, 10]
Every Cycle from Cycle 3 Day 1	1	[–10, 10]
Final Observation	NA	The last non-missing observation collected after the first dose of study drug to the last dose of study drug + 30 days

Note: Hematology and chemistry samples will be collected at visits specified in the protocol. Vital signs will be measured at visits specified in the protocol.

Table 3. Time Windows for Analysis of PROMIS Cancer Fatigue SF 7a, EORTC QLQ-C30, and EQ-5D-5L

Scheduled Visit	Nominal Cycle Rx Day	Time Window (Cycle Rx Day Range)
Baseline	≤ 1	See the baseline definition (Section 6.0)
Every Other Cycle from Cycle 3 Day 1	1	[–10, 10]

Note: PROMIS Cancer Fatigue SF 7a, EORTC QLQ-C30, and EQ-5D-5L will be assessed at visits specified in the protocol.

7.0 Subject Disposition

The screen failure reasons will be summarized for the screen failure subjects. Analyses for the subject disposition will be performed on the full analysis set.

The number and percentage of subjects will be summarized overall, by treatment arm, by country, and by site, for each of the following categories:

- Subjects randomized into the study
- Subjects who received at least one dose of study drug
- Subjects who randomized but not treated with any component of study drug

Additionally, the number and percentage of subjects will be summarized overall, and by treatment arm for each of the following categories:

- Subjects who discontinued venetoclax/placebo due to all reasons and primary reason
- Subjects who discontinued LDAC due to all reasons and primary reason
- Subjects who discontinued study due to all reasons and primary reason
- Subjects enrolled under each Protocol Amendment

There will be no statistical comparison for the subject disposition between the two treatment arms.

8.0 Demographics, Baseline Characteristics, Medical History, Prior Medications, Concomitant Medications, and Post-treatment Therapies

8.1 General Consideration

All demographic, baseline characteristics, medical history, prior and concomitant medication summaries will be performed by treatment arm, on full analysis set as described in Section 5.1.

8.2 Demographic and Baseline Characteristics

All baseline characteristic summary statistics and analyses are based on characteristics prior to the first dose of study drug or date of randomization for non-treated subjects.

Distributions of the continuous demographic and baseline characteristic variables will be summarized by treatment arm with the number of non-missing observations, mean, standard deviation, and median, as well as the minimum and maximum values.

For the categorical demographic and baseline characteristic variables, the frequency and percentages of subjects within each category will be summarized by treatment arm. The number of subjects with missing information will also be summarized.

There will be no statistical comparison for the demographic and baseline characteristics between the two treatment arms.

The following demographic and baseline characteristics will be summarized:

Demographics:

- age (years) and age categories (18 – < 75 years, ≥ 75 years)
- gender (male/female)
- race (white, black or African American, Asian, and other)
- region (US, EU, China, JP, and ROW)
- height (cm)

- weight (kg)

Baseline and Disease-Related Characteristics:

- ECOG performance status (0, 1, 2, 3)
- AML Status (de novo, secondary)
- Type of secondary AML (therapy related, post MDS/CMML)
- AML with myelodysplasia related changes (AML-MRC) (Yes, No)
- Cytogenetics risk (favorable, intermediate, poor)
- Bone marrow blast count (< 30%, ≥ 30% – < 50%, ≥ 50%)
- Bone marrow blast count (%)
- CTC grade of neutropenia
- Neutrophils value ($\times 10^9$/L)
- CTC grade of anemia
- Hemoglobin value (G/L)
- CTC grade of thrombocytopenia
- Platelet count ($\times 10^9$/L)
- Reasons for being not eligible for standard induction therapy
- Prior HMA used (Yes, No)
- Antecedent hematologic history of MDS (Yes, No)
- Baseline RBC transfusion dependence (Yes, No)
- Baseline platelet transfusion dependence (Yes, No)
- Molecular marker via central lab (FLT3, IDH1/2, TP53, NPM1)
- Hepatic impairment (Yes, No)
- Renal impairment (Yes, No)

The cross tabulation of the following demographic and baseline characteristics captured in IVRS/IWRS and CRF will be generated:

- Age categories (18 – < 75, ≥ 75 years)
- AML status (de novo, secondary)

- Cytogenetics risk (favorable, intermediate, poor)
- Region (US, EU, China, JP, and ROW)

8.3 Medical History

Medical history data will be coded for conditions/diagnoses by system organ class and preferred term. The number and percentage of subjects with a particular condition/diagnosis will be summarized by system organ class and preferred term. Subjects reporting more than one condition/diagnosis within a system organ class will be counted only once for that system organ class. There will be no statistical comparison for the medical history between the two treatment arms.

8.4 Prior Medications, Concomitant Medications, and Post-treatment Therapies

A prior medication is defined as any medication taken on or prior to the first dose of study drug. A concomitant medication is defined as any medication that started prior to the first dose of study drug, and continued to be taken after the first dose of study drug, or any medication that started after the first dose of study drug, but not after the last dose of study drug. All medications are considered prior medications for subjects who were randomized but did not receive any study drug. A post-treatment therapy for the treatment of AML is defined as any medications taken after the discontinuation of study drug and entered via Post Treatment eCRFs.

The number and percentage of subjects who have taken medications will be summarized by generic drug name coded by WHO Drug dictionary 2017Q1 or a more recent version for prior medications, prior oncology therapies, concomitant medication, and post-treatment therapies. The number and percentage of subjects who have taken TLS prophylaxis agents including intravenous hydration on or prior to the first dose of study drug and concomitant to the study treatment will be summarized separately. The number and percentage of subjects who have taken transfusions on or prior to first dose of study

drug will be summarized by transfusion type. There will be no statistical comparison between the two treatment arms for the aforementioned summaries.

For summaries of concomitant medications, if an incomplete or missing start date was collected for a medication, the medication will be assumed to be a concomitant medication, or prior medication for subjects who did not receive any study drug, unless there is evidence that confirms that the medication was not a concomitant medication (e.g., the medication end date was prior to the first dose of study drug).

A subject who reports the use of two or more medications will be counted only once in the summary of "Any Concomitant Medication." A subject who reports two or more uses of the same medication will be counted only once in the total for the associated generic drug name. Similar rules apply to prior medications and post-treatment therapy as well.

9.0 Study Treatment Exposure and Compliance

The duration of exposure to venetoclax/placebo will be summarized by treatment arm, for safety analysis set as described in Section 5.1. Duration of exposure is defined for each subject as (last dose date – first dose date) + 1. Duration of exposure will be summarized using the following statistics (months): sample size (N), mean, standard deviation, median, and range. In addition, the number and percentage of subjects exposed to venetoclax/placebo will be summarized for the following categories of exposure duration:

- 0 to 4 weeks (0 to 28 days),
- > 4 weeks to 8 weeks (28 to 56 days),
- > 8 weeks to 12 weeks (57 to 84 days),
- > 12 weeks to 16 weeks (85 to 112 days),
- > 16 weeks to 20 weeks (113 to 140 days),
- > 20 weeks to 24 weeks (141 to 168 days),
- > 24 weeks to 28 weeks (169 to 196 days),
- > 28 weeks to 32 weeks (197 to 224 days),
- > 32 weeks to 36 weeks (225 to 252 days),

- > 36 weeks to 52 weeks (253 to 364 days),
- > 52 weeks (> 364 days).

The number of cycles that subjects are exposed to study drug will be summarized by treatment arm. There will be no statistical comparison for the study treatment exposure between the two treatment arms.

10.0 Efficacy Analysis

10.1 General Considerations

Unless otherwise noted, for all statistical analysis, statistical significance will be determined by a two-sided *P* value ≤ 0.05 (when rounded to three decimal places).

When the 133st death in the full analysis set occurs, there will be a final review of the eCRF data. After the data collection is completed and reviewed for completeness and all data management quality assurance (QA) and quality control (QC) procedures are performed, the study blind will be broken and clinical database data will be extracted for documentation and statistical analyses of the efficacy and safety data. Efficacy analyses will be performed on the full analysis set.

Unless otherwise specified, the primary analysis of all response/progression related endpoints (e.g., CR/CRi) will be based on the investigator assessment. Sensitivity analyses will be performed based on Independent Review Committee's assessment.

For the analysis of all response/progression related endpoints, MRD response rate, and transfusion independence rate, the data collected after the data cutoff date or initiation of post treatment therapy, whichever comes earlier, will be excluded.

Censoring Dates for Prematurely Blind Broken Subjects

For all efficacy endpoints except OS, if the subject's blind has prematurely broken by sites before "Cutoff" date, the date of premature blind break will be set as "Cutoff" date for the subject.

10.2 Primary Efficacy Analysis

10.2.1 Overall Survival

Overall survival will be defined as the number of days from the date of randomization to the date of death. All events of death will be included, regardless of whether the event occurred while the subject was still taking study drug, or after the subject discontinued study drug. If a subject has not died, then the data will be censored at the date the subject was last known to be alive on or before the cutoff date. The date of the last known alive will be determined by selecting the last available date of the following study procedures for a subject: start date of adverse event, bone marrow collection, disease assessment, vital signs assessment, clinical laboratory collection, study drug administration, start date of concomitant medicine, transfusion procedure, survival follow-up, biospecimen sample collection, quality of life assessments, and performance status. All subjects in the full analysis set will be included in the analysis.

The distribution of overall survival will be estimated for each treatment arm using Kaplan-Meier methodology and compared between treatment arms using the log-rank test, stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). The hazard ratio between treatment arms will be estimated using the Cox proportional hazards model, stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75).

10.3 Secondary Efficacy Analyses

10.3.1 CR + CRi Rate

CR + CRi rate will be defined as the proportion of subjects who achieve a complete remission or complete remission with incomplete blood count recovery (CR + CRi) at any time point during the study per the modified IWG criteria for AML. Subjects who are randomized but have no IWG disease assessment will be considered to be non-responders in the calculation of CR + CRi rate. All subjects in the full analysis set will be included in the analysis.

CR + CRi rate will be compared between treatment arms using Cochran-Mantel-Haenszel (CMH) test, stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). In addition, the 95% confidence interval for CR + CRi rate based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided. The risk difference and 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

10.3.2 CR + CRh Rate

CRh (Complete remission with partial hematologic recovery) is a derived response based on bone marrow blast and hematology lab values. A subject achieves a CRh when meeting the following criteria:

- Bone marrow with < 5% blasts and
- Peripheral blood neutrophil count of > $0.5 \times 10^3/\mu L$ and
- Peripheral blood platelet count of > $0.5 \times 10^5/\mu L$ and
- A 1 week (≥ 7 days) platelet transfusion-free period prior to the hematology lab collection.

For a bone marrow sample collected before the last cycle of study treatment, the hematology lab results collected from the date of the bone marrow sample collection up to the Day 1 of a subsequent cycle of study treatment will be used for CRh analysis. For a bone marrow sample collected during or after the last cycle of study treatment, the hematology lab results collected within 14 days after bone marrow sample collection date will be used for CRh analysis.

CR + CRh rate will be defined as the proportion of subjects who achieve a complete remission (CR) or complete remission with partial hematologic recovery (CRh) at any time point during the study. Subjects who are randomized but have no disease assessment will be considered as non-responders in the calculation of CR + CRh rate. All subjects in the full analysis set will be included in the analysis.

CR + CRh rate will be compared between treatment arms using Cochran-Mantel-Haenszel (CMH) test stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). In addition, the 95% confidence interval for CR + CRh rate based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided. The risk difference and the 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

10.3.3 CR + CRh and CR + CRi rate by the Initiation of Cycle 2

CR + CRh rate by the initiation of Cycle 2 will be defined as the proportion of subjects who achieved a complete remission (CR) or complete remission with partial hematologic recovery (CRh) by the initiation of Cycle 2. Subjects who are randomized but have no disease assessments by the initiation of Cycle 2 will be considered to be non-responders in the calculation of CR + CRh rate by the initiation of Cycle 2. For subjects who discontinue the treatment before initiation of Cycle 2, all disease assessments evaluated before the cutoff date and the initiation of post treatment therapy, whichever occurs earlier, will be included. All subjects in the full analysis set will be included in the analysis.

CR + CRh rate by the initiation of Cycle 2 will be compared between treatment arms using Cochran-Mantel-Haenszel (CMH) test stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). In addition, the 95% confidence interval for CR + CRh rate by the initiation of Cycle 2 based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided. The risk difference and the 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

Similar analyses will be performed for CR + CRi rate by the initiation of Cycle 2.

10.3.4 Complete Remission (CR) Rate

CR rate will be defined as the proportion of subjects who achieved a complete remission (CR) at any time point during the study per the modified IWG criteria for AML. Subjects

who are randomized but have no IWG disease assessment will be considered to be non-responders in the calculation of CR rate. All subjects in the full analysis set will be included in the analysis.

CR rate will be compared between treatment arms using Cochran-Mantel-Haenszel (CMH) test, stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). In addition, the 95% confidence interval for CR rate based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided. The risk difference and 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

10.3.5 Transfusion Independence Rates

Post baseline transfusion independence rate will be calculated as the proportion of subjects who achieved transfusion independence post baseline. Transfusion independence is defined as a period of at least 56 days (≥ 56 days) with no transfusion after the first dose of study drug and on or before the last dose of study drug + 30 days or before death or before the initiation of post-treatment therapy, whichever is earlier. In addition, the rate of conversion will be calculated as proportion of subjects being post-baseline transfusion independent from baseline transfusion dependence. The transfusion independence rate will be evaluated for 1) RBC and 2) platelets. All subjects in the full analysis set will be included in the analysis. Subjects who did not receive any study drug are considered as transfusion dependent.

The post-baseline transfusion independence rates will be compared between two arms using CMH test stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). In addition, 95% confidence interval based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided for the post-baseline transfusion independence rate. The risk difference and 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

The rates of conversion from baseline transfusion dependence to post-baseline transfusion independence for 1) RBC and 2) platelets will be estimated and the 95% confidence interval based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided. The risk difference and 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

10.3.6 Minimal Residual Disease (MRD) Response Rates

The proportion of subjects achieved 1) CR + CRi and MRD response 2) CR + CRh and MRD response will be calculated among all subjects in the full analysis. The MRD response will be defined using a threshold of less than 10^{-3} of residual blasts per leukocytes as measured in bone marrow. Subjects who are randomized but have no MRD assessment will be considered as non-responders for the calculation of MRD response rate. All subjects in the full analysis set will be included in the analysis.

The proportion of subjects who achieved CR + CRi and MRD response or CR + CRh and MRD response will be compared between two arms using CMH test stratified by age (18 – < 75, ≥ 75) and AML status (de novo, secondary). In addition, 95% confidence interval based on the binomial distribution (Clopper-Pearson exact method) by treatment arms will be provided. The risk difference and 95% confidence interval for the risk difference (exact unconditional confidence limits) will also be provided.

10.3.7 Quality of Life (QoL)

Any subject who does not have baseline measurement or any post-baseline measurements in the full analysis set will not be included in QoL analyses. Post-baseline measurements will be obtained according to the visit window as in Table 3.

PROMIS Cancer Fatigue SF 7a

PROMIS instruments measure concepts such as pain, fatigue, physical function, depression, anxiety and social function. (https://www.assessmentcenter.net/documents/PROMIS%20Fatigue%20Scoring%20Manual.pdf). Fatigue will be assessed using the PROMIS Cancer Fatigue SF that has been

developed for use in oncology populations. PROMIS Cancer Fatigue SF 7a is a seven item questionnaire that assesses the impact and experience of fatigue over the past 7 days. All questions employ the following five response options: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, and 5 = Always.

Scores will be computed according to the procedures outlined in the PROMIS Fatigue scoring manual, available at https://www.assessmentcenter.net/Manuals.aspx. A linear mixed effects regression model with an appropriate covariance structure will be fitted to the longitudinal data to test for differences between the two treatment arms. For the linear mixed effects regression model, stratification factors (age (18 – < 75, ≥ 75) and AML status (de novo, secondary)) and treatment arm will be included as fixed factors. Furthermore, time and treatment by time interaction will be included in the model. The repeated correlation structure in the timepoints will be assessed by using the Bayesian Information Criterion (BIC). The following covariance structures will be explored: Unstructured (TYPE = UN), compound symmetry (TYPE = CS) and first-order autoregressive (TYPE = AR(1)). The type resulting in model convergence and the lowest BIC will be used for analysis. Change in the PROMIS Fatigue score from baseline will be compared between the two treatment arms for each post baseline visit except the final observation.

EORTC QLQ-C30

The QLQ-C30 is a 30-item subject self-report questionnaire composed of both multi-item and single scales, including five functional scales (physical, role, emotional, social, and cognitive), three symptom scales (fatigue, nausea and vomiting, and pain), a global health status/quality of life scale, and six single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties). Subjects' rate items on a four-point scale, with 1 as "not at all" and 4 as "very much."

Scores will be computed according to procedures outlined in the EORTC QLQ-C30 scoring manual, available at http://groups.eortc.be/qol/manuals. A linear mixed effects regression model with an appropriate covariance structure will be fitted to the longitudinal

data to test for differences between the two treatment arms. For the linear mixed effects regression model, stratification factors (age (18 – < 75, ≥ 75) and AML status (de novo, secondary)) and treatment group will be included as fixed factors. Furthermore, time and treatment by time interaction will be included in the model. The repeated correlation structure in the timepoints will be assessed by using the Bayesian Information Criterion (BIC). The following covariance structures will be explored: Unstructured (TYPE = UN), compound symmetry (TYPE = CS) and first-order autoregressive (TYPE = AR(1)). The type resulting in model convergence and the lowest BIC will be used for analysis. Change in the PROMIS Fatigue score from baseline will be compared between the two treatment arms for each post baseline visit except the final observation.

10.3.8 Event-free Survival

Event-free Survival (EFS) will be defined as the number of days from randomization to the date of progressive disease, confirmed morphologic relapse from CR or CRi, treatment failure defined as failure to achieve CR, CRi or MLFS after at least 6 cycles of study treatment collected on study drug completion eCRF or death from any cause. The confirmation is required for morphologic relapse (MR). Based on the IWG criteria implemented in the protocol, if the bone marrow blasts are between 5 - 10%, a repeated bone marrow assessment should be conducted to distinguish between bone marrow regeneration and relapse. For MR with bone marrow blast ≤ 10% and followed by a non-PD/non-MR response prior to start of post-treatment therapy, this MR is considered an unconfirmed MR and therefore, the subject is still considered to be a responder.

If the specified event does not occur, subjects will be censored. The detailed censoring rule is described in Table 4. Data for subjects without any disease assessments performed after randomization will be censored at the time of randomization.

Table 4. Event/Censor and Corresponding Event/Censor Time for EFS

Situation	Event/Censor	Event Time/Censor Time
PD, confirmed morphologic relapse (MR),* and treatment failure events on or prior to any post-treatment therapy and data cutoff date, whichever is earliest	Event	Earliest PD, MR or treatment failure event date
Death event on or prior to data cutoff date regardless whether it occurred after any post-treatment therapy initiated	Event	Death date
No PD, confirmed MR,* treatment failure, or death events, and without post-treatment therapy initiated on or prior to the data cutoff date	Censor	Last disease assessment date (bone marrow or hematology lab collection date) on or prior to the data cutoff date
No confirmed MR,* PD, or treatment failure on or prior to post-treatment therapy initiated on or prior to the data cutoff date, and no death on or prior to data cutoff	Censor	Start date of post-treatment therapy

* An un-confirmed MR is defined as a MR with bone marrow blast ≤ 10% and followed by non-PD/Non-MR prior to the post-treatment therapy. Otherwise, MR is considered to be a confirmed MR.

The distribution of EFS will be estimated for each treatment arm using Kaplan-Meier methodology and compared between treatment arms using the log-rank test stratified by AML status (de novo, secondary) and age (18 – < 75, ≥ 75). Median EFS will be calculated and 95% confidence interval for median EFS will be presented by treatment arm. The hazard ratio between treatment arms will be estimated using the Cox proportional hazards model stratified by age (18 – < 75, ≥ 75) and AML status (de novo, secondary).

10.4 Exploratory Efficacy Endpoint

10.4.1 Additional Quality of Life Analyses

EQ-5D-5L

The EQ-5D-5L has five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These dimensions are measured on a five level scale: no problems, slight problems, moderate problems, severe problems, and extreme problems.

The scores for the 5 dimensions are used to compute a single utility index score ranging from zero (0.0) to 1 (1.0) representing the general health status of the individual. The EQ-5D-5L also contains a visual analog scale (VAS) to assess the subject's overall health.

Additional exploratory analyses comparing the effects of venetoclax + cytarabine versus placebo + LDAC will be performed on the following PRO measures: EQ-5D-5L, and the subscales/items from the EORTC QLQ-C30 and PROMIS Cancer Fatigue SF 7a. Scores will be calculated as per the scoring manuals for all scales/items of the EORTC QLQ-C30, PROMIS Cancer Fatigue SF 7a, the EQ-5D-5L utility score, and the EQ-5D VAS score at each assessment. Linear mixed effects regression models similar to that described in Section 10.3.7 will be used to test for differences between treatment arms. Within-group changes from baseline and between-group comparisons at each assessment will also be assessed. Additional analyses may include an assessment of time to deterioration and time to improvement.

10.4.2 Duration of CR + CRi

Duration of CR + CRi will be defined as the number of days from the date of first response (CR or CRi) per the modified IWG criteria for AML to the earliest evidence of confirmed morphologic relapse, progressive disease or death due to disease progression. If a subject is still responding at the data cutoff date, then the subject will be censored. The disease assessment data after the onset of any post-treatment therapy will not be included in the duration of CR + CRi analysis. The detailed censoring rule is described in Table 5. Duration of CR + CRi will be analyzed for all randomized subjects who achieve the best response of CR or CRi.

Table 5. Event/Censor and Corresponding Event/Censor Time for Duration of CR + CRi

Situation	Event/Censor	Event Time/Censor Time
Confirmed morphologic relapse (MR)* or progressive disease (PD) on or prior to any post-treatment therapy and data cutoff date, whichever is earliest	Event	Earliest MR or PD event date
Death due to disease progression on or prior to data cutoff date regardless whether it occurred after any post-treatment therapy initiated	Event	Death date
No confirmed MR,* PD, or death due to disease progression events and without post-treatment therapy initiated on or prior to the data cutoff date	Censor	Last disease assessment date (bone marrow or hematology lab collection date) on or prior to the data cutoff date
No confirmed MR,* or PD on or prior to post-treatment therapy initiated on or prior to the data cutoff date, no death due to disease progression on or prior to data cutoff date	Censor	Start date of post-treatment therapy

* An un-confirmed MR is defined as a MR with bone marrow blast ≤ 10% and followed by non-PD/Non-MRprior to the post-treatment therapy. Otherwise, MR is considered to be a confirmed MR.

The distribution of duration of CR + CRi will be estimated for each treatment arm using Kaplan-Meier methodology. Median duration of CR + CRi will be calculated and 95% confidence interval for median duration of CR + CRi will be presented by treatment arm.

Analysis of modified duration of CR + CRi that includes death from all causes by the cutoff date will be performed as a sensitivity analysis.

10.4.3 Duration of CR

Duration of CR will be defined as the number of days from the date of first CR per the modified IWG criteria for AML to the earliest evidence of morphologic relapse, progressive disease, or death due to disease progression. The censoring rule described in Table 5 will be applied to the analysis of duration of CR.

Analysis of modified duration of CR that includes death from all causes by the cutoff date will be performed as a sensitivity analysis.

10.4.4 Duration of CR + CRh

Duration of CR + CRh will be defined as the number of days from the date of first response (CR or CRh) to the earliest evidence of morphologic relapse, progressive disease or death due to disease progression. The censoring rule described in Table 5 will be applied to the analysis of duration of CR + CRh.

Analysis of modified duration of CR + CRh that includes death from all causes by the cutoff date will be performed as a sensitivity analysis.

10.5 Efficacy Analysis per Independent Review Committee

An Independent Review Committee (IRC) will evaluate disease assessment data on all randomized subjects. The analysis for the following efficacy endpoints, described in Section 10.3 and Section 10.3.8, will be conducted based on this IRC review:

- CR rate
- CR + CRi rate
- CR + CRi rate by the initiation of Cycle 2
- EFS

10.6 Additional Efficacy Analysis

In addition to the stratified log-rank test for the primary and secondary efficacy endpoints, the following analyses may be performed for the comparison of OS, CR rate, CR + CRi rate, CR + CRh rate, CR + CRi rate by the initiation of Cycle 2, and CR + CRh rate by the initiation of Cycle 2 between the two treatment groups.

- Un-stratified log-rank test and the Cox proportional hazards model for OS
- CMH test for CR rate and CR + CRi rate using IRC data
- Fisher's exact test for CR rate and CR + CRi rate using investigator assessment data and IRC data

- Modified analyses for OS, CR rate, CR + CRi rate, CR + CRi rate by the initiation of Cycle 2, CR + CRh rate, and CR + CRh rate by the initiation of Cycle 2 to use all data available in the extracted database (no "Cutoff" date).

10.7 Handling of Multiplicity

The fixed sequence testing procedure will be performed with a significance level of 0.05 for the primary endpoint OS and key secondary efficacy endpoints sequentially. If statistical test is not significant for the primary efficacy endpoint of OS, then statistical significance will not be declared for any of the secondary efficacy endpoints. The Lan-DeMets alpha spending function with O'Brien-Fleming boundary will be used at the interim and final analysis to ensure that the false positive rate for each primary or key secondary efficacy endpoint is 0.05 or less.

For CR + CRi rate, CR + CRi rate by the initiation of Cycle 2, CR + CRh rate, CR + CRh rate by the initiation of Cycle 2, CR rate, transfusion independence rates, and MRD response rate, the information fraction to be used at IA (100 death events) is calculated as detailed in Table 6.

Table 6. Calculation of Information Fraction (IF)

Endpoint	Information Fraction
CR + CRi rate	Portion of subjects with at least 6 months follow up since randomization
CR + CRh rate	Portion of subjects with at least 6 months follow up since randomization
MRD response rate	Portion of subjects with at least 6 months follow up since randomization
CR + CRi rate by the initiation of Cycle 2	Portion of subjects who started Cycle 2 Day 1 dosing or discontinued treatment by the end of Cycle 1
CR + CRh rate by the initiation of Cycle 2	Portion of subjects who started Cycle 2 Day 1 dosing or discontinued treatment by the end of Cycle 1
CR rate	Portion of subjects with at least 9 months follow up since randomization
Transfusion independence rates (RBC/Platelet)	Portion of subjects with at least 9 months follow up since randomization

Table 7 (for EU and EU reference countries) and Table 8 (for US and US reference countries) present the hierarchical ranking and alpha spending for each ranked primary and secondary endpoints at IA and FA.

Table 7. Alpha-spending Boundary (One-sided p-value) for Each Ranked Endpoint for EU and EU Reference Countries

Endpoint		Timing of Analysis	
		IA	FA
1	OS	As specified in Table 2	As specified in Table 2
2	CR + CRi rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
3	Post-baseline platelet transfusion independence[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
4	CR + CRi rate in IDH1/2 subgroup[a,b]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
5	CR rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
6	CR + CRi by the initiation of Cycle 2[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
7	Post-baseline RBC transfusion independence[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
8	MRD and CR + CRi response rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
9	EORTC QLQ-C30	0.0001	0.025
10	PROMIS Cancer Fatigue SF 7a	0.0001	0.025
11	OS in IDH1/2 subgroup[b]	0.0001	0.025
12	OS in FLT3 subgroup[b]	0.0001	0.025
13	CR + CRi rate in FLT3 subgroup[b]	0.0001	0.025
14	EFS	0.0001	0.025

a. The Lan DeMets with O'Brien-Fleming approach will be applied to determine the efficacy boundaries of the following ranked secondary endpoints with the information fraction define in Table 6.

b. If the size of identified subgroup population is less than 30 or randomization ratio between Venetoclax + LDAC arm and placebo + LDAC arm is higher than 3:1 at the time of IA and FA, the endpoint will be ranked below EFS in the fixed sequence testing procedure.

Table 8. Alpha-spending Boundary (One-sided p-value) for Each Ranked Endpoint for US and US Reference Countries

Endpoint		Timing of Analysis	
		IA	FA
1	OS	As specified in Table 2	As specified in Table 2
2	CR + CRh rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
3.	CR + CRi rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
4	Post-baseline platelet transfusion independence[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
5	CR + CRh rate in IDH1/2 subgroup[a,b]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
6	CR rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
7	CR + CRh by the initiation of Cycle 2[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
8	Post-baseline RBC transfusion independence[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
9	MRD and CR + CRh response rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
10	MRD and CR + CRi response rate[a]	To be calculated with IF as specified in Table 6	To be calculated with IF as specified in Table 6
11	PROMIS Cancer Fatigue SF 7a	0.0001	0.025
12	EORTC QLQ-C30	0.0001	0.025
13	OS in IDH1/2 subgroup[b]	0.0001	0.025
14	OS in FLT3 subgroup[b]	0.0001	0.025
15	CR + CRh rate in FLT3 subgroup[b]	0.0001	0.025
16	EFS	0.0001	0.025

a. The Lan DeMets with O'Brien-Fleming approach will be applied to determine the efficacy boundaries of the following ranked secondary endpoints with the information fraction define in Table 6.

b. If the size of identified subgroup population is less than 30 or randomization ratio between Ven + LDAC arm and placebo + LDAC arm is higher than 3:1 at the time of IA and FA, the endpoint will be ranked below EFS in the fixed sequence testing procedure.

10.8 Efficacy Subgroup Analyses

To evaluate the impact of demographic and baseline characteristics on efficacy, subgroup analyses will be performed, but not limited to, for CR rate, CR + CRi rate, CR + CRh rate, CR + CRi rate by the initiation of Cycle 2, CR + CRh rate by the initiation of Cycle 2, and overall survival. The subgroups in the full analysis set including, but not limited to, those defined below will be used for these analyses:

1. Gender (Male, Female)
2. Age (18 – < 65 years, 65 – < 75 years, ≥ 75 years)
3. AML status (de novo, secondary)
4. Type of secondary AML (therapy related, post MDS/CMML)
5. Region (US, EU, China, JP, ROW)
6. Baseline ECOG (< 2, ≥ 2)
7. Subjects who received prior HMA for MDS (Yes, No)
8. Cytogenetic risk categorization (Favorable, Intermediate, Poor)
9. Molecular marker(FLT3, IDH1/2, TP53, NPM1)
10. AML with Myelodysplasia related changes (AML-MRC)

11.0 Safety Analysis

Safety assessments will only include subjects in the safety analysis set as described in Section 5.1. Safety will be assessed by treatment received for the safety analysis set.

11.1 Analysis of Treatment-Emergent Adverse Events

All summaries/analyses involving AEs will include treatment-emergent adverse events (TEAE) only, unless otherwise specified. TEAEs are defined as any adverse event with onset after the first dose of study drug and no more than 30 days after the last dose of

study drug. Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent. If an incomplete or missing onset date was collected for an AE, the AE will be assumed to be treatment-emergent unless there is evidence that confirms that the AE was not treatment-emergent (e.g., the AE end date was prior to the date of the first dose of study drug).

For summaries of AEs related (reasonable possibility) to study drug, at each level of summation (overall, SOC, and PT) each subject is counted only once. If a subject has an AE with unknown relationship, then the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same event with a relationship of "no reasonable possibility" present. The only exception is if the subject has another occurrence of the same AE with the relationship of "reasonable possibility." In this case, the subject will be counted under the reasonable possibility category.

Adverse event data will be summarized and presented using primary MedDRA system organ classes (SOCs) and preferred terms (PTs) according to the MedDRA coding dictionary version 21.0 or higher.

The number and percentage of subjects experiencing treatment-emergent adverse events will be summarized for the following adverse event summaries:

- Any treatment-emergent adverse event.
- Any treatment-emergent adverse event with NCI toxicity (CTCAE V4.03) grade 3, 4, or 5 adverse events.
- Any treatment-emergent serious adverse event.
- Any treatment-emergent adverse event with reasonable possibility related to venetoclax/placebo by the investigator.
- Any treatment-emergent adverse event leading to discontinuation of venetoclax/placebo.
- Any treatment-emergent adverse event leading to venetoclax/placebo interruption.

- Any treatment-emergent adverse event leading to venetoclax/placebo reduction.
- Any treatment-emergent adverse event with reasonable possibility related to LDAC by the investigator.
- Any treatment-emergent adverse event leading to discontinuation of LDAC.
- Any treatment-emergent adverse event leading to LDAC interruption.
- Any treatment-emergent adverse event leading to LDAC reduction.
- Any treatment-emergent adverse event leading to death.

In addition, the treatment emergent AEs and serious AEs for selected grouped preferred terms (PTs) including, but not limited to, those listed in Table 9, will be summarized.

Table 9. Selected Adverse Events

Selected Adverse Events	Search Criteria
Tumor Lysis Syndrome (AE)	SMQ – "Tumor Lysis Syndrome" (narrow)
Grade ≥ 3 neutropenia	PT terms – "neutropenia," "neutrophil count decreased," "febrile neutropenia," "agranulocytosis," "neutropenic infection," and "neutropenic sepsis"
Grade ≥ 3 infection, including opportunistic infections	SOC of "infections and infestations"
Haemorrhages	SMQ – "Haemorrhages" (narrow)
Anemia	PT terms – "Anaemia" and "Haemoglobin decreased"
Thrombocytopenia	PT terms – "Thrombocytopenia" and "Platelet count decreased"

There will be no statistical comparison for TEAEs between two treatment arms.

11.2 Deaths

The number of subject deaths will be summarized (1) for all deaths in this study regardless of the number of days after the last dose of study drug, (2) for deaths occurring within 30 days of the first dose of study drug, (3) for deaths occurring within 60 days of the first dose of study drug, (4) for deaths occurring within 30 days of the last dose of

study drug, and (5) for deaths occurring more than 30 days of the last dose of study drug. There will be no statistical comparison for above analyses.

11.3 Analysis of Laboratory and Vital Signs Data

The value for baseline used in laboratory is defined in Section 6.0. Post baseline visits windows are specified in Table 2. There will be no statistical comparison for laboratory data between two treatment arms.

11.3.1 Analysis of Mean Changes from Baseline in Clinical Laboratory and Vital Signs Data

Changes from baseline will be analyzed for each chemistry, hematology, and vital sign parameters at scheduled post-baseline visits and the final observation (Table 2). Descriptive statistics including mean, standard deviation and median for baseline, and mean, standard deviation, median, minimum and maximum for change from baseline will be presented for baseline, each scheduled post-baseline visit, and final observation by each treatment arm. Laboratory tests to be summarized are included in Table 10 (if data is available).

Table 10. Clinical Laboratory Tests

Hematology	Clinical Chemistry
Hematocrit Hemoglobin Red blood cell (RBC) count White blood cell (WBC) count Neutrophils Bands (if detected) Blast count Lymphocytes Monocytes Basophils (if detected) Eosinophils (if detected) Platelet count (estimate not acceptable)	Blood urea nitrogen (BUN) Creatinine Calculated or Measured creatinine clearance Total bilirubin Serum glutamic-pyruvic transaminase (SGPT/ALT) Serum glutamic-oxaloacetic transaminase (SGOT/AST) Alkaline phosphatase Sodium Potassium Calcium Inorganic phosphorus Uric acid Total protein Glucose Albumin Lactate dehydrogenase (LDH) Chloride Bicarbonate
Coagulation	
Prothrombin time (PT) Activated partial thromboplastin time (aPTT)	

11.3.2 Analyses of Shift from Baseline in Clinical Laboratory Data

For shifts relative to National Cancer Institute Common Toxicity Criteria for Adverse Events (NCI CTCAE version 4.03), baseline and post-baseline laboratory observations will be categorized as grade 0, grade 1, grade 2, grade 3, or grade 4 according to NCI CTCAE grade version 4.03.

The baseline and final grades will be defined respectively as the grade of the last measurement collected on or prior to the first dose of study drug (or randomization for non-treated subjects), and as the last post-baseline measurement collected no more than 30 days after the last dose of study drug.

The maximum NCI toxicity grade value is the value with highest NCI toxicity grade collected after the first dose of study drug and within 30 days following the last dose of

study drug. In cases where multiple values are collected on the same day, the maximum grade value will be selected as the value for that day.

For each variable, shift tables will be generated that cross tabulate the number of subjects with baseline values of grade 0, grade 1, grade 2, grade 3, or grade 4 versus maximum post-baseline/final observations of grade 0, grade 1, grade 2, grade 3, or grade 4. All treated subjects will be included in the cross tabulation regardless whether baseline or post-baseline measurements are collected.

The separate laboratory shifts tables, based on the two criteria below will be generated for each laboratory tests related to CTCAE:

1. Shifts from Grade 0 (Normal) at baseline to grade 1 - 4 Post-baseline (maximum) and worsening from an abnormal baseline value of at least one grade up post-baseline (maximum)
2. Shits from Grade 0 - 2 at baseline to grade 3 or 4 Post-baseline (maximum) and from grade 3 at baseline value to Grade 4 post-baseline (maximum).

Detailed listings of data for subjects experiencing NCI CTCAE grade 3 to 4 blood chemistry and hematology values will be provided. All measurements collected, regardless of the number of days after the last dose of study drug will be included in these listings.

Number and percentage of subjects with liver Enzyme value meeting the criteria for potential drug-induced liver injury (ALT > 3 × ULN or AST > 3 × ULN and TBILI > 2 × ULN within 72 hours of each other) will be presented.

Number and percentage of subjects meeting the Howard criteria for Laboratory TLS will be presented. Laboratory TLS requires two or more metabolic abnormalities must be present during the same 24-hour period. The evaluation period for TLS is after the first dose of study drug during Cycle 1 until 7 days from the first dose of study drug.

- uric acid > 476 mcmol/l,
- potassium > 6 mmol/l,
- inorganic phosphorus > 1.5 mmol/l
- calcium < 1.75 mmol/l.

11.4 Safety Subgroup Analysis

Safety analysis as described in Section 11.1, Section 11.2, and Section 11.3, may also be summarized for the subgroups in the safety analysis set including, but not limited to, those defined below:

- Region (US, EU, China, JP, ROW)
- Gender (Male, Female)
- Age (18 – < 65 years, 65 – < 75 years, ≥ 75 years)
- Hepatic impairment at baseline (Yes, No)
- Renal impairment at baseline (Yes, No)

12.0 Pharmacokinetic (PK) Analyses

Plasma concentrations of study drug will be listed for each subject by arm and scheduled visit. Summary statistics will be computed for each arm and dose level by scheduled visit. For Chinese subjects enrolled in China, pharmacokinetic parameter values will be tabulated for each subject, and summary statistics will be computed for each pharmacokinetic parameter. Samples with significant sampling time deviations will be excluded from summary statistics calculations.

13.0 References

1. Cheson BD, Bennett JM, Kopecky KJ, et al. International working group for diagnosis, standardization of response criteria, treatment outcomes, and reporting standards for therapeutic trials in acute myeloid leukemia. 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. J Clin Oncol. 2003;21(24):4642-9.

2. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-47.
3. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika. 1983;70(3):659-63.
4. Howard SC, Jones DP, Pui CH. The tumor lysis syndrome. N Engl J Med. 2011;364(19):1844-54.
5. Xu T, Qin Q, Wang X. Defining information fraction in group sequential clinical trials with multiple endpoints. Contemp Clin Trials Commun. 2018;10:77-9.

14.0 Appendix

List of Abbreviations

AML	acute myeloid leukemia
BCL	B cell lymphocyte
BTD	Breakthrough Therapy Designation
CI	confidence interval
CR	complete remission
CRi	complete remission with incomplete blood count recovery
CRh	complete remission with partial hematologic recovery
CTC	Common Terminology Criteria
ECOG	Eastern Cooperative Oncology Group
EFS	event-free survival
EORTC-QLQ-C30	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core
FEV1	forced expiratory volume in 1 second
FLT3	FMS-like tyrosine kinase 3
GFR	glomerular filtration rate
HCT	hematopoietic stem cell transplantation
HCT-CI	Hematopoietic Stem Cell Transplant Co-Morbidity Index
HMA	hypomethylating agent
DMC	data monitoring committee
IRC	independent review committee
ITD	internal tandem duplication FLT3
IWG	International Working Group
LDAC	low-dose cytarabine
MDS	myelodysplastic syndrome
MLFS	morphologic leukemia-free state
MR	morphologic relapse
MRD	minimal residual disease
OS	overall survival
PD	progressive disease
PR	partial remission
PROMIS	Patient Reported Outcomes Measurement Information System

QD	once daily
TLS	tumor lysis syndrome
US	United States

Document Approval

Study M16043 - Statistical Analysis Plan Version 3 - 30Nov2018 (E3 16.1.9)

Version: 1.0 **Date:** 30-Nov-2018 08:45:20 PM **Company ID:** 11302018-00F9F683FD0A23-00001-en

Signed by:	Date:	Meaning Of Signature:
[redacted]	30-Nov-2018 07:23:18 PM	Approver
	30-Nov-2018 07:25:26 PM	Approver
	30-Nov-2018 07:42:51 PM	Author
	30-Nov-2018 07:44:48 P	Approver
	30-Nov-2018 07:50:33 PM	Approver
	30-Nov-2018 08:45:19 PM	Approver