Statistical Analysis Plan for Study M19-708

A Randomized, Open Label, 2-Arm, Multicenter, Phase 3 Study of Venetoclax and Azacitidine Versus Best Supportive Care as Maintenance Therapy for Patients with Acute Myeloid Leukemia (AML) in First Remission After Conventional Chemotherapy (VIALE-M)

Date: 21 September 2020

Version 1.0

Table of Contents

1.0 Introduction 5
2.0 Study Design and Objectives 5
2.1 Objectives and Hypotheses 5
2.2 Study Design Overview 6
2.3 Treatment Assignment and Blinding 7
2.4 Sample Size Determination 8
3.0 Endpoints 9
3.1 Primary Endpoint(s) 9
3.2 Secondary Endpoint(s) 9
3.3 Other Efficacy Endpoint(s) 10
3.4 Safety Endpoints 11
4.0 Analysis Populations 11
5.0 Subject Disposition 11
6.0 Study Drug Duration and Compliance 12
7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications 12
7.1 Demographics and Baseline Characteristics 13
7.2 Medical History 14
7.3 Prior and Concomitant Medications 15
8.0 Efficacy Analyses 16
8.1 General Considerations 16
8.2 Handling of Missing Data 16
8.3 Primary Efficacy Endpoint(s) and Analyses 17
8.3.1 Primary Efficacy Endpoint(s) 17
8.3.2 Handling of Missing Data for the Primary Efficacy Endpoint(s) 17
8.3.3 Primary Efficacy Analysis 17
8.3.4 Additional Analyses of the Primary Efficacy Endpoint(s) 18
8.4 Secondary Efficacy Analyses 18
8.4.1 Key Secondary Efficacy Analyses 18
8.4.2 Supportive Secondary Efficacy Analyses 21
8.5 Additional Efficacy Analyses 21

8.6 Efficacy Subgroup Analyses 23
9.0 Safety Analyses 24
9.1 General Considerations 24
9.2 Treatment-Emergent Adverse Events 24
9.2.1 Adverse Event Overview 25
9.2.2 Treatment-Emergent Adverse Events by SOC and/or PT 26
9.2.3 SAEs and Deaths 26
9.2.4 Adverse Events of Special Interest 27
9.3 Analysis of Laboratory Data 28
9.4 Safety Subgroup Analyses 29
10.0 Interim Analyses 30
10.1 Data Monitoring Committee 31
11.0 Overall Type-I Error Control 32
12.0 Version History 33
13.0 References 33

List of Tables

Table 1. Event/Censor and Corresponding Event/Censor Time for RFS 17
Table 2. AESI Search Criteria 27
Table 3. Selected Adverse Events 27
Table 4. Stopping Boundaries for IAs and FA of RFS 30
Table 5. Testing Sequence and Alpha-Spending Boundaries (One-Sided P value) for the Primary and Key Secondary Endpoints 32
Table 6. SAP Version History Summary 33

List of Figures

Figure 1. Study Schematic 7

List of Appendices

Appendix A. Protocol Deviations 35

Appendix B. Time Windows for Analysis of EQ 5D-5L, EORTC QLQ-C30, and PROMIS Cancer Fatigue SF 7a 36
Appendix C. Time Windows for Analysis of Clinical Hematology and Chemistry Parameters 37
Appendix D. Conversion of leukocyte differential counts from percentage (%) value to absolute value 38
Appendix E. Data Cutoff Date 39
Appendix F. Hepatic and Renal Function 40

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for Venetoclax Study M19-708 protocol (version 2.0) titled "Randomized, Open-label, 2-Arm, Multicenter, Phase 3 Study of Venetoclax and Azacitidine Versus Best Supportive Care as Maintenance Therapy for Patients with Acute Myeloid Leukemia in First Remission After Conventional Chemotherapy (VIALE-M)."

The analyses of pharmacokinetic endpoints and biomarker exploratory endpoints will not be covered in this SAP.

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

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

2.0 Study Design and Objectives

2.1 Objectives and Hypotheses

The primary objectives of the study are to determine the recommended Phase 3 dose (RPTD) of venetoclax in combination with azacitidine (AZA) as maintenance therapy in subjects with AML who have achieved CR or CRi with conventional chemotherapy (Part 1) and to evaluate if venetoclax in combination with AZA based on Part 1 recommended dose as maintenance therapy improve relapse-free survival (RFS) comparing to best support care (BSC) in subjects with AML who have achieved CR or CRi with conventional chemotherapy (Part 2).

The secondary objectives for Part 2 are

- To evaluate if venetoclax in combination with AZA as maintenance therapy improves OS in comparison to BSC.

- To evaluate the safety of venetoclax in combination with AZA as maintenance therapy in comparison to BSC.
- To evaluate if venetoclax in combination with AZA as maintenance therapy improves the minimal residual disease (MRD) conversion rate among subjects who are MRD-positive at study initiation in comparison to BSC.
- To evaluate if venetoclax in combination with AZA as maintenance therapy when compared to BSC delays time to deterioration in Global Health Status (GHS)/Quality-of Life (QoL) score as measured by procedures outlined in the European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire-Core 30-item (EORTC QLQ-C30) scoring manual.
- To evaluate if venetoclax in combination with AZA as maintenance therapy has impact on fatigue when compared to BSC based on Patient-Reported Outcome (PRO) assessment of the Patient Reported Outcomes Measurement Information System (PROMIS) Fatigue Short Form (SF) 7a.

Exploratory

- To explore biomarkers predictive of venetoclax activity. These analyses may be part of a multi-study assessment to compare responses to the therapies and/or disease state.
- To evaluate the impact of venetoclax-based therapy on remaining subscales/items from the EORTC QLQ-C30, version 3.0 and the European Quality-of-Life-5 dimesional-5-level (EQ-5D-5L) questionnaire.
- To evaluate if venetoclax in combination with AZA as maintenance therapy impacts health care utilization in comparison to BSC.

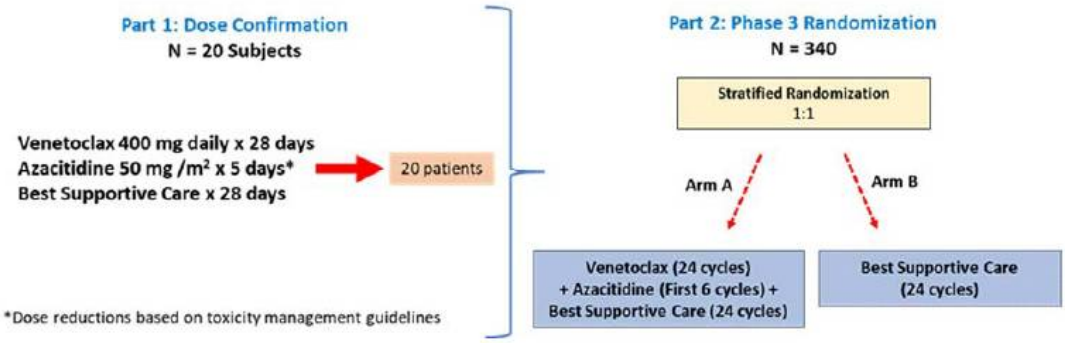
2.2 Study Design Overview

This is a Phase 3, randomized, open-label, 2-arm, multicenter study with an initial dose confirmation component to establish the recommended dose of venetoclax in combination with AZA in this treatment setting. This study will establish the benefit and risk profile of venetoclax + AZA versus BSC as maintenance therapy in subjects with AML after first remission resulting from conventional (including both induction and consolidation)

chemotherapy. BSC is defined as the best supportive care and expectant management according to institutional standards and will exclude any AML directed therapy. BSC will be determined for each subject by the investigator. Approximately 360 subjects will be enrolled in the trial. Approximately 20 subjects will be enrolled into Part 1 (Dose Confirmation) and approximately 340 subjects into Part 2 (Randomization).

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



2.3 Treatment Assignment and Blinding

This study has two parts:

- Part 1 Dose Confirmation (see protocol V2.0 Section 4.1)
- Part 2 randomized, open label, 2-arm

In Part 2 of the study, subjects will be randomized 1:1 to one of the following two arms:

- Arm A: Receive combination of Venetoclax and AZA for up to the first 6 cycles and Venetoclax (monotherapy) from Cycle 7 to Cycle 24 .
- Arm B: Receive best supportive care (BSC) only.

Randomization will be stratified by:

[redacted]

2.4 Sample Size Determination

This study has two parts. Part 1 is a dose confirmation study and Part 2 is a randomized study. The sample size for Part 1 is dependent upon the dose levels utilized.

The sample size of 340 subjects for Part 2, the randomization portion, is based on the following calculations for the primary endpoint of RFS:

[redacted]

- interim futility analysis of RFS at 40% of RFS events with O'Brien-Fleming boundary
- interim efficacy analysis of RFS at 75% of RFS events with O'Brien-Fleming boundary
- 1:1 randomization ratio for the venetoclax in combination with AZA arm (Arm A) and the BSC arm (Arm B)

[redacted]

The assumed hazard ratio of [redacted] describes a relevant and clinically meaningful improvement over BSC. In the AZA-AML-001 study (Dombret et al), the combined CR + CRi rate for azacitidine monotherapy in AML patients was 28% and the median OS for azacitidine in this population was 10.4 months. DiNardo et al reported in a single arm study a CR + CRi rate for azacitidine combined with venetoclax (400 mg daily for 28 days) of 76%. In this study, the median OS was not reached with a median follow up duration of 15.1 months (DiNardo et al, Blood 2019). As RFS results are attributable to the most recent treatment and not due to subsequent therapies as may be seen with an OS endpoint, the assumed hazard ratio for this study appears to be reasonable based on what

was observed in the treatment effect of venetoclax 400 mg in combination with azacitidine in previous reported study (DiNardo et al, Blood 2019).

A total of [redacted] will provide approximately 90% power to detect a statistically significant difference between the 2 arms for RFS at 1-sided alpha level of 0.025 using the log-rank test.

A total of approximately 340 subjects will be randomized in the study to obtain the [redacted]

The planned sample size of 340 subjects in Part 2 will also provide approximately [redacted]

3.0 Endpoints

3.1 Primary Endpoint(s)

Relapse Free Survival (RFS) is defined as time from randomization to the earliest occurrence of an RFS event measured in days. An RFS event includes the following: a relapse (as evaluated by Independent Review Committee (IRC)) or death from any cause. If a subject does not experience an RFS event at the data cut-off, the subject's data will be censored at subject's last disease assessment date defined as last of bone marrow or complete blood count assessment date on or prior to data cut-off, or randomization date if the subject does not have any post-baseline disease assessment.

3.2 Secondary Endpoint(s)

Secondary endpoints include:

- Overall Survival (OS) is defined as time from date of randomization to death from any cause measured in days. Subjects who have not died will be censored at subject's last known date to be alive.
- Time to deterioration in GHS/QoL score, as measured by procedures outlined in the EORTC QLQ-C30 scoring manual, is defined as time from randomization to death from any cause, or the first time decrease of 5 points or more from baseline in GHS/QoL score, whichever occurs first. Subjects without any of the specified events will be censored at subject's last EORTC QLQ-C30 assessment.
- Overall impact on fatigue score, measured by procedures outlined in the PROMIS Fatigue SF 7a scoring manual, will be evaluated by mean changes from baseline to all post-baseline visits (specified in Appendix B).
- MRD conversion rate is defined as the proportion of subjects deemed MRD positive ($\geq 10^{-3}$) at study initiation who converted to MRD of $< 10^{-3}$ in the bone marrow during the maintenance therapy. Subjects who are deemed MRD positive at baseline but do not have any post-baseline MRD results will be included in the analysis as non-conversion events.

3.3 Other Efficacy Endpoint(s)

The exploratory efficacy endpoints are:

- Change in subject reported signs, symptoms and impact of AML as measured by the EORTC-QLQ-C30, European Quality-of-Life-5 dimesional-5-level (EQ-5D-5L).
- Health Care Resource Utilization: Information will be collected on each hospitalization including reason for admission (e.g., disease relapse, AML-related illness, treatment-related AE) and days of hospitalization by treatment setting (e.g., inpatient unit, special care unit, etc.).

3.4 Safety Endpoints

Safety and tolerability will be assessed by evaluating adverse events (AEs), physical examinations and changes in laboratory data and vital signs for the entire study duration.

4.0 Analysis Populations

The following population sets will be used for the analyses.

The Intent-to-Treat (ITT) Population includes all randomized subjects. The ITT Population will be used for all efficacy analyses. Subjects will be included in the analysis according to the study arm that they are randomized to.

The Safety Analysis Set consists of all subjects who received at least 1 dose of study drug (venetoclax and AZA) in Part 1 or Arm A in Part 2, and all randomized subjects in Arm B in Part 2. Arm A subjects in Part 2 who did not receive any dose of venetoclax and AZA will not be included in the Safety Analysis Set.

5.0 Subject Disposition

A summary of subject accountability will be provided where the number and percentage of subjects in each of the following categories will be summarized for each study arm:

- Subjects who enrolled in the study;
- Subjects who took at least one dose of venetoclax;
- Subjects who took at least one dose of AZA;
- Subjects who prematurely discontinued venetoclax (all reasons and primary reason);
- Subjects who prematurely discontinued AZA (all reasons and primary reason);
- Subjects who prematurely discontinued study (all reasons and primary reason).

In addition, reasons for screen failures will also be summarized. A separate summary of subject disposition for Part 1 subjects will be provided.

6.0 Study Drug Duration and Compliance

Exposure to two study drugs will be summarized for subjects enrolled in Part 1 and in Arm A of Part 2 separately for the Safety Analyses Set. The descriptive statistics (number of non-missing observations, mean, standard deviation, median, and range) will be summarized for each of the following variables:

- Duration of study drug defined for each subject as (last dose date – first dose date) + 1;
- Number of cycles that subjects were exposed to study drug;
- Average dosed days per cycle defined for each subject as (total number of days subject received study drug)/(the number of cycles subject was exposed to study drug).
- Dose intensity accounting for dose reduction and interruption for venetoclax defined as actual total venetoclax dose of each subject divided by the planned total venetoclax dose of venetoclax per protocol (e.g., planned dose reduction due to CYP3A use) for each subject.

The number and percentage of subjects with dose reduction or dose interruption of venetoclax will be summarized by the number of occurrences of dose reduction or interruption reported by investigators as no, 1 time, 2 times, or > 2 times.

In addition, the duration of study drug/number days on study drug based on 28 days of interval or by cycle may be summarized.

The number and percentage of subjects who received CYP3 inhibitors will be summarized by type moderate or strong CYP3 inhibitor.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Unless otherwise specified, all summaries in this section will be presented for Part 1 and Part 2 separately.

Demographics, baseline or disease characteristics, medical history, and prior and concomitant medications will be summarized for the ITT population overall, by study arm, and by dose level for Part 1. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

All demographics and baseline disease characteristic summary statistics and analyses are based on the latest one collected prior to the first dose of study drug or date of randomization for untreated subjects. The following demographic and baseline disease characteristics will be summarized:

Demographics:

- age (years) and age categories (< 60 years, ≥ 60 years)
- gender (male/female)
- race (white, black or African American, Asian, and other)
- region ([redacted], or other regions requested by regulatory agencies)
- height (cm)
- weight (kg)

Baseline Disease-Related Characteristics:

- AML type (De Novo, Secondary AML)
- Secondary AML (Post MDS/CMML/AML, Therapy-related AML)
- Previous treatment with venetoclax (Yes, No)
- Previous treatment with AZA (Yes, No)

- MRD status prior to randomization ($\geq 10^{-3}$, $< 10^{-3}$, Unknown/Missing)
- NCCN cytogenetic risk (Intermediate, Poor)
- Best response after conventional Induction chemotherapy prior to randomization (CR, CRi)
- Best response after conventional Consolidation chemotherapy prior to randomization (CR, CRi)
- Number of Consolidation cycles (1 cycle, 2 cycles, 3 cycles, > 3 cycles)
- ECOG performance status (0, 1, 2)
- Molecular marker at diagnosis (FLT3, IDH1/2, TP53, NPM1, if data is available)
- Hepatic impairment (Yes, No)
- Renal impairment (Yes, No)
- Stratification factor: relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others)
- Bone marrow blast count at diagnosis ($< 30\%$, $\geq 30\% - < 50\%$, $\geq 50\%$)
- Bone marrow blast count at diagnosis (%)
- NCI CTC (National Cancer Institute Common Terminology Criteria, version 5.0) grade of neutropenia at baseline
- Neutrophils value at baseline ($\times 10^9$/L)
- NCI CTC grade of thrombocytopenia at baseline
- Platelet count at baseline ($\times 10^9$/L)
- NCI CTC grade of anemia at baseline
- Hemoglobin value at baseline (G/L)

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized overall and by study arm. The system organ class (SOC) will be presented in

alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic medication name coded by World Health Organization (WHO) dictionary.

For subjects enrolled in Part 1 or Arm A in Part 2, a prior medication is defined as any medication taken prior to the date of the first dose of study drug. A concomitant medication is defined as any medication that started prior to the date of the first dose of study drug and continued to be taken after the first dose of study drug or any medication that started on or after the date of the first dose of study drug, but not after the date of the last dose of study drug.

For subjects enrolled in Arm B in Part 2, a prior medication is defined as any medication taken prior to the date of randomization. A concomitant medication is defined as any medication that started prior to the date of randomization and continued to be taken after randomization or any medication that started on or after the date of randomization, but not after the end of Cycle 24, disease relapse, or study discontinuation, whichever is the earliest.

The number and percentage of subjects taking medications will be summarized by generic medication name for both prior and concomitant medications. A subject who reports the use of two or more medications will be counted only once in the summary of any prior or 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.

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 to the contrary.

8.0 Efficacy Analyses

8.1 General Considerations

The data cutoff date for the primary analysis of RFS will be the date when the [redacted] as assessed by IRC is observed in the ITT population. When the [redacted] in the ITT population 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 clinical database will be extracted for documentation and statistical analyses of the efficacy and safety. Unless otherwise specified, data after the cutoff date will be excluded from statistical analyses.

Unless otherwise specified, the primary analyses of all relapse related endpoints (e.g., RFS) will be based on IRC assessment. Sensitivity analyses will be performed based on the investigator assessment.

In order to perform a region-specific subgroup analysis to support regional filing, region will be used as a stratification factor in the randomization. However, region will not be included in the stratified efficacy analysis since it is not considered as a prognostic factor.

Unless otherwise specified, age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others) will be used in all stratified efficacy analyses. The stratification factor values under which the subject is randomized by the IVRS/IWRS will be used in the efficacy analyses.

Unless otherwise specified, all efficacy analyses will be conducted in the ITT population as defined in Section 4.0 and all tests will be 1-sided at an overall alpha level of 0.025 (when rounded to four decimal places).

8.2 Handling of Missing Data

Details on the handling of missing data are described in the analysis section corresponding to each endpoint.

8.3 Primary Efficacy Endpoint(s) and Analyses

8.3.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoint is IRC-assessed relapse-free survival (RFS). RFS is defined as the time from randomization to death from any cause or disease relapse, whichever occurs first. Criteria for disease relapse are described in the IRC charter. All RFS events included in the analysis will be based on IRC adjudication. The primary analysis of RFS will occur when the [redacted] is observed in the ITT population. All subjects in the ITT population will be included in the RFS analysis.

8.3.2 Handling of Missing Data for the Primary Efficacy Endpoint(s)

RFS is a time to event endpoint. The detailed event definition and censoring rule for RFS is described in Table 1.

Table 1. Event/Censor and Corresponding Event/Censor Time for RFS

Situation	Event/Censor	Event Time/Censor Time
Relapse or Death	Event	Earliest date on or prior to the data cutoff date
Still in remission	Censor	Last adequate disease assessment date (bone marrow or complete blood count assessment date) on or prior to the data cutoff date

Subjects without any specified RFS event will be censored at the date of last disease assessment on or prior to the cutoff date. The date of last adequate assessment will be determined by the date of last bone marrow assessment or complete blood count (CBC) assessment (white blood cell count, platelet count, hemoglobin, blasts).

Subjects without any disease assessment after randomization will be censored at the date of randomization.

8.3.3 Primary Efficacy Analysis

The distribution of RFS will be estimated for each study arm using Kaplan-Meier methodology and compared between the two study arms using the log-rank test, stratified

by age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others). The hazard ratio and corresponding 95% CI will be estimated using the Cox proportional hazards model, stratified by age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others).

8.3.4 Additional Analyses of the Primary Efficacy Endpoint(s)

The following sensitivity analyses of RFS will be performed:

- Stratified log-rank test and stratified Cox proportional hazards model based on RFS data from investigator assessment.
- Stratified Wilcoxon test, Un-stratified log-rank test, and the Cox proportional hazards model for RFS testing.
- Stratified log-rank test and covariate-adjusted proportional hazards Cox model will include potential prognostic factors not used in the randomization stratification.
- Modified primary RFS definition with censoring subjects who received other non-protocol treatment for AML prior to relapse.

8.4 Secondary Efficacy Analyses

8.4.1 Key Secondary Efficacy Analyses

The key secondary endpoints include overall survival (OS), time to deterioration in GHS/QoL, PROMIS Cancer Fatigue SF 7a and MRD conversion rate. All subjects in the ITT population will be included in the secondary efficacy analyses.

Overall Survival (OS)

Overall survival is defined as the time from randomization to death from any cause. All events of death that occur prior to the cutoff date will be included in the analysis of OS. All subjects in the ITT population will be included in the analysis of OS.

Subjects who have not died will be censored at the last known alive date on or before the cutoff date. The last known alive date will be determined by selecting the last available date of the following study procedures: start date of adverse event, bone marrow collection, disease assessment, vital signs assessment, clinical laboratory collection, study drug administration, start date of concomitant medicine, survival follow-up, biospecimen sample collection, quality of life assessments, and performance status.

The distribution of endpoint OS will be estimated for each study arm using Kaplan-Meier methodology and compared between the two study arms using the log-rank test, stratified by age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others). The hazard ratio and corresponding 95% CI between the two study arms will be estimated using the Cox proportional hazards model, stratified by age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others).

Time to Deterioration in GHS/QoL

The EORTC-QLQ-C30 is a 30-item patient-reported questionnaire composed of both multi item and single scales including 5 functional scales (physical, role, emotional, social, and cognitive), 3 symptom scales (fatigue, nausea and vomiting, and pain), a global health status/QoL scale, and 6 single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties). Subjects rate items on a 4-point scale ranging from 1 to 4 (1 = Not at All, 2 = A Little, 3 = Quite a Bit, and 4 = Very Much). The EORTC-QLQ-C30 was developed and validated for use in a cancer patient population, and its reliability and validity are highly consistent across different language-cultural groups. A change of 5 to 10 points is considered a small change and the lower bound (5) will be used to define the minimum important difference. A change of ≥ 10 to < 20 points is considered a moderate change.

Time to deterioration in GHS/QoL is defined as the time from randomization to death from any cause or the first-time deterioration of ≥ 5 points in GHS/QoL score as measured by EORTC QLQ-C30. All subjects in the ITT population will be included in the analysis.

Subjects without any of the specified events will be censored at the date of the last GHS/QoL assessment. Subjects without any GHS/QoL assessment after the randomization will be censored at the date of randomization.

The distribution of endpoint time to deterioration in GHS/QoL will be estimated for each study arm using Kaplan-Meier methodology and compared between the two study arms using the log-rank test, stratified by age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive (≥ 10^{-3}), others). The hazard ratio and corresponding 95% CI between the two study arms will be estimated using the Cox proportional hazards model, stratified by age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive (≥ 10^{-3}), others).

PROMIS Cancer Fatigue SF 7a

PROMIS is a system of highly reliable, precise measures of patient-reported health status for physical, mental, and social well-being. PROMIS instruments measure concepts such as pain, fatigue, physical function, depression, anxiety and social function. Fatigue will be assessed using the PROMIS Fatigue SF that has been developed for use in oncology populations. PROMIS Fatigue SF 7a is a 7-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. For Question 7, scores will be reversed. The recommended minimum important difference range is 3 to 5 points.

Patient reported outcome on fatigue will be assessed using the global fatigue score as measured by PROMIS Cancer Fatigue SF 7a. Scores will be computed according to the PROMIS Cancer Fatigue SF 7a scoring manual.

A linear mixed effects regression model with an appropriate covariance structure will be used to fit the longitudinal data. In the linear mixed effects regression model, stratification factors age (< 60, ≥ 60 years) and relapse risk (poor risk cytogenetic or MRD-positive (≥ 10^{-3}), others), will be included as fixed factors. The model will also

include time, study arm, and study arm by time interaction. The following covariance structures will be explored: Unstructured (TYPE = UN), compound symmetry (TYPE = CS) and first-order autoregressive (TYPE = AR(1)). The covariance structure resulting in model convergence and the lowest Bayesian Information Criterion (BIC) will be used for analysis. Change from baseline scores from all post-baseline visits described in Appendix B will be compared between the two study arms based on the linear mixed effects model. Subjects without baseline score or any post-baseline score will be excluded from the linear mixed effects regression analysis.

MRD Conversion Rate

MRD conversion rate is defined as percentage of subjects with MRD $\geq 10^{-3}$ at randomization who convert to MRD $< 10^{-3}$ after initiation of therapy. Subjects whose MRD is $< 10^{-3}$ or missing at randomization will be excluded from the analysis. Subjects who are deemed MRD positive at baseline but do not have any post-baseline MRD result will be included in the analysis as non-conversion events.

The MRD conversion rates between the two study arms will be compared using Miettinen-Nurminen method stratified by age group (< 60, ≥ 60 years). The difference and corresponding 95% CI in MRD conversion rate between the two study arms will also be provided using Miettinen-Nurminen method stratified by age group (< 60, ≥ 60 years). In addition, the 95% confidence interval of MRD conversion rate in each study arm will also be provided using the binomial distribution (Clopper-Pearson exact method).

8.4.2 Supportive Secondary Efficacy Analyses

OS will be analyzed using all data collected in the extracted database. The same analysis method described in Section 8.4.1 will be applied.

8.5 Additional Efficacy Analyses

All subjects in the ITT population will be included in the additional efficacy analyses.

EORTC QLQ-C30

Exploratory analyses will be performed on the subscales/items from the EORTC QLQ-C30. Scores will be calculated as per the scoring manuals. Linear mixed effects regression models similar to that described for the analysis of PROMIS Cancer Fatigue SF 7a in Section 8.4.1 will be used to compare between the two study arms. Subjects without baseline score or any post-baseline score will be excluded from the linear mixed effects regression analysis.

EQ-5D-5L

The EQ-5D-5L is a generic preference instrument that has been validated in numerous populations.[6,7] The EQ-5D-5L has 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These dimensions are measured on a 5-point 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 to 1 representing the general health status of the individual. The EQ-5D-5L status will be converted into a single preference-weighted health utility index score by applying country-specific weights (if available) or US weights (if not available). The EQ-5D-5L also contains a visual analog scale (VAS) to assess the subject's overall health. The minimum important difference for the EQ-5D-5L utility index score in cancer patients is 0.08, and the minimum important difference for EQ-5D-5L VAS is 7.[7]

Exploratory analyses will be performed on EQ-5D-5L. Scores will be calculated as per the scoring manuals. Linear mixed effects regression models similar to that described for the analysis of PROMIS Cancer Fatigue SF 7a in Section 8.4.1 will be used to compare between the two study arms. Subjects without baseline score or any post-baseline score will be excluded from the linear mixed effects regression analysis.

Health Care Resource Utilization

Number and percentage of subjects admitted to hospital during evaluation period will be summarized and compared between the two study arms using Fisher's exact test.

Percent of hospitalization stay over evaluation period will be summarized and compared between the two study arms using ANOVA model. The percent of hospitalization stay is defined as total duration of hospitalization divided by duration of evaluation period for each subject. In addition, reasons for hospital admission and treatment setting (e.g., inpatient unit, special care unit, etc.) will also be summarized.

The evaluation period is from the date of the first dose of study drug or date of randomization (randomized but not received study drug) to 30 days of the end of Cycle 24, disease relapse, study discontinuation or death, whichever is earlier.

8.6 Efficacy Subgroup Analyses

To evaluate of the impact of demographics and baseline characteristics on efficacy, subgroup analyses on efficacy endpoints including, but not limited to, RFS and OS will be performed for subgroups including, but not limited to, those defined below:

- Age (< 60; ≥ 60)
- Gender (Male, Female)
- Previous treatment with venetoclax (Yes, No)
- Previous treatment with azacitidine (Yes, No)
- MRD prior to randomization ($\geq 10^{-3}$, $< 10^{-3}$, Unknown)
- Previous treatment with venetoclax or azacitidine (Yes, No)
- Previous treatment with venetoclax (Yes, No)
- Previous treatment with azacitidine (Yes, No)
- NCCN cytogenetic risk (Intermediate, Poor)
- Best response from conventional chemotherapy prior to randomization (CR, CRi)
- ECOG performance status (0, 1, 2)
- Molecular marker ([redacted])
- Hepatic impairment (Yes, No)
- Renal impairment (Yes, No)

- Stratification factor: relapse risk (poor risk cytogenetic or MRD-positive ($\geq 10^{-3}$), others)
- Bone marrow blast count at diagnosis (< 30%, ≥ 30% – < 50%, ≥ 50%)
- region (Stratification regions ([redacted]), or other regions requested by regulatory agencies)
- CYP3A inhibitors use leading to venetoclax dose reduction (Yes, No)
- Post-study treatment (Yes, No)

9.0 Safety Analyses

9.1 General Considerations

Safety analyses will only include subjects in the Safety Analysis Set as defined in Section 4.0. Safety summaries will be presented for Part 1 and Part 2 separately and will present for overall, for each study arm in Part 2, and for each dose level in Part 1. Unless otherwise specified, data after the cutoff date will be excluded from statistical analyses.

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

9.2 Treatment-Emergent Adverse Events

Treatment-emergent AEs are defined as any AE with the onset that is on or after the first dose of study drug until 30 days after the last dose of the study drug for subjects in Part 1 and ARM A of Part 2. 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 subjects randomized to Arm B, AEs with an onset date on or after the date of randomization up to 30 days after the end of Cycle 24, disease relapse or study discontinuation, whichever is the earliest will be included in the summaries of treatment-emergent AEs.

All treatment-emergent AEs will be summarized overall, for venetocalx combination with azacitidine as well as for venetoclax monotherapy separately.

9.2.1 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any treatment-emergent AE
- Any treatment-emergent AE with NCI toxicity (CTCAE V5.0) grade ≥ 3
- Any treatment-emergent AE with NCI toxicity (CTCAE V5.0) grade 3 or 4
- Any treatment-emergent AE with reasonable possibility related to venetoclax as assessed by the investigator (for Arm A and Part 1)
- Any treatment-emergent AE with reasonable possibility related to azacitidine as assessed by the investigator (for Arm A and Part 1)
- Any treatment-emergent serious AE
- Any treatment-emergent AE leading to discontinuation of venetoclax (for Arm A and Part 1)
- Any treatment-emergent AE leading to dose interruption of venetoclax (for Arm A and Part 1)
- Any treatment-emergent AE leading to dose reduction of venetoclax (for Arm A and Part 1)
- Any treatment-emergent AE leading to discontinuation of azacitidine (for Arm A and Part 1)
- Any treatment-emergent AE leading to dose interruption of azacitidine (for Arm A and Part 1)

- Any treatment-emergent AE leading to dose reduction of azacitidine (for Arm A and Part 1)
- Any treatment-emergent AE leading to death
- All deaths

9.2.2 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported. Treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the total active group.

9.2.3 SAEs and Deaths

SAEs (including deaths) will be summarized by SOC and PT and in listing format.

The number of subject deaths and causes of death will be summarized

(1) all deaths from all causes.

(2) deaths occurring within 30 days of the first dose of study drug (Arm A or Part 1).

(3) deaths occurring within 30 days of the last dose of study drug (Arm A or Part 1).

(4) deaths occurring within 30 days of the randomization (Arm B).

(5) deaths occurring within 30 days of the end of Cycle 24, disease relapse or study discontinuation, whichever is the earliest (Arm B).

9.2.4 Adverse Events of Special Interest

Adverse events of special interest (AESI) will be summarized by SOC and PT and will be based on standardized or company MedDRA queries (SMQs or CMQs). AESI is Tumor Lysis Syndrome (TLS) and will be identified using the following search criteria:

Table 2. AESI Search Criteria

Area of Safety Interest	Search Criteria
Tumor Lysis Syndrome (AE)	SMQ – "Tumor Lysis Syndrome" (narrow)

Number and percent of subjects experienced treatment-emergent TLS occurring within 7 days from first dose for Arm A (or randomized date for Arm B) will be summarized.

In addition, treatment-emergent AEs and serious AEs for the following selected grouped preferred terms (PTs) including, but not limited to, those listed in Table 3, will be summarized.

Table 3. Selected Adverse Events

Selected Adverse Events	Search Criteria
grade ≥ 3 Neutropenia	PT terms – "neutropenia," "neutrophil count decreased," "febrile neutropenia," "agranulocytosis," "neutropenic infection," and "neutropenic sepsis"
grade ≥ 3 Infection, including opportunistic infection	SOC of "infections and infestations"
Haemorrhages	SMQ – "Haemorrhages" (narrow)
Anaemia	PT terms – "Anaemia" and "Haemoglobin decreased"
Thrombocytopenia	PT terms – "Thrombocytopenia" and "Platelet count decreased"
Leukopenia Neutropenia and neutrophil count decrease	PT terms – "Leukopenia" and "White blood cell count decreased" PT teams - "neutropenia" and "neutrophil count decreased"

9.3 Analysis of Laboratory Data

Unless otherwise specified, analysis of laboratory data will include post-baseline measurements within 30 days of the last dose of the study drug for Arm A or Part 1 or within 30 days of the end of Cycle 24, disease relapse, or study discontinuation whichever is the earliest for subjects in Arm B.

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized. Change from baseline will be summarized and presented for each lab test at scheduled post-baseline visits defined in Appendix C. Baseline is defined as the last non-missing observation before the first administration of study drug or randomization if no study drug is given.

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 grade is defined as the grade of the last measurement collected on or prior to the first dose of study drug (or randomization for non-treated subjects). In cases where multiple values are collected on the same day, the maximum grade value will be selected as the value for that day. The maximum NCI toxicity grade value is the value with highest NCI toxicity grade collected after the first dose of study drug.

For each lab test, cross tabulate will be generated for the number of subjects with baseline values of grade 0, grade 1, grade 2, grade 3, grade 4, or missing grade versus maximum post-baseline values of grade 0, grade 1, grade 2, grade 3, grade 4, or missing grade. All subjects in the safety analysis set 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. Shifts 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).

For above shift tables, baseline grade of 0 (normal) will be imputed for all subjects with at least one post-baseline but missing a baseline value for each lab test.

Detailed listings of all data for subjects experiencing NCI CTCAE grade 3 to 4 blood chemistry and hematology values will be provided.

DILI will be assessed using laboratory data, including alanine aminotransferase (ALT), aspartate aminotransferase (AST) and total bilirubin (TBILI), based on Hy's law criteria. Number and percentage of subjects with liver enzyme value meeting the Hy's law 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.

9.4 Safety Subgroup Analyses

Safety analyses described in Section 9.2 (adverse events) may be summarized for the subgroups including, but not limited to, those defined below:

- Age (< 60, ≥ 60)
- Gender (Male, Female)
- region ([redacted], or other regions requested by regulatory agencies)
- Hepatic impairment at baseline (Yes, No)

- Renal impairment at baseline (Yes, No)

Definition of hepatic impairment at baseline and renal impairment at baseline are described in Appendix F.

10.0 Interim Analyses

Interim review of safety data from Part 2 of the study will be conducted at the following time points:

- The first analysis to review the safety data will occur when 20 subjects have been on study for 3 cycles in Part 2 of the study
- The subsequent reviews of safety data will occur approximately every 6 months after the first review of safety data.

Two planned interim analysis (IAs), one for futility and one for efficacy, for the RFS endpoint will be conducted:

- RFS Futility IA (IA1) at 40% RFS ([redacted]);
- RFS Efficacy IA (IA2) at 75% RFS ([redacted]).

The Lan-DeMets error spending function with O'Brien-Fleming boundary will be used to determine the futility and efficacy stopping boundaries for the RFS endpoint (Table 4).

Table 4. Stopping Boundaries for IAs and FA of RFS

Analysis Time	Number of RFS Events	Efficacy Stopping Boundaries (One-sided P-value/ Observed HR)	Futility Stopping Boundaries (One-sided P-value/ Observed HR)
RFS IA1	[redacted]	NA	≥ 0.621/≥ 1.066
RFS IA2	[redacted]	≤ 0.010/≤ 0.703	NA
RFS FA	[redacted]	≤ 0.022/≤ 0.770	NA

IA = interim analysis; FA = final analysis; HR = hazard ratio; RFS = relapse-free survival; NA = not applicable

At the 40% RFS futility IA1, the probability of declaring futility is 36% assuming a hazard ratio of 1 and is 2.3% [redacted]. At the 75% RFS efficacy IA2, the probability of declaring success is 70% [redacted] and is 85% [redacted].

10.1 Data Monitoring Committee

The interim analyses will be performed by an independent Statistical Data Analysis Center. An Independent Data Monitoring Committee (IDMC) composed of persons independent of AbbVie and with relevant expertise in their field will review the above interim analyses.

A separate IDMC Charter will be prepared outside of the protocol and will describe the roles and responsibilities of the IDMC members, frequency of data reviews, and relevant safety data to be assessed.

The IDMC will make a recommendation based on RFS and OS to either to stop for success or futility at each interim analysis or continue the study as planned. The planned stopping boundaries are described in Table 4 for RFS. The actual stopping boundaries at IA2 and FA of RFS will be derived using Lan-DeMets alpha spending function based on the observed number of RFS events in the extracted database. The actual stopping boundaries will be presented in the IDMC charter addendum. At RFS IA2, if the IDMC informs the Sponsor that the RFS data is statistically significant and in favor of the venetoclax arm, upon notification from the Sponsor's Internal Executive Review Committee and after consultation with regulatory authorities as required to ensure maturity of data for registration, the Sponsor will then prepare regulatory submissions globally based on the available data while the study will continue for 100% OS evaluation. If the results of the RFS analysis are not statistically significant in favor of the venetoclax arm, the study will continue to follow protocol-specified procedures until the final analysis.

Details of the IDMC review will be presented in the IDMC Charter.

11.0 Overall Type-I Error Control

The fixed sequence testing procedure will be performed with a significance level of 0.025 (one-sided) for the primary efficacy endpoint and key secondary efficacy endpoints sequentially. The ranking and alpha spending of the key secondary efficacy endpoints are described in Table 5. If the statistical test is not significant for the primary efficacy endpoint, then statistical significance will not to be declared for any of the key secondary endpoints. The Lan-DeMets alpha spending function with O'Brien Fleming boundary will be used at the efficacy interim analysis to ensure that the false positive rate for RFS is controlled at the one-sided 0.025 level.

The analysis of OS will be performed at the time of each RFS analysis and at the time of the final OS analysis. [redacted] will provide approximately 80% power to detect a statistically significant difference between the two study arms at one-sided alpha level of 0.025 using the log-rank test (which corresponds to an increase in median of OS from 24 months in the BSC arm to 34 months in the venetoclax arm).

Table 5. Testing Sequence and Alpha-Spending Boundaries (One-Sided P value) for the Primary and Key Secondary Endpoints

Testing Sequence	Endpoint	Timing of Analysis			
		RFS IA1	RFS IA2	RFS FA	OS FA
1	RFS	0.0001	As specified in Table 4	As specified in Table 4	NA
2	OS	0.0001	0.0001	OBF[b]	OBF[b]
3	Time to deterioration in QoL[a]	0.0001	0.025	NA	
4	PROMIS Cancer Fatigue SF 7a[a]	NA	0.025	NA	
5	MRD[a]	NA	0.025	NA	

RFS = relapse-free survival; FA = final analysis; HR = hazard ratio; IA = interim analysis; NA = not applicable; OS = overall survival

a. Will be tested one time using the data at 75% RFS interim (IA2).

b. OBF (O'Brien-Fleming boundary) will be calculated based on [redacted] (100% information).

12.0 Version History

Table 6. SAP Version History Summary

Version	Date	Summary
1.0 Draft	20 AUG 2019	Original draft version prepared for agencies' feedback
1.0	21 SEP 2020	Incorporated agencies' comments

13.0 References

1. Pollyea DA, Pratz KW, Jonas BA, et al. Venetoclax in combination with hypomethylating agents induces rapid, deep, and durable responses in patients with AML ineligible for intensive therapy. Abstract Presented at: 2018 ASH Annual Meeting & Exposition; 02 December 2018; San Diego, CA. Abstract 285.
2. Dombret H, Seymour JF, Butrym A, et al. International Phase 3 study of azacitidine vs conventional care regimens in older patients with newly diagnosed AML with > 30% blasts. Blood. 2015;126(3):291-9.
3. DiNardo CD, Pratz K, Pullarkat V, et al. Venetoclax combined with decitabine or azacitidine in treatment-naive, elderly patients with acute myeloid leukemia. Blood. 2019;133(1):7-17.
4. National Comprehensive Cancer Network (2016). Acute Myeloid Leukemia (version 2.2016). Available from: https://www.nccn.org/professionals/physician_gls/pdf/aml.pdf.
5. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials oncology. J Natl Cancer Inst. 1993;85(5):365-76.
6. Cella D, Yount S, Rothrock N, et al. The patient-reported outcomes measurement information system (PROMIS): progress of an NIH roadmap cooperative group during its first 2 years. Med Care. 2007;45 (5 Suppl 1):S3-11.

7. Yost KJ, Eton DT, Garcia SF, et al. Minimally important differences were estimated for six PROMIS Cancer scales in advanced-stage cancer patients. J Clin Epidemiol. 2011;64(5):507-16.
8. Oppe M, Devlin NJ, van Hout B, et al. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. Value Health. 2014;17(4):445-53.
9. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007;5:70.
10. Ramalingam SS, Kummar S, Sarantopoulos J, et al. Phase I study of vorinostat in patients with advanced solid tumors and hepatic dysfunction: a National Cancer Institute Organ Dysfunction Working Group study. J Clin Oncol. 2010;28(29):4507-12.
11. Ferri FF. Ferri's Best Test: A Practical Guide to Laboratory Medicine and Diagnostic Imaging. Philadelphia: Saunders, an imprint of Elsevier Inc.; 2015.
12. Vogel H, Maas J, Gebauer A, editors. Drug Discovery and Evaluation: Methods in Clinical Pharmacology. New York: Springer; 2011.
13. FDA guidance for industry Pharmacokinetics in Patients with Impaired Renal Function – Study Design, Data Analysis, and Impact on Dosing and Labeling March 2010

Appendix A. Protocol Deviations

Subjects who reported any protocol deviation will be summarized.

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

Scheduled Visit	Nominal Cycle Rx Day	Time Window (Cycle Rx Day Range)
Baseline	≤ 1	Last non-missing value on or prior to the first dose of study drug for Part 1 or Arm A in Part 2; Last non-missing value on or prior to randomization for Arm B in Part 2
Each Other Cycle from Cycle 3 Day 1	1	[–10, 10]

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

Appendix C. Time Windows for Analysis of Clinical Hematology and Chemistry Parameters

Scheduled Visit	Nominal Cycle Rx Day	Time Window (Cycle Rx Day Range)
Baseline	≤ 1	Last non-missing value on or prior to the first dose of study drug for Part 1 or Arm A in Part 2; Last non-missing value on or prior to randomization for Arm B in Part 2
Cycle 1 Day 8	8	[5, 11]
Cycle 1 Day 15	15	[12, 18]
Cycle 1 Day 22	22	[19, 25]
Cycle 2 Day 1	1	[–3, 3]
Cycle 2 Day 8	8	[5, 11]
Cycle 2 Day 15	15	[12, 18]
Cycle 2 Day 22	22	[19, 25]
Cycle 3 Day 1	1	[-3,10]
Cycle 4 onward Day 1	1	[–10, 10]

Note: Clinical hematology and chemistry samples will be collected at visits specified in the ops manual.

Appendix D. Conversion of leukocyte differential counts from percentage (%) value to absolute value

All efficacy and safety analyses will include leukocyte differential values in % and absolute values. The following conversion process will be performed, if sites only provided % values for leukocyte differential counts without absolute counts on the same collection date and time of the total leukocyte counts.

1. The conversion is only for leukocyte differential counts in % values which were collected on the same date and at same time as total leukocyte counts.
2. Leukocyte differential absolute value ($\times 10^9$/L) = Total leukocyte count ($\times 10^9$/L) × leukocyte differential value (%)/100
3. The following low normal ranges will be used for leukocyte differential counts in % value converted to absolute count values for the shift tables.

Lab Test: **Total Leukocytes and differential leukocyte counts**	**Normal Range (SI units)**
Total Leukocyte ($\times 10^9$/L)	4.5 - 11
Neutrophils (%)	40 - 70
Lymphocytes (%)	22 - 44

*Source: Kratz A et al. Laboratory Reference Values. N Engl J Med 2004; 351:1548-63.

Lower limit of normal of 1.8×10^9/L for neutrophil is calculated from total leukocyte count of 4.5×10^9/L, according to formula on item 2 above. Lower limit of normal of 0.99×10^9/L for lymphocytes is calculated from total leukocyte count of 4.5×10^9/L. Above the lower limit of normal range values are used to classify CTCAE grade for neutrophil and lymphocyte counts.

Appendix E. Data Cutoff Date

Data cutoff date for each statistical summary or test will be determined and applied.

Appendix F. Hepatic and Renal Function

Hepatic Function based on National Cancer Institute Organ Dysfunction Working Group Classification of Hepatic Dysfunction:[10-13]

	Bilirubin (mg/dL)	**AST (IU/L)**
Normal	≤ 1.0	≤ 40
Mild Impairment	≤ 1.0	> 40
Moderate Impairment	> 1.0 and ≤ 1.5	Any
	> 1.5 and ≤ 3.0	Any
Sever Impairment	> 3.0	Any

Renal Function based on FDA guidance for industry Pharmacokinetics in Patients with Impaired Renal Function – Study Design, Data Analysis, and Impact on Dosing and Labeling:

	CrCL (mL/min)
Normal	≥ 90
Mild Impairment	60 ≤ CrCL< 90
Moderate Impairment	30 ≤ CrCL< 60
Sever Impairment	CrCL < 30

If serum creatinine is in µmol/L, it can be converted to mL/min following:

$$CrCL = \frac{(140 - age) \times (weight\ in\ kg) \times (1.23\ if\ male, 1.04\ if\ female)}{serum\ creatinine\ in\ \mu mol/L}$$