STATISTICAL ANALYSIS PLAN

Version 3.0, dated 11 July 2022

A Phase 3 Multicenter, Open-label, Randomized Study of ASP2215 (Gilteritinib), Combination of ASP2215 Plus Azacitidine and Azacitidine Alone in the Treatment of Newly Diagnosed Acute Myeloid Leukemia with FLT3 Mutation in Patients Not Eligible for Intensive Induction Chemotherapy

ISN/Protocol: 2215-CL-0201

IND number: 117,548

Sponsor:

Astellas Pharma Global Development, Inc. (APGD)
1 Astellas Way
Northbrook, IL 60062

Table of Contents

I. LIST OF ABBREVIATIONS AND KEY TERMS

List of Abbreviations

Abbreviations	Description of abbreviations
AE	Adverse event
AESI	Adverse events of special safety interest
ALP	Alkaline Phosphatase
ALT	Alanine transaminase
AML	Acute myeloid leukemia
ANC	Absolute neutrophil count
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
APGD	Astellas Pharma Global Development
ASCM	Analysis Set Classification Meeting
ASP2215	Astellas Compound code for 2215
AST	Aspartate Transaminase
ATC	Anatomical Therapeutic Chemical
AUC_t	Area under the curve at t hours
BFI	Brief Fatigue Inventory
BMI	Body Mass Index
BSA	Body surface area
C_{max}	Maximum concentration
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019
CR	Complete remission
CR/CRh	Complete remission and complete remission with partial hematological recovery
CRc	Composite complete remission
CRh	Complete remission with partial hematologic recovery
CRi	Complete remission with incomplete hematologic recovery
CRp	Complete remission with incomplete platelet recovery
CS	Classification Specifications
CSR	Clinical Study Report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
C_{trough}	Concentration immediately prior to dosing at multiple dosing
CV	Coefficient of variation
DBP	Diastolic blood pressure
DLT	Dose limiting toxicity

Abbreviations	Description of abbreviations
DMC	Data Monitoring Committee
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDTA	Ethylenediaminetetraacetic acid
EFS	Event-free survival
EQ-5D-5L	EuroQol Group-5 dimension-5 level
FAB	French-American-British
FACIT-Dys-SF	Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms
FACT-Leu	Functional Assessment of Cancer Therapy-Leukemia
FAS	Full analysis set
FLT3	FMS-like tyrosine kinase
FSI	First subject in
GMR	Geometric mean ratio
H	High
HAR	High Allelic Ratio
HRQoL	Health-Related Quality of Life
HSCT	Hematopoietic stem cell transplant
ICF	Informed consent form
ICH	International Conference on Harmonization
IDAC	Independent data analysis center
IRT	Interactive Response Technology
ISN	International Study Number
ITD	Internal tandem duplication
IU/L	International units/liter
L	Low
LAR	Low Allelic Ratio
LFS	Leukemia-free survival
LLOQ	Lower limit of quantification
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligrams
Min	Minute
mL	Milliliter
mmHg	millimeters of mercury
MRD	Minimal residual disease
msec	milliseconds

Abbreviations	Description of abbreviations
MUGA	Multigated acquisition scan
N	Number
NCI	National Cancer Institute
NE	Not Evaluable
NR	Non-Response
OS	Overall Survival
PD	Protocol Deviation
PGx	Pharmacogenomics
PK	Pharmacokinetic
PKAS	Pharmacokinetic analysis set
PR	Partial remission
PT	Preferred Term
QD	quaque die, a Latin phrase meaning “every day”
QTc	QT interval corrected for heart rate
QTcF	Fridericia-corrected QT interval
RBC	Red blood cell
RR	Interval between 2 consecutive r waves on an ECG
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical analysis plan
SAS	Statistical Analysis System
SBP	Systolic blood pressure
SOC	System Organ Class
TBD	To be determined
TEAE	Treatment Emergent Adverse Event
TF	Treatment Failure
TLF	Tables, Listings and Figures
TKD	Tyrosine kinase domain
t_{max}	Time of maximum concentration
ULN	Upper limit of normal
US	United State
VAS	Visual analogue scale
WHO	World Health Organization
WHO-DD	World Health Organization – Drug Dictionary

List of Key Terms

Terms	Definition of terms
Baseline	Observed values/findings which are regarded as the observed starting point(s) for comparison.
Enroll	To register or enter into a clinical trial. NOTE: Once a subject has been enrolled, the clinical trial protocol applies to the subject.
Intervention	The drug, therapy or process under investigation in a clinical study that is believed to have an effect on outcomes of interest in a study (e.g., health-related quality of life, efficacy, safety, pharmacoeconomics).
Investigational period	Period of time where major interests of protocol objectives are observed and where the test drug or comparative drug (sometimes without randomization) is usually given to a subject and continues until the last assessment after completing administration of the test drug or comparative drug.
Post investigational period	Period of time after the last assessment of the protocol. Follow-up observations for sustained adverse events and/or survival are done in this period.
Screening period	Period of time before entering the investigational period, usually from the time of starting a subject signing consent until just before the test drug or comparative drug (sometimes without randomization) is given to a subject.
Randomization	The process of assigning trial subjects to treatment or control groups using an element of chance to determine assignments in order to reduce bias.
Screening	A process of active consideration of potential subjects for enrollment in a trial.
Screen failure	Potential subject who did not meet 1 or more criteria required for participation in a trial.
Screening period	Period of time before entering the investigational period, usually from the time of starting a subject signing consent until just before the test drug or comparative drug (sometimes without randomization) is given to a subject.
Study period	Period of time from the first site initiation date to the last site completing the study.
Variable	Any quantity that varies; any attribute, phenomenon or event that can have different qualitative or quantitative values.

1 INTRODUCTION

This Statistical Analysis Plan (SAP) contains a more technical and detailed elaboration of the principal features of the analysis described in the protocol and includes detailed procedures for executing the statistical analysis to fulfil the objectives of the study.

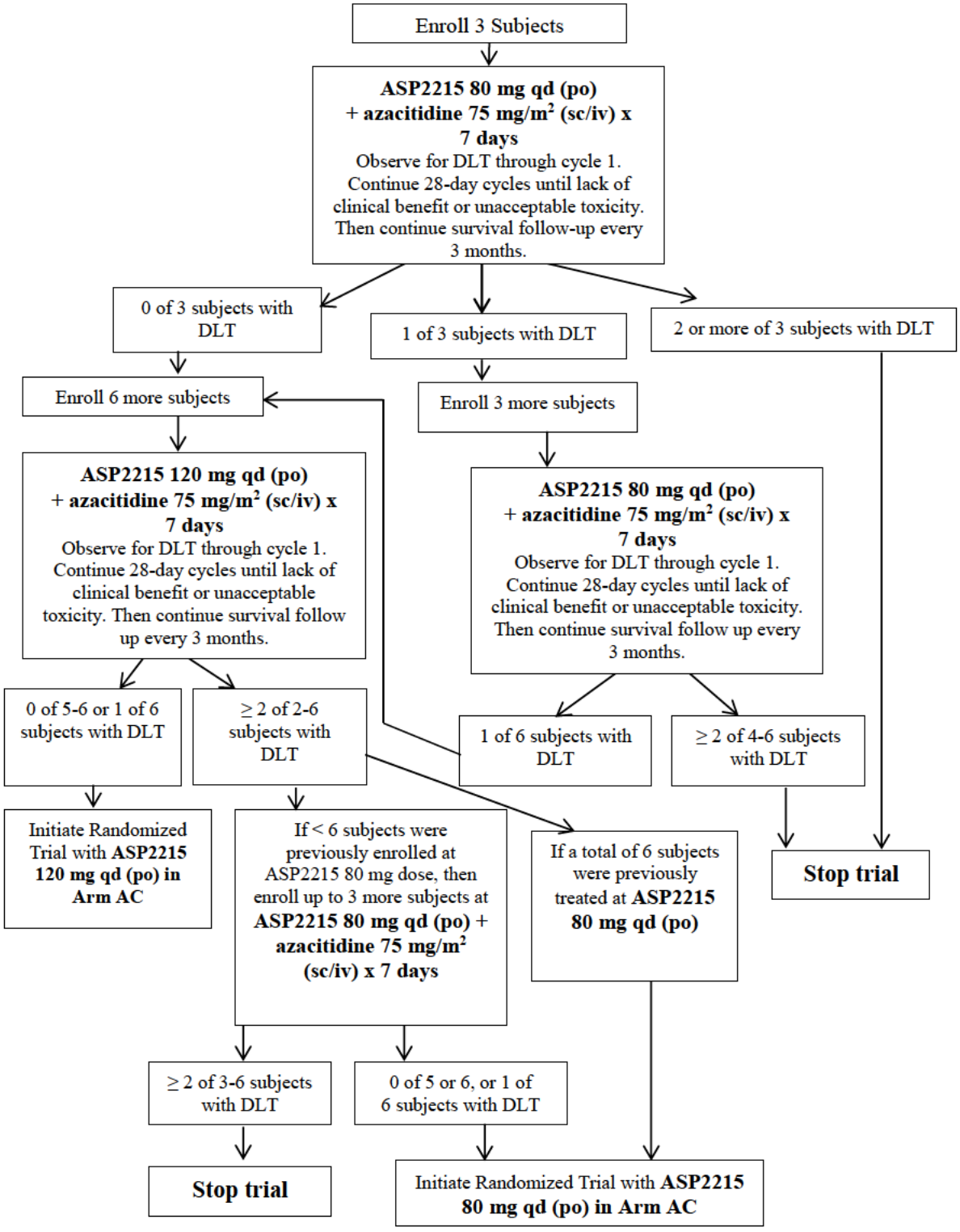
The SAP will be finalized and approved prior to any of the following: study unblinding, database hard lock, interim analysis, or accumulation of substantial amount of data in an open-label study to ensure lack of bias. If needed, revisions to the approved SAP may be made prior to database hard lock. Revisions will be version controlled.

Any changes from the analyses planned in the final SAP will be documented in the Clinical Study Report (CSR).

Analysis set classification will be finalized prior to study database lock.

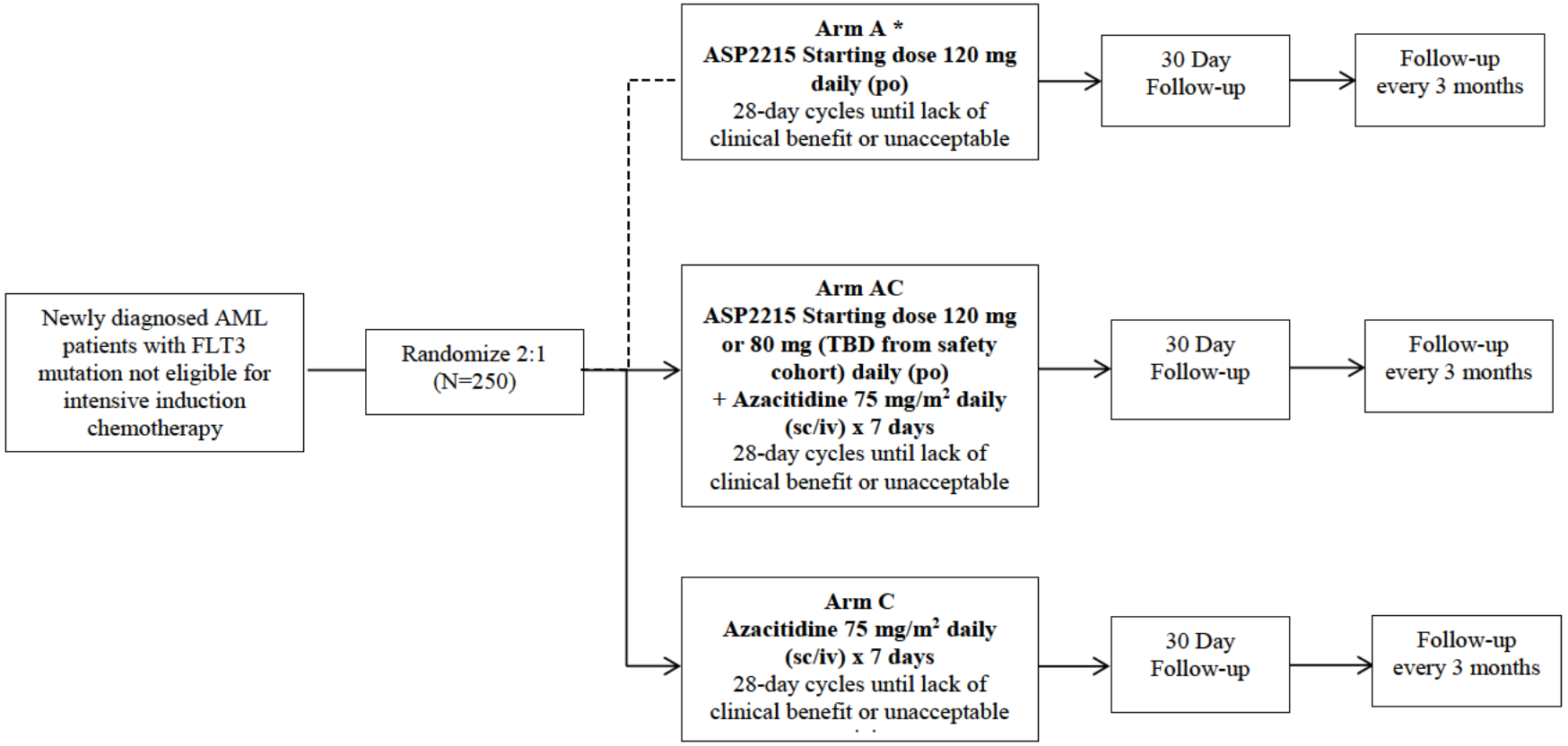
2 FLOW CHART AND VISIT SCHEDULE

Flow Chart 1: Safety Cohort



DLT: dose limiting toxicity; po: oral administration; qd: every day.

Flow Chart 2: Subject Flow in Randomized Trial



AML: acute myeloid leukemia; FLT3: FMS-like tyrosine kinase; TBD: to be determined; po: oral administration; qd: every day; sc: subcutaneous injection; iv: intravenous infusion.

* Protocol versions 6.0 and earlier included a 1:1:1 randomization ratio to receive Arm A (Gilteritinib monotherapy), AC or C. Randomization to Arm A was removed in protocol version 7.0. Subjects previously randomized to Arm A should continue following treatment and assessments as outlined in the protocol.

Table 1 Schedule of Assessments

Assessments	Screening (Days -14 to -1)[w]	Cycle 1 [z]					Cycle 2 [z]		Subsequent Cycles [z]
		D 1	D 4 ±1	D 8 ±1	D 9	D 15	D 1 ±1	D 15 ±1	D 1 ±2
Signed ICF	X								
Medical and disease history	X								
Randomization [v]		X[w]							
Physical examination [a, b]	X	X	X	X		X	X	X	X
Vital signs [b]	X	X	X	X		X	X	X	X
ECOG performance [b]	X	X [b]				X	X	X	X
Prior and concomitant medications	X [c]	X	X	X		X	X	X	X
Pregnancy test for WOCBP [d]	X	X					X		X
Chest X-ray (or CT of chest)	X[x]								
12-lead ECG - All Subjects [b, e]	X	X		X [f]	X [f]	X	X		X
12-lead ECG - Dense PK Sampling subset – additional time points			X [g]			X [g]			
Clinical laboratory tests (serum chemistry, hematology, coagulation, urinalysis) [b, h]	X [i]	X	X	X		X	X	X	X
Thyroid Function Test	X								X [b, u]
Coagulation profile (PT/INR, d-dimer, fibrinogen)	X								
MUGA or ECHO [j]	X								
FLT3 mutation status (bone marrow or whole blood)	X[k]								
Bone marrow aspiration and/or biopsy for disease assessment and MRD analysis (or whole blood)	X[k]						X[l]		X[l]
AE/SAE assessment [m]	X	X	X	X		X	X	X	X
Pharmacokinetic sampling – Arms A* and AC (whole blood samples for plasma pharmacokinetics) [n]		X		X		X	X		X
Pharmacokinetic sampling – Dense PK Sampling subset – all arms (whole blood samples for plasma pharmacokinetics) – additional time points [o]			X			X			
PGx (whole blood and buccal swab) [p]		X							
Patient reported outcome tools - Brief Fatigue Inventory [q]		X		X		X	X	X	X
Patient reported outcome tools – all others [q, r]		X					X		X

Table continued on next page

Assessments	Screening (Days -14 to -1)[w]	Cycle 1 [z]					Cycle 2 [z]		Subsequent Cycles [z]
		D 1	D 4 ±1	D 8 ±1	D 9	D 15	D 1 ±1	D 15 ±1	D 1 ±2
Resource utilization		X					X		X
IRT Transaction [y]	X	X					X		X
ASP2215 dosing at the clinic [s]		X		X		X	X	X	X
Azacitidine dosing at the clinic [t]		X	X [o]				X		X

AE: adverse event; C: cycle; CR: complete remission; CRc: composite complete remission; CRi: complete remission with incomplete hematologic recovery; CRp: complete remission with incomplete platelet recovery; CT: computed tomography; D: day; ECG: electrocardiogram; ECHO: echocardiogram; ECOG: Eastern Cooperative Oncology Group; FLT3: FMS-like tyrosine kinase; ICF: informed consent form; INR: international normalized ratio; IRT: interactive response technology; MRD: minimal residual disease; MUGA: multigated acquisition scan; PK: pharmacokinetic; PGx: pharmacogenomics; PT: prothrombin time; QTcF: Fridericia-corrected QT interval; SAE: serious adverse event; WOCBP: women of childbearing potential.

a. Height measurement performed only at screening. Weight measurement should be performed at screening and on day 1 of each cycle.
b. Obtained predose of study drug (ASP2215 and/or azacitidine).
c. Includes medications taken within 28 days prior to C1D1.
d. WOCBP must have a negative serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of human chorionic gonadotropin) within 72 hours prior to the start of study treatment. If screening pregnancy test was performed within 72 hours prior to start of study treatment, then it does not need to be repeated at day 1.
e. Screening ECG is required. ECG assessment will be evaluated before dosing on C1D1, C1D8, C1D15 and day 1 of each subsequent cycle. Predose assessments should be taken within 1 hour before study treatment (ASP2215 and/or azacitidine) administration. The 12-lead ECGs will be recorded in triplicate (3 separate ECGs with 10 minutes resting prior to first ECG and at least 5 minutes apart per time point) and transmitted electronically for central reading. The mean QTcF of the triplicate ECG tracings based on central reading will be used for all treatment decisions. If the mean triplicate QTcF is > 500 ms at any time point, the ECG will be repeated (within 2 hours if identified on machine read or as soon as possible if identified by central read). See [Section 5.4.5]. ECGs can be repeated during screening period.
f. If the mean QTcF from C1D1 to C1D8 has increased > 30 ms with no other known etiology, a confirmatory ECG should be performed on C1D9. If the C1D8 and C1D9 ECGs confirm the > 30 ms increase from C1D1 in QTcF, then the investigator should assess if ASP2215 dose modification should occur as per the dose interruption or reduction guideline in [Section 5.1.2]. On C1D1, it is recommended that the ECG is taken as early as possible in the morning and transmitted immediately. In addition, it is recommended that the C1D9 visit is scheduled later in the day in order to allow for receipt and assessment of the C1D8 central read ECG. This also allows for a subject to be contacted if the C1D9 ECG is no longer required.
g. For subjects participating in the dense PK sampling subset only, additional ECGs will be performed on C1D4 and C1D15. The 12-lead ECGs will be recorded in triplicate (3 separate ECGs with 10 minutes resting prior to first ECG and at least 5 minutes apart per time point) and transmitted electronically for central reading. Triplicate ECGs are to be performed prior to obtaining the time-matched PK sample, therefore must be started at least 10 to 15 minutes before the PK draw.:
   See Table 3.
h. Uric acid will be tested on days 1, 4, 8 and 15 in cycle 1. Urinalysis is only required at screening. Additional laboratory tests may be performed according to institutional standard of care.

Footnotes continued on next page

i. Subjects may be screened and randomized from local labs. However, samples must also be submitted for central read. Labs can be repeated during screening period.
j. MUGA scans or ECHO (per standard of care) are to be performed at screening for subjects with history of New York Heart Association Class 3 heart failure. NOTE: MUGA scans are not applicable to Germany.
k. FLT3 mutation status must be determined from central read results. At screening, 2 bone marrow aspirate samples are required: one will be sent to Invivoscribe (central FLT3 mutation testing laboratory) and the other to Hematogenix (central disease assessment laboratory). A bone marrow aspirate is preferred for FLT3 assessment. However, if a bone marrow aspirate sample is unavailable at screening -
    - a whole blood sample can be sent to Invivoscribe for FLT3 testing, provided there are measureable leukemic cells present and
    - the bone marrow biopsy from initial diagnosis and a whole blood sample should be sent to Hematogenix for disease assessment.

    Aspirate from initial diagnosis can be sent to Invivoscribe if it was collected in a sodium heparin tube, stored at 2 to 8°C and can be sent within 5 days of collection and testing can occur within 7 days of sample collection. Subjects in the safety cohort only should still send a FLT3 sample to the central laboratory; however, the results are not required to enroll.
l. Subsequent bone marrow samples are required during, C2D1 and C3D1. For subjects who do not achieve a CR, CRp or CRi, the bone marrow assessments will be repeated on day 1 of every 2 subsequent cycles. For subjects who achieve a CRc (CR, CRp or CRi), bone marrow will be repeated 1 month after the date of remission and every 3 subsequent cycles or if there is suspicion of relapse in the whole blood. Bone marrow samples are also required at the end-of-treatment visit and as clinically indicated. If bone marrow aspirate is unavailable then an EDTA tube of whole blood along with bone marrow core biopsy (block or slides) should be collected instead. Post screening bone marrow samples only need to be sent to Hematogenix and remaining bone marrow aspirate and/or whole blood samples will be used for MRD analysis and other biomarker analyses.
m. For Safety Cohort and First 6 Japanese Subjects on Arm AC (either in safety cohort or randomization portion of the study): DLT observation period will occur from day 1 through day 28 of cycle 1 only.
n. For subjects in randomized portion of the trial only: PK samples for ASP2215 will be collected for all subjects in Arm A* and Arm AC at predose (within 1 hour before ASP2215 administration) on C1D1, C1D8, C1D15 and on day 1 of each subsequent cycle. See Table 3.
o. Dense PK is applicable to randomization portion of the trial only, and will include the first 12 non-Japanese and 12 Japanese subjects randomized at the sites participating in the dense PK sampling subset. For subjects participating in dense PK sampling subset only, additional PK samples will be collected on C1D4 for ASP2215 (Arms A* and Arm AC) and/or azacitidine (Arm AC and Arm C) at predose (within 1 hour before ASP2215 for Arm A and AC and within 1 hour before azacitidine administration for Arm C), and at 0.25 (± 5 min), 0.5 (± 5 min), 1 (± 10 min), 2 (± 10 min), 4 (± 20 min), and 6 (± 20 min) hours post dose. Subjects on Arms AC and C should receive azacitidine via subcutaneous injection only for cycle 1. Intravenous infusion will be allowed after cycle 1. For subjects on Arm AC who will have 2 tubes collected per time point, the ASP2215 PK sample will be drawn before the azacitidine sample. On day 4, azacitidine will be administered immediately following ASP2215 administration. Dense PK samples will also be collected for ASP2215 for subjects on arms A* and AC on C1D15 at 4 hours (± 20 min) post dose. See Table 3.
p. Whole blood and buccal swab collected at cycle 1 day 1 predose for optional PGx study.
q. If possible, patient reported outcome measures should be performed prior to any other assessments on that visit day.
r. Includes EuroQol Group 5-dimension 5-level instrument, Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Form, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items.
s. ASP2215 is taken daily at home except for visits marked where it will be taken at the clinic.

Footnotes continued on next page

* Protocol versions 6.0 and earlier included a 1:1:1 randomization ratio to receive Arm A (ASP2215 monotherapy), AC or C. Randomization to Arm A was removed in protocol version 7.0. Subjects previously randomized to Arm A should continue following treatment and assessments as outlined in the protocol.

t. Azacitidine regimen (75 mg/m^2) to be administered daily by subcutaneous injection or intravenous infusion on days 1 to 7 of each 28-day cycle. Refer to the pharmacy manual for administration instructions. The route of administration of azacitidine outside of label instructions is not recommended and is a clinical decision of the investigator. On all visits indicated for azacitidine administration, the dosing should occur in clinic. For subjects in Arm AC, whenever possible, azacitidine will be administered immediately following ASP2215 administration. Subjects in the dense PK sampling subset on Arms AC and C should receive azacitidine daily by subcutaneous injection only for cycle 1.
u. Thyroid Function Tests is performed during screening visit, C3D1 and will be repeated after every 2 cycles of therapy thereafter (C5D1, C7D1, C9D1 etc.).
v. Japan only – for the first 6 subjects enrolled or randomized to the combination arm in Japan (either in safety cohort or randomization portion of the study), only 1 subject may begin study drug administration per day (i.e., no 2 subjects will receive their initial dose of study medication on the same day).
w. Safety eligibility assessments (clinical labs and ECG) can be repeated to ensure the patient meets eligibility during screening period. The screening period can be extended to repeat safety assessments with the approval of the Medical Monitor.
x. A chest X-ray (or CT of chest) does not need to be repeated if a chest x-ray result performed within 2 weeks prior to start of screening is available to assess subject eligibility.
y. Randomization in the IRT can occur a day prior to C1D1. For the purposes of drug preparation and dispensing activities during subsequent visits, IRT transaction may be done prior to the visit and do not need to fall within the protocol visit window.
z. Unscheduled visits may be performed at any time during the study whenever necessary to assess for AEs or follow-up on AEs, or if deemed necessary by the investigator. Unscheduled visits can include AE assessment, additional assessments (e.g., laboratory testing, ECG etc) as deemed appropriate by the investigator. If dose interruption of study treatment exceeds more than 28 days (exception HSCT), then safety assessments (clinical labs and ECG) will be repeated every 28 days as an unscheduled visit from the time of interruption of the study treatment to the resumption of the study treatment by the subject.

Table 2 Posttreatment Schedule of Assessments

Assessments	Pre-HSCT Visit/ EOT [a]	30-day Follow-up (+ 7 days) [b]	Long-term Follow-up [c] (+/- 7 days)
Physical examination	X [d]		
Vital signs	X [d]		
ECOG performance	X [d]		
Pregnancy test for WOCBP	X		
12-lead ECG	X		
Clinical laboratory tests (serum chemistry, hematology, coagulation)	X [d]		
Thyroid Function Test	X		
Bone marrow aspiration and/or biopsy for disease assessment and MRD analysis(or whole blood) [e]	X		
Concomitant medications [f]	X	X [f]	
AE/SAE assessment	X	X [g]	X [h]
Patient reported outcome tools – EQ-5D-5L [i]	X	X	X
Patient reported outcome tools – all others [i, j]	X		
Resource utilization	X	X	
Survival and subsequent antileukemic treatments and their outcomes		X	X
IRT Transaction	X		

AE: adverse event; ECG: electrocardiogram; ECOG: Eastern Cooperative Oncology Group; EDTA: ethylenediaminetetraacetic acid; EOT: end-of-treatment; EQ-5D-5L: EuroQol Group 5-dimension 5-level instrument; FLT3: FMS-like tyrosine kinase; HSCT: Hematopoietic stem cell transplant; IRT: interactive response technology; MRD: minimal residual disease; SAE: serious adverse event; WOCBP: women of childbearing potential.

a. End-of-treatment visit is to be performed within 7 days after last dose of study treatment (ASP2215 and/or azacitidine). For subjects who will undergo HSCT and plan to resume ASP2215 treatment after HSCT, a pre-HSCT visit will be performed. For subjects in Arm C who will undergo HSCT, an EOT visit will be performed followed by the Posttreatment Schedule of Assessments for long term follow-up.

b. Telephone contact with the subject is sufficient unless any assessment must be repeated for resolution of treatment-related AEs.

c. Telephone contact every 3 months for up to 3 years. Additional contacts will be made during interim analysis/final analysis or as needed.

d. Assessment does not need to be repeated if collected at a regularly scheduled visit within 3 days before the end-of-treatment visit.

e. Bone marrow aspiration and biopsy are not required if a sample was obtained within 14 days before end-of-treatment visit. If bone marrow aspirate is unavailable then an EDTA tube of whole blood along with bone marrow core biopsy (block or slides) should be collected instead. Post screening bone marrow samples only need to be sent to Hematogenix and remaining bone marrow aspirate and/or whole blood samples will be used for MRD analysis and other biomarker analyses.

f. Concomitant medications should be collected for reported or ongoing AE/SAEs through 30 days after the last dose of study treatment (ASP2215 and/or azacitidine) for subjects who have discontinued treatment. For subjects who undergo HSCT, concomitant medications should be collected for reported or ongoing AE/SAEs through start of conditioning treatment or 30 days after the last dose of study treatment (ASP2215 and/or azacitidine), whichever comes first. For subjects who resume ASP2215 treatment after HSCT, concomitant medications collection will resume upon the resumption of ASP2215 treatment and continue until 30 days after the last dose of study drug.

Footnotes continued on next page

g. For subjects who plan to proceed to HSCT, AE (included SAE) collection will continue until the start of the HSCT conditioning regimen or until 30 days after the last dose of study treatment (ASP2215 and/or azacitidine), whichever comes first. However, the following AE/SAEs will continue to be collected until 30 days after the last dose of study treatment (ASP2215 and/or azacitidine), regardless of the time of the HSCT conditioning regimen:
   - Any study drug related AE that is ongoing will be followed until resolved
   - Any SAE that is deemed to be related to study drug by the investigator
   - Any event of veno-occlusive disease (VOD) of the liver, cardiac failure, grade 3 or higher QT prolongation, rhabdomyolysis, drug-induced liver injury, or posterior reversible encephalopathy syndrome (PRES)
   - Adverse events leading to death

   For subjects who resume ASP2215 treatment after HSCT, AE (including all SAE) collection will resume upon the resumption of ASP2215 treatment and continue until 30 days after the last dose of study drug.

h. Only SAEs related to ASP2215 alone, azacitidine alone or ASP2215 + azacitidine will be collected.

i. If possible, patient reported outcome measures should be performed prior to any other assessments on that visit day.

j. Includes Brief Fatigue Inventory, Functional Assessment of Chronic Illness Therapy-Dyspnea-Short Form, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items.

Table 3 Sampling Time points for ECG, PK and Dense PK for Each Treatment Arm in Randomization Cohort

ASSESSMENT	CYCLE 1					CYCLE 2		SUBSEQUENT CYCLES
	D 1	D 4	D 8	D 9	D 15	D 1	D 15	D 1
ECG [a] [All Subjects]	Predose		Predose	Predose [only if C1D8 is ≥ 30 ms from baseline]	Predose	Predose	Predose	Predose
ECG [a, b] [Dense PK Subjects- Arm A & AC]		Predose Postdose: 4 hrs (-30 min)			Postdose: 4 hrs (-30 min)			
ECG [a, b] [Dense PK Subjects- Arm C]		Predose Postdose: 4 hrs (-30 min)						
PK Samples - ASP2215 [c] [Subjects in Arms A and AC]	Predose		Predose		Predose	Predose		Predose
Dense PK Samples - ASP2215 [c] [Subjects in Arms A and AC]		Predose *Postdose:* ▪ 0.25 hrs (± 5 min), ▪ 0.5 hrs (± 5 min), ▪ 1 hrs (± 10 min), ▪ 2 hrs (± 10 min), ▪ 4 hrs (± 20 min), ▪ 6 hrs (± 20 min)			*Postdose:* ▪ 4 hrs (± 20 min)			
Dense PK Samples – Azacitidine [c] [Subjects in Arms C and AC]		Predose *Postdose:* ▪ 0.25 hrs (± 5 min), ▪ 0.5 hrs (± 5 min), ▪ 1 hrs (± 10 min), ▪ 2 hrs (± 10 min), ▪ 4 hrs (± 20 min), ▪ 6 hrs (± 20 min)						

C: cycle; D: day; ECG: electrocardiogram; PK: pharmacokinetic.

a. Triplicate ECG predose assessments should be taken within 1 hour before study treatment (ASP2215 and/or azacitidine) administration and at least 10 to 15 minutes before the predose PK draw (includes dense PK).

b. Triplicate ECG post dose are to be performed prior to obtaining the time-matched PK sample, therefore must be started at least 10 to 15 minutes before the PK draw.

c. Predose PK samples will be collected after triplicate ECG and prior to administration of ASP2215 and/or azacitidine. Both predose triplicate ECG and predose PK sample should be collected within 1 hour prior study treatment (ASP2215 and/or azacitidine) administration.

3 STUDY OBJECTIVES AND DESIGN

3.1 Study Objectives

3.1.1 Primary Objective

The primary objective is to determine the efficacy superiority of ASP2215 plus azacitidine versus azacitidine as measured by overall survival (OS).

3.1.2 Secondary Objectives

3.1.2.1 Key Secondary Objectives

The key secondary objective is to determine the efficacy superiority of ASP2215 plus azacitidine versus azacitidine as measured by event-free survival (EFS).

3.1.2.2 Additional Secondary Objectives

Additional secondary objectives are to evaluate the safety and efficacy of ASP2215 plus azacitidine versus azacitidine in terms of:

- Best response
- Complete remission (CR) rate
- Complete remission with partial hematologic recovery (CRh) rate
- CR/CRh rate
- Composite complete remission (CRc) rate
- Duration of remission
- Leukemia-free survival (LFS)
- Transfusion conversion rate; transfusion maintenance rate
- Patient reported fatigue (Brief Fatigue Inventory [BFI])
- Adverse events (AEs), clinical laboratory results, physical examinations, vital signs, electrocardiograms (ECGs) and Eastern Cooperative Oncology Group (ECOG) performance scores

3.1.3 Exploratory Objectives

The exploratory objectives are to:

- Evaluate the efficacy of ASP2215 plus azacitidine versus azacitidine in terms of:
    - Transplantation rate
    - Minimal residual disease (MRD)
    - FLT3 gene mutation status
        - Mutation types and frequency
        - Relationship to efficacy and safety
        - Mechanisms of acquired resistance
    - Exploratory (predictive) biomarkers of ASP2215 activity
    - Patient reported dyspnea (Functional Assessment of Chronic Illness Therapy-Dyspnea-Short Form [FACIT-Dys-SF])

  - Patient reported signs, symptoms and impacts of AML (Functional Assessment of Cancer Therapy-Leukemia [FACT-Leu] and dizziness and mouth sore items)
  - Health-related quality of life assessed by the EuroQol Group 5-dimension 5-level (EQ-5D-5L) instrument
  - Resource utilization including hospitalization, blood transfusion, antibiotic intravenous infusions, medication for AEs and opioid usage
- Characterize the pharmacokinetics of ASP2215 and azacitidine given as single agents and/or as combination treatment
- Evaluate and compare the pharmacokinetics of ASP2215 and azacitidine in a subset of Japanese and Non-Japanese subjects

3.2 Study Design

This is a phase 3 multicenter, open-label, randomized study to compare the efficacy and safety of ASP2215 plus azacitidine versus azacitidine in newly diagnosed FLT3 mutated AML subjects not eligible for intensive induction chemotherapy.

3.2.1 Safety Cohort

Prior to initiation of the randomized trial, 8 to 12 subjects will be enrolled to a safety cohort to evaluate the safety and tolerability of ASP2215 given with azacitidine in the study population. Groups of 3 to 6 subjects in a cohort may be enrolled at the same time. The subjects will be initially treated with ASP2215 80 mg daily (days 1-28) (with dose reductions or increases permitted after cycle 1) and azacitidine 75 mg/m^2 daily (days 1-7).

The Sponsor, principal investigators and, if appropriate, expert consultants, will review safety data through the DLT observation period in the safety cohort. The DLT observation period will be from day 1 through day 28 of cycle 1. Evaluable subjects are defined as subjects who experience a DLT or in the absence of DLT, receive at least 23/28 doses of ASP2215 and at least 5/7 doses of azacitidine. Subjects who are not evaluable for reasons other than DLT will be replaced. Based on review of the safety cohort data, the decision to initiate the randomized trial, at the targeted dose (120 mg) or the initial dose (80 mg) dose for ASP2215 in the combination arm will be made by the Sponsor in consultation with the investigators.

Dose escalation rules are as follows:

- At ASP2215 80 mg daily plus azacitidine 75 mg/m2 daily (days 1-7):
  - If 0 of 3 subjects experiences a DLT, 6 subjects will be enrolled at ASP2215 120 mg daily plus azacitidine 75 mg/m2 daily (days 1-7).
  - If 1 of 3 subjects experiences a DLT, up to an additional 3 subjects will be enrolled at ASP2215 80 mg.
    - If 1 of 6 subjects experiences a DLT, 6 subjects will be enrolled in ASP2215 120 mg daily plus azacitidine 75 mg/m2 daily (days 1-7).
    - If ≥ 2 of 4 to 6 subjects experience DLT, the trial will be stopped.
  - If 2 or more of 3 subjects experience a DLT, the trial will be stopped.
- At ASP2215 120 mg daily plus azacitidine 75 mg/m2 daily (days 1-7):

- If 0 out of 5 or 6 subjects or 1 out of 6 subjects experiences a DLT, the randomization will be initiated with ASP2215 120 mg dose for the combination arm.
- If ≥ 2 of 2 to 6 subjects experience a DLT,
  - If less than 6 subjects were previously treated at ASP2215 80 mg dose, then up to 3 more subjects will be enrolled at ASP2215 80 mg daily plus azacitidine 75 mg/m2 daily (days 1-7).
    - If 0 out of 5 or 6 subjects or 1 of 6 subjects experiences a DLT, randomization will be initiated with ASP2215 80 mg dose for the combination arm.
    - If ≥ 2 of 3 to 6 subjects experience a DLT, the trial will be stopped.
  - If a total of 6 subjects were previously treated at ASP2215 80 mg dose, randomization will be initiated with ASP2215 80 mg dose for the combination arm.

If the safety cohort is stopped, the randomized trial will not open. However, an alternative dosing schedule might be explored via an amendment.

3.2.2 Randomized Trial

Approximately 250 subjects will be randomized in a 2:1* ratio to receive ASP2215 plus azacitidine (Arm AC) or azacitidine only (Arm C). The randomization will be stratified based on age group described below.

- Age ≥ 75 years
- Age < 75 years

Subjects will enter the screening period up to 14 days prior to the start of treatment. Subjects will be administered treatment over 28-day cycles. The dose and duration of study treatments are outlined in Section 5.1.1 of the protocol.

ASP2215 dose will be 120 mg for Arm A*, and 120 mg for Arm AC based on safety cohort outcome. Dose increases and reductions are permitted for ASP2215 and azacitidine. Note: Japan only - for the first 6 subjects enrolled or randomized in Japan to the ASP2215+AZA (either in safety cohort or randomization portion of the study), only 1 subject may begin study drug administration in a day (i.e., no 2 subjects will receive their initial dose of study medication on the same day). A subset of sites are participating in the dense PK substudy. The first 12 non-Japanese and 12 Japanese subjects randomized at those sites will participate in the dense PK subset. Dense PK subjects on Arms AC and C should receive azacitidine via subcutaneous injection only for cycle 1. Intravenous infusion will be allowed after cycle 1.

* Protocol versions 6.0 and earlier included a 1:1:1 randomization ratio to receive Arm A (Gilteritinib monotherapy), AC or C. Randomization to Arm A was removed in protocol version 7.0. Subjects previously randomized to Arm A should continue following treatment and assessments as outlined in the protocol.

For all subjects taking ASP2215, ASP2215 plus azacitidine or azacitidine alone, treatment should continue until the subject no longer receives clinical benefit from therapy in the

opinion of the investigator or until unacceptable toxicity occurs or the subject meets another treatment discontinuation criterion.

Dose adjustments for ASP2215 and azacitidine are described in Section 5.1.2 of the protocol.

Subjects in safety cohort, Arm A* and Arm AC proceeding for HSCT can remain on the study and resume treatment with ASP2215 only after HSCT if certain conditions are met [Refer to Section 5.1.5 of the protocol Resumption of Treatment After Hematopoietic Stem Cell Transplantation]. If the subject discontinues the study during or post HSCT, then they should follow the Posttreatment Schedule of Assessments (Table 2) for long term follow-up.

Subjects in Arm C proceeding for HSCT will discontinue treatment by performing an EOT visit and should follow the Posttreatment Schedule of Assessments (Table 2) for long term follow-up.

Subjects will have an end-of-treatment (EOT) visit within 7 days after last dose of study treatment (ASP2215 and/or azacitidine), followed by a 30-day follow-up for safety, after which the subjects will enter the long-term follow-up period of up to 3 years for collection of subsequent AML treatment, EQ-5D-5L, remission status and survival (cause of death and date of death). See Posttreatment Schedule of Assessments (Table 2).

A formal interim analysis by an Independent Data Monitoring Committee (IDMC) will be performed when approximately 50% (i.e., death events = 70) of the planned total number of deaths (i.e., death events = 140) by any cause have occurred. The interim analysis will be utilized to determine if randomized Arm AC has a more favorable or unfavorable outcome compared to Arm C. If the interim analysis demonstrates a more favorable or unfavorable outcome for Arm AC based on OS, enrollment to the study may be stopped. However, any subject continuing to derive clinical benefit as assessed by the investigator will be allowed to continue treatment until they meet a discontinuation criterion as outlined in [Section 6 Discontinuation of the protocol] or upon marketing authorization and commercial availability of ASP2215 for this indication in the country of residence if applicable.

The EFS will be evaluated at the time of OS interim analysis, only if the randomized Arm AC has favorable outcome (i.e., 2-sided p-value <0.003) compared to Arm C with respect to OS. By the time of OS interim analysis with 70 events, 88 EFS events are expected (the actual number of events may vary).

3.3 Randomization

Enrollment, randomization and study drug assignment will be performed via Interactive Response Technology (IRT). Prior to the initiation of the study treatment, the site staff will

* Protocol versions 6.0 and earlier included a 1:1:1 randomization ratio to receive Arm A (ASP2215 monotherapy), AC or C. Randomization to Arm A was removed in protocol version 7.0 and subjects will be randomized in a 2:1 ratio to Arm AC and Arm C stratified by age group (≥ 75 years, < 75 years). Subjects previously randomized to Arm A should continue following treatment and assessments as outlined in the protocol.

contact the IRT in order to determine the randomly assigned treatment. Specific procedures for randomization through the IRT are contained in the study procedures manual.

4 SAMPLE SIZE

This is an open-label, randomized study. One interim analysis and one final analysis are planned. This is a group sequential design based on OS using the O'Brien-Fleming boundaries as implemented by Lan-DeMets alpha/beta spending method (East®).

The interim analysis will occur when approximately 50% (i.e., death events = 70) of the planned total number of deaths (i.e., death events = 140) by any cause have occurred from the randomized Arm AC and Arm C. The IDMC will evaluate the test of OS and inform the Sponsor the result if the randomized Arm AC has favorable outcome (i.e., 2-sided P value < 0.003) or unfavorable outcome (i.e., 2-sided P value ≥ 0.724, non-binding futility boundary) compared to Arm C with respect to OS, the study may be stopped due to efficacy or futility, respectively. Otherwise, the study will continue without impact. The final analysis will be performed after the planned 140 death events have been observed from the randomized Arm AC and Arm C. Additionally, OS will be tested at 2-sided 0.049 significant level for efficacy.

A total of approximately 250 subjects will be randomized to the ASP2215 plus azacitidine (AC) and azacitidine (C) arms. The planned 140 death events will provide at least 80% power to detect a difference in OS between Arm AC and Arm C, assuming 16.7 months median survival time from Arm AC and 10 months median survival time from Arm C (hazard ratio = 0.60) at the overall 2-sided 0.05 significance level.

Based on the planned sample size and final OS analysis timing, 176 EFS events are expected, which will provide above 80% power to detect a hazard ratio of 0.60 in EFS (11.2 months median EFS for Arm AC and 6.7 months for Arm C). The EFS will be evaluated at the time of OS interim analysis, only if the OS result is positive at the interim analysis. By the time of OS interim analysis with 70 events, 88 EFS events are expected (the actual number of events may vary). An O'Brien-Fleming stopping boundary based on Lan-DeMets alpha spending method will be used for EFS. Based on a projected number of events of 88 at the interim, the efficacy stopping boundary is 2-sided nominal alpha of 0.003 for interim analysis and 0.049 for the final analysis. The actual rejection boundary for EFS may vary according to the actual number of EFS events that occur at the interim analysis.

The sample size for the safety cohort is not based on a statistical power calculation. The planned number of evaluable subjects up to 12 would provide adequate information for the objectives of the safety cohort.

5 ANALYSIS SETS

In accordance with International Conference on Harmonization (ICH) recommendations in guidelines E3 and E9, the following analysis sets will be used for the analyses.

The determination of whether subjects are included or excluded from analysis sets will be made prior to database hard-lock. Inclusions and exclusions from the pharmacokinetic analysis set may occur after unblinding. Detailed criteria for analysis sets will be laid out in Classification Specifications (CS) if manual classification is required.

The data from all randomized/registered subjects will be included in the data tables, listings, and figures.

5.1 Full Analysis Set (FAS)

The Full Analysis Set (FAS) is defined as the intention to treat set, which will consist of all subjects who are randomized and all registered subjects from the safety cohort and will be used for efficacy analyses. Subjects will be analyzed based on the randomized treatment arms (Arms A, C, AC). In addition, subjects randomized to Arm AC and all subjects from the safety cohort will be analyzed as Total Arm AC.

The FAS will be used for summaries and analyses of efficacy data, patient reported outcomes, resource utilization, as well as selected demographic and baseline characteristics. In addition, FAS will be used for exploratory biomarker variables.

5.2 Safety Analysis Set (SAF)

The safety analysis set (SAF) consists of all subjects who received at least one dose of study drug (ASP2215 or azacitidine). A subject erroneously receiving a treatment different from the randomized treatment will be assigned to the treatment group that subject received at first dose.

The SAF will be used for summaries of demographic and baseline characteristics and all safety and tolerability related variables.

5.3 Pharmacokinetics Analysis Set (PKAS)

The pharmacokinetic analysis set (PKAS) consists of the randomized subjects in SAF and provide at least one sample for the measurement of drug concentrations and both the date and time of dosing and PK sampling are known. Exclusion of subjects from the PKAS will be considered by the pharmacokineticist on a case-by-case basis. Any formal definitions for exclusion of subjects or time points from the PKAS will be documented by pharmacokineticist using CPED document.

The PKAS will be used for all tables and graphical summaries of the PK data. Subjects with dense PK sampling will be identified in the PK data. PKAS will be used for PK-PD analyses.

6 ANALYSIS VARIABLES

If variables are derived by referring to the randomization date for the randomized subjects, and the same variables will be derived by referring to registration date for the subjects in the safety cohort if appropriate.

6.1 Efficacy Endpoints

6.1.1 Primary Efficacy Endpoint

Overall survival (OS) is the primary efficacy endpoint. OS is defined as the time from the date of randomization until the date of death from any cause (death date – randomization date + 1). For a subject who is not known to have died by the end of study follow-up, OS is censored at the date of last contact (date of last contact – randomization date + 1). The date of last contact is the latest date that the subject is known to be alive at the analysis cutoff date. The last contact date will be derived for subjects alive at the analysis cutoff date. Subjects with death or last known alive date beyond the analysis cutoff date will be censored at the analysis cutoff date.

Date of last contact is defined as the latest of the following dates: treatment discontinuation date, 30-day follow-up date, last dosing administration date (starting/stopping dose dates), last disease assessment date (including bone marrow, lab, ECG, ECOG, vital signs, CM, hospitalization, PRO, resource utilization assessment dates), AE (starting/stopping dates), last known alive date from long term follow-up, and randomization date.

During the interim analysis and final analysis, only deaths occurring on or prior to the cutoff date are counted as OS events.

As a sensitivity analysis, OS will be defined similarly as the above primary analysis; however, subjects who undergo an HSCT will be censored at the time of HSCT (HSCT date – randomization date +1).

6.1.2 Secondary Efficacy Endpoints

6.1.2.1 Key Secondary Efficacy Endpoints

- Event-free survival (EFS)

  EFS is defined as the time from the date of randomization until the date of documented relapse from CR (defined below in Section 6.1.2.3 Response Definitions), treatment failure (failing to achieve CR within 6 cycles of treatment) or death from any cause, whichever occurs first [earliest of (relapse date, treatment failure date, death date) – randomization date + 1]. For a subject who is not known to have relapse from CR or treatment failure or death, EFS is censored at the date of last relapse-free disease assessment (last relapse-free disease assessment date – randomization date +1).

  The following sensitivity analyses for EFS will be conducted:

  - Defining EFS as the time from randomization until the date of documented relapse from CR (for subjects who achieved CR within 6 cycles of treatment), or death from any cause, whichever comes first
  - Defining EFS as the time from randomization until the date of documented relapse from CR (for subjects who achieved CR within 6 cycles of treatment), or treatment failure (the date of permanent discontinuation of all study treatments or the end of 6 cycles of treatment), or death from any cause, whichever is earlier

Scenarios (at the EFS data cutoff date)	**EFS indicator [Event Type]; EFS date; Time to EFS (days)**		
	Primary Analysis	**Sensitivity Analysis 1**	**Sensitivity Analysis 2**
Randomized, but not treated	Non-Event [Censor]; Date of randomization; EFS = 1	Same as primary analysis	Same as primary analysis
Randomized and treated			
No post-baseline disease assessment	Event [type=TF]; Date at randomization; EFS = 1	Non-Event [Censor]; Date of randomization; EFS = 1	Event [type = TF]; Date at min(last dose, C7D1); EFS=date of min(last dose, C7D1) – date of randomization + 1
At least 1 post-baseline disease assessment, do not achieve CR within 6 cycles of treatment			
Die prior to completing 6 cycles of treatment	Event [type=death]; Date of death; EFS = date of death – date of randomization + 1	Same as primary analysis	Same as primary analysis
Alive without completing 6 cycles of treatment due to early discontinuation or being ongoing	Event [type=TF]; Date at randomization; EFS = 1	Non-Event [Censor]; Date of last clinical assessment ≤ C7D1; EFS = date of last clinical assessment ≤ C7D1 – date of randomization + 1	Event [type = TF]; Date at min(last dose, C7D1); EFS=date of min(last dose, C7D1) – date of randomization + 1
Alive and completing 6 cycles of treatment	Event [type=TF]; Date at randomization; EFS = 1	Non-Event [Censor]; Date of last clinical assessment ≤ C7D1; EFS = date of last clinical assessment ≤ C7D1 – date of randomization + 1	Event [type = TF]; Date at min(last dose, C7D1); EFS=date of min(last dose, C7D1) – date of randomization + 1

Table continued on next page

At least 1 post-baseline disease assessment, achieve 1st CR within 6 cycles of treatment			
Die without relapse	Event [type=death]; Date of death; EFS = date of death – date of randomization + 1	Same as primary analysis	Same as primary analysis
Relapse	Event [type=relapse]; Date at relapse; EFS = date of relapse – date of randomization + 1	Same as primary analysis	Same as primary analysis
Alive, no relapse or death	Non-Event [Censor]; Date of last relapse free clinical assessment; EFS = date of last clinical assessment – date of randomization + 1	Same as primary analysis	Same as primary analysis

Note: TF = Treatment Failure; 6 cycle of treatments = from first dose date until first date of C7D1 visit or 168 days of treatment (6*28) if first date of C7D1 visit is missing.

Disease assessment date refers to the date of bone marrow aspiration or biopsy assessment. In the event that central bone marrow assessment is not performed, or bone marrow is not adequate, local bone marrow assessment date will be used. If no aspirate or biopsy is available and subject is evaluated based on blast count from peripheral blood at a visit that bone marrow is expected to be collected, the date when the peripheral blood sample is drawn will be used.

6.1.2.2 Other Secondary Efficacy Endpoints

- Complete remission (CR) rate

  CR is defined below in Section 6.1.2.3 Response Definitions. CR rate is defined as the number of subjects who achieve the best overall response of CR divided by the number of subjects in the analysis population. Subjects with unknown or missing response, or who provide no information on response at the end of treatment will be included in the denominator when calculating rates.

- Composite complete remission (CRc) rate

  CRc rate is defined as the number of subjects with best response of CR, CRp, or CRi (i.e., CR+ CRp + CRi) divided by the number of subjects in the analysis population.

- CRp rate, CRi rate – Defined similarly as CR rate.

- CR/CRh rate

CR/CRh is defined as the number of subjects who achieve CR or CRh divided by the number of subjects in the analysis population.

- CRh rate

  CRh rate is defined as the number of subjects who achieve CRh at any post-baseline visits and do not have best response of CR divided by the number of subjects in the analysis population.

- Duration of remission

  Duration of remission includes duration of CRc, duration of CR, duration of CRh, duration of CR/CRh and duration of response (CRc + PR).

  Duration of CRc is defined as the time from the date of first CRc until the date of first documented relapse for subjects who achieve CRc (relapse date – first CRc disease assessment date + 1). Subjects who die without report of relapse are considered non-events and censored at their last relapse-free disease assessment date (last relapse-free disease assessment date – first CRc disease assessment date + 1). Other subjects who achieve CRc and do not relapse on study are considered non-events and censored at the last relapse-free disease assessment date.

  Duration of CR and duration of CRh are defined similarly as duration of CRc. And duration of CR and duration of CRh will be only derived for subjects with best response of CR and CRh, respectively.

  Duration of CR/CRh is defined similarly as duration of CRc for subjects who achieve either CR or CRh. For subjects who achieve both CR and CRh, the first date of CR/CRh will be used as the starting date of the duration.

  As a sensitivity analysis, duration of CR and duration of CR/CRh will be defined similarly as the above; however, deaths without documented relapse are considered as event.

  Duration of response is defined as the time from the date of either first CRc or PR until the date of first documented relapse of any type (i.e. the date of first NR after CRc or PR) for subjects who achieve CRc or PR (relapse date – first CRc or PR disease assessment date + 1). Subjects who die without report of relapse are considered non-events and censored at their last relapse-free disease assessment date (last relapse-free disease assessment date – first CRc or PR disease assessment date + 1). Other subjects who do not relapse on study are considered non-events and censored at the last response relapse-free disease assessment date.

- Leukemia-free survival (LFS)

  LFS is defined as the time from the date of first CRc, defined below in Section 6.1.2.3 Response Definitions, until the date of documented relapse (excluded relapse from PR) or death for subjects who achieve CRc (relapse date or death date – first CRc disease assessment date + 1). For a subject who achieves CRc and is not known to have relapsed

or died, LFS is censored on the date of last relapse-free disease assessment date (last relapse-free disease assessment date – first CRc disease assessment date + 1).

- Time to Remission

  Time to remission includes time to CRc, time to CR, time to first CR/CRh, time to best CR/CRh, and time to response (CRc + PR).

  Time to CRc is defined as the time from the date of randomization until the date of first CR, CRp, or CRi (first CR or CRp or CRi disease assessment date – randomization date +1). For subjects who achieve multiple responses (CR, CRp or CRi), the earliest date among the first response dates will be used.

  Time to CR is defined similarly as time to CRc.

  Time to first CR/CRh is defined similarly as time to CRc for subjects who achieved either CR or CRh. For subjects who achieve both CR and CRh, the first CR date or CRh date, whichever occurs first, will be used.

  Time to best CR/CRh is defined similarly as time to CRc for subjects who achieved either CR or CRh. For subjects who achieve both CR and CRh, the first CR date will be used.

  Time to response is defined similarly as time to CRc for subjects who achieved either CRc or PR.

- Transfusion conversion rate and transfusion maintenance rate

  For the purpose of defining transfusion conversion rate and transfusion maintenance rate, transfusion status (independent vs. dependent) at baseline period and post-baseline period are defined in the following for subjects who took at least one dose of study drug:

  Baseline transfusion status:

  - Baseline period is defined as the period from 28 days prior to the first dose to earliest of (28+first dosing date, max(date of last dose of ASP2215, initial dosing date of the last cycle of azacitidine +27), date of death, or date of data cutoff).
  - Subjects are classified as baseline transfusion independent if there is no RBC or platelet transfusions within the baseline period; otherwise, the subject is baseline transfusion dependent.

  Post-baseline transfusion status:

  - Post-baseline period is defined as the period from 29 days post first dose until earliest of (max(date of last dose of ASP2215, initial dosing date of the last cycle of azacitidine +27) , date of death, or date of data cutoff).
  - For subjects with post-baseline period of at least 56 days, subjects are classified as post-baseline transfusion independent if there is at least one consecutive 56 days without any RBC or platelet transfusion within post-baseline period.

- For subjects with post-baseline period <56 days, if there is no RBC or platelet transfusion within post-baseline period, post-baseline transfusion status is not evaluable.
- For subjects without post-baseline period, post-baseline transfusion status is not evaluable.
- Otherwise, the subject is considered as post-baseline transfusion dependent.

Both transfusion conversion rate and maintenance rate are defined for subjects who has evaluable post-baseline transfusion status.

Transfusion conversion rate is defined as the number of subjects who were transfusion dependent at baseline period but become transfusion independent at post-baseline period divided by the total number of subjects who were transfusion dependent at baseline period.

Transfusion maintenance rate is defined as the number of subjects who were transfusion independent at baseline period and still maintain transfusion independent at post-baseline period divided by the total number of subjects who were transfusion independent at baseline period.

Additionally, transfusion conversion rate and transfusion maintenance rate will be defined by including all dosed subjects including those who had not-evaluable post-baseline transfusion status into denominator.

- Brief fatigue inventory (BFI)

  BFI [Mendoza et al, 1999] was developed to assess the severity of fatigue and the impact of fatigue on daily functioning in patients with fatigue due to cancer and cancer treatment. The BFI short form has 9 items and a 24-hour recall. A global fatigue score is computed by averaging the scores of the 9 items measured on the numeric rating scale. The average can be computed only if at least 5 of the 9 items are answered. A higher score indicates a higher degree of fatigue.

6.1.2.3 Response Definitions

Response to treatment will be defined per modified Cheson criteria [2003] as outlined below.

Response at a post-baseline visit will be derived using myeloblast counts from centrally evaluated bone marrow aspirate if it's adequate. In case of non-adequacy, myeloblast counts from centrally evaluated bone marrow biopsy will be used. If neither central bone marrow aspirate nor biopsy is available, myeloblast counts will be imputed with locally evaluated bone marrow aspirate, if not available, locally evaluated bone marrow biopsy assessments. Centrally evaluated hematology results including ANC, platelet count and blast count in peripheral blood will be used in response derivation. Missing central hematology results will be imputed with local hematology results as collected on the eCRF.

Response will be derived for all post-baseline visits on or after 21 days from first dosing date.

- Best Response

Best response for a subject is defined as the best measured response (in the order of CR, CRp, CRi, and treatment failure) from all post-baseline visits.

Subjects with best responses of CR, CRp, and CRi will be considered responders. Subjects who do not achieve at least a best response of CRi will be considered non-responders.

- Complete Remission (CR)

  Response at a post-baseline visit is classified as CR for having bone marrow regenerating normal hematopoietic cells and achieving a morphologic leukemia-free state and having an ANC $\geq 1 \times 10^9$/L and platelet count $\geq 100 \times 10^9$/L and normal marrow differential with < 5% myeloblast counts and missing or ≤2% peripheral blood blast counts and being red blood cell (RBC) and platelet transfusion independent (defined as 1 week without RBC and platelet transfusion prior to the disease assessment). There should be no evidence of extramedullary leukemia. There is no evidence of Auer rods.

- Complete Remission with Partial Hematologic Recovery (CRh)

  Response at a post baseline visit is classified as CRh if marrow myeloblast counts < 5%, partial hematologic recovery ANC $\geq 0.5 \times 10^9$/L and platelets $\geq 50 \times 10^9$/L, no evidence of extramedullary leukemia and cannot be classified as CR. The blast counts in peripheral blood must be ≤ 2% or missing.

- CR/CRh

  Response at a post baseline visit is classified as CR/CRh if it fulfills criteria for CR or CRh at the visit.

- Complete Remission with Incomplete Platelet Recovery (CRp)

  Response at a post-baseline visit is classified as CRp, for achieving CR except for incomplete platelet recovery ($< 100 \times 10^9$/L).

- Complete Remission with Incomplete Hematologic Recovery (CRi)

  Response at a post-baseline visit is classified as CRi, for fulfilling all the criteria for CR except for incomplete hematological recovery with residual neutropenia $< 1 \times 10^9$/L with or without complete platelet recovery. RBC and platelet transfusion independence is not required.

- Composite Complete Remission (CRc)

  Response at a post baseline visit is classified as CRc if it fulfills criteria for CR, CRp or CRi at the visit.

- Partial Remission (PR)

  If marrow myeloblasts between <5% and 25% and ≤2% or missing peripheral blood blast and a decrease from baseline of at least 50% in the marrow myeloblasts and no evidence of extramedullary leukemia, the response is classified as PR. If marrow myeloblasts <5% and

≤2% or missing peripheral blood blast and no evidence of extramedullary leukemia, the response is classified as PR even with Auer rods.

- Non-Response (NR)/Not Evaluable (NE)

  Response at a post-baseline visit will be classified as not evaluable (NE) if no bone marrow assessments are performed, no blasts from peripheral blood are observed (blast ≤ 2% or missing) and extramedullary leukemia is missing or not present. In any case, response at a post-baseline visit cannot be classified as CR, CRp, CRi, PR, or NE, it will be classified as non-response (NR).

- Relapse

  Relapse at a post-baseline visit after CR, CRp, or CRi is defined as a reappearance of leukemic blasts (>2%) in the peripheral blood or ≥ 5% myeloblasts in the bone marrow not attributable to any other cause or reappearance or new appearance of extramedullary leukemia.

  Relapse at a post-baseline visit after PR is similarly defined with reappearance of significant numbers of peripheral blasts (>2%) or an increase in the percentage of myeloblasts in the bone marrow to > 25% not attributable to any other cause or reappearance or new appearance of extramedullary leukemia.

- Treatment Failure (defined by lack of CRc)

  Treatment failure for a subject is defined as lack of CRc (CR, CRp or CRi), and is determined at the end of treatment.

6.1.3 Exploratory Endpoints

- Transplantation

  Transplantation will be evaluated by transplantation rate, which is defined as the percentage of subjects undergoing HSCT during the study period and may be defined as the percentage of subjects undergoing HSCT during the study period from Protocol Version 9 if appropriate.

- Minimal residual disease (MRD)

  MRD will be assessed by molecular assessment as the ratio of FLT3-ITD sequencing reads to total FLT3 sequencing reads after baseline. At a time point, a subject may have several ITD mutations characterized by different mutation length in base pair. The individual mutation ratio and the sum of the mutation ratios at each time point for each subject will be used in summary.

  The MRD status of being positive or negative will be defined as: the MRD status is positive if the sum of FLT3-ITD mutation ratios is larger than 10^{-4} at a time point. Otherwise, the MRD status is negative.

- Baseline FLT3 gene mutation type
  - Mutation types (internal tandem duplication [ITD] alone or tyrosine kinase domain [TKD] [D835/I836] alone, or both ITD and TKD) and frequency
  - Subgroup analysis of efficacy and safety based on baseline status
- Baseline FLT3 Mutation Status
  - Subjects with FLT3-ITD mutations will be classified based on low allelic ratio (< 0.5, LAR) or high allelic ratio (≥ 0.5, HAR). Allelic ratio will be determined by the LeukoStrat CDx FLT3 Mutation Assay and calculated as the ratio of the peak area of the FLT3-ITD mutant signal, if present, divided by the peak area of the FLT3 wild type signal, if present
  - FLT3 mutation status can be classed as FLT3-TKD, FLT3-ITD low allelic ratio (< 0.5, LAR), and FLT3-ITD high allelic ratio (≥ 0.5, HAR)
  - These factors will be used as stratification factors in Section 6.4.6 and subgroup analysis in Section 7.8
- Patient Reported Outcome (PRO) Measures

  PRO measures include BFI (secondary endpoint) and the following exploratory endpoints:

  - Patient reported dyspnea via Functional Assessment of Chronic Illness Therapy-Dyspnea Short Form [FACIT-Dys-SF],
  - Patient reported signs, symptoms and impacts of AML via Functional Assessment of Cancer Therapy-Leukemia [FACT-Leu],
  - Stand-alone dizziness and mouth sores items,
  - Health related quality of life via EuroQol Group 5 dimension 5 level [EQ-5D 5L]).

  PRO measures are collected via electronic PRO device.

  - Functional Assessment of Chronic Illness Therapy-Dyspnea-Short Forms (FACIT-Dys-SF)

    The FACIT-Dys-SF [Choi et al, 2011] was developed to assess dyspnea severity and related functional limitations. It comprises of two parts. Part I asks patients about ten common daily tasks and Part II lists the same ten activities and inquiries about the difficulty completing them. It has a 7-day recall period and 20 items. The FACIT-Dys-SF is scored with 2 domains: dyspnea and function limitations. The higher the score is, the worse the dyspnea or functional limitations are.

    The derivation algorithm is as follows.

    - If the answer to the item is “0” or “.”, then the item is considered to be not answered.

- For each domain, the raw score will be calculated if at least 6 out of the 10 items were answered. Otherwise, the raw score is set to be missing.
- For each domain, the raw score is equal to rounding (10*[the total of the scores/number of questions answered]) to integer.
- The raw score for each domain will be converted to scale score based on the following conversion table:

Dyspnea		Functional Limitations	
Raw Score	Scale Score	Raw Score	Scale Score
10	27.7	10	29.7
11	32.8	11	34.9
12	36.1	12	38.0
13	38.6	13	40.3
14	40.6	14	42.1
15	42.3	15	43.8
16	43.8	16	45.2
17	45.2	17	46.5
18	46.4	18	47.8
19	47.6	19	49.0
20	48.8	20	50.1
21	50.0	21	51.2
22	51.1	22	52.3
23	52.1	23	53.4
24	53.2	24	54.4
25	54.2	25	55.4
26	55.2	26	56.4
27	56.2	27	57.4
28	57.2	28	58.4
29	58.1	29	59.4
30	59.2	30	60.4
31	60.2	31	61.4
32	61.2	32	62.4
33	62.3	33	63.5
34	63.5	34	64.7
35	64.8	35	66.0
36	66.1	36	67.3
37	67.7	37	68.9
38	69.5	38	70.7
39	71.9	39	73.0
40	75.9	40	76.7

- Functional Assessment of Cancer Therapy-Leukemia (FACT-Leu)

  The FACT-Leu [Cella et al, 2012] is a 44-item questionnaire designed to measure health-related quality of life (HRQoL) in leukemia. The FACT-Leu contains some of the most common patient reported signs, symptoms, and impacts of AML. The questionnaire includes a global score and 5 subscales including physical well-being (PWB), social/family well-being (SWB), emotional well-being (EWB), functional well-being (FWB) and a leukemia subscale (LeuS). In addition to the 5 subscales, the FACT-Leu can be used to calculate the Functional Assessment of Cancer Therapy-General (FACT-G) score, the FACT-Leu Trial Outcome Index (FACT-Leu TOI), and the FACT-Leu total score which is comprised of all 44 items included in the questionnaire. The FACT-Leu has a 7-day recall period. A higher score indicates a better quality of life.

  For Questions 1-7, 16, 18-21, 29-38, 41-45, the derived score=4-raw score. For each of 5 domains, the score is calculated as number of items*(the total of the scores/number of items answered) if more than half of the items were answered. Otherwise, the score is set to be missing. The leukemia global (total) score is the sum of the scores of the 5 domains.

- Dizziness and mouth sore items

  Dizziness and mouth sore items which utilize a 5-point scale (0 to 4). A lower score indicates less experience of the symptom. The higher the score is, the worse dizziness or mouth sore is.

  There is no derivation for dizziness score and mouth sore score. The data collected will be summarized directly.

- EuroQol Group-5 Dimension-5 Level Instrument (EQ-5D-5L)

  The EQ-5D-5L is a common measure of patients' HRQoL. The EQ-5D-5L consists of the EuroQol Group-5 Dimension descriptive system and the EuroQol Group Visual Analogue Scale (VAS).

  The EuroQol Group-5 Dimension descriptive system consists of 5 items assessing 5 key dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems and extreme problems.

  The VAS records the respondent's self-rated health status on a graduated (0 - 100) scale, where the endpoints are labeled 'worst imaginable health state' (0) and 'best imaginable health state' (100) with higher scores for higher HRQoL.

  There is no derivation for EQ-5D-5L. The data collected will be summarized directly.

- Resource utilization, including hospitalization, blood transfusion, antibiotic intravenous infusions, medication for AEs and opioid usage

  There are eCRF pages to collect start date and stop date for hospitalization and blood transfusion, where the duration of hospitalization and blood transfusion can be derived.

Antibiotic intravenous infusions, medications for AEs, and opioid usage will be collected in the eCRF of the concomitant medication. The types and duration of usage for these medications will be derived. The ATC codes for these medications will be footnoted in the table.

- Pharmacogenomics (PGx)

  PGx research may be conducted in the future to analyze or determine genes of relevance to clinical response, PK, and toxicity/safety issues.

6.2 Safety Variables

Safety will be assessed by evaluation of the following variables:

- Treatment-emergent adverse events (TEAEs; frequency, severity, CTCAE grade, seriousness, and relationship to study drug)

  TEAE is defined as an AE observed after starting administration of the study treatment (ASP2215 or azacitidine) until 30 days from the last study treatment. If the AE occurs on Day 1 and the onset check box is marked "Onset after first dose of study drug" or the onset check box is left blank, then the AE will be considered treatment emergent. If the AE occurs on Day 1 and the onset check box is marked "Onset before first dose of study drug", then the AE will not be considered treatment emergent. If a subject experiences an event both during the pre-investigational period and during the investigational period, the event will be considered a TEAE only if it has worsened in severity (i.e., it is reported with a new start date). All AEs collected that begin within 30 days after taking the last dose of study drug are counted as TEAE, except for subjects that undergo HSCT without leaving the study and plan to resume ASP2215 treatment after HSCT. For these subjects, TEAE is defined as AEs observed after starting administration of the study treatment until the last dose before pre-HSCT visit plus 30 days, and AEs that begin after resumption of ASP2215 and within 30 days after the last dose of ASP2215 are counted as TEAE. Any AEs with onset dates completely missing will be considered TEAEs in summaries. AEs with partially missing onset dates will be assumed TEAEs unless the available portion of the date indicates that the onset was strictly before start of study medication or 30 days after the last study treatment. Missing or partial AE onset date will be imputed per Section 7.12.1.

- A drug-related TEAE is defined as any TEAE with at least possible relationship (possibly or probably related) to study treatment as assessed by the investigator or with missing assessment of the causal relationship.

- Serious adverse events (SAEs) include AEs that are flagged as serious by the investigator on an eCRF, or upgraded by the Sponsor based on review of the Sponsor's list of Always Serious term.

- Adverse events of special safety interest (AESI) are defined in the Safety Review Plan for ASP2215.

- Clinical laboratory variables (hematology, biochemistry including liver function tests and thyroid function test, coagulation, and urinalysis)
- Vital signs (systolic and diastolic blood pressure, pulse rate, and body temperature)
- 12-lead electrocardiogram (ECG)
- ECOG performance scores

6.2.1 Safety Cohort and First 6 Japanese Subjects on Arm AC (either in the Safety Cohort or Randomization Portion of the Study) - Definition of DLT

A DLT is defined as any of the following events that occur during the observation period and that is considered to be possibly or probably related to study regimen. The observation period for DLT will be from the start of treatment until day 28 of the first cycle.

- Any grade ≥ 3 nonhematologic or extramedullary toxicity with the following exceptions:
  - Anorexia or fatigue
  - Grade 3 nausea and/or vomiting if not requiring tube feeding or total parenteral nutrition, or diarrhea if not requiring or prolonging hospitalization that can be managed to grade ≤ 2 with standard antiemetic or antidiarrheal medications used at prescribed dose within 7 days of onset
  - Grade 3 mucositis that resolves to grade ≤ 2 within 7 days of onset
  - Grade 3 fever with neutropenia, with or without infection
  - Grade 3 infection
- Prolonged myelosuppression defined as ANC ≤ 0.5×10^9/L for more than 21 days from the onset of severe neutropenia in the absence of evidence of active leukemia in the marrow or blood.
- Any toxicity that requires a dose reduction (protocol Section 5.1.2)

6.3 Pharmacokinetics Variables

The following PK variables will be included in the analysis:

- Plasma concentration data of ASP2215
- Plasma concentration of ASP2215 and azacitidine from dense PK subset (Japanese and Non-Japanese subjects from Arm AC and Arm C)

6.4 Other Variables

6.4.1 Body Mass Index (BMI)

BMI = weight (kg) / [height (m)]^2

6.4.2 First and Last Dose Dates

The study will define start and stop dates for ASP2215, AZA and Overall for every subject.

For ASP2215:

- Start date is the start date of the first ASP2215 record

- Stop date is maximum dates from the last ASP2215 record if the last record is the final record or the subject discontinued treatment or data cutoff date

For AZA:

- Start date is the start date of the first Aza injection or infusion record
- Stop date is the stop date from the last Aza injection or infusion record if the subject discontinued treatment or data cutoff date

For Overall:

- If a subject is in Arm A (ASP2215 alone) or in Arm C (AZA alone), the overall start and stop dates are the same as ones from their study drug.
- If a subject is on ASP2215+AZA,
  - the start date is the minimum date from ASP2215 and AZA
  - the stop date is missing if either ASP2215 stop date or AZA stop date does not exist; otherwise the maximum date of the stop dates from ASP2215 and AZA.

6.4.3 Study Drug Exposure

Dose exposure for ASP2215

- Number of cycles

  Number of cycles = total number of cycles – number of cycles without any ASP2215
- Duration of exposure (days)

  Duration of exposure = earliest of (date of last dose of ASP2215, date of death, or date of data cutoff) – date of first dose of ASP2215 + 1 – (on-study HSCT period for subjects undergo on-study HSCT)
- Number of dosing days (days)

  Number of dosing days = number of days with ASP2215 dose > 0 mg
- Cumulative dose (mg)

  Cumulative dose = total dose of ASP2215 taken during the study
- Average daily dose (mg/day)

  Average daily dose = cumulative dose (mg)/number of dosing days (days)
- Dose intensity (mg/day)

  Dose intensity = cumulative dose (mg)/ duration of exposure (days)
- Planned dose intensity (mg/day)

  Planned daily dose = 120 mg/day for Arm A and Arm AC and 80 or 120 mg/day for safety cohort
- Relative dose intensity (%)

  Relative dose intensity = dose intensity (mg/day)/planned daily dose (mg/day)*100%
- Compliance

  Compliance = number of dosing days (days)/ duration of exposure (days)*100%

Dose exposure for azacitidine

- Number of cycles

  Number of cycles = total number of cycles – number of cycles without any azacitidine
- Duration of exposure (days)

  Duration of exposure = earliest of (initial dosing date of the last cycle of azacitidine + 27, date of death, or date of data cutoff) – date of first dose of azacitidine + 1
- Number of dosing days (days)

  Number of dosing days = number of days with azacitidine dose > 0 mg
- Cumulative dose (mg)

  Cumulative dose = total dose of azacitidine taken during the study
- Average daily dose (mg/day)

  Average daily dose = cumulative dose (mg)/number of dosing days (days)
- Dose intensity (mg/day)

  Dose intensity =cumulative dose (mg)/duration of exposure (days)
- Planned dose intensity (mg/day)

  Planned dose intensity =sum (7 (days/cycle)* 75 (mg/m^2/day)* BSA (m^2) at visit/28 (days/cycle))/number of cycles
- Relative dose intensity (%)

  Relative dose intensity = dose intensity (mg/day)/ planned dose intensity (mg/day) *100%
- Compliance

  Compliance =4* number of dosing days (days)/ duration of exposure (days)*100%

Dose exposure for Total ASP2215+ azacitidine

- Number of cycles

  Number of cycles =number of cycles with ASP2215 and/or azacitidine
- Duration of exposure (days)

  Duration of exposure = earliest of (max(date of last dose of ASP2215, initial dosing date of the last cycle of azacitidine +27) , date of death, or date of data cutoff) – min (date of first dose of ASP2215, date of first dose of azacitidine)+1– (on-study HSCT period for subjects undergo on-study HSCT)
- Number of dosing days (days)
  Number of dosing days = days with ASP2215 and/or azacitidine
- Cumulative dose (mg)

  Cumulative dose = NA (not applicable)
- Average daily dose (mg/day)= NA
- Dose intensity (mg/cycle)= NA
- Planned dose intensity (mg/cycle)= NA

- Relative dose intensity (%)=NA
- Compliance = NA

6.4.4 Duration of AML at time of randomization/registration

Duration of AML will be calculated in days using the following formula:

(Randomization/registration date – date of initial diagnosis of AML) + 1

Partial date of initial diagnosis of AML will be imputed per Section 7.12.1.

6.4.5 Prior and Concomitant medications and non-medication therapy

- Previous and concomitant treatment

  Previous treatment is defined as treatment administered before the date of first dosing (exclusive).

  Concomitant treatment is defined as treatment administered between the date of first dose (inclusive) and the date of last dose (inclusive) of study drug.

  For subjects that undergo HSCT without leaving the study and plan to resume ASP2215 after HSCT, concomitant medication is defined as medication with at least one dose taken between the date of first dose (inclusive) and the date of last dose (inclusive) before pre-HSCT visit, or between resumption of ASP2215 (inclusive) and the date of last dose (inclusive).

- Previous and concomitant transfusion

  Previous transfusion is defined as transfusion received before the date of first dose of study drug, i.e., transfusion completed before the date of the first dose.

  Concomitant transfusion is defined as transfusion received between the date of the first dose (inclusive) and the date of last dose (inclusive) of study drug.

  For subjects that undergo HSCT without leaving the study and plan to resume ASP2215 after HSCT, concomitant transfusion is defined similarly as concomitant medication.

6.4.6 Stratification Variables Used for Analysis

- Age Group
  - Age ≥ 75 years
  - Age < 75 years
- Risk Group
  - Favorable or intermediate cytogenetic risk
  - Unfavorable cytogenetic risk or secondary AML (regardless of cytogenetic risk)
- Baseline FLT3 Mutation Status
  - FLT3-TKD
  - FLT3-ITD low allelic ratio (< 0.5, LAR)
  - FLT3-ITD high allelic ratio (≥ 0.5, HAR)

Cytogenetic risk group used for stratification are defined by the presence of cytogenetic abnormalities and based on common clinical criteria [National Comprehensive Cancer Network, 2015] and presented in [Table 4].

Table 4 Cytogenetic Risk

Cytogenetic Risk Category	Definitions
Favorable	• Core binding factor: inv(16) or t(16;16) or t(8;21). Other cytogenetic abnormalities in addition to these do not alter better risk status.
Intermediate	• Normal cytogenetics • +8 alone • t(9;11) • Other non-defined cytogenetic abnormality
Unfavorable	• Complex (≥ 3 clonal chromosomal abnormalities) • Monosomal karyotype • -5, 5q-, -7, 7q- • 11q23 – non t(9;11) • inv(3), t(3;3) • t(6;9)

t(15;17) and t(9;22) are not included in the table as they are not eligible for this trial.

Subjects of non-secondary AML with unknown cytogenetic risk category will be classified into the group of favorable or intermediate cytogenetic risk for stratification analysis.

7 STATISTICAL METHODOLOGY

7.1 General Considerations

Since the study enrollment was terminated on 17Dec2020 due to futility at the interim analysis for OS by IDMC, all statistics from the final analysis are summary statistics without any formal testing.

For continuous variables, descriptive statistics will include the number of subjects (n), mean, standard deviation, median, minimum and maximum. When needed, the use of other percentiles (e.g., 10%, 25%, 75% and 90%) will be mentioned in the relevant section. In addition, for plasma concentrations, the coefficient of variation and the geometric mean will also be calculated. Frequencies and percentages will be displayed for categorical data. Percentages by categories will be based on the number of subjects with no missing data, i.e., will add up to 100%.

Summaries based on FAS (e.g., disposition, baseline and efficacy data) will be presented by randomized treatment groups (Arms A, C, AC), ASP2215+AZA (Safety Cohort), ASP2215+AZA(Total), unless specifically stated otherwise. Safety analysis and other summaries based on SAF will be presented by actual treatment received. Pharmacokinetic summaries based on PKAS will be presented by actual treatment received. For subjects with dose increase/decrease, actual treatment refers to the first dose received before dose change.

All statistical comparisons will be made using two-sided tests at α=0.05 significance level unless specifically stated otherwise. Formal statistical tests will be only carried out between Arm AC vs Arm C.

All data processing, summarization, and analyses will be performed using SAS® Version 9.3 or higher on Unix. Specifications for table, figure, and data listing formats can be found in the TLF specifications for this study.

Baseline is defined as the last available measurement before or at the randomization/registration date for efficacy assessment and before or at the first dose date of study drug for safety assessment. Unless otherwise specified, all summaries will be presented by treatment group.

For the definition of subgroups of interest please refer to Section 7.9.

7.2 Study Population

7.2.1 Disposition of Subjects

The following subject data will be presented:

- Number and percentage of subjects with informed consent, screening, discontinued before randomization/registration, randomized/registered (total only);
- Number and percentage of subjects in each analysis set by treatment group and total;
- Number and percentage of subjects who completed or discontinued treatment by primary reason for treatment discontinuation and by treatment group and total;
- Number and percentage of subjects who completed the 30-day follow-up evaluation by status and by treatment group and total;
- Number and percentage of subjects who completed the long term follow-up evaluation by status and by treatment group and total; and

All subject disposition data will be presented in listings.

7.2.2 Protocol Deviations

Protocol deviations (PDs) as defined in the study protocol (Section 8.1.6 Protocol Deviations) will be assessed for all randomized/registered subjects. The number and percentage of subjects with the following protocol deviation criteria will be summarized for each criterion and overall, by treatment group and total as well as by investigative site. Subjects deviating from a criterion more than once will be counted once for the corresponding criterion. Any subjects who have more than one protocol deviation will be counted once in the overall summary.

A data listing will be provided by site and subject.

The protocol deviation criteria will be uniquely identified in the summary table and listing. The unique identifiers will be as follows:

PD1 - Entered into the study even though they did not satisfy entry criteria,

PD2 - Developed withdrawal criteria during the study and was not withdrawn,

PD3 - Received wrong treatment or incorrect dose,

PD4 - Received excluded concomitant treatment.

7.2.3 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics will be summarized by treatment group and total using descriptive statistics. This will be done for the subjects in FAS and SAF.

Descriptive statistics for age, weight, body mass index (BMI), body surface area (BSA) and height at study entry will be presented. Frequency tabulations for sex, ethnicity, race, age group from IRT stratification, age group (<65, 65 – 74, >=75), EudraCT age category, baseline ECOG, baseline FLT3 mutation by type (TKD, ITD, TKD and ITD), cytogenetic risk status, secondary AML, and region will be presented.

Stratification factors (defined in SAP Section 6.4.6) will be tabulated by treatment group and total for FAS.

Number and percentage of subjects allocated to treatment in each country and site will be presented by treatment group and total for the SAF.

Medical history other than AML and conditions existing at Baseline will be coded in MedDRA and summarized by System Organ Class (SOC) and Preferred Term (PT) as well as by PT alone, by treatment group and total for the SAF and FAS. Baseline conditions are defined as those ongoing at the time of informed consent or arise following the time of informed consent and before the first dose of study drug. For ongoing medical conditions, Common Terminology Criteria for Adverse Events (CTCAE) grade will be provided in listing.

Frequency tabulations for targeted (AML) disease history including duration of AML disease in months, antecedent hematological disorder, type of hematological disorder, central nervous system leukemia, leukemia secondary to prior therapy, prior treatment for subjects with leukemia secondary to prior therapy, AML subtype as classified by World Health Organization (WHO, 2008) classification and French-American-British (FAB) classification, risk status with specific cytogenetic patterns will be presented by treatment group and total for the FAS and SAF.

Results from lumbar puncture, baseline extramedullary leukemia and MUGA scan, if performed, will be provided in listing.

All data will be presented in listings.

7.2.4 Previous and Concomitant Medications

Previous and concomitant medications are coded with World Health Organization – Drug Dictionary (WHO-DD), and will be summarized by therapeutic subgroup (Anatomical Therapeutic Chemical [ATC] 2nd level) and chemical subgroup (ATC 4th level) and

preferred WHO name for SAF. Summary will be sorted by alphabetical order by Therapeutic Subgroup and Chemical Subgroup and Preferred WHO name. In addition, concomitant medications will be summarized by preferred term alone. Summary will be presented by treatment group and total for the SAF in decreasing order of frequency based on the total number of subjects who took each medication and alphabetical order by preferred WHO name.

Subjects taking the same medication multiple times will be counted once per medication and investigational period. A medication which can be classified into several chemical and/or therapeutic subgroups is presented in all chemical and therapeutic subgroups.

All data will be presented in listings.

7.2.5 Previous and Concomitant Transfusions

Frequency tabulations of subjects who received transfusions and blood product will be presented for previous transfusion and concomitant transfusion by treatment group and total for SAF. Descriptive statistics will be presented for number of transfusion unit received per subject by treatment group and total for SAF.

All data will be presented in listings.

7.2.6 Previous and Concomitant Non-Medication Therapy

Frequency tabulations of subjects with previous and concomitant non-medication therapy and reason for use will be presented by treatment group and total for SAF. Number of previous non-medication therapy received per subject will be summarized by treatment group and total using descriptive statistics for SAF.

All data will be presented in listings.

7.2.7 Hospitalization Records

Under resource utilization, duration of hospitalization will be summarized by treatment group using descriptive statistics. Frequency tabulations of subjects with hospitalization, type of hospitalization and reason for hospitalization will be presented by treatment group.

All data will be presented in listings.

7.2.8 Clinical/Diagnostic Procedures

Clinical/diagnostic procedures will be presented in listings.

7.2.9 New AML Therapy

7.2.9.1 Post Treatment HSCT

Frequency tabulations of subjects with post treatment HSCT will be presented by treatment group. In addition, graft type and outcome of the transplant will be tabulated.

7.2.9.2 Subsequent Antileukemic Treatments and Outcomes

Frequency tabulations of subjects with subsequent antileukemic treatment, relapse prior to subsequent AML therapy, regiments, response to subsequent AML treatment and relapse

after new regimen will be presented by treatment group. Descriptive statistics will be presented for duration of subsequent AML therapy and duration of response to subsequent AML therapy if applicable.

All data will be presented in listings.

7.3 Study Drugs Exposure and Compliance

7.3.1 Exposure

The following information on drug exposure will be presented by actual treatment group for the SAF:

- Descriptive statistics for number of cycles, duration of exposure in days, number of dosing days, cumulative dose, average daily dose, dose intensity, planned intensity and relative dose intensity
- Number and percentage of subjects with dose increase, decrease or interruption
- Duration of exposure will be summarized in two ways:
    - Descriptive statistics will be presented by treatment group.
    - Exposure time will be categorized according to the following categories:
        - less than or equal to 7 days
        - at least 8 days, less than 28 days
        - at least 28 days, less than 84 days
        - at least 84 days, less than 168 days
        - 168 days or more
        - Unknown.

Counts and percentages of subjects in each of these categories will be summarized by treatment group for the SAF.

7.3.2 Treatment Compliance

Overall compliance with the dosing schedule will be examined for subjects in the SAF.

Percent of overall compliance will be summarized by treatment group in two ways:

- Descriptive statistics will be presented.
- Percent of compliance will be categorized according to the following categories:
    - less than 50%
    - at least 50%, less or equal to 75%
    - greater than 75%
    - Unknown.

All data will be presented in listings.

7.4 Analysis of Efficacy

As of 17Dec2020, study enrollment was terminated due to unfavorable of IA results. Details are provided in Section 7.11. No formal testing of OS will be performed for the final analysis.

7.4.1 Analysis of Primary Endpoint

Using the same notation of AC and C, as described in Section 4, the hypothesis for the primary endpoint (OS) is as follows:

H_{01}: OS of Arm AC is not different from that of Arm C
H_{11}: OS of Arm AC is different from that of Arm C

The primary efficacy endpoint of OS will be analyzed on the FAS for the randomized Arms AC and C using the stratified log-rank test with strata of age group per IRT, risk groups and FLT3 mutation status with potential of strata pooling (Section 7.12.4).

The interim analysis will occur when approximately 50% (i.e., death events = 70) of the planned total number of deaths (i.e., death events = 140) by any cause have occurred from the randomized Arm AC and Arm C. If Arm AC has favorable outcome (i.e., 2-sided P value < 0.003) or unfavorable outcome (i.e., 2-sided P value ≥ 0.724, non-binding futility boundary) compared to Arm C with respect to OS, the study may be stopped early for efficacy or futility respectively. Otherwise the study will continue without impact.

The final analysis will be performed after the planned 140 death events from the randomized Arm AC and Arm C have been observed. The hypothesis testing on OS will be performed at the overall 2-sided 0.049 significance level for efficacy.

The SAS code to implement stratified log-rank test will be similar to that shown below:

```
PROC LIFETEST;
    TIME time * status (1);
    STRATA stratification variables / group = treatment;
RUN;
```

Stratification variables are defined in Section 6.4.6 and strata pooling is defined in Section 7.12.4.

The 2-sided p-value for the above hypothesis test will be calculated and presented. The hazard ratio of the treatment effect along with 95% confidence interval will be calculated by the stratified Cox proportional hazard model with the same stratification variables.

The SAS codes to perform the Cox proportional hazard model with strata will be similar to that shown below:

```
PROC PHREG;
    Class treatment (ref='C') stratification variables;
    MODEL time * status (1) = treatment/RL;
```

STRATA *stratification variables*;

RUN;

The proportional hazards assumption will be evaluated graphically using the LOG(-LOG) survival function versus time plot and scaled Schoenfeld residuals by time plot, and by testing the interaction term between the treatment arm and LOG (time) in the Cox model. If the hazards appear to be non-proportional, piecewise Cox models will be used to explore the changes in hazard ratio over time.

Kaplan-Meier (KM) survival plots will be used to describe the OS in each treatment group. Median OS and 95% confidence interval (CI), survival rates at 6, 12, and 24 months and 95% CI will be estimated from the KM plots using the SAS code similar to that shown below:

PROC LIFETEST TIMELIST=(XX XX XX);

TIME time * status (1);

STRATA *stratification variables* / group = treatment;

RUN;

Where TIMELIST option contains timepoints of interest to estimate survival rates.

Median follow-up time and 95% CI will be obtained using the reverse KM method.

The sensitivity analyses for the primary efficacy endpoint will be performed as described below:

- With different analysis sets or different censoring on OS
  - The same analyses as primary analysis but on SAF
  - The same analyses as primary analysis on FAS, but with subjects who are censored at the time of HSCT as defined in Section 6.1.1
- With strata of age group
  - Similar analyses as primary analysis, but with strata of age group on the FAS
- Unstratified analyses
  - Similar analyses as primary analysis without stratification factors on FAS
- Additional sensitivity analysis will be performed using a test for equality of two survival functions based on weighted differences of Kaplan–Meier curves with estimation of difference of Restricted Mean Survival Time (RMST) and its 95% CI (Uno et.al, 2015) by a cutoff time when at least 10% of the randomized subjects are still in survival followup for both arms. The 10% criteria is determined to ensure that reasonable number of subjects are still at risk at the cutoff time to ensure a robust estimate of the mean survival time within this timeframe.
- Additional sensitivity analyses may be performed to assess the impact of COVID-19 and initiation of subsequent AML therapy

The same analyses as the above including the same strata will be repeated for comparison of ASP2215+AZA (Total) vs Arm C for the final analysis. No hypothesis testing of OS will be done for Arm A*. Descriptive summary statistics will be done for the OS data of Arm A*.

* Protocol versions 6.0 and earlier included a 1:1:1 randomization ratio to receive Arm A (Gilteritinib monotherapy), AC or C. Randomization to Arm A was removed in protocol version 7.0. Subjects previously randomized to Arm A should continue following treatment and assessments as outlined in the protocol.

7.4.2 Analysis of Secondary Endpoints

7.4.2.1 Key Secondary Efficacy Analysis

The key secondary efficacy endpoint of EFS will be analyzed on FAS using the same manner as primary endpoint (OS). The hypothesis for the key secondary endpoint is as follow.

- H_{02}: EFS of Arm AC is not different from that of Arm C
- H_{12}: EFS of Arm AC is different from that of Arm C

To control for overall type I error at the 2-sided 0.05 significance level, the hypothesis testing on EFS of the randomized Arm AC vs Arm C will be performed only if the null hypothesis on the primary analysis of OS is rejected at its corresponding significant level at interim analysis or final analysis. The rejection boundary of EFS will be based on the actual number of events at the interim and final analyses, and the group sequential testing method as described in Section 7.10.

The 2-sided p-value for the above hypothesis test will be calculated using stratified log-rank test with strata of age group per IRT, risk groups and FLT3 mutation status with potential of strata pooling (Section 7.12.4). The hazard ratio of the treatment effect along with 95% CI will be calculated by the stratified Cox proportional hazard model, and presented. KM survival plots will be used to describe the EFS in each treatment group. Median EFS and 95% CI, survival rates at 6, 12, and 24 months and 95% CI will be estimated, and presented.

The sensitivity analyses for the key secondary efficacy endpoint will be performed as described below:

- With different analysis sets or different censoring on EFS
  - Similar analyses as primary analyses for EFS on SAF
  - Similar analyses as primary analyses for EFS on FAS, with subjects who are censored at the time of HSCT
- With strata of age group
  - Similar analyses as primary analysis for EFS with strata of age group on FAS
- Unstratified analyses
  - Similar analyses as primary analysis without stratification factors FAS
- Similar analysis as primary analyses for EFS with different timing and censoring on the FAS
  - Defining EFS as the time from randomization to relapse from CR (for subjects who achieved CR), or death from any cause, whichever comes first

  - Defining EFS as setting time to treatment failure as the date of permanent discontinuation of all study treatment or the end of 6 cycles of therapy, whichever is earlier
- Similar analysis as primary analyses for EFS by defining relapse from CRc, treatment failure (failing to achieve CRc within 6 cycles of treatment) or death from any cause
- Additional sensitivity analyses may be performed to assess the impact of COVID-19 and initiation of subsequent AML therapy

The same analyses including the same strata will be repeated for comparison of ASP2215+AZA (Total) vs Arm C for the final analysis. No hypothesis testing of EFS will be done for Arm A. Descriptive summary statistics will be done for the EFS data of Arm A.

7.4.2.2 Other Secondary Efficacy Analysis

Age group per IRT will be used as the only stratification variable for other secondary efficacy analysis if applicable.

CR rate will be analyzed using the Cochran-Mantel-Haenszel (CMH) test to control of stratification variable (age group per IRT) on the FAS for randomized Arm AC and Arm C. The SAS code to implement this test will be similar to that shown below:

```
PROC FREQ;
    WHERE treatment in ('AC','C');
    TABLES stratification variables * response * treatment/cmh alpha=0.05;
RUN;
```

CR rate, stratified two-sided P value along with 2-sided exact 95% CI based on binomial distribution will be calculated for each treatment group using SAS code similar to that shown below:

```
PROC FREQ;
    TABLES response/BIONOMIAL EXACT alpha=0.05;
    BY treatment;
RUN;

PROC FREQ;
    WHERE treatment in ('AC','C');  TABLES response * treatment / EXACT alpha=0.05;
RUN ;
```

CRc rate, CR/CRh rate, CRh rate, transfusion conversion rate and transfusion maintenance rate may be analyzed in the same manner as CR rate.

Best overall response will be tabulated by responder (CR, CRp and CRi) versus non-responders (PR, NE and NR) by treatment.

LFS and duration of remission will be analyzed in the same manner as primary analyses for OS in the FAS, except for using age per IRT for stratified analysis. The hazard ratio of the treatment effect along with 95% CI will be calculated. KM curve will be created for each treatment group. Median time and 95% CI, LFS rates at Months 6, 12, and 24 if appropriate and 95% CI will be estimated from the KM curve.

Time to first and best CR/CRh will be summarized using descriptive statistics (mean, standard deviation, minimum, maximum and median) for subjects who achieved either CR or CRh in FAS.

Time to CR, CRc, response (CRc+PR) will be summarized by treatment group using descriptive statistics (mean, standard deviation, minimum, maximum and median) for subjects who achieved remission in the FAS.

BFI global fatigue score (average of all 9 items) and change from baseline will be summarized using descriptive statistics by treatment group at each visit on the FAS. Additionally, change from baseline will be calculated as the post-baseline value minus the baseline value and summarized in the same way. Change from baseline BFI score will be analyzed using Analysis of Covariance (ANCOVA) including treatment, age group per IRT as fixed factors, and baseline BFI score as covariate. LSmean difference in change from baseline and p-value will be presented using Arm C as reference.

The same analyses including only age group per IRT will be repeated for comparison of ASP2215+AZA (Total) vs Arm C for the final analysis. Only descriptive summary statistics will be done for Arm A.

7.4.3 Analysis of Exploratory Endpoints

FLT3 mutation status (TKD, ITD LAR, ITD HAR), FLT3 mutation type (FLT3-ITD alone, TKD (D835/I836) and ITD with TKD (D835/I836) will be summarized by the number and percentage of subjects in each category by treatment group at baseline and end of treatment on the FAS. Efficacy endpoint (OS, EFS, LFS, CR rate, CRc rate) and safety endpoint (TEAE, drug-related TEAE, lab parameter of interest) may be analyzed in subgroup based on baseline FLT3 mutation status and baseline FLT3 mutation status by type.

The individual mutation ratio will be calculated as the sum of all ITD mutation ratios at each visit for each subject. MRD mutation ratio will be summarized by treatment group by visit.

MRD negative is defined as the individual mutation ratio is less than 10^{-4} at a visit. The proportion of subjects with negative MRD will be analyzed using the Cochran-Mantel-Haenszel (CMH) with age group per IRT as stratification factor on FAS for Arm AC verse Arm C. MRD negative rate, stratified two-sided P value along with 2-sided exact 95% CI based on binomial distribution will be calculated for each treatment group by visit.

Transplantation rate, incidence of resource utilization counts including hospitalization, blood transfusion, antibiotic intravenous infusions, medication for AEs and opioid medication will be summarized by treatment group on the FAS. The difference between treatment groups may be tested using CMH test while controlling age group per IRT.

Resource utilization including durations of hospital stays, medications, blood transfusions, antibiotic intravenous infusions, medication for AEs and opioid medication will be summarized using descriptive statistics by treatment group on the FAS will summarize. The difference between treatment groups may be tested with ANOVA while controlling of age group per IRT as fixed factors.

The following, FACIT-Dys-SF (domain scores), FACT-Leu (general total score, the leukemia global score, subscale scores and the leukemia trial outcome index), dizziness and mouth sore items, and EQ-5D-5L (EQ-VAS and utility score), and their corresponding change from baseline will be summarized by treatment group using descriptive statistics. ANCOVA model will be used to analyze the change from baseline including treatment, age group per IRT as fixed factors, and baseline score as covariate. LSmean difference in change from baseline and p-value will be presented using Arm C as reference.

In addition, shift from baseline to each post-baseline visit will be provided for the 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) of EQ-5D-5L.

The same analyses will be repeated for comparison of ASP2215+AZA (Total) vs Arm C for the final analysis. Only descriptive summary statistics will be done for Arm A.

7.5 Analysis of Safety

Safety endpoints are AEs, laboratory tests, vital signs, ECGs and ECOG performance status.

All analyses of safety will be presented by treatment group for SAF, unless specified otherwise.

7.5.1 Adverse Events

Any adverse event (AE) recorded on treatment including within 30 days from the last study treatment will be classified as treatment-emergent AE (TEAE) and will be summarized.

Serious TEAEs include SAEs upgraded by the Sponsor based on review of the Sponsor's list of Always Serious terms if any upgrade was done.

The coding dictionary for this study will be MedDRA. It will be used to summarize AEs by SOC and PT. AEs will be graded using National Cancer Institute's Common Terminology Criteria for AEs (NCI-CTCAE).

An overview table to report the number and percentage of subjects and an overview table to report number of events and events adjusted by patient year from drug exposure will include the following details:

- TEAEs
- Drug-related TEAEs
    - ASP-related TEAEs
    - AZA-related TEAEs
- Serious TEAEs and Astellas upgraded serious TEAEs
- Drug-related serious TEAEs and Astellas upgraded drug-related serious TEAEs

  - ASP-related SAE
  - AZA-related SAE
- TEAEs leading to death
- Drug-related TEAEs leading to death
  - ASP-related TEAEs leading to death
  - AZA-related TEAEs leading to death
- TEAEs leading to withdrawal of treatment
  - TEAEs leading to withdrawal of ASP
  - TEAEs leading to withdrawal of AZA
- Drug-related TEAEs leading to withdrawal of treatment
  - ASP-related TEAEs leading to withdrawal of ASP
  - AZA-related TEAEs leading to withdrawal of AZA
- TEAEs leading to dose reduction
  - TEAEs leading to ASP reduction
  - TEAEs leading to AZA reduction
- Drug-related TEAE leading to dose reduction
  - ASP-related TEAEs leading to ASP reduction
  - AZA-related TEAEs leading to AZA reduction
- TEAEs leading to drug interruption
  - TEAEs leading to ASP interruption
  - TEAEs leading to AZA interruption
- Drug-related TEAE leading to drug interruption
  - ASP-related TEAEs leading to ASP interruption
  - AZA-related TEAEs leading to AZA interruption
- Grade 3 or higher TEAEs
- Drug-related Grade 3 or higher TEAEs
  - ASP-related Grade 3 or higher TEAEs
  - AZA-related Grade 3 or higher TEAEs
- Any deaths

The number and percentage of subjects with TEAEs and the number of events and events adjusted by patient year from drug exposure, as classified by SOC and PT will be summarized for each treatment group. Summaries will be provided for the list above.

The number and percentage of subjects with TEAEs, as classified by SOC and PT will also be summarized for each treatment group for the following:

- TEAEs excluding serious adverse events that equal to or exceed a threshold of 5% in any treatment group,
- Common TEAEs that equal to or exceed a threshold of 10% in any treatment group
- Drug-related common TEAEs that equal to or exceed a threshold of 10% in any treatment group
  - ASP-related common TEAEs that equal to or exceed a threshold of 10% in any treatment group

- AZA-related common TEAEs that equal to or exceed a threshold of 10% in any treatment group

The number and percentage of subjects with TEAEs, as classified by PT only, will be summarized for each treatment group for the following:

- TEAEs
- Drug-related TEAEs,
  - ASP-related TEAEs
  - AZA-related TEAEs

The number and percentage of subjects with adverse events of special safety interest (AESIs), as classified by AESIs category and PT, will be summarized by treatment group for the following:

- TEAEs with special safety interest
- Grade 3 or higher TEAEs with special safety interest
- Drug-related Grade 3 or higher TEAEs with special safety interest
  - ASP-related Grade 3 or higher TEAEs with special safety interest
  - AZA-related Grade 3 or higher TEAEs with special safety interest

AE summary tables will include subject counts as opposed to AE counts. If a subject experiences more than one episode of a particular AE, the subject will be counted only once for that AE. If a subject has more than one AE that codes to the same preferred term, the subject will be counted only once for that preferred term. Similarly, if a subject has more than one AE within a SOC, the subject will be counted only once in that SOC.

The number and percentage of subjects with TEAEs (serious TEAEs), as classified by SOC and PT will also be summarized by NCI-CTCAE severity grade and by relationship to study drug. In the subject count, if a subject has multiple TEAEs with the same SOC or PT, but with differing severity grade or relationship, then the subject will be counted only once with the worst severity grade and highest degree of relationship, however, if any of the severity grade or relationship values are missing then the subject will be counted only once with missing severity grade or relationship. Drug related TEAEs will be presented in a similar way by severity grade only. Serious TEAE will be presented in a similar way by relationship to study drug.

All AEs, deaths, SAEs and withdrawals due to adverse events will be displayed in listings.

DLT will be displayed in a listing for the safety cohort.

TEAE and drug related TEAE may be analyzed by subgroup of FLT3 mutation status at baseline.

7.5.2 Clinical Laboratory Evaluation

The baseline value is the last measurement taken prior to the first study drug administration.

Quantitative clinical laboratory variables, i.e., hematology, biochemistry will be summarized using mean, standard deviation, minimum, maximum and median for each treatment group at

each visit. Additionally, within-subject change from baseline will be calculated as the post-baseline value minus the baseline value and summarized in the same way. Each laboratory result will be classified as low (L), normal (N), or high (H) at each visit according to the laboratory supplied reference ranges.

The number and percentage of subjects below and above reference range will be summarized for each treatment group at each visit.

For hematology and biochemistry two types of shift tables will be presented:

- Shift tables of changes based on reference range defined category (low, normal, high) from baseline to each visit as well as worst finding during the treatment period, and
- Summary shifts of changes based on reference range defined category from baseline to each visit as well as worst finding during the treatment period (shift from high or normal to low, shift from low or normal to high, categorized increase [shift from low to normal or from normal to high], categorized no change [value stays in the same reference range], categorized decrease [shift from high to normal or from normal to low]).

Laboratory results will also be graded using NCI-CTCAE, where possible. Parameters that have criteria available for both low and high values, i.e., hypo- and hyper-, will be summarized for both criteria. The same subject can be counted for both values if the subject has different laboratory values meeting each criterion. NCI-CTCAE grade of laboratory evaluations will be summarized by number and percentage of subjects at each visit. Shift tables of NCI-CTCAE grade change from baseline to worst post-baseline grade will also be presented. The number and percentage of subjects with grade 3 or 4 laboratory test result will be summarized by treatment group and laboratory parameter (the name of the adverse event associated with the abnormal laboratory test result will be presented).

Laboratory results based on central assessment will be used for summaries as described above. Laboratory results based on local assessment and bone marrow results will be listed only. The laboratory results may be displayed in figures.

7.5.2.1 Liver Enzymes and Total Bilirubin

The following potentially clinically significant criteria in liver function tests for Alkaline Phosphatase (ALP), Alanine Transaminase (ALT), total bilirubin, Aspartate Transaminase (AST) and their combination are defined. The subject's highest value during the study will be used.

Parameter	Criteria
ALT	> 3xULN > 5xULN > 10xULN > 20xULN
AST	> 3xULN > 5xULN > 10xULN > 20xULN
ALT or AST	> 3xULN
Total Bilirubin	> 2xULN
ALP	> 1.5xULN
ALT and/or AST AND Total Bilirubin(*)	(ALT and/or AST > 3xULN) and (Total bilirubin > 2xULN)

(*) Combination of values measured within same sample

The number and percentage of subjects with potentially clinically significant values in liver enzymes and total bilirubin will be presented by treatment group.

7.5.3 Vital Signs

The baseline value is the last measurement taken prior to the first study drug administration.

Vital signs (systolic blood pressure [SBP], diastolic blood pressure [DBP], pulse rate, and body temperature) will be summarized using mean, standard deviation, minimum, maximum and median by treatment group and visit. Additionally, change from baseline will be calculated as the post-baseline value minus the baseline value and summarized by treatment group and visit.

Tables for potentially clinically significant vital signs will be generated using baseline value and highest post-baseline value for each subject for each treatment group. The results of vital signs may be displayed in figures.

The following potentially clinically significant criteria are defined:

Vital Sign Variable	**Criteria**
SBP	≥180 mmHg AND ≥20 mmHg change from baseline
DBP	≥105 mmHg AND ≥15 mmHg change from baseline
Pulse Rate	≥120 bpm AND ≥15 bpm change from baseline

7.5.4 Electrocardiograms (ECGs)

12-lead ECGs will be recorded in triplicate at the scheduled time points. Each ECG tracing will be taken at least 5 minutes apart and transmitted electronically to central reading. The

mean of the triplicate ECGs from central read should be used for all final treatment decisions, AE reporting and in the summary for analysis at each visit.

ECG variables including changes from baseline will be summarized using mean, standard deviation, minimum, maximum and median for each treatment group at each visit. The minimum value and maximum values among all post-baseline visits and corresponding changes will be summarized if appropriate.

Number and percentage of subjects with abnormal results as assessed by central read for the overall interpretation will be tabulated by treatment group at each visit. A shift analysis table showing shift in overall ECG interpretation from baseline to each visit will be provided. The worst of the three overall ECG interpretations will be used as the time-specific overall ECG interpretation for a subject.

The QT interval corrected for heart rate by Fridericia's formula, QTcF, is defined as: $QTc\ (F) = QT/(RR)^{0.33}$, where RR interval is inversely proportional to heart rate (approximately RR = 60/heart rate).

The QTcF interval will be summarized using frequency tables for each treatment group at each visit for values of clinical importance using the range criteria below.

	QTcF Interval Criteria Value (msec)	
	Cumulative Category	**Interval Category**
Normal	≤ 450	≤ 450
Borderline	> 450	> 450 to ≤ 480
Prolonged	> 480	> 480 to ≤ 500
Clinically significant	> 500	> 500

The QTcF interval will also be summarized by the frequency of subjects with a change from baseline of clinical importance using the criteria identified below. These summaries will be provided for each treatment group at each visit.

	Change from Baseline	
Variable	**Cumulative Category**	**Interval Category**
QTcF Interval (msec)	<0	<0
	≥ 0	≥ 0 to ≤ 30
	> 30	> 30 to ≤ 60
	> 60	> 60

The minimum and maximum values among all post-baseline visits and corresponding changes will be summarized by the categories defined above as appropriate.

7.5.5 Pregnancies

A detailed listing of all pregnancies will be provided.

7.5.6 Eastern Cooperative Oncology Group (ECOG) Performance Scores

Number of percent of subjects for each category of the ECOG performance scores at each assessment time will be provided by treatment group. Negative change scores indicate an improvement and positive scores indicate a decline in performance.

ECOG performance scores will also be summarized using shift table from baseline to post-baseline score for each treatment group.

7.6 Analysis of PK

Descriptive statistics include number of subjects, mean, standard deviation, coefficient of variation (CV), geometric mean, geometric CV, minimum, median and maximum. PK analysis will be conducted on the PKAS.

Plasma concentrations of ASP2215 will be presented by treatment group by visits and nominal time points using descriptive statistics.

For PKAS with ASP2215 C_{trough} values mean (SD) plasma concentration-time profiles (linear and semi-logarithm scales) for treatment arms A and AC on the same graph, and overlay (spaghetti) plots, separate for treatment arm A and AC, plasma concentration-time profiles will be produced (0-6 months)

For dense PK subgroup standard graphics including mean (SD) plasma concentration-time profiles (linear and semi-logarithm scales) and overlay (spaghetti) plots plasma concentration-time profiles will be produced for each analyte (ASP2215 treatment arms A and AC, azacitidine treatment arms C and AC on study Day 4, 0-6 hours).

7.6.1 Dense PK

For subjects participating in the dense PK subset, pharmacokinetic parameters will be estimated using noncompartmental methods. PK parameters include area under the curve (AUC_t), C_{max}, C_{trough} and t_{max} for both ASP2215 and azacitidine.

Plasma concentrations of ASP2215 and azacitidine will be summarized using descriptive statistics by treatment group and by visits and nominal time points for Japanese, Non-Japanese and overall. PK parameters will be summarized using descriptive statistics by treatment group for Japanese, Non-Japanese and overall. For t_{max}, only n, median, minimum and maximum will be presented.

Descriptive statistics for ASP2215 C_{trough} values from Cycle 1 Day 1 through Cycle 6 Day 1 will be provided for combined treatment arms A and AC.

If appropriate, to evaluate the effect of ASP2215 on the PK of azacitidine, an analysis of variance (ANOVA) model with treatment (Arm C / Arm AC) as fixed effects will be fitted on natural logarithm-transformed AUC_t and C_{max}.

To evaluate ethnicity differences, the same ANOVA model will be used by race (Japanese and non-Japanese), if applicable.

Within the ANOVA, the LS means of each treatment, LS means of differences between azacitidine in combination with ASP2215 (arm AC) and azacitidine alone (arm C), and 90% CIs for the differences will be estimated.

The LS means for AUCt and Cmax will be back-transformed to produce the geometric LS means. The geometric LS mean ratios and their corresponding 90% CIs for each pharmacokinetic parameter will be presented by back-transforming and expressed as percentages.If applicable, similar analysis may be used to evaluate the effect of azacitidine on the PK of ASP2215. Similar analysis by race may be used to evaluate ethnicity differences.

7.6.2 ASP2215 C_{trough} PK

Assessment of change in ASP2215 C_{trough} (120 mg dose) over time (C1D15-C6D1). Graphical analysis (box plot) with table of linear least squares regression analysis parameters (slope and intercept) along with summary statistics for the slope estiamate, from PKAS C_{trough} for participants with 4 or more values.

To assess the effect of treatment on the Ctrough of ASP2215, an analysis of variance (ANOVA) model with treatment (Arm A / Arm AC) as fixed effects will be fitted on natural logarithm-transformed Ctrough. Within the ANOVA, the least squares (LS) mean differences between azacitidine in combination with ASP2215 (arm AC) and ASP2215 alone (arm A) along with 90% CI on the differences will be estimated for Cycle 1 Day 15 and Cycle 2 Day 1 separately.

The LS means for Ctrough will be back transformed to produce the geometric LS means and presented with the number of participants for each treatment (Arm A / Arm AC). The geometric LS mean ratios and their corresponding 90% CIs for the pharmacokinetic parameter will be presented by back-transforming and expressed as percentages.

To assess the effect of ethnicity on the pharmacokinetics of ASP2215 Ctrough, an analysis of variance (ANOVA) model with ethnicity (Japanese / Non-Japanese) as fixed effects will be fitted on natural logarithm-transformed Ctrough . Within the ANOVA, the least squares (LS) mean differences between Non-Japanese and Japanese groups along with 90% CI on the differences will be estimated for Cycle 1 Day 15 and Cycle 2 Day 1 separately. The LS means for Ctrough will be back transformed to produce the geometric LS means and presented with the number of participants for each ethnic group (Japanese / Non-Japanese). The geometric LS mean ratios and their corresponding 90% CIs for the pharmacokinetic parameter will be presented by back-transforming and expressed as percentages.

These analyses will be performed separately for arm A, arm AC and both arms combined A+AC. Age and body weight may be used as covariates in the model if deemed appropriate.

A separate population pharmacokinetic analysis will be performed. Data from this study may be pooled with other studies for analysis. Details of this analysis will be specified in a separate population pharmacokinetic analysis plan.

7.7 Analysis of Pharmacodynamics

Not applicable.

7.8 PK-PD Analysis

Individual mean C_{trough} value is calculated as the mean C_{trough} of C1D15 and C2D1. ASP2215 C_{trough} group is defined based on median ASP2215 C_{trough} value. If individual subject's mean C_{trough} value of C1D15 and C2D1 is greater or equal to median ASP2215 C_{trough} value, then it's defined as high ASP2215 C_{trough}, otherwise it is defined as low ASP2215 C_{trough}.

Composite study day discontinuation (DPR) is defined as the earliest day of treatment discontinuation, first pause of ASP2215 ≥7 days, or first reduction in ASP2215 dose.

All analyses will be based on the first dose of actual treatment of ASP2215 120 mg group.

Demographics

Demographic and other baseline characteristics will be summarized by ASP2215 C_{trough} group using descriptive statistics.

Discontinuation

Graphical analysis of linear relationship of DPR and mean ASP2215 C_{trough} value of 120 mg dose group. Kaplan Meyer plot will be used to describe the probability of DPR by median ASP2215 C_{trough} group.

Efficacy

Graphical assessment (box plot) of clinical response (CRc, PR, NR, NE) by mean ASP2215 C_{trough} value for each patient.

Kaplan Meyer plot of OS by median ASP2215 C_{trough} group.

Safety

Graphical assessment (box plot with summary parameters) of neutropenia or thrombocytopenia (AE toxicity grade of 0-2, 3-4) by mean ASP2215 C_{trough} value.

7.9 Subgroups of Interest

Primary efficacy endpoint (OS) and key secondary endpoint (EFS) will be summarized by treatment group for the subgroups defined on the basis of the categorized variables listed below if the subgroup is deemed to have enough data. Other efficacy endpoints and treatment-emergent AEs may be summarized by all or selected subgroup variables. The subgroup analyses with modeling will not adjust any other factors besides treatment group. The percentage of subjects with events and the median time may be displayed for each subgroup for OS, EFS, and LFS. Absolute difference in the percentage of responders for Arm AC vs Arm C may be calculated for each subgroup without adjusting any other factors for CR, CR/CRh and CRc. The same analyses including the same strata will be repeated for comparison of ASP2215+AZA (Total) vs Arm C.

Grouping Variables	Subgroups
Age	≥ 75 years < 75 years
Age group 2	≥ 75 years [65, 75) years < 65 years
Sex	Female Male
Race	White Black or African American Asian Other
Baseline ECOG	0-1 ≥ 2
Region	North America Europe Asia/Pacific
Risk group	Favorable or intermediate Unfavorable cytogenetic risk or secondary AML
Baseline FLT3 Mutation Type	ITD alone ITD with TKD (D835/I836) TKD (D835/I836) alone
Baseline FLT3 Mutation Status	ITD LAR ITD HAR TKD

7.10 Other Analyses

Not applicable.

7.11 Interim Analysis (and Early Discontinuation of the Clinical Study)

A formal interim analysis for the randomized study is planned when approximately 50% (i.e., death events = 70) of the planned total number of deaths (i.e., death events = 140) by any cause have occurred from the randomized Arms AC and C.

The interim analysis will be conducted by the IDMC and inform Sponsor the result. If the randomized Arm AC has favorable outcome (i.e., 2- sided P value < 0.003) or unfavorable outcome (i.e., 2-sided P value ≥ 0.724) compared to Arm C with respect to OS, the study may be stopped due to efficacy or futility respectively. Otherwise, the study will continue without impact.

In addition, the EFS will be evaluated at the time of OS interim analysis, only if the OS result is positive at the interim analysis. By the time of OS interim analysis with 70 events, 88 EFS events are expected (the actual number of events may vary). An O'Brien-Fleming stopping boundary based on Lan-DeMets alpha spending method will be used for EFS. Based on a projected number of events of 88 at the interim, the efficacy stopping boundary is 2-sided nominal alpha of 0.003 for interim analysis and 0.049 for the final analysis. The actual rejection boundary for EFS may vary according to the actual number of EFS events that occur at the interim analysis.

The information above is summarized in the following table:

	Interim Analysis			Final Analysis	
	Number of events	2-sided P for futility	2-sided P for efficacy	Number of events	2-sided P for efficacy
OS	70	0.724	0.003	140	0.049
EFS	88	NA	0.003	176	0.049

If the study is not stopped after the interim analysis, a final analysis will occur after 100% of planned death events have been observed. If the trial is stopped early, provisions will be made for subjects who continue to derive benefit on their assigned treatment arm based on the investigator's assessment.

Production of tables and listings of the interim analysis will be performed by independent data analysis center (IDAC). No member of the clinical trial team will have access to the tables and listings created only for the IDMC. For more details consult the IDMC charter.

Safety data including AEs, clinical laboratory values, vital signs, ECG, and ECOG performance score will also be reviewed by the IDMC at the interim analysis. Other safety data reviews during the trial will be conducted by the IDMC on a periodic basis. The procedures for the IDMC safety review will be described in the IDMC charter.

Interim analysis was performed on 15Dec2020 based the first 70 death events. There was no significant difference in the median OS of Arm AC compared with that of Arm C. The 2-sided P value was 0.753, which crossed the predefined non-binding futility boundary of 0.724. The IDMC evaluated the analysis result of OS along with the safety data and recommended to stop enrolment for the study. The study was stopped for enrollment on 17 December 2020. No formal testing of OS will be performed for the final analysis.

7.12 Handling of Missing Data, Outliers, Visit Windows, and Other Information

7.12.1 Missing Data

Every effort will be made to resolve incomplete dates for death and disease relapse. If a partial date cannot be resolved, the most conservative imputation methods will be used to complete the missing information.

For primary endpoint OS, missing or incomplete death date will be imputed as the earliest feasible date on or after the date of last contact as the examples shown in the table below. The date of last contact will be obtained as described in Section 6.1.

Incomplete Date of Death (YYYY MMM DD)	**Date of Last Contact (YYYY MMM DD)**	**Imputed Date of Death (YYYY MMM DD)**
2005 APR ??	2005 MAR 31	2005 APR 01
2005 ??? 13	2005 MAR 31	2005 APR 13
2005 ??? ??	2005 MAR 31	2005 MAR 31
???? APR ??	2005 MAR 31	2005 APR 01
???? APR 13	2005 MAR 31	2005 APR 13
???? ??? ??	2005 MAR 31	2005 MAR 31

Partial relapse dates will be imputed to the first day of the month of the missing parameter but not earlier than the last disease assessment date. A month and year must be present or the date will remain missing.

Missing centrally evaluated bone marrow assessment will be imputed with local bone marrow assessment as described in Section 6.1.2.3. Non-responder imputation will be used for binary response variables.

Missing or partial start and stop dates of adverse events and concomitant medication will be imputed using the following algorithm:

- Imputation rules for partial or missing stop dates:
  - If the month and year are present, then impute as the last day of that month.
  - If only the year is present, impute as December 31 of that year.
  - If the stop date is entirely missing, assume the event or medication is ongoing.
- Imputation rules for partial or missing start dates:

Start Date		**Stop Date**						
		Complete: *yyyymmdd*		**Partial: *yyyymm***		**Partial: *yyyy***		**missing**
		< 1st dose	**≥ 1st dose**	**< 1st dose *yyyymm***	**≥ 1st dose *yyyymm***	**< 1st dose *yyyy***	**≥ 1st dose *yyyy***	
Partial: *yyyymm*	= 1st dose *yyyymm*	2	1	2	1	n/a	1	1
	≠ 1st dose *yyyymm*		2		2	2	2	2
Partial: *yyyy*	= 1st dose *yyyy*	3	1	3	1	n/a	1	1
	≠ 1st dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1 = Impute as the date of first dose; 2 = Impute as the first of the month; 3 = Impute as January 1 of the year; 4 = Impute as January 1 of the stop year

The imputed dates will be used to determine whether an AE is/is not treatment emergent. Listings of AEs and concomitant medications will present the actual partial dates; imputed dates will not be shown.

In the case of partial date of initial diagnosis of AML, the date will be imputed to the first day of the month. A month and year must be present or the date will remain missing.

In the case of partial starting date of subsequent AML therapy, the date will be imputed to the first day of the month but not earlier than the last dosing date of the study drug. A month and year must be present or the date will remain missing.

Concentrations below the lower limit of quantification (BQL) in PK should be treated as missing when the terminal elimination rate constant is evaluated. Otherwise BQL should be treated as zero in the estimation of individual pharmacokinetic parameters.

7.12.2 Outliers

All values will be included in the analyses.

7.12.3 Visit Windows

Visit windows are allowed for certain visits per the schedule of assessments. Subject data will not be excluded from analyses due to the subject's failure to comply with the visit schedule. CRF visit will be used for analysis. In the case of multiple observations at a specific visit, the observation which is closest to the target date will be used. If the observations have the same distance to the target visit, the latest one will be used. If more than one observation is made on the same day, an average value if continuous or the worst value if categorical will be included in the analysis.

The visit windows for assessments are described in the following table.

eCRF visit	**Visit Window**
Screening	Day-14 – Day -1
Cycle 1 Day 4	C1D4 ± 1
Cycle 1 Day 8	C1D8 ± 1
Cycle 1 Day 15	No Window
Cycle 2 Day 1	C2D1 ± 2
Cycle 2 Day 15	C2D15 ± 1
Cycle X Day 1	CXD1 ± 2
End of Treatment Visit	Last dose date + 7

Scheduled visit will be calculated using number of days relative to the first dose date based on 28-day cycles. In the case of multiple observations at a specific visit, the observation which is closest to the target date will be used. If the observations have the same distance to the target visit, the latest one will be used.

7.12.4 Pooling Strata

In the primary (OS) and key secondary efficacy (EFS) analyses, if all the events in one stratum combination are from one treatment group or the Cox proportional hazard model

does not converge due to small event size, this stratum combination may be pooled in the order of FLT3 mutation status (collapsing TKD and LAR together first and then remove this factor), risk group (remove this factor) and age group (remove this factor) until the issue is resolved or the normal (un-stratified) Cox proportional hazard model is applied. All the stratification factors can be found in Section 6.4.6.

7.12.5 Blinding

Although the study is an open label study, to increase the credibility of study results, the sponsor statistician's access to the randomized treatment assignment information will be limited except for independent statistician. This will reduce potential bias due to the sponsor knowing the treatment effect due to unintentional efficacy and safety summary by treatment. On the other hand, the clinical data should be used appropriately for clinical operation, data cleaning and generating statistical programs. Thus, we will follow the procedures specified below:

- The randomized treatment code will not be transferred to statistical programmers, supporting study statisticians and study statisticians before the database lock. Instead, data sets with scrambled treatment code will be transferred for preparing analysis programs.
- The study statisticians, supporting statisticians and statistical programmers will have no access to the randomized treatment before the database lock.
- Study manager and other study team members may have the access to the treatment information at the individual subject level.
- No treatment difference will be summarized during the study, except for outputs for safety review by IDMC and planned interim analysis or requested analyses by IDMC and performed by IDAC.

8 DOCUMENT REVISION HISTORY

Version	Date	Changes	Comment/rationale for change
1.0	05-27-2016	NA	
2.0	11 Nov 2020	1. Incorporate changes in Protocol Amendments 2. Incorporate responses to the FDA information requests	1. To remove ophthalmologic assessment; to add derivation algorithm for PRO endpoints; to reflect remove treatment Arm A and change randomization ratio to 2:1; to update HR assumption to 0.6 and update sample size to 250 with 70 OS events in the interim analysis and 140 OS events in the final analysis; to include EFS in the interim analysis if OS result is positive; to include additional stratification factor and pooling strategy for strata; to update analysis for exploratory endpoints. 2. To modify EFS definition and include additional sensitivity analyses of EFS endpoint.
3.0	15 Jun 2022	1. Update the PK analysis 2. Add PK-PD analysis 3. Add futility stopping information from the interim analysis	1. To provide details of PK analysis 2. To include PK-PD analysis section 3. To show no formal testing for the final analysis

9 REFERENCES

Bretz F, Posch M, Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedure using weighted Bonferroni, Simes, or parametric tests. Biom J. 2011;53(6):894-913.

Cella D, Jensen SE, Webster K, Hongyan D, Lai JS, Rosen S, et al. Measuring health-related quality of life in leukemia: the Functional Assessment of Cancer Therapy--Leukemia (FACT-Leu) questionnaire. Value Health. 2012;15:1051-8.

Cheson BD, Bennett JM, Willman CL, et al. 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):4642-4649.

Choi SW, Victorson DE, Yount S, Anton S, Cella D. Development of a conceptual framework and calibrated item banks to measure patient-reported dyspnea severity and related functional limitations. Value Health. 2011;14:291-306.

ICH Harmonized Tripartite Guideline E 3. Structure and Content of Clinical Study Reports, November 1995. (www.ich.org; Guidelines; "Efficacy" Topics)

ICH Harmonized Tripartite Guideline E 9. Statistical Principles for Clinical Trials, February 1998. (www.ich.org; Guidelines; "Efficacy" Topics)

Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika. 1983;70:659-63.

Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, et al. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. Cancer. 1999;85:1186-96.

Sievers EL, Larson RA, Stadtmauer EA, et al. Efficacy and safety of Gemtuzumab Ozogamicin in Patients with CD33-Positive Acute Myeloid Leukemia in First Relapse. J Clin Oncol. 2001;19:3244-3254.

Hajime Uno, Lu Tian, Brian Claggettc and L. J.Wei A versatile test for equality of two survival functions based on weighted differences of Kaplan–Meier curves. Statistics in Medicine 2015

Hajime Uno, Janet Wittes, Haoda Fu, Scott D. Solomon, Brian Claggett, Lu Tian, Tianxi Cai, Marc A. Pfeffer, Scott R. Evans and Lee-Jen Wei. Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies. Annals of Internal Medicine 2015. 163(2), 127–12.

10 APPENDICES

10.1 Appendix 1: Key Contributors and Approvers

List of Key Contributors and Approvers

Key Contributors

The following contributed to or reviewed this Statistical Analysis Plan as relevant to their indicated discipline or role.

Primary author (s)

PPD	Oncology Biostatistics Data Science Astellas Pharma Global Development, Inc.

Contributors and Reviewers

PPD	Oncology Biostatistics Data Science Astellas Pharma Global Development, Inc.
PPD	Development Therapeutic Area - Oncology Medical Science Astellas Pharma Global Development, Inc.
PPD	Development Therapeutic Area - Oncology Medical Science Astellas Pharma Global Development, Inc.
PPD	Pharmacovigilance Medical Safety Astellas Pharma Global Development, Inc.
PPD	Clinical Sciences Astellas Pharma Global Development, Inc.
PPD	Clinical Pharmacology & Exploratory Development Astellas Pharma Global Development, Inc.
PPD	Clinical Pharmacology & Exploratory Development Astellas Pharma Global Development, Inc.

PPD	HEOR Astellas Pharma Global Development, Inc
PPD	Medical Writing Astellas Pharma Global Development, Inc

Author and Approver Signatories

(E-signatures are attached at end of document)

PPD , Oncology Biostatistics/Global Data Science was the study statistician for this study and the primary author of this Statistical Analysis Plan

PPD , Oncology Biostatistics/Global Data Science was the global statistical lead (GSTATL) for this project and biostatistics peer reviewer of this Statistical Analysis Plan

This Statistical Analysis Plan was approved by:

PPD
PPD , Development Therapeutic Area - Oncology/Global Medical Science