STATISTICAL ANALYSIS PLAN

Final Version 4.0, dated 31-MAY-2022

A Phase 3 Open-label, Multicenter, Randomized Study of ASP2215 versus Salvage Chemotherapy in Patients with Relapsed or Refractory Acute Myeloid Leukemia (AML) with FLT3 Mutation

ISN/Protocol: 2215-CL-0301
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
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
aPTT	Activated Partial Thromboplastin Time
ASCM	Analysis Set Classification Meeting
ASP2215	Astellas Compound code for 2215
AST	Aspartate Transaminase
ATC	Anatomical Therapeutic Chemical classification system
BFI	Brief Fatigue Inventory
BQL	Blow Quantification Limit
CI	Confidence Intervals
CMH	Cochran-Mantel-Haenszel
CR	Complete Remission
CRc	Composite Complete Remission
CRh	Complete Remission with partial hematological recovery
CRi	Complete Remission with incomplete hematological recovery
CRp	Complete Remission with incomplete platelet recovery
CRF	Case Report Form
CS	Classification Specifications
CSR	Clinical Study Report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DBP	Diastolic Blood Pressure
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EDTA	Ethylenediaminetetraacetic acid
EFS	Event-Free Survival
EQ-5D-5L	EuroQol Group-5 Dimension-5 Level Instrument
FAB	French-American-British
FACIT-Dys-SF	Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms

Abbreviations	Description of abbreviations
FACT-Leu	Functional Assessment of Cancer Therapy-Leukemia
FAS	Full Analysis Set
FLAG-IDA	Fludarabine, cytarabine and granulocyte colony-stimulating factor with idarubicin
FLT3	FMS-like TyrosineKkinase
GD	Global Development
GVHD	Graft-versus-host disease
H	High
HSCT	Hematopoietic Stem Cell Transplant
IAP	Interim Analysis Plan
ICF	Informed Consent Form
ICH	International Conference on Harmonization
IDMC	Independent Data Monitoring Committee
IND	Investigational New Drug
INR	International Normalization Ratio
IRT	Interactive Response Technology
ISN	International Study Number
ITD	Internal Tandem Duplication
ITT	Intention to Treatment Set
IU/L	International units/liter
IV	Intravenous
L	Low
LFS	Leukemia-Free Survival
MEC	Mitoxantrone, etoposide and intermediate-dose cytarabine
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligrams
Min	Minute
mL	Milliliter
mmHg	millimeters of mercury
MMRM	Mixed-Effect Model Repeated Measure
mRAS	Modified Response Analysis Set
msec	milliseconds
MUGA	Multigated acquisition scan
N	Number
N	Normal
NCI	National Cancer Institute
NE	Not evaluable
NR	No Response
NYHA	New York Heart Association
OS	Overall Survival
PD	Protocol Deviation
PGx	Pharmacogenetics

Abbreviations	Description of abbreviations
PK	Pharmacokinetic
PKAS	Pharmacokinetic Analysis Set
PPS	Per Protocol Set
PR	Partial Remission
PRO	Patient reported outcome
PT	Preferred Term
PTT	Partial Thromboplastin Time
QD	quaque die, a Latin phrase meaning “every day”
QTc	QT interval corrected for heart rate
QTcF	QT interval corrected for heart rate using Fridericia’s factor
RAS	Response Analysis Set
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
SOP	Standard Operating Procedure
TEAE	Treatment Emergent Adverse Event
TKD	Tyrosine kinase domain
TLF	Tables, Listings and Figures
ULN	Upper limit of normal
VAS	Visual analogue scale
WBC	White Blood Cell
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 starting point 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 randomization.
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 signed consent but did not meet 1 or more criteria required for participation in a trial and did not randomize to the trial.
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 of the primary and secondary endpoints and other data.

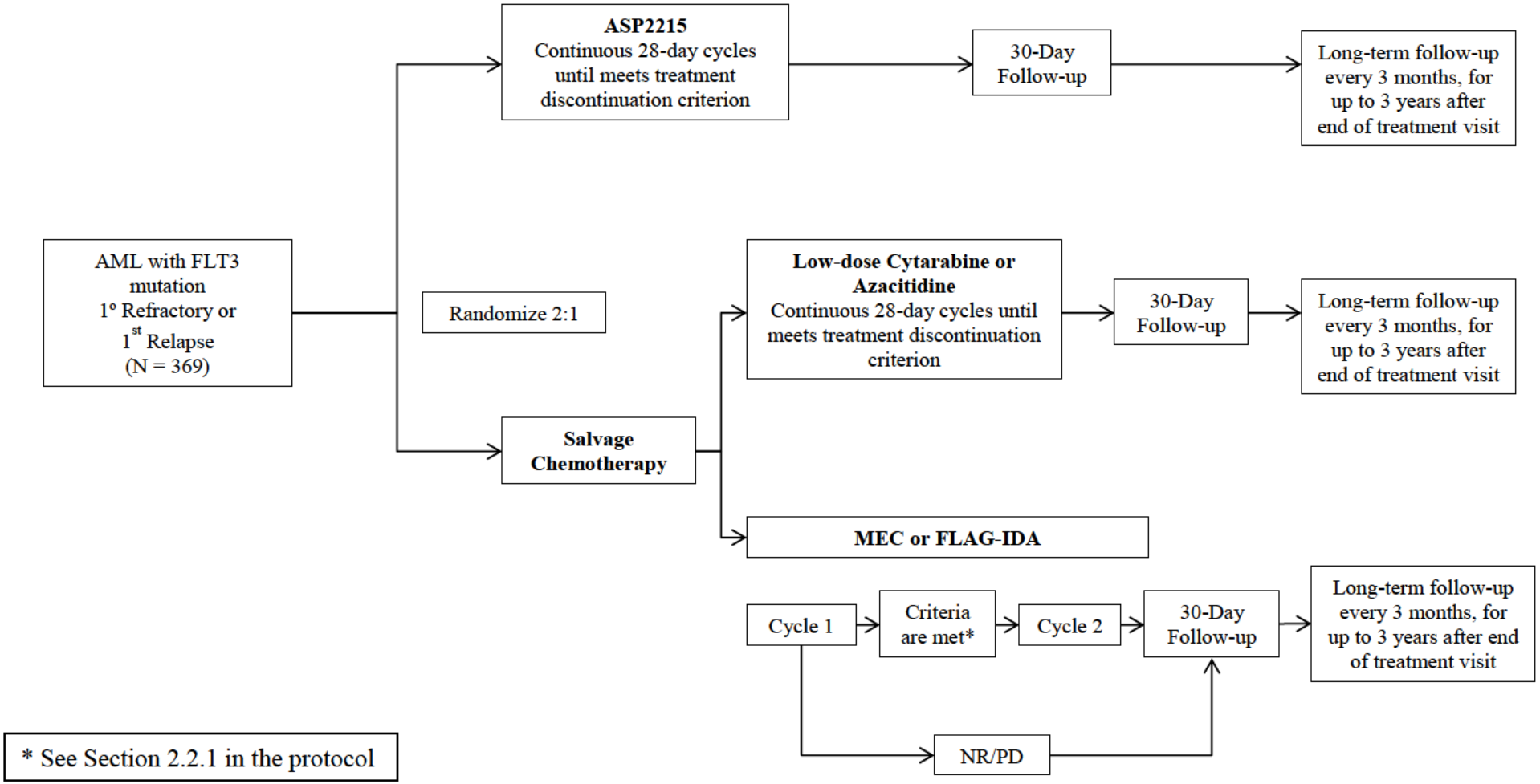
The SAP is finalized and signed prior to any of the following: study unblinding, database hard lock, interim analysis, or accumulation of substantial amount of data in an open-label study to ensure lack of bias. For operational efficiency an earlier time is usually targeted and wherever possible, the SAP should be developed in parallel with protocol finalization. For Phase 2-4 studies the SAP should be developed and approved before First Subject In (FSI). If the expected interval between FSI and soft-lock is less than 12 weeks, then the SAP should be approved by 12 weeks prior to the planned date of soft-lock. If needed, revisions to the approved SAP may be made prior to database hard lock. Revisions will be version controlled.

This statistical analysis is coordinated by the responsible biostatistician of GD-US. Any changes from the analyses planned in the SAP will be justified in the Clinical Study Report (CSR).

SAP version 3.0 dated 27 SEP 2018 has described the final analysis of primary, secondary, and exploratory efficacy endpoints with the data cutoff date of 17 SEP 2018. There will be no efficacy analysis performed for the end-of-study report. The End of Study analysis will be performed on safety data only and COVID-19 related data will be listed.

2 FLOW CHART AND VISIT SCHEDULE

Flow Chart



1°: primary; AML: acute myeloid leukemia; FLT3: FMS-like tyrosine kinase; FLAG-IDA: fludarabine, cytarabine and granulocyte colony-stimulating factor with idarubicin; MEC: mitoxantrone, etoposide and intermediate-dose cytarabine; NR: no response; PD: progressive disease

Table 1 Schedule of Assessments for ASP2215 Arm

Activity	Screening (Day -14 to -1)	Cycle 1					Cycle 2		Subsequent Cycles
		Day 1	Day 4 ± 1	Day 8 ± 1	Day 9	Day 15 ± 1	Day 1 ± 2	Day 15 ± 1	Day 1 ± 2
Signed ICF	X								
Medical and Disease History	X								
Randomization		X[p]							
Physical Examination[b]	X	X[a]	X	X		X	X	X	X
Vital Signs	X	X[a]	X	X		X	X	X	X
ECOG Performance	X	X[a]				X	X	X	X
Prior and Concomitant Medications	X[c]	X	X	X		X	X	X	X
Pregnancy Test for Woman of Childbearing Potential	X[d]	X					X		X
Chest X-ray (or CT of chest)[o]	X								
12-lead ECG[e]	X[g]	X		X[s]	X[s]	X	X		X
Clinical Laboratory Tests (chemistry, hematology, coagulation, urinalysis)[f]	X[g]	X[a]	X[a]	X[a]		X[a]	X[a]	X[a]	X[a]
Thyroid Function Test[t]	X								X[t]
Coagulation Profile (PT/INR, D-dimer, fibrinogen)	X								
MUGA or ECHO[h]	X								
Ophthalmologic Assessment[i]	X						X		X
FLT3 Mutation Status[j] (bone marrow aspirate or whole blood)	X								
Bone Marrow Aspiration and/or Biopsy	X[k]						X[k]		X[k]
AE/SAE Assessment	X	X	X	X		X	X	X	X
PK (whole blood samples for plasma PK)		X[l]		X[l]		X[l]	X[l]		X[l]
PGx[m]		X							
Patient Reported Outcome Tools[q, r]		X[a]		X[q]		X[q]	X	X[q]	X
EQ-5D 5L[r]		X[a]					X		X
Resource Utilization		X[a]					X		X
IRT Transaction Required[p]	X	X					X		X
ASP2215 Dosing at the Clinic[n]		X	X	X		X	X	X	X

AE: adverse event; CR: complete remission; CRc: composite complete remission; CRi: complete remission with incomplete hematologic recovery; CRp: complete remission with incomplete platelet recovery; CT: computed tomography; ECG: electrocardiogram; ECHO: echocardiogram; ECOG: Eastern Cooperative Oncology Group; EDTA: ethylenediaminetetraacetic acid; EQ-5D-5L: EuroQol Group-5 Dimension-5 Level Instrument;

FLT3: FMS-like tyrosine kinase; ICF: Informed Consent Form; INR: international normalization ratio; IRT: interactive response technology; MUGA: multigated acquisition scan; PGx: pharmacogenomics; PK: pharmacokinetic; PT: prothrombin time; SAE: serious adverse event

Footnotes continued on next page

a. Obtained predose.

b. Height measurement performed only at screening. Weight measurement should be performed at screening and day 1 of each cycle.

c. Includes medications taken within 28 days prior to cycle 1 day 1.

d. Woman of childbearing potential 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.

e. Screening ECG is required. ECG assessment will be evaluated at predose of cycle 1 day 1, cycle 1 day 8, cycle 1 day 15 and day 1 of each subsequent cycle. Predose assessments should be taken within 1 hour before drug administration. The 12-lead ECGs will be recorded in triplicate (3 separate ECGs 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 final treatment decisions and AE reporting. If the mean of the triplicate QTcF is > 500 ms at any time point (by either value on ECG tracing printout or central reading), then triplicate ECGs will be repeated (within 2 hours if based on value on ECG tracing printout and as soon as possible if based on central reading). If the repeat ECG confirms a mean of the triplicate QTcF > 500 ms, dosing of ASP2215 will be interrupted for up to 14 days. While ASP2215 may be interrupted temporarily based on value on ECG tracing printout, the central reading should be used for final treatment decisions. Cardiology consult will be obtained as medically indicated. If QTcF resolves to ≤ 480 ms (grade 1 or less) by central reading within 14 days, the subject may resume dosing at the reduced dose.

f. Urinalysis only required at screening. Uric acid will be tested on days 1, 4, 8 and 15 in cycle 1. Additional laboratory tests should be performed according to institutional standard of care.

g. Subjects may be screened and randomized from local labs only. However, samples must also be submitted for central read. Labs and/or ECG can be repeated during screening period.

h. MUGA scans or ECHO (per standard of care) are to be performed at screening for subjects with history of congestive heart failure New York Heart Association Class 3 or 4 (unless MUGA scans or ECHO performed either within 1 month prior revealed left ventricular ejection fraction ≥ 45%).

i. Ophthalmologic assessment to be performed by visual acuity measurement and ophthalmoscopy during the screening period, day 1 (± 7 days) of cycle 2, day 1 (± 7 days) of every 2 cycles thereafter, and when clinically indicated. In symptomatic subjects, the ophthalmologic assessment should also include slit lamp biomicroscopy, visual fields performed by Humphrey method and optical coherence tomography.

j. FLT3 mutation status will be assessed from bone marrow sample taken at the screening visit. If bone marrow sample is unavailable (e.g., dry tap), the whole blood sample taken at the screening visit will be used. Subjects must be screened by the central laboratory. All subjects including those with rapidly proliferative disease must have screening sample sent to the central lab. If central result is negative, central FLT3 testing can be repeated during screening period.

k. Bone marrow samples are required during screening, cycle 2 day 1 and cycle 3 day 1. For subjects who do not achieve a CRc (CR, CRp or CRi), the bone marrow assessments will be repeated at day 1 of every 2 subsequent cycles. For subjects who achieve a CRc (CR, CRp or CRi), bone marrow sampling will be repeated on 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 pre-HSCT visit/end of treatment visit and as clinically indicated. If bone marrow aspirate is unobtainable (e.g., dry tap), an additional EDTA tube of whole blood should be collected instead. Bone marrow aspirate is required, and bone marrow biopsy is preferred. In case of inadequate aspirate, bone marrow biopsy is required.

l. PK samples for ASP2215 will be collected on cycle 1 day 1 predose, cycle 1 day 8 predose and at cycle 1 day 15 and day 1 predose of each subsequent cycle (1 hour before drug administration). See Section 7.6 of the protocol.

m. Whole blood and buccal swab collected at day 1 for optional pharmacogenomic study.

n. ASP2215 is taken daily at home except for clinic days when it will be taken at the clinic.

Footnotes continued on next page

o. Chest X-ray (or CT of chest) does not need to be repeated if performed within 2 weeks prior to start of screening.

p. For the purposes of drug preparation and dispensing activities, IRT transaction may be done prior to the visit and do not need to fall within the protocol visit window.

q. Includes Brief Fatigue Inventory, Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items. The BFI will be administered at cycle 1 day 1 predose, cycle 1 day 8 (± 1 day), day 15(± 1 day), cycle 2 day 1 (± 2 days), day 15(± 1 day) and all subsequent cycles day 1 (± 2 days). Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items will be administered at cycle 1 day 1 predose, cycle 2 day 1 (± 2 days) and all subsequent cycles day 1 (± 2 days).

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

s. A cycle 1 day 8 ECG will be taken and the central read results will be provided to the site 24 hours after receipt of the tracing. A confirmatory ECG should be performed on cycle 1 day 9 if the mean QTcF from cycle 1 day 1 to cycle 1 day 8 has increased > 30 ms with no other known etiology, based on the central read ECG. On cycle 1 day 8, 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 cycle 1 day 9 visit is scheduled later in the day in order to allow for receipt and assessment of the cycle 1 day 8 central read ECG. This also allows for a subject to be contacted if the cycle 1 day 9 ECG is no longer required. If the cycle 1 day 9 ECG is still required, the result of the central read ECG will be received on cycle 1 day 10, in which the investigator should assess if the ASP2215 dose modification should occur as per the dose interruption or reduction guideline in Section 5.1.2 of the protocol.

t. Thyroid function tests will be repeated after every 2 cycles of therapy (C3D1, C5D1, C7D1, etc.).

Table 2 Schedule of Assessments for Chemotherapy Arm

Activity	Screening (Day -14 to -1)	Cycle 1				Cycle 2		Subsequent Cycles
		Day 1	Day 4 ± 1	Day 8 ± 1	Day 15 ± 1	Day 1 ± 2	Day 15 ± 1	Day 1 ± 2
Signed ICF	X							
Medical and Disease History	X							
Randomization		X[p]						
Physical Examination[b]	X	X[a]	X	X	X	X	X	X
Vital Signs	X	X[a]	X	X	X	X	X	X
ECOG Performance	X	X[a]			X	X	X	X
Prior and Concomitant Medications	X[c]	X	X	X	X	X	X	X
Pregnancy Test for Woman of Childbearing Potential	X[d]	X				X		X
Chest X-ray (or CT of chest)[o]	X							
12-lead ECG[e]	X[g]	X			X	X		X
Clinical Laboratory Tests (chemistry, hematology, coagulation, urinalysis)[f]	X[g]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]
Thyroid Function Test[s]	X							X[s]
Coagulation Profile (PT/INR, D-dimer, fibrinogen)	X							
MUGA or ECHO[h]	X							
Ophthalmologic Assessment	X[i]					X[i]		X[i]
FLT3 Mutation Status[j] (bone marrow aspirate or whole blood)	X							
Bone Marrow Aspiration and/or Biopsy	X[k]				X[k]	X[k]		X[k]
AE/SAE Assessment	X	X	X	X	X	X	X	X
PGx[l]		X						
Patient Reported Outcome Tools[q, r]		X[a]		X[q]	X[q]	X	X[q]	X
EQ-5D 5L[r]		X[a]				X		X
Resource Utilization		X[a]				X		X
IRT Transaction Required[p]	X	X				X		X
LoDAC or Azacitidine Dosing		See Footnote [m]						
MEC or FLAG-IDA Dosing		See Footnote[n]						

AE: adverse event; CR: complete remission; CRc: composite complete remission; CRi: complete remission with incomplete hematologic recovery; CRp: complete remission with incomplete platelet recovery; CT: computed tomography; ECG: electrocardiogram; ECHO: echocardiogram; ECOG: Eastern Cooperative Oncology Group; EDTA: ethylenediaminetetraacetic acid; EQ-5D-5L: EuroQol Group-5 Dimension-5 Level Instrument; FLAG-IDA: fludarabine, cytarabine and granulocyte colony-stimulating factor with idarubicin; FLT3: FMS-like tyrosine kinase; ICF: Informed Consent Form; INR: international normalization ratio; IRT: interactive response technology; LoDAC: low-dose cytarabine; MEC: mitoxantrone, etoposide and intermediate-dose cytarabine; MUGA: multigated acquisition scan; PGx: pharmacogenomics; PT: prothrombin time; SAE: serious adverse event

Footnotes continued on next page

a. Obtained predose.
b. Height measurement performed only at screening. Weight measurement should be performed at screening and day 1 of each cycle.
c. Includes medications taken within 28 days prior to cycle 1 day 1.
d. Woman of childbearing potential 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.
e. Screening ECG is required. ECG assessment will be evaluated at predose of cycle 1 day 1, cycle 1 day 15 and day 1 each subsequent cycle. Predose assessments should be taken within 1 hour before drug administration. The 12-lead ECGs will be recorded in triplicate (3 separate ECGs 10 minutes resting prior to first ECG and at least 5 minutes apart per time point) and transmitted electronically for central reading. See Section 7.5.5 of the protocol.
f. Urinalysis only required at screening. Uric acid will be tested on days 1, 4, 8 and 15 in cycle 1. Additional laboratory tests should be performed according to institutional standard of care.
g. Subjects may be screened and randomized from local labs only. However, samples must also be submitted for central read. Labs and/or ECG can be repeated during Screening period.
h. MUGA scans or ECHO (as per standard of care) are to be performed at screening for subjects with history of congestive heart failure New York Heart Association Class 3 or 4 (unless MUGA scans or ECHO performed either within 1 month prior revealed left ventricular ejection fraction ≥ 45%).
i. Ophthalmologic assessment to be performed by visual acuity measurement and ophthalmoscopy during the screening period, day 1 (± 7 days) of cycle 2, day 1 (± 7 days) of every 2 cycles thereafter, and when clinically indicated. In symptomatic subjects, the ophthalmologic assessment should also include slit lamp biomicroscopy, visual fields performed by Humphrey method and optical coherence tomography.
j. FLT3 mutation status will be assessed from bone marrow sample taken at the screening visit. If bone marrow sample is unavailable (e.g., dry tap), the whole blood sample taken at the screening visit will be used. Subjects must be screened by the central lab. All subjects including those with rapidly proliferative disease must have screening sample sent to the central lab. If central result is negative, central FLT3 testing can be repeated during screening period.
k. For MEC and FLAG-IDA, bone marrow samples are required during screening and cycle 2 day 1. Also, an additional bone marrow sample is required at cycle 1 day 15 or later, per institutional guidelines, to assess the need for a second cycle. For LoDAC or azacitidine, bone marrow samples are required during screening, and at cycle 2 day 1 and at cycle 3 day 1. For subjects who do not achieve a CRc (CR, CRp or CRi), the bone marrow assessments will be repeated at day 1 of every 2 subsequent cycles. For subjects who achieve a CRc (CR, CRp or CRi), bone marrow sampling will be repeated at 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 unobtainable (e.g., dry tap), an additional EDTA tube of peripheral blood should be collected instead. Bone marrow aspirate is required, and bone marrow biopsy is preferred. In case of inadequate aspirate, bone marrow biopsy is required.
l. Whole blood and buccal swab collected at day 1 for optional pharmacogenomic study.
m. LoDAC or azacitidine dosing may continue past cycle 2.
n. Additional clinic visits are allowed per institutional guideline for subjects receiving MEC (days 1 through 5) or FLAG-IDA (days 1 through 6). MEC and FLAG-IDA are administered for up to 2 cycles depending on response and safety assessments as described in Section 5.1 of the protocol.
o. Chest X-ray (or CT of chest) does not need to be repeated if performed within 2 weeks prior to start of screening.
p. For the purposes of drug preparation and dispensing activities, IRT transaction may be done prior to the visit and do not need to fall within the protocol visit window.

Footnotes continued on next page.

q. Includes Brief Fatigue Inventory, Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items. The BFI will be administered at cycle 1 day 1 predose, cycle 1 day 8 (± 1 day), day 15 (± 1 day), cycle 2 day 1 (± 2 days), day 15 (± 1 day) and all subsequent cycles day 1(± 2 days). Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items will be administered at cycle 1 day 1 predose, cycle 2 day 1 (± 2 days) and all subsequent cycles day 1 (± 2 days).

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

s. For subjects receiving LoDAC or azacitidine, thyroid function tests will be repeated after every 2 cycles of therapy (C3D1, C5D1, C7D1, etc.).

Table 3 Post-treatment Schedule of Assessments

Activity	Pre-HSCT Visit/ End of Treatment Visit[a]	30-Day Follow-up (+ 7 days Post Pre-HSCT Visit/Post End of Treatment Visit)	Long-term Follow-up (+/- 7 days)[h]
Physical Examination	X[b]		
Vital Signs	X[b]		
ECOG Performance	X[b]		
Concomitant Medications	X[l]		
Pregnancy Test for Woman of Childbearing Potential	X		
12-lead ECG	X		
Ophthalmologic Assessment[j]	X		
Clinical Laboratory Tests (chemistry, hematology, coagulation)	X[b]		
Thyroid Function Tests	X		
Bone Marrow Aspiration and/or Biopsy	X[c]		
FLT3 Mutations[d] (bone marrow or whole blood)	X		
Patient Reported Outcome Tools[i,k]	X		
EQ-5D-5L[k]	X	X	X
Resource Utilization	X		
AE/SAE Assessment	X	X[e,f]	X[g]
IRT Transaction Required	X		
Survival and Subsequent Antileukemic Treatments and Their Outcomes		X[f]	X

AE: adverse event; ECG: electrocardiogram; ECOG: Eastern Cooperative Oncology Group; EDTA: ethylenediaminetetraacetic acid; EQ-5D-5L: EuroQol Group-5 Dimension-5 Level Instrument; FLT3: FMS-like tyrosine kinase; HSCT: hematopoietic stem cell transplant; IRT: interactive response technology; SAE: serious adverse event

a. End of treatment visit is to be performed within 7 days after treatment discontinuation, and before initiation of any other systemic antileukemic treatment or conditioning regimen for HSCT.

b. Does not need to be repeated if collected at a regularly scheduled visit within 3 days of the end of treatment visit.

c. Bone marrow aspiration and/or biopsy for morphology are preferred, but biopsy may be omitted if the aspirate is considered to be adequate. If bone marrow aspirate is unobtainable (e.g., dry tap), an additional EDTA tube of peripheral blood should be collected instead.

d. FLT3 mutation analysis will be performed on the bone marrow samples collected post study treatment.

Footnotes continued on next page

e. For subjects who plan to proceed to HSCT and resume ASP2215 treatment after HSCT, AE collection will continue until the start of the HSCT conditioning regimen and AE collection will resume upon the resumption of ASP2215 treatment until 30 days after the last dose of study drug. For subjects who do not plan to resume ASP2215 treatment after HSCT, AE collection will continue until the start of the HSCT conditioning regimen or 30 days after the last dose of study drug, whichever comes first. However, the following AE/SAEs will continue to be collected until 30 days after the last dose of study drug, 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 PRES
   - Adverse events leading to death

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

g. Only SAE data that is related to ASP2215 will be collected.

h. Telephone contact every 3 months. Ad hoc contact will be required during interim analysis.

i. Includes Brief Fatigue Inventory, Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items. The BFI will be administered at pre-HSCT/end of treatment visit. Functional Assessment of Chronic Illness Therapy–Dyspnea-Short Forms, Functional Assessment of Cancer Therapy-Leukemia and dizziness and mouth sores items will be administered at pre-HSCT/end of treatment visit.

j. Ophthalmologic assessment to be performed by visual acuity measurement and ophthalmoscopy, at the Pre-HSCT/end of treatment (+/- 7 days). In symptomatic subjects, the ophthalmologic assessment should also include slit lamp biomicroscopy, visual fields and optical coherence tomography (OCT).

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

l. Concomitant medications should be collected for reported AE/SAEs through 30 days post dose for subjects who have discontinued. For subjects who undergo HSCT, concomitant medications should be collected for reported AE/SAEs through start of conditioning treatment or 30 days post dose, whichever comes first.

3 STUDY OBJECTIVES AND DESIGN

3.1 Study Objectives

3.1.1 Primary Objectives

The primary objectives are to:

- Determine the clinical benefit of ASP2215 therapy in subjects with FMS-like tyrosine kinase (FLT3) mutated AML who are refractory to or have relapsed after first-line AML therapy as shown with overall survival (OS) compared to salvage chemotherapy.
- Determine the efficacy of ASP2215 therapy as assessed by the rate of complete remission and complete remission with partial hematological recovery (CR/CRh) in subjects with FLT3 mutated AML who are refractory to or have relapsed after first-line AML therapy.

3.1.2 Secondary Objectives

The key secondary objectives are to:

- Determine the overall efficacy in event-free survival (EFS) of ASP2215 compared to salvage chemotherapy.
- Determine the overall efficacy in complete remission (CR) rate of ASP2215 compared to salvage chemotherapy.

The secondary objectives are to evaluate the safety and efficacy of ASP2215 therapy versus salvage chemotherapy in terms of:

- Leukemia-free survival (LFS)
- Duration of remission
- Complete remission with partial hematologic recovery (CRh) rate
- Composite complete remission (CRc=CR+CRp+CRi) rate
- Transfusion conversion rate; Transfusion maintenance rate
- Transplantation rate
- Patient reported fatigue (Brief Fatigue Inventory [BFI])
- Adverse events (AEs), safety labs, vital signs, ophthalmologic exams, electrocardiograms (ECGs) and Eastern Cooperative Oncology Group (ECOG) performance scores
- Evaluation of ASP2215 (and metabolites as appropriate) plasma concentration and population pharmacokinetics (PK)

3.1.3 Exploratory Objectives:

Evaluate the safety and efficacy of ASP2215 therapy versus salvage chemotherapy in terms of:

- Pharmacogenomics (PGx)
- FLT3 gene mutation status
  - mutation types and frequency
  - relationship to efficacy and safety
  - mechanisms of acquired resistance
- Exploratory (predictive) biomarkers of ASP2215 activity
- Resource utilization in this study population including hospitalization, blood transfusion, antibiotic intravenous infusions, medication for AEs and opioid usage
- Patient reported dyspnea (Functional Assessment of Chronic Illness Therapy-Dyspnea-Short Forms [FACIT-Dys-SF])
- Patient reported signs, symptoms and impacts of AML (Functional Assessment of Cancer Therapy-Leukemia [FACT-Leu], dizziness and mouth sore items)
- EuroQol Group-5 Dimention-5 Level Instrument (EQ-5D-5L)

3.2 Study Design

3.2.1 Study Design

This is a phase 3, open-label, multicenter, randomized study to compare the efficacy and safety of ASP2215 therapy to salvage chemotherapy in FLT3-mutated AML subjects who are refractory to or have relapsed after first-line AML therapy. Approximately 140 centers in North America, Europe, Asia and rest of the world will participate in this study.

Three hundred sixty nine subjects will be randomized. The randomization of the 369 subjects will be in a 2:1 ratio to receive ASP2215 or salvage chemotherapy. Subjects will enter the screening period up to 14 days prior to the start of treatment. Prior to randomization, the Investigator will preselect a salvage chemotherapy regimen for each subject; options will include LoDAC, azacitidine, MEC or FLAG-IDA. The randomization will be stratified by response to first-line therapy and preselected salvage chemotherapy. Subjects will be administered treatment over continuous 28-day cycles and per institutional guidelines for chemotherapy product preparation and administration. The dose and duration of study treatments are outlined in Section 5.1.1 of the protocol.

For subjects taking ASP2215, LoDAC or azacitidine, treatment should continue until the subject meets a treatment discontinuation criterion.

Subjects receiving MEC or FLAG-IDA will receive 1 cycle of therapy and will be assessed for response on or after day 15, per institutional guidelines. If the bone marrow cellularity is

20% or greater with at least a 50% reduction in blasts, the subject may receive a second cycle of the same chemotherapy. If bone marrow cellularity is between 5% and 20%, the investigator should make the decision whether the subject should receive another treatment cycle or be observed for recovery. If bone marrow cellularity is 5% or less, the subject will be observed for recovery. Subjects achieving CR, CRi or CRp may receive a second cycle of chemotherapy at the Investigator's discretion. Subjects with no response (NR) or progressive disease following cycle 1 will discontinue study treatment.

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

Subjects who have a donor identified and achieve a response allowing them to undergo HSCT per each institution's assessment can undergo HSCT without leaving the study. However, ASP2215 should be stopped and a pre-HSCT visit should be performed prior to starting the conditioning regimen for HSCT. ASP2215 can be resumed after stem cell transplantation if the following conditions are met:

- Subject is between 30 - 60 days post HSCT
- Subject has had successful engraftment as demonstrated by absolute neutrophil count (ANC) ≥ 500/mm3 and platelets ≥ 20000/mm3 without transfusions
- Subject does not have ≥ grade 2 acute graft-versus-host-disease (GVHD)
- Subject is in CRc

For subjects resuming treatment, subjects will follow the procedures listed under subsequent cycles day 1 in the Schedule of Assessments. Subjects who do not resume ASP2215 will be followed for OS endpoint.

After treatment discontinuation, subjects will have a pre-HSCT/end of treatment visit within 7 days after treatment discontinuation, followed by a 30-day follow-up visit in which a telephone contact with the subject is sufficient unless any assessment must be repeated for resolution of treatment-related AEs. After which the subjects will enter the long-term follow-up period for collection of patient reported outcome (PRO) using EQ-5D-5L, subsequent AML treatment, remission status and survival (cause of death and date of death). The long-term follow-up will be every 3 months, for up to 3 years from the subject's end of treatment visit.

Two interim analyses by an Independent Data Monitoring Committee (IDMC) will be conducted. The first interim analysis is planned when approximately 141 subjects are randomized into ASP2215 arm and at least 112 days (4 treatment cycles) post first dose or randomization (for subjects who received no study drug). The first interim analysis will be performed to evaluate the efficacy endpoint of CR/CRh rate and the study conduct will not be impacted by the result of the CR/CRh rate.

The second interim analysis will be performed when approximately 50% of the planned total number of deaths (death events =129 of planned 258 death events) by any cause have occurred. The second interim analysis will be utilized to determine whether the study should be terminated earlier than planned if ASP2215 has more favorable or harmful outcome than

the salvage chemotherapy group. Details of the interim analyses are described in Section 7.10.

The End of Study analysis will be performed on safety data only.

3.3 Randomization

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 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.

Subjects will be randomized in a 2:1 ratio to receive ASP2215 or salvage chemotherapy. Randomization will be stratified by response to first-line AML therapy and preselected salvage chemotherapy:

Response to first-line therapy:

- Relapse within 6 months after allogeneic HSCT
- Relapse after 6 months after allogeneic HSCT
- Primary refractory without HSCT
- Relapse within 6 months after CRc and no HSCT
- Relapse after 6 months after CRc and no HSCT

Preselected chemotherapy:

- High intensity chemotherapy (FLAG-IDA, MEC)
- Low intensity chemotherapy (LoDAC or azacitidine)

4 SAMPLE SIZE

This is a group sequential design based on co-primary endpoint of OS using the O'Brien-Fleming boundaries (non-binding) as implemented by Lan-DeMets alpha/beta spending method (East®).

The overall 0.025 one-sided type I error rate is allocated by 0.0005 and 0.0245 (corresponding to 0.001 and 0.049 for two-sided type I error rates) for the two co-primary efficacy endpoints of CR/CRh and OS, respectively. The type I error (alpha) in the first interim analysis is a nominal alpha which is arbitrarily selected for acknowledgement of the CR/CRh rate evaluation at the first interim and will not be recycled in the second interim analysis and final analysis.

Two interim analyses and one final analysis are planned. The first interim analysis is planned when approximately 141 subjects are randomized into ASP2215 arm and at least 112 days (4 treatment cycles) post first dose or randomization (for subjects who received no study drug). The second interim analysis is planned when approximately 129 death events have occurred and the final analysis is planned when approximately 258 death events have occurred.

OS:

Approximately 369 subjects (the planned sample size with 10% dropout rate) will be randomized in a 2:1 ratio to receive ASP2215 or salvage chemotherapy (246 subjects in the ASP2215 treatment arm and 123 subjects in the salvage chemotherapy arm). The planned 258 death events will provide about 90% power to detect a difference in OS between the ASP2215 arm with 7.7 months median survival time and salvage chemotherapy arm with 5 months median survival time (hazard ratio = 0.65) at the overall one-sided 0.0245 significance level.

CR/CRh rate:

The co-primary endpoint of CR/CRh rate will be evaluated only at the first interim analysis. One hundred and forty one subjects randomized to ASP2215 arm (211 subjects in total: 141 in the ASP2215 arm and 70 in the salvage chemotherapy arm) with a minimum follow-up of 4 treatment cycles are considered to achieve a maximum width of 15.78% for the two-sided 95% exact confidence interval (CI) when the CR/CRh is expected to be in the 5% to 30% range as summarized in the following table. A sample size of 141 subjects provides 80% power to exclude a CR/CRh rate of 12% using the two-sided 95% exact CI when the CR/CRh rate of ASP2215 is assumed to be 21%.

Table 1: Observed CR/CRh with Exact 95% CI (N=141 in ASP2215 arm)

Observed CR/CRh (n and %)	95% Exact CI
43 (30.50%)	(23.03%, 38.80%)
36 (25.53%)	(18.57%, 33.55%)
29 (20.57%)	(14.23%, 28.18%)
28 (19.86%)	(13.62%, 27.41%)
27 (19.15%)	(13.01%, 26.62%)
26 (18.44%)	(12.41%, 25.84%)
25 (17.73%)	(11.82%, 25.05%)
24 (17.02%)	(11.22%, 24.26%)
23 (16.31%)	(10.63%, 23.46%)
22 (15.60%)	(10.04%, 22.66%)
15 (10.64%)	(6.08%, 16.94%)
8 (5.67%)	(2.48%, 10.87%)

EFS and CR rate:

The planned sample size with 258 EFS events will provide about 90% power to detect the difference in EFS (6 months median EFS for ASP2215 arm and 3.9 months for salvage chemotherapy arm with hazard ratio = 0.65) and > 90% power to detect a difference in CR rate between ASP2215 (with 25% CR rate) and the salvage chemotherapy (with 10% CR rate) at the overall 1-sided 0.0245 significance level.

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.

Safety Analysis Set (SAF) will be used for the analyses of safety variables. Pharmacokinetic Analysis Set (PKAS) will be used for pharmacokinetic analyses. The data from all randomized subjects will be included in the data listings.

The ITT, SAF and PKAS analysis sets will be used for the final end-of-study report.

5.1 Intention to Treatment Set (ITT)

The Intention to Treatment Set (ITT) consists of all subjects who are randomized. The subjects will be analyzed based on the randomized treatments.

The ITT will be used for primary analyses of efficacy data. Selected demographic and baseline characteristics, patient reported outcomes and resource utilization.

5.2 Full Analysis Set (FAS)

The Full Analysis Set (FAS) consists of all randomized subjects with FLT3 mutation based on central test. The subjects will be analyzed based on the randomized treatments.

The selection of subjects for the FAS will be confirmed in the Analysis Set Classification Meeting (ASCM).

The FAS will be used for sensitivity analyses of efficacy data Selected demographic and baseline characteristics will be summarized for the FAS.

5.3 Safety Analysis Set (SAF)

The Safety Analysis Set (SAF) consists of all subjects who received at least one dose of study treatment (ASP2215 or salvage chemotherapy). The subjects will be analyzed based on the actual treatment received.

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

5.4 Pharmacokinetics Analysis Set (PKAS)

The Pharmacokinetics Analysis Set (PKAS) consists of the subset of the SAF who received at least one dose of ASP2215 for whom at least 1 plasma concentration datum is available and both the date and time of dosing on the day of PK sampling and the date and time of sampling are known. Additional subjects may be excluded from the PKAS at the discretion of the pharmacokineticist. Any formal definitions for exclusion of subjects or time-points from the PKAS will be documented in the Classification Specifications.

The PKAS is used for all tables and graphical summaries of the PK data.

6 ANALYSIS VARIABLES

6.1 Efficacy Endpoints

There will be no efficacy analysis performed for the end of study report. Summaries of the primary, secondary, and exploratory efficacy results are available in the primary CSR dated 24 JAN 2019 for the data with cutoff date 17 SEP 2018.

6.1.1 Co-Primary Efficacy Endpoint of OS (Second Interim and Final only)

Overall survival (OS) is one of the co-primary efficacy endpoints. OS will be evaluated at the second interim and final analysis. No formal evaluation on OS will be conducted at the first interim. 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 by the cutoff date. The last contact date will be derived for subjects alive at the analysis cutoff date. Subjects with last contact date beyond the analysis cutoff date will be censored at the analysis cutoff date.

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).

In addition, as another sensitivity analysis, OS will be defined similarly as the above primary analysis, however, subjects who initiate any new anti-leukemia therapy will be censored at the time of first new anti-leukemia therapy (first new anti-leukemia therapy date – randomization date +1).

During second interim analysis and final analysis, only death events occurring on or prior to the cutoff date are counted as OS events. Subjects with death date after the cutoff date will be censored at the cutoff date.

6.1.2 Co-Primary Efficacy Endpoint of CR/CRh Rate (First Interim only)

The co-primary efficacy endpoint of CR/CRh rate will be evaluated only at the first interim analysis. Complete remission and complete remission with partial hematologic recovery (CR/CRh) rate is defined as the number of subjects who achieve either CR or CRh (as defined in Section 6.1.3.3) at any of the post-baseline visit divided by the number of subjects in the analysis population.

As a sensitivity analysis, CR/CRh rate by cycle 4 will be defined as the number of subjects who achieve CR/CRh by the end of 4 cycles treatment divided by the number of subjects in the analysis population. As another sensitivity analysis, CR/CRh rate prior to HSCT will be defined as the number of subjects who achieve CR/CRh prior to HSCT divided by the number of subjects in the analysis population.

6.1.3 Secondary Efficacy Endpoints

6.1.3.1 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 (excluding relapse after PR), treatment failure or death from any cause within 30 days after the last dose of study drug, whichever occurs first [earliest of (relapse date, treatment failure date, death date) – randomization date + 1].

  If a subject experiences relapse or death within 30 days after the last dose of study drug, the subject is defined as having EFS event related to either "relapse" or "death", and the event date is the date of relapse or death.

  If a subject discontinues the treatment and fails to achieve any of the response of CR, CRp or CRi during the treatment period (subject with best response of PR or NR), the subject is defined as having EFS event related to treatment failure, and the event date is the randomization date. Subjects that discontinue the treatment with post-treatment disease assessment and best response of NE will be censored.

  For a subject who is not known to have relapsed 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). Subject is not censored at HSCT. For subjects who are censored, last relapse-free disease assessment date refers to the subject's last disease assessment date. Subject without post-treatment disease assessment or without study treatment will be censored at randomization date.

  EFS will be tested at the second interim and final analysis only when OS is rejected. No formal evaluation on EFS will be conducted at the first interim. During second interim analysis and final analysis, only death, relapse, treatment failure occurring on or prior to the cutoff date are counted as EFS events. Subjects without these events before cutoff date will be censored at the last relapse-free disease assessment date post-baseline on or before the cutoff date.

  Disease assessment date refers to the date of central 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 the date the peripheral blood sample is drawn will be used. If none of the above is available and subject is evaluated based on the presence of extramedullary leukemia, the assessment date of extramedullary leukemia will be used.

  The following sensitivity analyses for EFS will be conducted:

  - EFS will be defined similarly as the above primary analysis, however the date of the first new anti-leukemia therapy or the last treatment evaluation (when new anti-leukemia therapy date is not available), will be used as the event date of treatment failure.

- Sensitivity analysis of EFS will be performed to evaluate the impact of loss to follow-up on EFS. EFS will be defined similarly as the above primary analysis, however, subjects who discontinue the treatment due to "Lost to follow up" will also be considered as an EFS event (See Section 7.4.3.1 for details of "Lost to follow up" event), and the event date is withdrawal date or last evaluation date as collected on end of treatment, end of 30-day follow up, and end of long-term follow up CRF page whichever occurs first.
- EFS will be defined similarly as the above primary analysis with the following differences:
  - Subjects who discontinue the treatment and achieve best response of NR only will be defined as treatment failure and the event date is the randomization date
  - Subjects who discontinue the treatment and achieve best response of CR, CRp, CRi or PR and experience relapse will be defined as relapse and the event date is the date of first NR after CR, CRp, CRi or PR.
- EFS will be defined similarly as the above primary analysis, however, subjects who discontinue the treatment with post-treatment disease assessment and achieve best response of PR, NR or NE will be defined as treatment failure and the event date is the randomization date.

- Complete remission (CR) rate

  CR rate is defined as the number of subjects who achieve the best response of CR divided by the number of subjects in the analysis population. As a sensitivity analysis, CR rate prior to HSCT will be defined as the number of subjects who achieve CR prior to HSCT divided by the number of subjects in the analysis population.

  CR rate will be tested at the second interim and final analysis only when both OS and EFS are rejected. No formal evaluation on CR rate will be conducted at the first interim.

6.1.3.2 Other Secondary Efficacy Endpoints

- Leukemia-free survival (LFS)

  LFS is defined as the time from the date of first CRc until the date of documented relapse (excluding 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 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).

- Complete Remission with Partial Hematologic Recovery (CRh) Rate

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

- Composite Complete Remission (CRc) Rate

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

- Duration of remission

  Duration of remission includes duration of CRc (DCRc), duration of CR/CRh (DCRCRh), duration of CRh (DCRh), duration of CR (DCR), and duration of response (DR) (i.e., 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 do not relapse on study are considered non-events and censored at the last relapse-free disease assessment date.

  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 will be used as the starting date of the duration.

  As a sensitivity analysis, Duration of CR/CRh will be defined similarly as the above, however deaths without report of relapse are considered as events.

  Duration of CRh is defined similarly as duration of CRc for subjects who achieve CRh.

  Duration of CR is defined similarly as duration of CRc for subjects who achieve CR.

  Duration of response is defined as the time from the date of either first CRc or PR until the date of documented relapse (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.

- Transfusion conversion rate and transfusion maintenance rate

  Transfusion conversion rate and transfusion maintenance rate will only be defined for subjects in ASP2215 arm.

  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 28 days post first dose. For subjects who are on treatment <28 days, baseline period is from 28 days prior to the first dose until the end of treatment.

- 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 (only defined for ASP2215 arm):

- Post-baseline period is defined as the period from 29 days post first dose until last dose.
- For subjects who are on treatment >=84 days, subjects are classified post-baseline transfusion independent if there is one consecutive 56 days without any RBC or platelet transfusion within post-baseline period.
- For subjects who are on treatment >28 days but <84 days, if there is no RBC or platelet transfusion within post-baseline period, post-baseline transfusion status is not evaluable.
- For subjects who are on treatment <=28 days, post-baseline transfusion status is not evaluable.
- Otherwise, the subject is considered 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.

As a sensitivity analysis, transfusion conversion rate and transfusion maintenance rate will be defined alternatively by considering all subjects who had not-evaluable post-baseline transfusion status as transfusion dependent.

- Transplantation rate

  Transplantation rate is defined as the percentage of subjects undergoing HSCT during the study period.

- Brief fatigue inventory (BFI)

  BFI 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. Other BFI scores will be derived per PRO SAP.

6.1.3.3 Response Definition

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

Response 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.

- Complete Remission (CR)

  For subjects to be classified as being in CR at a post-baseline visit, they must have bone marrow regenerating normal hematopoietic cells and achieve a morphologic leukemia-free state and must have an absolute neutrophil count (ANC) $\geq 1 \times 10^9$/L and platelet count $\geq 100 \times 10^9$/L, and normal marrow differential with < 5% blasts, and they will be red blood cell (RBC) and platelet transfusion independent (defined as 1 weeks without RBC transfusion and 1 week without platelet transfusion). There must be no presence of Auer rods. There should be no evidence of extramedullary leukemia. The blast counts in peripheral blood must be ≤ 2%.

- Complete Remission with Incomplete Platelet Recovery (CRp)

  For subjects to be classified as being in CRp at a post-baseline visit, they must achieve CR except for incomplete platelet recovery ($< 100 \times 10^9$/L).

- Complete Remission with Incomplete Hematological Recovery (CRi)

  For subjects to be classified as being in CRi at a post-baseline visit, they must fulfill 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)

  For subjects to be classified as being in CRc at a post-baseline visit, they must either achieve CR, CRp or CRi at the visit.

- Complete Remission with Partial Hematologic Recovery (CRh)

  At a post baseline visit, subjects will be classified as CRh if they have marrow blasts < 5%, partial hematologic recovery ANC >= 0.5 x 109/L and platelets >= 50 x 109/L, no evidence of extramedullary leukemia and cannot be classified as CR. The blast counts in peripheral blood must be ≤ 2%

- Partial Remission (PR)

For subjects to be classified as being in PR at a post-baseline visit, they must have bone marrow regenerating normal hematopoietic cells with evidence of peripheral recovery with no (or only a few regenerating) circulating blasts and with a decrease of at least 50% in the percentage of blasts in the bone marrow aspirate with the total marrow blasts between 5% and 25%. A value of less or equal than 5% blasts is also considered a PR if Auer rods are present. There should be no evidence of extramedullary leukemia.

- Not Evaluable (NE)/No Response (NR)

  In the situation where no bone marrow assessments are performed or myeloblast value is missing, blast value from peripheral blood is missing or <=2%, and extramedullary leukemia is missing or not present, the response will be classified as not evaluable (NE). In any case response cannot be categorized as CR, CRp, CRi, PR or NE, it will be categorized as NR

- Relapse

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

  Relapse after PR is similarly defined with reappearance of significant numbers of peripheral blasts and an increase in the percentage of blasts in the bone marrow aspirate to > 25% not attributable to any other cause or reappearance or new appearance of extramedullary leukemia.

- Best Response

  Best response is defined as the best measured response to treatment for all visits (in the order of CR, CRp, CRi, PR, NR and NE) post-baseline. Subjects who achieve the best responses of CR, CRp, CRi, or PR will be classified as responders. Subjects who do not achieve at least a best response of PR will be classified as non-responders.

6.1.4 Exploratory Endpoints

- Pharmacogenomics
- FLT3 gene mutation status
  - mutation types (FLT3-ITD mutation and D835/I836 TKD mutation) and frequency
  - relationship to efficacy and safety
  - mechanisms of acquired resistance
- Exploratory (predictive) biomarkers of ASP2215 activity
- Resource utilization, including hospitalization, blood transfusion, antibiotic intravenous infusions, medication for AEs and opioid usage
- Functional Assessment of Chronic Illness Therapy-Dyspnea-Short Forms (FACIT-Dys-SF)

The FACIT-Dys-SF was developed to assess dyspnea severity and related functional limitations. It has a 7-day recall period and 20 items. The FACIT-Dys-SF is scored with 2 domains: dyspnea and function limitations.

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

  The FACT-Leu is designed to measure leukemia-specific signs, symptoms and the impact of AML on patients. The 44-item scale has global and domain scores including physical well-being, social/family well-being, emotional well-being, functional well-being and additional leukemia-specific concerns. The FACT-Leu contains some of the most common patient reported signs, symptoms, and impacts of AML. The FACT-Leu has a 7-day recall period.

- Dizziness and mouth sore items

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

  The EQ-5D-5L is being used as a measure of respondents' health related quality of life. The EQ-5D-5L consists of the EuroQol Group-5 Dimension descriptive system and the EuroQol Group VAS.

  The EuroQol Group-5 Dimension descriptive system comprises of 5 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 'best imaginable health state' and 'worst imaginable health state' with higher scores for higher health related quality of life.

6.1.5 Other Efficacy Variables

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

- Best response rate – Defined as the number of subjects who achieve the best response of CR or CRp or CRi or PR divided by the number of subjects in the analysis population.

- Time to remission

  Time to remission includes time to CRc (TTCRc), time to CR (TTCR), time to first CR/CRh (TTFCRCRH), time to best CR/CRh (TTBCRCRH) and time to response (TTR).

  Time to CRc is defined as the time from the date of randomization until the date of first CRc (first CRc disease assessment date – randomization date +1). TTCRc will only be evaluated for subjects who achieved CRc.

  Time to CR is defined similarly as time to CRc for subjects who achieve CR.

  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 as the time from the date of randomization until the date of either first CRc or PR (first CRc or PR disease assessment date – randomization date +1). TTR will only be evaluated for subjects who achieved CRc or PR.

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 adverse event observed after starting administration of the study treatment (ASP2215 or salvage chemotherapy). If the adverse event 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 adverse event will be considered treatment emergent. If the adverse event occurs on Day 1 and the onset check box is marked "Onset before first dose of study drug", then the adverse event 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 as TEAE only if it has worsened in severity (i.e. it is reported with a new start date). All adverse events collected that begin within 30 days after taking the last dose of study drug will also be 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 adverse events observed after starting administration of the study treatment until the last dose before pre-HSCT visit plus 30 days, and adverse events that begin after resumption of ASP2215 and within 30 days after the last dose of ASP2215 will also be 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 7.11.1.
  - 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 adverse events that are flagged as serious by the investigator on 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 (as specific in Section 10.3, Appendix 3).
- Clinical laboratory variables (hematology, biochemistry including liver function test and thyroid function test)

- Vital signs (systolic and diastolic blood pressure, pulse rate, and body temperature)
- 12-lead electrocardiogram (ECG)
- Ophthalmologic assessment (visual acuity measurement)
- ECOG performance scores

6.3 Pharmacokinetics Variables

Plasma concentration data of ASP2215 will be used in pharmacokinetic analysis.

6.4 Other Variables

- Duration of exposure
  Duration of exposure to each component of a study drug will be calculated in days, using the following formula:

  Last date of exposure – First dose date + 1 – (on-study HSCT period for subjects undergo on-study HSCT)

  For ASP2215, last date of exposure = last dose date of ASP2215

  For chemotherapy, last date of exposure = (initial dose date of the last cycle + 28 – 1) or death date if death occur within last cycle

  When the start or stop date is missing, then the exposure will be treated as missing.

  Duration of exposure for a chemotherapy regimen, i.e. MEC and FLG-IDA, will be the maximum (duration of exposure of all treatment components).

  When the last date of exposure is beyond cutoff date, the cutoff date will be used as the last date of exposure.
- Number of dosing days

  Number of days with non-zero dosing.
- Number of cycles

  Number of cycles for each component of study drug refers to total number of cycles with non-zero dosing in the cycle.

  Number of cycles for a chemotherapy regimen refers to total number of cycles with non-zero dosing of any treatment component in the cycle.
- Cumulative dose (dose unit)

  Cumulative dose will be calculated using the following formula for each type of study drug:
  - ASP2215: Sum of [(stop date – start date + 1) * dose captured on dosing CRF] across all records
  - LoDAC: Sum of (number of dose taken * 20mg) across all records and all cycles. If number of dose take is “unknown”, it will be treated as missing.
  - Azacitidine: Sum of (number of dose taken * actual delivered dose captured on dosing CRF) across all records and all cycles. If number of dose taken is “unknown”, it will be treated as missing.

- Each component of MEC and FLAG-IDA: Sum of (actual delivered dose captured on dosing CRF) across the days when study drug was administered. If answer to "entire infusion administered" is No, actual delivered dose will be treated as missing.

- Average daily dose (dose unit/days)

  Cumulative dose (dose unit)

  Number of dosing days (days)

- Dose intensity

  Dose intensity will be calculated in mg/day for ASP2215 as:
  Cumulative dose (mg)

  Duration of exposure (days)

  Dose intensity will be calculated in dose unit/cycle for chemotherapy as:
  Cumulative dose (dose unit)
  ---------- x 28 (days/cycle)
  Duration of exposure (days)

- Relative dose intensity

  Relative dose intensity for ASP2215 will be calculated as:
  Dose intensity (mg/day)
  ---------- x 100%
  120 (mg/day)

  Relative dose intensity for each component of chemotherapy will be calculated as:
  Dose intensity (dose unit/cycle)
  ---------- x 100%
  Planned dose intensity (dose unit/cycle)

  Where planned dose intensity (dose unit/cycle) will be calculated based on protocol specified dose:

  - LoDAC: 40 mg/day * 10 days/cycle
  - Azacitidine: 75 mg/m^2/day * BSA * 7 days/cycle
  - MEC

    Mitoxantrone: 8 mg/m^2/day * BSA * 5 days/cycle

    Etoposide: 100 mg/m^2/day * BSA * 5 days/cycle

    Cytarabine: 1000 mg/m^2/day * BSA * 5 days/cycle
  - FLAG-IDA

G-CSF: 300 μ g/m^2/day * BSA * 5 days/cycle

Fludarabine: 30 mg/m^2/day * BSA * 5 days/cycle

Cytarabine: 2000 mg/m^2/day * BSA * 5 days/cycle

Idarubicin: 10 mg/m^2/day * BSA* 3 days/cycle

Average of BSA for all dosed cycles will be used in the calculation of planned dose intensity.

- Relative dose intensity for a chemotherapy regimen will be the average of relative dose intensity of all treatment components.
- Duration of AML

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

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

  Partial date of initial diagnosis of AML will be imputed per Section 7.11.1.
- Previous and concomitant medication

  Previous medication is defined as medication with at least one dose taken before the date of the first dose of study drug (exclusive).

  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) of study drug.

  For subjects that undergo HSCT without leaving the study and plan to resume ASP2215 treatment 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).
- Prior Use of FLT3 inhibitor
  Prior Use of FLT3 inhibitor is defined as "Yes" if subjects received prior AML therapy of Midostaurin, Sorafenib or Quizartinib.
- 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 first dose (exclusive).

  Concomitant transfusion is defined as transfusion received 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 treatment after HSCT, concomitant transfusion is defined as transfusion received 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).

7 STATISTICAL METHODOLOGY

7.1 General Considerations

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 specified 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 ITT (e.g., disposition) will be presented by randomized treatment group, 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 initial dose received before dose change.

All statistical comparisons will be made using two-sided tests at the α=0.05 significance level unless specifically stated otherwise. All null hypotheses will be of no treatment difference, all alternative hypotheses will be two-sided, unless specifically stated otherwise.

All data processing, summarization, and analyses will be performed using SAS® Version 9.3 or higher on Statistical Computing Environment. 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 prior to the first dose of study drug. Unless otherwise specified, all summaries will be presented by treatment groups.

For the definition of subgroups of interest please refer to Section 7.8.

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, re-screening, discontinued before randomization, randomized (overall only);
- Number and percentage of subjects randomized in each analysis set, by treatment group and overall;
- Number and percentage of subjects completed and discontinued treatment, by primary reason for treatment discontinuation for randomized subjects, by treatment group;
- Number and percentage of subjects completed the 30-day follow-up evaluation, by 30-day follow-up status for randomized subjects and by treatment group;
- Number and percentage of subjects completed the long term follow-up evaluation, by long term follow-up status for randomized subjects and by treatment group.

7.2.2 Protocol Deviations

Protocol deviations as defined in the study protocol (Section 8.1.6 Protocol Deviations) will be assessed for all randomized subjects. The number and percentage of subjects meeting any criteria will be summarized for each criterion and overall, by treatment group and total as well as by study 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.

The visit-based assessments affected by COVID-19 pandemic such as the assessments performed out of window, performed at alternative location or virtual assessment due to COVID-19 will be provided in data listings by site and subject.

The non-visit-based assessments affected by COVID-19 pandemic such as treatment discontinuation, adverse event, dose changes due to COVID-19 will be provided in data listings by site and subject.

7.2.3 Demographic and Other Baseline Characteristics

The demographics and other baseline characteristics were reported in the primary CSR dated 24 JAN 2019 for the data with cutoff date of 17 SEP 2018. Those will not be summarized for the end of study report.

Demographic and other baseline characteristics will be summarized by descriptive statistics.

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

Descriptive statistics for age, weight, body surface area (BSA) and height at study entry will be presented by treatment group. Frequency tabulations for sex, ethnicity, age group (defined in Section 7.8), race, region, baseline central FLT3 status, baseline ECOG, prior use of FLT3 inhibitor and cytogenetic risk status will be presented by treatment group.

Frequency tabulations for AML disease history including AML subtype as classified by World Health Organization (WHO) classification and French-American-British (FAB) classification, risk status, antecedent hematological disorder, central nervous system leukemia, rapidly progressing disease, local FLT3 mutation status, local FLT3-ITD mutation status, local FLT3 point mutation status will be presented by treatment group.

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. 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 prior transplant including number of prior transplant, graft type, donor relatedness, match type and outcome of transplant will be presented by treatment group.

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

7.2.4 Previous and Concomitant Medications

The prior medications were reported in the primary CSR dated 24 JAN 2019 for the data with cutoff date of 17 SEP 2018. Only concomitant medications will be summarized for the end of study report.

Previous 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 by treatment group.

As with previous medication, concomitant medication will be summarized for each treatment group by therapeutic subgroup (ATC 2nd level) and chemical subgroup (ATC 4th level) and preferred WHO name for the SAF. 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.

Concomitant medication will also be summarized by treatment group and PT and presented in decreasing order of frequency based on the total number of subjects took each medication.

7.2.5 Previous and Concomitant Transfusions

The prior transfusions were reported in the primary CSR dated 24 JAN 2019 for the data with cutoff date of 17 SEP 2018. Only concomitant transfusions will be summarized for the end of study report.

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

7.2.6 Prior AML Chemotherapy

The prior AML chemotherapy was reported in the primary CSR dated 24 JAN 2019 for the data with cutoff date of 17 SEP 2018. This data will not be summarized for the end of study report.

Frequency tabulations of subjects with prior AML chemotherapy, response to first line therapy, regimen, type of treatment, prior use of FLT3 inhibitor, and best response to prior AML therapy will be presented by treatment group. Descriptive statistics will be presented for duration of response to prior AML therapy.

7.2.7 Non-Medication Therapy

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

7.2.8 Clinical/Diagnostic Procedures

Frequency tabulations of subjects with clinical/diagnostic procedures and reason for use will be presented by treatment group. Number of clinical/diagnostic procedures received per subject will be summarized using descriptive statistics.

7.2.9 Subsequent AML Therapy

Frequency tabulations of subjects with subsequent AML therapy, regimens, relapse prior to subsequent AML therapy, reason of starting subsequent AML therapy, and response to subsequent AML treatment will be presented by treatment group. Descriptive statistics will be presented for duration of subsequent AML therapy.

7.3 Study Drugs

7.3.1 Exposure

The following information on drug exposure will be presented for each treatment regimen for the SAF:

- Descriptive statistics for cumulative amount of the drug subject was exposed to, number of dosing days, number of dosing cycles, average daily dose, dose intensity, and relative dose intensity; and
- Number and percent of subject with dose increases, decreases or interruptions.

Duration of exposure will be summarized in two ways.

- Descriptive statistics will be presented by treatment regimen.
- Exposure time will be categorized according to the following categories by treatment regimen:
  - less than or equal to 5 days
  - at least 6 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 for each treatment regimen for the SAF.

Listing of subjects with dose reduction and dose escalation will also be provided.

7.4 Analysis of Efficacy

There will be no efficacy analysis performed for the end of study report.

The significance levels at each interim and final analyses for each co-primary and secondary endpoints are specified in Table 4 below:

Table 4 Summary of Timing, Sample Size and Decision Guidance at the Planned Analyses

Analysis	Criteria for conduct of analysis (Projected timing)	Endpoint/ Analysis Set	Efficacy Boundary*		Futility Boundary*	
			p-value (1-sided) at the Boundary	Approx. Observed HR at Boundary	p-value (1-sided) at the Boundary	Approx. Observed HR at Boundary
First Interim Analysis: CR/CRh rate	When 141 subjects are randomized into ASP2215 arm and at least 112 days (4 treatment cycles) post first dose or randomization (for subjects who received no study drug)	CR/CRh rate /2215 subjects in RAS	NA (0.0005 nominal)	NA	NA	NA
Second Interim Analysis: OS; EFS when null hypothesis of OS is rejected; CR rate when null hypotheses of both EFS and OS are rejected	Approx. 129 OS events were observed	OS/ITT	0.00147	0.57	0.38674	0.95
		EFS/ITT	0.01519	0.67	0.30218	0.91
		CR rate/ITT	0.01519	NA	0.30218	NA
Final Analysis: OS; EFS when null hypothesis of OS is rejected; CR rate when null hypotheses of both EFS and OS are rejected	Approx. 258 OS events were observed	OS/ITT	0.02402	0.77	NA	NA
		EFS/ITT	0.01357	0.75	NA	NA
		CR rate/ITT	0.01357	NA	NA	NA

*: P-value at both efficacy and futility boundaries (except the first interim) are based on 50% information fraction for OS, EFS and CR rate, and need update based on actual observed number of events at the second interim and final analysis.

7.4.1 Analysis of Co-Primary Endpoint of OS (Second Interim and Final only)

The co-primary efficacy endpoint of OS will be analyzed on ITT for primary analysis. In order to compare the OS between ASP2215 and the salvage chemotherapy, the null hypothesis will be constructed:

- H_{01}: OS in ASP2215 is worse or equal to the OS in salvage chemotherapy

The accompanying alternative hypothesis is:

- H_{11}: OS in ASP2215 is better than the OS in salvage chemotherapy

The co-primary endpoint of the OS will be tested only at second interim and final analyses. The one-sided p-value for the above hypothesis test will be performed using the stratified log-rank test (primary test) with strata (per IRT) to control for response to first-line AML therapy and preselected salvage chemotherapy.

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

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

The hazard ratio of the treatment effect along with 95% confidence interval will be calculated by the stratified Cox proportional hazard model. The same stratification factors will be applied to both the stratified log-rank test and the stratified Cox proportional hazard model. The same Cox proportional hazard model will be included as a sensitivity analysis for testing H_{01}.

The SAS code to calculate the hazard ratio of the treatment effect along with 95% confidence interval will be similar to that shown below:

```
PROC PHREG;
     MODEL time * status (1) = treatment/RL;
     STRATA stratification variables;
 RUN;
```

Stratification variables will include response to first-line AML therapy and preselected salvage chemotherapy.

Kaplan-Meier survival plots will be used to describe the OS in each treatment group. Median OS and 95% confidence interval, survival rates at 6, 12, 24 and 36 months and 95% confidence interval will be estimated from the Kaplan-Meier curve using the SAS code similar to that shown below:

```
PROC LIFETEST DATA=XX TIMELIST=(XX XX XX);
     TIME time * status (1);
     STRATA treatment;
RUN;
```

Where TIMELIST option contains timepoint of interest to estimate survival rate.

Sensitivity analyses for OS include:

- The same analysis as primary analysis but on FAS
- The same analysis as primary analysis but on PPS

- Stratified Cox proportional hazard model with strata to control for response to first-line AML therapy and preselected salvage chemotherapy on ITT
- The same analysis as primary analysis but with OS censoring at HSCT as defined in Section 6.1.1
- The same analysis as primary analysis but with OS censoring at first new anti-leukemia therapy as defined in Section 6.1.1
- Stratified Cox proportional hazard model with strata to control for response to first-line AML therapy and preselected salvage chemotherapy and initiation of new anti-leukemia therapy as a time-dependent binary covariate on ITT
- Additional sensitivity analysis will be performed using a new 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 pre-specified cutoff time at 18 months. The 18-month time-point was determined to ensure that reasonable number of patients were still at risk at the cutoff.
- The same analysis as primary analysis will be performed by dose adjustment for subjects whose dose was increased (200 mg dose), decreased (80 mg dose) or remained at 120 mg in ASP2215 arm.

7.4.2 Analysis of Co-Primary Endpoint of CR/CRh Rate (First Interim only)

The co-primary efficacy endpoint of CR/CRh rate will be evaluated at the first interim analysis only. The two-sided 95% exact confidence interval of CR/CRh rate will be calculated for approximately 141 subjects who are randomized into ASP2215 arm and at least 112 days (4 treatment cycles) post first dose or randomization (for subjects who received no study drug), i.e., 2215 subjects in RAS. The lower limit will be used to compare with the benchmark of CR/CRh rate of 12%.

CR/CRh rate along with the two-sided 95% exact confidence interval based on binomial distribution will be calculated using SAS code similar to that shown below:

```
PROC FREQ;
     TABLES outcome/BIONOMIAL (EXACT);
RUN;
```

The sensitivity analysis for CR/CRh rate will be performed to assess the robustness of the CR/CRh rate as follows:

- The same analysis as primary analysis but only including subjects in mRAS
- The same analysis as primary analysis but only including subjects who took at least one dose of ASP2215
- The same analysis as primary analysis but only including subjects who had at least one post baseline bone marrow assessment
- Apply the same analysis method used for primary analysis to evaluate CR/CRh rate by cycle 4 (as defined in Section 6.1.2)

- Apply the same analysis method used for primary analysis to evaluate CR/CRh rate prior to HSCT (as defined in Section 6.1.2)

7.4.3 Analysis of Secondary Endpoints

7.4.3.1 Key Secondary Efficacy Analysis

The key secondary efficacy endpoint of EFS will be analyzed in the same manner as the co-primary endpoint of OS. In order to compare the EFS between ASP2215 and the salvage chemotherapy, the null hypothesis will be constructed:

- H_{02}: EFS in ASP2215 is worse or equal to EFS in salvage chemotherapy

The accompanying alternative hypothesis is:

- H_{12}: EFS in ASP2215 is better than EFS in salvage chemotherapy

To maintain the overall 1-sided Type I error rate at the 0.0245 significance level, the key secondary efficacy endpoint of EFS will be tested only at second interim and final analyses and only if the null hypothesis on OS (H_{01}) is rejected. The test significance levels of EFS endpoint at the second interim and final analyses are specified in Table 4.

The one-sided p-value for the above hypothesis test will be performed using the stratified log-rank test with strata (per IRT) to control for response to first-line AML therapy and preselected salvage chemotherapy.

The hazard ratio of the treatment effect along with 95% confidence interval will be calculated using stratified Cox proportional hazard model. The same stratification factors will be applied to both the stratified log-rank test and the stratified Cox proportional hazard model. The same Cox proportional hazard model will be included as a sensitivity analysis for testing H_{03}.

Kaplan-Meier survival plots will be used to describe the EFS in each treatment group. Median EFS and 95% confidence interval, EFS rates at 6, 12, 24 and 36 months and 95% confidence interval will be estimated from the Kaplan-Meier curve. The number and percentage of EFS event (relapse, death or treatment failure) and censoring will be summarized for each treatment group.

The sensitivity analyses for the key secondary efficacy endpoints of EFS include:

- The same analysis as primary analysis but on FAS
- The same analysis as primary analysis but on PPS
- Stratified Cox proportional hazard model with strata to control for response to first-line AML therapy and preselected salvage chemotherapy on ITT
- The same analysis as primary analysis but with the date of the first new anti-leukemia therapy after the end of study treatment or the last treatment evaluation date (when the date of new anti-leukemia is not available) will be used as the event date of treatment failure as defined in Section 6.1.3.1
- The same analysis as primary analysis but with “Lost to follow up” also considered as an EFS event as defined in Section 6.1.3.1. “Lost to follow up” event includes

subjects who discontinued either study treatment or study follow up (including 30 day and long term follow up) due to either "LOST TO FOLLOW-UP" or "WITHDRAWAL BY SUBJECT". For subjects who discontinued the study treatment due to "WITHDRAWAL BY SUBJECT", only those who revoked the authorization to follow up will be included.

- The same analysis as primary but consider the subjects who discontinued the treatment and only achieved best response of NR as treatment failure as defined in Section 6.1.3.1.
- The same analysis as primary but consider the subjects who discontinued the treatment with post-treatment disease assessment and achieved best response of PR, NR or NE as treatment failure as defined in Section 6.1.3.1.
- The same analysis as primary but consider using long term follow up data of Death and new AML therapies. The new AML therapies captured on CRF page will only be considered as event when the patient did not relapse per derived response.

The key secondary efficacy endpoint of CR rate will be analyzed using the Cochran-Mantel-Haenszel (CMH) test to control for response to first-line AML therapy and preselected salvage chemotherapy (Per IRT) on ITT. In order to compare the CR rate between ASP2215 and the salvage chemotherapy, the null hypothesis will be constructed:

- H_{03}: CR rate in ASP2215 arm is lower than or equal to the CR rate in salvage chemotherapy arm

The accompanying alternative hypothesis is:

- H_{13}: CR rate in ASP2215 arm is higher than in salvage chemotherapy

To maintain the overall 1-sided Type I error rate at the 0.0245 significance level, the key secondary efficacy endpoint of CR rate will be tested only at second interim and final analyses and only if the null hypothesis on OS (H_{01}) and EFS (H_{02}) are rejected hierarchically. The test significance levels of CR rate endpoint at the second interim and final analyses are specified in Table 4.

The treatment difference along with 95% confidence interval will be calculated using Mantel-Haenszel estimate.

The SAS code to implement the above test will be similar to that shown below:

```
PROC FREQ;
     TABLES stratification variables * response * treatment/cmh;
RUN;
```

CR rate along with the two-sided 95% exact confidence interval based on binomial distribution will be calculated for each treatment group using SAS code similar to that shown below:

```
PROC FREQ;
```

```
      TABLES outcome/BIONOMIAL (EXACT);

      BY treatment;

RUN;
```

The sensitivity analysis for CR rate will be performed as following:

- The same analysis as primary analysis but on FAS
- The same analysis as primary analysis but on PPS
- The same analysis as primary analysis but only including subjects who received at least one dose of study treatment
- The same analysis as primary analysis but only including subjects who had at least one post baseline bone marrow assessment
- Apply the same analysis method used for primary analysis to evaluate CR rate prior to HSCT (as defined in Section 6.1.3.1)
- Un-stratified Fisher's exact test on ITT
- The same analysis as primary analysis for CR rate will be performed by dose adjustment in ASP2215 arm.

7.4.3.2 Other Secondary Efficacy Analysis

At the first interim analysis, Duration of CR/CRh (DCRCRh) will be summarized descriptively by median, corresponding 95% confidence interval and range as estimated from the Kaplan-Meier curve. DCRCRh will be analyzed for subjects who achieved either CR or CRh in the same analysis set of CR/CRh rate. Duration of CRh (DCRh) will also be summarized similarly as duration of CR/CRh but for subjects who achieved CRh only.

Transfusion status (independent vs. dependent) at baseline period and post-baseline period will be summarized in a two-by-two contingency table.

Transfusion conversion rate and transfusion maintenance rate will be calculated based on the transfusion status. 95% confidence interval will be presented for both transfusion conversion and maintenance rates.

LFS and duration of remission (including DCRc, DCR, and DR) will be analyzed for subjects who achieved remission using the stratified log-rank test with strata to control for response to first-line AML therapy and preselected salvage chemotherapy. The hazard ratio of the treatment effect along with 95% confidence interval will be calculated. The survival curve and the median will be estimated using the Kaplan-Meier method and will be reported along with the corresponding 95% confidence interval.

CRc rate and transplantation rate will be analyzed in the same manner as CR rate using CMH test to control for response to first-line AML therapy and preselected salvage chemotherapy. The number and percentage of subject with CRc and transplantations will be summarized for each treatment group along with exact 95% confidence interval based on binomial distribution. The treatment difference along with 95% confidence interval will also be calculated. The CRc rate will be summarized by dose adjustment in ASP2215 Arm.

BFI global fatigue score will be summarized using mean, standard deviation, minimum, maximum and median by treatment group at each visit. Additionally, a within-subject change will be calculated as the post-baseline measurement minus the baseline measurement and summarized in the same way. Change from baseline BFI score will be analyzed using Analysis of Covariance (ANCOVA) including treatment as a fixed factor, baseline BFI score, response to first-line AML therapy and preselected salvage chemotherapy as covariates.

7.4.4 Analysis of Exploratory Endpoints

Central FLT3 mutation status at screening, including subgroups of FLT3-ITD mutation and D835/I836 TKD mutations, will be summarized by the number and percentage of subjects in each category by treatment group. An exploratory analysis of FLT3 mutation type and selected clinical efficacy endpoints will be conducted (see Section 7.8).

Additional biomarker exploratory analysis on AXL status, FLT3 related biomarkers and other genes mutations will be performed and the details will be included in a separate Biomarker Statistical Analysis Plan (Biomarker SAP).

Incidence of resource utilization including hospitalization, blood transfusion, antibiotic intravenous infusions, medication for AEs and opioid medication during the entire study period will be summarized by treatment group. The difference between treatment groups will be tested using CMH test while controlling for response to first-line AML therapy and preselected salvage chemotherapy. Duration of hospital stays, medication for AEs and opioid medication will be summarized by treatment group using descriptive statistics (mean, standard deviation, minimum, maximum and median). The difference between treatment groups will be tested with ANOVA while controlling for response to first-line AML therapy and preselected salvage chemotherapy.

FACIT-Dys-SF domain scores will be summarized by treatment group at each visit using descriptive statistics (mean, standard deviation, minimum, maximum and median).Additionally, a within-subject change will be calculated as the post-baseline score minus the baseline score and summarized in the same way. ANCOVA model will be used to analyze the change in the FACIT-Dys-SF domain scores from baseline to post-baseline visits including treatment as a fixed factor, baseline score, response to first-line AML therapy and preselected salvage chemotherapy as covariates.

FACT-Leu global score and domain scores will be summarized by treatment group at each visit using descriptive statistics (mean, standard deviation, minimum, maximum and median). Additionally, a within-subject change will be calculated as the post-baseline score minus the baseline score and summarized in the same way. ANCOVA model will be used to evaluate change from baseline to post-baseline visits for the global and domain scores, of the FACT-Leu including treatment as a fixed factor, baseline score, response to first-line AML therapy and preselected salvage chemotherapy as covariates. The same analytic approach will be used for the dizziness and mouth sore items.

EQ-5D-5L VAS will be summarized by treatment group at each visit using descriptive statistics (mean, standard deviation, minimum, maximum and median). Additionally, a

within-subject change will be calculated as the post-baseline score minus the baseline score and summarized in the same way. ANCOVA model will be used to evaluate change from baseline to post-baseline visits for the EQ-5D-5L VAS including treatment as a fixed factor, baseline score, response to first-line AML therapy and preselected salvage chemotherapy as covariates. Shift table showing shift in each dimension score from baseline to each post-baseline visit will be provided for the 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression).

Additional exploratory analysis of patient reported outcomes (BFI, FACIT-Dys, FACT-Leu, dizziness, mouth sore items) and heath outcome (EQ-5D-5L) will be performed and the details will be included in a separate Patient Reported Outcome Statistical Analysis Plan (PRO SAP).

Additional exploratory analysis of resource utilization with regard to health economic model will be performed and the details will be included in a separate Health Economic Statistical Analysis Plan.

7.4.5 Analysis of Other Variables

Best Response will be summarized by treatment group. The number and percentage of subjects in each category will be presented. The best response rate by prior FLT3 inhibitor use will also be summarized.

At the first interim analysis, Time to first and best CR/CRh (TTCRCRh) will be summarized using descriptive statistics (mean, standard deviation, minimum, maximum and median) for subjects who achieved either CR or CRh in the same analysis set of CR/CRh rate.

TTCRc, TTCR, TTR will be summarized by treatment group using descriptive statistics (mean, standard deviation, minimum, maximum and median) for subjects who achieved remission in the ITT.

7.4.6 Analysis of CR/CRh rate, Duration of CR/CRh and Time to first and best CR/CRh at Final Analysis

The CR/CRh rate, Duration of CR/CRh and Time to first and best CR/CRh will be summarized descriptively on both arms at the final analysis.

The CR/CRh rate will be evaluated in the same manner as CR rate using CMH test to control for response to first-line AML therapy and preselected salvage chemotherapy on ITT. The number and percentage of subjects with CR/CRh will be summarized for each treatment group along with exact 95% confidence interval based on binomial distribution. The treatment difference along with 95% confidence interval will be calculated.

The sensitivity analysis for CR/CRh rate will be performed to assess the robustness of the CR/CRh rate as follows:

- The same analysis as primary analysis but on FAS
- The same analysis as primary analysis but on PPS

- The same analysis as primary analysis but only including ITT subjects who took at least one dose of study drug
- The same analysis as primary analysis but only including ITT subjects who had at least one post baseline bone marrow assessment
- Apply the same analysis method used for primary analysis to evaluate CR/CRh rate prior to HSCT (as defined in Section 6.1.2)
- The same analysis as primary analysis for CR/CRh rate will be performed by dose adjustment in ASP2215 arm.

Duration of CR/CRh (DCRCRh) and Time to first and best CR/CRh (TTCRCRh) will be summarized in the same manner as Duration of CR (as specified in Section 7.4.3.2) and Time to CR (as specified in Section 7.4.5), respectively, for subjects who achieved either CR or CRh on ITT. Median follow-up of DCRCRh estimated from reverse Kaplan-Meier curve will be provided.

As a sensitivity analysis, Duration of CR/CRh will also be analyzed with deaths without report of relapse considered as events (as defined in Section 6.1.3.2). In addition, Duration of CR/CRh will also be analyzed by subjects with or without transplantation during the study period.

7.5 Analysis of Safety

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

7.5.1 Adverse Events

All adverse event (AE) recorded on treatment including within 30 days from the last study treatment will be summarized.

Summaries and listings of SAEs and 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,
- Serious TEAEs,
- Drug-related serious TEAEs,
- TEAEs leading to death
- Drug-related TEAEs leading to death
- TEAEs leading to withdrawal of treatment,

- Drug-related TEAEs leading to withdrawal of treatment,
- Grade 3 or higher TEAEs,
- Drug-related Grade 3 or higher TEAEs,
- Any death

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:

- TEAEs
- Drug-related TEAEs,
- Serious TEAEs,
- Drug-related serious TEAEs,
- TEAEs leading to death
- Drug-related TEAEs leading to death
- TEAEs leading to withdrawal of treatment,
- Drug-related TEAEs leading to withdrawal of treatment,
- Grade 3 or higher TEAEs,
- Drug-related Grade 3 or higher TEAEs

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,

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

- TEAEs
- Drug-related TEAEs,

The number and percentage of subjects with TEAE of special safety interest (AESI), as classified by AESI category and PT will be summarized for each 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

AE summary tables will include subject counts as opposed to AE counts. If a subject experiences more than one episode of a particular AE, that subject will be counted only once for that event. If a subject has more than one AE that code 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 body system, the subject will be counted only once in that body system.

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

7.5.2 Clinical Laboratory Evaluation

The baseline visit is the last measurement taken prior to initial study drug administration.

Quantitative clinical laboratory variables, i.e., hematology and biochemistry based on central assessment will be summarized using mean, standard deviation, minimum, maximum and median for each treatment group at each visit. Additionally, a within-subject change will be calculated as the post-baseline measurement minus the baseline measurement 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.

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. 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 local assessment and bone marrow results will be listed only.

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 investigational period 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 during the investigational period will be presented by treatment group.

7.5.3 Vital Signs

The baseline visit is the last measurement taken prior to initial study drug administration.

Vital signs (systolic blood pressure, diastolic blood pressure, pulse rate, and body temperature) will be summarized using mean, standard deviation, minimum, maximum and median by treatment group and visit. Additionally, a within-subject change will be calculated per visit as the post-baseline measurement minus the baseline measurement and summarized by treatment group and visit.

Tables for potentially clinically significant vital signs will be generated using baseline value and highest value obtained during treatment for each subject for each treatment group.

The following potentially clinically significant criteria are defined for each parameter:

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 5 minutes apart. ECGs will be read at the site for clinical decision making and transmitted to a central reviewer. Data from the central reviewer will be used in summary presentations. The three values of each ECG parameter within a time point from the central reviewer will be averaged to determine time-specific parameter for a subject, and used in summaries.

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 treatment visit and time point for values of clinical importance using the range criteria below.

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

Number and percent of subjects with maximum postbaseline QTcF > 480 msec and the time to maximum post baseline QTcF value > 480 msec will be summarized by treatment group.

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 status at each assessment time will be provided by treatment group. Negative change scores indicate an improvement and positive scores indicate a decline in performance.

7.5.7 Ophthalmologic Assessment

Quantitative ophthalmologic variable (logmar score for visual acuity) will be summarized using mean, standard deviation, minimum, maximum and median by treatment group at each visit for each eye.

7.6 Analysis of PK

PK analysis will be conducted on the PKAS.

Plasma concentrations of ASP2215 will be summarized by treatment group and where appropriate by nominal time points using descriptive statistics, including number of subjects,

mean, standard deviation, minimum, median, maximum, geometric mean, and coefficient of variation (CV).

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 PD

Not applicable.

7.8 Subgroups of Interest

There will be no efficacy analysis and subgroup analysis performed for the end of study report.

Co-primary efficacy endpoints (OS and CR/CRh rate) and key secondary endpoints (EFS and CR rate) will be summarized by treatment groups at the appropriate interim/final analysis timing for the subgroups defined on the basis of the categorized variables listed below:

Grouping variable	Subgroups
Age group	< 65 years
	>=65 years
Sex	Female
	Male
Race	White
	Black or African American
	Asian
	Other
Baseline ECOG	0-1
	>=2
Region	North America
	Europe (including Turkey and Israel)
	Asia
Central FLT3 Mutation Type	FLT3-ITD alone
	FLT3-TKD alone
	FLT3-ITD & TKD
	Others (Unknown/Missing/Negative)
Response to First-line Therapy	Relapse within 6 months after allogeneic HSCT
	Relapse after 6 months after allogeneic HSCT
	Primary refractory without HSCT
	Relapse within 6 months after CRc and no HSCT
	Relapse after 6 months after CRc and no HSCT
Salvage Chemotherapy	High intensity chemotherapy (FLAG-IDA, MEC)
	Low intensity chemotherapy (LoDAC or azacitidine)
Prior Use of FLT3 inhibitor	Yes
	No
Cytogenetic Risk Status	Favorable
	Intermediate
	Unfavorable
	Other

Pooled stratification factors, if used for stratified analysis, will also be included in subgroup analysis.

7.9 Other Analyses

7.9.1 PK-PD Analysis

There will be no exploratory PKPD analysis performed for the end of study report.

Exploratory PKPD analysis will be performed on safety and efficacy variables. Graphical approaches will be performed on selected variables. Modeling (including mixed effect modeling) may be performed if warranted.

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

To evaluate whether ASP2215 is particularly beneficial or harmful compared to the benchmark data or the salvage chemotherapy group while the study is ongoing, two interim analyses are planned. At the interim analyses, production of tables and listings will be performed by an independent data analysis center (IDAC). The interim analysis outputs will be reviewed by IDMC. No member of the clinical study team will have access to the tables and listings created for the IDMC. For more details consult the IDMC Charter.

First Interim Analysis

The first interim analysis is planned when approximately 141 subjects are randomized into ASP2215 arm and at least 112 days (4 treatment cycles) post first dose or randomization (for subjects who received no study drug. At the first interim analysis, only the co-primary endpoint of CR/CRh rate will be evaluated for subjects in ASP2215 arm. The two-sided 95% exact confidence interval of CR/CRh rate will be calculated for approximately 141 subjects who are randomized into ASP2215 arm and at least 112 days (4 treatment cycles) post first dose or randomization (for subjects who received no study drug), i.e., ASP2215 group in RAS. The IDMC will inform Sponsor the outcome on CR/CRh endpoint as favorable or unfavorable at the first interim analysis, using the decision rules specified in table below:

Decision Rules based on lower limit of two-sided 95% exact confidence interval of CR/CRh rate in ASP2215 arm at First Interim Analysis

CR/CRh Rate with two-sided 95% exact confidence interval	
Lower limit <=12%	Lower limit > 12%
Unfavorable	Favorable

The Sponsor may prepare a regulatory submission for conditional approval using the first interim analysis data. However, the study conduct will not be impacted by the first interim analysis result.

A nominal 1-sided p-value 0.0005 (i.e., 2-sided p-value 0.001) which is arbitrarily selected, will be spent to acknowledge the single-arm CR/CRh rate evaluation at the first interim analysis and will not be recycled in the second interim analysis and final analysis. No formal hypothesis testing will be conducted for CR/CRh rate for the study, only the descriptive statistics including two-sided 95% exact CI of CR/CRh rate will be provided to IDMC at the first interim. The hypothesis testing for the primary and secondary endpoints of OS, EFS and CR rate at the secondary interim analysis and final analysis are well controlled at an overall 1-sided type I error rate 0.0245 based on a sequential testing procedure.

Second Interim Analysis

The second interim analysis will be performed when approximately 50% of the total planned death events (death event = 129) have occurred for the study. OS will be tested at 1-sided 0.00147/0.38674 significant level for efficacy/futility according to the O'Brien-Fleming type alpha/beta spending function. The IDMC may recommend terminating the trial for favorable or unfavorable results at the second interim analysis based on OS endpoint, using the decision rules specified in table below:

Decision Rules based on 1-sided P-Value obtained at Second Interim Analysis

Overall Survival		
P-value ≥ 0.38674	0.38674>P-value ≥ 0.00147	P-value < 0.00147
Unfavorable: Trial may be stopped for Futility	Trial continues	Favorable: Trial may be stopped for Efficacy

The futility bound of this study is non-binding and is considered guidance rather than strict bound. Depending on the recommendations of the IDMC, the Sponsor may prepare a regulatory submission, if the outcome on OS at the second interim analysis is favorable.

If the study is not stopped after the second interim analysis, a final analysis based on OS will occur after 100% of the total planned death events (death events = 258) have been observed. The 1-sided significance level for the final analysis is 0.02402 for OS.

According to the sequential testing procedure on controlling overall type I error rate for multiple endpoints, EFS will be tested when the null hypothesis of OS is rejected at the second interim or at the final analyses. CR rate will be tested when the null hypotheses of both OS and EFS are rejected hierarchically at the second interim or at the final analyses. Their significance levels at the second interim and final analyses will be based on Pocock/ O'Brien-Fleming alpha/beta spending function. The details about the significance levels at each interim and final analyses for each co-primary and secondary endpoints are specified in Section 7.4 Table 4.

Details for the two formal interim analyses and the IDMC safety review procedures will be included in either Interim Analysis Plan (IAP) or IDMC Charter. The statistical analysis plan for the interim analyses will be described in interim analysis plan (IAP). Safety data including AE, lab, vital signs and ECG will be reviewed by IDMC on a periodic basis. The procedures for IDMC safety review will be described in the IDMC charter.

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

7.11.1 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 6.1.1.

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 6.1.3.3. 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	n/a	1	n/a	1	1
	≠ 1st dose *yyyymm*		2	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.

7.11.2 Outliers

All values will be included in the analyses.

7.11.3 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.

7.11.4 Pooling Strata

For OS and EFS analyses, if all events are from one treatment group in at least one stratum combination, or the Cox proportional hazard model does not converge due to small event size in some stratum combinations, the stratum combinations will be pooled per following steps until the issue is resolved or the normal (un-stratified) Cox proportional hazard model is applied.

Step1: Pooling within preselected chemotherapies

Within each preselected chemotherapy (high intensity or low intensity), if the criterions described above don't meet, pool the 5 levels of response to first-line therapy into 3 levels:

- Relapse after allogeneic HSCT
- Primary refractory without HSCT
- Relapse after CRc and no HSCT

After the above pooling, if the criterions still don't meet, pool the 3 levels into 2 levels:

- Relapse after allogeneic HSCT or after CRc without HSCT
- Primary refractory without HSCT

After the above pooling, if the criterions still don't meet, pool the 2 levels into 1 level:

- Relapse after allogeneic HSCT or after CRc without HSCT or Primary refractory without HSCT

Step2: Pooling across preselected chemotherapies

If after Step1 pooling, the criterions still don't meet, pool the preselected high and low dose chemotherapies into one level:

- High intensity chemotherapy (FLAG-IDA, MEC) or Low intensity chemotherapy (LoDAC or azacitidine)

After the above pooling, re-evaluate the pooling of response to first-line therapy after merging the subjects from both preselected high and low intensity chemo groups, i.e., start the pooling from the 5 levels and follow the similar order described in Step1 until the criterions meet.

For CR rate analysis using CMH test, similar pooling strategy will be applied when all the CR subjects are from one treatment group appears in at least one stratum combination. For CR/CRh rate analysis using CMH test, if all the CR/CRh subjects are from one treatment group in at least one stratum combination, the same pooled strata used for CR will be applied.

All sensitivity analyses for the above primary and key secondary endpoints will apply the same pooled strata used for the primary analysis. In the case that the criterions don't meet, the un-stratified analysis will be used. For other secondary efficacy endpoints, if stratified analysis cannot be conducted based on original stratification factors, un-stratified analysis will be used.

7.11.5 Blinding

Although the study is an open label study, to maintain trial integrity and increase the credibility of study results, the sponsor statisticians' access to the randomized treatment assignment information will be limited. 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 (more details can be found in Section 10.2):

- The study statistician, supporting statisticians and statistical programmers will have no access to the randomized treatment information before the database lock. The randomized treatment code will not be transferred to study statistical programmers, study statistician and supporting statisticians before the database lock. Instead, datasets with scrambled treatment code will be used to prepare analysis programs.
- Study manager and other study team members may have the access to the treatment information at the individual subject level.
- No by treatment summary and treatment difference will be generated during the study, except planned interim analysis conducted by IDAC, or early submission package preparation conducted by an external independent data analysis team.

8 DOCUMENT REVISION HISTORY

Version	**Date**	**Changes**	**Comment/rationale for change**
1.0	24-Aug-2015	NA	Document finalized
2.0 (Amendment 1)	23-Sep-2017	Updated Schedule of Assessment and footnote	Updated due to the following protocol amendments: Protocol amendment 4 dated 9-Dec-2015; Protocol amendment 7 dated 8-Aug-2016
		Updated Section 3 Study objectives and design: • Added another primary objective to evaluate CR/CRh rate; • Added secondary objective to evaluate CRh rate; • Added secondary objective to evaluate transfusion conversion and maintenance rates • Added another interim analysis for IDMC to evaluate the CR/CRh rate	Updated due to Protocol amendment 8 dated 20-Sep-2017
		Updated Section 3 Study design: Modified description on high dose chemo therapy	Updated due to Protocol amendment 7 dated 8-Aug-2016
		Updated Section 4 Sample size to add the sample size justification to evaluate CR/CRh at the first interim analysis	Updated due to protocol amendment 8 dated 20-Sep-2017
		Updated Section 5 Analysis Sets: • Added RAS and mRAS for the purpose of evaluate CR/CRh at first interim analysis; • Changed the primary efficacy analysis set from FAS to ITT	Updated due to protocol amendment 8 dated 20-Sep-2017
		Updated Section 6 Analysis variables: • Added co-primary endpoint of CR/CRh rate; • Added the definition of transfusion conversion and maintenance rates;	Updated due to protocol amendment 8 dated 20-Sep-2017

Version	**Date**	**Changes**	**Comment/rationale for change**
		• Modified response definition per protocol change; • Added duration of CR/CRh; • Added time to CR/CRh; • Modified the treatment failure event time to be randomization time for EFS endpoint; • Deleted the adverse event during HSCT period;	
		Updated Section 6 Analysis variables: Clarified the definition of last contact day for OS endpoint;	Previous definition of last contact day didn't cover all possible dates to support subject alive status
		Updated Section 6 Analysis variables: Clarified the derivation of exposure related variables including duration of exposure, cumulative dose, dose intensity and relative dose intensity	Previous definition may not be appropriate for chemotherapy
		Updated Section 6 Analysis variables: Clarified the exclusion of HSCT period when defining the concomitant medication and transfusion	Previous definitions were not clear
		Updated Section 7 Statistical Methodology: Added analysis method for clinical/Diagnostic procedure and subsequent AML therapy	Due to new CRF pages added
		Updated Section 7 Statistical Methodology: Removed treatment compliance analysis due to unreliable drug accountability data	Updated due to Protocol amendment 7 dated 8-Aug-2016
		Updated Section 7 Statistical Methodology: • Added a table to summarize the timing, sample size and decision guidance for all planned analyses; • Changed the primary analysis method from Cox proportional hazard model to stratified log-rank test;	Updated due to protocol amendment 8 dated 20-Sep-2017

Version	**Date**	**Changes**	**Comment/rationale for change**
		• Added several new sensitivity analyses for OS, EFS and CR rate; • Changed the hypothesis from 2-sided to 1-sided for OS, EFS and CR rate • Added the entire section to describe the CR/CRh rate analysis method; • Added the analysis for duration of CR/CRh and time to CR/CRh; • Removed the analysis of AE, SAE during the on-study HSCT period; • Added the grade 3 and 4 summary for selected lab test parameters; • Added the CR/CRh endpoint for subgroup analysis; • Modified the category for central FLT3 mutation type and region for subgroup variables; • Modifed the interim analysis section; • Modified the blinding section	
		Updated Section 7 Statistical Methodology: Added how to handle partial date in subsequent AML therapy; Modified the pooling strata section	For clarification purpose
		Updated Section 7 Statistical Methodology: Removed the visit window table	It is redundant
3.0 (Amendment 2)	27-SEP-2018	Updated Section 5.4.1 Reasons for Exclusion From PPS to add two criteria: - Neither central FLT3 mutation nor local FLT3 mutation with rapidly progressing disease at baseline -No active relapse at baseline	To clarify the detailed Reasons for Exclusion From PPS criteria.
		Updated Section 5.7 Pharmacokinetics Analysis Set (PKAS) to clarify that the PKAS is for ASP2215 patients only and both	To clarify the detailed PKAS criteria.

Version	**Date**	**Changes**	**Comment/rationale for change**
		dosing date/time and sampling date/time must be available.	
		Updated Section 6.1.3.1 and 7.4.3.1 to clarify the definition of EFS and add additional sensitivity analysis: -Clarify that only death from any cause within 30 days after the last dose of study drug will be considered as EFS event. -Clarify that only subject that discontinued the treatment and has achieved best response of PR or NR will be considered as treatment failure in primary EFS analysis. -Added the new sensitivity analysis to define only subject that discontinued the treatment and has achieved only best response of NR as treatment failure -Added the new sensitivity analysis to define only subject that discontinued the treatment and has achieved best response of PR or NR or NE as treatment failure	Clarify the EFS definitions regarding to handling of deaths (for EFS event of death) and PR (for EFS event of treatment failure).
		Updated Section 6.1.3.1 and 7.4.3.1 to add sensitivity of CR rate prior to HSCT.	Add the sensitivity analysis to keep consistent with analysis on CR/CRh rate.
		Updated Section 6.1.3.2 to add the alternate definition of Duration of CR/CRh to consider deaths without report of relapse as events for sensitivity analysis.	Add the sensitivity analysis per regulatory review comments.
		Updated Section 6.1.3.2: -To add sensitivity of Transfusion conversion rate and transfusion maintenance rate to include transfusion status of not-evaluable post-baseline as transfusion dependent -To clarify transfusion conversion rate and transfusion maintenance rate will only be defined for subjects in ASP2215 arm	Add the sensitivity analysis per regulatory review comments.
		Updated Section 6.2 and added Section 10.3 to clarify Adverse events of special safety interest (AESI) searching strategy	Include AESI for more comprehensive analysis on AEs.

Version	**Date**	**Changes**	**Comment/rationale for change**
		Updated Section 6.4: -To clarify the post-baseline BSA per cycle will be used to calculate planned dose intensity -To correct typos in dose unit -To add the definition of Prior Use of FLT3 inhibitor -To clarify end date for concomitant medications and concomitant transfusions	Clarify the details in variables defined in this section.
		Updated Section 7.2.3 and 7.8 to add two subgroups prior use of FLT3 inhibitor and cytogenetic risk status for demographics and subgroup analysis.	Add two new subgroups per knowledge learned in IA1.
		Updated Section 7.4.4 to clarify: -The additional biomarker exploratory analysis will be covered by a separate Biomarker SAP. -The exploratory analysis of resource utilization, patient reported outcomes (BFI, FACIT-Dys, FACT-Leu, dizziness, mouth sore items) and heath outcome (EQ-5D-5L) will be conducted in the separate Patient Reported Outcome Statistical Analysis Plan (PRO SAP) and Health Economic Statistical Analysis Plan.	Adjust the contents covered by main SAP and other supplemental analysis plan.
		Added a new Section 7.4.6 to clearly describe CR/CRh related summaries at final analysis.	Clarify how CR/CRh rate will be summarized at primary hardlock.
		Updated Section 7.5.1 to include additional summaries of AEs and AE summaries by patient-year.	Update to keep consistent with ISS.
		Updated 7.11.1 to include the imputation rule of partial date of initial diagnosis of AML	Clarify the missing date handling.
		Updated 7.11.4 to clarify that strata pooling will only be performed for OS, EFS, CR rate (and CR/CRh rate) as appropriate.	Clarify details in strata pooling.
		Added Section 10.4 to cover Region Specific Analyses of Japan	Add this section to fulfill PMDA request.
4.0 (Amendment 3)	27-MAY-2022	Specified that there will be no efficacy analysis performed for the end-of-study report.	Summaries of the primary, secondary, and exploratory efficacy endpoints results were already produced and

Version	**Date**	**Changes**	**Comment/rationale for change**
			available in the primary CSR dated 24 JAN 2019.

9 REFERENCES

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, Kopecky KJ, Büchner T, Willman CL, Estey EH, 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)

Karnofsky DA, Burchenal JH. “The Clinical Evaluation of Chemotherapeutic Agents in Cancer." In: MacLeod CM (Ed), *Evaluation of Chemotherapeutic Agents.* Columbia Univ Press. 1949;196.

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: Approvers

Author and Approver Signatories

(E-signatures are attached at end of document)

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

PPD Global Data Science was the GSTATL for this study and biostatistics peer reviewer of this Statistical Analysis Plan

This Statistical Analysis Plan was approved by:

PPD

Global Medical Science

10.2 Appendix 2: Blinding Process

As specified in SAP Section 7.11.5, 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 in study database will be limited. This will reduce potential bias due to the sponsor knowing the treatment effect due to unintentional efficacy and safety summary by treatment.

For the purpose of controlling unintentional assessment of efficacy and safety summary by treatment, the study statisticians and programmers will have no access to treatment codes until the time of final database lock (DBL). For this study, treatment codes can be obtained from two sources: IRT vendor and EDC system. The following procedures should be applied when handling data extraction, data transferring and TLFs preparation to maintain the blinding during the entire study conducting period before the final DBL.

- The treatment codes collected in EDC system will not be extracted until the final DBL. The treatment variable will be blocked during any data extraction before the final DBL. The option for blocking treatment variable will be set up before the first data extraction.
- Dummy treatment codes will be created by study programmers for preparing the SDTM, ADaM and TLFs for interim analysis validation and final data analysis.
- Raw data (SDTM/ADAM if needed) without treatment codes (or with dummy codes) will be provided by the study team to IDAC to conduct interim analysis or other IDMC requested analysis, and may be provided to an external independent data analysis team to conduct analysis for submission based on first interim analysis data before final DBL.
- Separate IRT transfer requests will be sent directly by Independent Contact Statistician to IRT vendor for transferring treatment codes. The Independent Contact statistician will transfer the treatment codes to IDAC during the formal interim analysis or other IDMC requested analysis, and may transfer the treatment codes to an external independent data analysis team to prepare the early submission package before final DBL.
- Dosing or other data (including PK data) which may reveal treatment assignment will be extracted/transferred to an unblinding folder which is only accessible to the unblinding programmers.

Study manager and other study team members may have the access to the treatment assignment information at the individual subject level. Efficacy and safety information will not be summarized by treatment during the study, except at the planned interim analysis or other IDMC requested analysis conducted by IDAC, or early submission package preparation conducted by an external independent data analysis team.

10.3 Appendix 3: Search Strategy for Adverse Events of Interest

Risk	**Search Strategy: MedDRA Version 23.0**
Anaphylactic reaction	Anaphylactic reaction (SMQ Broad)
Cardiac failure	Cardiac failure (SMQ Narrow)
Creatine phosphokinase increased	Rhabdomyolysis/ myopathy (SMQ Narrow)
	Blood creatine phosphokinase abnormal (PT=10005468 and Grade >=3)
	Blood creatine phosphokinase increased (PT=10005470 and Grade>=3)
	Blood creatine phosphokinase MM increased (PT=10005477 and Grade>=3)
	PT: Myalgia
	PT: Myositis
	PT: Muscular weakness
Diarrhea	Noninfectious diarrhoea (SMQ Broad)
Differentiation Syndrome*	PT: Acute interstitial pneumonitis, Acute kidney injury, Acute lung injury, Acute pulmonary oedema, Acute respiratory distress syndrome, Acute respiratory failure, Anuria, Atypical pneumonia, Blood creatinine increased, Blood pressure systolic decreased, Body temperature increased, Capillary leak syndrome, Cardiopulmonary failure, Cardiorenal syndrome, Cardiorespiratory distress, Cough, Differentiation syndrome, Dyspnoea, Febrile neutropenia, Fluid overload, Fluid retention, Generalised oedema, Hepatorenal failure, Hydraemia, Hypervolaemia, Hypotension, Lower respiratory tract infection, Lower respiratory tract inflammation, Lung infection, Lung infiltration, Multiple organ dysfunction syndrome, Noncardiogenic pulmonary oedema, Oedema, Oedema peripheral, Pericardial effusion, Pleural effusion, Pneumonia, Pneumonitis, Prerenal failure, Pulmonary congestion, Pulmonary oedema, Pulmonary toxicity, Pyrexia, Renal failure, Renal impairment, Renal injury, Respiratory arrest, Respiratory distress, Respiratory failure, Weight increased PTs
Gastrointestinal obstruction	Gastrointestinal obstruction (SMQ Narrow)
Gastrointestinal perforation	Gastrointestinal perforation (SMQ Narrow)
Liver transaminase increased	Liver related investigations, signs, and symptoms (SMQ Narrow)

Pancreatitis	Acute pancreatitis (SMQ Broad)
Pericarditis/Pericardial effusion	HLT Noninfectious pericarditis
	PT Pericardial effusion
PRES	Noninfectious encephalopathy/delirium (SMQ Narrow)
QT Prolongation	Torsade de pointes/QT prolongation (SMQ Narrow)
Teratogenicity and Embryo-Fetal Deaths	SMQ Broad-All Pregnancy

*Only AEs occur within the first 90 days

10.4 Appendix 4: Japan Specific Analysis

10.4.1 Introduction

This appendix describes the planned regional specific analyses in addition to the analyses defined by the main body of this SAP. The regional analyses specified in this appendix are for the PMDA submission only.

10.4.2 Region Specific Analyses of Japan

The following regional specific analyses of Japan will be performed as appropriate:

- All analyses as described in the main body of this SAP will be repeated for the subgroups of Japan vs Non-Japan
- All listings will be presented by sites ordered as Japan followed by Non-Japan sites

Below shows a couple of examples where the regional specific analyses of Japan will be handled differently from what are described in the main body of this SAP. The exceptions will be reflected in the corresponding TLF specifications, where the detailed Specifications for table, figure, and data listing formats can be found.

Example(s) of Region Specific Analyses of Japan with Difference

Section(s)	Change(s)	Comment/rationale for change
7.2.3	Geographical Region will not be included in the summary of demographic and other baseline characteristics	The summary of demographic and other baseline characteristics will be repeated for the subgroups of Japan vs Non-Japan
7.4	The endpoints of OS, EFS, CR rate and CR/CRh rate will be evaluated using un-stratified analysis only.	Due to the small number of Japan patients, all these efficacy analysis will be performed using un-stratified analysis instead of stratified analysis.
7.8	Subgroup analyses won't be performed for Geographical Region	The subgroup analyses will be repeated for the subgroups of Japan vs Non-Japan