Official Title of Study:

A Phase 1b/2 Open-Label, Randomized Study of 2 Combinations of Isocitrate Dehydrogenase (IDH) Mutant Targeted Therapies Plus Azacitidine: Oral AG-120 Plus Subcutaneous Azacitidine and Oral AG-221 Plus SC Azacitidine in Subjects With Newly Diagnosed Acute Myeloid Leukemia Harboring an IDH1 or an IDH2 Mutation, Respectively, Who Are Not Candidates to Receive Intensive Induction Chemotherapy

PROTOCOL(S) AG-221-AML-005

NCT Number: NCT02677922

Document Date (Date in which document was last revised): March 25, 2020

STATISTICAL ANALYSIS PLAN

A PHASE 1B/2 OPEN-LABEL, RANDOMIZED STUDY OF 2 COMBINATIONS OF ISOCITRATE DEHYDROGENASE (IDH) MUTANT TARGETED THERAPIES PLUS AZACITIDINE: ORAL AG-120 PLUS SUBCUTANEOUS AZACITIDINE AND ORAL AG-221 PLUS SC AZACITIDINE IN SUBJECTS WITH NEWLY DIAGNOSED ACUTE MYELOID LEUKEMIA HARBORING AN IDH1 OR AN IDH2 MUTATION, RESPECTIVELY, WHO ARE NOT CANDIDATES TO RECEIVE INTENSIVE INDUCTION CHEMOTHERAPY

STUDY DRUGS: **AG-120, AG-221**

PROTOCOL NUMBER: **AG-221-AML-005**

DATE FINAL: **25Mar2020**

Prepared by:

[redacted]

on behalf of

Celgene Corporation

86 Morris Avenue

Summit, NJ 07901

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SIGNATURE PAGE

STATISTICAL ANALYSIS PLAN (SAP) AND SAP AMENDMENT APPROVAL SIGNATURE PAGE		
SAP TITLE	AG-221-AML-005 Statistical Analysis Plan	
SAP VERSION, DATE	Version 1.0, 25 March 2020	
SAP AUTHOR	[redacted]	*{See appended electronic signature page}*
	Printed Name and Title	Signature and Date
PROTOCOL TITLE	A PHASE 1B/2 OPEN-LABEL, RANDOMIZED STUDY OF 2 COMBINATIONS OF ISOCITRATE DEHYDROGENASE (IDH) MUTANT TARGETED THERAPIES PLUS AZACITIDINE: ORAL AG-120 PLUS SUBCUTANEOUS AZACITIDINE AND ORAL AG-221 PLUS SC AZACITIDINE IN SUBJECTS WITH NEWLY DIAGNOSED ACUTE MYELOID LEUKEMIA HARBORING AN IDH1 OR AN IDH2 MUTATION, RESPECTIVELY, WHO ARE NOT CANDIDATES TO RECEIVE INTENSIVE INDUCTION CHEMOTHERAPY	
INVESTIGATIONAL PRODUCTS	AG-120, AG-221	
PROTOCOL NUMBER	AG-221-AML-005	
PROTOCOL VERSION, DATE	Final: 29-Sep-2015, Amendment 1.0: 04-Feb2016, Amendment 2.0: 25-Oct-2016	
SIGNATURE STATEMENT	By my signature, I indicate I have reviewed this SAP and find its contents to be acceptable.	

Statistical Therapeutic Area Head

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date _____

Lead Clinical Research Physician / Clinical Research Physician

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date _____

Lead Product Safety Physician

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date

SIGNATURE PAGE

STATISTICAL ANALYSIS PLAN (SAP) AND SAP AMENDMENT APPROVAL SIGNATURE PAGE		
SAP TITLE	AG-221-AML-005 Statistical Analysis Plan	
SAP VERSION, DATE	Version 1.0, 25 March 2020	
SAP AUTHOR	[redacted]	*{See appended electronic signature page}*
	Printed Name and Title	Signature and Date
PROTOCOL TITLE	A PHASE 1B/2 OPEN-LABEL, RANDOMIZED STUDY OF 2 COMBINATIONS OF ISOCITRATE DEHYDROGENASE (IDH) MUTANT TARGETED THERAPIES PLUS AZACITIDINE: ORAL AG-120 PLUS SUBCUTANEOUS AZACITIDINE AND ORAL AG-221 PLUS SC AZACITIDINE IN SUBJECTS WITH NEWLY DIAGNOSED ACUTE MYELOID LEUKEMIA HARBORING AN IDH1 OR AN IDH2 MUTATION, RESPECTIVELY, WHO ARE NOT CANDIDATES TO RECEIVE INTENSIVE INDUCTION CHEMOTHERAPY	
INVESTIGATIONAL PRODUCTS	AG-120, AG-221	
PROTOCOL NUMBER	AG-221-AML-005	
PROTOCOL VERSION, DATE	Final: 29-Sep-2015, Amendment 1.0: 04-Feb2016, Amendment 2.0: 25-Oct-2016	
SIGNATURE STATEMENT	By my signature, I indicate I have reviewed this SAP and find its contents to be acceptable. [redacted] I authored this document 04 Jun 2020 17:57:48 +00:00 DocuSign	

Statistical Therapeutic Area Head

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date _____

Lead Clinical Research Physician / Clinical Research Physician

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date _____

SIGNATURE PAGE

STATISTICAL ANALYSIS PLAN (SAP) AND SAP AMENDMENT APPROVAL SIGNATURE PAGE	
SAP TITLE	AG-221-AML-005 Statistical Analysis Plan
SAP VERSION, DATE	Version 1.0, 25 March 2020
SAP AUTHOR	[redacted] *{See appended electronic signature page}* Printed Name and Title — Signature and Date
PROTOCOL TITLE	A PHASE 1B/2 OPEN-LABEL, RANDOMIZED STUDY OF 2 COMBINATIONS OF ISOCITRATE DEHYDROGENASE (IDH) MUTANT TARGETED THERAPIES PLUS AZACITIDINE: ORAL AG-120 PLUS SUBCUTANEOUS AZACITIDINE AND ORAL AG-221 PLUS SC AZACITIDINE IN SUBJECTS WITH NEWLY DIAGNOSED ACUTE MYELOID LEUKEMIA HARBORING AN IDH1 OR AN IDH2 MUTATION, RESPECTIVELY, WHO ARE NOT CANDIDATES TO RECEIVE INTENSIVE INDUCTION CHEMOTHERAPY
INVESTIGATIONAL PRODUCTS	AG-120, AG-221
PROTOCOL NUMBER	AG-221-AML-005
PROTOCOL VERSION, DATE	Final: 29-Sep-2015, Amendment 1.0: 04-Feb2016, Amendment 2.0: 25-Oct-2016
SIGNATURE STATEMENT	By my signature, I indicate I have reviewed this SAP and find its contents to be acceptable.
Statistical Therapeutic Area Head Signature: *{See appended electronic signature page}* [redacted] Date: 2020.04.15 22:58:51 -04'00' Printed Name: [redacted] Date:	
Lead Clinical Research Physician / Clinical Research Physician Signature: *{See appended electronic signature page}* Printed Name: [redacted] Date:	

SIGNATURE PAGE

STATISTICAL ANALYSIS PLAN (SAP) AND SAP AMENDMENT APPROVAL SIGNATURE PAGE	
SAP TITLE	AG-221-AML-005 Statistical Analysis Plan
SAP VERSION, DATE	Version 1.0, 25 March 2020
SAP AUTHOR	[redacted] *{See appended electronic signature page}* Printed Name and Title — Signature and Date
PROTOCOL TITLE	A PHASE 1B/2 OPEN-LABEL, RANDOMIZED STUDY OF 2 COMBINATIONS OF ISOCITRATE DEHYDROGENASE (IDH) MUTANT TARGETED THERAPIES PLUS AZACITIDINE: ORAL AG-120 PLUS SUBCUTANEOUS AZACITIDINE AND ORAL AG-221 PLUS SC AZACITIDINE IN SUBJECTS WITH NEWLY DIAGNOSED ACUTE MYELOID LEUKEMIA HARBORING AN IDH1 OR AN IDH2 MUTATION, RESPECTIVELY, WHO ARE NOT CANDIDATES TO RECEIVE INTENSIVE INDUCTION CHEMOTHERAPY
INVESTIGATIONAL PRODUCTS	AG-120, AG-221
PROTOCOL NUMBER	AG-221-AML-005
PROTOCOL VERSION, DATE	Final: 29-Sep-2015, Amendment 1.0: 04-Feb2016, Amendment 2.0: 25-Oct-2016
SIGNATURE STATEMENT	By my signature, I indicate I have reviewed this SAP and find its contents to be acceptable.

Statistical Therapeutic Area Head

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date _____

Lead Clinical Research Physician / Clinical Research Physician [redacted]

Signature *{See appended electronic signature page}*

Printed Name [redacted] Date [redacted]

Lead Product Safety Physician

Signature [redacted]

Printed Name [redacted] Date 17-April-2020

1. LIST OF ABBREVIATIONS

Table 1: Abbreviations and Specialist Terms

Abbreviation or Specialist Term	Explanation
ADaM	Analysis Data Model
AE	Adverse Event
ALT	Alanine Aminotransferase (SGPT)
AML	Acute Myeloid Leukemia
ANC	Absolute Neutrophil Count
AST	Aspartate Aminotransferase (SGOT)
ATC	Anatomical Therapeutic Chemical
[redacted]	
AZA	Azacitidine
[redacted]	
BL	Baseline
BM	Bone Marrow
BMI	Body Mass Index
BSA	Body Surface Area
BPM	Beats per Minute
C1D1	Cycle 1, Day 1
CI	Confidence Interval
CR	Morphologic Complete Remission
CRh	Morphologic Complete Remission with Partial Hematologic Recovery
CRi	Morphologic Complete Remission with Incomplete Neutrophil Recovery
CRp	Morphologic Complete Remission with Incomplete Platelet Recovery
CRc	Cytogenetic Complete Remission
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DDS	Dose Determining Set

Abbreviation or Specialist Term	Explanation
DEP	DLT – Evaluable Population
DLT	Dose Limiting Toxicity
DMC	Data Monitoring Committee
[redacted]	
DRT	Dose Review Team
EAP	Evaluable Analysis Population
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EFS	Event-Free Survival
ELN	European LeukemiaNet
EORTC	European Organization for Research and Treatment of Cancer
EOT	End of Treatment
FAP	Full Analysis Population
FCBP	Females of Childbearing Potential
HDL-C	High - Density Lipoprotein Cholesterol
[redacted]	
HI	Hematologic Improvement
HI-E	Hematologic Improvement Erythroid Response
HI-N	Hematologic Improvement Neutrophil Response
HI-P	Hematologic Improvement Platelet Response
[redacted]	
HRQoL	Health-Related Quality-of-Life
HSCT	Hematopoietic Stem Cell Transplantation
IC	Intensive Chemotherapy
IDAC	Intermediate-Dose Cytarabine
IDH1	Isocitrate Dehydrogenase Isoform 1
IDH2	Isocitrate Dehydrogenase Isoform 2

Abbreviation or Specialist Term	Explanation
ITT	Intent-to-Treat
[redacted]	
IWG	International Working Group
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LDL-C	Low-Density Lipoprotein Cholesterol
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MDS	Myelodysplastic Syndrome
MLFS	Morphologic Leukemia-Free State
MPN	Myeloproliferative Neoplasms
MR	Morphologic Relapse after CR/CRi/CRp
[redacted]	
MUGA	Multi-Gated Acquisition
NCA	Non-Compartmental
NCI	National Cancer Institute
ORR	Overall Response Rate
OS	Overall Survival
PB	Peripheral Blood
PD	Progressive Disease
PK	Pharmacokinetics
PO	Orally (lat: “per os”)
PR	Partial Remission
PT	Preferred Term
Q1	First Quartile
Q3	Third Quartile
QTcB	Heart-Rate Corrected QT with Bazett’s Correction
QTcF	Heart-Rate Corrected QT with Fredericia’s Correction
RBC	Red blood Cells

Abbreviation or Specialist Term	Explanation
RCD	Recommended Combination Dose
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Stable Disease
SGPT	Serum Glutamic Pyruvic Transaminase
[redacted]	
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TTO	Time Trade-off
ULN	Upper Limit of Normal
[redacted]	
VAS	Visual Analogue Scale
WBC	White Blood Cells
WHO	World Health Organization

2. INTRODUCTION

This statistical analysis plan (SAP) describes the analyses and data presentations for Celgene's protocol AG-221-AML-005 "A PHASE 1B/2 OPEN-LABEL, RANDOMIZED STUDY OF 2 COMBINATIONS OF ISOCITRATE DEHYDROGENASE (IDH) MUTANT TARGETED THERAPIES PLUS AZACITIDINE: ORAL AG-120 PLUS SUBCUTANEOUS AZACITIDINE AND ORAL AG-221 PLUS SC AZACITIDINE IN SUBJECTS WITH NEWLY DIAGNOSED ACUTE MYELOID LEUKEMIA HARBORING AN IDH1 OR AN IDH2 MUTATION, RESPECTIVELY, WHO ARE NOT CANDIDATES TO RECEIVE INTENSIVE INDUCTION CHEMOTHERAPY" which was issued on 29Sep2015 and amended on 04Feb2016 and 25Oct2016. It contains definitions of analysis populations, derived key efficacy variables, and statistical methods for the analysis of efficacy and safety data.

Throughout this SAP, the treatment arms will be referred to as AG-120 500 mg + AZA 75 mg/m^2 in Dose-finding, AG-120 500 mg + AZA in Expansion, AG-221 100 mg + AZA 75 mg/m^2, and AG-221 200 mg + AZA 75 mg/m^2 during Phase 1b dose-finding and AG-120 expansion stage; the treatment arms will be referred to as AG-221 100 mg + AZA 75 mg/m^2 and AZA alone during Phase 2 randomized stage. Treatment arms AG-120 + AZA 75 mg/m^2 in Dose-finding and AG-120 + AZA in Expansion refer to the combination of AG-120 Orally (PO) plus azacitidine (AZA) subcutaneously (SC). Treatment arm AG-221 + AZA refers to the combination of AG-221 PO plus azacitidine SC. Treatment arm AZA alone refers to azacitidine SC alone.

This SAP provides a comprehensive and detailed description of the strategy, rationale, and statistical techniques to evaluate the efficacy and safety endpoints. The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to the analysis of study data prior to database lock for final analysis. There is one formal interim analysis but no pre-specified statistical stopping rule and type 1 error rate adjustment. However, the DMC will review data periodically in order to monitor the drug toxicity and treatment inferiority/superiority. There will be several interim looks throughout the trial. The SAP may be amended as necessary to accommodate a protocol amendment and will be finalized and signed off prior to the clinical database lock for the primary and final analysis. All statistical analyses detailed in this SAP will be conducted using SAS® Version 9.4.

3. STUDY OBJECTIVES

3.1. Primary Objectives

For the Phase 1b Dose-finding Stage:

- To assess the safety and tolerability of the combination treatments of oral AG-120 when administered with subcutaneous (SC) azacitidine and oral AG-221 when administered with SC azacitidine in subjects with newly diagnosed acute myeloid leukemia (AML) with an IDH1 or an IDH2 mutation, respectively, who are not candidates to receive intensive induction chemotherapy (IC).
- To establish the recommended combination dose (RCD) of oral AG-120 and oral AG-221 when administered with SC azacitidine.

For the Phase 1b AG-120 Expansion Stage:

- To assess the safety and tolerability of the combination treatments of oral AG-120 when administered with SC azacitidine in subjects with newly diagnosed AML with an IDH1 mutation who are not candidates to receive intensive IC.

For the Phase 2 AG-221 Randomized Stage:

- To assess the efficacy of oral AG-221 when administered with SC azacitidine versus SC azacitidine alone in subjects with newly diagnosed AML with an IDH2 mutation, who are not candidates to receive intensive IC.

3.2. Secondary Objectives

For the Phase 1b Dose-finding Stage:

- To assess the preliminary efficacy of the combination treatments of oral AG-120 when administered with SC azacitidine and oral AG-221 when administered with SC azacitidine in subjects with newly diagnosed AML with an IDH1 or an IDH2 mutation, respectively, who are not candidates to receive intensive IC.

For the Phase 1b AG-120 Expansion Stage:

- To assess the preliminary efficacy of the combination treatments of oral AG-120 when administered with SC azacitidine in subjects with newly diagnosed AML with an IDH1 mutation, who are not candidates to receive intensive IC.
- To characterize the pharmacokinetics (PK) of oral AG-120 when administered with SC azacitidine.

For the Phase 2 AG-221 Randomized Stage:

- To evaluate the safety of oral AG- 221 when administered with SC azacitidine.
- To characterize the PK of oral AG-221 when administered with SC azacitidine.
- To evaluate the effect of oral AG-221 when administered with SC azacitidine versus SC azacitidine alone on health-related quality-of-life (HRQoL) outcomes.

4. INVESTIGATIONAL PLAN

4.1. Overall Study Design and Plan

This Phase 1b/2 study is an open-label, randomized, multicenter trial to evaluate the safety and efficacy of oral AG-120 + SC azacitidine and oral A-221 + SC azacitidine in subjects with newly diagnosed AML with an IDH1 or an IDH2 mutation, respectively. The study population consists of subjects who are not candidates to receive intensive IC. The study comprises a Phase 1b dose-finding and AG-120 expansion stage and a Phase 2 randomized stage.

Phase 1b Dose-Finding Stage

The Phase 1b segment is a combination dose-finding study to identify the RCD of the oral AG-120 + SC azacitidine in subjects with an IDH1 mutation or oral AG-221 + SC azacitidine in subjects with an IDH2 mutation when administered in combination with SC AZA. Subjects are considered eligible for this trial if they have newly diagnosed AML with IDH1 or IDH2 mutations and present with co-morbidities, declining performance status, or other factors that in the investigators judgment make them not candidates to receive intensive IC.

The Phase 1b segment will evaluate the safety and clinical activity of oral AG-120 or AG-221 administered with SC azacitidine in this population. For the AG-120 combination, upon declaration of the RCD by the dose review team (DRT), an AG-120 expansion cohort of up to 15 subjects will be enrolled at RCD for further safety evaluation and PK sampling.

The Phase 1b dose-finding segment will use a standard "3 + 3" design. Each dose cohort will enroll a minimum of 3 subjects. Study drug will start on Cycle 1 Day 1 (C1D1) at which time daily dosing will begin for AG-120 and AG-221 for days 1-28 of each 28-day cycle. AZA will be administered SC for 7 days of each 28-day cycle starting at Cycle 1 Day 1. If there are multiple subjects in the screening process at the time the third subject within a cohort begins treatment, up to 2 additional subjects may be enrolled with documented approval of the medical monitor.

The safety of dosing during Phase 1b will be evaluated by the DRT (Celgene medical monitor, Celgene lead safety physician, Celgene biostatistician, other Celgene functional area representatives or designees, as appropriate and all active site investigators and/or designees [at sites with a subject who has received study drug]). The DRT will review the emerging safety data from each cohort.

Dose Limiting Toxicity (DLT)-evaluable Subjects

Dose Determining Set (DDS): Subjects who take at least one dose of study drug in phase 1b dose-finding stage and either have a DLT during Cycle 1 regardless of amount of study drug exposure, or have no DLT and complete at least 75% of AG-120 or AG-221 doses (21 out of 28 days) and a minimum of 5 doses of azacitidine, and at least 50% of the planned combination doses for AG-120 or AG-221 and azacitidine administered together (in the same day for 4 out of 7 days) in the first 28 days from Cycle 1 Day 1 and are also considered by the Clinical Study Team to have sufficient safety data available to conclude that a DLT does not occur during Cycle 1. A subject diary will be used during outpatient treatment to record details around AG-120 and AG-221 dosing.

Safety Evaluation for Combination Therapy

This study will use a standard "3 + 3" design for AG-120 and AG-221 dose determination, identifying the RCD dose level. Each dose cohort will plan to enroll 3 DLT-evaluable subjects, starting with Dose Level 1. Dose escalation or de-escalation decisions will be made independently for each IDH inhibitor + SC azacitidine therapy. For AG-120, there is no dose escalation, but 1 dose de-escalation is allowed to Dose Level -1. For AG-221, there is 1 dose escalation to Dose Level 2 and 1 dose de-escalation to Dose Level -1. Due to the lack of overlapping toxicities, AZA will be administered, by the site staff, at 75 mg/m^2 dose for 7 days of the 28-day cycle. The DRT will review all safety data available for each cohort to determine whether a dose adjustment of AZA is warranted for the combination.

Phase 1b AG-120 Expansion Stage

An expansion cohort of approximately 15 IDH1 patients will be enrolled to the AG-120 combination. Subjects enrolled in the AG-120 expansion will receive the AG-120 + azacitidine at the RCD.

Phase 2 AG-221 Randomized Stage

Phase 2 AG-221 Randomized Stage will be referred to "Phase 2" in the SAP body text. The Phase 2 segment of the study is an open label, randomized, 2-arm design to evaluate the efficacy and safety of oral AG-221 with SC azacitidine versus SC azacitidine alone. Eligible subjects must have newly diagnosed AML with IDH2 mutations and, who are not candidates to receive intensive IC based on the investigators assessment.

All subjects discontinued from study treatment for any reason other than withdrawal of consent for follow-up will continue to be assessed for AEs, concomitant medications, concomitant procedures, transfusions, [redacted] response, hematologic improvement, subsequent AML therapies, and survival.

- All subjects discontinued from study treatment for any reason except withdrawal of consent for follow-up or disease progression will continue to be assessed during the Follow-up period of the study for response until disease progression.
- All subjects discontinued from study treatment for any reason except withdrawal of consent for follow-up will continue to be assessed for subsequent AML therapies, and survival.

Length of Study

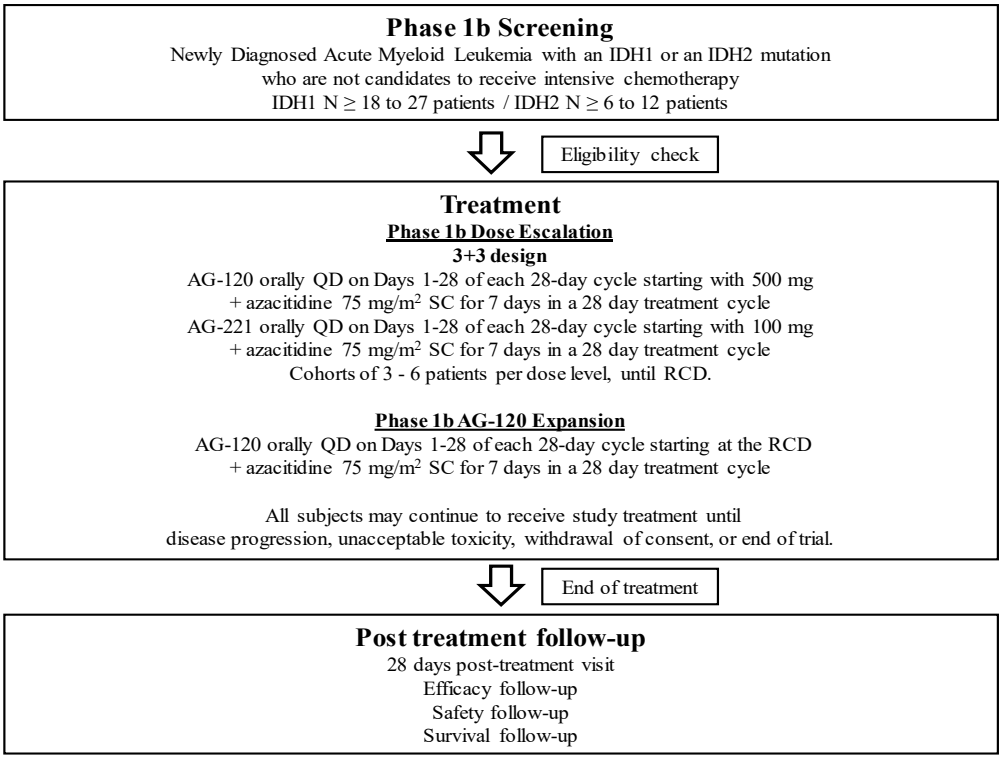
The full length of the study is expected to be approximately 60 months including recruitment, screening, treatment, and follow up for Phase 1b and Phase 2. For a single subject, the expected duration of the Phase 1b segment of the study is approximately 13 months, including a screening period for up to 28 days, and the expected duration of the Phase 2 segment of the study is approximately 30 months, including a screening period for up to 28 days.

The End of Trial is defined as either the date of the last visit of the last subject to complete the post-treatment follow-up, or the date of receipt of the last data point from the last subject that is required for primary, secondary, [redacted] analysis, as pre-specified in the protocol, whichever is the later date.

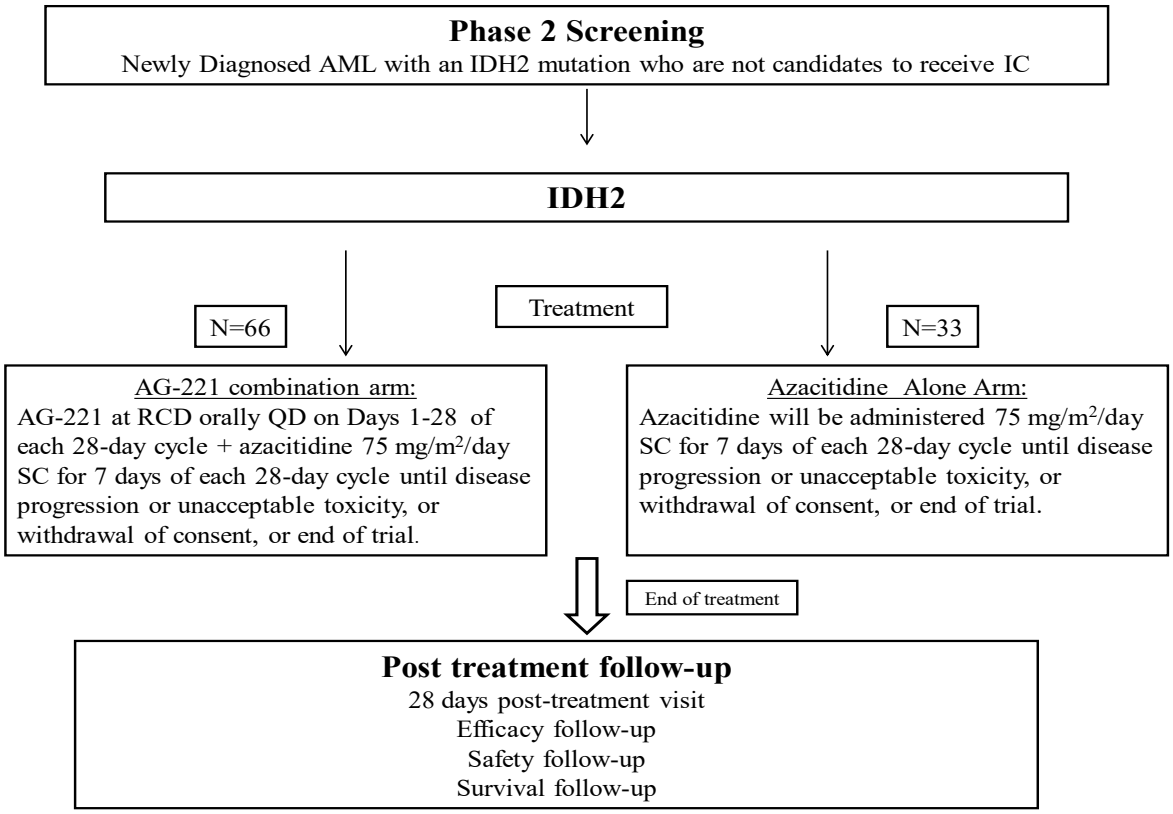
The study schematic is presented in Figure 1 and the table of events is provided in Appendix 17.3 .

Figure 1: Overall Study Design for Phase 1b and Phase 2

Phase 1b:



Phase 2:



Key: AML = acute myeloid leukemia; IDH1 = isocitrate dehydrogenase isoform 1; IDH2 = isocitrate dehydrogenase isoform 2; IC-intensive chemotherapy, RCD= recommended combination dose; PD = progressive disease; QD = once a day; SC = subcutaneously.

4.2. Study Endpoints

4.2.1. Primary Endpoints

Phase 1b:

The primary endpoints for the Phase 1b part are RCD and safety and tolerability.

Phase 2:

The primary endpoint for the Phase 2 part is overall response rate (ORR, as assessed by the investigator), which is defined as rate of morphologic complete remission (CR) + morphologic complete remission with incomplete neutrophil recovery (CRi) + morphologic complete remission with incomplete platelet recovery (CRp) + morphologic leukemia-free state (MLFS) + partial remission (PR) according to modified International Working Group (IWG) AML response criteria (Appendix 17.4).

4.2.2. Secondary Endpoints

The secondary endpoints for Phase 1b and Phase 2 are listed in Table 2.

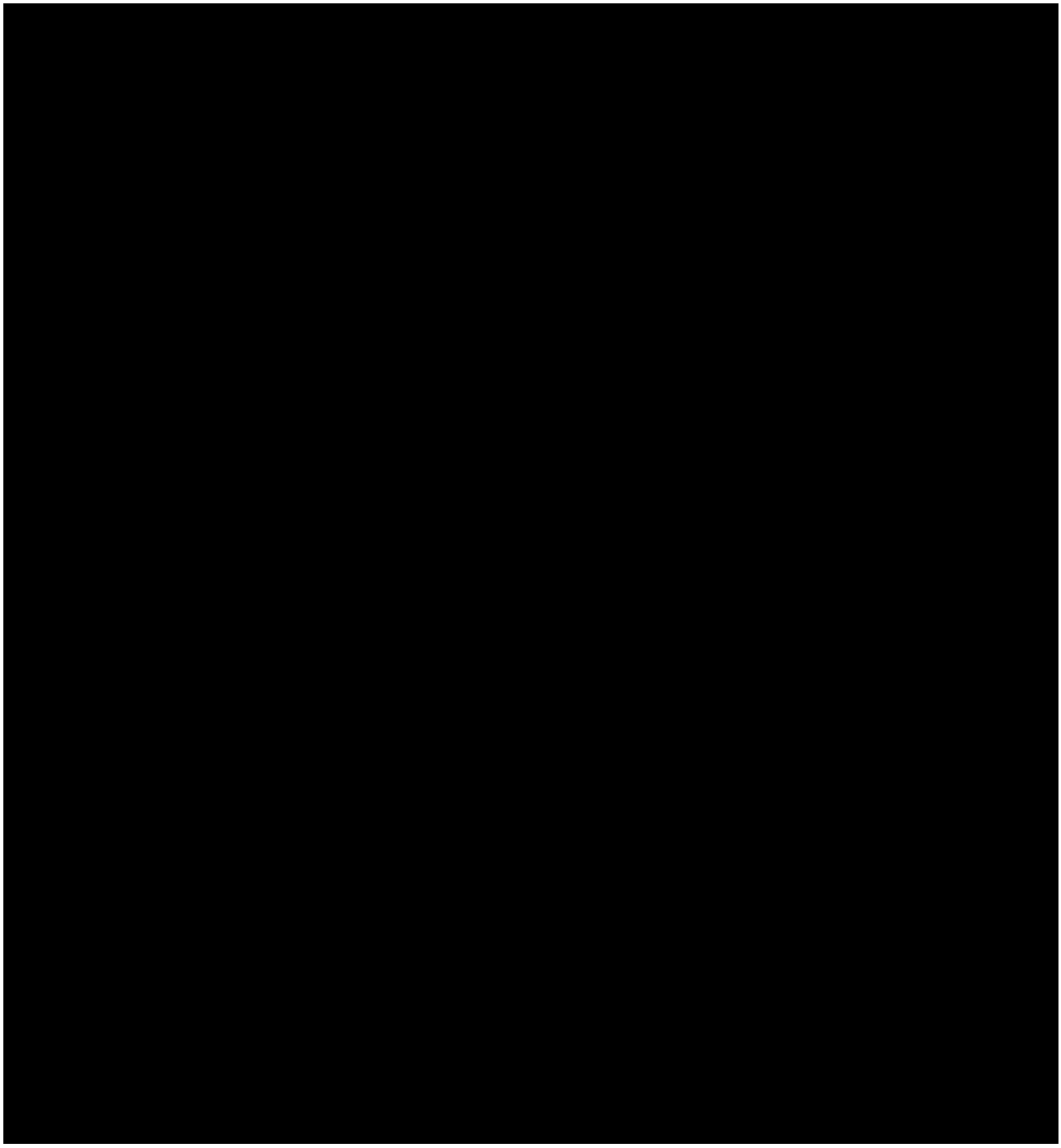
Table 2: Secondary Endpoints

Name	Description	Timeframe
Phase 1b		
Overall response rate (as assessed by the investigator)	Rate of CR + CRi + CRp + MLFS + PR according to modified IWG AML response criteria.	~ 13 months
Sponsor Derived CR and CRh	Rate of CR/ CRh and CR + CRh based on laboratory data.	~ 13 months
PK parameters	Plasma concentrations and pharmacokinetic parameters of AG-120.	~ 13 months
Phase 2		
Event-free survival	Time from randomization to documented morphologic relapse, PD according to modified IWG AML response criteria or death from any cause, whichever occurs first.	~ 30 months
Safety and tolerability	Type, frequency, severity, seriousness and relationship of adverse events (AEs) to study treatments.	~ 30 months
Complete remission rate	Rate of CR according to modified IWG AML response criteria.	~ 30 months

Name	Description	Timeframe
Phase 2, continued		
Sponsor Derived CR	Rate of CR based on laboratory data.	~ 30 months
Sponsor Derived CR and CRh	Rate of CR +CRh based on laboratory data.	~ 30 months
Hematologic improvement rate	Rate of HI-N + HI-P + HI-E according to IWG MDS HI criteria.	~ 30 months
Duration of response	Time from the first documented MLFS/CR/CRi/CRp/PR to documented morphologic relapse, PD according to modified IWG AML response criteria or death due to any cause, whichever occurs first.	~ 30 months
Time to response	Time from first dose of study drug to first documented CR/CRi/CRp/ MLFS/PR according to modified IWG AML response criteria.	~ 30 months
Time to sponsor assessed CR and CRh	Time from first dose of study drug to first documented CR / CRh.	~ 30 months
Duration of sponsor assessed CR and CRh	Time from the first documented CR/ CRh to documented morphologic relapse, progression.	~ 30 months
Overall survival	Time from randomization to death due to any cause.	~ 30 months
One-year survival	The probability of survival at 1 year from randomization.	~ 30 months
PK parameters	Plasma concentrations and pharmacokinetic parameters of AG-120 or AG-221.	~ 30 months
HRQoL outcomes	European Organization for Research and Treatment of Cancer Quality-of-Life questionnaire (EORTC QLQ-C30) and EuroQoL Group EQ-5D-5L instrument.	~ 30 months

Abbreviations: AML = acute myeloid leukemia; CR = morphologic complete remission; CRh = morphologic complete remission with partial hematologic recovery; CRi = morphologic complete remission with incomplete neutrophil recovery; CRp = morphologic complete remission with incomplete platelet recovery; ECG = electrocardiogram; EORTC= European Organization for Research and Treatment of Cancer; HI = hematologic improvement; HI-E = hematologic improvement erythroid response; HI-N = hematologic improvement neutrophil response; HI-P = hematologic improvement platelet response; HRQoL = Health-related Quality of Life;

IWG = International Working Group; MDS = myelodysplasic syndromes; MLFS = morphologic leukemia-free state; PK=pharmacokinetic; PR = partial remission; PD = progressive disease.



4.2.4. Safety Endpoints

Safety assessments include physical examination, vital signs, electrocardiogram (ECG), hematology, serum chemistry, cardiac markers, fasting lipid panel, pregnancy testing (for females of childbearing potential [FCBPs] only), adverse events (AEs), concomitant medications, concomitant procedures and transfusions. After screening, echocardiogram (ECHO)

or multi-gated acquisition (MUGA) scan, urinalysis and coagulation will be repeated as clinically indicated.

4.3. Stratification, Randomization, and Blinding

Randomization only applies to the Phase 2 part and will be carried out using an Interactive Web Response System (IWRS). Subjects with an IDH2 mutation will be randomized to receive oral AG-221 + SC azacitidine (Arm 1) versus SC azacitidine (Arm 2) in 2:1 ratio. Arm 1 will include a minimum of 66 subjects and Arm 2 will include a minimum of 33 subjects, (99 subjects total in both arms).

For the Phase 2 part, all subjects randomized will be stratified by primary (i.e., de novo) or secondary (progression of MDS or myeloproliferative neoplasms [MPN], or therapy-related) AML according to the WHO classification.

During Phase 1b and Phase 2 a retrospective central pathology review of all bone marrow aspirates and/or biopsies and peripheral blood smears, which are collected in subjects of the AG-221 (Phase 1b dose-finding and Phase 2 randomized stage) treatment arms during screening, will be conducted by personnel blinded to subject treatment to confirm eligibility. Bone marrow aspirate and/or biopsy and peripheral blood smear collected after the start of study treatment must be available for both local and central pathology review. The retrospective central pathology review will require a set of duplicate slides for each bone marrow collection time point including bone marrow aspirate (BMA), peripheral blood smear, and bone marrow biopsy (BMB) if performed. The central pathology review will be conducted by personnel blinded to study treatment.

4.4. Sample Size Determination

The Phase 1b dose-finding stage will enroll a total of approximately 24 subjects. Additional subjects may be enrolled in a dose level to replace subjects who are not evaluable for the primary endpoint, or for further exploring safety, PK, PK/pharmacodynamics, or preliminary clinical activity.

In Phase 1b AG-120 expansion stage, to ensure acceptable toxicity at the RCD, up to 15 subjects will be accrued into the AG-120 expansion phase. Based on a sample size of 18 AML subjects treated at RCD (i.e., 15 subjects in the dose expansion and around 3 subjects in the dose escalation), there is 95% probability of detecting 1 or more AEs with an underlying rate of 15%, and 85% probability of detecting 1 or more AEs with an underlying rate of 10%.

In Phase 2 AG-221 randomized stage, subjects with AML with IDH2 mutation will be randomized to receive oral AG-221 + SC azacitidine versus SC azacitidine in a 2:1 ratio. A minimum of 99 subjects will be randomized in this phase with 66 subjects in the oral AG-221 + SC azacitidine arm, and 33 subjects in the SC azacitidine monotherapy arm. The statistical comparisons will be conducted for oral AG-221 + SC azacitidine versus azacitidine monotherapy. Assuming an ORR of 30% in the azacitidine monotherapy and an ORR of 50% for oral AG-221+ SC azacitidine arm, the sample size of 66 subjects in AG-221 + SC azacitidine arm and 33 patients in SC azacitidine monotherapy arm will provide 75% power to detect an 20% difference in ORR at a two-sided type I error rate of 0.2. The multiple comparisons were not considered in the sample size calculation.

[redacted]

5. GENERAL STATISTICAL CONSIDERATIONS

5.1. Reporting Conventions

The following reporting conventions apply generally to tables, listings, and figures:

- Data from all study centers will be combined for analysis;
- In general, analyses for Phase 1b subjects will be done by AG-120 + AZA and AG-221 + AZA and by dose cohort in dose-finding, total in dose-finding, dose cohort in expansion (only for AG-120 + AZA), and total for AG-120 + AZA. For Phase 2, AG-221+AZA versus pooled AZA alone will be compared.
- P-values will be rounded to 4 decimal places; P-values that round to 0.0000 will be presented as '<0.0001' and p-values that round to 1.000 will be presented as '>0.9999
- Confidence intervals (CIs) will be presented as two-sided 95% CIs unless specified differently in specific analysis;
- Summary statistics will consist of the number and percentage of subjects (or cycles, if appropriate) in each category for discrete variables, and the sample size, mean, median, standard deviation, first quartile (Q1), third quartile (Q3), minimum, and maximum for continuous variables;
- All mean and median values will be formatted to one more decimal place than the measured value. Standard deviation values will be formatted to two more decimal places than the measured value;
- All percentages will be rounded to one decimal place. The number and percentage of responses will be presented in the form XX (XX.X%), where the percentage is in the parentheses. In case of zero frequency, the percentage will not be displayed;
- All listings will be sorted for presentation in order of treatment arm, study center, subject, and date/time of procedure or event and will be presented separately for each phase. Subjects who are screened but not randomized will be listed separately as a non-randomized group;
- All analysis and summary tables will have the analysis population sample size (i.e., number of subjects);
- The day of the first dose of any study drug will be defined as Day 1.
- Unless noted otherwise, baseline value will be defined as the last non-missing value prior to the start of study treatment on Day 1 of Cycle 1. For subjects who are randomized but not treated, the baseline will be the assessment value on or prior to randomization date.
- To summarize efficacy/safety data by visit, the post-baseline results will be summarized by the scheduled visit or the derived visit as appropriate.
- Calculation of Cycles:

The start date of each treatment cycle will be determined based on study drug exposure records and visit date for each subject. The start date of the first cycle will be the earliest date the subject receives any study drug. For the other cycles, the start date will be based on scheduled study visit date for each subject. Once the start dates, e.g., S_1, S_2, S_3… are determined, the end date of each cycle is calculated as the day before the start date of the following cycle, i.e., $E_i = S_{i+1} - 1$. The cycle number for each date of interest, e.g., adverse event (AE) or lab will be calculated based on event start date and the cycle window start and end dates. If an event date is on or after S_i and before S_{i+1}, the corresponding cycle number will be i.

In case cycle information is not appropriately collected in the electronic case report forms (eCRFs), then cycle may be derived using study day relative to the first dose day. All by cycle analysis will use the defined cycle.

5.2. Analysis Populations

5.2.1. Intent-to-Treat Population

The intent-to-treat (ITT) population includes all subjects who are randomized to treatment, regardless of whether they received treatment or not. This population is defined for the Phase 2 part only and will be used in most of the data analysis except for safety data analysis.

5.2.2. Modified Intent-to-Treat Population

The modified intent-to-treat (mITT) population includes all subjects who have met all inclusion and exclusion criteria and experienced no major protocol deviations during the study, received at least one cycle of study treatment and had at least one post-randomization treatment response assessment performed. Major protocol deviations leading to exclusion from the mITT population are defined as any of the following:

- Had at least one inclusion/exclusion criteria violation;
- Disease status at screening not confirmed by retrospective central review;
- Initial study treatment given is not the one assigned in randomization;
- Received protocol-prohibited concomitant medication.

This population is defined for the Phase 2 part only.

5.2.3. Safety Population

For Phase 1b part, safety population includes all subjects who were enrolled and received at least one dose of study treatment. Subjects will be classified according to the treatment received, where treatment received is defined as the assigned dose level/schedule if it was received at least once, or the first dose level/schedule received if assigned treatment was never received. The safety population will be used for all safety analyses and drug exposure.

For Phase 2, the safety population includes all randomized subjects who received at least 1 dose of study treatment. Subjects will be analyzed according to the initial treatment actually received.

5.2.4. Pharmacokinetics Population

The PK population includes all subjects who enroll and receive at least one dose of study drug (AG-221/AG-120) and have at least 1 measurable concentration datum of study drug.

5.2.5. Full Analysis Population

Full analysis population (FAP) includes all subjects who were enrolled and received at least 1 dose of study treatment. Subjects will be classified according to the assigned dose level and schedule. FAP is the primary analysis population and will be the default analysis set for all analyses except the safety analyses, unless otherwise specified. This population is defined for the Phase 1b part only.

5.2.6. DLT-Evaluable Population

Subjects who take at least one dose of study drug in phase 1b dose-finding stage and either have a DLT during Cycle 1 regardless of amount of study drug exposure, or have no DLT and complete at least 75% of AG-120 or AG-221 doses (21 out of 28 days) and a minimum of 5 doses of azacitidine, and at least 50% of the planned combination doses for AG-120 or AG-221 and azacitidine administered together (in the same day for 4 out of 7 days) in the first 28 days from Cycle 1 Day 1 and are also considered by the Clinical Study Team to have sufficient safety data available to conclude that a DLT does not occur during Cycle 1. A subject diary will be used during outpatient treatment to record details around AG-120 and AG-221 dosing. Subjects in this population set will be denoted as evaluable for DLT assessment and RCD estimation. This population is defined for the Phase 1b Dose-finding stage only.

5.2.7. Evaluable Analysis Population

For Phase 1b, Evaluable Analysis Population (EAP) includes all subjects in the FAP for whom the baseline response assessment and at least one post baseline response assessment at Day 28 or later are available and evaluable. The clinical activity of AG-221/AG-120 combined with azacitidine will be primarily assessed in the FAP. Additional efficacy analyses may be conducted for the EAP.

For Phase 2, EAP is defined as all subjects in the ITT for whom the baseline response assessment and at least one post baseline response assessment at Day 28 or later are available and evaluable. The clinical activity of AG-221 combined with azacitidine will be primarily assessed in the ITT. Additional efficacy analyses may be conducted for the EAP.

6. SUBJECT DISPOSITION

The total number of subjects screened for both Phase 1b and Phase 2 will be presented. The number and percentage of subjects and the failed inclusion/exclusion criteria for subjects who were screened but not enrolled or randomized will be included in the summary of screen failures. Percentage will be based on all screened subjects.

Subject disposition will be reported separately for Phase 1b and Phase 2. The disposition of subjects will be summarized with counts and percentages by dose cohort for Phase 1b or treatment arm for Phase 2 in the following four analysis populations, respectively:

- FAP (Phase 1b)
- ITT population (Phase 2)
- mITT population (Phase 2)
- Safety population (Phase 2)

Reasons for treatment discontinuation will be summarized with the following categories:

- Death
- Adverse event
- Pregnancy
- Progressive disease
- Lack of efficacy
- Withdrawal by subject
- Non-compliance with study drug
- Lost to follow up
- Study terminated by sponsor
- Transition to commercially available treatment
- Physician decision
- Disease relapse
- Symptomatic deterioration
- Allogeneic HSCT
- Protocol violation (to be specified on eCRF)
- Other (to be specified on the eCRF)

Reasons for study discontinuation will be summarized with the categories below. Same categories for study discontinuation will be summarized for subjects who continue into the follow-up phase.

- Death

- Adverse event
- Progressive disease
- Lack of efficacy
- Withdrawal by subject
- Non-compliance with study drug
- Lost to follow up
- Study terminated by sponsor
- Transition to commercially available treatment
- Physician decision
- Disease relapse
- Symptomatic deterioration
- Other (to be specified on the eCRF)

Summary of subjects enrolled by region, country and site will be provided based on all enrolled and randomized population and FAP population (Phase 1b), respectively, and ITT population (Phase 2).

Subject disposition based on the FAP population (Phase 1b) and ITT population (Phase 2) will be listed, respectively. A listing of reasons for exclusion from the mITT population will be provided based on ITT population (Phase 2). A separate listing will be provided for subjects who were not enrolled/randomized (screen failures) with reasons for screening failure.

7. PROTOCOL DEVIATIONS/VIOLATIONS

The protocol deviations/violations will be identified by sites and assessed by clinical research physician or designee following company standard operational procedure. These protocol deviations/violations will be summarized by dose cohort for the FAP population (Phase 1b) and by treatment arm for the ITT population (Phase 2).

Protocol deviations and violations will be reviewed prior to database lock for CSR. Major protocol deviations will be used to determine the mITT population for Phase 2 subjects (defined in Section 5.2.2).

Protocol deviations and violations will be summarized using frequency tabulations. A by-subject listing of subjects with all protocol deviations/violations in the FAP population (Phase 1b) and ITT population (Phase 2) will be provided.

8. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Demographics and baseline characteristics, prior/concomitant medications and procedures will be reported separately for Phase 1b and Phase 2. Demographic and baseline disease characteristics will be summarized by dose cohort for FAP (Phase 1b) and by treatment arm and overall for the ITT, mITT and Safety populations (Phase 2). In addition, for Phase 2, the Pearson Chi-square test (for categorical parameters with more than 2 categories), Fisher exact test (for binary parameters and parameters which have more than 20% of cells with expected frequencies less than 5) and the Wilcoxon sum rank test (for continuous parameters) will be provided as descriptive purpose. Due to the limited sample size and multiplicity, it may not be suitable to use those p-values to assess the comparability of two randomized treatment arms in each setting for demographic, baseline disease characteristics, and AML diagnosis history.

Individual subject listings will be provided to support the summary tables.

8.1. Demographics

Age (year), height (cm), weight (kg), body mass index (kg/m^2), body surface area (m^2), and other continuous baseline characteristics will be summarized using descriptive statistics (N, mean, standard deviation, Q1, Q3, median, minimum, maximum), while age group (<65 years, ≥ 65-<75 years, ≥75 years), gender, ethnicity, race, geographic region and other categorical variables will be provided using frequency tabulations (count, percent) by dose cohort groups (Phase 1b) and by treatment arm and overall (Phase 2).

Age or year of birth will be recorded on eCRF. Where age is not recorded, age will be calculated as follows: Age = year of informed consent – year of birth.

Body mass index (BMI) will be calculated as follows: BMI (kg/m^2) = weight in kg / (height in m)^2.

If not collected in eCRF, body surface area (BSA) will be calculated as follows: BSA (m^2) = weight (kg)$^{0.425}$ x height (cm)$^{0.725}$/139.2.

8.2. Baseline Disease Characteristics and AML Diagnosis History

The number and percentage of subjects in each of the following categories will be summarized by dose cohort (Phase 1b) and treatment arm (Phase 2) and overall (Phase 1b and Phase 2 separately). Continuous variable will be summarized by mean, standard deviation, median, Q1, Q3, minimum, and maximum.

8.2.1. Baseline Disease Characteristics

Baseline value is defined as the last non-missing data on or before the first dose of the study drug, unless otherwise specified. Baseline values from the parameters below, but not limited to, will be tabulated:

1. IDH mutation type: IDH1 or IDH2, and the sub-type (R172, R140, both R172 and R140) for IDH2 gene mutated subjects;
2. Eastern Cooperative Oncology Group (ECOG) performance status;

3. Eligibility for preselected to receive intensive chemotherapy;
4. Bone marrow blasts (%) as a continuous variable and a category variable (Category I: < 20%; 20% to < 30%; 30% to < 50%; ≥50%; Category II: <5%; 5% to < 25%; 25% to < 70%; ≥ 70%) from the screening BM aspirate sample;
5. Bone marrow blasts (%) from the screening BM biopsy sample (0-4%, 5-9%, 10-19%, >=20%);
6. Blasts percentage from PB smear sample;
7. Cytogenetic Risk Status: better-risk, intermediate-risk, poor-risk, failure (local, central);
8. Baseline values of the following lab parameters: Hemoglobin (<80 g/L , ≥ 80 g/L), platelet count (< 50 x10^9/L, 50 to <100 x10^9/L; ≥ 100 x10^9/L), absolute neutrophil count (ANC) (< 0.5 x10^9/L, 0.5 to < 1 x10^9/L, ≥ 1.0 x10^9/L), white blood cells (WBC) (<15 x10^9/L, 15 to < 30 x10^9/L, ≥ 30 x10^9/L), creatinine clearance (< 45 mL/min, 45 to < 60 mL/min, ≥ 60 mL/min); These parameters will also be summarized as continue variables;
9. Number of units of red blood cell (RBC), whole blood, platelet, plasma, and other (specify) transfusions;
10. Left ventricular ejection fraction (LVEF) by echocardiogram (ECHO) or multi-gated acquisition (MUGA) scan;
11. ECHO/MUGA findings (Normal, Abnormal and Unevaluable).

8.2.2. AML Diagnosis History

1. AML initial diagnosis: primary (i.e., de novo) or secondary (myelodysplastic syndrome (MDS) progression or myeloproliferative neoplasms (MPN), or therapy-related) according to WHO classification. If AML is secondary, number and percentage of subjects who has an antecedent hematologic disorder will be summarized, and type of the antecedent hematologic disorder will be tabulated too;
2. Time (months) from the initial diagnosis of AML, calculated from time of initial diagnosis to randomization date for phase 2 and the enrollment date for phase 1b.

8.2.3. Chromosomal Abnormalities at Baseline

The cytogenetics in each risk status (better-risk, Intermediate-risk, Poor-risk) at baseline will be summarized for Phase 2 only.

8.3. Medical History and Concomitant Disease

The medical and surgical history, as well as concomitant disease will be summarized by the Medical Dictionary for Regulatory Activities (MedDRA) Version 22.0 by system organ class (SOC) and preferred term (PT). Active concomitant disease will be summarized in separate tables.

8.4. Prior/Concomitant/Follow-up Medications/Procedures/Therapies

Prior medications/procedures are defined as medications/procedures/transfusions that were started prior to start of study drug or had a missing start date. If start date is missing and end date is before start of study drug will be considered as prior only.

Concomitant medications/procedures are defined as non-study medications/procedures/transfusions that are started on or after the start of study drug but before the end of the study treatment period or started before the start of the study drug and ended or remain ongoing during the study treatment period. The end of the study treatment period is defined as the earlier date of last dose of study drug + 28 days and the date of death.

Medications/procedures with missing start date but end date on or after start of the study drug and before the end of the study treatment period or with both missing start date and end date is also considered as concomitant medications/procedures/transfusions.

Follow-up medications/procedures are defined as non-study medications/procedures/transfusions that are started on or after the date of the end of the study treatment period.

All non-study medications and therapies will be coded by Anatomical Therapeutic Chemical (ATC) code and PT using the World Health Organization Drug Dictionary latest version. All procedures and surgeries will be coded by the Medical Dictionary for Regulatory Activities (MedDRA) Version 22.0 by system organ class (SOC) and preferred term (PT).

8.4.1. Prior/Concomitant/Follow-up Medications

All prior, concomitant and follow-up medications will be summarized in frequency tabulations (subject counts and percentages) and by World Health Organization ATC third level and PT. Listings will also be provided for general medications and concomitant QT prolonging medications.

Prior/concomitant/follow-up medications will be summarized and listed based on FAP for Phase 1b subjects, and ITT and Safety populations for Phase 2 subjects.

8.4.2. Prior/Concomitant/Follow up Procedures/Surgeries

The prior, and concomitant procedures/surgeries will be summarized in frequency tabulations (subject counts and percentages) by SOC and PT. Listings will also be provided.

Prior/concomitant/follow-up procedures/surgeries will be summarized and listed based on FAP for Phase 1b subjects, and ITT and Safety populations for Phase 2 subjects.

8.4.3. Prior/Subsequent Anti-cancer Therapies

This section includes both AML therapies and non- AML therapies. AML therapies include radiation, procedure/surgery, systemic and stem cell transplant. Non-AML therapies are for the other hematologic or non-hematologic malignancies.

While the eCRF form is called subsequent AML therapies, the form applies to Prior therapies (therapy started prior to the date of the first dose of study drug) and Subsequent therapies (Follow-up, cancer treatments that the subject started after the end of study drug) to treat the subject's cancer (disease under study and any other cancers).

Prior and subsequent therapies will be summarized separately in tables. All tables described in this section will be based on FAP for Phase 1b subjects, and ITT populations for Phase 2 subjects.

8.4.3.1. Prior/Subsequent Radiation Therapies

The number and percentage of subjects who had any prior/subsequent radiation therapy will be presented by dose cohort for the FAP (Phase 1b) and by treatment arm for the ITT population (Phase 2). Summaries will include number of subjects receiving any radiation therapy, as well as the frequency by type of radiation therapy received (i.e., external beam, radio-immuno therapy, brachytherapy, etc.), intent of therapy, and settings. For subjects received radiation therapy, treatment site, type of radiation therapy, location of external beam, start/stop date, dose, number of fractions, intent of therapy, and settings of therapy will be presented in a listing.

8.4.3.2. Prior/Subsequent Anti-Cancer Procedures

The prior/subsequent anti-cancer procedures will be provided in listings.

8.4.3.3. Subsequent Systemic Anti-Cancer Therapies for AML

The number and percentage of subjects with any subsequent systemic anti-cancer therapies (e.g. chemotherapy, immunotherapy) for AML will be presented by dose cohort (Phase 1b) or treatment arm (Phase 2). Summaries will include number of subjects receiving any prior/subsequent systemic anti-cancer therapy for AML, type of therapy received, intent of therapy, and disease status. A subject may have multiple lines and regimens of systemic anti-cancer treatment for AML. A subject which has multiple intents, types and outcomes will be counted multiple times per category. For subjects with systemic anti-cancer therapies, detailed information will be presented in a listing.

8.4.3.4. Subsequent Stem Cell Transplants for AML

The number and percentage of subjects with any follow-up stem cell transplants for AML will be presented by dose cohort (Phase 1b) or treatment arm (Phase 2). For subjects with stem cell transplants, type and disease status at the time of HSCT will be summarized in tables and the detailed information will be presented in a listing.

8.4.3.5. Prior/Subsequent Systemic Anti-Cancer Therapies Other than AML

The number and percentage of subjects with any subsequent systemic anti-cancer medications for diseases other than AML will be summarized in frequency tabulations by ATC third level and PT, and by dose cohort (Phase 1b) or treatment arm (Phase 2). The type of diseases other than AML will be summarized in separate tables. The detailed information will be presented in a listing.

8.5. Prior/On-treatment/Follow-up Transfusions

Prior transfusions are defined as any type of transfusions received in 8 weeks (56 days) prior to the first dose date. On-treatment transfusions are defined as transfusions that are started on or after the start of study drug but before the last dose date + 28 days. Follow-up transfusions are defined as transfusions that are started on or after the last dose date + 28 days.

Subjects who received prior transfusions are considered transfusion dependent at baseline. The number of subjects receiving any transfusion product, and the number of subjects receiving any transfusion of particular type (RBCs, whole blood, platelets, plasma, other) will be summarized by dose cohort (Phase 1b) or treatment arm (Phase 2). On-treatment transfusions and follow-up transfusions will be summarized separately in tables.

In Phase 2, subjects who achieved post-baseline 56-day RBC transfusion independence, i.e., without RBC transfusion for at least 56 consecutive days post baseline during treatment exposure period, will be summarized by baseline RBC transfusion dependence status (Yes vs. No), within each category of best response (e.g., CR, CRi/CRp, PR/MLFS, SD, PD, etc.) and overall.

The post-baseline 56-day Platelet transfusion independence will be analyzed similarly.

For parameters RBC and Platelet in Phase 2, the number of subjects who are transfusion dependent at baseline, the changes from dependence to post baseline transfusion independence during treatment period will be summarized by treatment arm. The time to transfusion independence and duration of transfusion independence will be summarized by treatment arm as well. In addition, duration of transfusion independence will be summarized by with/without transfusion at baseline. For subjects which have multiple transfusion independent periods, time to the first transfusion independence period and duration of longest transfusion independence period will be summarized.

Transfusion type, number of units, reasons for taking transfusion, and dates of transfusions will be listed in a listing.

9. STUDY TREATMENTS AND EXTENT OF EXPOSURE

All analyses of treatment exposure will be conducted using the safety population, separately for Phase 1b and Phase 2. Exposure to AG-120, AG-221, and AZA will be included in the treatment exposure summaries. Duration of treatment and number of cycles will be calculated for all treatment arm.

Descriptive statistics for duration of treatment exposure, duration of treatment, average prescribed daily dose, average calculated daily dose, average daily dose, cumulative dose, dose intensity and relative dose intensity, will be presented by dose cohort groups (Phase 1b) and by treatment arm (Phase 2). Individual subject listings will be provided to support the tables.

9.1. Duration of Treatment Exposure and Duration of Treatment

Duration of treatment exposure and duration of treatment will be summarized by study drug compound(s) in each dose cohort (Phase 1b) or treatment arm (Phase 2) (e.g. for Phase 1b, dose cohort AG-120 500 mg +AZA, the data will be presented by AG-120, AZA and AG120+AZA).

Duration of treatment exposure (months) is defined as: (treatment end date – treatment start date + 28 days)/30.4. The treatment start date is the date of the first dose of study drug, and the treatment end date is the last dose date. If death date or clinical data cut-off date is after treatment end date and earlier than 28 days after treatment end date, duration of treatment exposure is defined as: (the earlier of death date and data cut-off – treatment start date + 1)/30.4.

Duration of treatment (months) is defined as: (treatment end date – treatment start date + 1)/30.4. The treatment start date is the date of the first dose of study drug, and the treatment end date is the last dose date of study drug. For subjects who are still on treatment at a data cut-off date, treatment end date will be the earlier of data cut-off date and last day that the subjects take study drug collected in eCRF.

See Section 5.1 for cycle length and cycle number calculations. Average cycle length (days), and number of treatment cycles will be summarized by study drug compound(s) in each dose cohort groups (Phase 1b) or treatment arm (Phase 2). For average cycle length, a single average value will be computed for each subject first and then the descriptive statistics will be computed for each dose cohort groups (Phase 1b) and treatment arm (Phase 2). Additionally, the count and percent of subjects by number of cycles of treatment will be provided for each dose cohort groups (Phase 1b) and treatment arm (Phase 2).

9.2. Cumulative Dose

Cumulative dose will be calculated for AG-120, AG-221, and AZA, separately for Phase 1b and Phase 2. Cumulative dose is defined as the sum of all doses taken during the treatment period.

Cumulative dose for AG-120 and AG-221 in mg = Sum of (administered dose in mg during the treatment period); Cumulative dose for AZA in mg= Sum of (administered dose in mg during the treatment period); Cumulative dose for AZA in mg/m^2= Sum of (prescribed dose in mg/m^2 during the treatment period);

The dosage will be counted as zero for days when the study drug is not taken.

9.3. Dose Intensity

Dose intensity during the treatment is defined as the cumulative dose divided by the duration of treatment or modified duration of treatment. Dose intensity will be calculated in the following way:

Dose intensity for AG-120 and AG-221 = [cumulative dose in mg]/[duration of AG120/AG221 treatment (days)]. Duration of AG120/AG221 treatment (days) is defined as: (AG120/AG221 last dose date – AG120/AG221 first dose date + 1).

Dose intensity for AZA = [cumulative dose in mg/m^2]/[modified duration of AZA treatment (days)]. Modified duration of AZA treatment (days) is defined as (start date of AZA last cycle – start date of AZA first cycle +28).

9.4. Relative Dose Intensity

Relative dose intensity is the dose intensity divided by the planned dose intensity.

For Phase 1b, the planned dose intensity for AG-120 and AG-221 will vary by dose cohort.

For Phase 2, the planned dose intensity for AG-221 will be determined in Phase 1b.

For both Phase 1b and Phase 2, the planned dose for AZA is 75 mg/m^2/day for 7 days of each 28-day treatment cycle, so the planned dose intensity is 18.75 mg/m^2/day;

Relative dose intensity for all drugs mentioned above will be categorized into < 75%, 75% to < 90%, 90% to < 120%, and ≥120%, and frequency counts will be provided by dose cohort (Phase 1b) or treatment arm (Phase 2).

9.5. Average Daily Dose and Average Prescribed Daily Dose for AZA

Average daily dose during treatment is defined as the cumulative dose in mg divided by number of days dosed. Average prescribed daily dose for AZA is defined as the cumulative dose in mg/m^2 divided by number of days dosed. A single average value will be computed for each subject, and then the descriptive statistics will be computed for each dose cohort (Phase 1b) or treatment arm (Phase 2).

9.6. Dose Modification

The number of subjects who have at least one dose modification (interruption/reduction), number of dose modifications per subject, and reasons for dose modifications will be summarized by dose cohort (Phase 1b) or treatment arm (Phase 2). Details of dose adjustments will be provided in a by subject listing.

The discontinuation of AG-120, AG-221, or AZA for subjects in the combination arms of the study is allowed. Subjects may continue treatment with single agent AG-120, AG-221, or AZA if in the investigator's assessment the subject continues to show clinical benefit and all protocol-specified criteria for continuing study treatment are met. Number of subjects who discontinue one drug (AG-120, AG-221 or AZA) in combined treatment arms of AG-120 + AZA or AG-221 + AZA will be summarized.

10. EFFICACY ANALYSIS

All efficacy analyses will be conducted separately for Phase 1b and Phase 2. All efficacy analysis will be performed on the FAP population (Phase 1b) and on the ITT population and mITT as well (Phase 2) unless otherwise specified. Key efficacy analyses will also be performed on EAP supportive evidence and to assess the robustness of the efficacy findings for both Phase 1b and Phase 2. Subjects will be analyzed according to either dose cohort (Phase 1b) or randomized treatment arm (Phase 2).

All statistical tests will be two-sided at the significance level of $\alpha = 0.05$ unless otherwise specified, and the corresponding p-values and two-sided 95% confidence interval (CI) for intended point estimates will be reported, unless specified otherwise.

10.1. Analysis of Primary Efficacy Endpoint

Phase 1b does not have a primary efficacy endpoint.

For Phase 2, the primary efficacy endpoint of ORR is defined as rate of CR + CRi+ CRp + MLFS + PR according to modified IWG AML response criteria as assessed by investigator. The analysis of the primary endpoint ORR will be based on the ITT population. Subjects will be analyzed as they were randomized, regardless of the actual treatment received. The ORR will be summarized by treatment arm. The treatment difference in ORR will be tested using the Chi-squared test or Fisher exact test if appropriate at a two-sided 0.2 level and in the ITT population. This test will provide the pivotal p-value for the comparison of the ORR of the oral AG-221 + SC azacitidine versus azacitidine mono therapy. Estimated response rate with 80% and 95% CIs will also be provided.

Response will also be summarized by the best overall response following the hierarchical order of CR, CRi/CRp, PR, MLFS, stable disease (SD), progressive disease (PD)/relapse, and not evaluable (NE). Cytogenetic Complete Remission (CRc) will be counted as CR thus will not be summarized separately. The best response of CRi and CRp are of the same rank and thus will be reported as a single category. The best response of PD, and relapse, will be grouped together. Sensitivity analyses will be performed on the mITT and EAP.

10.2. Analyses of Secondary Efficacy Endpoints

For Phase 1b, ORR and Sponsor Derived CR/CRh (based on laboratory dataset) are secondary efficacy endpoints and will be analyzed based on the FAP population and on the EAP as sensitivity analysis. These endpoints will be presented descriptively with 95% and 90% CI. Other measures of clinical activity, including duration of response, duration of CR/CRh, duration of remission, time to response, time to remission, time to CR/CRh, event-free survival (EFS), and overall survival may be summarized as appropriate in the FAP population.

For Phase 2, time-to-event secondary efficacy variables are displayed as part of Table 2 (EFS, duration of response, OS, duration of remission (CR), and duration of sponsor assessed CR and CRh) will be analyzed by the KM method if appropriate to estimate time-to-event curves. In addition, Cox proportional hazards regression models will be generated to obtain hazard ratios and their two-sided 95% CI, and the p-value will be based on 2-sided log-rank test as appropriate. Time to response, time to remission, and time to sponsored assessed CR and CRh will be summarized

descriptively. Counts and percentages will be used to describe categorical variables. Rates, such as Complete remission rate, Hematologic improvement rate, sponsor derived CR and CR + CRh, will be analyzed similarly to ORR. All secondary efficacy endpoints for Phase 2 will be analyzed based on ITT population and on the mITT as supportive evidence and to assess the robustness of the efficacy findings. Subjects will be analyzed according to randomized treatment arm.

10.2.1. Event-free survival

EFS is assessed by investigator based on modified International Working Group (IWG) AML response criteria (Section 17.4).

Event-free survival (EFS) is defined as the interval from either the date of the first dose for subjects in Phase 1b or the date of randomization for subjects in Phase 2 to the date of documented morphologic relapse, progression, or death from any cause, whichever occurs first. Censoring rules are provided in Table 4. The analysis of the EFS for Phase 2 will be conducted using the log-rank test in the ITT population, and mITT population. The Cox proportional hazards regression model will be used to estimate the hazard ratio and the corresponding 95% confidence interval for oral AG-221 with SC azacitidine relative to SC azacitidine alone.

Additionally, the number of events, the number of censored and reasons for censoring will be summarized (n, percent) for this endpoint. KM curves will be provided by treatment arm in figures.

Table 4: Censoring Rules for EFS

Value of Event-free Survival Date (ADT)	Censored (Y, N)	Derivation
ADT = min(death date, MR date, progression date)	N	If a subject died, had MR, or had disease progression by investigator review, and the subject has less than two consecutive missing response assessments prior to a visit documented death, MR, or progression which is before new AML therapy
ADT = the last event-free assessment date before missing assessments/ randomization date	Y	If a subject died, had MR, or had disease progression by investigator review, and the subject has two or more consecutive missing response assessments prior to a visit documented death, MR, or progression
ADT = the last adequate assessment date/randomization date	Y	If a subject did not die, did not have MR, or did not have progression, and did not receive any subsequent anticancer therapy. If there was no post-baseline response assessment, then ADT = the randomization date.
ADT = the last adequate assessment date on or prior to start of subsequent anticancer therapy/randomization date	Y	If a subject did not die, did not have MR, or did not have progression on or prior to start of subsequent anticancer therapy. If there was no post-baseline response assessment, then ADT = the randomization date.
ADT = the last adequate assessment date on or prior to start of subsequent anticancer therapy/randomization date	Y	If a subject died, had MR, or had disease progression by investigator review after receiving subsequent anticancer therapy. If there was no post-baseline response assessment on or prior to start of subsequent anticancer therapy, then ADT = the randomization date.

ADT = analysis date; N = no; Y = yes, MR = morphologic relapse.
Note: Event-free response refers to a response that was neither morphologic relapse (MR), progressive disease (PD) nor death.

Sensitivity analyses for EFS will be done with the following censoring rules:

- Subjects with two or more consecutive missing response assessments prior to an event date (death, MR or disease progression) will be considered as events at the event date by ignoring the missing assessments; (EFS1)

- Subjects with subsequent AML therapy prior to an event date will be considered as events at the event date; (EFS2)
- Analysis of EFS with subjects censored at the time of transplantation after study treatments; (EFS3)

Other sensitivity rules may be added when data is available.

10.2.2. Duration of Response

Duration of response is defined as time from first documented CR/CRi/CRp/PR/MLFS to documented morphologic relapse, PD, or death due to any cause, whichever occurs first. Subjects without morphologic relapse, PD, or death due to any cause will be censored at the date of the last response assessment.

The duration of response will be analyzed using the KM method. The 25th percentile, median and 75th percentile time (including two-sided 95% CI) will be summarized for each treatment arm. Only subjects with documented CR/CRi/CRp/PR/MLFS response as determined by investigator assessment will be included in the analysis.

10.2.3. Duration of Remission

Duration of remission is defined as time from first documented CR to documented morphologic relapse, PD, or death due to any cause, whichever occurs first. Subjects without morphologic relapse, PD, or death due to any cause will be censored at the date of the last response assessment.

The duration of remission will be analyzed using the KM method. The 25th percentile, median and 75th percentile time (including two-sided 95% CI) will be summarized for each treatment arm. Only subjects with documented CR as determined by investigator assessment will be included in the analysis.

10.2.4. Overall Survival

Overall survival (OS) is defined as time from first date of study drug (Phase 1b) or randomization (Phase 2) to death due to any cause. This time to event endpoint will be analyzed, similarly to the EFS analyses described in 10.2.1. Sensitivity analysis for OS may be explored if needed.

10.2.5. One-Year Survival

One-year survival is defined as the probability of survival at 1 year from randomization.

One-year survival will be evaluated by KM estimate with 95% CI by treatment arm. Difference between treatment arms and corresponding two-sided 95% CI will also be displayed based on log-rank test. The CI for the difference in the 1-year survival probabilities will be derived using Greenwood's variance estimate.

10.2.6. Complete Remission Rate

Complete remission rate is defined as rate of CR according to modified IWG AML response criteria.

Number and percentage of subjects with complete remission will be summarized by treatment arm. Comparisons between two treatment arms will be analyzed by Fisher's exact test if chi-square test is not appropriate.

10.2.7. Hematologic Improvement Rate

Hematologic improvement rate is defined as rate of neutrophil response (hematologic improvement neutrophil response [HI-N]) + platelet response (hematologic improvement platelet response [HI-P]) + erythroid response (hematologic improvement erythroid response [HI-E]) according to IWG MDS HI criteria (Section 17.5).

Number and percentage of subjects with hematologic improvement will be summarized by treatment arm for Phase 2. Comparisons between two treatment arms will be analyzed by Fisher's exact test.

10.2.8. Time to Response

Time to response is defined as time from the first dose of study drug for Phase 1b subjects and the date of randomization for Phase 2 subjects to the first documented CR/CRi/CRp/PR/MLFS according to modified IWG AML response criteria.

Time to response will be analyzed using the summary statistics for each treatment arm. Only subjects with documented CR/CRi/CRp/PR/MLFS response as determined by investigator assessment will be included in the analysis.

10.2.9. Time to Remission

Time to remission is defined as time from the first dose of study drug for Phase 1b subjects and the date of randomization for Phase 2 subjects to the first documented CR according to modified IWG AML response criteria.

Time to remission will be analyzed using the summary statistics for each treatment arm. Only subjects with documented CR as determined by investigator assessment will be included in the analysis.

10.2.10. Sponsor Derived CR

Sponsor derived CR rate is defined as rate of CR based on laboratory data.

Number and percentage of subjects with sponsor derived complete remission will be summarized by treatment arm. Comparisons between two treatment arms will be analyzed by Fisher's exact test.

10.2.11. Sponsor Derived CR and CRh

Sponsor derived CR and CRh rate is defined as rate of CR and CRh based on laboratory data. CRh is defined as less than 5% blasts in a BM aspirate sample with marrow spicules plus ANC > 0.5 x 10^9/L (1,000/μL) & Platelet count > 50 x 10^9/L (100,000/μL).

Number and percentage of subjects with sponsor derived CR and CRh will be summarized by treatment arm. Comparisons between two treatment arms will be analyzed by Fisher's exact test.

10.2.12. Duration of Sponsor Assessed CR and CRh

Duration of sponsor assessed CR and CRh is defined as time from first documented CR/CRh to documented morphologic relapse, PD, or death due to any cause, whichever occurs first. Subjects without morphologic relapse, PD, or death due to any cause will be censored at the date of the last response assessment.

The duration of response will be analyzed using the KM method. The 25th percentile, median and 75th percentile time (including two-sided 95% CI) will be summarized for each treatment arm. Only subjects with documented CR/CRh response as determined by sponsor assessment will be included in the analysis.

10.2.13. Time to Sponsor Assessed CR and CRh

Time to sponsor assessed CR and CRh is defined as time from first dose of study drug for Phase 1b subjects and the date of randomization for Phase 2 subjects to first documented CR/CRh.

Time to sponsor assessed CR and CRh will be analyzed using the summary statistics for each treatment arm. Only subjects with documented CR/CRh response as determined by sponsor assessment will be included in the analysis.

10.2.14. EORTC QLQ-C30 and EQ-5D-5L

For Phase 2, the analyses will address the mean differences by treatment arm on the EORTC QLQ-C30 scale and subscale scores and the treatment arm differences in the proportion of subjects who achieve minimal clinically important differences. The scoring for the EORTC QLQ-C30 and methods to address missing values will be accomplished according to directions provided by the instrument developer. The analyses of the EORTC QLQ-C30 will be based on all randomized subjects who completed the baseline assessment and had at least one post-baseline assessment with the EORTC QLQ-C30. The observed case method and the descriptive statistics will be used to summarize the observed scores and the change from baseline score by visit and the treatment arm for each domain of the QLQ-C30. In addition, frequency tables for categorized change from baseline scores (Improving, No Change, and Worsening) will be presented by treatment arm. A change of at least 10 points on the standardized domain scores will be required for it to be considered meaningful (Osoba, 1998).

For Phase 2, EQ-5D-5L will be scored according to the instrument guidance and analyzed accordingly. The EQ-5D-5L consists of 2 pages – the EQ-5D-5L descriptive system and the EQ Visual Analogue scale (EQ VAS). The descriptive system comprises of 5 dimensions: 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. Responses from the 5 dimensions are coded so that a '1' indicates no problem on that dimension, and '5' indicates the most serious problem. The responses for the five dimensions can be combined in a five-digit number describing the respondent's health state. For instance, a state of 11111 indicates no problem with any of the 5 dimensions, but a state of 55555 indicates the most difficulty on all 5. Missing and un-interpretable values (e.g. two responses are given for a single dimension) which are not recoverable through queries will be coded as '9'. These health states will be converted to a single index value using the crosswalk link function based on the existing value sets derived based on United Kingdom time trade-off (TTO) valuation techniques.

EQ VAS is a respondent's self-rated today's health scale which is recorded on a 20 cm vertical, visual analogue scale with endpoints labeled "the best health you can imagine" and "the worst health you can imagine". The scale is numbered from 0 to 100 with 0 corresponding to the worst imaginable health state and 100 corresponding to the best imaginable health state.

The analyses of the EQ-5D-5L will be based on all randomized subjects who completed the baseline assessment and had at least one post-baseline assessment with the EQ-5D-5L. The observed case method and the describable statistics will be used to summary the EQ VAS and index values and the change from baseline by visit and the treatment arm.

[redacted]

10.3. Subgroup Analysis

Appropriate subgroup analyses will be conducted for primary and selected secondary efficacy endpoints by the following subgroups for Phase 2:

1. Age groups (< 65 years, ≥ 65 to < 75 years, ≥ 75 years);
2. Gender (male and female);
3. Region (US, Non-US);
4. Race (white vs non-white);
5. Baseline cytogenetic risk status (better-risk, intermediate-risk, poor-risk, failure);
6. For IDH2 mutated subjects, IDH2 gene mutation sub-type (R140Q vs. R172K);
7. AML initial diagnosis: (primary or secondary);
8. WHO classification of AML;
9. ECOG performance status at baseline (Grade 0, 1, 2);
10. European LeukemiaNet (ELN) risk classification (favorable, intermediate and adverse).

A forest plot will be provided displaying the hazard ratios (for time to event endpoints) and odds ratios (for ORR) and their 95% CIs, obtained from Cox proportional hazards and logistic regression models, respectively, for each subgroup.

Subgroup analyses will be conducted for the following efficacy endpoints:

- Overall Response Rate;
- Overall Survival.

11. SAFETY ANALYSIS

All analyses of safety data will be conducted using the safety population unless otherwise specified, separately for Phase 1b and Phase 2. Safety summaries will be provided by dose cohort (Phase 1b), and treatment arm (Phase 2), and overall (Phase 1b and Phase 2 separately).

For summaries of safety data, the treatment period is defined as follows:

date of first dose through date of last dose + 28 days.

No inferential statistical analyses will be conducted on the safety data.

11.1. Adverse Events

Adverse events (AEs) will be analyzed in terms of treatment-emergent adverse events (TEAEs) which are defined as any AEs that begin on or after the start of study drug through 28 days after the last study treatment. All AEs will be coded using the Medical Dictionary for Regulatory Affairs® (MedDRA) Version 22.0. The severity will be graded by the study personnel based on National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03.

A treatment-related TEAE is defined as TEAE that is assessed by the Investigator to be possibly related to AG120/AG221 or AZA.

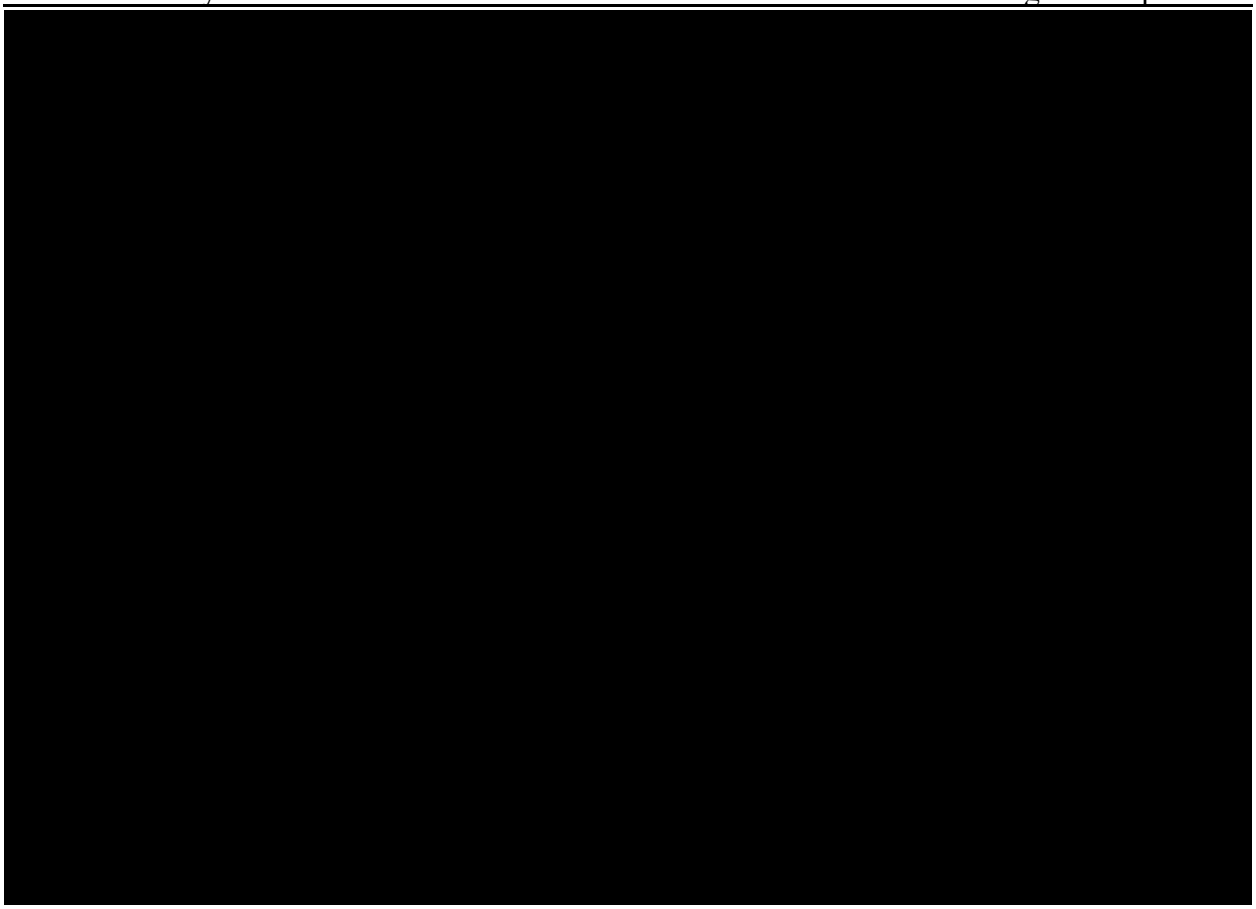
If a subject experiences the same AE more than once with different toxicity grade, then the event with the highest grade will be tabulated in "by grade" tables. If a subject experiences multiple AEs under the same PT (SOC), then the subject will be counted only once for that PT (SOC).

The number of subjects with at least one TEAE will be summarized. The incidence of TEAEs will be summarized by MedDRA SOC and PT. Tables summarizing the incidence of TEAEs will be generated for each of the following:

- Summary of TEAEs
- All TEAEs;
- TEAEs Exposure Adjusted Incidence Rate (EAIR), only for Phase 2;
- TEAES by Cycle, only for Phase 2;
- TEAE by decreasing frequency of PT;
- TEAEs related to study drug;
- TEAEs by maximum severity;
- TEAEs with CTCAE grade 3 or;
- Related TEAEs with CTCAE grade 3 or 4;
- Serious TEAEs;
- Serious TEAEs related to study drug;
- Serious TEAEs EAIR, only for Phase 2;
- Serious TEAE by decreasing frequency of PT, only for Phase 2;

- TEAEs leading to study drug withdrawn;
- TEAEs related to study drug and leading to study drug withdraw;
- TEAEs with leading to dose reduction;
- TEAEs related to study drug and leading to study drug dose reduction;
- TEAEs leading to study drug dose interruption;
- TEAEs related to study drug and leading to study drug dose interruption;
- All TEAEs resulting in death;
- Related TEAEs resulting in death;
- TEAEs leading to death EAIR, only for Phase 2;
- TEAEs for the following baseline subgroups (provided when the number of subjects are sufficient), only for Phase 2:
  - Gender (male and female);
  - Age (< 65 years, ≥ 65);
  - Baseline ECOG performance status (Grade 0, 1, 2);
  - Baseline ANC value ($\leq 0.5 \times 10^9$/L, $0.5 < 1 \times 10^9$/L, $\geq 1 \times 10^9$/L)
  - Baseline creatinine clearance value (<45, 45-<60, ≥60 mL/min);

Listings for the corresponding summary tables will be presented separately. Non-treatment-emergent AEs will also be included in one of the listings.



11.3. Death

System organ classes (SOCs) and preferred terms (PTs) are coded using the MedDRA dictionary (Version 22.0). Cause of death will be coded as well using the same MedDRA dictionary. Cause of death will be summarized in the Safety population with the following categories:

- On-treatment death: Deaths within 28 days of the last dose of study drug.
- Post-treatment death: Deaths on or after 28 days of the last dose of study drug.
- Overall: All death

On-treatment deaths, post-treatment deaths and all deaths will also be summarized by SOC and PT. A listing will be provided including deaths for all screened subjects.

11.4. Clinical Laboratory Evaluations

Clinical laboratory results will be summarized descriptively by dose cohort (Phase 1b) or treatment arm (Phase 2), which will also include a display of change from baseline. Laboratory values outside of the normal ranges will be identified. Normal ranges will be used to determine the categories of High, Low, and Normal for laboratory tests that have no severity grade.

Clinically significant hematologic and non-hematologic laboratory abnormalities that meet Grade 3 or Grade 4 criteria according to the NCI CTCAE 4.03 will be summarized. The worst grade during the treatment period will be summarized by dose cohort (Phase 1b) and treatment

arm (Phase 2). Frequency distributions of CTC grade that shift from baseline to the worst grade during treatment period will be presented by treatment arm when data is applicable. Graphical display of select laboratory parameters over the course of the study will be provided where useful to assist in the interpretation of results.

Listings of all clinical laboratory data from central laboratory with abnormal flags will be provided by subjects and tests. Listings will also be provided for the local laboratory data.

11.4.1. Hematology

For hematologic parameters, the change from baseline will be summarized by study visit and by treatment arm for Phase 2.

Shifts in CTC grades from baseline to worst grade, separated by worst (high) and worst (low) assessments, will be summarized by dose cohort (Phase 1b) and treatment arm (Phase 2) for hemoglobin, platelet counts, ANC, WBC, and RBC.

In addition, figures for the mean of hemoglobin, platelet counts, ANC, WBC and RBC by visit will be provided by treatment arm for Phase 2.

11.4.2. Serum Chemistry

For serum chemistry parameters, the change from baseline will be summarized by study visit and by treatment arm for Phase 2.

Shifts in CTC grades from baseline to worst grade will be summarized by dose cohort (Phase 1b) and treatment arm (Phase 2).

The proportion of subjects with clinically significant post-baseline change in hepatic, renal function and electrolytes will be assessed based on the criteria presented in the Table 5 .

Table 5 Serum Chemistry Changes of Interest

Chemistry Laboratory Test	Parameters for any value and the last value
ALT	ALT ≥× 3ULN (presented as ≥×3- < × 5 ULN, ≥×5- < ×8 ULN and ≥×8ULN)
AST	AST≥×3ULN (presented as ≥×3- < × 5 ULN, ≥5- < ×8 ULN and ≥×8ULN)
Bilirubin Total	Bilirubin Total ≥ × 2ULN Bilirubin Total ≥ × 3ULN
Total Bilirubin and ALT	Composite of Total Bilirubin ≥ ×2ULN and ALT≥ × 3ULN (ALT presented as ≥×3 - < × 5 ULN, ≥5 < × 8 ULN and ≥×8 ULN) concurrent and within 1 cycle of Total Bilirubin elevation and at any time after start of treatment
Total Bilirubin and AST	Composite of Total Bilirubin ≥ × 2ULN and AST ≥× 3ULN, (AST presented as ≥3 < × 5 ULN, ≥5 < × 8 ULN and ≥8 ULN) concurrent and within 1 cycle of Total Bilirubin elevation and at any time after start of treatment
Potassium	Hypokalemia < 3.0 mmol/L and >15% decrease from BL Hyperkalemia > 6.0 mmol/L and >15% increase from BL
Potassium and Uric acid	Composite of concurrent hyperkalemia (potassium >ULN and >25% increase from BL) and hyperuricemia (uric acid >ULN and >25% increase from BL)
Phosphorus and Uric acid	Composite of concurrent hyperphosphatemia (phosphorus >ULN and >25% increase from BL) and hyperuricemia (uric acid >ULN and >25% increase from BL)
Calcium and Uric acid	Composite of concurrent hypocalcemia (calcium <ULN and >25% decrease from BL) and hyperuricemia (Uric acid >ULN and >25% increase from BL)
Magnesium	Hypomagnesemia Grade ≥2 and >25% decrease from BL

11.4.3. Urinalysis

The results will only be displayed in a listing.

11.4.4. Coagulation

The results will only be displayed in a listing.

11.4.5. Fasting Lipid Panel

Fasting lipid panel includes total cholesterol, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides. Lipid profiles will be described by concentration of the lipids at scheduled visits and change from baseline by treatment arm for Phase 2. The proportion of subjects with clinically meaningful post-baseline increase in lipids will be summarized based on criteria presented in Table 6 for both Phase 1b and Phase 2.

Table 6: Fasting Lipid Changes of Interest

Laboratory Test	Parameters for any value and the last value
LDL cholesterol	>ULN and >10% increase from BL
LDL-C/HDL-C ratio	>3.5 and >10% increase from BL

11.4.6. Cardiac Markers

Cardiac markers include troponin T. The results will only be displayed in a listing.

11.4.7. Females of Childbearing Potential

The results will only be displayed in a listing.

11.5. Vital Sign Measurements

Evaluation of vital signs includes recording height (only at screening), weight, temperature, systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, and respiratory rate at scheduled visits.

Vital sign measurements will be listed for each subject and by scheduled visit. Descriptive statistics for vital signs (except height), both observed values and changes from baseline, will be summarized by treatment arm for Phase 2.

In addition, a shift table will show the number of subjects with high, normal, low values by treatment arm and by parameter at each visit. Normal ranges are defined as below (anything lower than the low limit is a low value, anything higher than the upper limit is a high value):

- SBP 100-140 mmHg
- DBP 60-90 mmHg
- Pulse 60-100 beats per minute (bpm)
- Respiration 12-20 bpm

11.6. Electrocardiograms

Electrocardiogram (ECG) parameters include heart rate, PR interval, QRS duration, RR interval, QT, QTc Fridericia value (QTcF) and QTc Bazett value (QTcB). QTcB is not recorded in the eCRF page and will be calculated in the following way: QTcB=QT (ms)/sqrt (RR (ms)/1000).

Recorded values and of ECG parameters and change from baseline values will be summarized at each time point by treatment arm for Phase 2 only.

In addition, at baseline and maximum post-baseline visits, the proportion of subjects having absolute QTcF and QTcB intervals of the following categories will be presented:

- ≤ 480 ms
- > 480 to ≤ 500 ms
- > 500 ms

At maximum post-baseline visits, the proportion of subjects who have an increase from baseline in QTcF and QTcB intervals of the following categories will be presented:

- ≤ 30 ms
- > 30 to ≤ 60 ms
- > 60 ms

The overall ECG interpretation will be summarized by presenting the shift from baseline to worst by treatment arm for all cycles. The overall ECG interpretation will be displayed in cross–tabulations for each treatment.

11.7. Echocardiogram/Multi-gated Acquisition

Data of echocardiogram (ECHO) and multi-gated acquisition (MUGA) measurements will be collected on the eCRF as part of the safety assessment. Data listing will be provided.

11.8. ECOG Performance Status

Shift tables will be provided for ECOG PS from baseline to worst values across all visits.

A by-subject listing will be presented for each phase of the study.

12. PHARMACOKINETIC ANALYSIS

The PK population will be used for all PK analyses.

12.1. Values below the Limit of Quantification or Missing

Pre-dose concentrations that are below the limit of quantification (BLQ) will be assigned a numerical value of zero for PK analysis. Post-dose concentrations that are BLQ and occur before the first quantifiable concentration will also be treated as zero for PK analysis. Post-dose concentrations that are BLQ but occur between two quantifiable concentrations will be treated as missing for PK analysis.

Concentrations assigned a value of missing will be omitted from the calculation of descriptive statistics. For concentration values that are BLQ, a concentration value of zero will be included for the computation of arithmetic mean, and a postdose zero concentration will be substituted with a 50% value of the lower limit of quantification, for the computation of geometric mean. If 50% or more of the values are BLQ at one time point, the mean, median, and geometric mean will be reported as BLQ, the maximum will be the greatest measurable value (or BLQ if no measurable values), and other descriptive statistics will not be calculated.

12.2. Plasma Sample Collection

Blood samples for intensive PK assessments of AG-221 in plasma will be collected at the following time points: predose (0 hour) and 2, 3, 4, 6, 8 and 24 hours post-dose in the first twelve subjects on C2D1 in each of treatment arms in Phase 2.

Blood samples for intensive PK assessments of AG-120 in plasma will be collected at the following time points: pre-dose and 0.5, 2, 3, 4 hours, 6, and 8 hours post-dose on Days 1 of Cycle 1 and 2 in Phase 1b.

Plasma concentrations of AG-221 or AG-120 will be summarized by nominal time points, including mean, standard deviation, % coefficient of variation (CV%), geometric mean, geometric CV%, minimum, median, and maximum. Mean (±standard deviation) plot of plasma concentrations will be presented in both linear scale and semi-logarithmic scale. Figure will be provided.

The subjects not being assessed for intensive PK in the AG-221 treatment arm during the Phase 2 segments will undergo sparse PK samples for measurement of AG-221 concentration. Listing of all sparse PK samples will be provided.

12.3. Pharmacokinetic Parameters

PK parameters will be determined using non-compartmental (NCA) methods, based on individual plasma concentration-time data for AG-120 from up to 15 subjects with intensive PK assessments on C1D1 and C2D1 in AG-120 treatment arms in Phase 1b.

PK parameters will be determined using non-compartmental (NCA) methods, based on individual plasma concentration-time data for AG-221 from ten to twelve subjects with intensive PK assessments on C2D1 in AG-221 treatment arms in Phase 2. In addition, Listings with sparse PK samples will be provided.

The following PK parameters will be calculated for AG-221 or AG-120 using concentration time data on C2D1 or C1D1/C2D1 from subjects with intensive PK assessments in Phase 2 or Phase 1b:

- AUC_{0-8}: Area under the plasma concentration-time curve from time zero to 8 hours, calculated using the linear trapezoid rule.
- AUC_{0-24} (for AG-221 only): Area under the plasma concentration-time curve from time zero to 24 hours, calculated using the linear trapezoid rule.
- C_{max}: Maximum observed plasma concentration, obtained directly from the observed concentration versus time data.
- T_{max}: Time of maximum observed plasma concentration, obtained directly from the observed concentration versus time data.

Additional PK parameters may be determined when appropriate.

All plasma PK parameter calculations will be performed using actual time points. The actual dose of AG-221 or AG-120 will be used for PK parameter calculation. If a dose of AG-221 or AG-120 is missed or is reduced/increased, the complete corresponding concentration-time profile may be excluded from the PK analysis and/or summary statistics after review by Celgene on a case-by-case basis.

PK parameters will be summarized descriptively (mean, standard deviation, CV%, geometric mean, geometric CV%, min, median, and max). Table will be provided.

13. QUALITY OF LIFE ANALYSIS

Two instruments, the European Organization for Research and Treatment of Cancer Quality-of-Life questionnaire (EORTC QLQ-C30) and EuroQoL Group EQ 5D-5L, will be used for evaluating HRQoL outcomes. Description of the analysis is provided in Section 10.2.14.

[redacted]

Data Monitoring Committee (DMC)

An external DMC will be convened. The DMC includes a group of physicians with experience in treating subjects with AML and a statistician, all of whom are not otherwise involved in the study conduct. During the course of the study, the DMC members will periodically review the safety data to ensure the safety of the study subjects. The detailed roles and responsibilities of DMC can be found in the DMC charter.

17. APPENDICES

17.1. Handling of Dates

Dates will be stored as numeric variables in the SAS analysis files and reported in DDMMMYYYY format (i.e., the Date9. datetime format in SAS). Dates in the clinical database are classified into the categories of procedure dates, log dates, milestone dates, outcome dates, and special dates.

- **Procedure Dates** are the dates on which given protocol-specified procedure are performed. They include the dates of laboratory testing, physical examinations, etc. They should be present whenever data for a protocol-specified procedure are present and should only be missing when a procedure are marked as NOT DONE in the database. Procedure dates will not be imputed.
- **Log Dates** are dates recorded in eCRF data logs. Specifically, they are the start and end dates for AE and concomitant medications/procedures. They should not be missing unless an event or medication is marked as *ongoing* in the database. Otherwise, incomplete log dates will be imputed according to the rules in Section 17.2 (e.g., for duration or cycle assignment, etc.). However, in listings, log dates will be shown as recorded without imputation.
- **Milestone Dates** are dates of protocol milestones such as randomization, study drug start date, study drug termination date, study closure date, etc. They should not be missing if the milestone occurs for a subject. They will not be imputed.
- **Outcome Dates** are dates corresponding to study endpoints such as survival, progression, etc. In most cases they are derived either from a milestone (e.g., the survival date is derived from the death date), or a procedure date (e.g., the progression date is derived from the date of the hematology lab that was used to determine progression). They may be subject to endpoint-specific censoring rules if the outcome did not occur but are not otherwise subject to imputation.
- **Special Dates** cannot be classified in any of the above categories and they include the date of birth. They may be subject to variable-specific censoring and imputation rules.

Dates recorded in comment fields will not be imputed or reported in any specific format.

17.1.1. Calculation Using Dates

Calculations using dates (*e.g.*, subject's age or relative day after the first dose of study drug) will adhere to the following conventions:

- Study days after the start day of study drug will be calculated as the difference between the date of interest and the first date of dosing of study drug (e.g., lenalidomide) plus 1 day. The generalized calculation algorithm for relative day is the following:
  - If TARGET DATE ≥ DSTART then STUDY DAY = (TARGET DATE – DSTART) + 1;
  - Else use STUDY DAY = TARGET DATE – DSTART.

Note that Study Day 1 is the first day of treatment of study drug. Negative study days are reflective of observations obtained during the baseline/screening period. Note: Partial dates for the first study drug are not imputed in general. All effort should be made to avoid incomplete study drug start dates. Some analyses may exclude subjects without study drug start date. The SAP may be amended on a case-by-case basis, including guideline and suggestion such as imputing Day 1 by consent date of treated subject.

- Age (expressed in days) is calculated: AGE = CONSENT – DATE of BIRTH + 1. In practice, age will be transformed to years by dividing the difference by 365.25 days, then truncating.
  - Preference is for using calculated age from clinical database. When not available, calculated age from eCRF or IWRS may be used
  - Partial birth date: impute missing day as 15th of the month; impute missing month as July; set missing age for missing year
- Intervals that are presented in weeks will be transformed from days to weeks by using (without truncation) the following conversion formula:

  WEEKS = DAYS /7
- Intervals that are presented in months will be transformed from days to months by using (without truncation) the following conversion formula:

  MONTHS = DAYS /30.4

17.2. Date Imputation Guideline

17.2.1. Impute Missing Date of Adverse Events/ Prior or Concomitant Medications

Partially missing AE start dates will be imputed in the Analysis Data Model (ADaM) dataset for AEs, but partially missing AE end dates will not be imputed in the same dataset. If the AE end date is complete with no missing year, month, or day, and the partially missing start date imputed by the rules below is after the AE end date, then the start date will be imputed by the AE end date.

Partially missing start/stop dates for prior/concomitant medications and partially missing start dates for prior/concomitant procedures will be imputed in the ADaM dataset for prior/concomitant medications/procedures. For prior/concomitant medications, if the stop date is complete with no missing year, month, or day, and the partially missing start date imputed by the rules below is after the stop date, then the start date will be imputed by the stop date.

Incomplete Start Date: If the stop date is not missing, and the imputed start date is after the stop date, the start date will be imputed by the stop date.

Missing day and month

- If the year is the **same** as the year of the first dosing date, then the day and month of the first dosing date will be assigned to the missing fields.
- If the year is **prior to** the year of first dosing date, then December 31 will be assigned to the missing fields.

- If the year is **after** the year of first dosing, then January 1 will be assigned to the missing fields.

Missing day only

- If the month and year are the **same** as the year and month of first dosing date, then the first doing date will be assigned to the missing day.
- If either the year of the partial date is **before** the year of the first dosing date or the years of the partial date and the first dosing date are the same but the month of partial date is **before** the month of the first dosing date, then the last day of the month will be assigned to the missing day.
- If either the year of the partial date is **after** the year of the first dosing date or the years of the partial date and the first dose date are the same but the month of partial date is **after** the month of the first dosing date, then the first day of the month will be assigned to the missing day.
- If the stop date is not missing, and the imputed start date is after the stop date, the start date will be imputed by the stop date.

Missing day, month, and year

- No imputation is needed, the corresponding AE will be included as TEAE if end date of AE is after the first dose date or the end date is also missing.

Incomplete Stop Date: If the imputed stop date is before the start date, then the imputed stop date will be equal to the start date.

Missing day and month

- If the year of the incomplete stop date is the **same** as the year of the last dosing date, then the day and month of the last dosing date will be assigned to the missing fields.
- If the year of the incomplete stop date is **prior to** the year of the last dosing date or prior to the year of the first dosing date, then December 31 will be assigned to the missing fields.
- If the year of the incomplete stop date is **prior to** the year of the last dosing date but is the same as the year of the first dosing date, then the first dosing date will be assigned to the missing date.
- If the year of the incomplete stop date is **after** the year of the last dosing date, then January 1 will be assigned to the missing fields.

Missing day only

- If the month and year of the incomplete stop date are the **same** as the month and year of the last dosing date, then the day of the last dosing date will be assigned to the missing day.
- If either the year of the partial date is **not equal to** the year of the last dosing date or the years of the partial date and the last dosing date are the same but the month of partial

date is **not equal to** the month of the last dosing date, then the last day of the month will be assigned to the missing day.

17.2.2. Medical History

Partially missing medical history start dates will be imputed in the ADaM dataset for medical history. The 16th of the month will be used to impute a partially missing start date that has only the day missing, and July 1st will be used to impute a partially missing start date that has both the month and day missing.

17.3. Table of Events

Events Table covers both Phase 1b & 2 (unless otherwise noted)	Screening Period	Treatment Period[a]								Follow-up Period[a]
	Screening	28-day cycles							EOT [b]	Follow-up
		Cycles 1 to 2				Cycles 3 to 4		Cycle 5 and beyond		
	Days -28 to -1	Day 1	Day 8	Day 15	Day 22	Day 1	Day 15	Day 1		
Informed Consent	× [c]	-	-	-	-	-	-	-	-	-
IWRS Subject Status Registration (Phase 1b/2)	×	× [d]	-	-	-	×	-	×	×	-
Inclusion & Exclusion Criteria	×	× [e, f]	-	-	-	-	-	-	-	-
Demographics	×	-	-	-	-	-	-	-	-	-
AML Diagnosis	× [g]	-	-	-	-	-	-	-	-	-
Local Testing of IDH1/ IDH2 Gene Mutations on BMA and PB (In addition, BMA and PB samples to be sent to central lab)	× [g]	-	-	-	-	-	-	-	-	-
Medical History	×	-	-	-	-	-	-	-	-	-
Prior Medications and Procedures	× [h]	-	-	-	-	-	-	-	-	-
EORTC QLQ-C30 and EQ-ED-5L	-	×	-	-	-	×	-	×	×	-
ECOG Performance Status	×	×	-	-	-	×	-	×	×	-
Physical Examination [i]	×	×	-	-	-	×	-	×	×	-
Vital Signs	×	×	×	×	×	×	×	×	×	-

Events Table covers both Phase 1b & 2 (unless otherwise noted)	Screening Period	Treatment Period[a]								Follow-up Period[a]
	Screening	28-day cycles							EOT [b]	Follow-up
		Cycles 1 to 2				Cycles 3 to 4		Cycle 5 and beyond		
	Days -28 to -1	Day 1	Day 8	Day 15	Day 22	Day 1	Day 15	Day 1		
ECHO / MUGA Scan [k]	× [l]	As clinically indicated								-
Height	×	-	-	-	-	-	-	-	-	-
Body Weight	×	×	-	-	-	×	-	×	×	-
BSA Calculation [j]	×	×	-	-	-	×	-	×	-	-
12-Lead ECG [k]	×	×	-	×	-	×	-	×	×	-
Pregnancy Test (FCBP only) [m]	× [n]	× [k, o]	-	-	-	× [k]	-	× [k]	× [k]	-
Urinalysis [n]	×	As clinically indicated								-
Coagulation Laboratory [n]	×									-
Hematology Laboratory [n]	×	×	×	×	×	×	×	×	×	× [p]
Chemistry Laboratory [n]	×	×	-	×	-	×	-	×	×	-
UGT1A1 Gene Mutation Test (for diagnosis of Gilbert's Syndrome; refer to Section 4.2 in Protocol) [n]	×	-	-	-	-	-	-	-	-	-
Cardiac Markers [n]	-	× [f]	× [q]							-
Fasting Lipid Panel [n]	-	× [f]	× [r]							-
Adverse Event	× [s]									

Events Table covers both Phase 1b & 2 (unless otherwise noted)	Screening Period	Treatment Period[a]								Follow-up Period[a]
	Screening	28-day cycles							EOT [b]	Follow-up
		Cycles 1 to 2				Cycles 3 to 4		Cycle 5 and beyond		
	Days -28 to -1	Day 1	Day 8	Day 15	Day 22	Day 1	Day 15	Day 1		
Concomitant Medications and Procedures	**Continuous until 28 days after the last study treatment**									
Transfusion Data Collection and Assessment [t]	**Assess and record on an ongoing basis (prior to each dose of IP) until the next AML therapy after discontinuation from study treatment, death, lost to follow-up, withdrawal of consent for further data collection, or the End of Trial, whichever occurs first. Clinical site staff should confirm if any transfusions were received by the subject (including any at outside local institutions in between study visits) prior to each IP administration via use of patient diary or other local procedure in place at the investigational site.**									
[redacted]										
BMA for Disease Assessment [v, g]	×	× [w]	-	-	-	× [x]	-	× [y]	× [z]	× [p]
[redacted]										
Blood for Pharmacodynamics [redacted] **[g, v]**	×	× [w]	-	×	-	× [x]	-	× [y]	× [z]	× [p]
[redacted]										
BMB [g, bb]	×	× [w]	-	-	-	× [x]	-	× [y]	× [z]	× [p]
PB Smear [g]	×	× [w]	-	-	-	× [x]	-	× [y]	× [z]	× [p]
Cytogenetics Testing [g]	×	× [w, cc]	-	-	-	× [x, cc]	-	× [y, cc]	× [z, cc]	× [p, cc]
Modified IWG Response and HI	-	× [w]	-	-	-	× [x]	-	× [y]	× [z]	× [p]

Events Table covers both Phase 1b & 2 (unless otherwise noted)	Screening Period	Treatment Period[a]								Follow-up Period[a]
	Screening	28-day cycles							EOT [b]	Follow-up
		Cycles 1 to 2				Cycles 3 to 4		Cycle 5 and beyond		
	Days -28 to -1	Day 1	Day 8	Day 15	Day 22	Day 1	Day 15	Day 1		
PB for Pharmacokinetics (AG-120 only in Phase 1b / AG-221 only in Phase 2)	-	× [ee]	-	-	-	× [ee]	-	-	-	-
IP Accountability [dd]	-	× [ff]	-	-	-	×	-	×	×	-
AG-120 / AG-221 Dispensation	-	×	-	-	-	×	-	×		-
Azacitidine Treatment Administration	-	**See Section 7.2 in Protocol for details**							-	-
Survival Follow-up	-	-	-	-	-	-	-	-	-	× [gg]
Subsequent AML Therapies	-	-	-	-	-	-	-	-	-	× [gg]

Abbreviations: AML = acute myeloid leukemia; β-hCG = β-subunit of human chorionic gonadotropin; BMA = bone marrow aspirate; BMB = bone marrow biopsy; BSA = body surface area; CR = morphologic complete remission; [redacted] ECHO = echocardiogram; eCRF = electronic case report form; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; EORTC= European Organization for Research and Treatment of Cancer; FCBP = female of childbearing potential; EOT = End of Treatment; HI = hematologic improvement; HSCT = hematopoietic stem cell transplantation; ICF = informed consent form; IDH2 = isocitrate dehydrogenase isoform 2; IP = investigational product; IWG = International Working Group; IWRS = interactive web response system; MRD = minimal residual disease; MUGA = multi-gated acquisition; PK = pharmacokinetics; PB = peripheral blood; RNA= Ribonucleic acid; SOC = standard of care; UGT1A1 = uridine diphosphate-glucuronosyltransferase 1 family, polypeptide A1

[a] One cycle (one month) is considered as 28 days (ie, 4 weeks). Unless noted otherwise, an administrative window of ± 3 days is permitted for all subsequent visits after the start of study treatment in each treatment cycle in the Treatment Phase. Day 1 of Cycles 2 and beyond may be delayed from Day 28 of the prior cycle in order for subjects to recover from toxicity and meet criteria for initiating each treatment cycle (Section 7.2 in Protocol). The study visit window for EOT or monthly-scheduled survival follow-up is ± 7 days.

[b] See Section 6.2.3 in Protocol for details.

[c] Including informed consent for mandatory genetic testing [redacted].

[d] The subject should start study treatment (ie, Day 1 of Cycle 1) within 3 days after enrollment or randomization.

[e] Subject should continue to be eligible for study entry prior to the start of study treatment on Day 1 of Cycle 1.

[f] Cycle 1 only.

[g] See Section 6.1 in Protocol for details regarding collecting BMA, BMB, PBS, and cytogenetics at screening for assessing AML diagnosis, collecting BMA and PB at screening for potential retrospective confirmation of mutational status, and collecting BMA and PB at screening for pharmacodynamics [redacted].

[h] All prior medications (prescription and non-prescription) taken and treatment procedures received from the 4-week period (ie, 28 days) prior to starting study treatment (including those prior to the start of study treatment on Day 1 of Cycle 1) and all prior anticancer therapies, regardless of discontinuation date of treatment.

[i] Source documented only.

[j] The BSA calculation is per the Dubois & Dubois formula: BSA (m^2) = weight (kg)$^{0.425}$ x height (cm)$^{0.725}$/139.2. The dose should be calculated on Day 1 of each treatment cycle. The dose during a treatment cycle should not be amended.

[k] These assessments will be done locally. See Section 6.2 in Protocol for details.

[l] If an assessment has been performed within 28 days prior to the start of study treatment, it does not need to be repeated.

[m] Pregnancy test is required for all FCBPs (see Section 4.2 in Protocol for the definition). Serum β-hCG pregnancy test (sensitivity of at least 25 mIU/mL) will be performed centrally at screening. For FCBP subjects, a local serum β- hCG pregnancy test (sensitivity of at least 25 mIU/mL) is to be done within 72 hours prior to study treatment administration on Day 1 of every treatment cycle in the Treatment Phase and at the EOT Visit. Negative results are required for study treatment administration.

[n] These assessments will be done centrally. See Section 6.2 in Protocol for details.

[o] A local serum pregnancy test (sensitivity of at least 25 mIU/mL) is to be done within 72 hours prior to the start of study treatment in the Treatment Phase for FCBP only (note that the screening central serum pregnancy test can be used as the test prior to the start of study treatment in the Treatment Phase if it is performed within the 72-hour timeframe).

[p] Whenever response/diagnosis is assessed in the Follow-up Period (ie, every 8 weeks [± 28 days]). See Section 6.3 in Protocol for details.

[q] Day 1 (± 14 days) of every 3rd cycle (eg, Cycles 3, 6, 9, etc) or more frequently per standard institutional practice. Not necessary for the EOT Visit if last performed within 14 days.

[r] Day 1 (± 28 days) of every 6th cycle (eg, Cycles 6, 12, 18, etc) or more frequently per standard institutional practice. Not necessary for the EOT Visit if last performed within 28 days.

[s] Continuous starting after signing ICF through 28 days after the last study treatment. See Section 6.4.1 in Protocol for details.

[t] Including type, number of units, reasons, and date of transfusions taken ≤ 8 weeks prior to the start of study treatment through 28 days after the last study treatment. Thereafter, transfusions will continue to be collected until the next AML therapy after discontinuation from study treatment, death, lost to follow-up, withdrawal of consent for further data collection, or the End of Trial, whichever occurs first.

[redacted]

[v] In addition to the frequency specified in the table, samples will also be collected if clinically indicated (eg, confirmation of CR/CRi/CRp, morphologic relapse after CR/CRi/CRp, or progression by a repeated bone marrow assessment at least 1 month later) or required for toxicity assessment. A sample of bone marrow and peripheral blood must also be sent at these time points for central pathology review.

[w] Within 7 days prior to Day 1 of Cycle 2. The assessment is not required at Day 1 of Cycle 1.

[x] Within 7 days prior to Day 1 of Cycle 3. The assessment is not required at Day 1 of Cycle 4.

[y] Within 7 days prior to Day 1 of Cycle 5 and Day 1 of every 2nd cycle thereafter (eg, Cycles 7, 9, etc).

[z] Not necessary for the EOT Visit if last performed within 28 days.

[redacted]

[bb] A bone marrow biopsy can be collected in conjunction with an aspirate if it is standard institutional practice. A bone marrow biopsy must be collected if adequate aspirate is not attainable.

[cc] A standard cytogenetic metaPhase preparation will be prepared if the bone marrow aspirate is obtained for assessing CR, morphologic relapse, or progressive disease and will be sent to the local laboratory for cytogenetic analysis.

[dd] Including diary cards. See Section 7.6 in Protocol for details.

[ee] Intensive PK sampling is performed as defined in Table 5 in Protocol for AG-120 subjects in the Phase 1b Expansion and Table 6 in Protocol at select sites for 6 - 12 in Phase 2. Sparse PK sampling will be performed at all other AG-221 patients at the selected sites as well as all other sites as defined in Table 7 in Protocol.

[ff] Cycle 2 only.

[gg] Every 4 weeks (± 7 days) for survival follow-up until death, lost to follow-up, withdrawal of consent for further data collection or the End of Trial, whichever occurs first. Subsequent AML therapies should be collected at the same time schedule. See Section 6.3 in Protocol for details.

[redacted]

17.4. International Working Group Acute Myeloid Leukemia Response Criteria

Hematologic Response According to modified IWG Criteria for AML	
Category	**Definition**
Morphologic Complete Remission (CR)	Defined as less than 5% blasts in a BM aspirate sample with marrow spicules and with a count of at least 200 nucleated cells. There should be no blasts with Auer rods and absence of extramedullary disease. Plus, all the following conditions should be met: • ANC >= 1 x 10^9/L (1,000/µL) • Platelet count >= 100 x 10^9/L (100,000/µL) • Independent of red cell transfusions for >= 1 week immediately before each response assessment
Morphologic Complete Remission with Incomplete Neutrophil Recovery (CRi) [a]	Defined as all criteria of morphologic CR except the following: • ANC < 1 x 10^9/L (1,000/µL)
Morphologic Complete Remission with Incomplete Platelet Recovery (CRp) [a]	Defined as all criteria of morphologic CR except the following: • Platelet count < 100 x 10^9/L (100,000/µL)
Morphologic Leukemia-free State (MLFS)	Defined as less than 5% blasts in a BM aspirate sample with marrow spicules and with a count of at least 200 nucleated cells. There should be no blasts with Auer rods and absence of extramedullary disease
Partial Remission (PR)	Defined as all hematologic criteria of morphologic CR with a > 50% decrease in the percentage of BM blasts to 5% to 25% (a blast count value of < 5% may also be considered a partial remission if Auer rods are present) [b]
Cytogenetic Complete Remission (CRc)	Defined as CR/CRi/CRp with a reversion to a normal karyotype in cases with an abnormal karyotype at baseline, based on evaluating >= 20 metaphase cells from BM
Morphologic Relapse after CR/CRi/CRp [a]	Defined as one of the following conditions: • Reappearance of >= 5% blasts in the BM not attributable to any other cause (eg, BM regeneration after consolidation therapy); or • Development of extramedullary disease
Not evaluable (NE) [a]	Defined as without a post-treatment response assessment
Stable Disease (SD) [a]	Defined as failure to meet any of the above criteria and not meeting the criteria of progressive disease (see below)
Progressive Disease (PD) [a]	Defined as one of the following conditions: • For subjects with 5 to 70% BM blasts at baseline: a > 50% increase of **BM blast count percentage** from baseline to >= 20%; or • For subjects with > 70% BM blasts at baseline: a doubling of **absolute blast count in PB** from baseline to >= 10 x 10^9/L (10,000/µL); or • Development of new extramedullary disease since last response assessment Progressive disease is to be confirmed by 2 consecutive response assessments separated by at least 1 month. The date of progressive disease is defined as the first date that one of three conditions listed above was met.

AML = acute myeloid leukemia; ANC= absolute neutrophil count; BM= bone marrow; PB=peripheral blood;;IWG= International Working Group. CRh = morphologic complete remission with partial hematologic recovery: defined as less than 5% blasts in a BM aspirate sample with marrow spicules plus ANC > 0.5 x 10^9/L (1,000/µL) & Platelet count > 50 x 10^9/L (100,000/µL).

[a] Modification to IWG response criteria.
[b] If the pre-treatment BM blast percentage was 50% to 100%, the percentage of blasts must decrease to a value between 5% and 25%; if the pre-treatment blast percentage was 20% to less than 49%, they must decrease by at least half to a value of more than 5%.

Notes: Deletions to the IWG response criteria are not shown.

[redacted]

17.5. Hematologic Improvement According to the International Working Group for Myelodysplastic Syndromes

Hematologic Improvement According to IWG Criteria	
Hematologic improvement[a]	**Response criteria (responses must last at least 8 week)**[b]
Erythroid Response (HI-E) (pre-treatment, < 11 g/dL)	• Hemoglobin increase by >= 1.5 g/dL • Relevant Reduction in units of RBC transfusions by an absolute number of at least 4 RBC transfusions/8 week compared with the pretreatment transfusion number in the previous 8 weeks Note: Only RBC transfusions given for a hemoglobin of <= 9.0 g/dL on treatment will count in the RBC transfusion response evaluation [b]
Platelet Response (HI-P) (pre-treatment, < 100 X 10^9/L)	• Absolute increase of >= 30 X 10^9/L for subjects starting with > 20 X 10^9/L platelets • Increase from < 20 X 10^9/L to > 20 X 10^9/L and by at least 100% [b]
Neutrophil Response (HI-N) (pre-treatment, < 1.0 X 10^9/L)	• At least 100% increase and an absolute increase > 0.5 X 10^9/L [b]
Progression or Relapse After HI [c]	At least 1 of the following: • At least 50% decrease from maximum response levels in granulocytes or platelets • Reduction in hemoglobin by >= 1.5 g/dL • Transfusion dependence

HI-E = hematologic improvement erythroid response; HI-N = hematologic improvement neutrophil response; HI-P = hematologic improvement platelet response; IWG = International Working Group; RBC = red blood cell.

Pre-treatment counts averages of at least 2 measurements (not influenced by transfusions, ie, no RBC transfusions for 2 weeks and no platelet transfusions for 1 week) >= 1 week apart (modification).
Modification to IWG [redacted] response criteria.

[c] In the absence of another explanation, such as acute infection, repeated courses of chemotherapy (modification), gastrointestinal bleeding, hemolysis, and so forth. It is recommended that the 2 kinds of erythroid and platelet responses be reported overall as well as by the individual response pattern.

Note: Deletions to the IWG response criteria are not shown. To convert hemoglobin levels from grams per deciliter to grams per liter, multiply grams per deciliter by 10.

[redacted]