STATISTICAL ANALYSIS PLAN

A Phase 1/2 Study of SL-401 as Consolidation Therapy for Adult Patients with Adverse Risk Acute Myeloid Leukemia in First CR, and/or Evidence of Minimal Residual Disease (MRD) in First CR

Protocol STML-401-0214

Protocol Number:	STML-401-0214
Protocol Version and Date:	Amendment 5: 27 March 2018 Amendment 4: 20 April 2017 Amendment 3: 13 July 2016 Amendment 2: 08 July 2015 Amendment 1: 09 January 2015 Original: 02 September 2014
Name of Test Drug:	SL-401 (Tagraxofusp)
Phase:	Phase 1/2
Methodology:	Non-randomized, open-label, dose escalation study, consisting of 2 stages
Sponsor:	Stemline Therapeutics, Inc. 750 Lexington Avenue, 11th Floor New York, NY 10022 Tel: (646) 502-2310 Fax: (646) 389-0968
Sponsor Representative:	[redacted]
Analysis Plan Date:	19 November 2020
Analysis Plan Version:	Final Version 1.0

APPROVAL SIGNATURE PAGE

Protocol Title: A Phase 1/2 Study of SL-401 in Consolidation Therapy for Adult Patients with Adverse Risk Acute Myeloid Leukemia in First CR, and/or Evidence of Minimal Residual Disease (MRD) in First CR

Sponsor: Stemline Therapeutics, Inc.
750 Lexington Avenue, 11th Floor
New York, NY 10022

Protocol Number: STML-401-0214

Document Date / Version: 19 November 2020/ Final Version 1.0

Veristat, LLC Author:

Approval Signature	Job Title
[redacted]	[redacted]

Sponsor Approval

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidances and guidelines.

I have discussed any questions I have regarding the contents of this document with the biostatistical author.

I also understand that any subsequent changes to the planned statistical analyses, as described herein, may have a regulatory impact and/or result in timeline adjustments. All changes to the planned analyses will be described in the clinical study report.

Sponsor Signatory:

Approval Signature	Job Title
[redacted]	[redacted]

TABLE OF CONTENTS

TABLES AND FIGURES INCLUDED IN THE TEXT

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Definition
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
AST	Aspartate aminotransferase
ATC	Anatomic therapeutic class
CI	Confidence interval
CLS	Capillary leak syndrome
CR	Complete response
CRi	Complete response [Incomplete blood count recovery]
CSC	Cancer stem cell
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DLT	Dose-limiting toxicity
DT	Diphtheria toxin
ECOG	Eastern Cooperative Oncology Group
ECG	Electrocardiogram
eCRF	Electronic case report form
IL-3	Interleukin-3
IL-3R	Interleukin-3 receptor
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MRD	Minimal residual disease
MTD	Maximum tolerated dose
OS	Overall survival
PCS	Potentially clinically significant
PK	Pharmacokinetic
PT	Preferred Term
RFS	Relapse-free survival
SAE	Serious adverse event
SAP	Statistical analysis plan
SCT	Stem cell transplant
SD	Standard deviation
SMQ	Standardized medical queries
SOC	System organ class

Abbreviation	Definition
TEAE	Treatment-emergent adverse event
ULN	Upper limit of normal
WHO	World Health Organization

1. INFORMATION FROM THE STUDY PROTOCOL

1.1. Introduction and Objectives

The overall purpose of study STML-401-0214 is to support the development of SL-401 and assess safety and clinical efficacy of the product in treatment of patients with acute myeloid leukemia (AML).

AML is characterized by the uncontrolled proliferation of immature myeloid cells in the bone marrow and peripheral blood, resulting in the development of anemia, neutropenia, and thrombocytopenia, and associated complications such as serious infections, bleeding, and fatigue. The median age at diagnosis is 67 years, and 5-year survival across all ages, treatments, and other prognostic groups is 24% (National Cancer Institute 2013). For AML patients who achieve hematologic remission following induction therapy, the evaluation of Minimal Residual Disease (MRD) provides additional information concerning the risk of relapse and mortality. The identification of MRD following AML remission has been associated with higher rates of relapse and diminished survival; for patients with intermediate risk AML in remission after 2 cycles of induction therapy, 1-year relapse rates were >60% for those with evidence of MRD versus <30% for those without MRD (Terwijn et al 2013). Approximately 70% of patients who receive first-line therapy and achieve a first complete response to treatment are expected to experience recurrent disease, and a subset of those who do not derive benefit from first-line treatment are candidates for subsequent therapy.

Although Complete Response (CR) rates after standard first-line induction therapy in AML are high, some subgroups of AML patients have limited tolerability of those regimens or have a high risk of relapse, even following the administered therapies (e.g. MRD). As a result, the National Comprehensive Cancer Network Guidelines (2013) recommend that AML patients with antecedent hematologic disease, age >60, or unfavorable cytogenetics are appropriate candidates for clinical studies with novel agents, even early in their disease. In addition, as there are no approved second-line AML treatments, there is unmet medical need for the relapsed/refractory population.

The target population for this clinical study is adult patients with adverse risk AML in first CR and/or with evidence of MRD in first CR.

1.1.1. Tagraxafusp (SL-401)

Diphtheria toxin (DT) Interleukin-3 (IL-3) fusion protein, named Tagraxofusp following Biologics Licensing Application (designated SL-401 during clinical development by Stemline Therapeutics, Inc. [“Stemline”]) is a novel biologic targeted therapy directed to the IL-3 receptor (IL-3R), a biomarker that is over-expressed on AML blasts and cancer stem cells (CSCs) relative to normal hematopoietic stem cells (Jordan, et al. 2006; Jordan, et al. 2000; Tehranchi, et al. 2010). Tagraxofusp is comprised of recombinant human IL-3 genetically fused to a truncated DT in which the binding domain of DT has been replaced with IL-3. The IL-3 domain of Tagraxofusp is able to target the agent to leukemia blasts and CSCs that over-express IL-3R, leading to receptor-mediated endocytosis and localization of Tagraxofusp to early endosomes. This allows Tagraxofusp to kill cells in a distinct manner from other cancer therapeutics. First, Tagraxofusp is a targeted therapy directed to the IL-3R that is present on CSCs and tumor bulk, but not on normal hematopoietic stem cells. Second, Tagraxofusp utilizes

a payload that is not cell cycle-dependent. Therefore, it is designed to kill not just highly proliferative tumor bulk, but also relatively quiescent CSCs. Lastly, Tagraxofusp utilizes a payload that may not be subject to multi-drug resistance mechanisms typically used by CSCs to evade traditional therapies. The payload also kills cells in a manner that is distinct from that of other available therapies, which is another reason why Tagraxofusp may be an effective addition to the therapeutic armamentarium against certain hematologic malignancies.

This study, STML-401-0214, was designed as a non-randomized, open-label, dose escalation, multicenter study, divided into 2 stages. Stage 1 (dose escalation) enrolled 9-18 patients with AML. Stage 2 (expansion) enrolled up to 20 additional patients with evidence of MRD.

1.1.2. Study Objectives

The following objectives are as defined in the Study Protocol (Amendment 5, 27 March 2018).

The primary objectives are to determine the maximum tolerated dose (MTD), or the maximum tested dose where multiple dose-limiting toxicities (DLTs) are not observed, of Tagraxofusp, and to characterize the safety profile of Tagraxofusp at the MTD or maximum tested dose.

Secondary objectives include:

- Evaluate the presence of MRD and changes in MRD status during and following Tagraxofusp therapy
- Estimate relapsed-free survival (RFS)
- Estimate overall survival (OS)
- Characterize the pharmacokinetics (PK) of Tagraxofusp
- Characterize the immunogenicity of Tagraxofusp.

Exploratory objectives are to characterize expression of IL-3R/CD123 (and other potentially relevant stem cell and disease markers) on leukemia cells in bone marrow (when feasible), to evaluate potential changes in IL-3R/CD123 (and other potentially relevant markers) expressing populations over time, and preliminary correlation of baseline IL-3R/CD123 (and other potentially relevant markers) expression and clinical efficacy (including changes in MRD status).

1.1.3. Purpose of this Document

This statistical analysis plan (SAP) is designed to outline the methods to be used in the analysis of study data in order to answer the study objective(s). Populations for analysis, data handling rules, statistical methods, and formats for data presentation are provided. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of the clinical study report (CSR) for this trial.

This SAP will also outline any differences in the currently planned analytical objectives relative to those planned in the study protocol.

1.2. Study Design

1.2.1. Synopsis of Study Design

This study is a non-randomized, open-label, dose escalation, multicenter study, divided into 2 stages: Stage 1, dose escalation and Stage 2, expansion. A cycle of therapy is 28 days.

Stage 1 (dose escalation) was designed to identify a MTD. During Stage 1, approximately 9-18 patients were planned to be treated with Tagraxofusp. The starting dose of Tagraxofusp was 7 μg/kg/day for 5 consecutive days every 28 days, with escalation to 9 and 12 μg/kg/day. Three to 6 patients were planned to be treated at each dose level. All enrolled subjects at each corresponding dose were enrolled and followed for their complete first cycle before additional patients could be enrolled. If 0/3 patients experienced a DLT, 3 new patients would be enrolled at the next sequential dose level. If 1/3 patients experienced a DLT, 3 additional patients would be enrolled at the same dose level. If 1/6 patients experienced a DLT, dose escalation could continue with dosing of a new cohort at the next higher dose level. If 2/3 or 2/6 patients experienced a DLT, the MTD was exceeded and further dose escalation would not occur. The MTD was defined as the dose preceding the dose level at which 2 or more patients experience a DLT during treatment Cycle 1. No intra-patient dose escalation was allowed.

During Stage 1, DLT was defined as any of the following occurring during the first cycle of therapy:

- Any treatment-emergent Grade 4 transaminase or creatine phosphokinase (CPK) elevation (confirmed within 24 hours of initial identification), regardless of duration or relationship to Tagraxofusp.
- Any treatment-emergent Grade 4 hematologic toxicity (unrelated to recurrent leukemia or prior AML therapy) lasting > 28 days after last infusion of Tagraxofusp.
- Any treatment-emergent Grade ≥3 non-hematologic toxicity (unrelated to recurrent leukemia), with the exception of Grade 3 laboratory toxicities that resolve to Grade ≤1 or baseline ≤28 days after the last infusion of Tagraxofusp , or the following Grade 3 toxicities if they resolve to Grade ≤1 or baseline ≤28 days after last infusion of Tagraxofusp: arthralgia, myalgia, fever responding to treatment, nausea and/or vomiting (excluding cases that require tube feeding, total parenteral nutrition or hospitalization) or diarrhea associated with suboptimal prophylaxis or treatment.

During Stage 2 (expansion), up to 20 additional patients with evidence of MRD as determined locally would be treated at the MTD or maximum tested dose at which multiple DLTs are not observed (identified as 12 μg/kg/day in Stage 1) so that up to 15 patients with evidence of MRD as determined centrally would be evaluable for safety and response at this dose.

1.2.2. Randomization Methodology

As this is a single-agent study, randomization is not applicable.

1.2.3. Stopping Rules and Unblinding

Tagraxofusp treatment could be discontinued for any of the following reasons:

- Patient withdrawal of consent
- Occurrence of unacceptable toxicity, including DLT
- Tagraxofusp related anaphylaxis or Grade ≥3 hypersensitivity reaction
- Requirement for >1 dose reduction unless there is evidence of MRD eradication or sustained AML remission (beyond Cycle 2), in which case additional dose reductions are permitted, however these reductions must be discussed with the Medical Monitor and documented in the context of ongoing AML response
- Disease recurrence/progression
- Intercurrent illness that prevents further administration of Tagraxofusp
- Patient non-compliance
- Occurrence of pregnancy
- Completion of 6 cycles of treatment. (The administration of additional cycles of Tagraxofusp [beyond 6 cycles] may be considered in the setting of MRD eradication, sustained AML remission, or other evidence of sustained anti-leukemia benefit in the opinion of the Investigator. Administration of Tagraxofusp beyond 6 cycles must be discussed with the Medical Monitor at which time the individual patient's potential risk/benefit of further treatment will be assessed).
- Investigator's decision

The reason for Tagraxofusp discontinuation and the date of discontinuation are recorded in the electronic case report form (eCRF).

1.2.4. Study Procedures

The schedules of assessments, as outlined in the study protocol, are provided in Section 7, Table 7-1 and Table 7-2.

1.2.5. Efficacy, Pharmacokinetic, and Safety Parameters

1.2.5.1. Efficacy Parameters

Efficacy assessments include rates of MRD eradication (conversion), RFS, and OS. Response/remission will be assessed using International Working Group criteria for AML (Cheson et al. 2003). Response criteria for AML are summarized in Section 8.1. All efficacy outcomes will be presented by subgroups defined by MRD Status and line of therapy in study (First-Line or Relapsed/Refractory). Definitions for each efficacy assessment are described in Section 4.3.

1.2.5.2. Safety Parameters

Safety evaluations performed during the study included physical examinations, Eastern Cooperative Oncology Group (ECOG) performance status, measurement of vital signs, 12-lead electrocardiograms (ECGs), clinical laboratory evaluations (hematology, serum chemistry, coagulation, and urinalysis), and monitoring of adverse events (AE) (including treatment serious AEs [SAEs]), DLTs, and concomitant medications (CM).

Further details on the definitions and analysis methods for safety endpoints are provided in Section 4.4.

1.2.5.3. Pharmacokinetic and Immunogenicity Parameters

Noncompartmental PK parameter summaries are described in a separate Modeling and Simulation Plan.

Summary of immunogenicity parameters are described in a separate plan.

2. PATIENT POPULATION

2.1. Population Definitions

The following patient populations will be evaluated and used for presentation and analysis of the data:

- Modified Intent-to-Treat (mITT) Population: All patients who are eligible based on the screening criteria and who received at least 1 dose of Tagraxofusp. Patients who meet the criteria for mITT will be considered "evaluable." Patients will be grouped according to the planned dose level at time of enrollment.

- Safety Population: All patients enrolled in the study who received at least 1 dose of Tagraxofusp. Patients will be grouped according to the actual dose level received.

The mITT population is the primary population for the analysis of efficacy parameters. Supportive subgroup efficacy analyses will be performed in the mITT population, as described in Section 3.8. The safety population is the primary population for the analysis of safety parameters.

2.2. Protocol Violations

All protocol violations will be presented in a data listing.

3. GENERAL STATISTICAL METHODS

3.1. Sample Size Justification

Approximately 21-33 adult patients diagnosed with AML that is considered high-risk disease who have achieved a first or second CR or complete remission with incomplete bone marrow recovery (CRi) within 6 months prior to enrollment and are not considered immediate candidates for allogenic stem cell transplant (SCT). The anticipated sample size is sufficient to evaluate the primary objectives.

During Stage 1, approximately 9-18 patients will be treated with Tagraxofusp.

In Stage 2, up to 20 additional patients with MRD as determined locally will enroll and receive Tagraxofusp at either the MTD or maximum tested dose, so that a total of 15 patients with evidence of MRD, as determined centrally, are treated at this dose. The assumptions governing sample size are as follows:

- Null hypothesis: eradication of MRD ≤5%
- Alternate hypothesis: eradication of MRD ≥20%
- Type 1 error: 17% one-sided
- Power: >80%

Because no available anti-leukemia therapies (other than allogeneic transplant) are believed to be associated with the eradication of evidence of MRD in this high-risk AML setting, the study would be considered positive if ≥2 of 15 patients in 1st remission are converted from MRD-positive to MRD-negative status. (In a situation where 1 of 15 patients has eradication of MRD, the results may be considered of interest, given the dearth of efficacious therapy in this setting).

At the time of database lock, in Stage 1, 18 patients were enrolled. In Stage 2, 9 patients with MRD were enrolled.

3.2. General Methods

All data listings that report an evaluation date will include a relative study day (Rel Day). Pre-treatment and on-treatment study days are numbered relative to the day of the first dose of study drug which is designated as Day 1. The preceding day is Day -1, the day before that is Day -2, etc. The last day of study drug is designated with an “L” (eg, Day 26L). Post-treatment study days are numbered relative to the last dose and are designated as Day 1P, Day 2P, etc.

All output will be incorporated into Microsoft Word files, sorted and labeled according to the International Council for Harmonisation (ICH) recommendations, and formatted to the appropriate page size(s).

Tabulations will be produced for appropriate demographic, baseline, efficacy, and safety parameters. For categorical variables, summary tabulations of the number and percentage of patients within each category (with a category for missing data) of the parameter will be presented. Two-sided 95% confidence intervals (CIs) will be computed using the Clopper Exact

method. For continuous variables, the number of patients, mean, median, standard deviation (SD), minimum, and maximum values will be presented. Time-to-event data will be summarized using Kaplan-Meier methodology using 25th, 50th (median), and 75th percentiles with associated 2-sided 95% CIs, as well as percentage of censored observations. Formal statistical hypothesis testing will be performed on a secondary endpoint, eradication or MRD rate among in high-risk AML enrolled in Stage 2.

3.3. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software Version 9.4, unless otherwise noted. Medical history and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 19.0. Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary (September, 2016).

3.4. Baseline Definitions

For all analyses, baseline will be defined as the most recent measurement prior to the first administration of study drug.

3.5. Methods of Pooling Data

Overall summaries of safety parameters will be presented by stage of enrollment for AML patient groups across all dose levels.

Efficacy analyses for patients with AML who received 12 µg/kg in Stage 2 will be conducted. Subgroup analyses by MRD status and line of therapy will also be performed.

3.6. Adjustments for Covariates

Due to small sample size, no formal statistical analyses that adjust for possible covariate effects are planned. Exploratory subgroup analyses will be performed for descriptive purposes, as described in Section 3.8.

3.7. Multiple Comparisons/Multiplicity

Multiplicity is not of concern for this Phase 1 study.

3.8. Subgroup Analyses

Subgroup analyses will be performed for efficacy analyses by MRD status and line of therapy in study (First-Line or Relapsed/Refractory).

3.9. Withdrawals, Dropouts, Loss to Follow-up

At the discretion of the Sponsor, additional patients may be enrolled to supplement patient data compromised due to premature study dropout or other reasons.

3.10. Missing, Unused, and Spurious Data

In general, there will be no substitutions made to accommodate missing data points. All data recorded on the eCRF will be included in data listings that will accompany the CSR.

When tabulating AE data, partial dates will be handled as follows. If the day of the month is missing, the onset day will be set to the first day of the month unless it is the same month and year as study treatment. In this case, in order to conservatively report the event as treatment-emergent, the onset date will be assumed to be the first day of treatment. If the onset day and month are both missing, the day and month will be assumed to be January 1, unless the event occurred in the same year as the study treatment. In this case, the event onset will be set to the first day of treatment in order to conservatively report the event as treatment-emergent. A missing onset date will be set to the first day of treatment.

3.11. Visit Windows

It is expected that all visits should occur according to the protocol schedule. All data will be tabulated per the evaluation visit as recorded on the eCRF even if the assessment is outside of the visit window. In data listings, the relative day of all dates will be presented (see Section 3.2).

3.12. Timing of Analysis

[redacted] Efficacy data were not summarized at that time. The respective analyses were described in a separate Statistical Analysis Plan.

There were no planned interim efficacy analyses. The Final, End of Study Analysis in support of the CSR will be supported from the analyses described in this document.

4. STUDY ANALYSES

4.1. Patient Disposition

Patient disposition will be tabulated and include the number screened, the number treated in total, the number in each patient population for analysis, the number treated in each Stage of the study, the number who withdrew prior to completing the study and reason(s) for withdrawal. The summary will be presented by dose level and stage.

A data listing of study completion information, including the reason for premature study withdrawal, if applicable, will be presented.

4.2. Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized and presented by dose level and Stage. Age, height, and weight will be summarized using descriptive statistics (number of patients, mean, SD, median, minimum, and maximum). The number and percentage of patients in each gender, ethnicity, race, and ECOG performance status category will be presented.

Medical history and results from baseline screens such as for pregnancy, and smoking history will be tabulated. Primary disease history will be summarized including CD123 status, AML baseline diagnosis, time since AML diagnosis, and adverse-risk AML and MRD disease elements. Prior radiotherapy and prior systemic therapy (for AML and not related to AML) will be summarized.

Demographic and baseline data for each patient will be provided in data listings.

4.3. Efficacy Evaluation

Data for all efficacy endpoints will be presented in by-patient listings. Subgroup analyses for all efficacy outcomes will be presented by MRD Status and line of therapy.

4.3.1. Eradication (Conversion) of MRD

The rate of MRD eradication (conversion) will be presented as the number and percentage of patients with evidence of MRD prior to initial treatment with Tagraxofusp for whom MRD cannot be detected upon subsequent assessments.

Rates of MRD eradication at a single assessment and at multiple time points following initiation of Tagraxofusp will be presented.

95% Clopper Exact CIs will be presented for each rate.

4.3.2. Relapse-Free Survival

Relapse-Free Survival is defined as the time from the date of first infusion of Tagraxofusp to the date of relapse (AML recurrence) or death from any cause, whichever occurred first. According to current WHO classification, a myeloblast population comprising greater than 5% of nucleated cells in the bone marrow or blood is required for diagnosis of AML relapse following CR/CRi

(if no peripheral blasts are present, then a confirmation aspirate is required ≥1 week subsequent to the aspirate at which >5% blasts were identified).

For patients who receive SCT, RFS will include time to progressive disease or death post-transplant. Patients who do not progress and are still alive at the time of analysis will be censored on the date of last treatment recorded prior to the analysis cut-off date. If patients start other anticancer therapies (with the exception of SCT or preparations for SCT), they will be censored at the latter of the date of last treatment with Tagraxofusp or date of last disease assessment recorded that occurred prior to the start of the new therapy.

The proportion of subjects receiving treatment who are alive and without evidence of AML recurrence will be presented. The distribution for RFS will be estimated by Kaplan-Meier methodology and the 25th percentile, median, 75th percentile, number and percentage of events and censored observations, and appropriate CIs will be presented. 6- and 12-month RFS rates will be evaluated.

4.3.3. Overall Survival

Overall survival is defined as the time from the date of first infusion of Tagraxofusp to the date of death from any cause. Patients still alive or lost to follow-up at the time of the analysis will be censored on the last date known to be alive prior to the analysis cut-off date, as determined by in- person visit or telephone contact. The overall distribution for OS will be estimated by Kaplan-Meier methodology in a similar manner to RFS.

4.4. Safety Analyses

All safety tabulations will be presented by dose received.

4.4.1. Study Drug Exposure

Duration of study drug exposure will be calculated as the number of days and number of cycles patients were administered study drug.

$$Duration\ of\ Study\ Drug\ Exposure = (Date\ of\ last\ dose - Date\ of\ first\ dose) + 1$$

Duration in cycles of drug exposure will be defined the total number of cycles of study drug initiated while the patient is on study. Total dose administered will be summarized overall and by cycle.

Relative Dose Intensity will be computed using the following definition:

$$Relative\ Dose\ Intensity\ (\%) = 100 * \frac{Sum(Cumulative\ Actual\ Dose\ Received\ by\ Cycle)}{Sum(Planned\ Dose\ to\ be\ Administered\ by\ Cycle)}$$

Exposure to Tagraxofusp will be presented in by-patient data listings.

4.4.2. Adverse Events

All AEs will be coded using the MedDRA coding system and displayed in tables and data listings using system organ class (SOC) and preferred term (PT).

Analyses of AEs will be performed for those events that are considered treatment-emergent, where treatment-emergent is defined per protocol as any AE with onset after the first administration of Tagraxofusp through 30 days after the last dose of Tagraxofusp, or any event that was present at baseline but worsened in intensity or was subsequently considered drug-related by the Investigator through 30 days after the last dose of Tagraxofusp.

The number and percentage of patients with any treatment-emergent AE (TEAE), with any TEAE assessed by the Investigator as related to treatment, with any TEAE with severity ≥ Common Terminology Criteria for Adverse Events (CTCAE) Grade 3, with any SAE, with any AE leading to discontinuation of study treatment, and with any AE leading to dose modification will be summarized by dose group and overall. In these tabulations, each patient will contribute only once (ie, the most related occurrence or the most intense occurrence) to each of the incidence rates in the descriptive analysis, regardless of the number of episodes.

Treatment-emergent AEs summarized by patient incidence rates are not tabulated by severity or relationship to study treatment; therefore, in such tabulations, a patient contributes only once to the count for a given AE SOC or PT. For tabulations that include classification by relationship to study treatment, AEs with missing relationship will be considered related to study drug.

Adverse events of special interest (AESI's) will be determined using MedDRA version 19.0 standardized medical queries (SMQs), high-level terms, or PTs, as follows:

- Possible hypersensitivity events, based on the SMQ Hypersensitivity (broad search).
- Vascular capillary leak syndrome (CLS), based on the MedDRA SMQ [to be provided] in addition to PTs of hypoalbuminaemia, blood albumin decreased, and proteinuria.
- Possible drug-induced liver injury events, based on the SMQ drug-related hepatic disorders (broad search).

Summary tables of AESIs produced will report AESIs as follows:

- AESIs overall and by PT,
- AESIs by PT overall and by grouped cycle (Cycle 1, Cycles 2-4, Cycles ≥5),
- ≥3 Grade AESIs overall and by PT,
- Serious AESIs overall and by PT,
- AESIs resulting in drug interruption overall and by PT,
- AESIs resulting in dose reduction overall and by PT,
- AESIs resulting in study drug withdrawn/ discontinued overall and by PT,

- AESIs resulting in death overall and by PT.

These tabulations will be completed for all AML patients, by dose group. AESIs will also be included in data listings.

4.4.2.1. Capillary Leak Syndrome

In addition to the summaries described in Section 4.4.2, summaries will be provided that include:

- Summary of number and grade of CLS events, exposure to Tagraxofusp prior to onset of CLS, and time to first onset and time to resolution of CSL events.

4.4.3. Laboratory Data

Clinical laboratory values will be expressed in Système International (SI) units.

The actual value and change from baseline (Day 1) to each on-study evaluation through cycle 6 will be summarized for each clinical laboratory parameter, including hematology, clinical chemistry, coagulation, and urinalysis. In the event of repeat values, the last non-missing value per study day/time will be used.

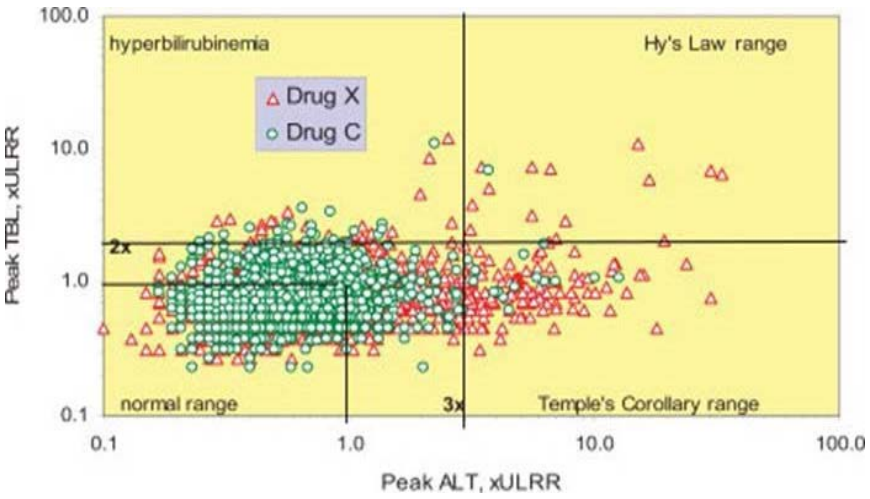
Shift tables of change in CTCAE grade of laboratory parameters from baseline to worst value and from baseline to last value on study will be presented. Both scheduled and unscheduled visits will be included in shift tables.

To assess for possible drug-induced liver injury (FDA 2009), a figure plotting peak alanine aminotransferase (ALT) versus peak total bilirubin (both on a logarithmic scale × upper limit of normal [ULN]) will be produced similar to that recommended by Watkins et al (Watkins, et al. 2008) so that values within the normal reference range (<ULN) for ALT and total bilirubin are found in the left lower quadrant and Hy's Law case candidates are in the upper right quadrant (ALT >3 × ULN and total bilirubin >2 × ULN). Patients with Gilbert's syndrome or cholestasis are typically found in the upper left quadrant, and patients with ALT elevations without significant hepatic abnormality (ie, without increased total bilirubin) are found in the lower right quadrant. The peak total bilirubin value plotted will be the peak within ± 7 days of the peak ALT value. If at least five patients are identified as potential Hy's Law Cases, a summary table will be produced for the number and percentage of patients who have values of aspartate aminotransferase (AST) (>3 to ≤5 × ULN, >5 to ≤10 × ULN, >10 to ≤20 × ULN, >20 × ULN), ALT (>3 to ≤5 × ULN, >5 to ≤10 × ULN, >10 to ≤20 × ULN, >20 × ULN), alkaline phosphatase (>1.5 × ULN), or total bilirubin (>1.5 to ≤2 × ULN, >2 × ULN) after initiation of study drug. In this table, a patient may be counted only one time. For instance, a patient with a result of AST of 11 × ULN will be counted in the category of >10 to ≤20 × ULN.

This plot will be repeated for peak aspartate aminotransferase (AST) by total bilirubin.

An example of the scatterplot is provided in Figure 4-1.

Figure 4-1: Hy's Law Candidates Scatterplot



Time to first onset of elevated AST and ALT will be analyzed using Kaplan-Meier analyses and be presented as figures.

All laboratory data will be provided in by-patient data listings.

A by-patient listing will also be presented for all laboratory values with CTCAE Grade ≥3.

4.4.4. Vital Signs and Physical Examination

The actual value and change from baseline (Day 1) to each on-study evaluation and to the last on study evaluation will be summarized for vital signs.

A summary table of the number and percent of patients with treatment-emergent potentially clinically significant (PCS) vital signs parameters will be tabulated based on the following criteria:

	PCS – Low if:			**PCS – High if:**		
Variable Name	**Observed Value is:**	**AND**	**Decrease from Baseline is:**	**Observed Value is:**	**AND**	**Increase from Baseline is:**
Systolic Blood Pressure	<90 mmHg		≥20 mmHg	>180 mmHg		≥20 mmHg
Diastolic Blood Pressure	<50 mmHg		≥10 mmHg	>105 mmHg		≥10 mmHg
Heart Rate	<50 bpm		≥15 bpm	>120 bpm		≥15 bpm

PCS= potentially clinically significant.

All tables summarizing vital sign measurements only include visits in which at least 10% of the analysis population had measurements. Vital sign measurements will be presented for each patient in a data listing.

All physical examination findings and ECOG performance status findings will be tabulated by visit and presented in data listings.

4.4.5. Electrocardiogram

ECG results will be summarized descriptively, including the number and percentage of patients with normal, abnormal, and clinically significant abnormal results at baseline and each study visit. Actual values and change from baseline will be summarized for QTc intervals. All tables summarizing ECG measurements only include visits in which at least 10% of the analysis population had measurements

The number and percentage of patients whose mean QTcF or QTcB value at any time point meets any of the following categories will be summarized:

- >450 msec
- >480 msec
- >500 msec
- increase from baseline >30 msec
- increase from baseline >60 msec

QTcF and QTcB will be derived using QT interval and heart rate using the formulas:

$$QTcF = \frac{QT\ Interval}{Heart\ Rate^{1/3}} \qquad QTcB = \frac{QT\ Interval}{Heart\ Rate^{1/2}}$$

ECG and echocardiogram/multigated acquisition scan (MUGA) data for each patient will be provided in data listings.

4.4.6. Concomitant Medications and Subsequent Treatments

Concomitant medications will be coded using the WHO Drug Dictionary. Results will be tabulated by anatomic therapeutic class (ATC) and PT.

Concomitant medications will be tabulated by dose group, where any medications that were not discontinued prior to the first dose of study drug will be included. If an end date is missing or the medication is ongoing at the time of first dose, the medication will be considered concomitant.

The use of concomitant medications and subsequent treatment(s) will be included in by-patient data listings.

4.4.7. Pre-Treatment Medications

Pre-treatment medications will be coded using the WHO Drug Dictionary. Results will be tabulated by ATC and PT. In addition, the number and percent of patients who receive pre-treatment medication at least once will be summarized by dose group.

Pre-treatment medications will be included in by-patient data listings.

5. CHANGES TO PLANNED ANALYSES

As of this date, there have been no changes between the protocol-defined statistical analyses and those presented in this statistical analysis plan. All changes from procedures outlined in this SAP will be summarized in the study report. Decisions to deviate from planned analyses will be documented at the time they are made.

6. REFERENCES

Cheson BD, Bennett JM, Kopecky KJ, Buchner T, Willman CL, Estey EH, et al. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. *J Clin Oncol.* 2003;21(24):4642-9.

FDA. Guidance for Industry: Drug-Induced Liver Injury: Premarketing Clinical Evaluation. 2009.

Jordan CT, Guzman ML, and Noble M. Cancer stem cells. *N Engl J Med.* 2006;355(12):1253-61.

Jordan CT, Upchurch D, Szilvassy SJ, Guzman ML, Howard DS, Pettigrew AL, et al. The interleukin-3 receptor alpha chain is a unique marker for human acute myelogenous leukemia stem cells. *Leukemia.* 2000;14(10):1777-84.

National Cancer Institute. SEER Stat Fact Sheets: Acute Myeloid Leukemia. 2013, Vol. 2017: Surveillance, Epidemiology, and End Results (SEER) Program.

Shankar G, Arkin S, Cocea L, Devanarayan V, Kirshner S, Kromminga A, et al. Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. *AAPS J.* 2014;16(4):658-73.

Tehranchi R, Woll PS, Anderson K, Buza-Vidas N, Mizukami T, Mead AJ, et al. Persistent malignant stem cells in del(5q) myelodysplasia in remission. *N Engl J Med.* 2010;363(11):1025-37.

Terwijn M, van Putten WLJ, Kelder A et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. J Clin Oncol 2013: 31(31):3889-97.

Watkins PB, Seligman PJ, Pears JS, Avigan MI, and Senior JR. Using controlled clinical trials to learn more about acute drug-induced liver injury. *Hepatology.* 2008;48(5):1680-9.

7. STUDY FLOW CHARTS

The following Study Flow Charts are from Protocol Amendment 8 and Amendment 9.

Table 7-1: Study Events Schedule for Cycle 1 (Study Day -14 to Study Day 28-35)

		Cycle 1					
	Study Day -14 to -4	**Study Day -1 to 0**	**Study Days 1-5 (Up to Study Day 10 if Infusion(s) Held) SL-401 Treatment**		**Study Day 8±3[n], 15±3, and 21±3**	**Study Day 28±3[o]**	**Study Day 35±3, Then Every 7±3 days**
Tests and Observations	***Screening***	***Pre-treatment***	***Pre-Infusion***	***Infusion***		***End of Cycle[o]***	***Delayed End of Cycle for Toxicity Resolution Only if Required***
Informed consent form	X						
Inclusion/exclusion criteria	X						
Medical history including prior therapy, concomitant medications	X						
Concomitant Medications		X			X	X	X
ECOG performance status	X					X	X
Physical examination	X	X				X	
Pregnancy test [a]	X						
Vital signs and weight[b]	X	X	X	X		X	X
12-lead ECG[c]	X	X	X	X (Infusions 1, 5)		X	X
ECHO or MUGA scan[d]	X						
Hematology[e]	X	X	X		X	X	X
Serum electrolytes[f]	X	X	X		X	X	X
Serum albumin[g]	X	X	X		X	X	X
Serum chemistry[h]	X	X	X		X	X	X
Coagulation parameters: PT/INR, aPTT	X		X		X	X	X
Urinalysis[i]	X		X (Infusion 1)		X	X	X
Tumor response assessment: Bone marrow aspiration + biopsy[j]	X						
Administration of premeds [k]			X				
SL-401 administration[l]				X			
Pharmacokinetic sampling[m]			X (Infusions 1, 5)	X (Infusions 1, 5)			
Immunogenicity sampling			X (Infusion 1)		X (day 15)	X	
Vision assessment	X				X (day 21)		
AE and SAE monitoring			X	X	X	X	X

a Urine or serum pregnancy test must be performed within 1 week prior to treatment in women of childbearing potential.

b Vital signs should be performed after patient is sitting for 3-5 minutes. If during dosing period, vital signs should be taken immediately prior to infusion, immediately after completion of infusion, and 30, 60, and 240 minutes post-infusion.

c All patients will have a 12-lead ECG performed at the screening visit and pre-treatment visit, as well as Day 28. During the days when patients are undergoing PK sampling (Cycles 1 & 3, infusions 1 and 5), an ECG will be performed at 3 distinct time points (triplicates) within 5 minutes prior to each PK sample collection pre-infusion and at 30 and 60 minutes post-infusion (see footnote (m) and Table 6).

d MUGA or 2-D ECHO to quantify LVEF. Must be completed within 28 days prior to start of first cycle of study drug.

e To be collected prior to SL-401 infusion if during dosing period. Hematology includes WBC count with differential, RBC count, hematocrit, hemoglobin and platelet count.

f To be collected prior to SL-401 infusion if during dosing period. Electrolytes include sodium, potassium, bicarbonate, chloride, BUN, creatinine and glucose.

g To be collected prior to SL-401 infusion if during dosing period. Serum albumin may be a component of the chemistry panel (h). See protocol for administration of albumin if serum albumin decreases to <3.0 g/dL during treatment days or in the immediate post-treatment period.

h To be collected prior to SL-401 infusion if during dosing period. Serum chemistry includes electrolytes (see above, footnote f) and the following: ALT, albumin, ALP, AST, bilirubin (total, direct, and indirect), calcium, creatine phosphokinase (CPK), magnesium, LDH, phosphate, total protein, and uric acid.

i To be collected prior to SL-401 infusion if during dosing period. Urinalysis includes appearance, color, pH, specific gravity, ketones, leukocytes, protein, glucose, bilirubin, urobilinogen, and occult blood.

j Morphology and differential WBC/blast count on aspirate. Baseline must be performed within 14 days prior to the first administration of SL-401. Subsequent bone marrow aspirates (+ biopsy) will be performed at the end of Cycles 2, 4 and 6 and every 3 months (± 1 month) thereafter through 12 months after the start of SL-401 and at the Investigator's discretion thereafter. If the end-of- Cycle 2/4/6 bone marrow aspirate (+ biopsy) is empty, hypocellular, or inadequate, a bone marrow examination should be repeated within 14 (± 7) days to document response. All bone marrow samples collected on study will be assessed centrally. Bone marrow (2 mL aliquot) will be prioritized for assessment of MRD at the [redacted]. Additional bone marrow (~5 mL) is to be aliquoted for tumor assessment and for translational evaluation at MD Anderson Cancer center. CD123 is to be assessed in all bone marrow samples (and, if applicable, tissue samples) by flow cytometry (i.e., CD123 should be added to the panel of markers assessed by flow cytometry of bone marrow aspirates) and immunohistochemistry and the results recorded and captured in the eCRF.

k Refer to Section 7.5.2 – Premedication and Administration.

l Following treatment with premedication, SL-401 will be administered as a 15-minute infusion for the first 5 consecutive days of a 28-day cycle. Individual SL-401 infusions may be delayed to allow for toxicity resolution; all infusions should be completed within 10 days. Patient must be monitored for 4 hours post infusion.

m Plasma samples (6 mL each) will be collected immediately prior to the start of the infusion of SL-401, immediately after end of infusion (time recorded), then 15, 30, 45, 60, 90, 120, 180, and 240 minutes after completion of the infusion during infusions 1 (i.e., Study Day 1) and 5 (i.e., Study Day 5) during Cycles 1 & 3 (see (c)).

n For patients who live a considerable distance from the Study Center, for whom weekly travel to the Study Center is not feasible, the Day 8, 15, and 21 (±3 days) laboratory assessments (blood and urine) may be submitted to a local laboratory, although the results must be evaluated by the Study team. Concomitant medication and AE monitoring on these days will take place via telephone contact. In situations where SL-401 is not administered over 5 consecutive days (i.e., treatment delays because of AEs or other factors) and an infusion is to be administered on Day 8, all pre-infusion/infusion evaluations should be conducted, in addition to concomitant medications recorded. Upon completion of the 5th Cycle 1 infusion (Day 8, 9 or 10), the blood/urine/concomitant medication and AE assessments should be planned for Day 15, as indicated in the table.

o The end-of-the cycle evaluations (Day 28 or thereafter) may also serve as the pre-infusion evaluations for Cycle 2; these do not need to be duplicated on successive days unless there is an abnormality or other clinically relevant reason for repeat evaluation.

Table 7-2: Study Events Schedule for Cycles 2-6 and Subsequent Follow-up

	Cycle 2-6							
	Days1-5 (Up to Day 10 if Infusion(s) Held) SL-401 Treatment		Day 8±3[m], 15±3 and 21±3	Day 28±3[n]	Day 28±3, Then Every 7±3 days	End of Treatment	Safety: Through 30 Days After Last Infusion	Survival: Every 90 Days After Last Infusion
Tests and Observations	***Pre-Infusion***	***Infusion***		***End of Cycle[n]***	***Delayed End of Cycle for Toxicity Resolution Only if Required***			
Concomitant Medications	X		X	X	X		X	
ECOG performance status				X	X	X		
Physical examination				X		X		
Vital signs and weight[a]	X	X		X	X	X		
12-lead ECG[b]	X	X (Cycle 3; infusions 1 & 5)		X	X			
Hematology[c]	X		X	X	X			
Serum electrolytes[d]	X		X	X	X			
Serum albumin[e]	X		X	X	X			
Serum chemistry[f]	X		X	X	X			
Coagulation parameters: PT/INR, aPTT	X		X	X	X			
Urinalysis[g]	X (Infusion 1)		X	X	X			
Tumor response assessment: Bone marrow aspiration + biopsy [h]				X (Cycle 2, 4, 6)		X[h]		
Administration of premeds [i]	X							
SL-401 administration[j]		X						
Pharmacokinetic sampling[k]	X (Cycle 3; Infusions 1 &5)	X (Cycle 3; Infusions 1 &5)						
Immunogenicity sampling	X (Infusion 1)			X				X[l]
Vision assessment			X (day 21)			X		
AE and SAE monitoring	X	X	X	X	X	X	X	
Long-term Follow-up[l]							X	X

a Vital signs should be performed after patient is sitting for 3-5 minutes. If during dosing period, vital signs should be taken immediately prior to infusion, immediately after completion of infusion, and 30, 60, and 240 minutes post-infusion

b All patients will have a 12-lead ECG performed at the screening visit and pre-treatment visit, as well as Day 28. Because PK evaluations will only be performed during Cycles 1 & 3, additional ECGs during the minutes/hours following SL-401 infusions are not required for Cycles 2, 4-6.

c To be collected prior to SL-401 infusion if during dosing period. Hematology includes WBC count with differential, RBC count, hematocrit, hemoglobin and platelet count

d To be collected prior to SL-401 infusion if during dosing period. Electrolytes include sodium, potassium, bicarbonate, chloride, BUN, creatinine and glucose.

e To be collected prior to SL-401 infusion if during dosing period. Serum albumin may be a component of the chemistry panel (f). See protocol for administration of albumin if serum albumin decreases to <3.0 g/dL during treatment days or in the immediate post-treatment period.

f To be collected prior to SL-401 infusion if during dosing period. Serum chemistry includes electrolytes (see above, footnote d) and the following: ALT, albumin, ALP, AST, bilirubin (total, direct, and indirect), calcium, creatine phosphokinase (CPK), magnesium, LDH, phosphate, total protein, and uric acid.

g To be collected prior to SL-401 infusion if during dosing period. Urinalysis includes appearance, color, pH, specific gravity, ketones, leukocytes, protein, glucose, bilirubin, urobilinogen, and occult blood.

h Morphology and differential WBC/blast count on aspirate. Baseline must be performed within 14 days prior to the first administration of SL-401. Subsequent bone marrow aspirates (+ biopsy) will be performed at the end of Cycles 2, 4 and 6 and every 3 months (± 1 month) thereafter through 12 months after the start of SL-401 and at the Investigator's discretion thereafter. If the end-of- Cycle 2/4/6 bone marrow aspirate (+ biopsy) is empty, hypocellular, or inadequate, a bone marrow examination should be repeated within 14 (± 7) days to document response. All bone marrow samples collected on study will be assessed centrally. Bone marrow (2 mL aliquot) will be prioritized for assessment of MRD at the [redacted]. Additional bone marrow (~5 mL) is to be aliquoted for tumor assessment and for translational evaluation at MD Anderson Cancer center. CD123 is to be assessed in all bone marrow samples (and, if applicable, tissue samples) by flow cytometry (i.e., CD123 should be added to the panel of markers assessed by flow cytometry of bone marrow aspirates) and immunohistochemistry and the results recorded and captured in the eCRF.

i Refer to Section 7.5.2 – Premedication and Administration.

j Following treatment with premedication, SL-401 will be administered as a 15-minute infusion for the first 5 consecutive days of a 28-day cycle. Individual SL-401 infusions may be delayed to allow for toxicity resolution, all infusions should be completed within 10 days. Patient must be monitored for 4 hours post infusion. (The administration of additional cycles of SL-401 (beyond 6 cycles) may be considered in the setting of MRD eradication, sustained AML remission, or other evidence of sustained anti-leukemia benefit in the opinion of the Investigator. Administration of SL-401 beyond 6 cycles must be discussed with the Medical Monitor at which time the individual patient's potential risk/benefit of further treatment will be assessed).

k Plasma samples (6 mL each) will be collected immediately prior to the start of the infusion of SL-401, immediately after end of infusion (time recorded), then 15, 30, 45, 60, 90, 120, 180, and 240 minutes after the start of the infusion during infusions 1 (i.e., Day 1) and 5 (i.e., Day 5) of Cycles 1 & 3. PK evaluation will not be performed for Cycles 2, 4-6.

l After the follow-up visit, patients will then be followed every 90 days for survival status for at least one year following the start of SL-401. The survival follow-up may be by telephone contact. Patients who undergo SCT will be followed for the occurrence of VOD as part of long-term follow-up. A blood sample for immunogenicity studies will be collected at least 16 weeks to up to 20 weeks after the last SL-401 dose.

m For patients who live a considerable distance from the Study Center, for whom weekly travel to the Study Center is not feasible, the Day 8, 15, and 21 (±3 days) laboratory assessments (blood and urine) may be submitted to a local laboratory, although the results must be evaluated by the Study team. Concomitant medication and AE monitoring on these days will take place via telephone contact. In situations where SL-401 is not administered over 5 consecutive days (i.e., treatment delays because of AEs or other factors) and an infusion is to be administered on Day 8, all pre-infusion/infusion evaluations should be conducted, in addition to concomitant medications recorded. Upon completion of the 5th Cycle 1 infusion (Day 8, 9 or 10), the blood/urine/concomitant medication and AE assessments should be planned for Day 15, as indicated in the table.

n The end-of-the cycle evaluations (Day 28 or thereafter) may also serve as the pre-infusion evaluations for the subsequent cycles; these do not need to be duplicated on successive days unless there is an abnormality or other clinically relevant reason for repeat evaluation.

8. RESPONSE CRITERIA

8.1. Tumor Response Criteria for Patients with AML

Response	Location	Criteria
Complete Remission (CR)	Marrow	• Normalization of blast percentage (≤5%) • No detectable Auer rods
	Peripheral Blood	• Normalization neutrophil count (≥1,000/μL) and platelet count (≥100,000/μL) • Absence of leukemic blasts
	Extramedullary	• No extramedullary disease (CNS or soft tissue)
CR with incomplete blood count recovery (CRi)	Marrow	• Normalization of blast percentage (≤5%)
	Peripheral Blood	• Incomplete recovery of neutrophil and/or platelet count • Absence of leukemic blasts
	Extramedullary	• No extramedullary disease (CNS or soft tissue)
Partial Remission (PR)	Marrow	• Decrease by ≥50% in blast percentage to 5 - 25% or to ≤5% with Auer rods present
	Peripheral Blood	• Normalization neutrophil count (≥1,000/μL) and platelet count (≥100,000/μL)
Stable Disease (SD)		• Failure to achieve at least a PR, but no evidence of progression for at least 8 weeks
Relapse after CR/CRi	Marrow	• Blast percentage >5% (if no peripheral blasts, then confirmation aspirate required ≥1 week later)
Relapse after PR	Marrow	• Blast percentage ≥25% (if no peripheral blasts, then confirmation aspirate required ≥1 week later)
Progressive Disease (PD)	Marrow	• ≥50% increase in blasts from baseline
	Peripheral Blood	One or more of the following: • ≥50% decrease from peak remission levels in platelets or granulocytes; • Reduction in hemoglobin concentration by at least 2 g/dL; • Transfusion dependence