Statistical Analysis Plan

Sponsor:	Rafael Pharmaceuticals, Inc.
Protocol Title:	Phase III Multicenter Open-Label Randomized Trial to Evaluate Efficacy and Safety of CPI-613® (devimistat) in Combination with High Dose Cytarabine and Mitoxantrone (CHAM) Compared to High Dose Cytarabine and Mitoxantrone (HAM) therapy and control sub-groups: combination of Mitoxantrone, Etoposide and Cytarabine (MEC) and combination of Fludarabine, Cytarabine, and Filgrastim (FLAG) in Older Patients (≥ 50 years) with Relapsed/Refractory Acute Myeloid Leukemia (AML)
Study Code:	AML003

The signatures on this form indicate approval of this document.

Table of Contents

1. List of Abbreviations and Definition of Terms

AE	Adverse Events
ALP	Alkaline Phosphatase
ALT	Alanine Amino Transferase
AML	Acute Myeloid Leukemia
ANC	Absolute Neutrophil Count
AST	Aspartate Amino Transferase
AUC	Area Under the Curve
BSA	Body Surface Area
BILI	Total Bilirubin
CEBPα	CCAAT Enhancer Binding Protein A
CFR	Code of Federal Regulations
CHAM	CPI-613 in Combination with High Dose Cytarabine and Mitoxantrone
CL	Clearance
C_{max}	The Maximum (Peak) Plasma Concentration
C_{min}	The Minimum (Trough) Plasma Concentration
CMH	Cochran-Mantel-Haenszel
CR	Complete Remission
CRc	Cytogenetic Complete Response
CRh	CR (Complete Remission) with Partial Hematologic Recovery
CRi	CR (Complete Remission) with Incomplete Recovery
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
FLAG	Fludarabine, Cytarabine, and Filgrastim
Flt3	FMS-Like Tyrosine Kinase 3
HAM	High Dose Cytarabine and Mitoxantrone
HAMA or HM	High Dose Cytarabine + Mitoxantrone + Asparaginase
HiDAC	High Dose Cytarabine
HMA	High Dose Cytarabine + Mitoxantrone + Asparaginase
HR	Hazard Ratio
IDH1/2	Isocitrate Dehydrogenase 1 and 2
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IHC	Immunohistochemistry
ITT	Intent-To-Treat
IxRS	Interactive Voice/Web Response Systems
KGDH	Alpha-Ketoglutarate Dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
MEC	Mitoxantrone, Etoposide and Cytarabine

Mito	Mitoxantrone
NCA	Non-Compartmental Analysis
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NPM	Nucleophosmin
OS	Overall Survival
PD	Pharmacodynamic
PDH	Pyruvate Dehydrogenase
PFS	Progression Free Survival
PK	Pharmacokinetic
PP	Per Protocol
PRO	Patient-Reported Outcome
PS	Performance Score
QLQ	Quality of Life Questionnaire
QTc	Corrected QT Interval
RNA	Ribonucleic Acid
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SGOT	Serum Glutamic Oxaloacetic Transaminase
SGPT	Serum Glutamic Pyruvic Transaminase
$T_{1/2}$	Elimination half-life
T_{max}	Time to Reach the Maximum Plasma Concentration
US	United States
Vd	Volume of Distribution
WBC	White Blood Cells

2. Introduction

This Statistical Analysis Plan was written for the clinical trial Rafael AML003. The ICH guidelines E3 "Structure and Content of Clinical Study Reports" and E9 "Statistical Principles for Clinical Trials" were used as a guide to the writing of the plan.

3. Study Design and Objectives

3.1 Study Objectives

3.1.1 Primary Objective

The primary objective is:

- To determine efficacy of CHAM in terms of Complete Remission (CR) and compare with HAM [control]. CR will be determined as per standard response criteria for AML (Döhner et al. 2017).

3.1.2 Secondary Objectives

The main secondary objectives are:

- To determine efficacy of CHAM in terms of 1. Overall Survival (OS) and 2. Complete Remission + Complete Remission with Partial Hematologic Recovery (CR+CRh) as the two key secondary objectives to compare with HAM [control]. OS and CR+CRh will be determined as per standard response criteria for AML (Döhner et al. 2017).
- Safety: The assessment of safety will be based mainly on the frequency of adverse events based on the Common Terminology Criteria for Adverse Events (CTCAE [v5.0]) grade. Adverse events will be coded according to MedDRA version 21.0. The safety outcomes will include the occurrence of at least one serious adverse event, of at least one grade 3/4 adverse event, and of at least one adverse event requiring the discontinuation of study treatment. 30-day mortality and 60-day mortality will be evaluated as part of the safety analysis. QTc intervals will be also evaluated as part of the safety analysis.

Other secondary objectives:

- Pharmacokinetic (PK): The goal of non-compartmental (NCA) pharmacokinetic analysis will be to evaluate Cmax, Cmin, AUC, T1/2, Tmax, CL and Vd for both CPI-613 and its metabolites CPI-2850 and CPI-1810. NCA PK analysis is out of the scope of this SAP and will not be performed by IDDI.

  Definitions of Pharmacokinetic Parameters:

  - Cmax: Maximum (peak) plasma drug concentration
  - Cmin: Minimum (trough) plasma drug concentration
  - AUC: Area under the plasma concentration-time curve
  - T1/2: Elimination half-life
  - Tmax: Time to reach maximum (peak) plasma concentration following drug administration
  - CL: Apparent total body clearance
  - Vd: Apparent volume of distribution

- Patient-Reported Outcome (PRO) by EORTC QLQ-C30: Administered at enrollment, at time of recovery marrow, at completion of consolidation therapy (when applicable), every 3 months following completion of consolidation therapy and at time of study discontinuation (when possible).

3.1.3 Exploratory Objectives

Exploratory objectives are:

- Efficacy and safety analyses per gene mutations (e.g. FLT3, IDH1/2, TP53, CEBPα, NPM1, etc.)
- PK/PD analyses for dose/exposure-response on efficacy and safety
- Gene expression analysis by RNA sequencing for baseline bone marrow samples to validate previously described response signature from the phase I study (CCCWFU 22112)
- Slides from baseline and time of progression biopsies will be used for IHC staining for PDKs, PDH, KGDH, SOD2 and CD79a

3.2 Study Design

This is a prospective, **multicenter**, **open label**, **randomized Phase III** study of **CHAM** compared to **HAM** and control sub-groups: combination of Mitoxantrone, Etoposide and Cytarabine (MEC) and combination of Fludarabine, Cytarabine, and Filgrastim (FLAG) in older patients (≥ 50 years) with relapsed/refractory AML.

There will be two study arms:

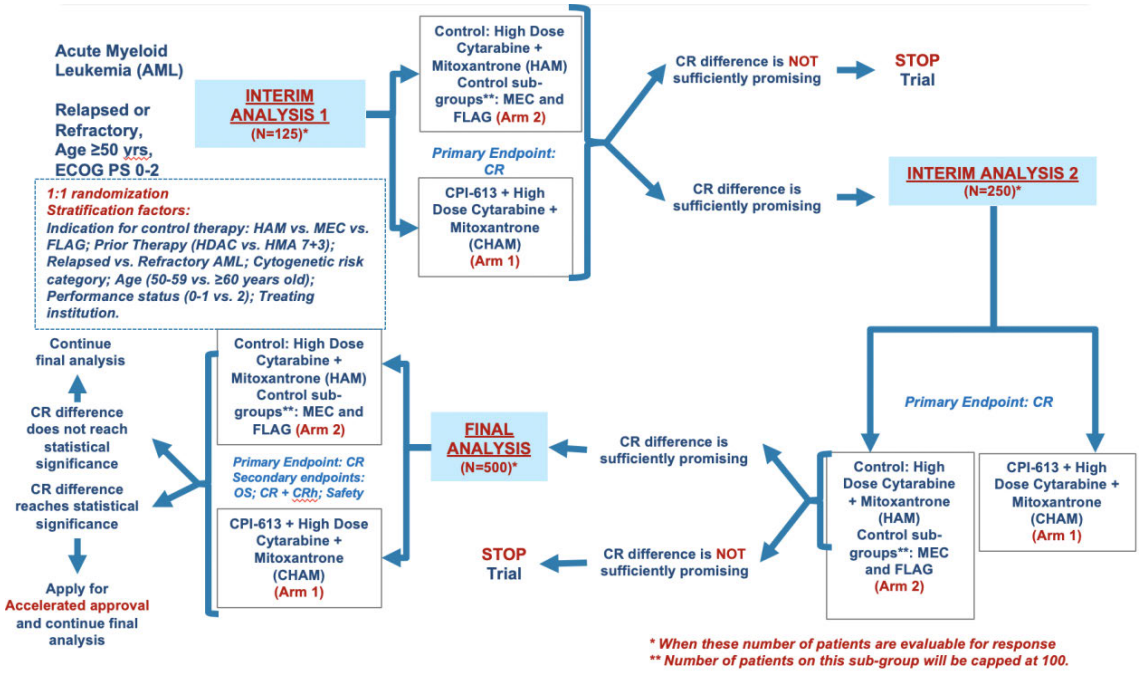
- Arm 1: CHAM (i.e. CPI-613 + High Dose Cytarabine and Mitoxantrone)
- Arm 2: HAM (i.e. High Dose Cytarabine and Mitoxantrone)
  - **Control sub-group 1:** Mitoxantrone, Etoposide and Cytarabine (MEC)
  - **Control sub-group 2:** Fludarabine, Cytarabine, and Filgrastim (FLAG)

Subjects will be randomized in 1:1 allocation ratio to treatment and control groups, respectively, in an open-label manner using a dynamic minimization procedure (additional details in section 4.5). The randomization number will be provided by IxRS. The randomization will be stratified by:

- Indication for control therapy: HAM vs. MEC vs. FLAG
- Prior therapy: HiDAC based vs. hypomethylator (azacytidine or decitabine, HMA) vs. 7+3 (7 days of cytarabine + anthracycline on the first 3 days)
- Relapsed vs. refractory AML
- Cytogenetic risk category (historical data might be used to define risk factors)
- Age (50-69 vs. ≥ 70 years old)
- Performance status (0-1 vs. 2)
- Treating Institution

The strata for control therapies of MEC and FLAG will be capped at 100 each. The overall study design is shown in Figure 1.

Figure 1 Rafael AML003 Schema



3.3 Sample Size Justification

The sample size calculation is based on an improvement in complete response rate (CR) from 26% in the control arm to 39% in the experimental arm (a 13% absolute increase, or a 50% relative increase), based on clinical efficacy data from Phase I study CCCWFU 22112 in AML patients. For a power of 80%, 500 patients need to be evaluated for response. This number allows for two interim analyses to be performed as detailed in the interim analysis plan.

The study is also powered to detect a clinically meaningful difference in overall survival (OS), with an expected median overall survival equal to 5.2 months in the control arm vs. 6.9 months in the experimental arm, i.e. a hazard ratio equal to 0.75 assuming exponential survival distributions. For a power of 80%, 394 events need to be observed. This number allows for one interim analysis but is merely added to safeguard the type I error as no interim analysis for OS is planned (as detailed in the interim analysis plan). Assuming an accrual rate of 15 patients per month and a common dropout rate of 10% at 3 years, a sample size of 500 patients will provide a power of 80% for the OS analysis 36 months after the first randomization. Note that all efficacy secondary endpoints will be analyzed at the final analysis after 394 deaths are observed.

Finally, the study has more than 86% power to detect a clinically meaningful difference in CR+CRh rate (complete remission with or without complete hematologic recovery) at the final analysis, with an expected rate equal to 33% in the control arm vs. 47% in the experimental arm. Similarly, as for OS, this sample size allows for 1 interim analysis although there is no intent to analyze CR+CRh at interim.

4. General Analysis Definitions

Data will be analyzed using SAS (Version 9.4 or higher).

Descriptive statistics will be tabulated as follows:

- Categorical data will be summarized in contingency tables presenting frequencies and percentages.
- Continuous data will be summarized using number of non-missing values (n), mean, standard deviation, median, minimum and maximum values.

The baseline tables will be created by treatment arm and overall. Tables containing post-randomization information will be created by treatment arm, unless otherwise specified.

Listings with individual values will be provided for all data presented in the tables, where noted. The listings will be presented by identifiers, such as treatment group, center and patient number.

One-sided tests will be used at a significance level equal to 0.025. Two-sided Exact Clopper Pearson confidence intervals will be computed for a coverage of 0.95.

Time to event outcomes ("survival times") will be described by treatment arm using the Kaplan-Meier method. Subjects who have not had the event of interest at the time of the analysis will be censored at the time of the last follow up. Summary statistics will be provided by treatment arm in terms of the number of events, number of censored observations, median and 95% confidence interval, and survival probabilities at specific time points (such as 1 year, 2 years, etc.). Survival curves will be plotted by treatment arm and compared with a log-rank test stratified by performance status, age and refractory versus relapsed disease. A stratified Cox regression model will be used to estimate the hazard ratio and its 95% CI, as well as to adjust the comparison for baseline covariates.

Binary outcomes will be described by proportions by treatment arm and compared with a Cochran-Mantel-Haenszel (CMH) test stratified by performance status, age and refractory versus relapsed disease. A logistic regression model will be used to adjust the comparison for baseline covariates.

4.1 Planned Analyses

4.1.1 First Interim Analysis

The first interim analysis will be performed when 125 patients are evaluable for response (had time for count recovery and bone marrow biopsy was obtained to assess remission status). The only endpoint analyzed at this interim analysis will be CR. The significance level to be used for this interim analysis will be determined using an O'Brien-Fleming type Lan-DeMets boundary for efficacy, and a Pocock type Lan-DeMets boundary for futility. Assuming an information fraction of 33% for CR at this interim analysis, the significance level for efficacy will be 0.0001 and it will be reached if the absolute difference in CR is larger than 26.7%. The significance level for futility is 0.48 and it will be reached if the absolute difference in CR is smaller than 0.5%. If, based on this interim analysis, the CR difference crosses the futility boundary, consideration will be given to stopping the trial; otherwise, the trial will proceed.

4.1.2 Second Interim Analysis

The second interim analysis will be performed when 250 patients are evaluable (had time for count recovery and bone marrow biopsy was obtained to assess remission status) for response. The only endpoint analyzed at this interim analysis will be CR. The significance level to be used for this interim analysis will be

determined using an O'Brien-Fleming type Lan-DeMets boundary for efficacy, and a Pocock type Lan-DeMets boundary for futility. Assuming an information fraction of 67% for CR at this interim analysis, the significance level for efficacy will be 0.006 and it will be reached if the absolute difference in CR is larger than 12.8%. The significance level for futility is 0.21 and it will be reached if the absolute difference in CR is smaller than 4.9%. If, based on this interim analysis, the CR crosses the futility boundary, consideration will be given to stopping the trial; otherwise, the trial will proceed.

4.1.3 Final Analysis

The final analysis will be performed when 500 patients are evaluable (had time for count recovery and bone marrow biopsy was obtained to assess remission status) for response. At that time, the primary endpoint of CR will be analyzed first, with a significance level determined using an O'Brien-Fleming type Lan-DeMets boundary for efficacy equal to 0.023, which it will be reached if the absolute difference in CR is larger than 8.3%. In case efficacy is declared for CR, it will be considered to file for accelerated approval based on CR while patients are further followed to collect OS data.

4.2 Definition of Populations

4.2.1 Intention-To-Treat (ITT) Population

The intent-to-treat (ITT) population (defined as all randomized patients) will be used for all analyses of efficacy and baseline characteristics. Patients will be analyzed according to their randomly assigned treatment regardless of treatment actually received.

4.2.2 Per-Protocol (PP) Population

The per-protocol (PP) population consists of randomized subjects who do not have any major protocol violation and received at least one dose of study treatment. The PP population will be used as a sensitivity analysis for efficacy endpoints.

4.2.3 Safety (SAF) Population

The safety population (SAF) includes all patients who receive at least one dose of study treatment. Patients will be analyzed according to treatment actually received. The SAF population will be used for safety analyses.

4.3 Subgroup Definitions

Subgroup analyses will be carried out with a descriptive intent. Treatment effects will be estimated and tested for important factors, including the factors used in the treatment allocation procedure (performance status, age and refractory versus relapsed disease), and displayed as forest plots. Interaction tests will be carried between treatment and each of these factors.

4.4 Multiple Comparison/Multiplicity

The randomized treatment arms will first be compared in terms of CR, using a Hochberg procedure to evaluate CR separately in the add-on ITT subpopulation (CHAM vs. HAM) and in the full ITT population. Then, conditional on the comparison of CR reaching statistical significance in either or both populations, the randomized treatment arms will be compared in terms of the two key secondary endpoints, OS

and CR+CRh, evaluated in the full ITT population. A Hochberg procedure will again be used to adjust the significance level to allow for multiple comparisons.

The Hochberg testing procedure will proceed as follows for the two populations (full ITT population and CHAM vs. HAM subpopulation): Let p1 and p2 be the p-values of the two populations. The two p-values will be arranged such that p1 ≤ p2. If p2 ≤ 0.023, then both key secondary endpoints will be declared significant. If p2 > 0.023 and p1 ≤ 0.0115, then the significance of the endpoint corresponding to p1 will be claimed.

The Hochberg testing procedure will proceed as follows (if applicable) for the two key secondary endpoints (OS and CR+CRh): Let p1 and p2 be the p-values of the two key secondary endpoints in the full ITT population. The two p-values will be arranged such that p1 ≤ p2.

- If statistical significance was reached in both populations and if p2 ≤ 0.023, then both endpoints will be declared significant. If p2 > 0.023 and p1 ≤ 0.0115, then the significance of the endpoint corresponding to p1 will be claimed.
- If statistical significance was reached in only one population and if p2 ≤ 0.0115, then both endpoints will be declared significant. If p2 > 0.0115 and p1 ≤ 0.00575, then the significance of the endpoint corresponding to p1 will be claimed.

4.5 Treatment Assignment and Treatment Arms

Eligible patients will be randomized to receive either

- Arm 1: CHAM (i.e. CPI-613 + High Dose Cytarabine and Mitoxantrone)
- Arm 2: HAM (i.e. High Dose Cytarabine and Mitoxantrone)
  - **Control sub-group 1:** Mitoxantrone, Etoposide and Cytarabine (MEC)
  - **Control sub-group 2:** Fludarabine, Cytarabine, and Filgrastim (FLAG)

Subjects will be randomized in 1:1 ratio to the experimental treatment or control, using a dynamic minimization procedure based on the methodology described by Pocock and Simon (Pocock and Simon, 1975). The minimization algorithm will use the variance range method to minimize overall imbalances between the treatment arms with respect to center and important prognostic factors (see below). A stochastic minimization will be used so that no treatment allocation is deterministic.

The stratification factors include:

- Indication for control therapy: HAM vs. MEC vs. FLAG
- Prior therapy: HiDAC based vs. hypomethylator (azacytidine or decitabine, HMA) vs. 7+3 (7 days of cytarabine + anthracycline on the first 3 days)
- Relapsed vs. refractory AML
- Cytogenetic risk category (historical data might be used to define risk factors)
- Age (50-69 vs. ≥ 70 years old)
- Performance status (0-1 vs. 2)
- Treating Institution

4.6 Calculated Variables

- Study day 1 is defined as the first date of drug intake. For those randomized who did not receive any study drugs, the date of randomization will be used.

- The baseline is defined as the last non-missing assessment done before start of treatment.
- Body Surface Area (m^2) = ([height (cm) × weight (kg)] / 3600)$^{0.5}$
- Months= days/30.4167

4.7 Methods to Be Used For Handling Missing Data

For the primary analyses, no imputations will be done, and data will be analyzed as observed, with patients without response assessment set to non-responder. Multiple-imputation will be used in secondary or sensitivity analyses, whenever necessary, to impute the missing efficacy endpoint data.

Missing Quality of Life (QOL) scales will be imputed as follows:

The missing items are imputed based on the average of other items that are answered by a patient for multi-item scales. None of the single item measures can be imputed using this method of imputation. In the event of missing full assessments, none of the subscales will be scored or imputed.

5. Study Patients

5.1 Disposition of Patients

The number of patients in each population will be tabulated by treatment arm and overall. Number screened and screen failed will be tabulated. Reason for screen failure will also be tabulated and listed.

The frequency of patients treated, of patients who discontinued the study treatment and of patients who terminated the study will be given for the ITT population. The primary reason for discontinuation of the study treatment and terminating the study will be summarized. The details of the patient's or investigator's decision will be included in a listing. The frequency of subjects on treatment, who completed treatment, and reason for discontinuation of treatment will be summarized by consolidation and maintenance therapy cycles.

5.2 Protocol Deviations

Protocol deviations classified as major will be determined before database lock. The major protocol deviations will be categorized by type and summarized by treatment arm, for the ITT population.

A listing of all major protocol deviations will be provided by treatment arm.

5.3 Inclusion and Exclusion Criteria

Enrolled patients who violate inclusion/exclusion entry criteria (described in Section 5 of protocol) will be summarized for those randomized.

A listing of all inclusion and exclusion criteria not met will be provided by treatment arm.

6. Demographic and other Baseline Characteristics

Descriptive statistics with respect to patient characteristics at baseline will be displayed for the ITT population. Stratification information will be tabulated by treatment arm.

The summary of demographic and baseline characteristics data will present the following variables:

- Age at Screening (years) and Age Category (50-55, 56-60, 61-65, 66-70, 71-75, 76-80, 81-85, 86-90, 90+)
- Race
- Ethnicity
- Sex
- ECOG Performance Status
- Height (cm)
- Weight (kg)
- BSA (m^2)

AML history including type, status, prior type of therapy, and time since diagnosis will be summarized by treatment arm and overall. Prior and concomitant medication will be summarized using WHODrug version 2018. Medical history and concomitant illness will be classified according to the terminology of CTCAE version 5.0 and coded using MedDRA version 21.0. Medical history will be tabulated by system organ class and preferred term for the ITT population.

7. Statistical Analysis

The efficacy analyses will be performed on the ITT population. In addition, Complete Remission (CR), Overall Survival (OS) and Complete Remission + Complete Remission with Partial Hematologic Recovery (CR+CRh) will be analyzed on the PP population.

7.1 Primary Analysis

7.1.1 Complete Remission (CR)

The treatment arm will be compared with control arm in terms of CR. The primary analysis will be a re-randomization test (Simon R., 1979) based on the CMH test-statistic, stratified by performance status, age and refractory versus relapsed disease. The re-randomization approach fixes all data except the treatment labels at their observed values, regenerates the randomization sequence using the minimization algorithm (Buyse M., 2000), and computes the test statistic corresponding to those reshuffled assignments. This process is repeated a large number of times, and a p-value is calculated as the proportion of re-randomized trials whose test statistic is at least as extreme as the observed one from the original assignments.

The trial design includes two interim analyses, and one final analysis. Both interim analyses are only performed for the primary endpoint CR rate with the intent to stop the trial if the difference in CR between arms is not sufficiently promising; there is no intention to stop the study early if efficacy is shown. The futility boundaries are conceived as non-binding boundaries allowing the DMC to decide independently at the timing of the interim analysis, taking all available data into account, whether the study should continue or stop.

7.2 Secondary Analyses

7.2.1 Overall Survival (OS) and Complete Remission + Complete Remission with Partial Hematologic Recovery (CR+CRh)

The primary analysis of these secondary efficacy endpoints will be a re-randomization test that calculates the p-value by re-randomizing and allocating patients to treatments. For OS, the re-randomization will use a stratified Cox proportional hazard test-statistic. For CR + CRh, the stratified CMH test-statistic will be used. The secondary endpoints will only be analyzed at the final analysis. Note that by design an interim analysis has been foreseen for OS, using an O'Brien-Fleming type Lan-DeMets boundary for efficacy.

When the number of required deaths is observed, all secondary endpoints will be analyzed at the time of final analysis. The number of required deaths for a final analysis of OS is 394.

7.2.2 Safety

The assessment of safety will be based mainly on the frequency of adverse events, which will be coded using the Medical Dictionary for Regulatory Activities (MedDRA Version 21.0) and will be graded according to the National Centre Institute Common Terminology Criteria for AEs (NCI-CTCAE criteria [v5.0]). Uncoded terms will be presented under "SOC uncoded", with their reported term.

Adverse events will be analyzed in terms of their type, incidence, severity and relationship to the study treatment.

A summary table will present, by treatment arm, the number and percentage of patients with at least one:

- AE
- Related AE
- AE leading to treatment discontinuation
- AE leading to dose interruption
- Grade 3/4 AE
- Serious AE
- AE leading to death

In addition, tabulations of the number of AE events (same categories as above) as well as severity of the events will be presented by system organ class and preferred term. Patients will only be counted once for each preferred term. In case a patient experienced the same event more than once, the worst severity will be presented.

Listings of all adverse events by treatment arm will be provided including the patient identifier, age, race, sex, verbatim, preferred term, severity, relationship to study treatment.

30-day mortality and 60-day mortality will also be evaluated as a part of the safety analysis using Kaplan-Meier for overall survival. The number of deaths will be tabulated together with the primary cause of death. The details of the 'other cause' will be included in the listing.

Hematology [hemoglobin, leukocytes (WBC), neutrophil count (ANC), lymphocytes, platelets] and biochemistry [creatinine, AST/SGOT, ALT/SGPT, alkaline phosphatase (ALP), total bilirubin (BILI)] lab data will be presented in shift tables from baseline to worst post-baseline CTC grade, along with listings.

Number and cause of death will be summarized and listed.

QTc interval at pre-infusion and at the end of infusion will be summarized by visit and treatment arm separately for induction, consolidation and maintenance therapies.

7.2.3 Patient-Reported Outcomes Analysis

Changes over time in QLQ-C30 Global Health Status/Quality of Life (QOL) scores will be summarized and listed for all time points and by treatment group. Changes over time in QLQ-C30 Global Health Status/Quality of Life (QOL) scores will also be compared between treatment groups using a restricted maximum likelihood-based mixed model for repeated measures under the assumption of missing at random. The minimal clinical important difference (MID) will help to interpret any observed differences between treatment groups. The MID for between-group differences on the QLQ-C30 Global Health Scale/QoL is 5 points.

Sensitivity analysis will be performed on PP population after imputations are made as per section 4.7.

7.3 Exploratory Analyses

Bone marrow aspirate/biopsy samples are collected from patients at baseline and RNA sequencing analyses will performed to determine the positive and negative predictive value of the response signature. The response summary will be displayed per FLT3, IDH1/2, TP53, CEBPα, and NPM1 gene mutation.

8. References

1. Döhner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*. 2010 Jan 21; 115(3): 453-74.
2. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics.* 1975; 31: 103-5.
3. Simon R. Restricted randomization designs in clinical trials. *Biometrics.* 1979; 35(2): 503-12
4. Buyse M. Centralized treatment allocation in comparative clinical trials. *Applied Clinical Trials.* 2000; 9(6): 32-37