STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 1, Open-label, Dose-escalation, Safety and Biomarker Prediction Study of Alvocidib and Cytarabine/Daunorubicin (7+3) in Patients with Newly Diagnosed Acute Myeloid Leukemia (AML)
Protocol Number:	TPI-ALV-101
IND Number:	057729
Study Drug:	Alvocidib (test agent; formerly flavopiridol) in combination with Ara-C (cytarabine) and mitoxantrone
Phase of Development:	Phase 1
Sponsor:	Sumitomo Dainippon Pharma Oncology, Inc. 3900 North Traverse Mountain Blvd, Suite 100 Lehi, UT USA 84043
Date of SAP:	28 July 2019 (Version 1.0)

Table of Contents

1. Synopsis 4
2. Abbreviations 5
3. Definitions 7
4. Introduction 8
   4.1. Study Endpoints 8
   4.2. Statistical Objectives 9
5. Sample Size 10
6. Treatment Assignment 10
7. Statistical Methods 10
   7.1. Analysis Populations 10
   7.2. Missing Values 11
   7.3. Visit Windows 11
   7.4. Interim Analysis 11
   7.5. Patient Disposition 12
   7.6. Baseline Characteristics 12
   7.7. Study Drug Exposure 12
   7.8. Efficacy Endpoints 13
      7.8.1. Remission Rates 13
      7.8.2. Progression-Free Survival 13
      7.8.3. Overall Survival 14
      7.8.4. Relapse-Free Survival Following CR 15
      7.8.5. Complete Remission Duration 15
      7.8.6. ECOG Performance Status 16
   7.9. Biomarker 16
   7.10. Safety Endpoints 17
      7.10.1. 30- and 60-Day Mortality 17
      7.10.2. Adverse Events 17
      7.10.3. Classification of Tumor Lysis Syndrome 18
      7.10.4. Clinical Laboratory Parameters 18
      7.10.5. Neutrophil Recovery 19
      7.10.6. Platelet Recovery 19

7.10.7. Concomitant Medication Use 19
8. Adverse Event and Medication Coding Dictionaries 20
9. Summary Tables 20
10. References 23

List of Tables

Table 1. Study Endpoints 8
Table 2. Cairo-Bishop Definition of Laboratory Tumor Lysis Syndrome 18
Table 3. Planned Summary Tables 20

List of Appendices

Appendix 1. Criteria for Grading Severity of Clinical Tumor Lysis Syndrome 24
Appendix 2. Sample Table Template 25

1. Synopsis

Protocol Title:

A Phase 1, Open-label, Dose-escalation, Safety and Biomarker Prediction Study of Alvocidib and Cytarabine/Daunorubicin (7+3) in Patients with Newly Diagnosed Acute Myeloid Leukemia (AML)

Study Design:

Open label, multicenter, 3+3 cohort dose escalation of alvocidib for induction therapy combined with 7+3 cytarabine and daunorubicin, followed by optional reinduction with alvocidib combined with 5+2 cytarabine and daunorubicin and/or up to 4 cycles of consolidation therapy with high-dose cytarabine (HiDAC)

Study Populations:

Patients with newly diagnosed, previously untreated AML excluding acute promyelocytic leukemia or core-binding factor AML

Induction Therapy (alvocidib with 7+3):

- Alvocidib on Days 1-3
    - 20-mg/m^2 IV bolus followed by 30-mg/m^2 IV infusion over 4 hours, or
    - 30-mg/m^2 IV bolus followed by 40-mg/m^2 IV infusion over 4 hours, or
    - 30-mg/m^2 IV bolus followed by 50-mg/m^2 IV infusion over 4 hours, or
    - 30-mg/m^2 IV bolus followed by 60-mg/m^2 IV infusion over 4 hours
- Cytarabine 100 mg/m^2/day continuous IV infusion on Days 5-11
- Daunorubicin 60-mg/m^2 IV bolus on Days 5-7

Optional Reinduction Therapy (alvocidib with 5+2):

- Alvocidib on Days 1-3 at the same dose used for induction
- Cytarabine 100 mg/m^2/day continuous IV infusion on Days 5-9
- Daunorubicin 60-mg/m^2 IV bolus on Days 5-6

Optional Consolidation Therapy (HiDAC up to 4 cycles):

- Cytarabine 3 g/m^2 IV q12h on Days 1, 3 and 5 for a total of 6 doses
- A reduced dose of 1.5 g/m^2 IV q12h was permitted for patients ≥60 years old

Sample Size:

A sufficient number of patients to complete a standard 3+3 cohort dose escalation for up to 4 dose levels plus up to 20 additional patients enrolled at the maximum tolerated dose (MTD) of alvocidib

Analysis Populations:

Intent-to-treat (primary for efficacy), safety, per-protocol

Response Assessments:

Local review of bone marrow and blood test results by the investigator at the study center where patient is enrolled to classify response to treatment following to European LeukemiaNet (ELN) criteria[1]

Missing Values:

Missing values will not be replaced by imputed values. Patients who do not supply bone marrow for response assessment are counted among those not achieving remission. Missing time-to-event endpoints will be censored using standard rules detailed below.

Interim Analysis:

No interim analyses. Ongoing review of clinical data performed as required for dose escalation/de-escalation decisions

Efficacy Endpoints:

CR rate (primary endpoint), combined CR rate, combined remission rate, stem cell transplant rate, overall survival, progression-free survival, relapse-free survival, CR duration

Safety Endpoints:

30- and 60-day survival, adverse events, time to neutrophil recovery, time to platelet recovery, prevalence of tumor lysis syndrome, clinical laboratory parameters, vital signs

Statistical Methods:

Descriptive statistics (including 95% confidence intervals) within and across alvocidib dose levels

2. Abbreviations

Standard abbreviations (e.g., AML = acute myeloid leukemia and mg = milligrams) are omitted intentionally.

AE	Adverse Event
AML	Acute Myeloid Leukemia
ANC	Absolute Neutrophil Count
ATC	Anatomical and Therapeutic Class
AUC	Area Under Curve
BM	Bone Marrow
BSA	Body Surface Area
CI	Confidence Interval
CM	Ara-C (Cytarabine) and Mitoxantrone
CMH	Cochran-Mantel-Haenszel (test)
CR	Complete Remission

CRi	Complete Remission with Incomplete Blood Count Recovery
CTCAE	Common Terminology Criteria for Adverse Events
DDE	Drug Dictionary Enhanced
DLT	Dose-Limiting Toxicity
DSMB	(Independent) Data Safety Monitoring Board
ECOG	Eastern Cooperative Oncology Group (performance status)
ELN	European LeukemiaNet
HiDAC	High-Dose Cytarabine
HR	Hazards Ratio
ITT	Intent-to-Treat (patient population)
IWG	International Working Group
KM	Kaplan-Meier (survivor function estimate)
LR	Logistic Regression
MedDRA	Medical Dictionary for Regulatory Activities
MRD–	Minimal Residual Disease Negative
MTD	Maximum Tolerated Dose
OS	Overall Survival
PBSC	Peripheral Blood Stem Cells
PFS	Progression-Free Survival
PH	Proportional Hazards (model)
PR	Partial Remission
PT	(MedDRA) Preferred Term
RFS	Relapse-Free Survival
ROC	Receiver Operating Characteristic (curve)
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SCT	Stem Cell Transplant/Transplantation
SMQ	Standardized MedDRA Query
SOC	(MedDRA) System Organ Class
TEAE	Treatment Emergent Adverse Event
TLS	Tumor Lysis Syndrome
WHO	World Health Organization

3. Definitions

Standard statistical terminology (e.g., type-1 error rate and significance level) is omitted intentionally. Efficacy endpoints are defined in Section 4.1 below (see Table 1).

Baseline Value: last measured value prior to the first dose of study drug

Combined CR: response assessment of CR or CRi

Combined Remission: response assessment of CR, CRi or PR

Cycle 1: period of time from the first dose of study drug until the start of study drug in cycle 2 or the start of alternative AML therapy other than supportive measures such as blood product transfusions (hence a patient whose first response assessment in cycle 1 is a CRi is considered to have achieved CR if neutrophils and platelets recover prior to initiation of cycle 2 or alternative AML therapy)

Combined CR Rate: percentage of patients who achieve a combined CR following treatment with alvocidib

Combined Remission Rate: percentage of patients who achieve a combined remission following treatment with alvocidib

CR Rate: percentage of patients who achieve a CR following treatment with alvocidib (see section 7.8.1)

Dose-Limiting Toxicity: Refer to the study protocol for the specific conditions and events that qualify as a dose-limiting toxicity (DLT)

Intent-to-Treat Population: all patients enrolled in the study regardless of whether they actually received study drug and regardless of whether evidence is found indicating they failed to meet study inclusion/exclusion criteria or had other protocol violations (see section 7.1)

Minimal Residual Disease Negative: Absence of gene fusion transcripts using standardized techniques such as multiparametric flow cytometry or next generation sequencing

Neutrophil Recovery: Absolute neutrophil count $\geq 1 \times 10^9$/L following the administration of study drug

Per-Protocol Population: The subset of safety patients who either (a) have at least one response assessment on study, or (b) die prior to their first scheduled response assessment; however, patients are excluded from the per-protocol population if they have an important protocol deviation, such as when evidence is found indicating they failed to meet any study inclusion/exclusion criterion (see section 7.1)

Platelet Recovery: Platelet count $\geq 100 \times 10^9$/L following the administration of study drug

Prestudy: Occurrence prior to the first administration of study drug

Safety Population: The subset of intent-to-treat patients who received at least one dose or partial dose of study drug (see section 7.1)

Shift Table: two-way frequency table pairing baseline value with the most extreme post-baseline result

Study Drug: alvocidib, Ara-C (cytarabine) and mitoxantrone

Treatment Failure: failing to achieve a CR, CRi or PR following induction therapy (see section 7.8.2)

4. Introduction

This Statistical Analysis Plan (SAP) details the planned methodology for summarizing data collected in the clinical study conducted under Tolero, Inc. protocol TPI-ALV-101. The current version of the study protocol is Amendment 6 dated 14-JUN-2019. Future protocol amendments (if any) will be reviewed to assess whether the changes necessitate modification of this SAP.

Since the study protocol is a companion document to this SAP, aspects in the protocol unrelated to statistical issues (e.g., patient eligibility criteria, descriptions of clinical materials, and criteria for defining treatment failure, response to treatment and disease relapse) are not repeated here.

4.1. Study Endpoints

Statistical analysis and/or summarization is planned for the following study endpoints. Complete details on how endpoints are calculated and compared across treatment arms are presented in the subsections below. A key to abbreviated terms appears in Section 2.

Table 1. Study Endpoints

Primary Efficacy Endpoint • CR Rate: percentage of patients who achieve a CR as assessed by the investigator following treatment with alvocidib
Secondary Efficacy Endpoints (not listed in order of importance) • CR_{MRD-} Rate: percentage of patients who achieve a CR and test negative for minimal residual disease (MRD) using standardized techniques such as multiparametric flow cytometry or next generation sequencing • Combined CR Rate: percentage of patients achieving a combined CR (CR or CRi) following treatment with alvocidib • Combined Remission Rate: percentage of patients achieving a combined remission (CR, CRi or PR) following treatment with alvocidib • Overall Survival: time from the date of first dose (Day 1) until death from any cause • Progression-Free Survival: time from the date of first dose (Day 1) until (a) treatment failure, (b) relapse after combined remission (CR, CRi or PR), or (c) death from any cause, whichever occurs first • Relapse-Free Survival Following CR (defined only for patients who achieve a CR during the study): time from the date of CR (Day 1) until (a) relapse after CR, or (b) death from any cause, whichever occurs first

Table 1. Study Endpoints

• Relapse-Free Survival Following Combined CR (defined only for patients who achieve a combined CR during the study): time from the earliest date of combined CR (Day 1) until (a) relapse after combined CR, or (b) death from any cause, whichever occurs first • Relapse-Free Survival Following Combined Remission (defined only for patients who achieve a combined remission during the study): time from the earliest date of combined remission (Day 1) until (a) relapse after combined remission, or (b) death from any cause, whichever occurs first • CR Duration (defined only for patients who achieve a CR during the study): time from the date of CR (Day 1) until relapse after CR • Combined CR Duration (defined only for patients who achieve a combined CR during the study): time from the earliest date of combined CR (Day 1) until relapse after combined CR • Combined Remission Duration (defined only for patients who achieve a combined remission during the study): time from the earliest date of combined remission (Day 1) until relapse after combined remission • Stem Cell Transplant Rate: percentage of patients proceeding directly to stem cell transplantation within 90 days after their last dose of study drug
Safety Endpoints (not listed in order of importance) • 30- and 60-Day Mortality • Adverse Event Incidence Rates • Time to neutrophil recovery ≥1x10^9/L • Time to platelet recovery ≥100x10^9/L • Presence and Severity of Tumor Lysis Syndrome • Clinical Laboratory Parameters • Vital Signs
Other Endpoints • Concomitant Medication Usage Rates

4.2. Statistical Objectives

The statistical objectives are threefold:

1. To define a maximum tolerated dose of alvocidib administered with cytarabine and daunorubicin in newly diagnosed AML patients
2. To describe the safety and efficacy of alvocidib administered with cytarabine and daunorubicin in newly diagnosed AML patients

3. To model the relationship between baseline MCL-1 dependence and response to treatment with alvocidib administered with cytarabine and daunorubicin

5. Sample Size

The sample size for this study is not predetermined. The number of patients enrolled depends on the occurrence and pattern of DLTs.

6. Treatment Assignment

Cohorts of 3 patients are assigned a dose of alvocidib according to standard 3+3 dose escalation/de-escalation rules. All patients receive the same, standard doses of cytarabine and daunorubicin.

7. Statistical Methods

Analyses will be performed using SAS statistical software. Default estimation methods in version 9.4 of SAS are always used unless an alternative is specified below. Data summaries will include the mean, standard deviation, median, minimum and maximum values for continuous data; the median, 25th and 75th percentiles, minimum and maximum values for time-to-event endpoints; and the number and percentage of patients in each category for categorical data. Pointwise 95% confidence intervals (CI) will also be estimated for the mean (continuous data), median (time-to-event endpoints) or percentage of patients (categorical data).

In general, data values will be obtained directly from the clinical database (e.g., birth date, treatment dates, AML relapse date and death date). Derivations for certain values not collected explicitly in the database are explained in relevant sections below (e.g., age and treatment failure date, sections 7.6 and 7.8.2, respectively). In some instances, it is beyond the scope of this SAP to explain schemes for patient classifications and data derivations (e.g., classifying patients in cytogenetic risk group according to published criteria). These details will be expounded in the clinical study report.

Durations of time expressed in months (e.g., times-to-event endpoints) will be reported in proportional “months” according to the formula

$$\text{months} = \frac{12}{365.25} \times \text{days}$$

Baseline value of a characteristic is defined as the last measured value prior to the first dose of study drug.

7.1. Analysis Populations

The intent-to-treat (ITT) patient population includes all patients enrolled in the study regardless of whether they actually received study drug and regardless of whether evidence is found indicating they failed to meet study inclusion/exclusion criteria or had other protocol violations. When the ITT patient population is analyzed, patients are grouped according to their assigned alvocidib dose regardless of actual dose received. The ITT patient population is the analysis population for the primary analyses of efficacy endpoints.

The safety patient population is the subset of ITT patients who received at least one dose of study drug. When the safety patient population is analyzed, patients are grouped according to the actual alvocidib dose received. The safety patient population is the analysis population for all analyses of safety data.

The per-protocol patient population is the subset of safety patients who either (a) have at least one response assessment on study, or (b) die prior to their first scheduled response assessment. However, patients are excluded from the per-protocol patient population if they have an important protocol deviation, such as when evidence is found indicating they failed to meet any study inclusion/exclusion criterion. When the per-protocol patient population is analyzed, patients are grouped according to their assigned alvocidib dose regardless of actual dose received. The per-protocol population is used for sensitivity (secondary, supportive) analyses of efficacy endpoints.

Each of the patient populations listed above may be further subdivided for certain analyses (e.g., by age, race or gender) to assess whether certain subpopulations benefited from treatment with alvocidib more than others or if certain subgroups experienced greater toxicity from exposure to alvocidib.

7.2. Missing Values

In general, missing values will not be replaced by imputed values.

- Patients who, for any reason, do not supply bone marrow for response assessment are counted among those not achieving remission
- Patients missing a time-to-event endpoint will have their value censored according to the rules set forth below (see sections 7.8.2 to 7.8.5)

Exceptions to the general rule of not replacing missing values are planned for the following items.

- The first day of the month will be used for a missing day component of a partial birthdate, and January will be substituted for a missing birth month (these substitutions are relevant only for calculating a patient's age)
- Partial start and end dates for adverse events will be replaced by calendar dates that maximize the duration of the adverse event (see section 7.10.2)

7.3. Visit Windows

Visit windows are not relevant for analyzing efficacy endpoints since these endpoints are either binary outcomes over the entire study period (e.g., CR rate) or measured on a continuous time scale (e.g., OS and PFS). Some safety assessments are summarized over time and within-cycle visit windows will be used (e.g., clinical laboratory parameters; see section 7.10.4).

7.4. Interim Analysis

No interim analyses are planned. However, ongoing review of clinical data was performed during patient accrual and follow-up as required for dose escalation/de-escalation decisions.

7.5. Patient Disposition

The numbers of patients enrolled (ITT population), dosed (safety population) and included in the per-protocol population will be summarized by frequency counts within and across alvocidib dose levels. Time on study will be summarized as a continuous variable (mean, standard deviation, median, minimum and maximum values). Reasons for discontinuation will be summarized as a categorical variable (number and percentage of patients in each category). Relevant data supporting these summaries will be listed by patient within alvocidib dose level.

7.6. Baseline Characteristics

Patient age in years on the date of the first dose of alvocidib will be calculated by dividing the number of days since birth (Day 1) by 365.25. If body surface area (BSA) expressed in m^2 is missing in the study database, then it will be calculated using the formula[6]

$$\text{BSA} = \sqrt{\frac{W \times H}{3600}}$$

where *W* is the patient's weight in kg and *H* is the patient's height in cm.

Summaries of baseline characteristics will include descriptive statistics appropriate for continuous, time-to-event and categorical variables within and across alvocidib dose levels (see section 7). Separate summaries will be produced for the ITT, safety and per-protocol patient populations.

Baseline characteristics include demographic data (e.g., age, sex, race, ethnicity, height, weight and BSA); AML history (e.g., clinical onset of AML [de novo AML, prior MDS or prior leukemogenic therapy], time since AML diagnosis, peripheral blood and bone marrow blast percentages, cytogenetic risk group, frontline AML therapy and response to frontline AML therapy [see section 6]); MCL-1 dependence; and other medical history (e.g., baseline ECOG performance status, hemoglobin, leukocyte and platelet counts, and blood chemistry parameters assessing liver and kidney function). Reported medical history terms will be mapped to Medical Dictionary for Regulatory Activities (MedDRA) preferred terminology and then summarized within and across alvocidib dose levels by incidence rates for each MedDRA primary system organ class (SOC) and preferred term (PT). Optionally, depending on the perceived relevance, prestudy medications will mapped to terminology in the World Health Organization (WHO) Drug Dictionary Enhanced (DDE) and then summarized within and across alvocidib dose levels by usage rates for each WHO DDE anatomical and therapeutic class (ATC) and preferred (i.e., standardized) drug name. Relevant data supporting these summaries will be listed by patient within alvocidib dose level.

7.7. Study Drug Exposure

Patients' total exposures to study drugs and average exposures per treatment cycle will be calculated and summarized within and across alvocidib dose levels by the mean, standard deviation, median, minimum and maximum values. Numbers of treatment cycles received will be summarized as a continuous variable (by the mean, standard

deviation, median, minimum and maximum values) and as a categorical variable (by the number and percentage of patients in each category). Reasons for noncompliance with protocol-specified dosing schedules will be summarized as a categorical variable. Relevant data supporting these summaries (including actual infusion durations and expected doses based on BSA calculated from a patient's latest reported height and weight) will be listed by patient within alvocidib dose level.

7.8. Efficacy Endpoints

Relevant data supporting efficacy analyses will be listed by patient within alvocidib dose level.

7.8.1. Remission Rates

The primary efficacy endpoint is CR rate as assessed by the investigator following treatment with alvocidib. Patients who, for any reason, do not supply bone marrow for response assessment are counted among those not achieving remission.

Summary statistics will include the number and percentage of patients rated as not achieving CR due to not providing a bone marrow biopsy/aspirate, proportion of patients who achieve a CR and exact 95% CI estimated without stratification. Between–treatment-arm statistics will include the Mantel-Haenszel odds ratio and 95% CI estimated with stratification.

A standard logistic regression (LR) model will be estimated relating the probability of achieving CR as a function of baseline MCL-1 dependence. Results will be presented graphically as follows. Patients will be divided into groups according to which range of values their baseline MCL-1 dependence falls:

$$(0\%\text{-}\leq d_1),\ (d_1\text{-}\leq d_2),\ (d_2\text{-}\leq d_3),\ \ldots,\ (d_8\text{-}\leq d_9),\ (d_9\text{-}100\%)$$

where $d_1,\ldots,d_9$ are deciles for baseline MCL-1 dependence regardless of treatment arm. Empirical CR rates will be calculated for patients within each range and plotted alongside CR probabilities estimated from the LR model.

Other remission rates (e.g., combined CR rate and combined remission rate) will be summarized similar to CR rate.

7.8.2. Progression-Free Survival

PFS time is defined for all patients and is measured from the date of first dose (Day 1) until (a) treatment failure, (b) relapse after combined remission, or (c) death from any cause, whichever occurs first.

Treatment failure is defined as failing to achieve a combined remission following induction therapy. The date of treatment failure is defined by the earliest bone marrow collection following the first induction attempt regardless of whether the patient received a second induction attempt (unless the second induction resulted in a combined remission in which case "treatment failure" does not apply).

Additional rules for defining PFS time are:

- A patient found to have relapsed after missing one or more scheduled response assessments will have his PFS time end on the scheduled date for the first missed response assessment following the last confirmation of combined remission
- A patient who misses the first scheduled response assessment and who never achieves combined remission will have his PFS time end on the scheduled date for the first response assessment (unless the patient died before this date)

Patients enrolled subsequent to protocol amendment 6 were scheduled to be followed for 2 years following the end-of-study visit, whereas patients enrolled prior to protocol amendment 6 were followed for only 30 days following the end-of-study visit unless they reconsented to the amendment-6, 2-year follow-up period. The definition of PFS time and the following rules for censoring PFS time apply to all patients regardless of follow-up duration.

Rules for censoring PFS time are:

- A patient with PFS continuing as of the date he is lost to follow-up will have his PFS time censored on the date of his last response assessment
- A patient with PFS continuing as of the database lock will have his PFS time censored on the date of his last response assessment

Summary statistics will include the number and percentage of censored PFS times, quartile PFS times, 95% CIs for the quartile PFS times, and KM "survival" percentages[10] with 95% CIs at select milestone times (e.g., 6, 12 and 18 months).

7.8.3. Overall Survival

OS time is defined for all patients and is measured from the date of first dose (Day 1) until death from any cause. Patients enrolled subsequent to protocol amendment 6 were scheduled to be followed for 2 years following the end-of-study visit, whereas patients enrolled prior to protocol amendment 6 were followed for only 30 days following the end-of-study visit unless they reconsented to the amendment-6, 2-year follow-up period. The definition of OS time and the following rules for censoring OS time apply to all patients regardless of follow-up duration.

Rules for censoring OS time are:

- A patient with OS continuing as of the date he is lost to follow-up will have his OS time censored on the date of his last assessment of any type
- A patient with OS continuing as of the database lock will have his OS time censored on the date of his last assessment of any type

Summary statistics will include the number and percentage of censored OS times, quartile OS times, 95% CIs for the quartile OS times, and KM survival percentages with 95% CIs at select milestone times (e.g., 6, 12 and 18 months).

7.8.4. Relapse-Free Survival Following CR

RFS time is defined only for patients who achieve a CR during the study and is measured from the date of CR (Day 1) until (a) relapse after CR, or (b) death from any cause, whichever occurs first.

An additional rule for defining RFS time is:

- A patient found to have relapsed after missing one or more scheduled response assessments will have his RFS time end on the scheduled date for the first missed response assessment following the last confirmation of CR

Patients enrolled subsequent to protocol amendment 6 were scheduled to be followed for 2 years following the end-of-study visit, whereas patients enrolled prior to protocol amendment 6 were followed for only 30 days following the end-of-study visit unless they reconsented to the amendment-6, 2-year follow-up period. The definition of RFS time and the following rules for censoring RFS time apply to all patients regardless of follow-up duration.

Rules for censoring RFS time are:

- A patient with RFS continuing as of the date he is lost to follow-up will have his RFS time censored on the date of his last response assessment
- A patient with RFS continuing as of the database lock will have his RFS time censored on the date of his last response assessment

Summary statistics will include the number and percentage of patients who achieve a CR, the number and percentage of censored RFS times, quartile RFS times, 95% CIs for the quartile RFS times, and KM “survival” percentages with 95% CIs at select milestone times (e.g., 6, 12 and 18 months).

Two endpoints analogous to relapse-free survival following CR will be analyzed using similar methods.

- Relapse-Free Survival Following Combined CR (defined only for patients who achieve a combined CR during the study): time from the earliest date of combined CR (Day 1) until (a) relapse after combined CR, or (b) death from any cause, whichever occurs first
- Relapse-Free Survival Following Combined Remission (defined only for patients who achieve combined remission during the study): time from the earliest date of combined remission (Day 1) until (a) relapse after combined remission, or (b) death from any cause, whichever occurs first

Response categories included in “combined CR” and “combined remission” are delineated in Section 3.

7.8.5. Complete Remission Duration

CR duration is defined only for patients who achieve a CR during the study and is measured from the date of CR (Day 1) until relapse.

An additional rule for defining CR duration is:

- A patient found to have relapsed after missing one or more scheduled response assessments will have his CR duration end on the scheduled date for the first missed response assessment following the last confirmation of CR

Patients enrolled subsequent to protocol amendment 6 were scheduled to be followed for 2 years following the end-of-study visit, whereas patients enrolled prior to protocol amendment 6 were followed for only 30 days following the end-of-study visit unless they reconsented to the amendment-6, 2-year follow-up period. The definition of CR duration and the following rules for censoring CR duration apply to all patients regardless of follow-up duration.

Rules for censoring CR duration are:

- A patient who dies while in remission will have his CR duration censored on the date of death
- A patient in remission on the date he is lost to follow-up will have his CR duration censored on the date of his last response assessment
- A patient in remission on the database lock will have his CR duration censored on the date of his last response assessment

Summary statistics will include the number and percentage of patients who achieve a CR, the number and percentage of censored CR durations, quartile CR durations, 95% CIs for the quartile CR durations, and KM “survival” percentages with 95% CIs at select milestone times (e.g., 6, 12 and 18 months).

Two endpoints analogous to CR duration will be analyzed using similar methods.

- Combined CR Duration (defined only for patients who achieve a combined CR during the study): time from the earliest date of combined CR (Day 1) until relapse after combined CR
- Combined Remission Duration (defined only for patients who achieve combined remission during the study): time from the earliest date of combined remission (Day 1) until relapse after combined remission

Response categories included in “combined CR” and “combined remission” are delineated in Section 3.

7.8.6. ECOG Performance Status

ECOG performance status will be summarized within and across alvocidib dose levels using shift tables.

7.9. Biomarker

The biomarker in this study is MCL-1 dependence measured by mitochondrial sensitivity to NOXA BH3 peptides. Relevant data supporting analyses of MCL-1 dependence will be listed by patient within alvocidib dose level. Summary of baseline biomarker values will include descriptive statistics appropriate for continuous variables (see section 7.6). An LR model will be fit to examine the relationship between baseline MCL-1 dependence and the independent binary variable CR for newly diagnosed AML patients receiving ACM. In addition, area under the curve (AUC) will be calculated (by treatment

arm) for the trapezoidal receiver operating characteristic (ROC) curve. This value will quantify the ability of MCL-1 dependence to predict CR.[11] Ninety-five percent CIs will be calculated by assuming estimated AUC follows a normal distribution.

$$\widehat{\text{AUC}} \pm z_{1-\alpha/2}\, SE_{\text{AUC}}$$

where the standard error SE_{AUC} is calculated by the jackknife method.[12] Similar analyses will be conducted for combined CR and combined remission.

7.10. Safety Endpoints

Summaries of safety endpoints will include data collected from the safety patient population. Relevant data supporting safety analyses will be listed by patient within alvocidib dose level.

7.10.1. 30- and 60-Day Mortality

Estimates of 30- and 60-day mortality and their pointwise 95% CIs will be derived from KM curves for OS by treatment arm (see section 7.8.3).

7.10.2. Adverse Events

Reported adverse event (AE) terms will be mapped to MedDRA preferred terminology. AEs suggestive of tumor lysis syndrome (TLS) will be flagged based in the standardized MedDRA query (SMQ) for TLS. All reported events will appear in AE listings, however only treatment-emergent adverse events will be summarized. A treatment-emergent adverse event (TEAE) is an AE that starts or increases in severity any time after the first administration of any study drug up to 30 days following the last administration of any study drug. AE severity was rated by the investigator according to Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 criteria.[13]

A high-level safety summary will display the numbers of patients within and across alvocidib dose levels who experience one or more AEs in each of the following categories:

- All TEAEs regardless of severity or presumed relationship to study drug
- TEAEs judged related to study drug
- All TLS SMQ TEAEs
- TLS SMQ TEAEs judged related to study drug
- Treatment-emergent serious adverse events (SAEs)
- TEAEs leading to a delay in the administration of study drug
- TEAEs leading to a reduction in the protocol-specified dose of study drug
- TEAEs leading to discontinuation of study drug
- TEAEs leading to withdrawal from the study
- TEAEs leading to death

The base summary of TEAEs will show incidence rates for each MedDRA primary SOC and/or PT by highest reported CTCAE severity grade and overall. A separate summary

will be produced each of the AE subsets listed above. Additional AE summaries may be produced using safety data from subsets of patients and/or characterized using additional SMQs.

Partial start and end dates for adverse events will be replaced by calendar dates that maximize the duration of the adverse event. The following steps will be followed.

For a partial start date:

1. Replace a missing month with January
2. Replace a missing calendar day with the first of the month
3. If the replacement date is prior to the first dose of study drug, then set the adverse event start date equal to the date of the first dose

For a partial end date:

1. Replace a missing month with December
2. Replace a missing calendar day with the last day of the month

The replacement date for a missing end date may exceed end of the adverse event reporting period specified in the study protocol.

An end date will not be estimated for adverse events marked continuing at the end of the study.

7.10.3. Classification of Tumor Lysis Syndrome

Cairo-Bishop definitions for laboratory TLS and clinical TLS will be used to diagnose and grade TLS.[14] Patients are classified as experiencing laboratory TLS if they have any two or more serum values of uric acid, potassium, phosphorus and/or calcium shown in Table 2 within 3 days before or 7 days after the initiation of study drug.

Table 2. Cairo-Bishop Definition of Laboratory Tumor Lysis Syndrome

Laboratory Parameter	Criteria
Uric Acid	≥476 µmol/L or 25% increase from baseline
Potassium	≥6.0 mmol/L or 25% increase from baseline
Phosphorus	≥1.45 mmol/L or 25% increase from baseline
Calcium	≤1.75 mmol/L or 25% decrease from baseline

Diagnosis of clinical TLS requires the presence of laboratory TLS and elevated serum creatinine levels, cardiac arrhythmia and/or seizures. Criteria for grading the severity of clinical TLS are presented in Appendix 1.

The numbers of patients experiencing laboratory TLS and the various severities of clinical TLS will be summarized within and across alvocidib dose levels.

7.10.4. Clinical Laboratory Parameters

Within-cycle patient minimum and maximum laboratory values (and changes from baseline) for analytes measured on a continuous scale will be summarized within and across alvocidib dose levels using descriptive statistics appropriate for continuous variables (see section 7.6). Ordinal categorical test results will be summarized within

and across alvocidib dose levels using shift tables. Additionally, for tests where CTCAE [version 4.03] severity criteria are specified, CTCAE severity grades will be summarized in shift tables.

7.10.5. Neutrophil Recovery

Time to neutrophil recovery is defined for all patients and is measured from the date of first dose (Day 1) until the first day the absolute neutrophil count (ANC) is $\geq 1x10^9$/L for 14 consecutive days. It is assumed that the ANC remains above the threshold value of $1x10^9$/L unless there is evidence that it did not. Therefore, a patient with ANC $\geq 1x10^9$/L on Day 45 and without another assessment for >14 days will have a time to neutrophil recovery of 45 days.

Rules for censoring time to neutrophil recovery are:

- A patient with ANC $<1x10^9$/L as of the date he is lost to follow-up will have his time to neutrophil recovery censored on the date of his last assessment of any type
- A patient with ANC $<1x10^9$/L as of the database lock will have his time to neutrophil recovery censored on the date of his last assessment of any type

Summary statistics will include the number and percentage of censored times to neutrophil recovery, quartile times to neutrophil recovery, 95% CIs for the quartile times to neutrophil recovery, and KM “survival” percentages with 95% CIs at select milestone times (e.g., 42, 49 and 56 days).

7.10.6. Platelet Recovery

Time to platelet recovery is defined for all patients and is measured from the date of first dose (Day 1) until the first day the platelet count is $\geq 100x10^9$/L for 14 consecutive days. It is assumed that the platelet count remains above the threshold value of $100x10^9$/L unless there is evidence that it did not. Therefore, a patient with platelet $\geq 100x10^9$/L on Day 60 and without another assessment for >14 days will have a time to platelet recovery of 60 days.

Rules for censoring time to platelet recovery are:

- A patient with platelet count $<100x10^9$/L as of the date he is lost to follow-up will have his time to platelet recovery censored on the date of his last assessment of any type
- A patient with platelet count $<100x10^9$/L as of the database lock will have his time to platelet recovery censored on the date of his last assessment of any type

Summary statistics will include the number and percentage of censored times to platelet recovery, quartile times to platelet recovery, 95% CIs for the quartile times to platelet recovery, and KM “survival” percentages with 95% CIs at select milestone times (e.g., 49, 56 and 63 days).

7.10.7. Concomitant Medication Use

Optionally, depending on the perceived relevance, concomitant medications will mapped to terminology in the WHO DDE and then summarized within and across

alvocidib dose levels by usage rates for each level-1 ATC term and preferred (i.e., standardized) drug name.

8. Adverse Event and Medication Coding Dictionaries

The MedDRA and WHO DD versions (editions) that will be used to standardize medical terminology are not known at this time, but both will be recent editions published within 2 years before the date the first patient is enrolled into the study and the versions used will be documented in the clinical study report.

9. Summary Tables

Suggested titles for summary tables to be included in the clinical study report are shown in Table 3. The actual table titles may be altered as needed for stylistic consistency and to support text in the clinical study report. In addition, some tables listed below may be combined under a single title. A blueprint for the general format of these tables is presented in Appendix 2.

Table 3. Planned Summary Tables

Summary Table [*comments*]	Patient Population		
	ITT	Safety	Per-Protocol
Patient Disposition	X		
Screening and Demographic Information	X	X	X
AML History	X	X	X
Genetic Mutations	X	X	X
Medical History	X	X	X
Pre-Existing Medical Conditions	X	X	X
Relevant Prestudy Medications	X	X	X
Study Drug Exposure	X	X	X
Complete Remission Rates – Independent Hematologists Review [*repeat as needed for subgroup analyses*]	X		X
Complete Remission Rates – Local Investigator Review [*repeat as needed for subgroup analyses*]	X		X
Combined CR and Combined Remission Rates – Independent Hematologists Review [*repeat as needed for subgroup analyses*]	X		X
Combined CR and Combined Remission Rates – Local Investigator Review [*repeat as needed for subgroup analyses*]	X		X
Progression-Free Survival [*repeat as needed for subgroup analyses*]	X		X
Progression-Free Survival Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Overall Survival [*repeat as needed for subgroup analyses*]	X		X

Summary Table [*comments*]	Patient Population		
	ITT	Safety	Per-Protocol
Relapse-Free Survival Following CR [*repeat as needed for subgroup analyses*]	X		X
Relapse-Free Survival Following CR Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Relapse-Free Survival Following Combined CR [*repeat as needed for subgroup analyses*]	X		X
Relapse-Free Survival Following Combined CR Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Relapse-Free Survival Following Combined Remission [*repeat as needed for subgroup analyses*]	X		X
Relapse-Free Survival Following Combined Remission Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Complete Remission Duration [*repeat as needed for subgroup analyses*]	X		X
Complete Remission Duration Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Combined Complete Remission Duration [*repeat as needed for subgroup analyses*]	X		X
Combined Complete Remission Duration Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Combined Remission Duration [*repeat as needed for subgroup analyses*]	X		X
Combined Remission Duration Censoring at Transplant [*repeat as needed for subgroup analyses*]	X		X
Stem Cell Transplant Rates	X		X
ECOG Performance Status Shift Table	X		X
30- and 60-Day Mortality		X	
Numbers of Patients with Adverse Events		X	
Treatment-Emergent Adverse Events: Incidence Rates by MedDRA System Organ Class		X	
Treatment-Emergent Adverse Events: Incidence Rates by MedDRA Preferred Term		X	
Treatment-Emergent Adverse Events Judged Related to Study Drug: Incidence Rates by MedDRA System Organ Class		X	
Treatment-Emergent Adverse Events Judged Related to Study Drug: Incidence Rates by MedDRA Preferred Term		X	

Summary Table [*comments*]	Patient Population		
	ITT	Safety	Per-Protocol
Treatment-Emergent Adverse Events in the Standardized MedDRA Query for Tumor Lysis Syndrome: Incidence Rates by MedDRA Preferred Term		X	
Treatment-Emergent Adverse Events in the Standardized MedDRA Query for Tumor Lysis Syndrome Judged Related to Study Drug: Incidence Rates by MedDRA Preferred Term		X	
Serious Adverse Events: Incidence Rates by MedDRA System Organ Class		X	
Serious Adverse Events: Incidence Rates by MedDRA Preferred Term		X	
Serious Adverse Events Judged Related to Study Drug: Incidence Rates by MedDRA System Organ Class		X	
Serious Adverse Events Judged Related to Study Drug: Incidence Rates by MedDRA Preferred Term		X	
Treatment-Emergent Adverse Events Leading to a Delay in the Administration of Study Drug: Incidence Rates by MedDRA Preferred Term		X	
Treatment-Emergent Adverse Events Leading to a Reduction in the Protocol-Specified Dose of Study Drug: Incidence Rates by MedDRA Preferred Term		X	
Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug: Incidence Rates by MedDRA Preferred Term		X	
Treatment-Emergent Adverse Events Leading to Withdrawal from the Study: Incidence Rates by MedDRA Preferred Term		X	
Fatal Adverse Events: Incidence Rates by MedDRA Preferred Term		X	
Tumor Lysis Syndrome Incidence Rates During Cycle 1 by Maximum Cairo-Bishop Severity Grade		X	
Tumor Lysis Syndrome Incidence Rates During Study by Maximum Cairo-Bishop Severity Grade		X	
Clinical Laboratory Test Results and Changes from Baseline		X	
Clinical Laboratory Test Result Shift Tables		X	
Vital Signs and Changes from Baseline		X	
Concomitant Medications		X	

10. References

1. Dohner H, Estey E, Grimwade D, et al. Diagnosis and Management of AML in Adults; 2017 ELN Recommendations from an International Expert Panel. *Blood*. Jan 26 2017;129:424-447.
2. O'Brien PC, Fleming TR. A Multiple Testing Procedure for Clinical Trials. *Biometrics*. 1979;35:549-556.
3. Grimwade D, Walker H, Oliver F, et al. The Importance of Diagnostic Cytogenetics on Outcome in AML: Analysis of 1,612 Patients Entered Into the MRC AML 10 Trial. *Blood*. Oct 1 1998;92:2322-2333.
4. Yue L. Testing Multiple Secondary Endpoints in Confirmatory Comparative Studies – A Regulatory Perspective. Presented at: 2nd Annual AdvaMed Statistical Issues for Medical Devices and Diagnostics; April 29, 2009; Bethesda, MD. Available at: http://www.advamedmtli.org/download/File/ Statistics_April_2009/Yue_lilly_pres.pdf. Accessed February 5, 2015.
5. Koch GG, Gansky SA. Statistical Considerations for Multiplicity in Confirmatory Protocols. *Drug Information J*. 1996;30:523-533. Available at: http:// ferran.torres.name/download/shared/general/Multiplicity%20dj302612.pdf. Accessed February 5, 2015.
6. Mosteller RD. Simplified Calculation of Body-Surface Area. *N Engl J Med*. 1987;317:1098
7. Mantel N, Haenszel W. Statistical Aspects of Analysis of Data from Retrospective Studies of Disease. *J Natl Cancer Inst*. 1959;22:719-748.
8. SAS Institute Inc. 2010. Base SAS® 9.2 Procedures Guide: Statistical Procedures, Third Edition. Cary, NC: SAS Institute Inc.
9. Cohen J. A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas*. 1960;20:37-46.
10. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc*. 1958;53:457-481.
11. Pierceall WE, Lena R, Medeiros BC, et al. Mcl-1 Dependence Predicts Response to Vorinostat and Gemtuzumab Ozogamicin in Acute Myeloid Leukemia. *Leuk Res*. 2014;38:564-568.
12. Efron B, Stein C. The Jackknife Estimate of Variance. *Annals Stat*. 1981;9:586-596.
13. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. U.S. Dept of Health and Human Services, National Institutes of Health, National Cancer Institute; May 28, 2009 (v4.03 June 14, 2010). Available at: http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf. Accessed February 16, 2015.
14. Cairo MS, Bishop M. Tumor Lysis Syndrome: New Therapeutic Strategies and Classification. *Br J Haematol*. 2004;127:3-11.

Appendix 1. Criteria for Grading Severity of Clinical Tumor Lysis Syndrome

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Renal	Creatinine ≤1.5 x ULN	Creatinine >1.5-3.0 x ULN	Creatinine >3.0-6.0 x ULN	Creatinine >6.0 x ULN	Death
Cardiac Arrhythmia	Intervention not indicated	Non-urgent medical intervention indicated	Symptomatic and incompletely controlled medically or controlled with device (e.g., defibrillator)	Life-threatening (e.g., arrhythmia associated with CHF, hypotension, syncope, shock)	Death
Seizure	—	One brief generalized seizure; seizure(s) well controlled by anticonvulsants or infrequent focal motor seizures not interfering with activities of daily living	Seizure in which consciousness is altered; poorly controlled seizure disorder; with breakthrough generalized seizures despite medical intervention	Seizures of any kind which are prolonged, repetitive or difficult to control (e.g., status epilepticus, intractable epilepsy)	Death

Clinical TLS requires one or more clinical manifestations (renal, cardiac, neuro) along with criteria for laboratory TLS
Maximal clinical manifestation defines the severity grade
ULN = upper limit of normal; 114.4 and 105.6 µmol/L will be used for males and females, respectively, if ULN is not available for an institution

Appendix 2. Sample Table Template

Tolero Pharmaceuticals ALV-101

Page 1 of 327

Table X.X
Patient Disposition
(Intent-to-Treat Population)

	Alvocidib 20/30 mg/m² (N=3)	Alvocidib 30/40 mg/m² (N=3)	Alvocidib 30/50 mg/m² (N=3)	Alvocidib 30/60 mg/m² (N=13)	Study Total (N=22)
Study Center					
Columbia University Medical Center	3 (100%)	1 (33.3%)	0	2 (15.4%)	6 (27.3%)
University of North Carolina	0	2 (66.7%)	3 (100%)	7 (53.8%)	12 (54.5%)
Johns Hopkins University	0	0	0	4 (30.8%)	4 (18.2%)
Country					
United States	3 (100%)	3 (100%)	3 (100%)	13 (100%)	22 (100%)
Protocol Version					
Amendment 2	1 (33.3%)	0	0	0	1 (4.5%)
Amendment 3	2 (66.7%)	3 (100%)	3 (100%)	0	8 (36.4%)
Amendment 4	0	0	0	13 (100%)	13 (59.1%)
Safety Analysis Population					
Yes	3 (100%)	3 (100%)	3 (100%)	13 (100%)	22 (100%)
Response Evaluated Analysis Population					
Yes	3 (100%)	3 (100%)	3 (100%)	13 (100%)	22 (100%)
Alvocidib and 7+3 Induction					
Yes	3 (100%)	3 (100%)	3 (100%)	13 (100%)	22 (100%)

[a] Number of weeks from the first dose of alvocidib to withdrawal from study

Source: Data Extract ddmmmyy; SAS (v9.4) ddmmmyy at hh:mm

Table X.X
Patient Disposition
(Intent-to-Treat Population)

	Alvocidib 20/30 mg/m² (N=3)	Alvocidib 30/40 mg/m² (N=3)	Alvocidib 30/50 mg/m² (N=3)	Alvocidib 30/60 mg/m² (N=13)	Study Total (N=22)
Alvocidib and 5+2 Reinduction					
Yes	0	1 (33.3%)	0	1 (7.7%)	2 (9.1%)
No	3 (100%)	2 (66.7%)	3 (100%)	12 (92.3%)	20 (90.9%)
HiDAC Consolidation Cycles					
0	1 (33.3%)	2 (66.7%)	1 (33.3%)	9 (69.2%)	13 (59.1%)
1	1 (33.3%)	0	0	3 (23.1%)	4 (18.2%)
2	0	0	1 (33.3%)	1 (7.7%)	2 (9.1%)
3	1 (33.3%)	0	1 (33.3%)	0	2 (9.1%)
4	0	1 (33.3%)	0	0	1 (4.5%)
Weeks on Study[a]					
Mean (S.D.)	19.7 (13.8)	15.0 (11.0)	15.0 (8.1)	8.1 (3.6)	13.6 (9.1)
Median	17.7	9.4	14.4	9.0	10.0
Min, Max	7.0, 34.4	7.9, 27.6	7.3, 23.4	3.9, 12.3	3.9, 34.4
Primary Reason Patient was Withdrawn from the Study					
Completed the Study	2 (66.7%)	2 (66.7%)	1 (33.3%)	2 (15.4%)	7 (31.8%)
Death	0	0	0	1 (7.7%)	1 (4.5%)
Relapsed Disease (following remission)	0	0	1 (33.3%)	1 (7.7%)	2 (9.1%)
Lack of Efficacy (failure to achieve remission)	0	1 (33.3%)	1 (33.3%)	1 (7.7%)	3 (13.6%)
Unknown	1 (33.3%)	0	0	0	1 (4.5%)
Missing	0	0	0	8 (61.5%)	8 (36.4%)

[a] Number of weeks from the first dose of alvocidib to withdrawal from study

Source: Data Extract ddmmmyy; SAS (v9.4) ddmmmyy at hh:mm

Table X.X
Patient Disposition
(Intent-to-Treat Population)

	Alvocidib 20/30 mg/m² (N=3)	Alvocidib 30/40 mg/m² (N=3)	Alvocidib 30/50 mg/m² (N=3)	Alvocidib 30/60 mg/m² (N=13)	Study Total (N=22)
Poststudy Bone Marrow Transplant					
Yes	1 (33.3%)	1 (33.3%)	1 (33.3%)	0	3 (13.6%)
No	2 (66.7%)	2 (66.7%)	2 (66.7%)	13 (100%)	19 (86.4%)

[a] Number of weeks from the first dose of alvocidib to withdrawal from study

Source: Data Extract ddmmmyy; SAS (v9.4) ddmmmyy at hh:mm