Official Title:	**An Open-label, Multicenter, Extension Study for Subjects who Participated in Prior Guadecitabine Clinical Studies**
NCT Number:	NCT03603964
Document Date:	Statistical Analysis Plan Version: 03 November 2021

STATISTICAL ANALYSIS PLAN

Protocol Title:	An Open-Label, Multicenter, Extension Study for Subjects Who Participated in Prior Guadecitabine Clinical Studies
Protocol Number:	SGI-110-12
Phase:	2
Compound Number:	Guadecitabine
Sponsor Name:	Astex Pharmaceuticals, Inc.
Legal Registered Address:	4420 Rosewood Drive, Suite 200 Pleasanton, CA 94588

Regulatory Agency Identifier Number(s):

IND Number:	102743
EudraCT Number:	2017-004603-52

Approval Date: 03 November 2021 (Version 1.0)

TABLE OF CONTENTS

VERSION HISTORY

This Statistical Analysis Plan (SAP) for study SGI-110-12 is based on the protocol dated 19JAN2018.

SAP Version	Approval Date	Change	Rationale
1.0	03NOV2021	Not Applicable	Original version

1.0 INTRODUCTION

The purpose of this document is to provide details of the planned analyses for SGI-110-12 study. The analyses conducted to support the primary objectives will focus on safety data, in particular adverse events. Efficacy data will also be evaluated to obtain long-term survival information.

Descriptions of the planned summaries and analyses of non-key study data such as baseline characteristics and demographics, important protocol deviations, and other items not critical for regulatory needs, can be found in Appendix 3 (Non-key Analysis Specifications).

Specifications for data tables, listings and figures are detailed in a separate document.

1.1 Objectives and Endpoints

Objectives	**Endpoints**
Primary	
• To provide ongoing treatment with Guadecitabine for subjects who benefitted from guadecitabine treatment in a previous Astex-sponsored clinical study and to obtain long-term safety information.	• Safety as measured by adverse events
Secondary	
• To obtain long-term survival information on subjects who participated in a previous Astex-sponsored guadecitabine clinical study.	• Overall Survival

1.2 Study Design

This is a multicenter, open-label extension study for subjects who participated in a previous Astex-sponsored guadecitabine clinical study (including but not limited to SGI-110-01, SGI-110-04, SGI-110-05, SGI-110-06, and SGI-110-07).

Subjects who were still receiving treatment with guadecitabine and in the opinion of the investigator were still benefitting from treatment at the time of database close of the original study will be eligible to participate in this extension study. Approximately 250 subjects could be enrolled.

Subjects will attend clinic visits on Days 1-5 of each 28-day cycle to receive treatment with guadecitabine. Data collection will be limited to treatment exposure, adverse events, concomitant medications, limited laboratory parameters, and survival status.

There is no predetermined schedule of analyses for this study. Analyses could have been performed during the course of the study depending on the needs of the clinical development program. The analyses described in the SAP are conducted at the conclusion of the study.

2.0 STATISTICAL HYPOTHESES

The primary focus of the study is to offer continued guadecitabine treatment to subjects who are benefiting from the treatment and to collect the safety information of guadecitabine. No formal statistical hypotheses will be tested in the analysis of the trial data. All analyses will be descriptive in nature.

3.0 SAMPLE SIZE DETERMINATION

The sample size for this extension study was based on the eligible subjects from prior guadecitabine studies. No formal sample size justification was necessary.

4.0 ANALYSIS SETS

For purposes of analysis, the following analysis sets are defined:

Analysis Set	**Description**
All Subjects	All screened subjects, including those who did not meet the study entry criteria or did not receive a study treatment.
Safety	All subjects who receive any amount of study treatment.
Efficacy	All subjects who receive any amount of study treatment.

5.0 STATISTICAL ANALYSES

5.1 General Considerations

No formal statistical hypotheses will be tested. All analyses will be descriptive in nature.

Continuous variables will be summarized with mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized with counts and percentages. The SAS® statistical package (version 9.4 or a later version) will be used for the analyses.

Summary tables of efficacy, and safety data will be provided.

Summaries of disposition will be based on the All Subjects analysis set (all screened subjects). Demographics and baseline characteristics, as well as efficacy summaries will be based on the Efficacy analysis set. Safety summaries will be based on the Safety analysis set. In this study, the Safety and Efficacy analysis sets are identically defined.

Unless otherwise specified, the baseline value is defined as the last value collected prior to initiation of the study treatment in the SGI-110-12 study. For variables which are collected on Day 1 of Cycle 1 (C1D1) without explicit timing, the C1D1 values will be used as the baseline values.

Month is defined as 30.4375 days for analyses conducted in this study.

Study day of an SGI-110-12 subject is defined in reference to the subject's C1D1 date in SGI-110-12 and calculated following the data programming standards as detailed in BDM Manual 02 - Astex Data Programming Standards.

Laboratory values recorded as an interval such as "≥x", "<x", or "2+" or recorded as below the limit of quantification (BLQ) will be handled, if necessary, for calculation purposes, following the data programming standards as detailed in BDM Manual 02 - Astex Data Programming Standards.

Incomplete medication and AE start and stop dates will be imputed conservatively following the data programming standards as detailed in BDM Manual 02 - Astex Data Programming Standards.

5.2 Participant Dispositions

The number and percentage (n, %) of subjects screened, enrolled (by prior study), treated, treatment discontinuation (with reason) and withdrawn from study (with reason) will be summarized. All screened subjects will be included in the disposition analysis.

5.3 Primary Endpoint(s) Analysis

The primary endpoint is safety as measured by adverse events. Please refer to Section 5.5.2 Adverse Events for the detail analyses.

5.4 Secondary Endpoint Analysis

5.4.1 Key/Confirmatory Secondary Endpoint

Long-term overall survival is a secondary endpoint. Survival duration for each subject starting from the time of randomization in the prior study will be determined.

5.4.1.1 Definition of Endpoint(s)

Overall Survival (OS)

Overall survival is defined as time, in days, from the time of randomization in the prior study to the date of death (regardless of cause). Subjects without a documented death date at the time of analysis will be censored at the last date known alive.

Survival time in days = (earliest of date of death or censoring – randomization date in the prior study)

5.4.1.2 Main Analytical Approach

The OS curves will be estimated using the Kaplan-Meier method (Kaplan et al 1958). The median (and quartiles) duration of OS and the associated 95% CI (based on the log-log transformation for the survival function) will also be generated.

In addition, OS curves will also be estimated using the Kaplan-Meier method by the disease types (treatment-naive AML, relapse/refractory AML, and relapse/refractory intermediate- or high-risk MDS (including CMML). The median (and quartiles) duration of OS and the associated 95% CI (based on the log-log transformation for the survival function) will be generated for each group of the disease types.

5.5 Safety Analyses

5.5.1 Extent of Exposure

Exposure to guadecitabine will be summarized by numbers of treatment cycles received, delayed, as well as duration of exposure (in months) in the SGI-110-12 study. These measures of exposure will be based on dose administrations and dose delays identified by the study site and collected on the SGI-110-12 study drug administration CRFs. Duration of exposure, in months, is defined as the last treatment date minus the first date of treatment + 1 divided by 30.4375.

Dose intensity, presented as the incidence of subjects receiving <80% of their intended dose, will be summarized by cycle. Dose intensity will be calculated as the actual total dose divided by planned total dose for each treatment cycle.

Both completed or partially completed dose cycles are counted in these summaries.

5.5.2 Adverse Events

Adverse event (AE) terms reported by study subjects or observed by investigators will be mapped to the appropriate system organ class (SOC) and preferred term (PT) according to the MedDRA version 21.0. Severity of AE will be graded by National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

Treatment-emergent AEs (TEAEs) are defined as events that occur or worsen on or after the date of the first study treatment (C1D1) until 30 days after the last dose of study treatment or the start of an alternative anti-cancer treatment, whichever occurs first. SAEs that occur more than 30 days beyond the last dose of study treatment or the start of an alternative anti-cancer treatment will also be considered treatment-emergent if the event is related to the study treatment.

An overall safety summary table containing counts and percentages of subjects with any TEAE, any TEAE Grade ≥3, TEAE resulting in permanent treatment discontinuation, TEAE resulting in drug interruption, TEAE resulting in dose reduction, any serious AE (SAE), and subcategories of SAEs (fatal and non-fatal) will be produced. Related TEAEs will be similarly summarized. Related events are those that the investigator considered to be suspected to be related to study treatment as described in the study protocol.

The number and percentage of subjects experiencing TEAEs will be summarized by MedDRA SOCs (sorted alphabetically) with PTs sorted by decreasing frequency within each SOC. Related TEAEs, TEAEs Grade ≥3, related TEAEs Grade ≥3, serious TEAEs, related serious TEAEs, AEs

with a fatal outcome, AEs resulting in permanent treatment discontinuation, AEs resulting in drug interruption, and TEAEs resulting in dose reduction will be summarized similarly. In addition, TEAEs, related TEAEs, and SAEs will be summarized by SOC, PT, and CTCAE grade where subjects with multiple occurrences of the same TEAE will be counted once based on the highest CTCAE grade.

All AE data collected in the study database will be listed, including those that are not treatment emergent.

5.5.3 Additional Safety Assessments (if applicable)

Laboratory Data

Laboratory values will be graded, if relevant and possible, by CTCAE version 4.03 in conjunction with the Harrison (18th edition) lab book normal values (Longo et al 2011).

Shift tables will display (1) shift from baseline grade to the worst grade during the study, and (2) shift from baseline grade to the last grade at the end of study.

Summaries will also be provided of the incidence of all new or worsening laboratory abnormalities (any CTCAE grade) and new or worsening CTCAE Grade ≥3 laboratory abnormalities by parameter.

A listing of potential Hy's Law cases will be provided.

In addition, for selected laboratory parameters, a figure of mean values by visit will be generated.

5.6 Interim Analyses

No interim analysis will be performed for this study. Data summaries may be provided during the course of the study to meet the needs of the clinical development program.

6.0 SUPPORTING DOCUMENTATION

6.1 Appendix 1: List of Abbreviations

Abbreviation	Definition
AE	adverse event
AML	acute myeloid leukemia
BDM	biostatistics and data management
CI	confidence interval
CTCAE	Common Terminology Criteria for Adverse Events
CMML	chronic myelomonocytic leukemia
MDS	myeloid dysplastic syndromes
MedDRA	medical dictionary for regulatory activities
OS	Overall Survival
PT	preferred term
SAE	serious adverse event
SOC	system organ class
TEAE	treatment-emergent adverse event

6.2 Appendix 2: Changes to Protocol-planned Analyses

Overall survival definition is changed from

"Time, in days, from the time of the **first dose of Guadecitabine** in the prior study to the date of death"

to

"Time, in days, from the time of **randomization** in the prior study to the date of death".

All subjects enrolled in the SGI-110-12 study were from the prior studies involving randomization. These prior studies used the randomization date as the start date for overall survival. The change of the SGI-110-12 start time of the overall survival from the first dose of Guadecitabine to the date of randomization is to maintain the definition consistence between SGI-110-12 and the prior randomized studies.

In addition, overall survival analysis by disease types is added.

6.3 Appendix 3: Non-key Analysis Specifications

Demographics and Other Baseline Characteristics

Subject demographics (age, sex, etc.) and baseline characteristics (total WBC counts, disease types, etc.) will be summarized using descriptive statistics.

Concomitant Medications

Medications will be coded by the WHO Drug Dictionary WHO-DDE B2 March 2018.

Concomitant medications are the medications taken with a start date on or after the start of the administration of study treatment (C1D1), or those with a start date before the start of the administration of study treatment (C1D1) and a stop date on or after the start of the administration of study treatment (C1D1). Medications taken beyond 30 days from the last dose of study treatment or after the start of an alternative anti-cancer (study disease only) treatment are not considered concomitant medications, unless they are used for treating a related SAE.

Concomitant medications will be summarized by WHO Drug's Therapeutic Subgroup (ATC level 2) and Who Drug Name, sorted alphabetically, using counts and percentages.

7.0 REFERENCES

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53(282):457–481.

Longo DL, Fauci AS, Kasper DL, Hauser SL, Jameson JL, Loscalzo J. Harrison's Principles of Internal Medicine, 18th edition. New York: McGraw-Hill, 2012.