STATISTICAL ANALYSIS PLAN
SL-401 in Patients with Acute Myeloid Leukemia or Blastic Plasmacytoid Dendritic Cell Neoplasm Protocol STML-401-0114

Protocol Number:	STML-401-0114
Protocol Version and Date:	Amendment 12: 17 January 2019 Amendment 11: 12 June 2018 Amendment 10: 27 March 2018 Amendment 9: 08 February 2017 Amendment 8: 08 February 2016 Amendment 7: 11 August 2015 Amendment 6: 01 December 2014 Amendment 5: 28 July 2014 Amendment 4: 06 June 2014 Amendment 3: 21January 2014 Amendment 2: 12 November 2013 Amendment 1: 21 October 2013 Original: 26 September 2013
Name of Test Drug:	SL-401(ELZONRIS® [Tagraxofusp-erzs])
Phase:	Phase 1/2
Methodology:	A 4-part, non-randomized, open-label, dose escalation and expansion, multicenter study
Sponsor:	Stemline Therapeutics, Inc. 750 Lexington Avenue New York, NY 10022 Tel: (646) 502-2310 Fax: (646) 389-0968
Sponsor Representative:	[redacted]
Analysis Plan Date:	19 April 2021
Analysis Plan Version:	Final Version 3.0

APPROVAL SIGNATURE PAGE

Protocol Title: SL-401 in Patients with Acute Myeloid Leukemia or Blastic Plasmacytoid Dendritic Cell Neoplasm

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

Protocol Number: STML-401-0114

Document Date / Version: 19 April 2021 / Final Version 3.0

Veristat, LLC Author:

Approval Signature	Job Title
[redacted]	[redacted]

Sponsor Approval

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

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

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

Sponsor Signatory:

Approval Signature	Job Title
[redacted]	[redacted]

TABLE OF CONTENTS

TABLES AND FIGURES INCLUDED IN THE TEXT

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Definition
AE	Adverse event
Allo-SCT	Allogeneic Hematopoietic Stem Cell Transplantation
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
AST	Aspartate aminotransferase
ATC	Anatomic therapeutic class
AUC	Area under the curve
BLA	Biological licensing application
BM	Bone marrow
BMCR	Bone marrow complete response
BPDCN	Blastic plasmacytoid dendritic cell neoplasm
C_{max}	Maximum plasma concentration
CI	Confidence interval
CR	Complete response
CRc	Complete response [Clinical] with minimal residual skin abnormality
CRF	Case report form
CRi	Complete response [Incomplete blood count recovery]
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DLT	Dose-limiting toxicity
DOR	Duration of response
DT	Diphtheria toxin
EBMTG	The European Blood and Marrow Transplantation Group
ECOG	Eastern Cooperative Oncology Group
ECG	Electrocardiogram
eCRF	Electronic case report form
FDA	Food and Drug Administration
ICH	International Conference on Harmonisation
IL-3	Interleukin-3
IL-3R	Interleukin-3 receptor
IRC	Independent Review Committee
IWG	International Working Group
hIL-3	Human interleukin-3
MedDRA	Medical Dictionary for Regulatory Activities

Abbreviation	Definition
mITT	Modified Intent-to-Treat
mSWAT	Modified Severity Weighted Assessment tool
MTD	Maximum tolerated dose
N	Number
ORR	Objective response rate
OS	Overall survival
PCS	Potentially clinically significant
PD	Progressive disease
PFS	Progression-free survival
PK	Pharmacokinetic
PR	Partial response
R/R	Relapsed/refractory
Rel Day	Relative study day
RFS	Relapse-free survival
SAE	Serious adverse event
SAP	Statistical analysis plan
SCT	Stem cell transplant
SD	Standard deviation
SI	International System of Units
SOC	System organ class
TEAE	Treatment-emergent adverse event
TTCR	Time to complete response (CR + CRc)
TTP	Time to progression
TTR	Time to response (CR + CRi + CRc + PR)
ULN	Upper limit of normal
US	United States
VOD	Veno-occlusive disease
WHO	World Health Organization

1. INFORMATION FROM THE STUDY PROTOCOL

1.1. Introduction and Objectives

The overall purpose of study STML-401-0114 is to support the development of SL-401 (ELZONRIS® [Tagraxofusp-erzs]; "tagraxofusp") and assess safety and clinical efficacy of the product in treatment of patients with first-line blastic plasmacytoid dendritic cell neoplasm (BPDCN), relapsed/refractory (R/R) BPDCN, R/R acute myeloid leukemia (AML), and firstline high risk AML. The primary focus of this document is the assessment of efficacy and safety of tagraxofusp in first-line BPDCN patients with supporting evidence from the remaining populations. Additional analyses will be performed on the R/R BPDCN, R/R AML, and firstline high risk AML patients.

AML is characterized by the uncontrolled proliferation of immature myeloid cells in the bone marrow (BM) and peripheral blood, resulting in the development of anemia, neutropenia, and thrombocytopenia, and associated complications such as serious infections. The median age at diagnosis is 67 years, and 5-year survival across all ages, treatments, and other prognostic groups is 24% (National Cancer Institute 2013).

BPDCN is an uncommon hematological malignancy that is characterized by the clonal proliferation of malignant plasmacytoid dendritic cells (Facchetti, et al. 2003). While the disease expression of BPDCN is similar to AML, BPDCN was characterized as a distinct cancer type as early as 1994 (Adachi, et al. 1994) though little was known of the disease biomarkers until the mid-2000s.

BPDCN most commonly affects middle-aged and older patients and is approximately 3.5-fold more common in men than women. The median age at diagnosis is 66 years and is lower for women than men. The true incidence and prevalence of BPDCN is not precisely known. However, based on a published report (Wang, et al. 2012), BPDCN constitutes approximately 0.44% of the incident cases of hematologic cancers annually, or approximately 700 and 1,000 incident cases annually in the United States (US) and Europe, respectively, based on published data of hematologic cancer cases.

The clinical features and evolution of BPDCN consist primarily of 2 patterns (Assaf, et al. 2007; Brody, et al. 1995; Herling and Jones 2007; Petrella, et al. 2005; Petrella, et al. 1999; Reichard, et al. 2005; Suzuki, et al. 2005). The first pattern (90% of cases) is characterized by an indolent onset dominated by cutaneous lesions followed by tumor dissemination. The second pattern shows features of an acute leukemia with systemic involvement from the outset, and multiple skin nodules are frequently present (Suzuki, et al. 2005). BPDCN most typically presents with skin lesions as well as extracutaneous malignant disease in the BM, blood, lymph nodes, spleen, and other organs. Despite the frequent appearance of a somewhat indolent clinical presentation at the outset, the course of BPDCN is highly aggressive and the median survival is approximately 12-14 months (Bekkenk, et al. 2004; Feuillard, et al. 2002; Herling and Jones 2007; Jacob, et al. 2003; Petrella, et al. 2005). The disease almost always results in a terminal leukemic phase with proliferation of BPDCN blasts in the BM and peripheral blood, leading to decreased blood cell counts with resultant infections, bleeding, and invariably death.

Diphtheria toxin (DT) Interleukin-3 (IL-3) fusion protein (tagraxofusp) is a novel biologic targeted therapy directed to the IL-3 receptor (IL-3R), a biomarker that is over-expressed on AML and BPDCN blasts and cancer stem cells (CSCs) relative to normal hematopoietic stem

cells (Jordan, et al. 2006; Jordan, et al. 2000; Tehranchi, et al. 2010). Tagraxofusp is comprised of recombinant human IL-3 genetically fused to a truncated DT, in which the native receptorbinding domain of DT has been replaced with IL-3. The IL-3 domain of tagraxofusp is able to target the agent to cells that over-express IL-3R, leading to receptor-mediated endocytosis and localization of tagraxofusp to early endosomes. This allows tagraxofusp to kill cells in a distinct manner from other cancer therapeutics. First, tagraxofusp is a targeted therapy directed to the IL-3R that is differentially present on tumor bulk and/or CSCs versus normal hematopoietic stem cells. Second, tagraxofusp utilizes a payload that is not cell cycle-dependent. Therefore, it is designed to kill not just highly proliferative tumor bulk, but also relatively quiescent CSCs. Tagraxofusp utilizes a payload that may not be subject to multi-drug resistance mechanisms typically used by CSCs to evade traditional therapies. The payload also kills cells in a manner that is distinct from that of other available therapies, which is another reason why tagraxofusp is an effective addition to the therapeutic armamentarium against certain hematologic malignancies.

Tagraxofusp in frozen liquid formulation was approved by the US Food and Drug Administration (FDA) in 2019 for treatment of BPDCN in adults and pediatric patients 2 years and older based primarily on the efficacy findings from this study, STML-401-0114. STML401-0114 was designed as a Phase 1/2 dose escalation and expansion study targeting patients with BPDCN or AML. Stage 1 (dose escalation) enrolled patients with BPDCN, R/R AML, and previously untreated AML considered high risk for progression and/or unfit for standard therapy. Stage 2 (expansion) enrolled additional patients with first-line BPDCN, patients with R/R BPDCN, and R/R and untreated high risk AML patients. Stage 3 was a confirmatory cohort to test pre-specified hypotheses regarding the efficacy of tagraxofusp among first-line BPDCN patients in preparation for the biological licensing application (BLA) filing. Stage 4 initiated enrollment after August 29, 2016 through January 7, 2019 and included R/R BPDCN patients and first-line BPDCN patients enrolled after March 17, 2017. Patients enrolled into Stage 4 were treated with lyophilized formulation of tagraxofusp. At the time of BLA filing and approval limited data were available from Stage 4.

The purpose of this document is to support an updated report on the clinical efficacy and safety of tagraxofusp, emphasizing both the Stage 4 results that were not previously reported and the overall conclusions from the duration of the trial. Emphasis is on the results from BPDCN patients. Updated data are included in this report from an additional 44 patients enrolled into Stage 4 (42 treated) after the data cut-off for the original clinical study report dated May 09, 2018.

1.1.1. Definitions

For the purpose of this document and related analyses:

- Patients who have not received any prior treatment for their disease are considered "first-line."
- Patients who have received any prior treatment, regardless of duration or documentation of prior response, are considered "relapsed/refractory," abbreviated R/R.
- The frozen-liquid formulation of tagraxofusp, abbreviated "frozen", was approved for treatment of BPDCN patients in 2019.

- The lyophilized formulation of tagraxofusp, abbreviated "lyo", is considered an experimental product.

1.1.2. Study Objectives

The following objectives are defined in the Study Protocol (Amendment 9, 09 February 2017). The response criteria referenced throughout the Study Objectives are provided in Section 8.

Primary objectives for Stages 1, 2, 3, and 4 are as follows:

Stage 1: The objectives of Stage 1 are to:

- Determine the maximum tolerated dose (MTD), or the maximum tested dose where multiple dose limiting toxicities (DLTs) are not observed, of tagraxofusp.

Stage 2: The objectives of Stage 2 are to:

- Determine the efficacy of tagraxofusp in patients with BPDCN, as assessed by objective response rate (ORR).
- Characterize the safety profile of tagraxofusp at the MTD or maximum tested dose in both patients with AML and BPDCN.

Stage 3: The objectives of Stage 3 are:

- Determine the complete response (CR) rate (ie, CR + CR with incomplete blood count recovery [CRi] + CR [clinical] with minimal residual skin abnormality [CRc]) in patients with first-line BPDCN. Note that these objectives are supported by prespecified statistical hypotheses and were evaluated using appropriate statistical inference during the initial reporting of this study.
- Characterize the safety profile of tagraxofusp in patients with first-line BPDCN.

Stage 4: The objectives of Stage 4 are to:

- Further characterize the efficacy of tagraxofusp in first-line and R/R patients with BPDCN following the completion of Stage 2 and Stage 3 as assessed by the rate of CR (CR + CRi + CRc).
- Further characterize the safety profile of tagraxofusp among first-line and R/R patients with BPDCN.
- Characterize the efficacy and safety of a lyophilized formulation of tagraxofusp among first-line and R/R patients with BPDCN.

Secondary objectives include:

- Determine the CR rate (CR + CRi + CRc) for Stage 1 and 2 patients, and ORR for Stage 3 patients.

- Estimate duration of response (DOR), progression-free survival (PFS) and overall survival (OS) in patients with BPDCN. Note that for patients that received a stem cell transplant (SCT), duration of response includes time pre- and post-SCT.

- Enable preliminary characterization of the estimated ORR in patients with R/R AML, including subsets of patients with R/R AML categorized by pre-treatment blast counts, cytogenetics, or CD123 measurement.

- Estimate DOR, PFS, and OS in patients with AML.

- Characterize the pharmacokinetics (PK) and immunogenicity of tagraxofusp.

Exploratory objectives are to characterize expression of IL-3R (and other potentially relevant stem cell and disease markers) on leukemia cells in peripheral blood and BM, to evaluate potential changes in IL-3R (and other potentially relevant markers) expression on populations over time, and preliminary correlation of baseline IL-3R (and other potentially relevant markers) expression and clinical efficacy (including response rates).

1.1.3. Study Objectives and Definitions Updates

[redacted] An analysis of Stage 1 and Stage 2 first-line BPDCN patients was performed to assess the relative clinical benefit of the response designations of CR, CRc, and CRi, with benefit defined as duration of response, as defined in Section 4.3.2. Durations of response for patients with best responses of CRc and CRi were compared to duration of response for patients with best response of CR to assess the likelihood that the durations of CR and CRc or CR and CRi would have been derived from the same probability distribution. From this analysis, it was determined that the achievement of CRc corresponded to a durability of response that is very similar to the durability of CR. Patients who attain a response of CRc would be expected to benefit from SL-401 treatment in a similar manner as those patients who achieve CR, thereby justifying the inclusion of the category of CRc in the primary efficacy analysis of Stage 3. The results of the analysis of comparability of CR and CRi did not provide sufficient evidence of similarity. [redacted] As a result, the definition of CR has been revised to include both results from patients who achieved either CR or CRc .

In addition to bone marrow clearance and amelioration of other components of disease, the skin component of clinical complete response was defined in the protocol to consist of marked clearance of all skin lesions from baseline, including only residual hyperpigmentation or abnormality with BPDCN identified on biopsy (or no biopsy performed). [redacted] for more specificity in the concept of marked reduction of skin abnormality in the context of CRc, and to be consistent with the operational characteristics for this designation used by the study investigators, marked reduction in skin manifestation for a response of CRc was defined as at least a 75% reduction in the mSWAT score from baseline. Further, residual disease was defined as a remaining skin abnormality of 10% mSWAT or less, with this residual

consisting of skin hyperpigmentation or other visual abnormality that was not considered by the investigator to be active disease. (Section 8.2)

1.1.4. Purpose of this Document

This statistical analysis plan (SAP) is designed to outline the methods to be used in the analysis of study data to answer the study objectives related to Stage 4 and provide additional safety summaries related to all enrolled patients. Populations for analysis, data handling rules, statistical methods, and formats for data presentation are provided. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of an updated clinical study report (CSR) for this trial, inclusive of Stage 4 results.

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

1.2. Study Design

1.2.1. Synopsis of Study Design

This study is a non-randomized, open-label, multicenter study divided into 4 stages: dose escalation (Stage 1), expansion (Stage 2), pivotal (Stage 3), and supportive efficacy and safety (Stage 4). The targeted patient populations are patients with:

- First-line BPDCN for primary efficacy and safety profile of tagraxofusp (Stages 1 to 4).
- R/R BPDCN for supportive efficacy and supportive safety profile of tagraxofusp (Stages 1, 2 and 4).
- R/R AML and high-risk first-line AML for supportive efficacy and supportive safety profile of tagraxofusp (Stages 1 and 2).

A cycle of therapy consists of administration of tagraxofusp at the identified dose for 5 consecutive days every 21 days (with the entire cycle duration lasting 21 days). tagraxofusp is supplied as tagraxofusp Injection frozen solution (1 mg/mL) for Stages 1, 2, and 3. A lyophilized formulation of tagraxofusp (for Injection, lyophilized powder, for solution; 1 mg/mL) was supplied for dosing in selected patients with AML in Stage 2 and for BPDCN patients enrolled in Stage 4 after 17 March 2017.

Stage 4 is planned to continue enrollment of first-line and R/R BPDCN patients to further characterize the efficacy and safety of tagraxofusp and to evaluate the lyophilized formulation in a subset of patients. Pooled analysis of Stage 4 data with other stages will be performed at the time of the primary analysis of study results.

1.2.2. Randomization Methodology

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

1.2.3. Stopping Rules and Unblinding

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

- Patient withdrawal of consent

- Occurrence of unacceptable toxicity, including DLT in Stage 1
- Tagraxofusp related anaphylaxis or Grade ≥3 hypersensitivity reaction
- Requirement for >1 dose reduction unless there is evidence of AML/BPDCN response (beyond Cycle 1), in which case additional dose reductions are permitted, however these reductions must be discussed with the Medical Monitor and documented in the context of ongoing AML/BPDCN response
- Disease recurrence/progression
- Intercurrent illness that prevents further administration of tagraxofusp
- Patient non-compliance
- Occurrence of pregnancy
- Investigator's decision

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

1.2.4. Study Procedures

The schedules of assessments, as outlined in the study protocol, are provided in Section 7, Table 7-1, Table 7-2, Table 7-3, and Table 7-4.

1.2.5. Efficacy, Pharmacokinetic, and Safety Parameters

1.2.5.1. Efficacy Parameters

The primary evidence of efficacy was derived from an analysis of the rate of CR (CR + CRc), accompanied by an evaluation of the durability of response (including time pre- and post-SCT, for those patients who received SCT), as key secondary endpoints.

These analyses will be performed on data BPDCN patients treated at the 12 µg/kg/day dose level:

1. Bone marrow CR (BMCR; refer to Section 4.3.1) and duration of BMCR
2. CR (CR + CRc) rate and duration of CR

4. ORR (CR + CRc + CRi + partial response [PR]) and duration of objective response

6. Proportion of patients who receive an SCT
7. PFS
8. OS

Additional exploratory efficacy analyses will be performed on the above endpoints on pooled data across Stages and on additional exploratory efficacy parameters. Exploratory efficacy parameters will include time to CR (CR + CRc) (TTCR) and time to response (CR + CRc + CRi

+ PR) (TTR), time to BMCR, time to progression, and relapse-free survival (RFS). In addition, an exploratory efficacy analysis will be performed to assess the rate and duration of CR + CRc + CRi.

Bone marrow response will be assessed using International Working Group (IWG) criteria for AML (for patients with AML or BPDCN) (Cheson, et al. 2003), skin-related response using a weighted BSA/modified Severity Weighted Assessment tool (mSWAT) score (for patients with BPDCN) (Olsen, et al. 2011), and lymph node and visceral response using a modified version of the Revised Response Criteria for Malignant Lymphoma (for patients with BPDCN) (Cheson, et al. 2007). For the patients with BPDCN, assessment of response and disease progression will include evaluation of all sites of disease involvement at baseline, which may include BM, skin, peripheral blood, lymph nodes, spleen, and liver. Response criteria for AML and BPDCN are summarized in Sections 16.1 and 16.2 of the study protocol, Amendment 9. As applicable, up to 6 of the largest dominant nodes or nodal masses should be identified as index lesions for radiographic response assessments. Further detail on the definitions and analysis methods for efficacy endpoints are provided in Section 4.3.

1.2.5.2. Safety Parameters

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

Further details on the definitions and analysis methods for safety endpoints are provided in Section 4.3.6.

1.2.5.3. Pharmacokinetic and Immunogenicity Parameters

Summary of noncompartmental PK parameters and PK exposure response to treatment summaries will be described in a separate Modeling and Simulation Plan.

Summary of immunogenicity parameters, including anti-drug antibodies (further characterized for specificity to DT and/or human interleukin-3 [hIL-3]), specific anti-hIL-3 antibodies, and tagraxofusp neutralizing antibodies will be performed.

Further details on these analyses are provided in Section 4.5.

2. PATIENT POPULATION

2.1. Population **Definitions**

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

- Modified Intent-to-Treat (mITT) Population: All patients who are eligible based on the screening criteria and who received at least 1 dose of tagraxofusp. For BPDCN patients to be eligible for mITT they must also have diagnosis of disease based on central pathology review to be "Confirmed" or "Inconclusive: Likely BPDCN". Patients who meet the criteria for mITT will be considered "evaluable." Patients will be grouped according to the planned dose level at time of enrollment.

- Immunogenicity Population: All patients enrolled in the study who received at least 1 dose of tagraxofusp and had sufficient samples to permit immunogenicity analysis.

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

The mITT population is the primary population for the analysis of efficacy parameters. The Immunogenicity population will be used for summary of immunogenicity results. Supportive subgroup efficacy analyses will be performed in the mITT population, as described in Section 3.9. When the subgroup is defined by immunogenicity parameters, the population will be mITT patients with sufficient samples collected to permit the respective analysis. The Safety population is the primary population for the analysis of safety.

2.2. Protocol Violations

At the discretion of the Sponsor, major protocol violations, as determined by a review of the data prior to database lock and the conduct of statistical analyses described by this document (including Stage 3 efficacy analysis). The Sponsor or designee will be responsible for producing the final protocol violations file (formatted as a Microsoft Excel file), in collaboration with Veristat and the data monitoring group as applicable. This file will be finalized prior to database lock.

All protocol violations will be presented in a data listing.

3. GENERAL STATISTICAL METHODS

3.1. Sample Size Justification

The sample size was originally planned to be approximately 40-50 patients with BPDCN, including approximately 40 previously untreated patients planned to be treated with the tagraxofusp dose as determined in the completed Stage 1 of this study (12 µg/kg/day). In addition, up to approximately 36 patients with R/R AML were planned to be enrolled in Stage 2 of the study. These sample sizes were not based on a formal statistical approach but were considered reasonable to accomplish study objectives for these phases of the trial.

In Stage 3, a sufficient number of patients were enrolled to ensure 10 evaluable patients would be included, with evaluable patients in the mITT population, defined as patients with a diagnosis

of BPDCN, as confirmed via central pathology review, and who received at least 1 dose of tagraxofusp. [redacted] The primary efficacy analysis for Stage 3 compared the lower bound of a 2-sided 95% Clopper Exact confidence interval (CI) on the observed CR rate to a clinically meaningless value of 10%. Assuming a CR rate of at least 60%, a minimal sample size of 10 first-line BPDCN patients provided at least 90% power for the primary efficacy assessment. Due to the criteria for evaluability in the BPDCN mITT population requiring central pathology review, 13 first-line BPDCN patients were enrolled in Stage 3 between October 26, 2016 and March 17, 2017. Statistical significance was determined when the lower bound of this CI fell above the rate of 10%.

After completion of enrollment of patients with first-line BPDCN in Stage 3, 39 first-line BPDCN patients were enrolled in Stage 4 (37 treated with study drug and 2 screen failures) and 5 R/R BPDCN enrolled in this study after 29 August, 2016 were enrolled in Stage 4.

3.2. General Methods

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

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

Tabulations will be produced for appropriate demographic, baseline, efficacy, safety and immunogenicity parameters. For categorical variables, summary tabulations of the number and percentage of patients within each category (with a category for missing data) of the parameter will be presented. Two-sided 95% CIs will be computed using the Clopper Exact method. For continuous variables, the number of patients, mean, median, standard deviation (SD), minimum, and maximum values will be presented. Time-to-event data will be summarized using Kaplan-Meier methodology using 25^{th}, 50^{th} (median), and 75^{th} percentiles with associated 2-sided 95% CIs, as well as percentage of censored observations. Formal statistical hypothesis testing will be performed on the primary efficacy endpoint, CR (CR + CRc) rate among all evaluable first-line BPDCN patients enrolled in Stage 3, with the hypothesis test performed by comparing the lower-bound of a 2-sided 95% Clopper Exact CI to the benchmark of 10%.

3.3. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software Version 9.3, unless otherwise noted. Medical history and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 19.0. Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary (September, 2016).

3.4. Baseline Definitions

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

3.5. Methods of Pooling Data

Summaries will be displayed by stage of enrollment and drug formulation for FL BPDCN patients treated at 12 µg/kg, for R/R BPDCN at 12 µg/kg and overall, and AML patient groups.

Efficacy analyses for patients with first-line and R/R BPDCN who received 12 µg/kg will be conducted for the patient pools described in Table 3-1.

Table 3-1: Planned Pooling of Data for Efficacy Analysis at 12 µg/kg/day

Disease and Line of Therapy	**FL BPDCN**				**R/R BPDCN**		
Patient Population	**mITT**				**mITT**		
Stages Included	1&2	3	4	3&4*	1&2	4	1,2 &4
CR + Duration	X	X	X	X	X	X	X
BMCR + Duration		X	X	X			
ORR + Duration	X	X	X	X	X	X	X
Percent SCT		X	X	X			X
PFS		X	X	X			X
OS		X	X	X			X
TTCR		X	X	X			
TTR		X	X	X			
TT BMCR		X	X	X			
RFS		X	X	X			
TTP		X	X	X			

Abbreviations: mITT = modified intent to treat population; PP = per protocol population; BPDCN = blastic plasmacytoid dendritic cell neoplasm; FL = first-line; R/R = relapsed/refractory; CR = complete response (CR + CRc); BMCR = bone marrow complete response; ORR = objective response rate; SCT = stem cell transplant; PFS = progression free survival; OS = overall survival; TTCR = time to complete response (CR + CRi + CRc); TTR = time to response (CR + CRi + CRc + PR); TT BMCR = time to bone marrow complete response; RFS = relapse free survival; TTP = time to progression..

^ Per protocol analyses to be performed if protocol violations lead to multiple patients being excluded from the Stage 3 PP population.

* Stages 3&4 may be pooled if the distributions are determined to be similar, as described in Section 3.5.1.

3.5.1. Pooling of Stages 3 and 4

Patients with first-line BPDCN in Stages 3 and 4 may be pooled for analysis if the relative distributions of CR and duration of CR are determined to be similar to one another and could arise from the same underlying distribution. The importance of this determination is to support the potential similarities between the frozen liquid formulation and the lyophilized formulation of tagraxofusp. If similarities are observed between the stages, pooling will provide increased sample size and therefore increased precision of estimates.

The assessments for similarity between the Stage 3 and Stage 4 results will proceed as follows:

1) Comparison of 90% confidence intervals surrounding CR rate, ORR rate, and percent SCT to determine similarities of coverage

2) Permutation test assessing whether the observed percentage difference in CR rate between Stage 3 and Stage 4 could occur by chance or if it is indicative of different distributions

3) Permutation test assessing whether the observed differences in duration of CR in Stages 3 and 4 could occur by chance or if it is indicative of different distributions

For the permutation tests, the null hypothesis is that Stages 3 and 4 results arise from the same distributions. For each respective parameter, 10,000 replicate datasets (including all observations from Stages 3 and 4 first-line BPDCN) will be drawn without replacement and summary statistics will be computed to develop the empirical null distribution of possible results. The observed result from Stage 3 and Stage 4 will be compared to the empirical null distribution. If the observed result falls within the 10^{th} to 90^{th} percentiles, it would be considered reasonable that Stages 3 and 4 arose from the same underlying distribution and therefore the pooling of the study results are appropriate.

3.6. Adjustments for Covariates

Due to small sample size, no formal statistical analyses that adjust for possible covariate effects are planned. Exploratory subgroup analyses will be performed for descriptive purposes, as described in Section 3.9.

3.7. Assessment of Clinical Benefit

As is noted in the literature, BPDCN is a rare hematologic malignancy that may represent less than 1% of cases presenting as AML, and has a multifaceted presentation involving manifestation in the BM and skin, and may also involve other organs including lymph nodes and other viscera. There are limited prospective data in this rare disease, with no standard of care treatment established. It is therefore important to assess the clinical benefit of potential new therapies through the analysis and interpretation of multiple clinically relevant endpoints.

Conventional chemotherapy alone historically has not resulted in long-term remissions in the majority of patients with BPDCN (Pemmaraju 2015; Pemmaraju 2014; Pemmaraju 2012); however, as noted in the literature, induction remission followed by bridge to stem cell transplant (SCT) has been shown to provide long-term remissions in certain patients and has been cited as a major, if not the key, therapeutic objective.

Allogeneic hematopoietic stem cell transplantation (allo-SCT) in BPDCN has been reported to be potentially beneficial in select BPDCN patients. Allo-SCT is ideally performed during first remission (Aoki, et al. 2015; Kharfan-Dabaja, et al. 2013; Riaz, et al. 2014; Roos-Weil, et al. 2013; Shapiro 2015). A large study using data from 86 first-line BPDCN patients recruited in the French network of BPDCN demonstrated a median OS for 34 patients who received SCT (40%) following CR (autologous SCT, 4 patients; or allo-SCT, 30 patients) of 49 months, against OS of 8 months for the 52 patients without SCT ($p<0.0001$). The beneficial effect of SCT appeared to persist independently of age in multivariate analysis (Poret 2015). The Leukemia and Lymphoma Society in October 2016 stated that allo-SCT may result

in remissions that are longer, especially if offered after the first remission, and recommended that BPDCN patients be evaluated for SCT as soon as feasible. Dalle and colleagues (Dalle, et al. 2010) found that BPDCN patients who had SCT had a median OS of more than 40 months versus approximately 12 months for patients who did not receive SCT, and concluded that BM transplantation should be considered as a first-line treatment option when feasible. Similar results were seen in an analysis of 39 BPDCN patients from the European Blood and Marrow Transplantation Group (EBMTG) registry who underwent SCT as reported by Roos-Weil and colleagues (Roos-Weil, et al. 2013).

It is not clear from the literature how many patients with BPDCN are eligible for SCT, but it is generally known that candidates for allo-SCT are typically young, have a limited number of comorbidities, and have a suitable donor. Although limited data are available, promising results also have been seen with autologous SCT (auto-SCT), with 4 year OS and PFS rates of 82% and 73%, respectively, in a small (n=11) number of patients who underwent auto-SCT in first remission (Aoki, et al. 2015).

Ideally, to optimize the chances of SCT success, a patient ideally needs to be in remission following induction therapy and be considered to be 'fit' for the procedure based on the SCT guidelines. Accordingly, given that BPDCN is largely a disease of the middle aged and elderly, therapies that have an acceptable safety profile, are quickly effective, and allow for postremission transplantation of older patients with BPDCN are urgently needed and would be clinically beneficial (Pemmaraju 2016).

Control of BPDCN through reduction of blast cells in the BM may allow a patient to be bridged to SCT and thus is clinically meaningful, whether manifested as a component of a global CR, ie, CR among all compartments of disease present at baseline for patients with marrow disease (≥5% blast cells) at baseline, or as a CR in the BM compartment (BMCR) with PR in other non-BM compartment(s). As noted in the study protocol, a patient with BMCR is defined as having a complete BM response (reduction to <5% blast cells), for patients with marrow disease (≥5% blast cells) at baseline, with either CR or PR in other non-BM compartments. Marrow disease, as in other aggressive hematologic malignancies, is a leading cause of morbidity and death in BPDCN and is also the major barrier to resumption of a meaningful quality of life (eg, due to the requirement for growth factor support, antibiotics, hospitalizations, etc). Further, resolution of disease in the BM, or maintenance of a lack of BM disease from baseline, may provide the opportunity for a bridge to SCT (ie, marrow must ideally be grossly disease-free for a patient to be considered for SCT), even if there is detectable disease in other compartments of BPDCN. Given that bridge to SCT is a major, if not the key, therapeutic objective of BPDCN therapy with evidence of long-term benefit in certain patients, we believe that clinically meaningful outcomes of induction therapy in BPDCN include either global CR or BMCR (with PR in other compartments) that lead to bridge to SCT (Leclerc, et al. 2017).

The proportion of patients who achieve a response of BMCR and the proportion of patients who receive SCT, should therefore, in our view, be considered as part of the overall key endpoints that provide evidence of clinical benefit of tagraxofusp in patients with first-line BPDCN. (Leclerc, et al. 2017).

In order to more completely demonstrate the clinical benefit of tagraxofusp, additional endpoints will be analyzed in a descriptive, supportive manner, including endpoints related to longer-term benefit such as PFS and OS, the proportion of patients who receive SCT and the extended duration of effect subsequent to SCT. These additional endpoint analyses will be

performed on the pivotal cohort of first-line BPDCN patients from Stage 3, as well as the larger pool of data from first-line BPDCN patients across all stages of the trial.

3.8. Efficacy Endpoint Analysis

3.8.1. Primary Endpoint

The primary evidence of efficacy is an analysis of the rate of CR, accompanied by an evaluation of the durability of response, including durability of these CRs, as a key secondary endpoint.

3.8.2. Secondary Endpoints

Formal hypothesis testing of secondary endpoints is not planned. The analyses of all secondary endpoints will be descriptive. The description of analyses to be performed by stage and pooling of stages is described in Section 3.5. Summaries of secondary endpoints to be evaluated and presented as follows:

1. BMCR and duration of BMCR
2. ORR (CR + CRc + CRi + partial response [PR]) and duration of objective response
3. Proportion of patients who receive an SCT
4. PFS
5. OS

Further detail on the definitions and analysis methods of primary and secondary efficacy endpoints is provided in Section 4.3.

3.8.3. Exploratory Endpoints

Exploratory endpoints will be assessed to further evaluate efficacy, dependent on the extent of data available. These endpoints will include several additional time to event endpoints, including TTCR, TTR, time to BMCR, TTP, and RFS. Additional exploratory analysis may compare baseline disease severity between Stages and assess the respective impacts of baseline disease severity and compartments of disease expression over time on outcome measures such as CR rate and ORR. In addition, an exploratory efficacy analysis will be performed to assess the rate and duration of CR as defined per the Study Protocol Amendment 9 (09 February 2017), including CR + CRc + CRi.

Further detail on the definitions and analysis methods of exploratory efficacy endpoints is provided in Section 4.3.

3.9. Subgroup Analyses

Efficacy analyses for Stage 4 and pooled Stages 3 and 4 as appropriate, will be summarized separately for those bridged to SCT and not bridged to SCT post-treatment. Efficacy analyses will also be summarized by baseline disease site (Skin only, Systemic only [Bone Marrow, Peripheral Blood, Lymph Node, or Visceral Disease], and Skin and Systemic).

Subgroup analyses will be performed for safety parameters in the FL and RR BPDCN populations, as specified in the below table. Subgroup analyses will be presented only when the minimum-sized subgroup classification includes at least 10% of the population.

Table 3-2: Subgroups and Definitions for Safety Analyses

Parameter	Description
Age	<65 vs. ≥65, <75 vs. ≥75 years of age
Sex	Male, Female
BMI	Obese: >30 kg/m^2 Overweight: 25-30 kg/m^2 Normal/Underweight: < 25 kg/m^2
Eastern Cooperative Oncology Group (ECOG) Performance Status	Baseline score of 0 Baseline score of 1 or 2

Subgroup analyses will be performed for the following safety summaries:

Overview of adverse events (AE)
Treatment emergent AEs by system organ class (SOC) and preferred term (PT) (occurring in at least 15% of population)

3.10. Withdrawals, Dropouts, Loss to Follow-up

Patients who are withdrawn or discontinue from the study will not be replaced.

3.11. Missing, Unused, and Spurious Data

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

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

3.12. Visit Windows

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

3.13. Timing of Analyses

3.13.1. Interim Analyses

A preliminary and informal interim analysis was performed on efficacy data from Stages 1 and 2 [redacted] This analysis was performed on the 32 BPDCN patients (19 first-line and 13 R/R) enrolled

in the study prior to 29 August, 2016, 3 of whom were treated with tagraxofusp 7 µg/kg/day and 29 were treated with tagraxofusp 12 µg/kg/day. The ORR, CR rate, PFS, and OS efficacy parameters were summarized. Selected safety summaries also were presented.

Note that no further enrollment of first-line BPDCN patients occurred until October 26, 2016. [redacted] first-line BPDCN patients enrolled on or after October 26, 2016 (with enrollment period later defined as between October 26, 2016 and March 17, 2017) will be considered as a separate, pivotal patient cohort for statistical analysis, as previously described in Section 1.2.1.

In Stages 2-4 of the study, the Data Safety Review Committee (DSRC) was to conduct a safety data review every 1-2-months.

3.13.2. Independent Review Committee

An Independent Review Committee (IRC) has been established to review tumor responses and duration of response for patients with BPDCN enrolled in Protocol STML-401-0114.

Response is best assessed by the treating investigator because disease-compartment responses, and in particular bone marrow responses, are used in the real-time management of the patient. It is also important to verify to the extent possible that the protocol-specified response criteria have been applied objectively and uniformly, therefore the IRC will apply the protocol specified response criteria to provide consistent and independent evaluation of tumor responses by cycle and overall. The IRC evaluation, which will be considered as supportive to the primary assessment of response by the investigator, was performed on 28-29 August 2017, prior to the primary efficacy, safety, PK, and immunogenicity analysis. The IRC is governed by a separate charter.

3.13.3. Primary Efficacy, Safety, and Immunogenicity Analysis

The primary analysis of efficacy, safety, and immunogenicity data occurred following the cutoff date of 25 September 2017, after all patients enrolled into Stage 3 completed at least 4 cycles of treatment with tagraxofusp, have experienced disease progression, or have withdrawn from the study. The primary analysis included efficacy, safety, PK, and immunogenicity analysis of Stage 3 data. In addition, supporting data from Stages 1 and 2 were analyzed at this time in support of the BLA.

3.13.4. Updated Reporting and Stage 4 Analysis

Following a data cutoff date of 13 March 2020 (database lock), updated reports are planned to support the ongoing efficacy and safety of Stages 1, 2, and 3 patients as well as to summarize the efficacy and safety of the lyophilized formulation of tagraxofusp, administered in Stage 4. A summary of efficacy, safety, PK, and immunogenicity analyses will be presented at this time.

The Updated Reporting and Stage 4 Analysis is described in this SAP.

4. STUDY ANALYSES

4.1. Patient Disposition

Patient disposition will be tabulated and include the number screened, the number treated in total, the number in each patient population for analysis, the number treated in each Stage of the study, the number receiving the frozen liquid and lyophilized drug formulation, the number on study at the time of data cutoff, the number who withdrew prior to completing the study and reason(s) for withdrawal. The summary will be presented by disease and line of therapy (first-line BPDCN, R/R BPDCN, AML) within Stage.

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

4.2. Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized and presented by disease and line of therapy (first-line BPDCN, R/R BPDCN, AML) within Stage. Age, height, and weight will be summarized using descriptive statistics (number of patients, mean, standard deviation, median, minimum, and maximum). The number and percentage of patients in each gender, ethnicity and race category will be presented.

Medical history and results from baseline screens such as for pregnancy, and smoking status will be tabulated. Primary disease history will be summarized including sites and extent of disease involvement, molecular markers, mutations, BPDCN baseline diagnosis (type of analysis, immunophenotype expression), and time since BPDCN diagnosis. Prior chemotherapy, prior immunotherapy, prior radiation therapy, and prior SCT will be summarized.

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

4.3. Efficacy Evaluation

Data for all efficacy endpoints will be presented in by-patient listings. Efficacy summaries will be presented by stage of enrollment and line of therapy. Subgroup analyses will be reported for Stage 4 (and Stages 3 and 4 if pooled) according to whether patients were bridged to SCT posttreatment and not bridged to SCT post-treatment. Efficacy analyses of responses will also be summarized by baseline disease site.

4.3.1. Response Rate

Complete response rate will be presented as the number and percentage of patients who achieved CR (CR + CRc) after treatment with tagraxofusp.

Bone marrow CR (BMCR) rate will be measured by pooling data from patients who had either CR or PR according to the IWG response criteria, together with BMCR (reduction to $\leq$5% BM blast cells), as a proportion of all patients with marrow disease (>5% BM blast cells) at baseline.

Objective response rate will be presented as the number and percentage of patients who achieve CR (CR + CRc), CRi, or PR after treatment.

For CR, BMCR, and ORR, 95% Clopper Exact CIs will be presented.

Permutation test results for Stage 3 and Stage 4 comparison will be reported as the observed results and relative percentile of the empirical null distribution for CR and ORR. Graphical plots may be presented.

4.3.2. Duration of Response

Duration of CR is defined as the time from when the criteria are first met for CR/CRc (whichever is recorded first) until the date that the criteria for relapse after CR/CRc are met. This could be the occurrence of PD or relapse. In the case that PR or SD follows CR/CRc and there is no evidence that response rebounds to CR/CRc, duration of CR will end at the time of first reduction of response to below CR/CRc. If CRi follows CR/CRc it will not be considered evidence of relapse. For patients who receive SCT after CR/CRc, duration of CR will include time to disease relapse post-transplant. Patients who are lost to follow-up or who do not relapse after CR/CRc as of the cut-off for analysis will be censored on the latter of the date of last treatment with tagraxofusp or date of last disease assessment recorded prior to the analysis cutoff date. Permutation test results for Stage 3 and Stage 4 comparison will be reported as the observed results and relative percentile of the empirical null distribution. Graphical plots may be presented.

Duration of BMCR is defined as the time from when criteria are first met for BMCR until the date that the criteria for relapse are met, including PR, SD, PD, or relapse in bone marrow compartment or SD, PD, or relapse in any other disease compartment. Duration of BMCR will include time to disease relapse post-transplant. Patients who are lost to follow-up or who do not relapse after BMCR as of the cut-off for analysis will be censored on the latter of the date of last treatment with tagraxofusp or date of last disease assessment recorded prior to the analysis cutoff date.

Duration of objective response is defined as the time from when criteria are first met for CR/CRc, CRi, or PR until the date that the criteria for relapse (including SD, PD, or relapse) after CR/CRc/CRi/PR is met. For patients who receive SCT after CR/CRc/CRi/PR, DOR will include time to disease relapse post-transplant. Patients who are lost to follow-up or who do not relapse after objective response as of the cut-off for analysis will be censored on the latter of the date of last treatment with tagraxofusp or date of last disease assessment recorded prior to the analysis cut-off date.

For duration of CR, BMCR, and ORR, a Kaplan-Meier analysis will be applied to estimate the 25^{th} percentile, median, and 75^{th} percentile of times to event, number and percentage of events and censored observations, and appropriate CIs.

4.3.3. Bridge to Stem Cell Transplant

The number and proportion of patients who receive SCT subsequent to achieving an investigator-determined outcome from tagraxofusp that is deemed by the investigator to enable a SCT, will be reported with 95% Clopper-Pearson Exact CIs. Summaries will be performed overall and by treatment outcome on tagraxofusp. Permutation test results for Stage 3 and Stage 4 comparison will be reported as the observed results and relative percentile of the empirical null distribution. Graphical plots may be presented.

4.3.4. Progression-Free Survival

Progression-free survival is defined as the time from the date of first infusion of tagraxofusp to the date of PD or death from any cause, whichever occurred first. For patients who receive SCT, PFS will include time to PD or death post-transplant. Patients who do not progress and are still alive at the time of analysis will be censored on the date of last treatment recorded prior to the analysis cut-off date. If patients start other anticancer therapies (with the exception of SCT), they will be censored at the latter of the date of last treatment with tagraxofusp or date of last disease assessment recorded that occurred prior to the start of the new therapy. The distribution for PFS will be estimated by Kaplan-Meier methodology and the 25th percentile, median, 75th percentile, number and percentage of events and censored observations, and appropriate CIs will be presented.

4.3.5. Overall Survival

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

4.3.6. Exploratory Endpoints

4.3.6.1. Time to Event Endpoints

Time to CR is defined as the time from initiation of treatment with tagraxofusp to CR (CR, or CRc, whichever is achieved first). Only patients who have achieved CR or CRc will be included in the analysis. Time to CR will be summarized using descriptive statistics.

Time to response is defined as the time from initiation of treatment with tagraxofusp to CR (CR + CRc), CRi, or PR, whichever occurs first. Only patients who achieve CR, CRc, CRi, or PR will be included in the analysis. The analysis of TTR will follow the same methods as described for TTCR.

Time to BMCR is defined as the time from initiation of treatment with tagraxofusp to BMCR, defined as the first evaluation of marrow blast cells ≤5%, for the patients who had evidence of BM disease (>5% BM blast cells) at baseline. Only patients who have BM disease at baseline and achieve BMCR will be included in the summary. The analysis of time to BM response will follow the same methods as described for TTCR.

Time to progression is defined as the time from first treatment with study drug disease progression. Patients who do not progress or die prior to response will be censored on the date of last disease assessment recorded prior to the analysis cut-off date. The distribution for TTP will be estimated by Kaplan-Meier methodology and the 25th percentile, median, 75th percentile, number and percentage of events and censored observations, and appropriate CIs will be presented.

Relapse free survival (RFS) is defined as the time from CR (CR/CRc) to disease progression or death. For patients who receive SCT after CR/CRc, RFS will include time to disease relapse post-transplant. Patients who do not achieve CR or die prior to CR will not be included in the analysis. Patients who have responded and are alive at the time of analysis will be censored on the latter of the date of last treatment with tagraxofusp or date of last disease assessment

recorded prior to the analysis cut-off date. The distribution for RFS will be estimated by Kaplan-Meier methodology and the 25th percentile, median, 75th percentile, number and percentage of events and censored observations, and appropriate CIs will be presented.

4.3.6.2. Protocol-Specified Rate of Complete Response

The Study Protocol Amendment 9 (09 February 2017)-defined rate and duration of CR, inclusive of CR, CRc, and CRi will be summarized by disease and line of therapy (first-line BPDCN, R/R BPDCN, AML).

4.3.6.3. Additional Analyses

Sites and extent of disease involvement, as reported by the investigators, will be listed by Stage over time. For continuous disease measures Stages may be compared using ANOVA regression. Frequencies of categorical disease measures may be compared between Stages using chi-square tests of independence.

For a subset of efficacy parameters, additional supportive analyses may be performed based on the IRC-determined assessments of disease response over time. Efficacy parameters may include CR rate and ORR, and will be analyzed in a similar manner as was described in Section 4.3.1 and Section 4.3.2.

4.4. Safety Analyses

With the exception of Subgroup Analyses (Section 3.9), all safety tabulations will be presented by disease (first-line BPDCN, R/R BPDCN, AML) for all patients in the Safety Population by stage of enrollment (all doses included). All safety tabulations will be presented for first-line BPDCN patients in the Safety Population who were treated at 12 µg/kg/day by stage of enrollment and for those treated with lyophilized formulation of tagraxofusp.

4.4.1. Study Drug Exposure

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

$$DDDDDDDDDDDDDDD\ DDoo\ SSDDDDuuuu\ DDDDDDDD\ EEEEEEDDEEDDDDEE = (DDDDDDEE\ DDoo\ llDDEEDD\ uuDDEEEE - DDDDDDEE\ DDoo\ ooDDDDEEDD\ uuDDEEEE) + 1$$

Duration in cycles of drug exposure will be defined the total number of cycles of study drug initiated while the patient is on study. Dose modifications, including frequency of dose interruptions and dose reductions will be summarized. Total dose administered will be summarized overall and by cycle.

Relative Dose Intensity will be computed using the following definition and summarized by disease and line of therapy by stage:

$$SSDDSS(CCDDSSDDllDDDDDDRREE\ AAAADDDDDDll\ DDDDEEEE\ RREEAAEEDDRREEuu\ bbuu\ CCuuAAllEE)$$

$$RREEllDDDDDDRREE\ DDDDEEEE\ IIDDDDEEDDEEDDDDuu\ (\%) = 100 *$$

$$SSDDSS(PPllDDDDDDEEuu\ DDDDEEEE\ DDDD\ bbEE\ AAuuSSDDDDDDEEDDEEDDEEuu\ bbuu\ CCuuAAllEE)$$

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

4.4.2. Adverse Events

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

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

The number and percentage of patients with any treatment-emergent AE (TEAE), common AEs (defined as occurring in ≥10% of the study sample), with any TEAE assessed by the Investigator as related to treatment, with TEAE occurring within 24 hours of infusion, with any TEAE with severity ≥ Common Terminology Criteria for Adverse Events (CTCAE) Grade 3, with any serious adverse event (SAE), with any AE leading to discontinuation of study treatment, with any AE to dose reduction, and with any AE leading to dose interruption will be summarized by disease, line of therapy, dose and overall. In these tabulations, each patient will contribute only once (ie, the most related occurrence or the most intense occurrence) to each of the incidence rates in the descriptive analysis, regardless of the number of episodes.

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

Incidence of TEAEs occurring within 24 hours of infusion will be summarized by SOC and preferred term.

Analysis of AEs over time will be conducted by presentation of AE incidence by cycle. For this analysis, incidence per cycle will be based on the onset of a particular preferred term event within the cycle, where onset is indicated by first or repeat occurrence or increase in severity of the event within that cycle. This analysis will be conducted for all AEs and for AEs of Grade ≥3 severity. For AEs of special interest, Kaplan-Meier analyses of time to first occurrence of Grade ≥2 severity will be conducted for FL BPDCN, R/R BPDCN, and AML patients.

AESI's will be determined using MedDRA version 19.0 standardized medical queries (SMQs), high-level terms, or preferred terms, as follows:

- Possible hypersensitivity events, based on the SMQ Hypersensitivity (broad search).

- Vascular capillary leak syndrome (CLS), based on the MedDRA SMQ [to be provided] in addition to preferred terms of hypoalbuminaemia, blood albumin decreased, and proteinuria.
- Visual acuity, based on the MedDRA preferred terms of visual acuity reduced, visual acuity reduced transiently, visual acuity tests abnormal.
- Post-transplantation veno-occlusive disease, based on the MedDRA preferred terms pulmonary veno-occlusive disease and veno-occlusive liver disease.
- Possible drug-induced liver injury events, based on the SMQ drug-related hepatic disorders (broad search).

Summary tables of AESIs produced will report AESI by preferred term, causality, severity grade, seriousness, and whether AESI led to discontinuation. These tabulations will be completed for all patients, by line of therapy for BPDCN patients, AML patients, and indication. AESIs will also be included in data listings.

4.4.2.1. Capillary Leak Syndrome

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

- Summary of number of CLS events, exposure to tagraxofusp prior to onset of CLS, and time to first onset and time to resolution of CSL events;
- Summarized incidence of albumin infusions and timing of infusions in relation to CLS onset;
- Plots of liver function and albumin laboratory values over time for patients who experience CLS;
- Comparison of demographic characteristics of patients who experience AESI of CLS to those who do not

Additionally, analyses of capillary leak syndrome will be provided based on an FDA algorithm, found in Section 8.3.

4.4.3. Laboratory Data

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

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

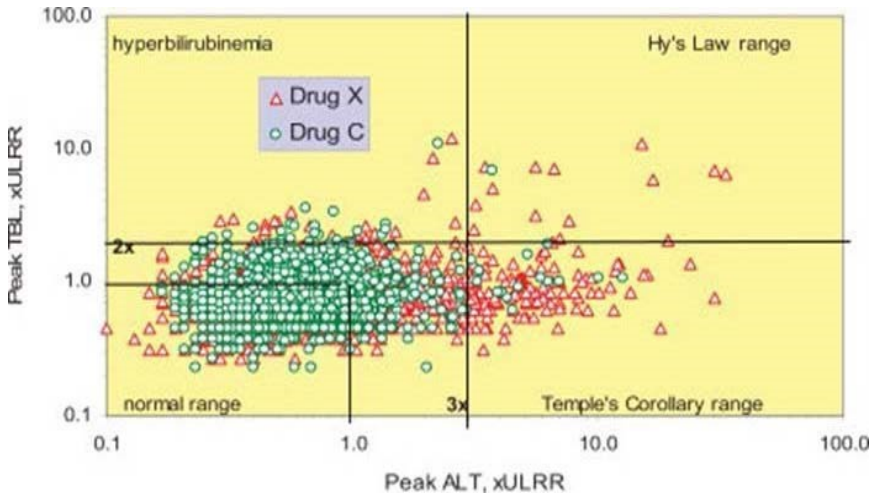
Shift tables of change in CTCAE grade of laboratory parameters from baseline to worst value and from baseline to last value on study will be presented. Both scheduled and unscheduled visits will be included in shift tables.

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

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

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

Figure 4-1: Hy's Law Candidates Scatterplot



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

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

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

4.4.4. Vital Signs and Physical Examination

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

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

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

PCS= potentially clinically significant.

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

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

4.4.5. Electrocardiogram

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

Based on 12-lead triplicate ECG results, the number and percentage of patients whose mean QTcF or QTcB value at any time point meets any of the following categories will be summarized:

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

Electrocardiogram and echocardiogram/MUGA data for each patient will be provided in data listings.

4.4.6. Concomitant Medications and Subsequent Treatments

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

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

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

4.4.7. Pre-Treatment Medications

Pre-treatment medications will be coded using the WHO Drug Dictionary. Results will be tabulated by ATC and preferred term. In addition, the number and percent of patients who receive pre-treatment medication at least once will be summarized for previously specified safety groups.

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

4.5. Summary of Immunogenicity Parameters

Immunogenicity parameters will be summarized for FL BPDCN, R/R BPDCN, all BPDCN, Stage 4 (lyophilized formulation) only, and all patients pooled across all doses and stages.

Analyses will be conducted in accordance with the principles outlined in (Shankar, et al. 2014) and will include:

- Incidence of patients with positive anti-drug antibodies, specific anti-hIL-3 antibodies, and neutralizing antibodies will be assessed by cycle uniquely through cycle 6 and cycles >6, collectively. Incidence will be summarized by number and percent of patients with positive antibody results.
- Mean fold increase from baseline and geometric mean titer are tabulated by scheduled assessments. Summary statistics will include mean, SD, median, min, and max.
- Incidence of subjects with a $\geq$ 10 fold increase in anti-drug antibody titer and/or a $\geq$ 4 fold increase in anti-hIL-3 antibody titer will be tabulated.
- Complete response, objective response, and complete response rates will be tabulated with 95% clopper exact CIs for subjects with a $\geq$ 10 fold increase in anti-drug antibody titer and/or a $\geq$ 4 fold increase in anti-hIL-3 antibody titer.
- Descriptive statistics will be provided for the count of AEs, AESIs, Immune-complex related AEs, serious AEs, and AEs leading to Discontinuation by antibody status and the fold increase populations previously described.

Immunogenicity parameters will be included in by-patient data listings.

5. CHANGES TO PLANNED ANALYSES

5.1. SAP Version 1.0 (28 April 2017)

In the Study Protocol (Amendment 9), in Section 12.6.3 Duration of Response is defined as the time from CR to disease progression. In case a patient does not progress, the protocol definition specifies that censoring will occur on the last radiologic assessment while on study. In this SAP Section 4.3.2 (Duration of Response) and Section 4.3.4 (PFS), censoring will occur on the date of last disease assessment, including BM, skin, or radiologic assessments, as appropriate and in accordance with the patient-specific disease presentation.

In the Study Protocol (Amendment 9), in Section 12.6.2 Bone Marrow Complete Response is defined as patients "who had either CR or PR with complete bone marrow response (reduction to <5% blast cells), as a proportion of all patients with marrow disease (≥5% blast cells) at baseline." Per the response criteria; however complete BM response requires a reduction to ≤5% blast cells. This notation was corrected in the SAP Section 4.3.1.

A footnote was added to the Response Criteria for BPDCN from Protocol (Amendment 9) to define marked clearance of skin disease and residual hyperpigmentation or abnormality identified on a biopsy (or no biopsy performed) as quantifiable changes to skin disease for identification of the achievement of CRc. (Section 8.2)

5.2. Changes from SAP Version 1.0 (28 April 2017) to Version 2.0

Changes from SAP Version 1.0 (28 April 2017) include:

- Section 1.1.3: [redacted] modification to the Study Protocol and SAP Version 1.0 definition to complete response from CR + CRc + CRi to include only CR + CRc. An assessment of rate and duration of CR + CRc + CRi was added as an exploratory analysis.

  In addition, the modification to the definition of CRc, including a quantifiable definition of "marked clearance" and "residual" were added to this SAP.

- Section 2: The population definitions were updated to improve clarity on their implementation in the analysis of the patient data.

- Section 3.5: Additional information was added to the Methods for Pooling Data to clarify the intent and implementation of analysis of the patient data.

- Section 3.9: Subgroup analyses were updated to better describe the analyses to be performed. The list of subgroups were modified to include only those relevant for analysis where at least 10% of patients fall within each group. In addition, the efficacy subgroup analyses were removed since SAP Version 1.0 as the analyses will not be performed.

- Section 3.14.3: The cutoff date for analysis is identified as 25 September 2017 to allow all surviving patients at least 6 months of follow-up.

- Section 4: Corrections and clarification of efficacy and safety analyses were added to the document.

- Corrections were made to exploratory time to event endpoint descriptions and planned analysis methods, including the addition of TTP as an exploratory endpoint.
- Time to event response endpoints of duration of CR, duration of BMCR, duration of OR, PFS, and RFS are censored on the latter of date of last treatment with tagraxofusp or date of last disease assessment prior to data cutoff. ○ In addition to other clarifications, the AESI section is updated for greater detail of the particular events of interest.
- Additional specification of CLS outputs to be included in CSR.

5.3. Changes from the SAP Version 2.0 (17 November 2017) to Version 3.0

Changes from SAP Version 2.0 (17 November 2017) include:

- Section 1.1: The purpose of this document was updated for an updated report on the clinical efficacy and safety of tagraxofusp, emphasizing both the Stage 4 results that were not previously reported and the overall conclusions from the duration of the trial. Details are included in Section 3.13.3 and 3.13.4.
- Section 1.2: Details on stages previous to Stage 4 were removed. Additional information on responses to be used in efficacy analyses were detailed in Section 1.2.5.1.
- Section 2.1: The Per-Protocol Population was removed. Sensitivity analyses were also removed.
- Section 3.5: Methods of pooling data were updated to emphasize Stage 4, including a pooling of Stages 3 and 4 (a permutation test described in Section 4.3.1, 4.3.2, and 4.3.3).
- Section 3.8: The efficacy endpoints were updated to emphasize stage 4 results and overall conclusions.
- Section 3.9: Subgroup analyses for efficacy among subjects that bridged to SCT and Not Bridged to SCT, as well as by baseline disease site, were added.
- Section 4.4.2.1: Analyses of CLS based on FDA algorithm were added.

6. REFERENCES

Adachi M, Maeda K, Takekawa M, Hinoda Y, Imai K, Sugiyama S, et al. High expression of CD56 (N-CAM) in a patient with cutaneous CD4-positive lymphoma. *Am J Hematol.* 1994;47(4):278-82.

Aoki T, Suzuki R, Kuwatsuka Y, Kako S, Fujimoto K, Taguchi J, et al. Long-term survival following autologous and allogeneic stem cell transplantation for blastic plasmacytoid dendritic cell neoplasm. *Blood.* 2015;125(23):3559-62.

Assaf C, Gellrich S, Whittaker S, Robson A, Cerroni L, Massone C, et al. CD56-positive haematological neoplasms of the skin: a multicentre study of the Cutaneous Lymphoma Project Group of the European Organisation for Research and Treatment of Cancer. *J Clin Pathol.* 2007;60(9):981-9.

Bekkenk MW, Jansen PM, Meijer CJ, and Willemze R. CD56+ hematological neoplasms presenting in the skin: a retrospective analysis of 23 new cases and 130 cases from the literature. *Ann Oncol.* 2004;15(7):1097-108.

Brody JP, Allen S, Schulman P, Sun T, Chan WC, Friedman HD, et al. Acute agranular CD4positive natural killer cell leukemia. Comprehensive clinicopathologic studies including virologic and in vitro culture with inducing agents. *Cancer.* 1995;75(10):2474-83.

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

Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25(5):579-86.

Dalle S, Beylot-Barry M, Bagot M, Lipsker D, Machet L, Joly P, et al. Blastic plasmacytoid dendritic cell neoplasm: is transplantation the treatment of choice? *Br J Dermatol.* 2010;162(1):74-9.

Facchetti F, Vermi W, Mason D, and Colonna M. The plasmacytoid monocyte/interferon producing cells. *Virchows Arch.* 2003;443(6):703-17.

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

Feuillard J, Jacob MC, Valensi F, Maynadie M, Gressin R, Chaperot L, et al. Clinical and biologic features of CD4(+)CD56(+) malignancies. *Blood.* 2002;99(5):1556-63.

Herling M, and Jones D. CD4+/CD56+ hematodermic tumor: the features of an evolving entity and its relationship to dendritic cells. *Am J Clin Pathol.* 2007;127(5):687-700.

Jacob MC, Chaperot L, Mossuz P, Feuillard J, Valensi F, Leroux D, et al. CD4+ CD56+ lineage negative malignancies: a new entity developed from malignant early plasmacytoid dendritic cells. *Haematologica.* 2003;88(8):941-55.

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

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

Kharfan-Dabaja MA, Lazarus HM, Nishihori T, Mahfouz RA, and Hamadani M. Diagnostic and therapeutic advances in blastic plasmacytoid dendritic cell neoplasm: a focus on hematopoietic cell transplantation. *Biol Blood Marrow Transplant.* 2013;19(7):1006-12.

Leclerc M, Peffault de Latour R, Michallet M, Blaise D, Chevallier P, Rohrlich PS, et al. Can a reduced-intensity conditioning regimen cure blastic plasmacytoid dendritic cell neoplasm? *Blood.* 2017;129(9):1227-1230.

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

Olsen EA, Whittaker S, Kim YH, Duvic M, Prince HM, Lessin SR, et al. Clinical end points and response criteria in mycosis fungoides and Sezary syndrome: a consensus statement of the International Society for Cutaneous Lymphomas, the United States Cutaneous Lymphoma Consortium, and the Cutaneous Lymphoma Task Force of the European Organisation for Research and Treatment of Cancer. *J Clin Oncol.* 2011;29(18):2598-607.

Pemmaraju N. Blastic plasmacytoid dendritic cell neoplasm. *Clin Adv Hematol Oncol.* 2016;14(4):220-2.

Pemmaraju NK, H.; Cortes, J.E.; Duvic, M.; Khoury, J.D.; Patel, K.; Daver, N.G.; O'Brien, S.M.; Pierce, S.; Garcia-Manero, G.; Jabbour, E; Jain, N.; Faderl, S.; Thomas, D.A.; Frankel, A.E.; Qazilbash, M.H.; Konopleva, M.; . 2015 Blastic Plasmacytoid Dendritic Cell Neoplasm (BPDCN): A Large Single-Center Experience: Analysis of Clinical and Molecular Characteristics and Patient Outcomes. ASH, Orlando, FL, 2015.

Pemmaraju NT, D.A.; Kantarjian, H.; Cortes, J.E.; Khoury, J.; Manning, J.T.; Medeiros, L.J.; O'Brien, S.M.; Daver, N.G.; Nazha, A.; Pierce, S.; Garcia-Manero, G.; Jabbour, E; Jain, N.; Qazilbash, M.H.; Faderl, S.; Frankel, A.E.; Konopleva, M. 2014 Characteristics of patients with blastic plasmacytoid dendritic cell neoplasm (BPDCN): Male predominance, propensity for extramedullary involvement, and poor outcomes. ASCO Annual Meeting, 2014. Vol. 32:5s. J Clin Oncol.

Pemmaraju NT, D.A.; Kantarjian, H.; O'Brien, S.M.; Daver, N.G.; Nazha, A.; Pierce, S.; Garcia-Manero, G.; Cortes, J.E.; Faderl, S. 2012 Analysis of outcomes of patients (pts) with blastic plasmacytoid dendritic cell neoplasm (BPDCN). ASCO Annual Meeting, 2012. Vol. 30. J Clin Oncol.

Petrella T, Bagot M, Willemze R, Beylot-Barry M, Vergier B, Delaunay M, et al. Blastic NKcell lymphomas (agranular CD4+CD56+ hematodermic neoplasms): a review. *Am J Clin Pathol.* 2005;123(5):662-75.

Petrella T, Dalac S, Maynadie M, Mugneret F, Thomine E, Courville P, et al. CD4+ CD56+ cutaneous neoplasms: a distinct hematological entity? Groupe Francais d'Etude des Lymphomes Cutanes (GFELC). *Am J Surg Pathol.* 1999;23(2):137-46.

Poret E, et al. 2015 How to Treat Blastic Plasmacytoid Dendritic Cell Neoplasm (BPDCN) Patients: Results on 86 Patients of the French BPDCN Network. American Society of Hematology, Orlando, FL, 2015.

Reichard KK, Burks EJ, Foucar MK, Wilson CS, Viswanatha DS, Hozier JC, et al. CD4(+) CD56(+) lineage-negative malignancies are rare tumors of plasmacytoid dendritic cells. *Am J Surg Pathol.* 2005;29(10):1274-83.

Riaz W, Zhang L, Horna P, and Sokol L. Blastic plasmacytoid dendritic cell neoplasm: update on molecular biology, diagnosis, and therapy. *Cancer Control.* 2014;21(4):279-89.

Roos-Weil D, Dietrich S, Boumendil A, Polge E, Bron D, Carreras E, et al. Stem cell transplantation can provide durable disease control in blastic plasmacytoid dendritic cell neoplasm: a retrospective study from the European Group for Blood and Marrow Transplantation. *Blood.* 2013;121(3):440-6.

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

Shapiro R, Sangle, N., Keeney, M., Chin-Yee, I.H., Hsia, C.C., et al. Blastic Plasmacytoid Dendritic Cell Neoplasm: A Review of Diagnosis, Pathology, and Therapy. *J Cell Sci Ther.* 2015;S8:008:2157-7013.

Suzuki R, Nakamura S, Suzumiya J, Ichimura K, Ichikawa M, Ogata K, et al. Blastic natural killer cell lymphoma/leukemia (CD56-positive blastic tumor): prognostication and categorization according to anatomic sites of involvement. *Cancer.* 2005;104(5):1022-31.

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

Wang JP, Zhang L, Madera RF, Woda M, and Libraty DH. Plasmacytoid dendritic cell interferon-alpha production to R-848 stimulation is decreased in male infants. *BMC Immunol.* 2012;13:35.

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

7. STUDY FLOW CHARTS

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

Table 7-1: Study Events Schedule for Cycle 1 (Study Day -14 to Study Day 21): Stage 1 (All Pts) & Stage 2 (AML Pts: Arm A)

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

Tumor response assessment: CT scan(k)	X						
Tumor response assessment: Skin, including biopsies and/or photographs(l)	X					X	
Tumor response assessment: Peripheral blood						X	
Translational assessment: Peripheral blood (m)			X (Infusions 1, 5)		X	X	
Administration of premeds (n)			X				
Tagraxofusp administration(o)				X			
		Cycle 1					
	Study Day -14 to -4	**Study Day -1 to 0**	**Study Days 1-5 (Up to Study Day 10 if Infusion(s) Held) tagraxofusp Treatment**			**Study Day 21±7**	**Study Day 28±7, Then Every 7±3 Days**
Tests and Observations	***Screening***	***Pre-treatment***	***Pre-Infusion***	***Infusion***	**Study Day 8±3[q] and 15±3**	***End of Cycle***	***Delayed End of Cycle for Toxicity Resolution Only if Required***
Pharmacokinetic sampling[p]			X (Infusions 1, 5)	X (Infusions 1, 5)			
Immunogenicity sampling[q]			X (Infusion 1)		X (Day 15)	X	
AE and SAE monitoring			X	X	X	X	X

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

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

(c) All patients will have a 12-lead Electrocardiogram (ECG) performed at the screening visit and pre-treatment visit, as well as Day 21 (and Day 28 only if delayed end of cycle) of each cycle. During the days when patients are undergoing PK sampling (Cycle 1, infusions 1 and 5; Cycle 3, infusions 1 and 5), an ECG will be performed at 3 distinct time points (triplicates) within 5 minutes prior to each PK sample collection pre-infusion and at 30 and 60 minutes post-infusion (see (p) and Table 10 in the protocol).

(d) Multigated Acquisition Scan (MUGA) or 2-D ECHO to quantify left ventricular ejection fraction (LVEF). Must be completed within 28 days prior to start of first cycle of study treatment.

(e) To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes White blood cell (WBC) count, differential white cell count, red blood cell count, hematocrit, hemoglobin and platelet count.

(f) To be collected prior to tagraxofusp infusion if during dosing period. Electrolytes include sodium, potassium, bicarbonate, chloride, blood urea nitrogen (BUN), and creatinine.

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

(h) To be collected prior to tagraxofusp infusion if during dosing period. Serum electrolytes and chemistry: Sodium, potassium, bicarbonate, chloride, BUN, creatinine, glucose, alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), aspartate aminotransferase (AST), bilirubin (total, direct, and indirect), calcium, creatine phosphokinase (CPK), magnesium, lactate dehydrogenase (LDH), phosphate, total protein, uric acid

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

(j) Morphology and differential WBC/blast count on aspirate. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. Subsequent bone marrow aspirates (± biopsy) will be performed 21 (±7) days after the start of Cycles 1 and 2, and at the Investigator's discretion thereafter until there is evidence of relapsed or progressive disease. If the Cycle 1 (Day 21 [±7]) bone marrow aspirate (±biopsy) is empty, hypocellular, or inadequate, a bone marrow examination should be repeated within 7 (±7) days to document response. During Stage 2, AML patients (and BPDCN patients with evidence of bone marrow involvement prior to study treatment) will also have bone marrow evaluations following Cycles 4 and 6, and at the Investigator's discretion prior to end of treatment and thereafter. Any additional bone marrow aspirate (±biopsy) remaining after the sample is aliquoted for tumor assessment will be collected for translational/correlative assessment. In general, the bone marrow aspirate will be prioritized for tumor assessment over translational/correlative assessment. However, submission of bone marrow aspirate material for translational/correlative assessment (5 mL, larger when feasible) is required both for a pre-treatment (baseline) specimen, and at the end of Cycle 1. If CD123 flow cytometry or IHC stain was performed on the bone marrow aspirations (± biopsy), the results should be recorded in the eCRF.

(k) BPDCN patients only. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. Subsequent CT scans will be performed 21 (±3) days after the start of Cycle 2 and 4 and 21 (±7) days after the start of every 4th cycle thereafter until there is evidence of relapsed or progressive disease. Baseline CT scans should be full-body, whereas follow-up scans should document response of index lesions.

(l) BPDCN patients only, with biopsies and/or photographs as indicated, for patients with skin involvement. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. Subsequent skin assessments will be performed 21 (±3) days after the start of Cycles 1 and 2 and 21 (±7) days after the start of every 2nd cycle thereafter (Cycles 4, 6, etc.) until there is evidence of relapsed or progressive disease.

(m) Cycle 1 and 3, and at end of treatment, Translational Assessment should be collected at Infusion 1 (Day 1), Infusion 5 (Day 5, 6, 7; if Infusion 5 does not occur by Day 8, please obtain this specimen on Day 8) and days 15, 21 (±3 days, but must occur prior to start of subsequent treatment cycle). (n) Refer to protocol Section 7.5.3.1 – Premedication

(o) Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion for the first 5 consecutive days of a 21-day cycle. Individual tagraxofusp infusions may be delayed to allow for toxicity resolution, all infusions should be completed within 10 days and fewer than 5 infusions are permitted in settings of incipient CLS or hepatotoxicity. Patients must be monitored for 4 hours post infusion.

(p) Plasma samples (6 mL each) will be collected immediately prior to the start of the infusion of tagraxofusp, immediately after end of infusion (time recorded), then 15, 30, 45, 60, 90, 120, 180, and 240 minutes after completion of the infusion during infusions 1 (ie, Study Day 1) and 5 (ie, Study Day 5) (see (c) and Table 10 in the protocol).

(q) Blood samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies according on Day 1 (pre-infusion), Day 15, and Day 21 (if patient is to receive Cycle 2, can be Day 1, pre-infusion for Cycle 2). If infusions are held, please collect a sample within 3±3 days after the last tagraxofusp infusion(s) of the cycle.

Table 7-2: Study Events Schedule for Cycles 2-6 and Subsequent Cycles: Stage 1 (All Pts) & Stage 2 (AML Pts: Arm A)

	Cycle 2+							
	Days 1-5 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment		Day 8±3(q) and 15±3	Day 21±3	Day 28±3, Then Every 7±3 days	End of Treatment	Safety: Through 30 Days After Last Infusion	Survival: Every 90 Days After Last Infusion
Tests and Observations	***Pre-Infusion***	***Infusion***		***End of Cycle***	***Delayed End of Cycle for Toxicity Resolution Only if Required***			
Medical history including prior therapy, concomitant medications								
Concomitant Medications			X	X	X			
ECOG performance status				X	X	X		
Physical examination			X	X		X		
Vital signs and weight[a]	X	X	X	X	X	X		
12-lead ECG[b]	X	X (Cycle 3, Infusions 1, 5)		X	X			
Hematology[c]	X		X	X	X			
Serum electrolytes(d)	X		X	X	X			
Serum albumin[e]	X		X	X	X			
Serum chemistry[f]	X		X	X	X			
Coagulation parameters: PT/INR, aPTT	X		X	X	X			
Urinalysis(g)	X (Infusion 1)		X	X	X			
Tumor response assessment: Bone marrow aspiration ± biopsy(h)				X (h)				
Tumor response assessment: CT scan(i)				X				
Tumor response assessment: skin, including biopsies and/or photographs(j)				X				

Tests and Observations	Cycle 2+ Days 1-5 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment: *Pre-Infusion*	Cycle 2+ Days 1-5 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment: *Infusion*	Cycle 2+ Day 8±3[q] and 15±3	Cycle 2+ Day 21±3: *End of Cycle*	Cycle 2+ Day 28±3, Then Every 7±3 days: *Delayed End of Cycle for Toxicity Resolution Only if Required*	End of Treatment	Safety: Through 30 Days After Last Infusion	Survival: Every 90 Days After Last Infusion
Tumor response assessment: peripheral blood				X [k]				
Translational assessment: Peripheral blood (l)	X (Infusions 1,5)		X	X		X		
Administration of premeds (m)	X							
Tagraxofusp administration(n)		X						
Pharmacokinetic sampling[o]	X (Cycle 3, Infusions 1, 5)	X (Cycle 3, Infusions 1, 5)						
Immunogenicity sampling[p]	X (Infusion 1)			X				
AE and SAE monitoring	X	X	X	X	X		X	
Long-term follow-up(q)							X	X

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

(b) All patients will have a 12-lead ECG performed at the screening visit and pre-treatment visit, as well as Day 21 (and Day 28 only if delayed end of cycle) of each cycle. During the days when patients are undergoing PK sampling (Cycle 1, infusions 1 and 5; Cycle 3, infusions 1 and 5), an ECG will be performed at 3 distinct time points (triplicates) within 5 minutes prior to each PK sample collection pre-infusion and at 30 and 60 minutes post-infusion (see footnote (n) and Table 10 in the protocol).

(c) To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count, differential white cell count, red blood cell count, hematocrit, hemoglobin and platelet count.

(d) To be collected prior to tagraxofusp infusion if during dosing period. Electrolytes include sodium, potassium, bicarbonate, chloride, BUN, and creatinine.

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

(f) To be collected prior to tagraxofusp infusion if during dosing period. Serum electrolytes and chemistry: Sodium, potassium, bicarbonate, chloride, BUN, creatinine, glucose, ALT, ALP, AST, bilirubin (total, direct, and indirect), calcium, CPK, magnesium, LDH, phosphate, total protein, uric acid

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

(h) Morphology and differential WBC/blast count on aspirate. Bone marrow aspirates (± biopsy) will be performed 21 (±7) days after the start of Cycle 2, and, in Stage 1, at the Investigator's discretion prior to end of treatment and thereafter until there is evidence of relapsed or progressive disease. In Stage 1, submission of aspirate (5 mL or greater) for translational/correlative assessment at timepoints subsequent to Cycle 1 is strongly encouraged, whenever feasible. During Stage 2, AML patients (and BPDCN patients with evidence of bone marrow involvement prior to study treatment) will also have bone marrow evaluations following Cycles 4 and 6 and then every 3 months from Months

6-12; every 6 months from 12 to 24 months; and every 12 months thereafter until there is evidence of relapsed or progressive disease. In Stage 2, any additional bone marrow aspirate (± biopsy) remaining after the sample is aliquoted for tumor assessment will be collected for translational/correlative assessment. The bone marrow aspirate will be prioritized for tumor assessment over translational/correlative assessment. If CD123 flow cytometry or IHC stain was performed on the bone marrow aspirations (± biopsy), the results should be recorded in the eCRF.

(i) BPDCN patients only. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. In Stage 1, subsequent CT scans will be performed 21 (±7) days after the start of Cycle 2 and 4 and 21 (±7) days after the start of every 4th cycle thereafter until there is evidence of relapsed or progressive disease. Baseline CT scans should be full-body, whereas follow-up scans should document response of index lesions.

(j) BPDCN patients only, with biopsies and/or photographs as indicated, for patients with skin involvement. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. Subsequent skin assessments will be performed 21 (±7) days after the start of Cycles 1 and 2 and 21 (±7) days after the start of every 2nd cycle thereafter (Cycles 4, 6, etc.) until evidence of relapsed or progressive disease. Quantification of disease burden via the mSWAT is required at the time of each skin assessment.

(k) Patients with evidence of peripheral blasts at baseline will have peripheral blood samples collected 21 (±7) days after the start of Cycles 1 and 2 and 21 (±7) days after the start of every 2nd cycle thereafter (Cycles 4, 6, etc.) until there is evidence of relapsed or progressive disease.

(l) Cycle 1 and 3, and at end of treatment, Translational Assessment should be collected at Infusion 1 (Day 1) and Infusion 5 (Day 5, 6, 7; if Infusion 5 does not occur by Day 8, please obtain this specimen on Day 8) and Days 15, 21 (±3 days, but must occur prior to start of subsequent treatment cycle). (m) Refer to protocol Section 7.5.3.1 – Premedication

(n) Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion for the first 5 consecutive days of a 21-day cycle. Individual tagraxofusp infusions may be delayed to allow for toxicity resolution; all infusions should be completed within 10 days and fewer than 5 infusions are permitted in settings of incipient CLS or hepatotoxicity. Patients must be monitored for 4 hours post infusion.

(o) Plasma samples (6 mL each) will be collected immediately prior to the start of the infusion of tagraxofusp, immediately after end of infusion (time recorded), then 15, 30, 45, 60, 90, 120, 180, and 240 minutes after the start of the infusion during infusions 1 (ie, Day 1) and 5 (ie, Day 5) of Cycle 3 (see footnote (b) and Table 10 in the protocol).

(p) Blood samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies according on Day 1 (pre-infusion), Day 15, and Day 21 (if patient is to receive additional cycles, can be Day 1, pre-infusion for next cycle). If infusions are held, please collect a sample within 3±3 days the last tagraxofusp infusion(s) of the cycle.

(q) After the follow-up visit, patients will then be followed every 90 days for survival status. The survival follow-up may be by telephone contact. If the patient is in CR/PR at the time of discontinuation, disease assessments should continue to be performed as described in protocol Section 8.11 on an every 6-week basis (±1 week) through 6 months post-C1D1 and then on an every 90-day basis or until, in the judgment of the Investigator, there is evidence of relapsed or progressive disease.

Table 7-3: Study Events Schedule for Cycles 1 & 2 (Study Day -14 to Study Day 21): BPDCN (Arm B) Patients in Stage 2

		Cycles 1 & 2					
	Study Day -14 to -4	**Study Day -1 to 0**	**Cycle Days 1-5 (Up to Study Day 10 if Infusion(s) Held) Tagraxofusp Treatment**			**Cycle Day 21±7**	**Cycle Day 28±7, Then Every 7±3 Days**
Tests and Observations	***Screening***	***Pre-treatment***	***Pre-Infusion***	***Infusion***	**Study Day 8±3[q] and 15±3**	***End of Cycle***	***Delayed End of Cycle for Toxicity Resolution Only if Required***
Informed consent form	X						
Inclusion/exclusion criteria	X						
Medical history including prior therapy, concomitant medications	X						
Concomitant Medications		X			X	X	X
ECOG performance status	X					X	X
Physical examination	X	X			X	X	
Pregnancy test [a]	X						
Vital signs and weight[b]	X	X	X	X	X	X	X
12-lead ECG[c]	X	X	X	X (Infusions 1, 5)		X	X
ECHO or MUGA scan[d]	X						
Hematology[e]	X	X	X		X	X	X
Serum electrolytes(f)	X	X	X		X	X	X
Serum albumin[g]	X	X	X		X	X	X
Serum chemistry[h]	X	X	X		X	X	X
Coagulation parameters: PT/INR, aPTT	X		X		X	X	X
Urinalysis(i)	X		X (Infusion 1)		X	X	X
Tumor response assessment: Bone marrow aspiration ± biopsy(j)	X					X	
Tumor response assessment: CT scan(k)	X					X (Cycle 2) (k)	

Tumor response assessment: Skin, including biopsies and/or photographs(l)	X					X	
Tumor response assessment: Peripheral blood						X	
Translational assessment: Peripheral blood (m)							
Administration of premeds (n)			X				
Tagraxofusp administration(o)				X			
Pharmacokinetic sampling(p)			X (Infusions 1, 5)	X (Infusions 1, 5)			
		Cycles 1 & 2					
	Study Day -14 to -4	**Study Day -1 to 0**	**Cycle Days 1-5 (Up to Study Day 10 if Infusion(s) Held) Tagraxofusp Treatment**			**Cycle Day 21±7**	**Cycle Day 28±7, Then Every 7±3 Days**
Tests and Observations	***Screening***	***Pre-treatment***	***Pre-Infusion***	***Infusion***	**Study Day 8±3(q) and 15±3**	***End of Cycle***	***Delayed End of Cycle for Toxicity Resolution Only if Required***
Immunogenicity sampling(q)			X (Infusion 1)		X (C1D15 only)	X	
Vision Assessment	X						
AE and SAE monitoring			X	X	X	X	X

(a) Urine or serum pregnancy test must be performed within 1 week prior to treatment in women of childbearing potential.
(b) Vital signs should be performed after patient is sitting for 3-5 minutes. If during dosing period, vital signs should be taken immediately prior to infusion, immediately after completion of infusion, and 30, 60, and 240 minutes post-infusion.
(c) All patients will have a 12-lead ECG performed at the screening visit and pre-treatment visit, as well as Day 21 (and Day 28 only if delayed end of cycle) of each cycle. During the days when patients are undergoing PK sampling (Cycle 1, infusions 1 and 5; Cycle 3, infusions 1 and 5), an ECG will be performed at 3 distinct time points (triplicates) within 5 minutes prior to each PK sample collection pre-infusion and at 30 and 60 minutes post-infusion (see (p) and Table 10 in the protocol).
(d) MUGA or 2-D ECHO to quantify LVEF. Must be completed within 28 days prior to start of first cycle of study treatment.
(e) To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count, differential white cell count, red blood cell count, hematocrit, hemoglobin and platelet count.
(f) To be collected prior to tagraxofusp infusion if during dosing period. Electrolytes include sodium, potassium, bicarbonate, chloride, BUN, and creatinine.
(g) To be collected prior to tagraxofusp infusion if during dosing period. Serum albumin may be a component of the chemistry panel (h). See protocol for administration of albumin if serum albumin decreases to <3.0 mg/dL during treatment days or in the immediate post-treatment period.
(h) To be collected prior to tagraxofusp infusion if during dosing period. Serum electrolytes and chemistry: Sodium, potassium, bicarbonate, chloride, BUN, creatinine, glucose, ALT, albumin, ALP, AST, bilirubin (total, direct, and indirect), calcium, CPK, magnesium, LDH, phosphate, total protein, uric acid

(i) To be collected prior to initial tagraxofusp infusion of the cycle. Urinalysis includes appearance, color, pH, specific gravity, ketones, leukocytes, protein, glucose, bilirubin, urobilinogen, and occult blood.

(j) Morphology and differential WBC/blast count on aspirate. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. For BPDCN patients in Stage 2, subsequent bone marrow aspirates (± biopsy) will be performed 21 (±7) days after the start of Cycles 1 and 2. If the Cycle 1 (Day 21 [±7]) bone marrow aspirate (± biopsy) is empty, hypocellular, or inadequate, a bone marrow examination should be repeated within 7 (±7) days to document response. If CD123 flow cytometry stain was performed on the bone marrow aspirations (± biopsy), the results should be recorded in the eCRF. Submission of aspirate (5 mL or greater) for translational/correlative assessment is required for all patients with evidence of bone marrow involvement prior to study treatment at 1) pre-treatment baseline, and 2) end of Cycle 1. Any additional bone marrow (± biopsy) remaining after the sample is aliquoted for tumor assessment will be collected for exploratory translational/correlative assessments. The bone marrow aspirate will be prioritized for tumor assessment (by the Investigator) over translational/correlative assessments.

(k) Baseline CT must be performed within 14 days prior to the first administration of tagraxofusp. For patients with baseline evidence of lymph node or visceral disease involvement, subsequent Stage 2 CT scans will be performed 21 (±7) days after the start of Cycles 2, 4 and 6, and 21 (±7) days after the start of every 4th cycle thereafter (approximately every 12 weeks) until there is evidence of relapsed or progressive disease. For patients with no baseline evidence of lymph node or visceral BPDCN involvement, subsequent scans should be performed at the end of Cycle 2 (±7 days; or at time of disease progression if PD occurs prior to end of Cycle 2), 21 (±7) days after the start of Cycle 6, and at Investigator's discretion thereafter. Baseline CT scans should be full-body, whereas follow-up scans should document response of index lesions.

(l) For patients with skin involvement. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. Subsequent skin assessments will be performed 21 (±7) days after the start of Cycles 1 and 2 and 21 (±7) days after the start of every 2nd cycle thereafter (Cycles 4, 6, etc.) until there is evidence of relapsed or progressive disease. Quantification of skin disease burden via the mSWAT instrument is required at the time of each skin assessment.

(m) In Cycle 1 and 3, and at end of treatment, Translational Assessment should be collected at Infusion 1 (Day 1), Infusion 5 (Day 5, 6, 7; if Infusion 5 does not occur by Day 8, please obtain this specimen on Day 8) and days 15, 21 (±3 days, but must occur prior to start of subsequent treatment cycle). (n) Refer to protocol section 7.5.3 – Premedication and Administration

(o) Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion for the first 5 consecutive days of a 21-day cycle. Individual tagraxofusp infusions may be delayed to allow for toxicity resolution; all infusions should be completed within 10 days and fewer than 5 infusions are permitted in settings of incipient CLS or hepatotoxicity. Patient must be monitored for 4 hours post infusion.

(p) Plasma samples (6 mL each) will be collected immediately prior to the start of the infusion of tagraxofusp, immediately after end of infusion (time recorded), then 15, 30, 45, 60, 90, 120, 180, and 240 minutes after completion of the infusion during infusions 1 (ie, Study Day 1) and 5 (ie, Study Day 5) of Cycles 1 & 3 (see Table 10 in the protocol).

(q) Blood samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies on Day 1 (pre-infusion), Day 15, and Day 21 (if patient is to receive Cycle 2, can be Day 1, pre-infusion for Cycle 2). If infusions are held, please collect a sample within 3±3 days after completion of the last tagraxofusp infusion of the cycle.

Table 7-4: Study Events Schedule for Cycles 3-6 and Subsequent Cycles: Stage 2 (BPDCN Patients)

	Cycle 3+							
	Days1-5 (Up to Day 10 if Infusion(s) Held) Tagraxofusp Treatment			**Day 21±7**	**Day 28±7, Then Every 7±3 days**			
Tests and Observations	***Pre-Infusion***	***Infusion***	**Day 8±3(q) and 15±3(q)**	***End of Cycle***	***Delayed End of Cycle for Toxicity Resolution Only if Required***	**End of Treatment**	**Safety: Through 30 Days After Last Infusion**	**Survival: Every 90 Days After Last Infusion**
Medical history including prior therapy, concomitant medications								
Concomitant Medications			X[r]	X	X			
ECOG performance status				X	X	X		
Physical examination				X		X		
Vital signs and weight[(a)]	X	X		X	X	X		
12-lead ECG[(b)]	X	X (Cycle 3, Infusions 1, 5)		X	X			
Hematology[(c)]	X		X[r]	X	X			
Serum electrolytes(d)	X		X[r]	X	X			
Serum albumin[(e)]	X		X[r]	X	X			
Serum chemistry[(f)]	X		X[r]	X	X			
Coagulation parameters: PT/INR, aPTT	X		X[r]	X	X			
Urinalysis(g)	X (Infusion 1)		X[r]	X	X			
Tumor response assessment: Bone marrow aspiration ± biopsy(h)				X(i)				
Tumor response assessment: CT scan(i)				X				
Tumor response assessment: skin, including biopsies and/or photographs(j)				X				

Tumor response assessment: peripheral blood				X(k)				
Translational assessment: Peripheral blood (l)	X(Infusions 1,5)			X		X		
Administration of premeds (m)	X							
Tagraxofusp administration(n)		X						
Pharmacokinetic sampling(o)	X (Cycle 3, Infusions 1, 5)	X (Cycle 3, Infusions 1, 5)						
Immunogenicity sampling(p)	X (Infusion 1)			X				
AE and SAE monitoring	X	X	Xr	X	X		X	
Long-term follow-up(q)							X	X

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

(b) All patients will have a 12-lead ECG performed at the screening visit and pre-treatment visit, as well as Day 21 (and Day 28 only if delayed end of cycle) of each cycle. During the days when patients are undergoing PK sampling (Cycle 1, infusions 1 and 5; Cycle 3, infusions 1 and 5), an ECG will be performed at 3 distinct time points (triplicates) within 5 minutes prior to each PK sample collection pre-infusion and at 30 and 60 minutes post-infusion (see (m) and Table 10 in the protocol).

(c) To be collected prior to tagraxofusp infusion if during dosing period. Hematology includes WBC count, differential white cell count, red blood cell count, hematocrit, hemoglobin and platelet count.

(d) To be collected prior to tagraxofusp infusion if during dosing period. Electrolytes include sodium, potassium, bicarbonate, chloride, BUN, and creatinine.

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

(f) To be collected prior to tagraxofusp infusion if during dosing period. Serum electrolytes and chemistry: Sodium, potassium, bicarbonate, chloride, BUN, creatinine, glucose, ALT, albumin, ALP, AST, bilirubin (total, direct, and indirect), calcium, CPK, magnesium, LDH, phosphate, total protein, uric acid

(g) To be collected prior to initial tagraxofusp infusion of the cycle. Urinalysis includes appearance, color, pH, specific gravity, ketones, leukocytes, protein, glucose, bilirubin, urobilinogen, and occult blood.

(h) Morphology and differential WBC/blast count on aspirate. In Stage 2, patients with evidence of bone marrow involvement prior to study will have bone marrow evaluations following Cycles 4 and 6 and then every 3 months from Months 6-12; every 6 months from 12 to 24 months; and every 12 months thereafter until there is evidence of relapsed or progressive disease. If CD123 flow cytometry stain was performed on the bone marrow aspirations (± biopsy), the results should be recorded in the eCRF. For patients with evidence bone marrow involvement at Baseline, any additional bone marrow (± biopsy) remaining after the sample is aliquoted for tumor assessment will be collected for exploratory translational/correlative assessments. The bone marrow aspirate will be prioritized for tumor assessment (by the Investigator) over translational/correlative assessments.

(i) Baseline CT must be performed within 14 days prior to the first administration of tagraxofusp. For patients with baseline evidence of lymph node or visceral disease involvement, subsequent Stage 2 CT scans will be performed 21 (±7) days after the start of Cycles 1, 2, 4 and 6, and 21 (±7) days after the start of every 4th cycle thereafter (approximately every 12 weeks) until there is evidence of relapsed or progressive disease. For patients with no baseline evidence of lymph node or visceral BPDCN involvement, subsequent scans should be performed at the end of Cycle 2 (±7 days; or at time of disease progression if PD occurs prior to end of Cycle 2), 21 (±7) days after the start of Cycle 6, and at Investigator's discretion thereafter. Baseline CT scans should be full-body, whereas follow-up scans should document response of index lesions.

(j) For patients with skin involvement. Baseline must be performed within 14 days prior to the first administration of tagraxofusp. Subsequent skin assessments will be performed 21 (±7) days after the start of Cycles 1 and 2 and 21 (±7) days after the start of every 2nd cycle thereafter (Cycles 4, 6, etc.) until there is evidence of relapsed or progressive disease. Quantification of skin disease burden via the mSWAT instrument is required at the time of each skin assessment.

(k) Peripheral blood samples will be collected 21 (±7) days after the start of Cycles 1 and 2 and, for patients with evidence of peripheral blasts at baseline, 21 (±7) days after the start of every 2nd cycle thereafter (Cycles 4, 6, etc.) until there is evidence of relapsed or progressive disease.

(l) In Cycle 1 and 3, and at end of treatment, Translational Assessment should be collected at Infusion 1 (Day 1) and Infusion 5 (Day 5, 6, 7; if Infusion 5 does not occur by Day 8, please obtain this specimen on Day 8) and Days 15, 21 (±3 days, but must occur prior to start of subsequent treatment cycle). (m) Refer to protocol section 7.5.3 – Premedication and Administration

(n) Following treatment with premedication, tagraxofusp will be administered as a 15-minute infusion for the first 5 consecutive days of a 21-day cycle. Individual tagraxofusp infusions may be delayed to allow for toxicity resolution; all infusions should be completed within 10 days and fewer than 5 infusions are permitted in settings of incipient CLS or hepatotoxicity. Patient must be monitored for 4 hours post infusion.

(o) Plasma samples (6 mL each) will be collected immediately prior to the start of the infusion of tagraxofusp, immediately after end of infusion (time recorded), then 15, 30, 45, 60, 90, 120, 180, and 240 minutes after the start of the infusion during infusions 1 (ie, Day 1) and 5 (ie, Day 5) of Cycles 1 & 3 (see Table 10 in the protocol).

(p) Blood samples (10 mL) will be collected for the detection of tagraxofusp reactive antibodies on Day 1 (pre-infusion), Day 15, and Day 21 (if patient is to receive additional cycles, can be Day 1, pre-infusion for next cycle). If infusions are held, please collect a sample within 3±3 days after the last tagraxofusp infusion of the cycle.

(q) After the follow-up visit, patients will then be followed every 90 days for survival status. The survival follow-up may be by telephone contact. If the patient is in CR/PR at the time of discontinuation, disease assessments should continue to be performed as described in protocol Section 8.12 on an every 6-week basis (±1 week) through 6 months post-C1D1 and then on an every 90-day basis or until, in the judgment of the Investigator, there is evidence of relapsed or progressive disease.

(r) For Cycles 5 and beyond (Stage 2), the Day 8 and 15 concomitant medication and AE evaluations may be conducted via telephone and the laboratory evaluations may be performed at local laboratories for patients who live a considerable distance from the study center.

8. RESPONSE CRITERIA

8.1. Tumor Response Criteria for Patients with AML

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

8.2. Tumor Response Criteria for Patients with BPDCN

Response	Location	Criteria
Complete Response (CR)	Marrow[b]	• Normalization of blast percentage (≤5%) [b]
	Peripheral Blood	• Normalization of neutrophil count (≥1,000/μL) and platelet count (≥100,000/μL) • Absence of leukemic blasts

	Skin[a]	• 100% clearance of all skin lesions from baseline; no new lesions in patients without lesions at baseline[a]
	Nodal Masses	• Regression to normal size on CT
	Spleen, Liver	• Not palpable, nodules disappeared
CR with incomplete blood count recovery (CRi)	Marrow[b]	• Normalization of blast percentage (≤5%) [b]
	Peripheral Blood	• Incomplete recovery of neutrophil and/or platelet count • Absence of leukemic blasts
	Skin[a]	• 100% clearance of all skin lesions from baseline; no new lesions in patients without lesions at baseline[a]
	Nodal Masses	• Regression to normal size on CT
	Spleen, Liver	• Not palpable, nodules disappeared
CR [clinical] with minimal residual skin abnormality (CRc)	Marrow[b]	• Normalization of blast percentage (≤5%) [b]
	Peripheral Blood	• Normalization of neutrophil count (≥1,000/μL) and platelet count (≥100,000/μL) • Absence of leukemic blasts
	Skin[a]	• Marked clearance of all skin lesions from baseline; residual hyperpigmentation or abnormality with BPDCN identified on biopsy (or no biopsy performed) [a, c]
	Nodal Masses	• Regression to normal size on CT
	Spleen, Liver	• Not palpable, nodules disappeared
Partial Response (PR)	Marrow[b]	• Decrease by ≥50% in blast percentage to 5 − 25% [b]
	Peripheral Blood	• Normalization of neutrophil count (≥1,000/μL) and platelet count (≥100,000/μL)
	Skin[a]	• 50% − <100% clearance of all skin lesions from baseline; no new lesions in patients without lesions at baseline[a]
	Nodal Masses	• ≥50% decrease in SPD of up to 6 largest dominant masses; no increase in size of other nodes
	Spleen, Liver	• ≥50% decrease in SPD of nodules (for single nodule in greatest transverse diameter); no increase in size of liver or spleen
Stable Disease (SD)		• Failure to achieve at least a PR, but no evidence of progression for at least 8 weeks
Relapse after CR/CRi/CRc	Marrow[b]	• Blast percentage >5% (if no peripheral blasts, then confirmation aspirate required ≥1 week later) [b]
	Peripheral Blood	• Presence of leukemic blasts
	Skin[a]	• Increase in skin score greater than the sum of nadir plus 50% baseline score[a]
	Nodal Masses	• Appearance of a new lesion(s) >1.5 cm in any axis, ≥50% increase from nadir in SPD of more than one node, or ≥50% increase from nadir in longest diameter of a previously identified node >1 cm in short axis
	Spleen, Liver	• >50% increase from nadir in the SPD of any previous lesions
	Marrow[b]	• Blast percentage ≥25% (if no peripheral blasts, then confirmation aspirate required ≥1 week later) [b]

	Skin[a]	• Increase in skin score greater than the sum of nadir plus 50% baseline score[a]
Relapse after PR[d]		
	Nodal Masses	• Appearance of a new lesion(s) >1.5 cm in any axis, ≥50% increase from nadir in SPD of more than one node, or ≥50% increase from nadir in longest diameter of a previously identified node >1 cm in short axis
	Spleen, Liver	• >50% increase from nadir in the SPD of any previous lesions
Progressive Disease (PD)[d]	Marrow[b]	• >50% increase in blasts from baseline (and blast percentage >5%)[b]
	Peripheral Blood	One or more of the following: • ≥50% decrease from peak remission levels in platelets or granulocytes; • Reduction in hemoglobin concentration by at least 2 g/dL; • Transfusion dependence
	Skin[a]	One or more of the following: • ≥25% increase in skin disease from baseline[a]
	Nodal Masses	• Any new tumors in patients without tumors at baseline • Appearance of a new lesion(s) >1.5 cm in any axis, ≥50% increase from nadir in SPD of more than one node, or ≥50% nadir in longest diameter of a previously identified node >1 cm in increase from short axis
	Spleen, Liver	• >50% increase from nadir in the SPD of any previous lesions

Abbreviations: CT, computed tomography; SPD, sum of the product of the diameters.
All parameters detailed above (including bone marrow, blood, skin [including mSWAT quantification], lymph nodes and viscera) are to be assessed both at baseline and stipulated subsequent timepoints. Responses are determined via comparison to baseline values (or post-treatment nadir values as stipulated in the above table).

[a] The percentage of clearance or increase in skin disease is calculated using the Modified Severity Weighted Assessment Tool (mSWAT), which is provided in protocol Section 16.3/Appendix C. Detailed guidance and examples regarding the calculation of an mSWAT assessment based on the size and nature of a patient's skin lesions are provided in a separate document.

[b] In settings in which there is a change in the blast population on bone marrow evaluation by means of flow cytometry or other molecular methodology without a similar degree of change in the morphologic blast percentage, the morphologic percentage should be utilized to determine response/progression; however the findings from flow cytometry (or other molecular methodology) should be recorded as part of the study record. [c] Marked clearance is defined as the resolution of ≥75% of a patient's skin lesions from baseline. Residual hyperpigmentation or abnormality identified on a biopsy (or no biopsy performed) is defined as residual skin abnormality, not indicative of active disease, that covers ≤10% of a patient's skin overall. [d] In settings in which there is preliminary but not conclusive evidence of disease progression (i.e., new skin lesions of indeterminate etiology, new lymph nodes ≤1.5cm, modest increase in bone marrow blast percentage subsequent to an initial marked reduction, appearance of new blast population on bone marrow evaluation by means of flow cytometry or other molecular methodology without increase in blast percentage consistent with PD), additional tagraxofusp cycles may be administered, provided that the Investigator documents that the evidence of potential disease progression is inconclusive and that the overall risk/benefit assessment favors additional investigational therapy. Additional tagraxofusp may also be administered in situations of mixed response BPDCN in which response/progression is not consistent between a given patient's disease sites, provided that the Investigator

documents that the overall risk/benefit assessment favors additional investigational therapy. In such situations, it is essential that relevant findings and assessments are documented, and that areas of potential disease progression are followed closely on subsequent assessments.

8.3. Capillary Leak Syndrome FDA Algorithm

Screen-positive subjects include those

a. With Preferred Term "Capillary leak syndrome" as an adverse event, OR

b. Those meeting 2 or more of the criteria in the table below with at least 2 criteria having a start date within 7 days of each other.

Criteria	**Variable**	**Value**
		Hypoalbuminaemia
≥ 1 of these (Hypoalbuminemia)	PT	
	Lab result	Albumin *(any value ≤ 3 g/dL)*
≥ 1 of these (Edema)	HLT	Oedema
	PT	Ascites
	PT	Eyelid oedema
	PT	Face oedema
	PT	Fluid overload
	PT	Fluid retention
	PT	Generalized oedema
	PT	Hypervolaemia
	PT	Laryngeal oedema
	PT	Oedema
	PT	Oedema peripheral
	PT	Periorbital oedema
	PT	Peripheral swelling
	PT	Swelling
	PT	Swelling face
	PT	Weight increased
	VS result	Weight (*any value > 5 kg from baseline)*
≥ 1 of these (Hypotension)	PT	Hypotension
	PT	Hypovolemia
	VS result	Systolic Blood Pressure *(any value < 90 mmHg)*
	Concomitant medication	Dopamine
	Concomitant medication	Ephedrine
	Concomitant medication	Epinephrine
	Concomitant medication	Norepinephrine
	Concomitant medication	Norepinephrine bitartrate
	Concomitant medication	Phenylephrine

	Concomitant medication	Vasopressin
≥ 1 of these (CRS/IRR)	PT	Cytokine release syndrome
	PT	Infusion related reaction
("Catch-all")	PT	Cardiac arrest
	PT	Cardiopulmonary failure
	PT	Multi-organ failure
	PT	Multiple organ dysfunction syndrome