NCT: NCT05305989

STATISTICAL ANALYSIS PLAN

Extended Treatment and Follow-up of Subjects Treated with Belumosudil in Study KD025-208 or Study KD025-213

Protocol Number:	KD025-217
Study Drug:	Belumosudil (KD025)
IND Number:	IND 125890
Phase:	2
Sponsor:	Sanofi
	55 Corporate Drive
	Bridgewater, NJ 08807
Version 1:	03Apr2024

SIGNATURE PAGE

[redacted]

Prepared by: [redacted]
Senior [redacted]
Tigermed

Date of Signature
(DD MMM YYYY)

[redacted]

Reviewed by: [redacted]
[redacted]

Sanofi

Date of Signature
(DD MMM YYYY)

[redacted]

~~Clinical Leader~~ [redacted]

Sanofi

Date of Signature
(DD MMM YYYY)

DOCUMENT HISTORY

Version	Author	Description
V1	[redacted]	Original Document

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Full Term
AE	Adverse Event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic Class
BID	Two times a day
BMI	Body Mass Index
cGVHD	Chronic Graft Versus Host Disease
CI	Confidence interval
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CM	Concomitant Medication
CR	Complete Response
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DOR	Duration of Responses
ECG	Electrocardiogram
FEV1	Forced Expiratory Volume in One Second
FFS	Failure Free Survival
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
KPS	Karnofsky Performance Scale
LR	Lack of Response
LSS	Lee Symptom Scale Score
LTFU	Long-term Follow-up
MDRD-4	Modification of Diet in Renal Disease
MedDRA	Medical Dictionary for Regulatory Activities
NE	Not Evaluable
NIH	National Institutes of Health
OS	Overall Survival
PFT	Pulmonary Function Test
PR	Partial Response
PT	Preferred Term
QD	Every day
QTcF	Corrected QT interval using Fridericia's formula
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment-emergent Adverse Event
TTNT	Time to Next Treatment

1 INTRODUCTION

This Statistical Analysis Plan (SAP) describes data-handling and statistical procedures to be used for Study KD025-217 as specified in protocol (Amendment No. 1, 08 September 2021): Extended Treatment and Follow-up of Subjects Treated with Belumosudil in Study KD025-208 or Study KD025-213.

The SAP was written in accordance with the recommendations outlined in the International Conference on Harmonisation (ICH) E9 Guideline entitled “Guidance for Industry: Statistical Principles for Clinical Trials” and the most recent ICH-E3 Guideline, entitled “Guidance for Industry: Structure and Content of Clinical Study Reports.”

2 STUDY SUMMARY

2.1 Study Objectives

2.1.1 Primary objective

- The objective of this study is to evaluate long-term safety and efficacy of belumosudil in subjects with Chronic Graft Versus Host Disease (cGVHD) who have previously been treated with belumosudil.

2.1.2 Secondary objectives

- Duration of Responses (DOR): Response is defined by the 2014 National Institutes of Health (NIH) Consensus Development Project on clinical trials in cGVHD and are assessed by investigators
- Change in Lee Symptom Scale Score (LSS)
- Time to Next Treatment (TTNT)
- Failure Free Survival (FFS)
- Overall Survival (OS)
- Percentage of subjects who have a best response of Partial Response (PR) and percentage of subjects who have a best response of Complete Response (CR)
- Response by individual organ
- Change in corticosteroid dose
- Change in cGVHD Global Severity Rating based on the Clinician-reported Global cGVHD Activity Assessment

2.2 Study Design

KD025-217 is a Phase 2, open-label, extended treatment and follow-up study in subjects with cGVHD who have been previously treated with belumosudil in Study KD025-208 or in Study KD025-213.

Subjects who have signed the Informed Consent Form (ICF) will be enrolled in Study KD025-217 if they have met any of the following conditions:

- Actively receiving belumosudil in Study KD025-208 or in Study KD025-213 (excluding adults in the Companion Study and excluding adolescents enrolled under KD025-213 Amendment 2 [01June 2020])
- Is in long-term follow-up (LFTU) on Study KD025-208 or Study KD025-213. Long-term follow-up will be defined as the period after ending treatment with belumosudil and until an FFS event occurs.
- Adult enrolled in the Companion Study as specified in Study KD025-213 Amendment 2 (01 June 2020) and received at least 6 months of treatment or is in LTFU.

After signing the ICF, at Baseline, subjects will undergo efficacy assessments by Clinician-reported Global cGVHD Activity Assessment, response assessment, pulmonary function tests (PFTs), LSS score, and documentation of corticosteroid dosage and/or other therapies for cGVHD.

Subjects will undergo safety assessments including symptom-directed Physical Exams (PE), vital signs, weight, height, Karnofsky Performance Scale (KPS), hematology and clinical chemistry laboratory testing, pregnancy testing, 12-lead Electrocardiogram (ECG), documenting concomitant medications, and monitoring for Adverse Events (AEs). Subjects will continue on their same dosing regimen they received in Study KD025-208 or Study KD025-213: belumosudil 200 mg QD, 200 mg BID, or 400 mg QD. If the Investigator reduced the subject's dose of belumosudil in either Study KD025-208 or Study KD025-213, the subject will continue to receive that same reduced dose treatment with belumosudil throughout Study KD025-217 or until the AE that led to the temporary dose reduction has resolved, in which case the dose treatment can resume at the original dose at the discretion of the Investigator. Subjects will begiven a study drug diary to record daily dosing. Subjects may be administered the first dose either in the clinic or self-administered at home.

At 3 months after Baseline and every 3 months thereafter (± 14 days), subjects will report to the clinic for the same efficacy and safety assessments and dispensing of belumosudil at Baseline.

(Height is only required at Baseline.) The study drug diaries for the past 3 months will be collected, and new study drug diaries dispensed. Any remaining belumosudil will be collected. Subjects may receive belumosudil on Study KD025-217 for up to an additional 2 years, at the discretion of the Investigator (and with guidance from the medical monitor), beyond what they had received in Study KD025-208 or Study KD025-213, cGVHD progression, or unacceptable toxicity.

Subjects will return to the clinic for a 28-day Follow-up Visit at 28 days (± 7 days) after the last dose of study drug. Subjects will have LFTU approximately every 3 months until an FFS event occurs. After the 28-day Follow-up Visit has been completed, subjects will be contacted approximately every 3 months by telephone, e-mail, or postal mail to confirm survival status, any initiation of new systemic cGVHD therapies, and relapse of underlying disease. Alternatively, this information may be obtained from the subject's medical records.

2.3 Visit Schedule and Study Assessment

The flow chart of visit schedule and study assessments is given in Appendix A of the KD025-217 Protocol.

3 STATISTICAL METHODS

3.1 General Methods

3.1.1 Computing Environment

All statistical analyses will be performed using SAS® Version 9.4 or higher for Windows.

3.1.2 Sample Size Justification

No formal power calculation will be performed to determine the study sample size.

3.1.3 General Considerations

General considerations for descriptive statistics, presentation, and analysis model (for efficacy analysis) used for continuous and categorical data are given below.

3.1.3.1 Reporting of Numerical Values

Continuous data will be described using descriptive statistics: number of observations (n), mean, standard deviation, median, minimum and maximum. Frequencies and percentages will be used for summarizing categorical data. When categorical data are presented, the percent will

be suppressed when the count is zero in order to draw attention to the non-zero counts. The denominator for all percentages, unless otherwise specified, will be the number of subjects in the specified analysis population or group.

Means, medians, standard deviations, and confidence intervals will be reported to one decimal place more than the data reported on the case report form (CRF) or by the laboratory/vendor. Minimum and maximum will be reported to the same number of decimal places displayed on the CRF or by the laboratory/vendor. P-Values will be reported to 4 decimal places.

3.1.4 Study Day

The study day for all assessments prior to the first study drug administration is calculated as the difference between the date of the measurement (e.g., assessment date, sample collection date, etc.) and the start date of study treatment. The day before the start of study treatment is Study Day -1.

The study day for all post assessments after the first study drug administration is calculated as the difference between the date of the measurement (e.g., assessment date, sample collection date, etc.) and the start date of study treatment in parent study KD025-208/KD025-213, plus one day. The first day of study treatment is Study Day 1.

3.1.5 Baseline

Baseline value for efficacy analyses is defined as the valid and last non-missing value obtained within 28 days prior to subject receiving the first study medication in parent study KD025-208/KD025-213.

Baseline value for safety analyses is defined as the “Baseline” visit from KD025-217.

Baseline can be the day before the first study medication or on the same day as the first study medication if a pre-dose assessment is available. Subjects without data on a parameter before the first study medication will have a missing baseline for this parameter.

3.1.6 Handling of Incomplete or Missing Data

Missing data will not be imputed in general and it will be reported as missing in all listings. For categorical variables, patients with missing data are not included in calculations of percentages unless otherwise specified.

3.1.6.1 Missing start and end dates for AE and concomitant medication (CM)

The assumption will be the worst or most conservative judgment when imputing AE and CM start and end dates. The purpose of imputing a start date is to help define whether the AE/CM started while taking study drug.

For a partial or missing start date:

- If the day is missing, the first day of the month will be imputed. If the missing day is the same as the month of first dose of study drug, then the first dose date will be imputed.
- If the day and month are missing, the first day of January will be imputed. If the year is the same as the first dose date, then the first dose date will be imputed.
- If the day is completely missing, the first dose date will be imputed. If the end date suggests it could have started prior to this, the first day of January of the same year as the end date will be imputed.
- When imputing a start date, the start date will ensure that the new imputed date is sensible, i.e., is prior to the end date of the AE or CM.

For a partial or missing end date:

- If the day is missing, the last day of the month or the last assessment date, whichever is earlier, will be imputed.
- If the day and month are missing, the 31st of December or the last assessment date, whichever is earlier, will be imputed.
- If the date is completely missing, there will be a need to look at whether the AE/CM is still ongoing before imputing a date. If the ongoing flag is missing, then it will be assumed that AE is still present, or CM is still being taken (i.e., do not impute a date). If the AE/CM has stopped, then the last assessment date will be imputed.

These data imputations are for categorization purpose only and will not be used in listings.

If the assessment of the relationship of the AE to belumosudil is missing, then it will be assumed that the AE is related to belumosudil and the AE considered as such in the frequency tables of possibly related AEs. No imputation should be done at the data level.

3.1.6.2 Missing event dates

Event date will be imputed only when day is missing, and the purpose of imputing an event date is to most conservatively calculate time to event. If the day is missing, the first (mid, last) day of the month will be imputed for undesired (neutral, desired) event. If the missing day is the same as the month of first dose of study drug, then the first dose date will be imputed for undesired event. These data imputations are for time to event calculation only and will not be used in the listings.

3.2 Analysis Populations

Full Analysis Set: The Full Analysis Set is defined as all subjects enrolled in the study. All safety and efficacy analyses will be performed on the Full Analysis Set.

There are 3 treatment groups for this study: belumosudil 200 mg QD, 200 mg BID, and 400 mg QD. Each subject's treatment group is determined by their originally assigned dosing cohort (Study KD025-208) or randomized arm (Study KD025-213).

3.3 Subject Disposition and Evaluability

A disposition of safety population, the number of subjects discontinuing from treatment, and the primary reason for discontinuation will be summarized. This summary will also be presented by exposure range.

Subject listing will be provided for the disposition of subjects who died during study.

3.4 Demographics and Baseline Characteristics

3.4.1 Demographics and Other Characteristics

Subject demographics information will be summarized. Descriptive statistics will be provided for age, height, weight, body mass index (BMI) and KPS. Frequencies and percentages will be tabulated for gender, child bearing potential, race, and ethnicity.

Demographics and other baseline characteristics will be summarized by treatment group in each population.

3.4.2 Medical History

Medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA) Version 24.1 terminology, and summarized by Preferred Term (PT). Subject listing will be provided.

3.5 Concomitant Medications

Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHODrug-Global-B3, version 202109). CM are defined as medications taken after the start of study treatment and during the study period, including those initiated before but ongoing at the start of study treatment. If a medication start date is partially or fully missing and it is unclear as to whether the medication is prior or concomitant, it will be assumed that the medication is concomitant.

Number and percentage of incidence of CM will be summarized according to Anatomical Therapeutic Class (ATC) and preferred drug name.

3.6 Treatment Exposure

The number of subjects will be summarized, along with descriptive summary of duration of treatment, exposure ranges and overall exposure.

The duration of treatment (days) is defined as: (Date of last dose – Date of first dose) + 1. It will also be classified into the following exposure ranges: 0 to 6 months; > 6 to 12 months, > 12 to 18 months, > 18 to 24 months, > 24 to 30 months, > 30 to 36 months and > 36 months.

Overall exposure will be summarized in total patient-years, derived as Exposure in patient (years) = Sum of duration of treatment in days (for all patients in each dose group) / 365.25.

3.7 Definition of Efficacy Endpoints

3.7.1 Primary Efficacy Endpoint: Overall Response Rate (ORR)

Kadmon Algorithmic Response Assessments is used in all of the overall response analyses. Individual organ response is assessed for Skin, Eyes, Mouth, Esophagus, Upper GI, Lower GI, Liver, Lungs, Joints and Fascia and Global Severity Rating by using individual scores or assessment results for the ten systems based on the APPENDIX A in the SAP. The overall response is calculated from the 9 organ (Skin, Eyes, Mouth, Esophagus, Upper GI, Lower GI, Liver, Lungs, and Joints and Fascia) response using the definition in the Table 2 below. Response is assessed with respect to the parent study baseline (Cycle 1, Day 1 (C1D1)) cGVHD assessment. The overall response at each assessment time point is categorized as Complete Response (CR), Partial Response (PR), or Lack of Response (LR), where LR includes the response status of unchanged, mixed, or progression as defined in Table 1.

The ORR is defined as the proportion of subjects with a best response meeting the overall response criteria assessment of CR or PR at any post-baseline response assessment.

If a treated subject is lost to follow up without any response assessment, this subject will be counted as a non-responder.

Point estimates, two-sided 95% confidence intervals (Clopper-Pearson (exact) method), by treatment groups will be reported. The number and percentage of subjects who have a best response of PR and number and percentage of subjects who have a best response of CR will be reported.

In addition to ORR assessed by investigators, analyses of ORR and DOR will be repeated with this investigator's assessment of overall response.

The following will also be presented:

Table 1 cGVHD Response Definitions

Response	**Definition**
Complete Response (CR)	Resolution of all manifestations of cGVHD in each organ or site
Partial Response (PR)	Improvement in at least one organ or site without progression in any other organ or site
Lack of Response (LR)	
Mixed (LR-M) *	Complete or partial response in at least one organ accompanied by progression in another organ
Unchanged (LR-U)	Outcomes that do not meet the criteria for complete response, partial response, progression or mixed response
Progression (LR-P)	Progression in at least one organ or site without a response in any other organ or site

*Considered progression for purposes of analysis

3.7.2 Secondary Efficacy Endpoints

3.7.2.1 Duration of Response (DOR)

DOR will be presented according to 5 methodologies:

- The primary definition of DOR is the time from first documentation of response to the time of first documentation of deterioration from best response (e.g., CR to PR, or PR to LR).
- The secondary definition of DOR is the time from first documentation of response to the time of first documentation of lack of response.
- The tertiary definition of DOR is the time from first documentation of response to the time of initiation of new systemic cGVHD therapy.
- The quaternary definition of DOR is the time from first documentation of response to the time of first documentation of lack of response (as the secondary definition) but with durations summed for multiple response/lack of response episodes.
- The FDA definition of DOR is the time from first documentation of response to the time of first documentation of organ-level progression, initiation of a new cGVHD therapy, or death from any cause, whichever occurs first

The DOR will be reported for early responders, and statistics will include:

- Kaplan-Meier plots and descriptive statistics of DOR. The censoring rules in Table 2 will be applied
- Landmark analyses: Number and percentage of subjects with response sustained for ≥ 3, ≥ 6, ≥ 9, ≥ 12, ≥ 15, ≥ 18, ≥ 24, ≥ 30, ≥ 36 months

Table 2 Censoring Rules for Duration of Response

DOR	Events	Censoring
Primary	• Deterioration from best response • Initiation of new systemic therapy for cGVHD • Death	• Last documented response assessment • If LR or initiation of new systemic therapy happens

Secondary	• Documented LR • Initiation of new systemic therapy for cGVHD • Death	immediately after two or more missed response assessments, the event date is set as 24 weeks after last documented response assessment prior this event
Tertiary	• Initiation of new systemic therapy for cGVHD • Death	Last response assessment or long term follow up assessment, whichever is the latest and available
Quaternary	• Documented LR • Initiation of new systemic therapy for cGVHD • Death • With summation of DOR from multiple response episodes	Same as censoring rules for primary and secondary
FDA defined	• Documentation of organ-level progression • Initiation of a new cGVHD therapy • Death	• Last documented response assessment • If organ-level progression or initiation of new systemic therapy happens immediately after two or more missed response assessments, the event date is set as 24 weeks after last documented response assessment prior this event

3.7.2.2 Change in Lee Symptom Scale Score (LSS)

The Lee cGVHD Symptom Scale consists 30 items of 7 domains: skin, eyes and mouth, breathing, eating and digestion, muscles and joints, energy, and mental and emotional. Each question is scored 0, 1, 2, 3 or 4.

A domain score will be calculated for each domain by taking the mean of all items completed if more than 50% were answered and normalizing to a 0 to 100 scale. A summary score will be calculated as average of all non-missing domain scores if more than 50% of them are non-missing. A higher score indicated more bothersome symptoms. 7 points difference on the summary score of cGVHD symptom scale was found to be clinically meaningful. In the above formulas, m is the number of items have non-missing value in a domain, and n is the number of domains have non-missing domain scores.

$$Raw\ \ Score = \frac{I_1 + \cdots + I_m}{m}, \qquad Domain\ \ Score = (Raw\ \ Score / 4) * 100$$

$$Summary\ \ Score = \frac{Domain\ \ Score_1 + \cdots + Domain\ \ Score_n}{n}$$

In the above formulas, m is the number of items have non-missing value in a domain, and n is the number of domains have non-missing domain scores.

Analyses will include:

- Both score and change-from-baseline values (summary score and domain scores) will be summarized as continuous variables by visit
- Number and percent of subjects with a ≥7-point reduction from baseline (C1D1)
- Number and percent of subjects with a ≥7- point reduction from baseline on at least 2 consecutive post-baseline assessments.
- Duration of a 7-PtR (DO7-PtR) (defined as time from documentation of the first ≥7-point reduction to the first documentation of less than 7-point reduction). If there are multiple episodes, then DO7-PtR will be measured as the sum of DO7-PtR from all episodes.

These analyses will be performed on Full Analysis Set.

3.7.2.3 Time to Next Treatment (TTNT)

The TTNT will be measured as the time from first treatment to the time of new systemic cGVHD treatment (which will be defined by clinical team review), censored by last response assessment or long term follow up assessment, whichever is the latest and available. TTNT will be analyzed

by the Kaplan-Meier survival method as well as the landmark analysis at 6, 12, 18, and 24 months.

3.7.2.4 Failure-free-survival (FFS)

FFS is defined as the absence of new cGVHD systemic therapy, non-relapse mortality and recurrent malignancy (i.e., underlying disease) and is censored by last response assessment or long term follow up assessment, whichever is the latest and available. Kaplan-Meier plots, descriptive statistics of FFS and landmark FFS at 6, 12, 18, 24, 30 and 36 months will be provided. In addition, analyses for the three components of FFS will also be provided.

3.7.2.5 Overall Survival (OS)

OS is defined as time from first dose of belumosudil to the date of death due to any cause and is censored by last response assessment or long term follow up assessment, whichever is the latest and available. Kaplan-Meier plots, descriptive statistics of OS and landmark OS at 6, 12, 18, 24, 30, 36 months will be provided.

3.7.2.6 Percentage of subjects who have a best response of PR and percentage of subjects who have a best response of CR

The percentage of subjects who have a best response of PR and percentage of subjects who have a best response of CR will be summarized separately by treatment groups.

3.7.2.7 Response by individual organ

For KD025-213 and KD025-217 organ involvement is captured in the CRF. For KD025-208, the database did not specifically capture baseline organ involvement. Organ involvement was determined as follows:

- All organs, except lungs:
  - If the baseline score is > 0, the organ will be defined as Involved, unless all response assessments are Not Evaluable (NE)
  - If the baseline score is 0, the organ will be defined as Not Involved, unless any response assessment of CR or PR which triggers clinical review to confirm baseline involvement.
- For the lungs:
  - If Forced Expiratory Volume in One Second (FEV1) ≥ 75%, the organ will be defined as Not Involved
  - If FEV1 < 75%, the organ will be defined as Involved, unless all responses are NE

The response assessment for the 9 individual organs (skin, eyes, mouth, esophagus, upper GI, lower GI, liver, lungs, and joints and fascia) and global severity rating is categorized into 5 response status: CR, PR, Progression, NE, Stable.

Number and percentage of subjects with best response of CR or PR and total (CR+PR) by individual organ system and global severity rating will be summarized and presented. With the total number of subjects in the Full Analysis Set with involvement of the given organ at baseline as denominator.

3.7.2.8 Change in corticosteroid dose

Corticosteroid doses will be presented as mg/kg/day prednisone equivalent dose. Any corticosteroid dose increases to ≥ 0.9 mg/kg PE during the study is counted as new systemic cGVHD therapy.

Descriptive statistics for the Full Analysis Set will be provided for:

- Both prednisone equivalent dose and change-from-baseline values (summary score and domain scores) will be summarized as continuous variables by visit
- Change and % change from baseline to the greatest corticosteroid dose reduction at any post-baseline visit
- Number and % of subjects who reduced systemic corticosteroid dose at any post-baseline visit
- Number and % of subjects who ever discontinued systemic corticosteroid usage at any post-baseline visit

If subjects are not using prednisone as the systemic corticosteroid, then the prednisone dose equivalent will be determined according to following conversion ratios2:

1 mg prednisone is equivalent to:

- 4mg Hydrocortisone
- 0.8mg Methylprednisolone
- 0.15mg Dexamethasone
- 1mg Prednisolone

- 0.8mg Triamcinolone

3.7.2.9 Change in cGVHD Global Severity Rating based on the Clinician-reported Global cGVHD Activity Assessment

The Clinician-reported global cGVHD Activity Assessment is a 0-10 point numeric rating scale with a score of 0 indicating “cGVHD symptoms not at all severe” and a score of 10 being “Most severe cGVHD symptoms possible”. The activities are assessed on at baseline and then every 3 months after.

Change from baseline in cGVHD severity based on the Clinician-reported global cGVHD Activity Assessment will be summarized as a categorical endpoint at all scheduled assessment visits.

3.8 Safety Analysis

Safety assessments include AEs, clinical laboratory evaluations (hematology and chemistry), vital sign measurements, and electrocardiograms. Unscheduled visits for safety assessments will not be presented in by visit summary tables but will be in listings and shift tables. All safety analyses will be performed using the safety population.

3.8.1 Adverse Events

AEs will be coded using the MedDRA dictionary (Version 24.1 or higher). Treatment-emergent AEs (TEAEs) are any AEs occurring or worsening in severity after the first administration of study medication until 28 days after EOT. All AEs will undergo 4 sets of ratings by Investigator as follows:

- Serious vs. non-serious [based on seriousness criteria of death, life threatening, hospitalization – initial or prolonged, IME, disability or permanent damage, congenital anomaly or birth defects, and required intervention]
- Grades using the 5-point Common Terminology Criteria for Adverse Events (CTCAE 4.03) scale (mild, moderate, severe, life-threatening, or death)
- Causality assessment with respect to study drug classified as either: related (definitely related, related, probably related, and possibly related) or not related (unlikely related, not related).
- The relationship of AEs to cGVHD or the underlying disease

The number (N) and percentage (%) of patients who experienced at least one TEAE will be summarized by dose group, including tabulation by:

- System Organ Class (SOC) and PT within each SOC in decreasing total frequency
- Decreasing frequency of PT
- Maximum severity
- Relationship to study drug

These analyses will be repeated for serious TEAEs, events leading to dose modification, events leading to study drug discontinuation, events related to study drug and events related to cGVHD or the underlying disease.

Subject listings will be provided for AEs, SAEs, AEs resulting in study drug discontinuation and deaths. Time to onset and duration of AEs will be included in listings, along with action taken and outcome.

3.8.2 Clinical Laboratory Evaluation

The summary statistics (including number, mean, SD, median, minimum and maximum) of all laboratory variables and changes from baseline will be calculated for each visit or study assessment by treatment group.

For parameters of white blood cell counts, neutrophils (absolute count), lymphocytes (absolute count), monocytes (absolute count), hemoglobin, platelets, ALP, ALT, aspartate aminotransaminase, gamma glutamyl transferase, total bilirubin, glomerular filtration rate, plots of mean/mean changes from baseline with the corresponding standard error will be displayed.

For shift tables, laboratory results will be classified using the CTCAE 4.03. All graded laboratory parameters will be summarized separately for hematology and biochemistry. Corresponding shift tables to compare baseline to the worst post-baseline grade within the treatment period will be provided.

A by patient by lab tests by visit listing will be generated for lab results with CTC grade >=3.

The estimated glomerular filtration rate will be calculated and summarized using the Modification of Diet in Renal Disease (MDRD-4) and Chronic Kidney Disease Epidemiology Collaboration 2021 (CKD-EPI$_{2021}$) methods. Below is the formula used for each respective method.

Method	Formula
MDRD-4	eGFR = 175 x $(S_{Cr})^{-1.154}$ x $(age)^{-0.203}$ x 0.742 [if female] x 1.212 [if Black]
CKD-EPI_{2021}	eGFR = 142 x $\min(S_{cr}/\kappa, 1)^{\alpha}$ x $\max(S_{cr}/\kappa, 1)^{-1.200}$ x 0.9938^{Age} x 1.012 [if female]

Where:

Scr = standardized serum creatinine in mg/dL

κ = 0.7 (females) or 0.9 (males)

α = -0.241 (female) or -0.302 (male)

min(Scr/κ, 1) is the minimum of Scr/κ or 1.0

max(Scr/κ, 1) is the maximum of Scr/κ or 1.0

3.8.3 Vital Signs

Descriptive statistics for vital signs (weight, temperature, systolic and diastolic blood pressure, pulse rate (beats/min), and respiratory rate (breath/min)) values and the change from baseline will be presented by treatment group for each scheduled assessment time point.

A by patient by visit listing will be generated for vital signs.

3.8.4 ECG

Descriptive statistics for ECG parameters (i.e., heart rate, PR interval, RR interval, QRS interval, QT interval, and QTcF interval) at each time point will be presented for the absolute values and change from baseline scores (QTcF is the QT interval using Fridericia's correction which is calculated by $QTcF = QT/RR^{1/3}$).

The number and percentage of subjects with observed absolute QTcF values at maximum value during treatment period that satisfy the following conditions will be presented by dose group: ≤ 450 ms; > 450 to 480 ms; > 480 to 500 ms; and > 500 ms.

The number and percentage of subjects having change from baseline QTcF at maximum value during treatment period that satisfy the following conditions will be presented by dose group: ≤ 0 ms; > 0 to ≤ 30 ms; > 30 to ≤ 60 ms; and > 60 ms.

A by patient by visit listing will be generated for ECG results.

3.9 Pharmacokinetic Analysis

No PK assessments for analysis will be performed in this study.

3.10 Pharmacodynamics Analysis

No pharmacodynamic assessments for analysis will be performed in this study.

3.11 Exploratory Objectives Analysis

No exploratory analyses will be conducted.

4 APPENDIX A: Kadmon Algorithmic Response Assessments

Baseline Organ Involvement

Organ	Organ Involvement
Esophagus	NIH Esophagus Score ≥ 1, not entirely explained by non-cGVHD cause
Upper GI	NIH Upper GI Score ≥ 1, not entirely explained by non-cGVHD cause
Lower GI	NIH Lower GI Score ≥ 1, not entirely explained by non-cGVHD cause
Lungs	FEV1 <75% predicted, not entirely explained by non-cGVHD cause If no PFTs available, use NIH lung score ≥ 1, not entirely explained by non-cGVHD cause
Eyes	NIH Eye Score ≥ 1, not entirely explained by non-cGVHD cause
Joints and Fascia	P-ROM Score <25 (< maximal score) or NIH Joints/Fascia Score ≥ 1, not entirely explained by non-cGVHD cause
Skin	NIH Skin Score ≥ 1, not entirely explained by non-cGVHD cause
Mouth	NIH Modified Oral Mucosal Rating Score ≥ 1, not entirely explained by non-cGVHD cause
Liver	Elevation > 2x ULN of one or more (ALT, ALP, Total Bilirubin), not entirely explained by non-cGVHD cause

Response Assessments

Organ	Complete Response	Partial Response	Unchanged	Progression	Non Evaluable
Esophagus	NIH Esophagus Score 0 after previous involvement	Decrease in NIH Esophagus Score by 1 or more points	No change or increase from 0 to 1 in NIH Esophagus Score	Increase in NIH Esophagus Score by 1 or more points, except 0 to 1	NIH esophagus score is Non evaluable
Upper GI	NIH Upper GI Score 0 after previous involvement	Decrease in NIH Upper GI Score by 1 or more points	No change or increase from 0 to 1 in NIH Upper GI Score	Increase in Upper GI Score by 1 or more points, except 0 to 1	NIH upper GI score is Non evaluable
Lower GI	NIH Lower GI Score 0 after previous involvement	Decrease in NIH Lower GI Score by 1 or more points	No change or increase from 0 to 1 in NIH Lower GI Score	Increase in Lower GI Score by 1 or more points, except 0 to 1	If NIH Lower GI score is Non evaluable

Lungs	%FEV1 ≥ 80% after previous involvement; or NIH Lung Symptom Score 0 after previous involvement if PFTs not available	Increase by 10% absolute value of %FEV1; or decrease in NIH Lung Symptom Score by 1 or more points if PTSs not available	Not in the other categories	Decrease by 10% absolute value of %FEV1 and %FEV1 < 75%; or increase in NIH Lung Symptom Score by 1 or more points, except 0 to 1 if PFTs not available	NIH Lung score is Non evaluable and %FEV1 is missing
Eyes	NIH Eye Score 0 after previous involvement	Decrease in NIH Eye Score by 1 or more points	No change or increase from 0 to 1 in NIH Eye Score	Increase in NIH Eye Score by 1 or more points, except 0 to 1	NIH eye score is Non evaluable
Joints and Fascia	NIH Joint and Fascia Score 0 and P-ROM score 25 after previous involvement in at least one measure	Decrease in NIH Joint and Fascia Score by 1 or more points or increase in P-ROM score by 1 or more points for any site	Not in the other categories	Increase in NIH Joint and Fascia Score by 1 or more points or decrease in P-ROM score by 1 or more points for any site	Both in NIH Joint and Fascia and P-ROM scores are Non evaluable
Skin	NIH Skin Score 0 after previous involvement	Decrease in NIH Skin Score from baseline by 1 or more points	No change or increase from 0 to 1 in NIH Skin Score	Increase in NIH Skin Score from baseline by 1 or more points, except 0 to 1	NIH skin score is Non evaluable
Mouth	NIH Modified OMRS 0 after previous involvement	Decrease in NIH Modified Oral Mucosa Rating Score of 2 or more points	No change in NIH Modified Oral Mucosa Rating Score	Increase in NIH Modified Oral Mucosa Rating Score of 2 or more points	NIH mouth score is Non evaluable

Liver	Normal ALT, ALP and total bilirubin after elevation of one or more at baseline	50% improvement in ALT, ALP, or total bilirubin without increase by 2x ULN for any other parameters of ALT, ALP or total bilirubin	Not in the other categories	Increase by 2x ULN for any one of ALT, ALP or total bilirubin	If ALT, ALP and total bilirubin are missing
Global	GSR 0	GSR decreases by 2 or more points	Not in the other categories	GSR increases by 2 or more points	Not applicable

Organs not involved at baseline can progress according to the definitions in this Response Assessments table, but cannot achieve a response (PR or CR).