NCT04680975

STATISTICAL ANALYSIS PLAN

A Phase 2, Open-label Multicenter Study to Evaluate the Efficacy and Safety of Belumosudil in Subjects with Diffuse Cutaneous Systemic Sclerosis (dcSSc)

Protocol Number:	KD025-215/ACT17634
Study Drug:	Belumosudil (KD025)
IND Number:	IND 140383
Phase:	2
Sponsor:	Kadmon, a Sanofi Company 450 East 29th Street New York, NY 10016
Version 2:	20 January 2023

SIGNATURE PAGE

Prepared by: [redacted] Senior Statistician Tigermed	Date of Signature (DD MMM YYYY)
Reviewed by: [redacted] Study Statistician Sanofi	Date of Signature (DD MMM YYYY)
Reviewed by: [redacted] MD, PhD Medical Monitor Sanofi	Date of Signature (DD MMM YYYY)

DOCUMENT HISTORY

Version	Author	Description
V1	[redacted]	Original Document
V2	[redacted]	Final Document

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	Adverse event
ADI	Actual dose intensity
ANCOVA	Analysis of covariance
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Class
BID	Twice daily
CI	Confidence interval
CM	Concomitant medication
CRF	Case report form
CRISS	Combined Response Index in Diffuse Cutaneous Systemic Sclerosis
CTCAE	Common Terminology Criteria for Adverse Events
dcSSc	Diffuse cutaneous systemic sclerosis
DL_{CO}	Diffusing capacity of the lungs for carbon monoxide
ECG	Electrocardiogram
EOT	End of Treatment
FEV_1	Forced expiratory volume (in the first second)
FVC	Forced vital capacity
GFR	Glomerular filtration rate
GGT	Gamma-glutamyl transferase
Hb	Hemoglobin
HLGT	High level group term
HRCT	High-resolution computerized tomography
IEC	Independent Ethics Committee
ILD	Interstitial lung disease
IND	Investigational New Drug
IRB	Institutional review board
ICH	International Conference on Harmonisation
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified intent-to-treatment
MMRM	Mixed-effects Model of Repeated Measures
Ms	Millisecond
mRSS	Modified Rodnan Skin Thickness Score
PA	Primary analysis

PD	Pharmacodynamic
PDI	Planned dose intensity
PE	Physical exam
PFTs	Pulmonary function tests
PK	Pharmacokinetic
PT	Preferred term
QD	Once daily
QTcF	Corrected QT interval using Fridericia's formula
RDI	Relative dose intensity
RV	Residual volume
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SHAQ-DI	Scleroderma Health Assessment Questionnaire-Disability Index
SOC	System organ class
SSc	Systemic sclerosis
TEAE	Treatment emergent adverse event
TLC	Total lung capacity
VAS	Visual Analog Scale
WHO	World Health Organization

1 INTRODUCTION

This Statistical Analysis Plan (SAP) describes data-handling and statistical procedures to be used for Study KD025-215/ACT17634 as specified in protocol (Amendment No. 2, 01 July 2021): A Phase 2, Open-label Multicenter Study to Evaluate the Efficacy and Safety of Belumosudil in Subjects with Diffuse Cutaneous Systemic Sclerosis (dcSSc)

Exploratory objectives analyses (biomarkers analysis, histology and gene expression analyses) will be described in separate documents.

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

- To evaluate the efficacy of belumosudil using the Combined Response Index in diffuse cutaneous Systemic Sclerosis (CRISS) at Week 24.

2.1.2 Secondary objectives

- To assess the CRISS at Weeks 8, 16, 36 and 52
- To evaluate the efficacy of belumosudil at Week 24 for:
    - Modified Rodnan skin thickness score (mRSS)
    - Forced vital capacity (FVC)
    - Physician global assessment
    - Patient global assessment
    - Scleroderma Health Assessment Questionnaire-Disability Index (SHAQ-DI)
- To evaluate the efficacy of belumosudil at Weeks 8, 16, 36 and 52 compared to baseline for all subjects for the parameters in the bullet above
- To assess the safety of belumosudil in subjects with diffuse cutaneous systemic sclerosis (dcSSc) by examining the percentage of subjects with Treatment-emergent AEs (TEAEs) (Common Terminology Criteria for Adverse Events [CTCAE] v5.0)

2.2 Study Design

KD025-215/ACT17634 is a Phase 2, open-label, multicenter trial in subjects with dcSSc.

Subjects who have signed an Institutional Review Board/Independent Ethics Committee-approved informed consent form and met all of the inclusion/exclusion criteria will be enrolled. A total of 12-15 subjects will receive orally administered belumosudil 200 mg twice daily (BID) for 52 weeks. Study drug will be collected at the end of Week 52. The primary endpoint will be analyzed using Week 24 data.

Efficacy will be assessed at baseline and throughout the 52 weeks of the treatment period using CRISS, mRSS, pulmonary function tests (PFTs), Physician Global Assessments, and Patient Global Assessments. Safety will be assessed throughout the study.

Subjects will undergo evaluations as outlined in the Study Assessment table. A 4-Week Follow-Up visit will occur 28 days (± 3 days) after the last dose of study drug.

Efficacy will be assessed throughout the 52-week dosing period using CRISS, mRSS, PFTs, physician global assessment, SHAQ-DI.

2.3 Visit Schedule and Study Assessment

The flow chart of visit schedule and study assessments is given in Appendix A of the Study 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

This is an exploratory study. Sample size and power are not driven by hypothesis testing. Twelve to 15 subjects will be enrolled to receive PO administered belumosudil 200 mg BID.

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 Continuous variables

Continuous variables will be described by using these descriptive statistics: number of observations (n), mean, standard deviation (SD), median, minimum and maximum.

The means, medians, and standard deviation (SD) 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.

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 event or measurement (e.g., adverse event [AE] onset date, 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 event or measurement (e.g., AE onset date, assessment date, sample collection date, etc.) and the start date of study treatment plus one day. The first day of study treatment is Study Day 1.

3.1.5 Baseline

Baseline value is defined as the valid and last non-missing value obtained within 29 days prior to subject receiving the first study medication, unless otherwise stated under the related assessment section. 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

There will be no imputation for missing efficacy values. 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.

3.2 Analysis Populations

Three populations will be employed in the analysis of study data:

- The Modified Intent-to-treat (mITT) Population will consist of all subjects who receive at least 1 dose of belumosudil 200 mg.

- The Safety Population is defined as all subjects who receive at least 1 dose of belumosudil 200 mg. In this study, the Safety Population is equivalent to the mITT Population.

The primary endpoint will be analyzed on the modified Intent-to-Treat (mITT) Population.

Demographics, subject disposition, baseline characteristics, and efficacy analyses will be summarized in mITT Population and by group.

All safety analyses will be performed on the safety population.

3.3 Subject Disposition and Evaluability

A disposition of all enrolled subjects will be summarized. The number of subjects screening failed, discontinuing from study treatment, discontinuing from study and the primary reason for discontinuation will be summarized.

3.4 Protocol Deviations

All protocol deviations will be identified and classified as major or minor before the clinical database lock and will be presented in a listing.

Major Deviation: Protocol deviation that may impact the accuracy, and/or reliability of the study data or that may impact subject rights, safety or well-being.

Minor Deviation: Protocol deviation that does not impact the accuracy, and/or reliability of the study data or subject rights, safety or well-being.

3.5 Demographics and Baseline Characteristics

3.5.1 Demographics and Other Characteristics

A summary table and a by patient listing will be generated for patient demographics and other characteristics. Variables to be included are: age, sex, race, ethnicity, height, weight, child bearing potential, and other variables as applicable based on indication/study design.

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

3.5.2 Medical History

Medical history will be summarized by primary system organ class (SOC) and preferred term (PT). Medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA, Version 22.0 or higher) terminology.

Medical history will be presented in a listing by patient and summarized in the Safety Population.

3.6 Concomitant Medications

Concomitant medications will be coded using World Health Organization (WHO) Drug Dictionary. Concomitant medications are all medications taken after the start of study consent, during the treatment period (day 1 until 28 days after last dose), including those started before but ongoing at the start of study treatment.

Concomitant medications will be summarized according to Anatomical Therapeutic Class (ATC) and preferred drug name.

Concomitant medications will be presented in a listing by patient and summarized in the Safety Population.

3.7 Treatment Compliance and Exposure

The relative dose intensity (RDI) will be used to assess the treatment compliance. The RDI is defined as:

$$\text{RDI (\%)} = 100 \times \text{ADI (mg/day)} / \text{PDI (mg/day)},$$

where ADI and PDI are the actual dose intensity and planned dose intensity, respectively:

$$\text{PDI (mg/day)} = \text{planned cumulative dose (mg)} / \text{duration of exposure (days)},$$

$$\text{ADI (mg/day)} = \text{actual cumulative dose (mg)} / \text{duration of exposure (days)}.$$

The planned cumulative dose is the planned daily dose amount multiplied by the duration of exposure, while the actual cumulative dose is the sum of actual total daily dose amount over the duration of exposure. The actual total daily dose will need the information of dose modifications (increased and reduced) and dose interruption (held) captured in CRF (Dose Modifications and Interruptions). If a subject does not take any study drug, the actual RDI by definition is zero.

The duration of exposure is defined as:

Duration of exposure (days) = (Date of last dose – Date of first dose) + 1, regardless of unplanned intermittent discontinuations.

One summary table will be generated for treatment exposure and compliance, including: treatment duration (weeks), treatment duration categories (0 to 12 weeks, 12 to 24 weeks, 24 to 28 weeks, 28 to 40 weeks, 40 to 52 weeks), actual cumulative dose, ADI, RDI, and RDI categories (>80%, <=80%; >95%, <=95%).

Another summary table will be generated for dose increased, dose reduced, dose held, and the reasons.

A by patient listing will be generated for the detailed exposure information.

It will be summarized in Safety Population and by group.

3.8 Definition of Efficacy Endpoints

3.8.1 Primary Efficacy Endpoint: CRISS

The primary efficacy endpoint is Combined Response Index in Diffuse Cutaneous Systemic Sclerosis (CRISS).

CRISS is a 2-step process for use in a clinical trial.

Step 1: patients who develop new onset of renal crisis, new onset or worsening of lung fibrosis, new onset of pulmonary arterial hypertension, new onset of left ventricular failure during the trial are considered as not improved and assigned a probability of improving equal to 0.0.

Step 2: For the remaining patients, step 2 involves computing the predicted probability of

improving for each individual, using the equation shown below.

$$\frac{exp\left[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt-glob} - 0.44 * \Delta_{MD-glob} - 3.41 * \Delta_{HAQ-DI}\right]}{1 + exp\left[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt-glob} - 0.44 * \Delta_{MD-glob} - 3.41 * \Delta_{HAQ-DI}\right]}$$

where Δ_{MRSS} indicates the change in mRSS from baseline to follow up, $\Delta_{FVC\%}$ denotes the change in FVC% predicted from baseline to follow up, $\Delta_{Pt-glob}$ indicates the change in patient global assessment, $\Delta_{MD-glob}$ denotes the change in physician global assessment, and Δ_{HAQ-DI} is the change in HAQ-DI. All changes are absolute change ($Time_2$-$Time_{baseline}$).

As patient global assessment and physician global assessment are based on a VAS of 0 (Extremely Poor)-100 (Excellent) in CRF. When calculating the CRISS score, these 2 scores should be divided by 10 firstly to make it in a 0-10 scale.

The outcome is a continuous variable between 0.0 and 1.0 (0-100%). A higher score indicates greater probability of improvement. A CRISS score ≥ 20% is considered as the minimal detectable difference and will be described as CRISS20 in the SAP. A CRISS score >= 60% is considered as the minimally important difference and will be described as CRISS60 in the SAP. This will be the primary variable for efficacy.

Analysis of the primary efficacy endpoint will be conducted using the mITT population. The number and percentage of patients achieving CRISS60 response at week 24 will be summarized.

The following analysis on CRISS will also be provided as supportive evidence for the primary analysis.

- The number and percentage of patients who develop new onset of renal crisis, new onset or worsening of lung fibrosis, new onset of pulmonary arterial hypertension, new onset of left ventricular failure during the trial are considered as not improved and assigned a probability of improving equal to 0.0.
- The numerical CRISS Score at week 24.

3.8.2 Secondary Efficacy Endpoints

3.8.2.1 CRISS

The numerical and categorical CRISS score will also be summarized for Weeks 8, 16, 36, and 52 on the mITT population.

3.8.2.2 mRSS

The mRSS is a measure of skin thickness rated with scores ranging from zero (0; normal) to three (3; severe skin thickening) across 17 different sites. The total score is the sum of the individual skin scores in the 17 body areas (e.g., face, hands, fingers; proximal area of the arms, distal area of the arms, thorax, abdomen; proximal area of the legs, and distal area of the legs, feet), giving a range of 0-51 units. A negative change from baseline demonstrates improvement.

MRSS total score, change from baseline, percent improvement will be summarized as a continuous endpoint at all scheduled assessment visits.

3.8.2.3 PFTs

Onset/progression of ILD will be evaluated by the effect of belumosudil in the change in predicted FVC, which is based on institutional standards and will be measured as part of PFTs, including FEV1, FVC, DL_{CO}, TLC, and RV.

FVC (%) and change from baseline will be summarized as a continuous endpoint at all scheduled assessment visits.

FEV1, FVC, DL_{CO}, TLC, RV and their change from baseline will be summarized as a continuous endpoint at all scheduled assessment visits.

3.8.2.4 Physician Global Assessment

Physician global assessment is physician's assessment of patient's overall health during the last week based on a VAS of 0 (Extremely Poor)-100 (Excellent). A positive change from baseline demonstrates improvement.

Physician global assessment score, change from baseline, percent improvement will be summarized as a continuous endpoint at all scheduled assessment visits.

3.8.2.5 Patient Global Assessment

Patient global assessment is patient's assessment of their overall health during the last week based on a VAS of 0 (Extremely Poor)-100 (Excellent). A positive change from baseline demonstrates improvement.

Patient global assessment score, change from baseline, percent improvement will be summarized as a continuous endpoint at all scheduled assessment visits.

3.8.2.6 Scleroderma Health Assessment Questionnaire-Disability Index (SHAQ-DI)

SHAQ-DI includes the general HAD-DI assessment and 6 scleroderma-specific Visual Analog Scale (VAS) items to explore the impact of participant's disease. The general HAD-DI assessment includes 8 sections: dressing, arising, eating, walking, hygiene, reach, grip, and activities. The 6 VAS items will be rated separately in scale of a 0−100 millimeters [mm], with higher scores indicating more severe disease. The 6 items are: 1) pain, 2) intestinal disease, 3)

breathing problem, 4) Raynaud syndrome, 5) finger ulcers, and 6) overall disease. A negative change from baseline demonstrates improvement.

SHAQ-DI total score, each of the 6 VAS scores and their change from baseline, percent improvement will be summarized as a continuous endpoint at all scheduled assessment visits.

3.8.2.7 Treatment Emergent Adverse Events

Treatment emergent adverse events in subjects with dcSSc at baseline will be presented by number (N) and percentage (%).

3.9 Safety Analysis

Safety assessments include AEs, serious adverse events (SAEs), vital sign measurements, clinical laboratory evaluations (hematology, chemistry, urinalysis, pregnancy tests) and electrocardiograms (ECGs). 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.

Clinically significant PE findings will be captured as AEs.

3.9.1 Adverse Events

AEs will be coded using the MedDRA dictionary (Version 22.0 or higher). TEAEs are any AEs occurring or worsening in severity after the first administration of study medication. All AEs (including SAEs) will be graded using the 5-point CTCAE V5.0 scale (mild, moderate, severe, life threatening, or death). Causality with study treatment will be classified as: definitely related; probably related; possibly related; unlikely related; not related.

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

- SOC and PTs within each SOC in decreasing total frequency
- PT in decreasing total frequency
- SOC, PTs, and maximum severity
- SOC, PTs, and relationship to study drug

These analyses will be repeated for Grade ≥3 TEAEs, SAE, TEAEs leading to dose modification, TEAES leading to dose interruption, and AE leading to study drug discontinuation.

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.

3.9.2 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 scheduled assessment visit.

For parameters of white blood cell counts, neutrophils (absolute count), lymphocytes (absolute count), monocytes (absolute count), Hb, platelets, ALP, ALT, AST, GGT, total bilirubin, GFR, 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 V5.0. 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.

3.9.3 Vital Signs

Descriptive statistics for vital signs (weight, heart rate, body temperature, systolic and diastolic blood pressure, pulse rate, and respiratory rate) values and the change from baseline will be presented for each scheduled assessment visit.

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

3.9.4 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 with triplicate ECGs will be presented for the values and change from baseline scores (QTcF is the QT interval using Fridericia's correction which is calculated by $QTcF = QT/RR^{1/3}$). In addition the maximum change from baseline will be calculated for each subject and summarized for each scheduled assessment visit.

The number and percentage of subjects with observed QTcF values that satisfy the following conditions will be presented:

- ≤ 450 millisecond (ms)

- >450 to ≤ 480 ms
- >480 to ≤ 500 ms
- > 500 ms

The number and percentage of subjects having change from baseline QTcF values that satisfy the following conditions will be presented:

- ≤ 0 ms
- >0 to ≤ 30 ms
- >30 to ≤ 60 ms
- > 60 ms

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

REFERENCES

1. Khanna D, Berrocal VJ, Giannini EH, Seibold JR, Merkel PA, Mayes MD, et al. The American College of Rheumatology provisional composite response index for clinical trials in early diffuse cutaneous systemic sclerosis. Arthritis Rheumatol. 2016;68(2):299–311.

2. Spiera R, Hummers L, Chung L, Frech TM, Domsic R, Hsu V, Furst DE, Gordon J, Mayes M, Simms R, Lafyatis R. Safety and efficacy of lenabasum in a phase II, randomized, placebo‐controlled trial in adults with systemic sclerosis. Arthritis & Rheumatology. 2020 Aug;72(8):1350-60.