Statistical Analysis Plan

A Randomized, Double-Blind, Vehicle-Controlled, Multicenter Study to Assess the Efficacy and Safety of Methyl aminolevulinate hydrochloride (MAL) 16.8% cream (CD06809-41) versus vehicle cream in the treatment of thin and moderately thick, non-hyperkeratotic, non-pigmented actinic keratosis of the face and scalp when using daylight photodynamic therapy (DL-PDT)

STATISTICAL ANALYSIS PLAN

RD.06.SPR.112199

Version Final 2.0

06-APR-2021

STATISTICAL ANALYSIS PLAN APPROVAL SIGNATURES

Author : PPD 06-APR-2021

PPD
GALDERMA R&D - PPD
PPD

Date
dd-MMM-yyyy

Reviewer: 06-APR-2021

PPD
GALDERMA R&D - PPD

Date
dd-MMM-yyyy

Statistical Analysis Plan

DOCUMENT HISTORY

Version	Date	Author	Summary of Changes
Draft version 0.1	30-MAY-2019	PPD	Initial version
Draft version 0.2	04-JUN-2019	PPD	PPD updated the SAP on the basis of the draft version 7.0 of the study protocol and implemented the changes suggested by PPD
Draft version 0.3	05-JUN-2019	PPD	PPD updated the SAP on the basis of the draft version 8.0 of the study protocol
Draft version 0.4	05-JUN-2019	PPD	PPD inserted the measures of compliance agreed with PPD
Draft version 0.5	07-JUN-2019	PPD	PPD updated the SAP on the basis of the draft version 9.0 of the study protocol
Draft version 0.6	12-JUN-2019	PPD	PPD updated the SAP on the basis of the final version 1.0 of the study protocol
Draft version 0.7	12-JUN-2019	PPD	PPD updated the SAP on the basis of the comments provided by PPD
Draft version 0.8	24-JUN-2019	PPD	PPD updated the SAP on the basis of the final version 1.1 of the study protocol
Draft version 0.9	28-JUN-2019	PPD	PPD updated the "Analysis Visits Definition" section
Draft version 0.10	22-AUG-2019	PPD	PPD updated the SAP on the basis of the version 1.4 of the study protocol
Draft version 0.11	06-SEP-2019	PPD	PPD updated the SAP on the basis of the final version 2.0 of the study protocol
Draft version 0.12	09-SEP-2019	PPD	PPD updated the SAP on the basis of the final version 2.1 of the study protocol
Final version 1.0	13-SEP-2019	PPD	PPD issued the final version 1.0
Draft version 1.1	09-JUN-2020	PPD	PPD updated the SAP in order to correct few typos and minor errors in data presentation, to include first and third quartiles in descriptive statistics and to include COVID-19 related information and issued the draft version 1.1

Draft version 1.2	01-JUL-2020	PPD	PPD updated the "Analysis Visits Definition" section, implemented the changes suggested by PPD and issued the draft version 1.2
Draft version 1.3	31-JUL-2020	PPD	PPD updated the study schema, sections 2, 4 and 9 of the SAP on the basis of the final version 3.0 of the study protocol and fixed few minor errors in the section 12
Draft version 1.4	24-MAR-2021	PPD	PPD implemented the following updates: • in section 8.2 it was clarified that non-randomized subjects who received at least one application of study drug by mistake will be listed separately. • in section 10.1.6 it was clarified that all proportions will be presented as percentages • in section 10.2.3 a summary for Fitzpatrick skin type groups was added • in section 10.6.1 it was clarified that confidence intervals of the difference in proportions within analysis centers will be calculated using the Newcombe method without a continuity correction, based on the two individual Wilson confidence intervals • in section 10.6.2.2 a sensitivity analysis to assess the impact of subjects of Site PPD on the primary efficacy results was added • in section 10.6.4.2 the upper threshold of 12 for lesions imputation was added • in section 10.6.4.2 it was clarified that for LOCF imputation baseline value will be carried forward if no post-baseline value is available • in section 10.7.3 the threshold for selecting common TEAE in the safety population and subgroups was changed from 5% to 1%; a table based on the original 5% threshold will be presented for safety population only (no subgroups summaries) • in section 10.7.6 it was added that for summary purposes, laboratory values that are listed as above or below particular thresholds will be numerically imputed as above or below that threshold, respectively, by the minimum measured amount for that parameter • in section 13.3 criterion for PCSV of Blood Urea Nitrogen was added and some errors were fixed
Final version 2.0	06-APR-2021	PPD	PPD issued the final version 2.0

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1 LIST OF ABBREVATIONS

AE	Adverse Event
AESI	Adverse Event of Special Interest
AK	Actinic Keratosis
ALA	5-aminolevulinic acid
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
CI	Confidence interval
COVID-19	Coronavirus Disease 2019
CR	Complete Response
CRO	Contract Research Organization
CS	Clinically Significant
CSR	Clinical Study Report
DL	Daylight
dL	deciliter
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
ERT	eResearch Technology, Inc.
ET	Early Termination
FST	Fitzpatrick Skin Type
GCP	Good Clinical Practice
IC	Informed Consent
ICF	Informed consent form
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IR	Irritant Reaction
IRT	Interactive Response Technology
ITT	Intent-To-Treat
L	liter
LOCF	Last Observation Carried Forward
MAL	Methyl Aminolevinulate Hydrochloride
MAR	Missing At Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram
MI	Multiple Imputations
mmol	millimole
µmol	micromole
MNAR	Missing Not At Random
NRS	Numeric Rating Scale
NT	Not Tested
OC	Observed Case
PDT	Photodynamic Therapy
PpIX	Protoporphyrin IX
PP	Per Protocol
SAE	Serious Adverse Event
SAF	Safety Population
SAP	Statistical Analysis Plan
TEAE	Treatment-emergent adverse event
U	unit
ULN	Upper Limit of Normal
WHO	World Health Organization

2 INTRODUCTION

This Statistical Analysis Plan (SAP) is based on the study protocol final version 3.0 dated 12-JUL-2020. This document describes all the analyses and reporting that will be required for a clinical report purpose and any resulting publications. This SAP has been developed prior to any examination of study data. The analyses and reporting will be performed after the completion of study. Any post hoc, or unplanned, exploratory analyses performed, if included, will be clearly identified as such in the final Clinical Study Report (CSR).

3 STUDY OBJECTIVES

3.1 Primary Objective

The primary objective is to evaluate the safety and efficacy of MAL DL-PDT in adult subjects with mild to moderate actinic keratoses on the face and scalp over a 14-week period, and to show superiority in efficacy of MAL cream over vehicle cream at 12 weeks after the last DL-PDT session (Visit 6).

3.2 Secondary Objectives

- Compare the lesion response between MAL DL-PDT and vehicle DL-PDT.
- Compare partial clearance of actinic keratoses (AKs) between MAL DL-PDT and vehicle DL-PDT.
- Compare subject assessment of pain (by NRS) between MAL DL-PDT and vehicle DL-PDT.
- Assess safety, including adverse events.
- Document the estimated PpIX effective dose for each subject DL-PDT exposure through satellite image data.

4 STUDY DESIGN

This is a phase 3, randomized, double-blind, vehicle-controlled, multicenter, parallel-group study in adult subjects with clinically-confirmed mild to moderate AKs on the face and the balding scalp, to be conducted at approximately 60 clinical centers in the United States. Clinical centers will be selected to ensure that diverse daylight conditions will be represented in the study to understand the effects of latitude, elevation, and climate.

Subjects will be randomized in a 2:1 ratio to receive one of the following two treatments of daylight photodynamic therapy (DL-PDT).

Group 1: CD06809-41 Methyl aminolevulinate hydrochloride (MAL) 16.8% cream.

Group 2: Vehicle cream.

There will be two treatment sessions at least 2 weeks up to 4 weeks apart. The Principal Investigators will apply a thin layer of MAL cream or vehicle cream to each lesion during each treatment. A total of 4-12 lesions will be treated (thin layer on each AK lesion and surrounding 5 to 10 mm of normal skin). At 30 minutes after cream application, subjects will go outside in daylight for 2 hours. After this time, the subject will return inside the investigative center and the cream will be removed by washing the skin with gentle skin cleanser.

This phase 3 study is designed to evaluate the safety and efficacy of MAL DL-PDT in adult subjects with mild to moderate actinic keratosis on the face and scalp over a 14-week period (12 weeks after last PDT session). Male and female subjects aged ≥ 18 years, who have at least 4 but no more than 12 clinically confirmed thin and moderately

thick (Grade I and II in the Olsen scale, respectively), non-hyperkeratotic, non-pigmented AK lesions located in on the face excluding eyelids, lips and mucosa (e.g., forehead, cheek, chin,), and balding scalp will be enrolled in the study.

The expected duration for each subject's participation in the study will be approximately 14 weeks. With potential visit rescheduling due to weather issues and incomplete visits, the duration of a subject's participation could take up to 20 weeks.

The planned duration of recruitment is approximately 9 months.

The planned duration of the clinical study is approximately 12 to 15 months.

Approximately 675 subjects will be screened for a total of 570 subjects to be randomized (380 in the MAL DL-PDT arm and 190 in the vehicle DL-PDT arm using a 2:1 randomization ratio) to have approximately 100 subjects of MAL DL-PDT arm complete the long-term follow-up study, who had complete response for the primary endpoint. Randomization is stratified by study centers using the Interactive Response Technology (IRT) System.

Approximately 60 study centers are planned in the United States with approximately 8 subjects enrolled per center.

All attempts will be made to keep the study center staff and subjects blinded throughout the study. Members of the study center staff, including the procedure operator, will not have access to the randomized treatment assignment.

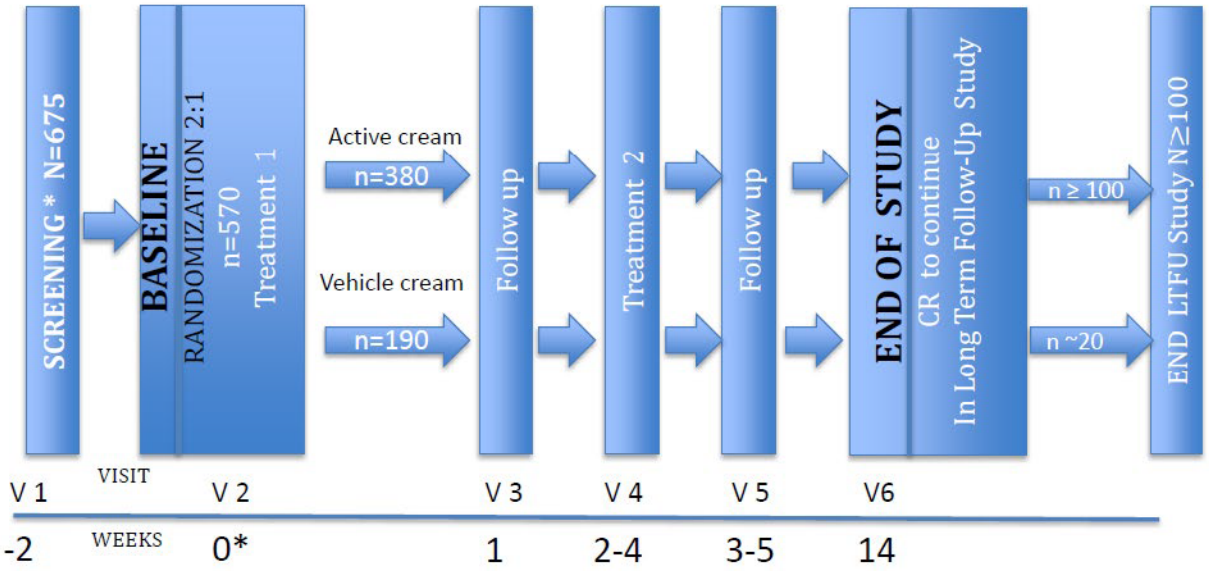
To ensure double-blind administration of study drug the study center pharmacist(s) or other qualified personnel will prepare all MAL 16.8% cream or vehicle cream treatments, according to the current version of the pharmacy manual provided by the IRT system.

The randomization code will remain blinded to all study centers and study team members until completion of the study and after the lock of the study database.

Personnel from Sponsor, CRO, and investigational centers directly involved with the ongoing conduct of the long-term follow-up study will not have access to any information that may lead to unblinding for the ongoing maintenance evaluation.

5 EFFICACY AND SAFETY ASSESSMENTS

Figure 1 - Study Schema



* At Visit 6 (Week 14 or Final Visit), 12 weeks after the last DL-PDT treatment those subjects who achieve Complete Response (CR) of treated lesions will be offered the opportunity to continue in the long-term follow-up study to evaluate recurrence of the treated lesions

Table 1 - Schedule of Assessments

Visit	1	2 [a]	2b [b]	3	4 [a]	4b [b]	5 [a,b]	6 [a,b]
Purpose	Medical history and labs	First DL-PDT session	If required	Follow-up 1 week after 1st DL-PDT (visit 2 or 2b)	Second DL-PDT session	If required	Follow-up 1 week after 2nd DL-PDT (visit 4 or 4b)	Final 12 weeks after last DL-PDT
Week	Screening	Baseline		Week 1	Week 2		Week 3	Week 14/ET [f]
Visit window	-14 to -5 days	0 to +14 days		-2 to +14 days	0 to +14 days		-2 to +14 days	-2 to +28 days
Informed Consent (IC for photography as needed)	X							
Demographics (including FST)	X							
Medical History	X							
Previous Therapies/Procedures	X							
Vital Signs/Physical Examination	X							X
Inclusion/Exclusion Criteria	X							X
Hematology/Blood Chemistry/UA/ECG	X							X
Pregnancy Test [c]	X	X	X		X	X		X
Weather assessment		X	X		X	X		
Photography (Selected sites) [d]		X						X
AK mapping + counting + grading [d]		X	X		X	X		X
Sunscreen application		X	X		X	X		
Lesion débridement and treatment application		X	X		X	X		
Geolocalized Satellite data and exposure time		X	X		X	X		
Study drug(s) Dispensing (D) and Accountability (A)		X	X		X	X		
Subject Assessment of Pain		X	X		X	X		
Subject Skin Aspect Assessment [d]								X
Subject Satisfaction Questionnaire					X	X		X [d]
Safety Visit Question				X			X	

Visit	1	2 [a]	2b [b]	3	4 [a]	4b [b]	5 [a,b]	6 [a,b]
Purpose	Medical history and labs	First DL-PDT session	If required	Follow-up 1 week after 1st DL-PDT (visit 2 or 2b)	Second DL-PDT session	If required	Follow-up 1 week after 2nd DL-PDT (visit 4 or 4b)	Final 12 weeks after last DL-PDT
Week	Screening	Baseline		Week 1	Week 2		Week 3	Week 14/ET [f]
Visit window	-14 to -5 days	0 to +14 days		-2 to +14 days	0 to +14 days		-2 to +14 days	-2 to +28 days
Adverse Events [e]	X	X	X	X	X	X	X	X
Concomitant Therapies/Procedures [d]		X	X	X	X	X	X	X
Subjects with Complete Response evaluation and IC for continuation into long-term follow-up study								X
Exit Form [f]								X

a) Visits 2 and 4 may be delayed for up to 2 weeks in case of unsuitable weather conditions at randomization/treatment outset. These postponements will be automatically added to the time of scheduled Visits 5 and 6

b) If subjects experience rain during the 2 hours of daylight exposure of either DL-PDT visit, they will be instructed to go indoors at the investigative site and the study drug will be washed off. The treatment will be considered incomplete and should be repeated at a minimum interval of 2 weeks, with a maximum interval of 4 weeks. There will be only one attempt at retreatment of an incomplete treatment. These visits will be Visits 2b and 4b. If the attempt at retreatment is also incomplete, the subject will continue in the study. Likewise these repeat visits will be added to the time of scheduled Visits 5 and 6.

c) Only for females of childbearing potential, urine at Visits 1, 2, 2b (if applicable), 4, 4b (if applicable) and 6

d) Should be performed earlier if subject discontinues before Visit 6

e) Adverse Events have to be collected from the time of the Informed Consent signature

f) Exit form should be completed after subject data collection has been completed for subjects in the study

6 EFFICACY AND SAFETY VARIABLES

6.1 Efficacy Variables

6.1.1 Lesion Response

At Visit 2 (before treatment), the Investigator will map the subject's lesions, count them, and report their severity according to criteria based on Olsen (1991) (see Table 2).

At Visits 2b (if applicable), 4, 4b (if applicable) before treatment and at Visit 6/ET, the severity of pre-existing, treated AK lesions that had been identified and treated previously will be recorded. If new AKs develop in the treatment area, they will be recorded, but not treated nor included into the efficacy evaluations.

Table 2 - Lesion Severity Grade Scale

Grade	Severity	Description
Grade 1	Mild	Slightly palpable, better felt than seen
Grade 2	Moderate	Moderately thick, easily felt and seen
Grade 3	Severe	Very thick and/or obvious actinic keratoses

The Investigator will be asked at Visits 2b (if applicable), 4, 4b (if applicable) and 5/ET to identify each lesion treated previously, and the lesion response as a CR or a non-CR, as described in Table 3.

Table 3 - Lesion Response Score

Response	Score	Description
Complete response (CR)	1	Complete disappearance of the lesion, visually and by palpation
Non-complete response (non-CR)	0	Non-complete disappearance of the lesion

If all of the treated lesions in the treatment area are assessed to be CR at 12 weeks after the last DL-PDT session (Visit 6), the subject will be assessed as a subject CR. If any of the treated lesions in the treatment area are assessed to be non-CR, the subject will be assessed as a non-complete responder. Those subjects who are assessed to be CR, i.e., all treated lesions cleared, at Visit 6 will be offered the opportunity to continue in the long-term follow-up study, to assess recurrence of lesions treated.

6.2 Safety Variables

Safety variables will be assessed for all subjects at the visit 1 (upon signing of the ICF) and at every subsequent visit.

6.2.1 Subject Assessment of Pain

At Visit 2, 2b (if applicable), and Visit 4, 4b (if applicable), after the treatment session (once illumination in daylight is completed), the subject will assess the maximal pain felt during daylight exposure. The subject will complete this assessment for all DL-PDT exposures, both complete and incomplete. The pain sensation will be assessed on an 11-point Numerical Rating Scale (NRS), where '0' is no pain at all and '10' is extreme pain as shown in Table 4. The center

personnel will collect and report the subject's self-assessment of pain in the Procedure Log. In the event that a subject requires an intervention for pain such as medication, it will then be recorded as an AE in the eCRF.

Table 4 - Numerical Rating Scale for Subject Self-assessment of Pain

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
No pain at all										**Extreme pain**

6.2.2 Safety Visit Question

At Visits 3 and 5, particular attention is paid to AEs and events which can be sequelae of treatment.

The Investigator will answer the following question at each of Visit 3 and Visit 5:

Does the subject exhibit signs and symptoms of possible contact sensitization? ☐ ***YES*** ☐ ***NO***

If YES then, according to the sections 9.2.4.2 and 9.4.4 of the study protocol, if the Investigator recognizes signs and symptoms of possible contact sensitization at other visits during the study, he/she should report these as an Adverse Event of Special Interest (AESI).

6.2.3 Clinical assessment of the subject's skin aspect

At Visit 6, for each lesion that has responded completely (see Table 3), the Investigator will assess the subject's skin aspect on the following signs: scarring, atrophy, induration, redness or change in pigmentation (see Table 5).

The clinical assessment of skin aspect will be graded according to the following scale.

Table 5 - Scale for Clinical Assessment of Subject's Skin Aspect

Clinical Assessment	Score	Description
Excellent	3	No scarring, atrophy or induration, and no or slight occurrence of redness or change in pigmentation compared to adjacent skin
Good	2	No scarring, atrophy or induration, but moderate redness or change in pigmentation compared to adjacent skin
Fair	1	Slight to moderate occurrence of scarring, atrophy or induration
Poor	0	Extensive occurrence of scarring, atrophy or induration

In the event of the score of 2 or less for the clinical assessment of subject's skin aspect, it will be recorded as an AE in the eCRF.

6.2.4 Adverse Events

Adverse events (AEs) will be recorded during each visit (Visits 1, 2, 3, 4, 5 and 6/ET) and at other visits caused by postponement of treatment or incomplete treatment.

All medical events, whether observed by the Investigator or reported by the subject, and whether or not thought to be treatment-related, will be considered AEs and recorded on the appropriate AE form.

An AE is defined as any untoward medical occurrence in a clinical study, in which a subject is administered a medicinal product that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not it is related to the medicinal (investigational) product. This includes an exacerbation of pre-existing conditions or events, intercurrent illnesses, drug interaction, or the significant worsening of the indication under investigation, that is not recorded elsewhere in the eCRF under specific efficacy assessments.

Refer to the study protocol for further details about AEs definition and management.

6.2.5 Clinical Laboratory Evaluation

The hematology laboratory analyses, clinical chemistry laboratory analyses, and urinalyses will be performed at a central laboratory. Samples for the analyses will be collected at screening visit and visit 6/ET and/or at any unscheduled visit. Reference ranges will be supplied by the central laboratory and used by the investigator to assess the laboratory data for clinical significance and pathological changes.

The Investigator or medically-qualified sub-investigator must review and evaluate laboratory values for each subject in a timely manner. Study centers should refer to the current version of the laboratory manual for laboratory values outside of normal limits. For each out-of-range laboratory result, the Investigator or designee will evaluate whether he/she considers it to be clinically significant (CS), defined as meeting at least 1 of the following conditions:

- The abnormality suggests a disease and/or organ toxicity, or
- The abnormality is of a degree that requires additional active management, e.g., discontinuation of the drug, close observation, more frequent follow-up assessments, or further diagnostic investigation.

If the Investigator observes a CS laboratory result, the test will be repeated as soon as possible and the subject will be monitored until the value returns to normal and/or an adequate explanation for the abnormality is found.

Investigators will also be allowed to repeat specific laboratory test(s) or procedure(s) where he/she suspects an inaccuracy or false result, and which may impact the safety of the subject or the interpretation of the study results; only after discussion with the medical monitor.

All CS out-of-range laboratory values at the screening visit will be recorded (report a diagnosis rather than the laboratory value, whenever possible). All CS out-of-range laboratory values after the screening visit are to be reported as an AE if this abnormality was not present at the screening visit or is assessed as having worsened since the screening visit (i.e., changed significantly from the screening visit). Whenever possible, the Investigator should provide a diagnosis of an AE when reporting the abnormal laboratory value.

Subjects should be reminded to be well hydrated before all visits for phlebotomy purposes. Subjects should fast for at least 8 hours before the visits when blood chemistry testing is planned, except for the screening visit. The screening visit laboratory values must be available before the baseline visit. Laboratory testing conducted in a nonfasting state will not be a protocol deviation.

Total blood volumes to be drawn at each visit are provided in the clinical laboratory manual. Additional samples may be required if medically indicated (e.g., at unscheduled visits for safety reasons, when an abnormal laboratory value is observed and requires a re-test).

6.2.5.1 Hematology

Hemoglobin, hematocrit, white blood cell count (with differential including eosinophils), red blood cell count, platelet count, mean cell volume and INR.

6.2.5.2 Blood Chemistry

Creatinine, AST, ALT, gamma glutamyltransferase, alkaline phosphatase, lactate dehydrogenase, total bilirubin, direct bilirubin, albumin, total protein, uric acid, sodium, potassium, calcium, chloride, glucose, urea, total cholesterol, triglycerides, low-density lipoprotein (LDL), high-density lipoprotein (HDL), creatinine phosphokinase (CPK). Creatinine phosphokinase isoenzyme test will be performed only if CPK is elevated to >2.5 × ULN. The Investigator should also contact the medical monitor in such situations.

For postmenopausal subjects (i.e., no menses for 12 consecutive months), postmenopausal status will be confirmed with a high follicle-stimulating hormone level in the postmenopausal range.

6.2.5.3 Urinalysis

pH, glucose, ketones, blood, protein, leukocytes, nitrites, bilirubin, urobilinogen and specific gravity.

6.2.6 Vital Signs

Vital signs will be evaluated at screening and visit 6/ET before blood sampling and/or at any unscheduled visit. Vital signs will include pulse rate, systolic and diastolic blood pressure (after the subject has been sitting for at least 15 minutes), and body temperature. All abnormal values at the screening visit, identified as CS by the Investigator will be recorded. Any CS changes from the screening visit will be recorded as an AE.

6.2.7 Physical Examination

Complete physical examination should be performed at the screening and visit 6/ET and/or at any unscheduled visit. The examination will include assessments of the head, ears, eyes, nose, throat, neck (including thyroid), skin/integumentary system, cardiovascular system, respiratory system, gastrointestinal system, musculoskeletal system, lymph nodes, and nervous system.

Investigators should assess all abnormal findings for clinical significance. All CS abnormal findings at the screening visit will be recorded in the medical history form. Any CS changes from the screening visit will be recorded as an AE.

6.2.8 Electrocardiograms

A 12-lead Electrocardiogram (ECG) will be performed and read centrally at screening and visit 6/ET and/or at any unscheduled visit using the ECG machine provided. Electrocardiograms for each subject should be obtained using the same electrocardiograph machine whenever possible. Electrocardiograms will be performed in the supine position at the time points described in the schedule of assessments and before any scheduled vital sign measurements and blood draws. Subjects should be monitored for potentially CS ECG results (refer to the current version of the central laboratory manual). Tests with abnormal results that are deemed CS should be repeated to ensure reproducibility of the abnormality. Any abnormalities considered to be CS by the Investigator are to be recorded as AEs and discussed with the medical monitor, as needed.

6.3 Other Variables

6.3.1 Satisfaction Questionnaires

At Visit 4 (4b if applicable, second DL-PDT) and visit 6/ET, the subjects will be asked to complete satisfaction questionnaires (see Appendices 1, section 15.1, and 2, section 15.2, of the study protocol).

The purpose of these questionnaires is to record subjects' feelings and opinions (efficacy, safety, visual aspect of the skin, convenience, etc.) concerning the treatments they received.

6.3.2 Weather Assessment

At Visit 2 (2b if applicable), and Visit 4 (4b if applicable), the visits at which the DL-PDT sessions are performed, the Investigator should ensure that the weather conditions are appropriate to begin the process of DL-PDT. Before randomizing/treating the subject, the Investigator should determine that subjects will be able to spend 2 hours comfortably outdoors in ambient conditions, after the AKs have been prepared for DL-PDT. The Investigator should also assess the likelihood of rain over the next 3 hours by consulting the local weather forecast using the internet (e.g., weather.gov). If it is raining or there is a significant probability of rain in the ensuing 3 hours, treatment should be postponed.

6.3.3 Subject Exposure Time and PpIX Effective Dose

The subject will be asked to stay outside under direct light exposure or shade for 2 consecutive hours.

The exposure should not start less than 2 hours before sunset. Trained personnel from the investigational center will be asked to note all changes in exposure conditions and their duration (Start/End) in order to be able to determine the effective duration of the daylight exposure for a given subject (Subject Exposure Time) (see section 8.4.4.3 of the study protocol). The location of the subject will be geolocalized during the Subject Exposure Time. Records of all assessments linked to Weather assessment and Subject Exposure Time will be reported in the Procedure Log.

From the geolocalization and time of the subject's daylight illumination, satellite data will be used to determine irradiance during the subject exposure time. The PpIX effective dose for each subject's daylight exposure will be calculated using these parameters.

6.3.4 Procedures for Suspected Sensitization (Rechallenge and Patch Ingredient Test)

If a subject experiences suspected skin sensitization (contact allergy), the following actions should be taken to characterize the event:

- Stop the study drug.
- Take a photo of the affected area and the non-affected surrounding skin.
- Document the event as an AESI and report the event to the Syneos Health Safety and Pharmacovigilance group within 24 hours as described for SAE in Section 9.2.4.3 of the study protocol.

a) In case of suspicion of allergic contact dermatitis

- After all signs and symptoms have resolved (after a minimum of 2 weeks), perform a rechallenge test with the assigned study drug. Patch tests will be supplied by Galderma.

- Ensure the subject has not been under any treatment with corticosteroids or antihistamines of any route of administration the week before testing.
- Ensure that the skin on the back has not been exposed to the sun or artificial ultraviolet sources the week before testing.
- Apply an appropriate quantity of the assigned study product to fill in the cupule of the test chamber to a naïve zone on the back either the right or left side of the center line (or the inner forearm if the back cannot be tested). If no test chamber is available on site, patch test units will be provided. The use of semi-occlusive conditions can be preferred depending on the irritant potential of the study product and the intensity of the reaction that was observed. The method to be used should be discussed with the Sponsor, see Table 6.
- Choose a skin site that was not previously involved in the inflammatory skin reaction on the back. Cover the test chamber for 48 hours with a hypoallergenic tape.
- Patient should be informed about avoiding exercise, showers, application of toiletries products, to keep the test system dry
- After 48 hours, remove the test chamber and evaluate the site and take photos after each reading:
  - ✓ at approximately 30 minutes after patch test removal (1st reading) and,
  - ✓ 24 to 48 hours later (i.e., 72 or 96 hours after application) (2nd reading)
  - ✓ if the result of the second reading is equivocal, the Investigator or at the Sponsor's request, may perform an optional 3rd reading at 96 to 120 hours later (i.e., 6 to 7 days after application of the patch)
  - ✓ Pictures of each reading and the reading results should be sent to the Syneos Health and Pharmacovigilance

Table 6 - Patch Test Procedure

Duration of study Product application	**1st reading + Photographs**	**2nd reading + Photographs**	**3rd reading (optional) + Photographs**
48 hours	48 hours after study product application (30 minutes after patch test removal)	72 to 96 hours after study product application (24 to 48 hours after patch test removal)	6 or 7 days after study product application (96 to 120 hours after patch removal)

- Use the following scoring system (Spiewak 2008, see Table 7) used by the International Contact Dermatitis Research Group (ICDRG) at each reading:

Table 7 - Suspected Sensitization - Challenge Reaction Grading Scale

Score	**Morphology**	**Interpretation**
-	No reaction	Negative
?	Faint, non-palpable erythema	Doubtful reaction
+	Palpable erythema (moderate edema or infiltrate), papules not present or scarce, vesicles not present	Weak positive reaction
++	Strong infiltrate, numerous papules, vesicles present	Strong positive reaction
+++	Erythema, infiltration, confluent vesicles, bullae or ulceration	Extreme positive reaction

Score	Morphology	Interpretation
IR	Inflammation sharply limited to the exposed area, lack of infiltrate, small petechiae, pustules, and efflorescences other than papules and vesicles	Irritant reaction
NT		Not tested

- At the last reading, the Investigator will provide an assessment regarding a possible sensitization reaction using the following scale (see Table 8):

Table 8 - Suspected Sensitization - Challenge Conclusion Scale

Score	Sensitization Reaction
0	Negative, might include an irritant reaction
1	Equivocal
2	Positive

- In case of absence of reaction after quotation with the International Contact Dermatitis Research Group (ICDRG scale), the subject may resume treatment if appropriate.
- If the rechallenge is positive or equivocal after quotation with the ICDRG scale, notify Syneos Health immediately. Except specific situations, a new series of patch tests will be initiated as directed by the Sponsor (with individual ingredients at different concentrations if applicable, and possibly negative and positive controls) after a minimum of an additional two weeks (but not later than 6 months) and after all signs and symptoms have resolved. If a rechallenge test is required, repeat this process above beginning with the test chamber step.

b) In case of suspicion of immediate contact skin reaction (such as urticaria)

A case-by-case approach will be applied and the procedure to follow will be discussed with the Sponsor.

7 EFFICACY ENDPOINTS

7.1 Primary Endpoint

- The primary endpoint is the subject complete response, defined as the proportion of subjects with complete clearance of all AK lesions treated, at 12 weeks after the last DL-PDT (Visit 6), comparing MAL cream with vehicle cream.

7.2 Secondary Endpoints

- Lesion complete response, defined as the percent reduction from baseline in the number of cleared treated lesions, at 12 weeks after the last DL-PDT treatment (Visit 6), comparing MAL cream with vehicle cream.
- Subject partial response, defined as the proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions, at 12 weeks after the last DL-PDT treatment (Visit 6), comparing MAL cream with vehicle cream.

8 POPULATIONS ANALYZED

The following analysis populations will be used to analyze the study data.

8.1 Intent-to-Treat Population

The intent-to-treat (ITT) population will comprise all randomized subjects. All primary and secondary efficacy endpoints will be analyzed based on the ITT population. The ITT population will be the primary population for all efficacy analyses.

8.2 Safety Population

The safety (SAF) population will consist of all randomized subjects who received at least one application of study drug. All safety data will be summarized based on the safety population.
All safety data for non-randomized subjects who used study drug mistakenly will be listed separately.
Randomized subjects for whom it is unclear whether they received at least one application of study drug will be included in the safety population as randomized.

8.3 Per Protocol Population

The per protocol (PP) population will comprise all subjects in the ITT population who have no major protocol deviations that would have a significant effect on the efficacy of the study treatment. The PP population will be used for the sensitivity analyses of primary endpoint to assess the robustness of study conclusion.

Potential major protocol deviations may include but are not limited to:

- Eligibility deviations (inclusion/exclusion criteria).
- Improper administration of study medication.
- Noncompliance with study medication due to 2 consecutive incomplete treatments or per the Investigator's discretion.
- Noncompliance with study procedures if the consequence of noncompliance would compromise either the subject's safety and/or the study integrity, primary endpoint, and/or is not in line with Good Clinical Practice (GCP)/ICH guidelines.

- Use of prohibited concomitant therapies.
- Administrative error:
  - ✓ Accidental unblinding;
  - ✓ Medication dispensing errors;
  - ✓ AK lesion counts and evaluation performed by a non-approved evaluator.

The final list of major protocol deviation criteria, subjects who have any major protocol deviations and subjects excluded from the per protocol (PP) population will be documented in the blind review memo before database lock. Decisions made regarding the exclusion of subjects and/or subject data from analyses will be made prior to database lock and unblinding and will be documented in blind review memo.

9 SAMPLE SIZE CONSIDERATION

A long-term follow up study is planned to gather data about recurrences at around the 1 year time point after the DL-PDT treatments.

Subjects who achieve a complete response at Visit 6, 12 weeks after the last DL-PDT treatment, will be offered the opportunity to be followed in a 9-month long-term follow-up study to have treated lesions assessed for recurrence. The sample size calculation is based on providing enough subjects to enable the detection of a treatment difference in the primary endpoint and ensuring approximately 100 subjects of MAL DL-PDT arm complete the long-term follow-up study, who had complete response for the primary endpoint.

Approximately 675 subjects will be screened for a total of 570 subjects to be randomized (380 in the MAL DL-PDT arm and 190 in the vehicle cream DL-PDT arm using a 2:1 randomization ratio) to have approximately 100 subjects of MAL DL-PDT arm complete the long-term follow-up study, who had complete response for the primary endpoint. This study will have more than 90% power to detect a 30% difference between MAL DL-PDT (45% complete response rate) and vehicle DL-PDT (15% complete response rate) and with a type I error of 0.00125 assuming 7% subjects will be non-evaluable at 12 weeks after the last DL-PDT for the primary endpoint, 70% of subjects achieving a complete response at 12 weeks after the last DL-PDT will be enrolled in the LT follow-up study and dropout rate of 10% during LT follow-up study.

10 STATISTICAL METHODS AND DATA CONSIDERATIONS

10.1 General Considerations

10.1.1 Baseline

For statistical analyses purpose, baseline is defined as the last measurement prior to the first DL-PDT treatment (regardless of treatment completion status).

Change from baseline and percent change from baseline will be calculated as:

$$Change\ from\ Baseline = PostBaseline\ Value - Baseline\ Value$$

$$Percent\ Change\ from\ Baseline = \begin{cases} Baseline\ Value \neq 0 \Rightarrow 100 \cdot \dfrac{PostBaseline\ Value - Baseline\ Value}{Baseline\ Value} \\ Baseline\ Value = 0 \Rightarrow Missing \end{cases}$$

10.1.2 Missing Dates Management

- Start Date Imputation of Adverse Events:
  - ✓ Imputation of adverse event end date has to be done before imputation of event start date.
  - ✓ Completely missing: Impute to the first study treatment date.
  - ✓ Missing day and month: Impute to January 1st, unless year is the same as year of first study treatment dose then impute to the first study treatment date.
  - ✓ Missing day: Impute to the 1st of the month, unless month and year are the same as month and year of first study treatment dose then impute to the first study treatment date.
  - ✓ If imputed event start date is after event end date (imputed or not), set the event start date to the imputed event end date.

- Start Date Imputation of Prior/Concomitant Therapies and Medical/Surgical Procedures:
  - ✓ Imputation of therapy/procedure end date has to be done before imputation of therapy/procedure start date.
  - ✓ Completely missing: Impute to the first study treatment date.
  - ✓ Missing day and month: Impute to January 1st, unless year is the same as year of first study treatment dose then impute to the first study treatment date.
  - ✓ Missing day: Impute to the 1st of the month, unless month and year are the same as month and year of first study treatment dose then impute to the first study treatment date.
  - ✓ If imputed therapy/procedure start date is after therapy/procedure end date (imputed or not), set the therapy/procedure start date to the imputed therapy/procedure end date.

- Start Date Imputation of Medical History Diseases:
  - ✓ Imputation of disease end date has to be done before imputation of disease start date.
  - ✓ Completely missing: Leave it missing.
  - ✓ Missing day and month: Impute to January 1st.
  - ✓ Missing day: Impute to the 1st of the month.
  - ✓ If imputed disease start date is after disease end date (imputed or not), set the disease start date to the imputed disease end date.

- End Date Imputation of Adverse Events, Prior/Concomitant Therapies and Medical/Surgical Procedures, Medical History Diseases:
  - ✓ Completely missing and flagged as being ongoing: Leave it missing.
  - ✓ Completely missing and not flagged as being ongoing: Impute to the last contact date.
  - ✓ Missing day and month: Impute to December 31st, unless year is the same as last contact date then impute to the last contact date.
  - ✓ Missing day: Impute to the last day of the month, unless year and month are the same as year and month of last contact date then impute to the last contact date.

10.1.3 Reference Start Date and Analysis Day

The first study treatment date will be the reference start date.

Analysis day will be calculated from the first study treatment date and will be used to show start/end day of assessments, events, therapies or procedures.

In the situation where the assessment/event/therapy/procedure/disease date is fully missing and cannot be imputed (i.e. when event, therapy or procedure is reported as being ongoing), analysis day will be missing.

10.1.4 Descriptive Statistics

For the descriptive statistics, unless otherwise noted, the categorical variables will be summarized by frequency and percentage (n, %) for each response category, and the continuous variables will be summarized using standard descriptive statistics (n, mean, standard deviation, median, first and third quartile, min and max).

10.1.5 Statistical Tests and Confidence Intervals

A two-sided type I error $\alpha = 0.00125$ will be used to declare statistical significance for the primary and the secondary endpoints. Confidence intervals (CIs) will be two-sided with both 99.875% coverage and 95% coverage for the primary and the secondary endpoints. The following flagging conventions will be applied for the p-values of all statistical testing:

- p-values < 0.00125 will be flagged with two asterisks (e.g. "0.00121 **");
- $0.00125 \leq$ p-values < 0.01 will be flagged with one asterisk (e.g. "0.00932 *").

10.1.6 Decimal Precision

Unless otherwise noted, the following rounding conventions will be applied:

- means, medians, first and third quartiles will be rounded and presented to one more decimal digit than the source data;
- confidence intervals of means and medians will be rounded and presented to one more decimal digit than the source data;
- standard deviations and standard errors will be rounded and presented to two more decimal digits than the source data;
- minima and maxima will be presented to the same number of decimal digits as the source data;
- proportions will be presented as percentages (not as fraction of unit);
- percentages greater than or equal to 0.1 will be rounded and presented to one decimal digit, percentages lower than 0.1 and greater than 0 will be presented as '<0.1', percentages equal to 0 will not be presented;
- confidence intervals of percentages will be rounded and presented to one decimal digit;
- standard errors of percentages will be rounded and presented to two decimal digits;
- p-values greater than or equal to 0.00001 will be rounded and presented to five decimal digits, whilst p-values lower than 0.00001 will be presented as '<.00001'.

10.1.7 Software Version

All analyses will be performed using SAS® software Version 9.4 or higher.

10.2 Study Subjects

10.2.1 Disposition of Subjects

The number and percentage of subjects (n, %) will be summarized by treatment group (randomized subjects only) and overall for the following categories:

- Subjects screened;
- Subjects eligible;
- Screen failures;
  - ✓ Reasons for screen failure;
- Screen failures due to COVID-19;
  - ✓ Reasons for screen failure due to COVID-19;
- Subjects randomized;
- Subjects randomized but discontinued before treatment;
  - ✓ Reasons for study discontinuation before treatment;
- Subjects randomized but discontinued before treatment due to COVID-19;
  - ✓ Reasons for study discontinuation before treatment due to COVID-19;
- Subject randomized and treated;
- Subject randomized and treated who completed the study;
- Subject randomized and treated who discontinued the treatment;
- Subject randomized and treated who discontinued the treatment due to COVID-19;
- Subject randomized and treated who discontinued the study;
  - ✓ Reasons for study discontinuation;
- Subject randomized and treated who discontinued the study due to COVID-19;
  - ✓ Reasons for study discontinuation due to COVID-19;
- Subjects randomized affected by COVID-19 related study disruptions;
- Subjects randomized affected by COVID-19 related study disruptions impacting efficacy;
- Subjects randomized affected by COVID-19 related study disruptions impacting safety.

All disposition data will be listed. A listing of all subjects affected by the COVID-19 related study disruptions inclusive of the description of how the individual's participation was altered will be provided.

10.2.2 Analysis Populations

The number and percentage (n, %) of subjects in each analysis population (ITT, SAF, PP) will be summarized by treatment group and overall. Analysis populations data will be listed.

10.2.3 Demographic and Baseline Characteristics

Descriptive summaries of the following demographic data and baseline characteristics will be presented by treatment group and overall using the ITT population.

- Age
- Age group
  - ✓ ≤ 39 years old
  - ✓ 40-64 years old
  - ✓ ≥ 65 years old

- Sex
  - ✓ Male
  - ✓ Female
- Ethnicity
  - ✓ Hispanic or Latino
  - ✓ Not Hispanic or Latino
  - ✓ Not Reported
  - ✓ Unknown
- Race
  - ✓ White
  - ✓ Black or African American
  - ✓ Asian
  - ✓ American Indian or Alaska Native
  - ✓ Hawaiian Native or Other Pacific Islander
  - ✓ Other
  - ✓ Multiple
- Fitzpatrick skin type
  - ✓ Type I
  - ✓ Type II
  - ✓ Type III
  - ✓ Type IV
  - ✓ Type V
  - ✓ Type VI
- Fitzpatrick skin type group
  - ✓ Type I, II and III
  - ✓ Type IV, V and VI
- Number of baseline AK lesions of grade 1
- Number of baseline AK lesions of grade 2
- Number of baseline AK lesions
- Classified number of baseline AK lesions
  - ✓ 4-8 AK lesions
  - ✓ 9-12 AK lesions
- Baseline AK grade
  - ✓ Grade 1 (mild) AK
  - ✓ Grade 2 (moderate) AK

All demographic data and baseline characteristics will be listed.

10.2.4 Accounting of Subjects

Number and percentage (n, %) of subjects for each clinical visit and each analysis visit will be presented by treatment group and overall using the ITT Population.

10.3 Protocol Deviations

Major protocol deviations, major protocol deviations leading to the exclusion from the PP population and major protocol deviations due to COVID-19 will be summarized using frequency and percentage (n, %) for each deviation coded term by treatment group and overall using the ITT population. All protocol deviations will be listed.

10.4 Medical History, Prior and Concomitant Therapies and Procedures

For statistical analysis purposes, prior therapies/procedures are defined as those ending before the day of first treatment; concomitant therapies/procedures are defined as those starting before the day of first treatment and ongoing on the day of first treatment and as those starting on the day of first treatment or after. If a subject does not undergo treatment, all therapies/procedures of that subject will be classified as prior.

Previous and concomitant therapies/medications will be coded using WHO Drug Dictionary (March 1, 2019, B3/C3 format). Medical history diseases and prior and concomitant medical/surgical procedures will be coded using MedDRA dictionary (version 22.0).

A summary table will be provided for each of the following using ITT population by treatment group and overall:

- Number and percentage (n, %) of subjects who had medical history diseases by System Organ Class and Preferred Term.
- Number and percentage (n, %) of subjects who had prior therapies/medications by ATC levels 2, 3 and Preferred Term.
- Number and percentage (n, %) of subjects who had concomitant therapies/medications by ATC levels 2, 3 and Preferred Term.
- Number and percentage (n, %) of subjects who had prior medical/surgical procedures by System Organ Class and Preferred Term.
- Number and percentage (n, %) of subjects who had concomitant medical/surgical procedures by System Organ Class and Preferred Term.

Listings of all medical history diseases, prior and concomitant therapies/medications and medical/surgical procedures will be provided.

10.5 Compliance

Since the subjects will receive MAL cream or vehicle cream at the investigational center, and have it applied by the investigator, treatment compliance will be assessed through the actual treatment sessions and adherence to the different steps of the procedure and timing. Non-compliances will be reported as protocol deviations.

Study duration (expressed as number of days) will be calculated as the date of end of participation minus the date of first treatment (or the date of randomization, for randomized subjects not undergoing treatment) plus one and will be summarized by descriptive statistics using ITT population by treatment group and overall.

Treatment duration (expressed as number of days) will be calculated as the date of last treatment (complete or incomplete) minus the date of first treatment (complete or incomplete) plus one and will be summarized by descriptive statistics using ITT population by treatment group and overall.

The number and percentage (n, %) of subjects who did not complete neither the first nor the second treatment, the number and percentage (n, %) of subjects who completed the first treatment only, the number and percentage (n, %) of subjects who completed the second treatment only and the number and percentage (n, %) of subjects who completed both first and second treatments will be summarized using ITT population by treatment group and overall.

All compliance data will be listed.

10.6 Efficacy Analysis

Primary inference for efficacy analysis will be based on the ITT population at week 14, 12 weeks after the last DL-PDT (Visit 6).

All efficacy data (including efficacy data of incomplete treatments) will be listed.

10.6.1 Analysis of the Primary Endpoint

The hypothesis test for the primary endpoint can be formally defined as follows:

$$\begin{cases} H_0: \Pi_{MAL} = \Pi_{Vehicle} \\ H_a: \Pi_{MAL} \neq \Pi_{Vehicle} \end{cases}$$

where Π_{MAL} is the proportion of subjects achieving a complete response at 12 weeks after the last MAL DL-PDT and $\Pi_{Vehicle}$ is the proportion of subjects achieving a complete response at 12 weeks after the last vehicle DL-PDT.

The hypothesis test for the primary endpoint will be evaluated on the ITT population at the two-sided significance level $\alpha = 0.00125$. Efficacy will be claimed if the between treatment difference on the primary endpoint is statistically significant with a p-value < 0.00125. This result will provide an acceptable level of evidence of efficacy, in presence of good internal consistency across primary and secondary endpoints in the absence of a second pivotal study.

The proportion of subjects with complete clearance of all AK lesions treated at each post-baseline analysis visit will be summarized descriptively by treatment groups (overall and by analysis centers). The proportion of subjects with complete clearance of all AK lesions treated at 12 weeks after the last DL-PDT (Visit 6) will be analyzed using the Cochran-Mantel-Haenszel (CMH) test stratified by analysis center (see definition in section 10.6.4.4). No continuity correction will be used. The p-value for the treatment comparison (MAL DL-PDT vs. vehicle DL-PDT) will be calculated from the general association statistic of the stratified CMH test. Difference in proportions between treatment groups and the 99.875% and 95% confidence intervals of the difference will be based on the large sample approximation method for binary data and will be calculated using the method described in the Huiping Zhang "Proportion difference and confidence interval based on Cochran-Mantel-Haenszel method in stratified multi-center clinical trial", 2016, PharmaSUG China 2016, Paper 25 [1]. In addition, the proportion of subjects with complete clearance of all AK lesions treated at 12 weeks after the last DL-PDT (Visit 6) for each treatment group will also be presented.

For the primary analysis, all subjects on the ITT population with missing data for the primary endpoint will be classified as non-responders regardless of treatment allocation (see section 10.6.4.2).

The consistency of the treatment effect across analysis centers will be evaluated using graphical methods (Forest plot). Confidence intervals of the difference in proportions within the analysis centers will be calculated using the Newcombe method without a continuity correction, based on the two individual Wilson confidence intervals.

A bar chart and line plots will be produced over time for each treatment group to summarize the complete responders.

10.6.2 Sensitivity Analyses of Primary Endpoint

To assess the robustness of the primary efficacy results, the following sensitivity analyses will be conducted:

1. The primary analysis will be repeated on the ITT population by imputing missing data using MI (Multiple Imputation) under the Missing At Random (MAR) assumption (see section 10.6.4.2 for further details).
2. The primary analysis will be repeated on the ITT population by imputing missing data using MI (Multiple Imputation) based on the Pattern-Mixture Model under the missing not at random (MNAR) assumption, by using the profiles from vehicle DL-PDT subjects with observed data to impute missing data (see section 10.6.4.2 for further details).
3. The primary analysis will be repeated on the ITT population by imputing missing data using MI (Multiple Imputation) based on the Pattern-Mixture Model under the missing not at random (MNAR) assumption in order to perform a tipping-point analysis (see section 10.6.4.2 for further details).
4. The primary analysis will be repeated on the ITT population by imputing missing data using Last Observation Carried Forward (LOCF).
5. Observed Case (OC) analysis on the ITT population.
6. The primary analysis will be repeated on the PP population.

10.6.2.1 Impact on the Primary Endpoint of COVID-19 related study disruptions

To assess the impact on the primary efficacy results of COVID-19 related study disruptions (e.g. treatment discontinuation, missing assessments, remote visits, etc.), the following sensitivity analysis will be conducted:

1. The primary endpoint will be analyzed using a logistic regression model with treatment, analysis center and occurrence flag of COVID-19 related study disruptions impacting efficacy results as fixed effects. All subjects on the ITT population with missing data for the primary endpoint will be classified as non-responders regardless of treatment allocation (see section 10.6.4.2).
2. The primary endpoint will be analyzed using a logistic regression model with treatment, analysis center, occurrence flag of COVID-19 related study disruptions impacting efficacy results and interaction of treatment and occurrence flag of COVID-19 related study disruptions impacting efficacy results as fixed effects. All subjects on the ITT population with missing data for the primary endpoint will be classified as non-responders regardless of treatment allocation (see section 10.6.4.2). The interaction of treatment and occurrence flag of COVID-19 related study disruptions impacting efficacy results will be deemed statistically significant if the p-value for the interaction term is lower than 0.1 and, in this case, the primary analysis will be repeated separately in the two subgroups of subjects affected and not affected by COVID-19 related study disruptions impacting efficacy results.

10.6.2.2 Impact on the Primary Endpoint of subjects of Site PPD

Site PPD (Dr. PPD) was non-responsive. This prevented Source Data Verification from being performed. The eCRFs of subjects screened and randomized at this site were not signed by the investigator and queries were not answered. Data cannot be verified and queried.

To assess the impact of subjects of Site PPD on the primary efficacy results, a sensitivity analysis for the primary endpoint will be performed on the ITT population by repeating the primary analysis with the exclusion of subjects randomized at this site.

10.6.3 Analysis of Secondary Endpoints

The hypothesis tests for the secondary endpoints can be formally defined as follows:

$$\begin{cases} H_0: \mu_{MAL} = \mu_{Vehicle} \\ H_a: \mu_{MAL} \neq \mu_{Vehicle} \end{cases}$$

where μ_{MAL} is the mean percent reduction from baseline in the number of cleared treated lesions at 12 weeks after the last MAL DL-PDT treatment (Visit 6) and $\mu_{Vehicle}$ is the mean percent reduction from baseline in the number of cleared treated lesions at 12 weeks after the last vehicle DL-PDT treatment (Visit 6) and

$$\begin{cases} H_0: \pi_{MAL} = \pi_{Vehicle} \\ H_a: \pi_{MAL} \neq \pi_{Vehicle} \end{cases}$$

where π_{MAL} is the proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions at 12 weeks after the last MAL DL-PDT treatment (Visit 6) and $\pi_{Vehicle}$ is the proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions at 12 weeks after the last vehicle DL-PDT treatment (Visit 6).

The hypothesis tests for the secondary endpoints will be evaluated on the ITT population at the two-sided significance level $\alpha = 0.00125$. The adjustment for multiple comparisons is described in section 10.6.4.5.

The number of cleared treated lesions at each analysis visit, the reduction from baseline and the percent reduction from baseline in the number of cleared treated lesions at each post-baseline analysis visit will be summarized descriptively by treatment groups (overall and by analysis centers). The percent reduction from baseline in the number of cleared treated lesions at 12 weeks after the last DL-PDT treatment (Visit 6) will be analyzed using an ANCOVA with treatment, analysis center (see definition in section 10.6.4.4) and baseline AK lesion count as fixed effects; the difference in percent reduction between MAL DL-PDT and vehicle DL-PDT, the 99.875% and 95% confidence intervals of the difference and the p-value will be generated from the ANCOVA model.

For analysis of percent reduction from baseline in the number of cleared treated lesions at 12 weeks after the last DL-PDT treatment (Visit 6), missing counts of lesions will be imputed using MI (Multiple Imputation) under the Missing At Random (MAR) assumption. It is expected that the pattern of missing data will be monotonic, with slight deviations being corrected by the Markov Chain Monte Carlo (MCMC) method of the MI procedure. Linear regression will be employed to model the missing lesion count data, with the following covariates included in the imputation model: treatment and non-missing data from earlier time points (see section 10.6.4.2).

Line plots will be produced over time for each treatment group to summarize the mean percent reduction from baseline in the number of cleared treated lesions.

The proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions at each post-baseline will be summarized descriptively by treatment groups (overall and by analysis centers). The proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions at 12 weeks after the last DL-PDT treatment (Visit 6) will be analyzed using the Cochran-Mantel-Haenszel (CMH) test stratified by analysis center (see definition in section 10.6.4.4). No continuity correction will be used. The p-value for the treatment comparison (MAL DL-PDT vs. vehicle DL-PDT) will be calculated from the general association statistic of the stratified CMH test. Difference in proportions between treatment groups and the 99.875% and 95% confidence intervals of the difference will be based on the large sample approximation method for binary data and will be calculated using the method described in the Huiping Zhang "Proportion difference and confidence interval based on Cochran-Mantel-Haenszel method in stratified multi-center clinical trial", 2016, PharmaSUG China 2016, Paper 25 [1]. In addition, the proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions at 12 weeks after the last DL-PDT treatment (Visit 6) for each treatment group will also be presented.

For the analysis of proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions at 12 weeks after the last DL-PDT treatment (Visit 6), all subjects of the ITT population with missing data will be classified as non-responders regardless of treatment allocation (see section 10.6.4.2).

A bar chart and line plots will be produced over time for each treatment group to summarize the partial responders.

10.6.4 Statistical and Analytical Issues

10.6.4.1 Adjustment for Covariates

The analysis of percent reduction from baseline in the number of cleared treated lesions at 12 weeks after the last DL-PDT treatment (Visit 6) will use an adjustment for the number of baseline AK lesions as described in section 10.6.3.

No other adjustment for covariates is planned for the primary and secondary efficacy analyses.

10.6.4.2 Handling of Dropouts or Missing Data

For the primary analysis, all subjects on the ITT population with missing data for the primary endpoint will be classified as non-responders regardless of treatment allocation.

For the sensitivity analysis of the primary endpoint and for the analysis of percent reduction from baseline in the number of treated AK lesions, missing data will be imputed using MI (Multiple Imputation) under the Missing At Random (MAR) assumption. The following steps will be followed:

1. For AK lesion count, the missingness pattern in the data will be evaluated. If the pattern is not monotone, the MCMC method of SAS PROC MI will be used to make it monotone. The single chain method will be used, with 200 burn-in iterations and 100 iterations between imputations. The minimum and maximum values for imputed variables will be set to 0 and 12 respectively, in order to force PROC MI to redraw another value for imputation when an intended imputed value is less than the 0 or greater than 12. Imputed values will be rounded to the nearest integer. The seed number will be set to 112199 and fifty (50) imputations will be created.
2. SAS PROC MI will be used for imputing missing values of data with monotone missing pattern. If the MCMC method of step 1 was previously employed, one imputation will be made using each of the fifty (50) MCMC-imputed datasets. If the MCMC method of step 1 was not previously employed, fifty (50) imputations will be created assuming the data are Missing At Random. The seed number will be set to 112199. These imputations will use the following models:
    a. For binary outcome, a logistic regression model will be used with covariates for treatment and non-missing AK lesion count from earlier scheduled time points including baseline.
    b. For AK lesion count, a linear regression model will be used with covariates for treatment and non-missing AK lesion count from earlier scheduled time points including baseline. The minimum and maximum values for imputed variables will be set to 0 and 12 respectively, in order to force PROC MI to redraw another value for imputation when an intended imputed value is less than the 0 or greater than 12. Imputed values will be rounded to the nearest integer.
3. The imputed datasets will be analyzed as specified in sections 10.6.1 and 10.6.3.
4. The resulting analysis on the imputed datasets will then be combined to produce a single set of statistics as follows:
    a. For binary outcome, the results from the CMH analysis will be combined using the procedure by Rubin [2] and Li et al. [3] to produce a pooled CMH statistic and p-value. The differences in proportions and standard errors will be combined using the SAS PROC MIANALYZE. The resulting pooled difference and standard error will be used to produce the confidence interval based on the large-sample approximation method for binary data without using continuity correction. Both these methods will be used as described in the Bohdana Ratitch, et al. "Combining Analysis Results from Multiply Imputed Categorical Data", 2013, PharmaSUG Proceedings, Paper SP-03 [4].

b. For AK lesion count, results from the ANCOVA analysis will be combined using the SAS PROC MIANALYZE.

The number of fifty (50) imputations was selected in order to prevent a power falloff due to choosing a number of imputations too small [5].

As additional sensitivity analysis of the primary endpoint, the multiple imputation process for complete response binary outcome will be repeated based on the Pattern-Mixture Model under the missing not at random (MNAR) assumption, by using the profiles from vehicle DL-PDT subjects with observed data to impute missing data. The process will be the same described for the MAR analysis, only the assumption about missing mechanism will be changed.

The multiple imputation process for complete response binary outcome will be repeated based on the Pattern-Mixture Model under the missing not at random (MNAR) assumption in order to perform a tipping-point analysis for the primary endpoint by varying independently the assumptions about the missing outcomes on the two arms and including scenarios where dropouts on MAL DL-PDT arm have worse outcomes than dropouts on Placebo DL-PDT arm with the aim of exploring the plausibility of missing data assumptions under which the conclusions change (i.e. under which there is no longer evidence of efficacy). The process will be the same described for the MAR analysis, only the assumption about missing mechanism will be changed and points 2, 3 and 4 will be repeated using two independent sequences of shift parameters (one sequence for each treatment arm) that adjust the imputed values for observations in the two treatment arms until the combined results provide no longer evidence of efficacy [6]. The plausibility of the shift parameters under which the conclusions change and, thus, the reliability of the results obtained under MAR assumption will be discussed.

Moreover, the primary analysis will be repeated on the ITT population by imputing missing data using Last Observation Carried Forward (LOCF). If no post-baseline value is available, baseline value will be carried forward.

10.6.4.3 Interim Analyses and Data Monitoring

No interim analysis is planned for this study.

10.6.4.4 Multicenter Studies

Prior to database lock, a review of the blinded data will be performed to determine the size of each center. If there are centers with a small number of randomized subjects, then these centers will be pooled in order for analyses to be carried out. The process of combining centers will be based on the ITT population, and same pooling will be repeated for PP population.

A small center is defined as a center which randomizes less than 12 subjects. First, centers will be sorted by latitude zone (based on geographic and climatic similarities), number of randomized subjects (descending order) and center number (ascending order). Pooling will start with combining the largest of the set of small centers of a latitude zone with the smallest center within the same latitude zone. If there is a further need to combine data (the size of the pooled centers includes less than 12 subjects), the next smallest center will be combined with the next largest of the small centers, until the criterion of a minimum of 12 subjects is met. The process will continue until all pooled centers have a minimum of 12 subjects within the same latitude zone. Any remaining small centers of a latitude zone will be pooled with the last pooled center within the same latitude zone. The pooled centers and the remaining original unpooled clinical centers will be referred to as 'analysis centers' and will be used as stratification factor in the statistical analyses.

If at the start of pooling any latitude zone has less than 12 subjects in the ITT population in total, then centers will be added to the list of small centers in another latitude and then combined as above. This decision will be documented in clinical report.

The consistency of the treatment effect across analysis centers for the primary endpoint will be evaluated using graphical methods (Forest plot).

10.6.4.5 Multiple Comparison/Multiplicity

In order to maintain the overall type I error rate at 0.00125, a predefined hierarchal testing procedure will be implemented to test the MAL cream against vehicle.

Testing will start from primary endpoint and descend in the order described as below for the secondary endpoints until the null hypothesis fails to be rejected at 0.00125 level of significance.

1. Lesion complete response, defined as the percent reduction from baseline in the number of cleared treated lesions, at 12 weeks after the last DL-PDT treatment (Visit 6), comparing MAL cream with vehicle cream.
2. Subject partial response, defined as the proportion of subjects with 75% or greater reduction from baseline in the number of cleared treated lesions, at 12 weeks after the last DL-PDT treatment (Visit 6), comparing MAL cream with vehicle cream.

10.6.4.6 Use of an Efficacy Subset of Patients

The classification of the protocol deviations and the exclusion of subjects from the PP population will be determined prior to breaking the study blind (see section 8.3).

The analysis on the PP population will allow evaluating the impact of major protocol deviations on the estimation of the treatment effect.

10.6.4.7 Active-Control studies intended to show equivalence

Not applicable.

10.6.4.8 Examination of Subgroups

All subgroup analyses are exploratory in nature.

To determine whether the treatment effect is consistent across various subgroups, the estimate of the between-treatment effect (with nominal 99.875% and 95% CI) for primary endpoint and secondary endpoints will be presented within each category of the following classification variables:

- AK lesions location
  - ✓ Face
  - ✓ Balding scalp
- Baseline AK grade
  - ✓ Grade 1 (mild) AK
  - ✓ Grade 2 (moderate) AK
- Sex
  - ✓ Male
  - ✓ Female
- Number of baseline AK lesions
  - ✓ 4-8 AK lesions
  - ✓ 9-12 AK lesions

- Fitzpatrick skin type
  - ✓ Type I, II and III
  - ✓ Type IV, V and VI

Differences in proportions and their nominal 99.875% and 95% confidence intervals by category for the classification variables listed above will be reported as well as presented graphically (Forest plot) on the ITT population.

Missing data will be imputed using the same methods foreseen for the primary analysis of the primary endpoint and for the analysis of the secondary endpoints (see section 10.6.4.2).

10.6.4.9 Analysis Visits Definition

Efficacy, safety and other variables will be summarized and analyzed by analysis visit. Analysis visit will be slotted according to the following table (see Table 9) to summarize the data by proper visit.

Table 9 - Analysis Visits Definition

Clinical Visit	Analysis Visit	Analysis Visit Number	Target Study Day
Visit 1	Baseline	1	1
Visit 2	Baseline	1	1
Visit 2b	Baseline	1	1
Visit 3	Week 1	2	8
Visit 4	Week 2-4	3	15
Visit 4b	Week 2-4	3	15
Visit 5	Week 3-5	4	22
Visit 6	Week 14/Final	5	99
Early Termination	Early Termination	6	---

Visit 2b and Visit 4b will occur only in case of incomplete treatment at Visit 2 and Visit 4 respectively. All efficacy assessments performed on Visits 2b and 4b will be assumed to fall at Visits 2 and 4 regardless of actual time.

Early termination visit assessments will be summarized as a separate visit in all by-visit outputs.

Unscheduled visits will be assigned to the same analysis visit of the last scheduled visit (including re-treatment visits) that occurred before the unscheduled visits.

All post-baseline safety assessments will be used for potentially clinically significant value (PCSV) determinations.

For efficacy and safety endpoints, if two or more assessments (include both scheduled and unscheduled assessments) are available for the analysis visits then all assessments will be listed and the following rules will be applied for determining the values to be used for the summary and analyses.

For efficacy endpoints:

- Baseline Analysis Visit: in presence of multiple assessments, the last assessment performed before the first DL-PDT treatment (regardless of treatment completion status) will be used for the summaries and analyses.
- Week 2-4 Analysis Visit: in presence of multiple assessments, the first assessment performed after the first DL-PDT will be used for the summaries and analyses.
- Week 14/Final Analysis Visit (where applicable): in presence of multiple assessments, the first assessment performed after the last DL-PDT will be used for the summaries and analyses.

For safety assessments:

- Clinical Laboratory Evaluation: in presence of multiple assessments on an analysis visit, the latest assessment will be used for the summaries and analyses.
- Vital Signs, ECG, Physical Examination: in presence of multiple assessments on an analysis visit, the assessment closest to the target visit date will be used for the summaries and analyses.
- Pain Assessment: in presence of multiple assessments on an analysis visit, the worst assessment (i.e. the one with the higher NRS score) will be used for the summaries and analyses.

For other assessments:

- Satisfaction Questionnaire: in presence of multiple assessments on an analysis visit, the first assessment performed after the last DL-PDT will be used for the summaries and analyses.

10.7 Safety Analysis

No formal inferential analysis is planned for the baseline and safety data, and only summary statistics will be provided.

10.7.1 Subject Assessment of Pain

Pain data will be summarized by treatment group using standard descriptive statistics (n, mean, standard deviation, median, first and third quartile, min, max) and using frequency and percentage (n, %) for each NRS score. Missing pain data will not be imputed. The Hodges-Lehmann estimator of the median difference in pain scores between treatment groups and the related 95% Moses distribution-free confidence interval will presented by visit for descriptive purposes only.

10.7.2 Extent of Exposure

The number of complete and incomplete treatments per subject and the number of complete treatments per subject will be summarized using descriptive statistics (n, mean, standard deviation, median, first and third quartile, min and max) by treatment group and overall.

All treatment applications data will be listed.

10.7.3 Adverse Events Analysis

All Adverse Events (AEs) will be coded using MedDRA dictionary (version 22.0).

A treatment-emergent adverse event (TEAE) is defined as an adverse event that occurred, or worsened, on or after the date of the first application of study drug.

All treatment-emergent adverse events will be listed in a by-subject listing which will include both the term reported on the eCRF (verbatim term) and the MedDRA System Organ Class and Preferred Term. Relative start and stop days will be included along with the actual onset and resolution dates. Pre-treatment AEs will be listed separately.

If relationship to study drug/procedures is missing the closest relationship, i.e. "Related", will be imputed. If severity is missing the greatest severity, i.e. "Severe", will be imputed.

A summary table of TEAEs will be provided for each of the following in the safety population by treatment group and overall:

- Number and percentage (n, %) of subjects with any TEAE and frequency of TEAEs overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE and frequency of TEAEs, that occurred in >1% of subjects of the safety population, overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any AESI and frequency of AESIs overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE by closest relationship to study drug and frequency of TEAEs by closest relationship to study drug overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE by closest relationship to study procedures and frequency of TEAEs by closest relationship to study procedures overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE related to the study drug and frequency of TEAEs related to the study drug by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE related to the study procedures and frequency of TEAEs related to the study procedures by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE by greatest severity and frequency of TEAEs by greatest severity overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE leading to study discontinuation and frequency of TEAEs leading to study discontinuation overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any TEAE leading to study drug discontinuation and frequency of TEAEs leading to study drug discontinuation overall and by System Organ Class and Preferred Term.

The same summaries of TEAEs will be provided for the following subgroups: baseline AK grade (Grade 1 AK and Grade 2 AK), sex (Male and Female), number of baseline AK lesions (4-8 AK lesions and 9-12 AK lesions) and Fitzpatrick skin type (Type I, II and III and Type IV, V and VI).

A summary table of TEAEs will be provided for the following in the safety population by treatment group and overall:

- Number and percentage (n, %) of subjects with any TEAE and frequency of TEAEs, that occurred in >5% of subjects of the safety population, overall and by System Organ Class and Preferred Term.

A summary table of serious TEAEs will be provided for each of the following in the safety population by treatment group and overall:

- Number and percentage (n, %) of subjects with any Serious TEAE and frequency of Serious TEAEs overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any Serious TEAE by closest relationship to study drug and frequency of Serious TEAEs by closest relationship to study drug overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any Serious TEAE by closest relationship to study procedures and frequency of Serious TEAEs by closest relationship to study procedures overall and by System Organ Class and Preferred Term;

- Number and percentage (n, %) of subjects with any Serious TEAE related to the study drug and frequency of Serious TEAEs related to the study drug by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any Serious TEAE related to the study procedures and frequency of Serious TEAEs related to the study procedures by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any Serious TEAE leading to study discontinuation and frequency of Serious TEAEs leading to study discontinuation overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any Serious TEAE leading to study drug discontinuation and frequency of Serious TEAEs leading to study drug discontinuation overall and by System Organ Class and Preferred Term;
- Number and percentage (n, %) of subjects with any Serious TEAE leading to death and frequency of Serious TEAEs leading to death overall and by System Organ Class and Preferred Term.

In order to evaluate the impact of COVID-19 on the safety of the subjects, a summary table of TEAEs and serious TEAEs will be provided for each of the following for the safety population by treatment group and overall:

- Number and percentage (n, %) of subjects with any TEAE related to COVID-19 and frequency of TEAEs related to COVID-19 by System Organ Class and Preferred Term.
- Number and percentage (n, %) of subjects with any Serious TEAE related to COVID-19 and frequency of Serious TEAEs related to COVID-19 by System Organ Class and Preferred Term.

Listings of all AEs leading to death, all serious AEs, all TEAEs leading to study discontinuation, all AESIs and all AEs due to COVID-19 will be provided.

10.7.4 Physical Examination Analysis

Physical examination assessments will be summarized in terms of number and percentage (n, %) of subjects with 'Normal', 'Abnormal/Not Clinically Significant' and 'Abnormal/Clinically Significant' results for each body system.

Summaries will be presented by treatment group and overall at baseline and visit 6/ET.

A listing of all physical examination assessments will be provided.

10.7.5 Vital Signs Analysis

Descriptive statistics (n, mean, standard deviation, median, first and third quartile, min and max) by treatment group and overall will be presented for vital signs at baseline and visit 6/ET and for their change from baseline to visit 6/ET.

Number and percentage (n, %) of subjects with Clinically Significant Abnormal Values (as identified by the Investigator) and with Potentially Clinically Significant Values (see section 13.2) will be summarized by treatment group and overall at baseline and visit 6/ET.

A listing of all vital signs will be provided.

10.7.6 Clinical Laboratory Tests Analysis

Descriptive statistics (n, mean, standard deviation, median, first and third quartile, min and max) by treatment group and overall will be presented for blood chemistry, hematology and quantitative urinalysis parameters at baseline and visit 6/ET and for their change from baseline to visit 6/ET.

For summary purposes, laboratory values that are listed as above or below particular thresholds will be numerically imputed as above or below that threshold, respectively, by the minimum measured amount for that parameter. For example, if a parameter is measured to two decimal places, and has a result of "> 5" then, for summary purposes, the value of 5.01 will be used.

Number and percentage (n, %) of subjects for each response category of qualitative urinalysis parameter will be presented by treatment group and overall at baseline and visit 6/ET.

Number and percentage (n, %) of subjects with laboratory test values below, within and above the laboratory reference ranges will be summarized by treatment group and overall at baseline and visit 6/ET.

Number and percentage (n, %) of subjects with Clinically Significant Abnormal Values (as identified by the Investigator) and with Potentially Clinically Significant Values (see section 13.3) will be summarized by treatment group and overall at baseline and visit 6/ET.

Shift tables from baseline to visit 6/ET describing shifts to abnormality will be provided as well. Only subjects with a baseline result and a result at visit 6/ET for the parameter will be considered.

A listing of all clinical laboratory test will be provided.

10.7.7 Electrocardiograms Analysis

The number and percentage (n, %) of subjects who have Investigator's interpretation of ECG 'Normal', 'Abnormal/Not Clinically Significant' and 'Abnormal/Clinically Significant' will be presented by treatment group and overall at baseline and visit 6/ET.

A listing of Investigator's interpretation of ECG results and a listing of ERT ECG results will be provided.

10.7.8 Safety Visit Question

The answers Yes/No to the question "Does the subject exhibit signs and symptoms of possible contact sensitization?" at Visit 3 and Visit 5 will be listed.

10.8 Other Assessments Analysis

10.8.1 Satisfaction Questionnaires Analysis

Number and percentage (n, %) of subjects for each answer of each question of the satisfaction questionnaires will be presented by treatment group and overall at visit 4 (4b if applicable, second DL-PDT) and visit 6/ET.

A listing of all satisfaction questionnaires data will be provided.

10.8.2 Weather Assessment

A listing of all weather assessment data will be provided.

10.8.3 Subject's Skin Aspect Analysis

Scores of clinical assessment of subject's skin aspect collected at Visit 6/ET will be summarized by treatment group and overall by reporting the number and percentage of subjects (n, %) for each score. For the aim of summary, the worst score among all the assessed lesions of each subject will be taken into account.

10.8.4 Subject Exposure Time and PpIX Effective Dose Analysis

Descriptive statistics (n, mean, standard deviation, median, first and third quartile, min and max) by treatment group and overall will be presented for subject exposure time and PpIX effective dose at visit 2 (2b, if applicable) and visit 4 (4b, if applicable).

A listing of all subject exposure times and PpIX effective doses will be provided.

10.8.5 Suspected Sensitization (Rechallenge and Patch Ingredient Test)

Rechallenge patch test and ingredient patch test information will be presented in separate listing.

11 CHANGES FROM THE PROTOCOL ANALYSIS PLAN

Any change from the protocol will be justified and fully documented.

If the blind review suggests changes to the principal features stated in the protocol, these have to be documented in a protocol amendment. Otherwise, it will suffice to update the statistical analysis plan with the considerations suggested from the blind review.

12 SHELLS OF TABLES, FIGURES AND LISTINGS AND REPORTING OUTPUT (GENERAL FEATURES)

Tables, figures and listings will be generated using SAS® and will be displayed on A4 size paper with landscape orientation, 2 cm for top and bottom margins, 0.8 cm for left and right margins and 8pt Courier New font.

The header section will comprise the sponsor's name, the protocol number, the delivery description, the TFL number, the TFL title, the population and the page number (Page X of Y). The footer section will include the TFL footnotes, the data extract date (if applicable), the date and time of the execution of the program, and the name of the program.

A clear, accurate and complete programming code will be developed to generate the statistical analyses, summary tables, figures and listings to be integrated in the report. Fluent use of precise titles and footnotes will be made to improve the understanding of summaries and document any assumption. Details of analysis specifications including but not limited to the SAS code will be documented on the shells.

13 APPENDICES

13.1 Shells for Table, Figure and Listings

The final list of tables, figures and listings and their shells for the reporting of this study will be available in a separate document that will be developed and will be finalized before database lock.

13.2 Potentially Clinically Significant Values of Vital Signs

Criteria for identifying vital signs values as Potentially Clinically Significant Values are the following.

Table 10 - Potentially Clinically Significant Values of Vital Signs

Parameter	Criteria		Change Relative to Baseline
Pulse Rate	≥ 120 beats/min	and an	increase ≥ 20 beats/min
	≤ 50 beats/min	and a	decrease ≥ 20 beats/min
Systolic Blood Pressure	≥ 160 mmHg	and an	increase ≥ 20 mmHg
	≤ 95 mmHg	and a	decrease ≥ 20 mmHg
Diastolic Blood Pressure	≥ 110 mmHg	and an	increase ≥ 10 mmHg
	≤ 45 mmHg	and a	decrease ≥ 10 mmHg
Body Temperature	≥ 38.3° C	and an	increase ≥ 1.1° C
	≥ 101° F	and an	increase ≥ 2° F

13.3 Potentially Clinically Significant Values of Laboratory Tests

Criteria for identifying laboratory test values as Potentially Clinically Significant Values are the following.

Table 11 - Potentially Clinically Significant Values of Laboratory Tests

Parameter	Sex	Criterion Values	
		Standard Units	SI Units
Blood Chemistry			
Aspartate Aminotransferase	Male and Female	> 3 x ULN	> 3 x ULN
Alanine Aminotransferase	Male and Female	> 3 x ULN	> 3 x ULN
Alkaline Phosphatase	Male and Female	> 400 U/L	> 400 U/L
Gamma Glutamyltransferase	Male and Female	> 3 x ULN	> 3 x ULN
Lactate Dehydrogenase	Male and Female	> 3 x ULN	> 3 x ULN
Urea	Male and Female	> 64 mg/dL	> 10.7 mmol/L
Blood Urea Nitrogen	Male and Female	> 29.9 mg/dL	> 10.7 mmol/L
Creatinine	Male and Female	> 2.0 mg/dL	> 176.8 µmol/L

Parameter	Sex	Criterion Values	
		Standard Units	SI Units
Creatinine Kinase	Male and Female	> 2.5 x ULN	> 2.5 x ULN
Uric Acid	Male	> 10.0 mg/dL	> 0.5948 mmol/L
	Female	> 8.0 mg/dL	> 0.4758 mmol/L
Total Bilirubin	Male and Female	> 2.0 mg/dL	> 34.2 µmol/L
Albumin	Male and Female	< 25 g/L	< 3.63 µmol/L
Sodium	Male and Female	< 130 mmol/L	< 130 mmol/L
		> 150 mmol/L	> 150 mmol/L
Potassium	Male and Female	< 3.0 mmol/L	< 3.0 mmol/L
		> 5.5 mmol/L	> 5.5 mmol/L
Calcium	Male and Female	< 7 mg/dL	< 1.75 mmol/L
		> 12 mg/dL	> 3.00 mmol/L
Chloride	Male and Female	< 90 mmol/L	< 90 mmol/L
		> 118 mmol/L	> 118 mmol/L
Glucose	Male and Female	< 50 mg/dL	< 2.78 mmol/L
		> 250 mg/dL	> 13.88 mmol/L
Total Cholesterol	Male and Female	> 300 mg/dL	> 7.77 mmol/L
Low-Density Lipoprotein	Male and Female	> 160 mg/dL	> 4.14 mmol/L
High-Density Lipoprotein	Male and Female	< 30 mg/dL	< 0.78 mmol/L
Triglycerides	Male and Female	> 300 mg/dL	> 3.39 mmol/L
Hematology			
Hematocrit	Male	≤ 37%	≤ 0.37 L/L
	Female	≤ 32%	≤ 0.32 L/L
Hemoglobin	Male	≤ 11.5 g/dL	≤ 115 g/L
	Female	≤ 9.5 g/dL	≤ 95 g/L
Platelets	Male and Female	≤ 75 10^3/µL	≤ 75 10^9/L
		≥ 700 10^3/µL	≥ 700 10^9/L
Leukocytes	Male and Female	≤ 2.8 10^3/µL	≤ 2.8 10^9/L
		≥ 16 10^3/µL	≥ 16 10^9/L
Eosinophils	Male and Female	> 1.0 10^3/µL	> 1.0 10^9/L
Neutrophils	Male and Female	< 1.5 10^3/µL	< 1.5 10^9/L

14 REFERENCES

1. Huiping Zhang, "Proportion difference and confidence interval based on Cochran-Mantel-Haenszel method in stratified multi-center clinical trial", PharmaSUG China 2016, Paper 25
2. Rubin D.B. 1987, "Multiple Imputation for Nonresponse in Surveys", New York: John Wiley and Sons.
3. Li, K.-H., Meng, X.-L., Raghunathan, T.E., and Rubin, D.B. 1991, "Significance levels from repeated p-values with multiply-imputed data", Statistica Sinica, 1(1), 65-92.
4. Bohdana Ratitch, et al. "Combining Analysis Results from Multiply Imputed Categorical Data", 2013, PharmaSUG Proceedings, Paper SP-03.
5. John W. Graham, Allison E. Olchowski, Tamika D. Gilreath, "How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory", Prev Sci (2007) 8:206–213 DOI 10.1007/s11121-007-0070-9
6. Yang Yuan, "Sensitivity Analysis in Multiple Imputation for Missing Data", SAS Institute Inc., Paper SAS270-2014