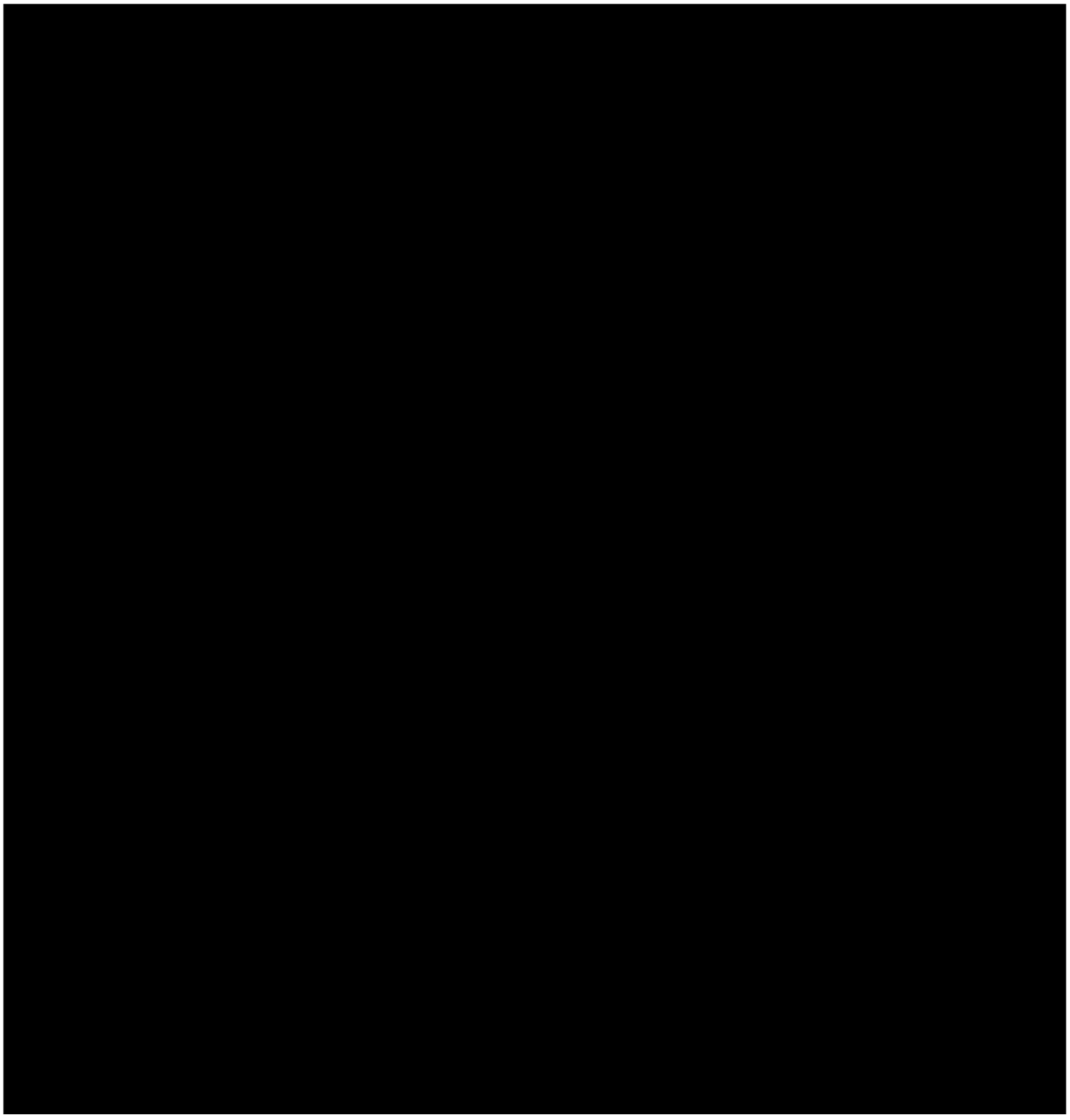
Statistical Analysis Plan

A Randomized, Evaluator-Blinded Clinical Study to Evaluate the Efficacy and Tolerability of an Investigational Light Therapy Mask on Subjects with Mild to Moderate Mottled Hyperpigmentation and Moderate to Severe Facial Wrinkles

Compound Name	PM Mask Treatment
Protocol Number	CO-170511132943-SACT
Phase	Not applicable
Version and Date	Final 23 May 2018

TABLE OF CONTENTS



1 INTRODUCTION

1.1 STUDY OBJECTIVES

As stated in the protocol:

The objective of this study is to evaluate the efficacy and tolerability of an investigational light therapy mask in comparison to a sham mask device over a 12-week usage period and after a 12-week regression period for subjects with mild to moderate mottled hyperpigmentation and moderate to severe facial wrinkles.

1.2 STUDY DESIGN

Extracted from the protocol:

This is a single-center, 2-cell, sham-controlled, randomized, evaluator-blinded clinical usage study utilizing three Expert Graders (i.e. the Principal Investigator [PI] plus two additional board-certified dermatologist Expert Graders; see Appendix XI). Up to approximately 125 subjects will be enrolled to finish with at least 96 subjects (targeting 48 subjects per cell). The target population is 35- to 70-year-old females with mild to moderate mottled hyperpigmentation on the face and moderate to severe facial wrinkles. Enrollment will be evenly distributed across all Fitzpatrick skin types (i.e. an approximately equal number of subjects will be enrolled per skin type per cell).

Subjects will be assessed at Screening, Baseline (Week 0), Week 1, Week 4, Week 12, and Week 24.

At Screening (Visit 1), each subject will be provided with an auxiliary cleanser (twice daily usage – morning and evening) and an auxiliary moisturizer (once daily usage in the morning after cleansing [plus additional evening usage (after completing the mask treatment, as applicable), as desired]) to use full-face for the duration of the study. At Baseline (Visit 2), each subject will be randomly assigned to also use either the active or sham light therapy mask (with respective activator; once daily usage for 10 minutes in the evening after cleansing) for the 12-week treatment period of the study. The active and sham masks will be identical in appearance; the Expert Graders will be blind to the IP assignment, and subjects will also be blinded to the extent possible (see section 6.6).

A 12-week regression period will begin at Week 12 (Visit 5); during the regression period, subjects will continue using the auxiliary products as described previously, but there will be no mask usage.

2 INTERIM ANALYSES

No interim analysis will be performed. Final analyses will be performed at the official database release.

3 ANALYSIS SETS

3.1 FULL ANALYSIS SET

[redacted] efficacy data will be evaluated for all intent-to-treat (ITT) subjects who used the investigational product and had baseline and at least one post-baseline data point.

3.2 'PER PROTOCOL' ANALYSIS SET

No Per Protocol Analysis will be performed.

3.3 SAFETY ANALYSIS SET

The safety analysis will be based on all randomized subjects who use investigational product.

3.4 OTHER ANALYSIS SETS

Demographic and baseline variables will be summarized for all randomized subjects.

4 ENDPOINTS AND COVARIATES

A) Efficacy: To be independently evaluated by the Principal Investigator (PI) and 2 other Expert Graders (note: only the PI will evaluate the subject at Screening [Visit 1]).

Fitzpatrick Wrinkle and Elastosis Scale

Fine lines, Periorbital wrinkles and Global wrinkling will be evaluated using the Fitzpatrick Wrinkle and Elastosis Scale. Half points are allowed:

Class	Score	Wrinkling	Degree of Elastosis
I	1-3	Fine wrinkles	Mild (fine textural changes with subtle skin lines)
II	4-6	Fine to moderate depth, moderate number of lines	Moderate (distinct popular elastosis, individual papules with yellow translucency, dyschromia)
III	7-9	Fine to deep wrinkles, numerous lines, with or without redundant skin	Severe (multipapular and confluent elastosis, thickened yellow and pallid cutis rhomboidalis)

Modified Griffiths Scale

Surface Roughness, Uneven Skin Tone, Mottled Hyperpigmentation, Sallowness or Yellowing, Lack of Radiance [redacted] will be evaluated using the Modified Griffiths Scale. Half points are allowed:

Rating (Score)	Category	Description
0	None	See below for parameter-specific descriptions.
1-3	Mild	
4-6	Moderate	
7-9	Severe	

Surface Roughness (using Modified Griffiths Scale):

Rating (Score)	Category	Description
0	None	Skin perfectly smooth.
1-3	Mild	Slight laxity of skin. Some surface roughness visible.
4-6	Moderate	Tactile and/or visible roughness demonstrable.
7-9	Severe	Topography of skin (visible and tactile) is undulating and rough.

Uneven Skin Tone (using Modified Griffiths Scale):

Rating (Score)	Category	Description
0	None	Uniform natural skin color; perfect evenness.
1-3	Mild	"Barely" to "slightly" to "slightly to mildly"…perceivable areas of redness, yellowness, or darkness.
4-6	Moderate	"Mildly to moderately" to "moderately" to "pronounced"…perceivable areas of redness, yellowness, or darkness.
7-9	Severe	"Pronounced" to "pronounced to significantly" to pronounced (severe)"… perceivable areas of redness, yellowness, or darkness.

Mottled Hyperpigmentation (using Modified Griffiths Scale):

Rating (Score)	Category	Description
0	None	Perfectly even in tone (no redness or hyperpigmentation).
1-3	Mild	Early color variation; no lentigines (freckles may be present).
4-6	Moderate	Early to moderate dyspigmentation/telangiectasia may be present; one or more lentigines.
7-9	Severe	Dyschromia (mottled and/or discrete) is likely, or it may be replaced by quite definite yellowing.

Sallowness or Yellowing (using Modified Griffiths Scale):

Rating (Score)	Category	Description
0	None	No evidence of yellowing; depending on Fitzpatrick skin type, the skin may be very pink and rosy.
1-3	Mild	No yellowing; depending on Fitzpatrick skin type, the skin may be pink and rosy to slightly pink and rosy.
4-6	Moderate	Early evidence of yellowing; one or more lentigines; depending on Fitzpatrick skin type, the skin may be slightly pink and rosy to sallow and pale.
7-9	Severe	Quite definite yellowing. Depending on Fitzpatrick skin type, the skin may be sallow and pale to very sallow and pale.

Lack of Radiance (using Modified Griffiths Scale):

Rating (Score)	Category	Description
0	None	Extremely bright, clear, radiant.
1-3	Mild	"Very" to "mildly" to "mildly to moderately"…bright, clear, radiant.
4-6	Moderate	Moderately bright, clear, radiant skin, but with some matte appearance to mildly dull/matte appearance to moderately dull/matte appearance.
7-9	Severe	"Moderately to pronounced" to "pronounced".

[redacted]

4.1 EFFICACY ENDPOINT(S)

The primary efficacy variable is the change from baseline in global wrinkling at Week 12.

The secondary efficacy variables are the change from baseline in the following endpoints at Week 1, Week 4, Week 12, and Week 24:

- Global wrinkling (excluding Week 12)
- Fine lines
- Periorbital wrinkles
- Surface Roughness
- Uneven Skin Tone
- Mottled Hyperpigmentation
- Sallowness or Yellowing
- Lack of Radiance

The tertiary efficacy variable is [redacted]

4.2 SAFETY ENDPOINTS

Adverse events (AE) will be captured from the time the informed consent is signed. Medical Dictionary for Regulatory Activities (MedDRA), current version at study start, will be used as the AE classification system.

AE data will be summarized as:

- The number and percentage of subjects experiencing adverse events, regardless of relationship to treatment.
- Pre-Treatment AEs (PTAEs)- AEs present prior to the initiation of the IP usage.
- Expected AEs
- AEs related (possible, probable, or very likely) to the product/study. The number and percentage of subjects experiencing related AEs will be summarized by treatment, system organ class, preferred term, and worst recorded severity.
- AEs non-related to the product/study
- SAEs related to the product/study
- SAEs non-related to the product/study

[redacted]

4.3 OTHER ENDPOINTS

[redacted]

4.4 COVARIATES

Model-based within-treatment and between-treatment comparisons of the efficacy variables will use the corresponding baseline value as a covariate.

4.5 DATA COMPUTATIONS AND DATA IMPUTATIONS

Efficacy scores from the 3 Expert Graders will be averaged for the baseline readings as well as for each post-baseline assessment time point before computing the change. For each subject, the change from baseline is then computed by subtracting post-baseline mean efficacy score from baseline mean efficacy score.

[redacted]

5 HANDLING OF MISSING VALUES

If global wrinkling is missing at Week 12 for more than 5% of the subjects, the missing value will be imputed by using the last observation carried forward (LOCF) method.

6 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

6.1 STATISTICAL HYPOTHESES

The primary efficacy variable is the change from baseline in global wrinkling at Week 12.

Specifically, the following hypothesis will be tested to assess within-treatment changes from baseline:

$$H_0: \mu_{Baseline} - \mu_{Post} \leq 0.5$$

against the one-sided alternative

$$H_1: \mu_{Baseline} - \mu_{Post} > 0.5$$

where $\mu_{Baseline}$ and μ_{Post} are baseline mean global wrinkling and post-baseline mean global wrinkling, respectively, for active mask or sham mask treatment.

The null hypothesis that two investigational product mean efficacy variable changes from baseline are equal will be tested against the alternative hypothesis that the investigational product mean efficacy variable changes from baseline are not equal.

In other words, the following hypothesis will be tested:

$$H_0: \mu_A - \mu_S = 0$$

against the two-sided alternative

$$H_1: \mu_A - \mu_S \neq 0$$

where μ_A and μ_S are the mean changes from baseline for active mask treatment and the sham mask treatment, respectively.

6.2 STATISTICAL DECISION RULES

Superiority to baseline will be concluded if at least two of the three Expert Graders rate the active mask superior to baseline with respect to global wrinkling at Week 12 and the mean improvement score in global wrinkling for the active mask exceeds 0.5 points at Week 12, that is the lower bound of the 95% confidence interval for the mean difference is above 0.5. (Note that under 8.2.1. Analysis of Primary Variable of the protocol it incorrectly reads that superiority to baseline will be concluded if the upper bound of the 95% confidence interval for the mean difference is below -0.5.)

6.3 STATISTICAL METHODS

Treatment means and between-treatment differences will be assessed by means of an ANCOVA. The model is described in detail in Section 6.4.2.

6.4 STATISTICAL ANALYSES

The Department of Quantitative Sciences [redacted] will be responsible for the statistical data analyses.

For continuous distributions, descriptive summaries will include number of subjects, mean, standard deviation, median, minimum and maximum values. Distributions of categorical variables will be summarized by presenting the number and percent of subjects in each response category.

6.4.1 Demographic and Baseline Characteristics

Demographic and baseline characteristics will be compared across treatment groups using Analysis of Variance (ANOVA) model (with term for treatment) for continuous variables or Chi-Square test for categorical variables. If the expected number of subjects within a specific category is sufficiently small, Fisher's exact test will be used in place of the Chi-Square test.

6.4.2 Primary Analysis

The primary efficacy variable is the change from baseline in global wrinkling at Week 12. Scores from the 3 Expert Graders will be averaged for the baseline readings as well as the Week 12 readings before computing the change. For each subject, the change from baseline is then computed by subtracting post-baseline mean global wrinkling from baseline mean global wrinkling. The mean change from baseline for each cell will be presented together with a two-sided 95% confidence interval.

The active mask will be compared with the sham mask. Treatment means and between-treatment differences will be assessed by means of an ANCOVA model with treatment and skin type group as factors and the corresponding averaged baseline score as a covariate. The two-sided 95% confidence interval for the treatment difference will be presented. Analysis of the primary efficacy variable will be based on the average scores as well separately for the scores from each Expert Grader.

For the secondary efficacy variables, scores from the 3 Expert Graders will be averaged for the baseline as well as each post-baseline visit before computing the change. The analyses of the secondary efficacy variables will be based on average scores and will be performed as for the primary efficacy variable.

The tertiary efficacy variable is the change from baseline at Week 1, Week 4, Week 12, and Week 24 for [redacted]

6.4.3 Secondary Analyses

No secondary analyses are planned.

6.4.4 Safety Analyses

Safety analyses will be based on the Safety Analysis Set, defined in section 3.3.

The number and percentage of subjects experiencing Pre-treatment AE (PTAEs) will be summarized by treatment, MedDRA system organ class and preferred term. This information will also be presented for the most commonly reported adverse events (≥5% in one or more treatment groups). PTAE is defined as any AE present prior to the initiation of the IP usage.

The number of subjects with PTAEs will be summarized by severity (mild, moderate and severe) for all adverse events. Subjects will be counted only once for each system organ class and preferred term by selecting the most severe event.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized by treatment, MedDRA system organ class and preferred term.

The number and percentage of subjects experiencing treatment-emergent AEs will also be summarized by treatment, MedDRA system organ class, preferred term and severity. Subjects will be counted only once for each system organ class and preferred term by selecting the most severe event.

The number and percentage of subjects experiencing treatment-related AEs will be summarized by treatment, MedDRA system organ class, and preferred term. Treatment-related AEs are events evaluated by the investigator as being possible, probable or very likely related to study products. AEs with an unknown relationship to treatment will be considered to be treatment-related.

The number and percentage of subjects experiencing treatment-related AEs will also be summarized by treatment, MedDRA system organ class, preferred term and severity.

The number and percentage of subjects who discontinued the study due to adverse events (based on study disposition) and the number and percentage of subjects experiencing serious adverse events will be summarized by treatment. These displays will include all AEs.

Listings of all adverse events will be provided. Treatment-emergent AEs and PTAEs, if present, will be listed separately.

[redacted]

7 SIGNATURES PAGE

APPENDIX 1 SUMMARY OF EFFICACY ANALYSES

Endpoint	Analysis Set	Statistical Method	Model/ Covariates	Missing Data
Mean efficacy change from Baseline	FAS	ANCOVA	Skin type group / Baseline mean	Excluded
Treatment differences in mean efficacy change from Baseline	FAS	ANCOVA	Treatment, skin type group / Baseline mean	Excluded

APPENDIX 2 SUMMARY TABLES

Disposition of subjects:

- Study completion status and subject evaluation

Demography and Baseline characteristics:

- Demographic and baseline characteristics, Fitzpatrick Skin Type
- Skin Type – Oiliness Level at Screening rated by subjects.

Efficacy variables:

- Fitzpatrick Wrinkle and Elastosis Scale. Analysis of the primary efficacy variable will be based on the average scores as well separately for the scores from each Expert Grader.
- Modified Griffiths Scale
- [redacted]
- [redacted]

Adverse events [redacted]:

- Summary of Adverse Events
- Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
- Summary of Commonly Reported Adverse Events
- Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Severity
- Summary of Treatment Related Adverse Events by System Organ Class and Preferred Term
- Summary of Treatment Related Adverse Events by System Organ Class, Preferred Term and Severity
- Summary of Pre-Treatment Adverse Events by System Organ Class and Preferred Term
- Summary of Pre-Treatment Adverse Events by System Organ Class, Preferred Term and Severity
- Summary of Adverse Events that Resulted in Study Discontinuation
- Summary of Serious Adverse Events
- [redacted]

APPENDIX 3 SUBJECT DATA LISTINGS

- Randomization code listing
- Study Completion and Subject Evaluation
- Subjects with Protocol Deviations
- Demographic and baseline characteristics
- Fitzpatrick Wrinkle and Elastosis Scale including individual mean values
- Modified Griffiths Scale including individual mean values
- [redacted]
- Subjects with Adverse Events
- Subjects with Serious Adverse Events
- Subjects with Adverse Events Leading to Investigational Product Withdrawn
- [redacted]