# Statistical Analysis Plan

## Clinical Trial Number: 43USSA1812

## Revision History

| Version | Date | Revision Author | Comments |
|---|---|---|---|
| 2.0 | 07 JUL 2020 | PPD | Incorporated changes from CIP V4 of protocol:<br>1. Month 6 visits are now month 7. Any patient with a month 6 visit will use that as a surrogate for their month 7 visit for secondary endpoint purposes<br>2. Many endpoints at month 9 visits are now at month 12 – including primary endpoint<br>3. PP population is now based on month 12 visit and primary endpoint<br>4. Changed BOCF to Multiple Imputation<br>5. Updated schedule of events<br>6. Device deficiencies will be summarized |
| 3.0 | 15 JAN 2021 | PPD | Added information on multiple imputation for the primary analysis (section 2.4.3 Multiple Imputation) |
| 4.0 | 23 MAR 2021 | PPD | 1. Amended the number of FST V-VI subjects to be at least 14 in alignment with the protocol amendment v5<br>2. Remove an erroneous mention to an unused endpoint, the WAS<br>3. Define how nodules, papules, and granuloma are defined from AEs<br>4. Clarify that AE's by injection technique include combinations of injection techniques |
| 5.0 | 30 JUL 2021 | PPD | 1. Moved "other effectiveness endpoints" to new section 2.4.6. Added information here on summarizing subjects with incorrect gain for dermal thickness.<br>2. Elaborated on how Snellen acuity will be summarized via Snellen decimal categories. |

A randomized, evaluator-blinded, no-treatment controlled, multicenter study to evaluate the effectiveness and safety of Sculptra Aesthetic for correction of cheek wrinkles

## SIGNATURE PAGE

PPD

## Table of Contents

# 1 Study Information

## 1.1 Background

This statistical analysis plan (SAP) describes the analysis variables and statistical procedures that will be used to analyze and report the results from Protocol 43USSA1812 (v3.0), dated 20 September 2019.

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

### 1.1.1 Study Design

This is a prospective, randomized, evaluator-blinded, no-treatment controlled study. Approximately 150 subjects will be randomized (2:1) to either treatment with Sculptra Aesthetic (Treatment Group) or no treatment (Control Group), 100 subjects to treatment with Sculptra Aesthetic and 50 subjects to no treatment. At least 15 subjects will be Fitzpatrick skin type (FST) IV, and at least 14 will be FST V-VI.

Effectiveness and safety data will be collected for up to 12 months following the initial treatment.

All subjects will be offered to participate in an open extension study beginning after successful completion of this study. Subjects in the no-treatment control group will be offered treatment with Sculptra Aesthetic in the extension study. Subjects in the treatment group will continue to be followed for additional data collection as described in the extension study protocol.

### 1.1.2 Number of Subjects

The study is planned to enroll approximately 150 subjects at approximately 13 centers in the United States.

## 1.2 Study Objectives

For the study to be deemed successful, the following criteria need to be met:

- The primary effectiveness endpoint, CCI [redacted] CCI [redacted], should be statistically significantly greater in the Treatment Group when compared to the Control Group.
- Acceptable safety profile in the treatment group.

### 1.2.1 Effectiveness Objectives

To evaluate the effectiveness of Sculptra Aesthetic versus a no-treatment control in the correction of cheek wrinkles.

### 1.2.2 Safety Objectives

To evaluate the safety of Sculptra Aesthetic in the correction of cheek wrinkles.

## 1.3 Effectiveness Assessments

For all assessments, baseline will be defined as the observation that is closest to but prior to study treatment on Day 1 for treated patients. For non-treated patients, baseline will be defined as the observation that is closest to but prior to randomization. Likewise, change from baseline will be calculated as the value at a given time point minus the baseline value.

### 1.3.1 [redacted]

The severity of cheek wrinkles will be assessed live by the Blinded Evaluator and the Treating Investigator using the validated 5-graded [redacted] during the study. The [redacted] scale is a validated photograph-based outcome instrument. A photo guide of the scale is provided in the protocol. The subject is to have a relaxed face during the assessment.

The [redacted] will be assessed live by the Blinded Evaluator at screening/baseline and at months 7, 9, and 12 after baseline. The [redacted] will be assessed live by the Treating Investigator at screening/baseline and at each treatment visit. However, the Treating Investigator [redacted] will not be included in any statistical analyses since these assessments are only made to decide whether further treatment is needed. See Table 1 for categorical scales used.

### 1.3.2 [redacted]

The severity of cheek wrinkles will be assessed live by the Blinded Evaluator using the validated 5-graded [redacted] during the study. The [redacted] scale is a validated photograph-based outcome instrument. A photo guide of the scale is provided in the protocol. The subject is to have a closed maximum smile during the assessment.

The [redacted] will be assessed live by the Blinded Evaluator at screening/baseline and at months 7, 9, and 12. The [redacted] will be assessed live by the Treating Investigator at screening/baseline and at each treatment visit. However, the Treating Investigator [redacted] will not be included in any statistical analyses since these assessments are only made to decide whether further treatment is needed. See Table 1 for categorical scales used.

CCI [redacted]

### 1.3.3 Global Aesthetic Improvement Scale (GAIS)

The aesthetic improvement of cheek wrinkles will be assessed by the Treating Investigator and the subject, independently of each other. Baseline photographs may be used as a comparison for the GAIS evaluation.

The Treating Investigator and the subject will, independently of each other, respond to the question: "How would you describe the aesthetic improvement today compared to the photograph taken before treatment?"

GAIS will be assessed by the Treating investigator and subject at months 7, 9, and 12. The following categorical scales will be used (Table 2).

CCI

[redacted]

### 1.3.4 Subject Satisfaction Questionnaire

Subject satisfaction with treatment result will be assessed by using a subject satisfaction questionnaire (Section 7.4 of the protocol) at all visits following the baseline visit for the Treatment Group.

### 1.3.5 FACE-Q "Satisfaction with Cheeks"

The FACE-Q is a patient-reported outcome instrument to evaluate the experience and outcomes of aesthetic facial procedures from the subject's perspective at all visits following the Baseline visit for the Treatment Group and at months 7, 9, and 12 for the Control Group

### 1.3.6 Subject Diary

A subject diary will be dispensed to all subjects in the Treatment Group for daily completion for 28 days beginning on injection day for each treatment with direct questioning for pain, tenderness, redness, bruising, swelling, itching, lumps/bumps and "other". Subjects will be specifically asked to record any of the following symptoms in the other section: changes in vision (i.e. loss of vision, blurriness, double vision, pain in or around your eye, blindness, blind spots, problems moving your eyes), skin changing color around the eyelids, crusty or scabby skin around the eyelids, pain, headache, fever, dizziness, confusion, weakness or numbness in the arms or legs, changes to consciousness or alertness, difficulty speaking/speech impairment and face droop.

Return to Social Engagement

Subjects will be asked to complete the following questions in the subject diary: Did you feel comfortable returning to social engagement today? If yes, what was the earliest time you felt comfortable to return to social engagement?

Returning to social engagement is defined as making public/social appearances, including but not limited to returning to work at a business office or other public work place; having dinner in a public restaurant; attending a social event/gathering such as dinner party, etc. Subjects will record the time they felt comfortable (with or without covering make-up) resuming social interactions, not necessarily the time for their first social interaction.

## 1.4 Effectiveness Endpoints

### 1.4.1 Primary Effectiveness Endpoint

The primary measure of effectiveness is the responder rate based on the [redacted], as assessed live by the Blinded Evaluator at month 12.

A responder is defined as a [redacted] with at least 1 grade improvement from baseline on both cheeks concurrently.

### 1.4.2 Secondary Effectiveness Endpoints

Secondary effectiveness endpoints include:

1. Responder rate based on the [redacted], as assessed live by the Blinded Evaluator at month 7 and 9. A responder is defined as a subject with at least 1 grade improvement in both cheeks concurrently from baseline.

2. Responder rate based on the [redacted], as assessed live by the Blinded Evaluator at month 7, 9 and 12. A responder is defined as a subject with at least 1 grade improvement in both cheeks concurrently from baseline.

3. Improvement rate based on the Independent Photographic Reviewer's assessment using random pairings of baseline and month 12 photographs.

   Treatment group: An improved subject is defined as a subject for whom the Independent Photographic Reviewer correctly identified the month 12 photograph in the pair of pre- and post-treatment photographs at rest.

   Control group: An improved subject is defined as a subject for whom the Independent Photographic Reviewer identified any of the photographs in the pair as post-treatment at rest.

4. Percentage of subjects having at least "Improved" according to the GAIS on both sides of the face combined, as assessed live by the subject and the Treating Investigator separately, at all visits following the Baseline visit for the Treatment Group and at month 7, 9 and 12 for the Control Group.

5. Percentage of subjects responding in each response category for each question in the subject satisfaction questionnaire at all visits following the Baseline visit for the Treatment Group.

6. Change from baseline in subject satisfaction using the Satisfaction with Cheeks FACE-Q questionnaire with Outcome Rasch transformed total scores as well as percentage of subjects in each response category for each of the individual questions at all visits following the

Baseline visit for the Treatment Group and at month 7, 9 and 12 for the Control Group.

7. Time to feeling comfortable returning to social engagements after treatment using subject diaries for 28 days after each treatment.

CCI

[redacted]

### 1.5 Safety Assessments

The methods for collecting safety data are described in Section 8 of the Clinical Study Protocol.

### 1.6 Safety Endpoints

Safety endpoints include:

1. Incidence, intensity, time to onset and duration of adverse events collected throughout the study period.

2. Incidence, intensity and number of days of pre-defined expected post-treatment events collected using subject diaries for 28 days from each treatment.

3. Safety assessment by a qualified staff member at all visits according to predefined methods at baseline and at all follow-up visits for the Treatment Group and the Control Group:
    - Cheek firmness, symmetry and function
    - Device palpability (baseline assessment excluded)
    - Mass formation
    - Cheek sensation
    - Visual Function Assessments

## 2 Statistical Methods

### 2.1 General Methods

Any change made to the finalized SAP before database lock will result in a SAP amendment. Otherwise, the change will be documented in the Clinical Study Report (CSR).

Some of the analyses detailed here may be more explicit or in some aspects different from those stated in the protocol. In case of differences, this SAP supersedes the statistical sections in the protocol.

### 2.1.1 Programming Conventions

[redacted] Statistical Solutions will have responsibility for performing analyses. All computations for statistical analyses will be performed using SAS® software, Version 9.4 or later. All SAS programs used in the production of statistical summary outputs will be validated with independent programming prior to finalization. In addition, all program outputs will be independently reviewed. The validation process will be used to confirm that all data manipulations and calculations were accurately done. Once validation is complete, a senior statistical reviewer will perform a final review of the documents to ensure the accuracy and consistency with this plan and consistency within tables. Upon completion of validation and quality review procedures, all documentation will be collected and filed by the project statistician or designee.

The electronic case report form (eCRF) data for all subjects will be provided in Standard Data Tabulation Model (SDTM) datasets. Analysis Data Model (ADaM) datasets will be developed from the SDTM datasets for use in table and figure production.

### 2.1.2 Reporting Conventions

The formats for the tables, listings, and figures described in this SAP will be provided in a companion document. Changes to the formats of these reports that are decided after the finalization of the SAP will not require an amendment. In addition, any additional supportive or CCI [redacted] analyses requested after SAP approval will not require amendment of the SAP. These additional analyses will be described in the CSR.

All study data from the eCRFs as well as derived variables will be provided in subject data listings. An indication of specific listings for each data type will not be indicated in the text of subsequent SAP sections. Data listings supplied as part of the CSR will be sorted by study center number concatenated with subject number, assessment dates, and/or time point.

The following conventions will be applied to all data presentations and analyses:

- Quantitative variables will generally be summarized by the number of subjects, mean, standard deviation (SD), median, minimum (min), and maximum (max). Unless otherwise specified, the minimum and maximum values will be displayed to the same number of decimal places as the raw data, the mean and median will be presented to one extra decimal place compared to the raw data, and the standard deviation will be displayed to two extra decimal places compared to the raw data
- Categorical variables will be summarized by the number and percentage of subjects (and number of events where appropriate) within each category. Unless otherwise specified, the percentage will be presented in parentheses to one decimal place. Frequency and percentage values of 0 will be presented as ‘0’ rather than ‘0 (0)’.
- All summary tables will include the analysis population sample size (i.e. number of subjects).
- Date variables will be formatted as DDMMMYYYY for presentation.

### 2.1.3 Data Transformations

The Rasch transformation scoring of the FACE-Q will be reported (Section 4).

## 2.2 Analysis Populations

The statistical analyses will be performed based on the following subject populations.

### 2.2.1 Intent-to-Treat (ITT) Effectiveness Population

Includes all subjects who were randomized based on the as randomized principle (i.e. according to the treatment they were randomized to).

### 2.2.2 Per-Protocol (PP) Effectiveness Population

Includes all subjects in ITT with a completed 12 month after baseline assessment of the Blinded Evaluator [redacted], and without any deviations considered to have substantial impact on the primary effectiveness outcome.

### 2.2.3 Safety Population

Includes all subjects who were treated with Sculptra Aesthetic or randomized to no treatment control group. Subjects are analyzed based on the as treated principle (i.e. according to the treatment actually received).

## 2.3 Study Subjects

Demographic endpoints and subject characteristics will be summarized using descriptive statistics by treatment group and overall based on the ITT population using the Observed Cases (OC). There are no planned inferential statistical analyses of demographic endpoints or subject characteristics.

### 2.3.1 Subject Disposition

The number of subjects screened will be shown in total and by study center.

The number of subjects in each study population (i.e. ITT, PP, and Safety) will be summarized by study center and in total.

The disposition of subjects will be summarized, including numbers of subjects who were:

- Randomized,
- Completed,
- Withdrawn (including primary reason for withdrawal).

These numbers will be summarized by study center. The number of subjects expected, completed, missed, and withdrawn will be summarized by scheduled visit.

- Expected = all subjects minus withdrawn subjects.
- Completed = subjects that showed up at that visit.
- Missed = expected subjects minus completed subjects.
- Withdrawn = all subjects who have withdrawn up to that visit.

All withdrawn subjects will be listed individually, by subject number, date and reason for withdrawal, and last visit performed.

### 2.3.2 Protocol Deviations

Subjects with any protocol deviations will be summarized by treatment group, overall, by site, and by type.

Depending on the seriousness of the deviation, a subject might be excluded from the PP population, which shall be documented prior to database lock. Reasons for exclusion from the PP population will be summarized.

### 2.3.3 Demographic Characteristics

Age, weight, and body mass index (BMI) will be summarized as continuous variables. Weight and BMI will additionally be summarized at each visit along with change from baseline.

Gender, age category (≥55 vs <55), race, ethnicity, Fitzpatrick skin type and baseline [redacted] (Blinded Evaluator and Treating Investigators) will be summarized as categorical variables.

### 2.3.4 Medical History, Medications, and Procedures

Prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHODD). Medical history, and prior and concomitant procedures/non-pharmacological treatments will be coded according to MedDRA.

Prior medications/procedures are the medications/procedures with stop dates prior to study treatment. Medications/procedures after the study treatment will be considered concomitant.

Subjects reporting medical history, cosmetic treatments/procedures, and prior and concomitant procedures/non-pharmacological treatments will be summarized by system organ class (SOC) and Preferred Term (PT).

Subjects reporting prior and concomitant medications will be summarized separately, by WHODD Anatomical Therapeutic Chemical (ATC) Class Level 3 (if Level 3 is not available, the highest class available will be used) and WHODD generic name.

## 2.4 Effectiveness Analysis

### 2.4.1 Datasets Analyzed

ITT is the primary population for all effectiveness analyses. The primary effectiveness analysis will be repeated using the PP analysis set if there is at least a 10% difference in the number of subjects between the PP and ITT sets. Safety analysis is performed based on the safety population set.

### 2.4.2 Handling of Missing Data

Number of missing values will be summarized and reported as appropriate.

For ITT analysis of the Blinded Evaluator [redacted] responder rate at month 12 (primary endpoint), missing values will be imputed via multiple imputation. Only patients with a baseline

score will be included in this analysis. Impact of missing data on the primary endpoint will be evaluated by performing sensitivity analysis based on the observed cases in the ITT set.

A small number of patients may have a month 6 visit rather than month 7. These patients' month 6 visit will be used as a conservative surrogate for their month 7 visit results, and pooled with the month 7 visit data.

All other effectiveness endpoints will be evaluated based on the observed cases in ITT.

Descriptive statistics of all safety data will be performed on observed cases.

### 2.4.3 Multiple Imputation

The multiple imputation (primary analysis only) will assume Missing Completely at Random (MCAR). Regardless of the actual pattern of missing data, the Full Conditionals (FCS) option of the MI procedure from the SAS® system will be used to generate ten sets of data with missing values imputed from observed data. The observed data used to impute missing Month 12 [redacted] values for each side of the face are as follows: treatment, side of face, and all previous [redacted] scores (including the following visits: screening, baseline, and months 7, 9, and 12). The FCS option with no arguments will be employed to model the missing left and right [redacted] scores separately, which is equivalent to a regression model which uses all variables as covariates. To achieve such a model, CCI [redacted]

The imputed datasets will be analyzed using the methodology described for the primary analysis. In terms of the continuous mapping of the [redacted] the multiple imputation, a responder will be defined as a subject with at least an improvement of 1 in both cheeks concurrently from baseline. Responder rates, difference in responder rates, and the associated confidence intervals will be combined by the MIANALYZE procedure of the SAS® system. The mean value of the Fisher's Exact Test p-values across imputations will be reported. The seed number to be used will be the last four digits of the protocol number (1812).

### 2.4.4 Primary Effectiveness Analysis

Responder rate based on the [redacted] as assessed live by the Blinded Evaluator at 12 months after baseline will be the primary effectiveness endpoint. CCI [redacted]. For a significant result, the two-sided p-value of the comparison of responder rates between the treated and untreated subjects at month 12 using the Fisher's exact test needs to be smaller than 0.05.

The null hypothesis will be that there is no relationship between responder rate and treatment. The alternative hypothesis will be that there is a relationship between responder rate and treatment.

The estimates of the responder rate in each treatment group will be presented along with the difference in responder rates. Corresponding confidence intervals for responder rates (via Clopper-Pearson) as well as difference in responder rates (via normal approximation) and p-value (via Fisher's exact test) for difference in responder rates will also be presented.

In order to investigate the poolability of the results of the primary analysis across different subgroups, subgroups defined by study center, age (≥55 vs <55), and FST (I-III vs IV-VI) will be used. Logistic regression models will be fitted which will include the subgroup factor as well as the corresponding interaction term (subgroup x treatment) in order to obtain least squares estimates of

the responder rates and confidence intervals at each level of the subgroup factor. These will be displayed in graphs.

2.4.5 Secondary Analysis

- Responder rate based on [redacted] by Blinded Evaluator at month 7 and 9: Fisher's exact test will be used to compare treated and untreated subjects. Responder rates and difference in responder rates will be displayed along with corresponding 95% confidence intervals and the p-value. The analysis will be the same as for the primary analysis.

- Responder rate based on [redacted] by Blinded Evaluator at month 7, 9 and 12: Fisher's exact test will be used to compare treated and untreated subjects. Responder rates and difference in responder rates will be displayed along with corresponding 95% confidence intervals and the p-value. The analysis will be the same as for the primary analysis.

- Improvement rate based on the Independent Photographic Reviewer's assessment using random pairings of baseline and month 12 photographs: improvement rates will be displayed along with 95% confidence intervals for each treatment group via Clopper-Pearson method.

- Percentage of subjects having at least "Improved" according to the GAIS on both sides of the face combined, as assessed live by the subject and the Treating Investigator separately, at all visits following the Baseline visit for the Treatment Group and at month 7, 9 and 12 for the Control Group: percentage will be displayed along with 95% confidence intervals for each treatment group via Clopper-Pearson method.

- Percentage of subjects responding in each response category for each question in the subject satisfaction questionnaire at all visits following the Baseline visit for the Treatment Group: percentage will be displayed for each category.

- Change from baseline in subject satisfaction using the Satisfaction with Cheeks FACE-Q questionnaire with Outcome Rasch transformed total scores at all visits following the Baseline visit for the Treatment Group and at month 7, 9 and 12 for the Control Group: for each treatment group, summary statistics and 95% confidence interval calculated for each treatment mean via t-distribution will be displayed.

- Proportion of subjects in each response category for each of the individual questions using the Satisfaction with Cheeks FACE-Q questionnaire with Outcome Rasch transformed total scores as well as at all visits following the Baseline visit for the Treatment Group and at month 7, 9 and 12 for the Control Group: percentages will be displayed for each category.

- Time to feeling comfortable returning to social engagement after treatment (defined as the subject diary date with the first "yes" for "Did you feel comfortable returning to social engagement today?" minus the date of the corresponding treatment) will be analyzed using Kaplan-Meier methods. The median time to feeling comfortable will be estimated and Kaplan-Meier plots will be created.

Time to feeling comfortable returning to social engagements will be calculated for each diary (treatment occasion) separately: after 1st treatment, 2nd treatment, etc. Subjects who never report feeling comfortable during each time period will have their response time censored at the time of their last diary entry in each period.

2.4.6 Other effectiveness endpoints:

CCI [redacted]

Some subjects' dermal thickness ultrasounds were taken using an incorrect gain (an ultrasound setting which brightens the image by a uniform margin). To ensure this does not compromise efficacy results, two summaries of dermal thickness will be provided – the first will be a summary using all subjects, regardless of their gain setting. The second summary will be a sensitivity analysis, which will compare two groups at each visit: subjects with correct gain vs. subjects with incorrect gain (correct gain defined as subjects who had the same gain at the baseline visit and the visit being analyzed; conversely, incorrect gain defined as subjects who had a different gain at the baseline visit compared to the visit being analyzed). Gain information at each subject-visit will be included in a listing.

## 2.5 Safety Analysis

Safety endpoints will be summarized using descriptive statistics by treatment group based on the safety population using the Observed Cases (OC). There are no planned inferential statistical analyses of safety endpoints.

### 2.5.1 Treatment Administration, Procedural Anesthetics, and Injection Concomitant Procedures

Treatment administration endpoints that will be summarized by treatment session include (but are not limited to):

- Injection volume
- Injection method
- Depth of injection.

The number of subjects with any procedural anesthetics will be summarized by type and location.

The number of subjects with any injection concomitant procedures will be summarized.

### 2.5.2 Pre-Defined Expected Post-Treatment Symptoms

Number and percentage of subjects reporting each pre-defined, expected, post-treatment symptoms, as collected in the 28-day diary, will be presented by treatment group in total and by maximum intensity. Number of days with the event will be presented by treatment group and category: 1, 2-7, 8-14, and 15-28 days

### 2.5.3 Adverse Events (AEs)

All AEs will be coded according to MedDRA and summarized by system organ class (SOC), preferred term (PT) and treatment.

AEs related to study product or injection procedure and unrelated AEs will be presented by maximum intensity, SOC and PT. For related AEs, the number of days to onset and the duration of event will be summarized by SOC and PT using mean, SD, min, max and median. Action taken for related AEs will also be summarized by SOC and PT. Serious AEs will be listed. AE's with late onset (> 21 days after most recent treatment) related to study product or injection procedure will be listed. Nodules, papules and granuloma, as identified by any AE with SOC containing "nodule", "papule", or "granuloma", will also be presented in a separate table including n and percent of subjects with the event as well as mean, SD, min, max and median of time to onset and duration.

In addition, a summary of all AEs will be provided, which will include (but is not limited to):

- number of subjects with at least one AE and number of events (in total as well as serious AEs)
- number of subjects with at least one related AE and number of events (in total as well as serious AEs)
- number of subjects with at least one un-related AE and number of events (in total as well as serious AEs)
- number of subjects who did not have an AE

To evaluate consistency of AEs across different subgroups, related AEs by SOC, PT, and maximum intensity will also be summarized by subgroups, such as study center, age group, FST, baseline blinded evaluator [redacted](both sides of face "moderate" vs. any side "severe"), and injection technique. The analysis of AEs stratified on different injection techniques will include all combinations of injection techniques.

### 2.5.4 Other Safety Analyses

Functionality, vision function assessments, sensation, firmness, symmetry, mass formation and palpation assessments, and device deficiencies will be analyzed descriptively as appropriate.

Snellen visual acuity will be summarized using Snellen decimal categories. Snellen ratios will be converted to decimals (e.g., 20/25=0.8, 20/30=0.67, 20/70=0.29) and categorized with Snellen decimal categories of >=0.8, 0.32 to <0.8, 0.125 to <0.32, 0.05 to <0.125, and <0.05. Each subject's eye will be categorized by the number of Snellen decimal categories worsened (0 if improved or no change), and the frequency of these changes will be summarized across subject-eyes.

For example, if (a,b) represent individual change in each subject's eyes, then possible values for a and b are 0+, -1, and -2, where 0+ = improvement or no change in Snellen decimal category, -1= worsening by 1 Snellen decimal category, -2 = worsening by 2 or more Snellen decimal categories. NA = Not available. Then, (-2, -2) represents a worsening of at least 2 Snellen decimal categories in both eyes, and (0+,-1) represents an improvement or no change in one eye and a worsening by 1 Snellen decimal category in other eye, etc.

## 2.6 Interim Analysis

There are no planned interim analyses for this study.

## 2.7 Determination of Sample Size

Based on what is seen in clinical studies of injectable fillers in the facial areas, it is reasonable to assume a responder rate of at least 70% in the Sculptra Aesthetic treatment group at month 12. For the no treatment control group, responder rates up to almost 30% have been observed in data on file. Based on this, it is assumed that the response rate will be maximum 35% in the no-treatment control group at month 12.

For the responder rate in [redacted], testing the difference using a two-sided test at the 5% significance level will have approximately 90% power to demonstrate difference between a responder rate of 70% in the Sculptra Aesthetic group, and a responder rate of 35% in the no treatment control group when the sample sizes are 84 and 42, respectively. Accounting for 15% drop-outs, approximately 150 subjects need to be randomized in a 2:1 ratio.

## 2.8 Changes in the Analysis Planned in the Protocol

There have been no substantial changes from the statistical methods described in the protocol.

# 3 Reference List

There are no other references beyond those that are included in the protocol.

# 4 Rasch Transformation of FACE-Q

CCI

[redacted]

# 5 Schedule of Events (Treatment and Control Groups)

**Table 1 Schedule of events for Treatment Group**

| | Visit 1 | Visit 2 | Visit 2a | Visit 3 | Visit 3a | Visit 4 | Visit 4a | Visit 5 | Visit 5a | Visit 6 & 7 | Visit 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Screening[1] | Baseline/ Treatment 1[1] | Telephone Contact | Treatment 2[4] (Optional) | Telephone Contact | Treatment 3[4] (Optional) | Telephone Contact | Treatment 4[4] (Optional) | Telephone Contact | Follow-up | Final visit/Early termination |
| | *(≤ 30 days of Baseline)* | *Day 1 Baseline* | *72 hours (±24 hrs) after Treatment 1* | *1 month (+5 weeks) after Treatment 1* | *72 hours(±24 hrs) after Treatment 2* | *1 month (+5 weeks) after Treatment 2* | *72 hours (±24 hrs) after Treatment 3* | *1 month (+5 weeks) after Treatment 3* | *72 hours (±24 hrs) after Treatment 4* | *Month 7 ( +2 weeks) Month 9 (±2 weeks ) after Baseline* | *Month 12 ( ±2 weeks) after Baseline* |
| Informed consent | X | | | | | | | | | | |
| Inclusion/Exclusion criteria | X | X[3] | | X[3] | | X[3] | | X[3] | | | |
| Demographics incl height and weight | X | X[9] | | | | | | | | X[9] | X[9] |
| Medical history | X | X[3] | | | | | | | | | |
| Previous therapies | X | | | | | | | | | | |
| Concomitant therapies | X | X | X | X | X | X | X | X | X | X | X |
| Urine pregnancy test[2] | X | X[3] | | X[3] | | X[3] | | X[3] | | | |
| Randomization | | X[3] | | | | | | | | | |
| Photography (2D and 3D) | | X[3] | | X | | X | | X | | X | X |
| Sculptra Aesthetic administration | | X | | X[4] | | X[4] | | X[4] | | | |
| Adverse events assessment | | X | X | X | X | X | X | X | X | X | X |
| Safety assessments[8] | X | X | | X | | X | | X | | X | X |
| Visual Function assessments[10] | X | X[3,7] | | X[3,6,7] | | X[3,6,7] | | X[3,6,7] | | X | X |
| Device deficiencies | | X | | X[6] | | X[6] | | X[6] | | | |
| Dispense subject diary | | X | | X[6] | | X[6] | | X[6] | | | |
| Collect subject diary | | | | X[3] | | X[3,5] | | X[3,5] | | X[5] | |
| Ultrasound (at selected sites) | | X[3] | | | | | | | | X | X |
| **Treating Investigator Assessments** | | | | | | | | | | | |
| GAIS | | | | X[3] | | X[3] | | X[3] | | X | X |
| Radiance, skin tightness and jawline contour | | | | | | | | | | X | X |
| [redacted] | X[3] | X[3] | | X[3] | | X[3] | | X[3] | | | |
| **Blinded Evaluator Assessments** | | | | | | | | | | | |
| [redacted] | X[3] | X[3] | | | | | | | | X | X |
| **Subject Assessments** | | | | | | | | | | | |
| GAIS | | | | X[3] | | X[3] | | X[3] | | X | X |
| Subject satisfaction questionnaire | | | | X[3] | | X[3] | | X[3] | | X | X |
| FACE-Q | | X[3] | | X[3] | | X[3] | | X[3] | | X | X |

1. *Screening and baseline visits may occur on the same day, if visits occur the same day assessments will not be duplicated*
2. *Females of childbearing potential*
3. *Pre-treatment*
4. *If treatment is not performed this is a follow-up visit*
5. *If applicable (i.e. if treatment was performed at the previous visit)*
6. *If applicable (i.e. if treatment is performed at this visit)*
7. *30 minutes post-treatment*
8. *Safety assessments include; Cheek firmness, symmetry and function, Device palpability (baseline assessment excluded), Mass formation, Cheek sensation*
9. *Weight only*
10. *Visual Function assessments include: Snellen visual acuity test, Extraocular muscle function test, Confrontation visual field*

*Abbreviations: GAIS (Global Aesthetic Improvement Scale), [redacted], ET (Early Termination)*

*Note: One month is defined as four weeks in the study.*

**Table 2 Schedule of events for Control Group/Untreated subjects**

| | Visit 1 | Visit 2 | Visit 3 & 4 | Visit 5 |
|---|---|---|---|---|
| | Screening[1] | Baseline | Follow-up | Final visit/Early termination |
| | *(≤ 30 days of Baseline)* | *Day 1 Baseline* | *Month 7( +2 weeks) Month 9 (±2 weeks) after Baseline* | *Month 12 ( ±2 weeks) after Baseline* |
| Informed consent | X | | | |
| Inclusion/Exclusion criteria | X | X | | |
| Demographics, incl height and weight | X | X[4] | X[4] | X[4] |
| Medical history | X | X | | |
| Previous therapies | X | | | |
| Concomitant therapies | X | X | X | X |
| Urine pregnancy test[2] | X | X | | |
| Randomization | | X | | |
| Photography (2D and 3D) | | X | X | X |
| Adverse events assessment | | X | X | X |
| Safety assessments[3] | X | X | X | X |
| Visual Function Assessments[5] | X | X | X | X |
| Ultrasound (at selected sites) | | X | X | X |
| **Treating Investigator Assessments** | | | | |
| GAIS | | | X | X |
| [redacted] | X | X | | |
| Radiance, skin tightness and jawline contour | | | X | X |
| **Blinded Evaluator Assessments** | | | | |
| [redacted] | X | X | X | X |
| **Subject Assessments** | | | | |
| GAIS | | | X | X |
| FACE-Q | | X | X | X |

1. *Screening and baseline visits may occur on the same day, if visits occur the same day assessments will not be duplicated*
2. *Females of childbearing potential*
3. *Safety assessments include; Cheek firmness, symmetry and function,), Mass formation, Cheek sensation*
4. *Weight only.*
5. *Visual Function assessments include Snellen visual acuity test, Extraocular muscle function test, Confrontation visual field*

***Abbreviations: GAIS (Global Aesthetic Improvement Scale), [redacted], ET (Early Termination)***

***Note: One month is defined as four weeks in the study.***

| GALDERMA EST. 1981 | Title<br>**43USSA1812 Statistical Analysis Plan - Sculptra Cheek Wrinkles** | Doc id<br>**MA-42004** |
|---|---|---|



# SIGNATURES PAGE

| Date | Signed by |
|---|---|
| 2021-08-30 11:59 | PPD |
| **Justification** | [redacted] |
| 2021-08-30 12:15 | PPD |
| **Justification** | [redacted] |
| 2021-08-30 12:36 | PPD y |
| **Justification** | [redacted] |
| 2021-08-30 14:55 | PPD |
| **Justification** | [redacted] |