Statistical Analysis Plan

Novan, Inc.

NI-MC304

A Phase 3 Multi-Center, Randomized, Double-Blind, Vehicle-Controlled, Parallel Group Study Comparing the Efficacy and Safety of SB206 and Vehicle Gel Once Daily in the Treatment of Molluscum Contagiosum

Protocol Version: 30Sep2020

Sponsor: Novan, Inc.
4105 Hopson Road
Suite 146
Morrisville, NC 27560

Prepared by: Christopher Tait
5909 Sea Otter Place, Suite 100
Carlsbad, CA 92010

Version	Date
Version 1	01 Oct 2020
Version 2	24 May 2021

Approval

Upon review of this document, including the table, listing, and figure shells, the undersigned approves the statistical analysis plan. The analysis methods and data presentation are acceptable.

Signature	Date
_____ Christopher Tait, PhD Principal Biostatistician Synteract	_____
_____ David Hebert, PhD Senior Director, Head of Biometrics Novan, Inc.	_____

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Full Notation
AE	adverse event
ATC	anatomical/therapeutic/chemical
BOTE	beginning of the end
CRO	contract research organization
DSMB	data safety monitoring board
e-consent	electronic consent
eCRF	electronic case report form
ET	early termination
ICH	International Council for Harmonisation
ITT	Intent-to-Treat
LSR	local skin reaction
MC	molluscum contagiosum
MedDRA	Medical Dictionary for Regulatory Activities
PP	Per-Protocol
QC	quality control
QD	once daily
SAP	statistical analysis plan
TEAE	treatment-emergent adverse event
TLFs	tables, listings, and figures

1. INTRODUCTION

This document outlines the statistical methods to be implemented during the analyses of data collected within the scope of Novan, Inc. protocol amendment 2 [A Phase 3 Multi-Center, Randomized, Double-Blind, Vehicle-Controlled, Parallel Group Study Comparing the Efficacy and Safety of SB206 and Vehicle Gel Once Daily in the Treatment of Molluscum Contagiosum]. The purpose of this plan is to provide specific guidelines from which the statistical analyses will proceed. Any deviations from this plan will be documented in the clinical study report.

2. STUDY DOCUMENTS

The following study documents are used for the preparation of the statistical analysis plan (SAP):

- Protocol Amendment 2, 30Sep2020
- Annotated electronic case report form (eCRF) version 1.0, 31Aug2020
- Data management plan version 1.0, 12Aug2020

3. STUDY OBJECTIVES

The objective of the study is to evaluate the efficacy and safety of SB206 10.3% once daily (QD) for the treatment of molluscum contagiosum (MC).

4. STUDY DESIGN AND PLAN

In response to the recent COVID-19 pandemic, it is imperative to include additional protections to enhance the assurance of subject safety while protecting the integrity of the clinical trial data. In alignment with the FDA's recommendation to utilize electronic and remote methods, the NI-MC304 study was designed to include electronic consent (e-consent) and remote technology for the study visits between Week 2 and Week 12, our primary endpoint, in an effort to minimize the amount of time that subjects and caregivers are required to be present in clinic, while also minimizing the number of contacts for our site personnel (FDA, 2020). Subjects/caregivers and investigators will practice at least one (simulated) remote assessment at the Baseline visit to familiarize each party with the technology before the first remote use. The frequency of planned on-site monitoring visits will be reduced, and more frequent remote monitoring visits will be incorporated to support protection of all team members. Enhanced central monitoring, telephone contact with the sites to review study procedures, trial participant status, and study progress, or remote monitoring of individual enrolled trial participants, will be performed as outlined in the monitoring plan. Additional information will be captured on any protocol deviations that occur due to COVID-19. Finally, a clinical trial continuity plan has been established for situations to outline the current modifications in case of further impact to the existing study design and plans. The plan incorporates contingency measures for additional disruptions due to COVID-19 and other unforeseen circumstances and provides a process for documenting any participants impacted by the circumstance and what was impacted. The SAP includes analyses that will be

performed to address the impact of the implemented contingency measures, and any necessary ad hoc analyses not incorporated into the SAP will be discussed in the clinical study report.

This is a Phase 3 multi-center, randomized, double-blind, vehicle-controlled, parallel group study to be conducted in approximately 850 subjects with MC. After obtaining informed consent/assent, subjects who satisfy entry criteria will be randomized 1:1 (active:vehicle) using an interactive web response system. Subjects receiving current treatment for MC at the time of the Screening Visit will enter a washout period of up to 14 days prior to randomization. In the event no washout period is required, Screening and Baseline visit activities may be combined into a single in-clinic visit. At randomization, subjects will be stratified by investigator type (dermatologist vs other), the subject's beginning-of-the end (BOTE) Inflammation Score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]) and number of randomly assigned subjects per household (1 subject per household vs 2 subjects per household). Subjects will be stratified to 5 strata first by subjects per household. Subjects from 1-subject households will be stratified by investigator type and baseline BOTE score. Subjects from 2-subject households will not be further stratified with respect to investigator type and baseline BOTE score because the overall sample that is expected for the stratum for households with two subjects is not large enough to support further stratification.

A maximum of two subjects from the same household may be randomized into the study. They must be randomized on the same day and both must individually meet all inclusion/no exclusion criteria. For subjects in the same household, Screening can occur on different days; however, the Baseline visit must occur on the same day. Households randomizing two subjects will receive the same treatment assignment for both subjects.

Subjects or their caregivers will apply SB206 10.3% or Vehicle Gel once daily for a minimum of 4 weeks and shall continue unless otherwise instructed by the investigator up to 12 weeks to all lesions identified at Baseline and new treatable lesions that arise during the course of the study. Subjects or their caregivers will continue to treat the area until the next scheduled visit even if the lesion(s) clear. At each in-clinic and remote visit, the investigator will count and record the number of active (raised, treatable) molluscum lesions per body area. Complete clearance should be confirmed by an in-clinic lesion count before treatment is discontinued. If the investigator determines all lesions are cleared at a visit, the investigator may instruct the subject to stop treatment. If treatment is stopped due to clearance, subjects will continue regularly scheduled visits through Week 12/ET and all procedures outside of study drug activities should be continued. If lesions recur or new lesions occur between visits after being stopped due to clearance, the subject or caregiver should re-initiate treatment to the recurring and new lesions. A new kit should be dispensed at Weeks 4 and 8 regardless of lesion count so that treatment can be restarted if lesions are observed between visits. An unscheduled visit is not required for the subject/caregiver to resume treatment. No study treatment is planned to be dispensed after the Week 12 visit.

Remote visits may be performed at Week 4 using remote technology. During the Baseline visit, the subject/caregiver will be instructed on the technology to be used for remote visits. After the

initial molluscum lesion count is completed, it will be repeated by the same assessor and the subject/caregiver using the subject/caregiver's electronic device that will be utilized to capture photographs for remote visits. This will provide the subject/caregiver an opportunity to practice with the technology and will allow for the assessor to understand lesion appearance in person compared to through the technology platform.

Subjects will be contacted via phone on Day 2 to collect subject information on early dose reactions. At Weeks 16 and 20, subjects will be contacted via phone to capture information regarding MC recurrence and adverse events (AEs); at Week 24, the subject will be seen at the site for a final study visit to assess scarring, keloid, and MC recurrence.

Subjects who discontinue the study prior to the Week 12 visit will be asked to complete the Week 12 visit assessments in the clinic: this will be recorded as an Early Termination (ET1) visit. Subjects who discontinue from the study after Week 12 but prior to Week 24 will be asked to come to the site to complete Week 24 assessments; this will be recorded as an ET2 visit.

Safety assessments include BOTE Inflammation Scores, local skin reaction (LSR) scores, AE collection, including scarring/keloid, and urine pregnancy tests. Safety assessments will be completed at specified visits through Week 12. Adverse events and concomitant medications will be reviewed and updated as needed at each visit through Week 24.

Inflammatory reactions around the MC lesions has been associated with imminent resolution of MC (sometimes referred to as "beginning-of-the-end" ["BOTE"] sign). The investigator (or designated evaluator) will assess the presence and overall degree of inflammatory reactions at MC lesions at Baseline (pre-dose) and Weeks 2 through 12 using the BOTE Inflammation Score in the table below. BOTE may be associated with itch, but not pain. BOTE is usually asymptomatic, self-limited, and localized to individual MC lesions and does not require discontinuation of study treatment or additional treatment. LSR is generally more diffuse, associated with significant itch or tenderness, may necessitate discontinuation of study treatment, and may need treatment for symptomatic relief (e.g., a topical corticosteroid or topical anesthetic). For very severe LSR, systemic corticosteroids may be considered. Investigators will assess the treatment area at each scheduled visit and use their medical judgement to differentiate between BOTE and LSR. BOTE Inflammation Score and LSR component scores will be recorded at each visit. BOTE should not be considered an AE. When LSRs are clinically significant at the application site, the investigator should report the condition as an AE(s).

Prior to the first application of study drug at Baseline, the investigator will assess the subject's skin as part of the physical examination, including the presence or absence of individual features of erythema, flaking/scaling, crusting, swelling, vesiculation/pustulation, and erosion/ulceration. At Baseline (at least 30 minutes after dosing) and Week 2 through Week 12, the investigators will rate LSRs on individual features including erythema, flaking/scaling, crusting, swelling, vesiculation/pustulation, and erosion/ulceration using the LSR Assessment Scale.

The following tables describe the scoring parameters for the BOTE Inflammation Score and LSR score.

BOTE Inflammation Score

Score	Global Assessment	Description
0	No inflammation	No evidence of local inflammation
1	Mild	Minimal erythema and/or edema
2	Moderate	Definite erythema and/or edema with or without hemorrhagic crusting
3	Severe	Erythema and edema with definite hemorrhagic crusting
4	Very severe	Strong reaction spreading beyond the treated area, bullous reaction, erosions

LSR Score

Score	Erythema	Flaking/ Scaling	Crusting	Swelling	Vesiculation/ Pustulation	Erosion/ Ulceration
0	Not present	Not present	Not present	Not present	Not present	Not present
1	Slightly pink	Mild, limited	Isolated crusting	Minimal, limited	Fine vesicles	Superficial erosion
2	Pink or light red	Moderate	Crusting < 50%	Mild, palpable	Scant transudate or exudate	Moderate erosion
3	Red, restricted to treatment area	Coarse	Crusting > 50%	Moderate	Moderate transudate or exudate	Marked, extensive
4	Red extending outside treatment area	Scaling extending outside treatment area	Crusting extending outside treatment area	Marked swelling extending outside treatment area	Marked transudate or exudate	Black eschar or ulceration

Adverse events will be assessed and collected after the initiation of study drug treatment through the end of the subject's last visit. Treatment-related AEs and all serious adverse events will be followed up until resolution or up to one year after last treatment, whichever is sooner.

Scar formation (scars/keloids/hypertrophic scars) will be assessed starting at the Week 4 visit through Week 24. The investigator will map locations of the molluscum lesions at Baseline. Additional lesions identified through Week 12 will be added to the map. Using the map as a guidance, the investigator will assess the treated areas for scar/keloid formation. All scars, including keloid/hypertrophic scars, that develop after the Baseline assessment should be captured as AEs and assessed for relatedness to study treatment.

If a subject's treatment is discontinued by the investigator or the subject because of an AE, that AE should be indicated as the reason for treatment discontinuation. All subjects will be encouraged to remain in the study and to complete all required study visits throughout the 24-week study duration.

When approximately 200 subjects are randomized, a data safety monitoring board (DSMB) will review all available unblinded safety data (including completed patch testing results). The DSMB will provide their recommendation for the study to proceed or not proceed.

5. DETERMINATION OF SAMPLE SIZE

Approximately 850 subjects, 6 months of age and older, with a minimum of 3 and a maximum of 70 MC lesions at Baseline will be randomized in a 1:1 (active:vehicle) scheme. The sample size assumptions for this study were informed by the integration of the completed phase 3 studies NI-MC301 and NI-MC302. Multiple replications of simulated random sampling of the integrated NI-MC301 and NI-MC302 data was performed to ensure the sample size assumptions align with the proposed design of NI-MC304. This includes the 1:1 ratio (active:vehicle) and the percentages observed in NI-MC301 and NI-MC302 for each stratum of each of the stratification factors: investigator type, number of randomized subjects per household, and baseline BOTE score that are planned for NI-MC304.

The analysis of the proportion of subjects with complete response of all treatable MC at Week 12 within the Intent-to-Treat (ITT) Population of the multiple replications of simulated random sampled integrated data yields rates of 20% for vehicle and 29.5% for SB206 10.3% QD with the covariate-adjusted treatment difference was 9.5%. A sample size of 850 subjects (425 subjects in the SB206 10.3% QD group and 425 subjects in the vehicle group) will provide 90% power for a 2-sided alpha test of size 0.05 to detect an absolute difference of 9.5% when the vehicle response rate is 20%.

Simulations were performed with 10,000 replications to confirm the accuracy of these power estimates (i.e., true power and simulation based estimate of power) and to provide simulated median, 10% percentile and lowest detectable treatment difference estimates for a sample size of 850 subjects.

At the two-sided alpha=0.05 level, the median detectable treatment difference is approximately 9.9% with a simulation based estimated power of 88%. The tenth percentile for detectable treatment difference is 7.1%, thus approximately 90% of the simulations with two-sided $p < 0.05$

produced an estimated treatment difference exceeding 7.1%. The lowest detectable treatment difference is 5.7%.

At two-sided alpha levels <0.05, for example, p=0.025 level, the median detectable treatment difference is approximately 10.2% with a simulation based estimated power of 82%. The tenth percentile for detectable treatment difference is 7.7% and a lowest detectable difference of 6.5%.

6. GENERAL ANALYSIS CONSIDERATIONS

The statistical analyses will be reported using summary tables, listings, and figures (TLFs). The International Council for Harmonisation (ICH) numbering convention will be used for all TLFs. Unless otherwise noted, all statistical testing will be two-sided and will be performed at the 0.05 significance level. Tests will be declared statistically significant if the calculated *P*-value is ≤0.05.

Continuous variable summaries will include the number of subjects (n) with non-missing values, mean, standard deviation (SD), median, minimum, and maximum. Other statistics such as quartiles, confidence intervals (CIs), and number of missing values may be added as appropriate.

Categorical variable summaries will include the frequency and percentage of subjects who are in the category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population within each treatment group, unless otherwise specified. The denominator for by-visit displays will be the number of subjects in the relevant study population with non-missing data at each visit.

All summary tables will be presented by treatment group. Baseline summaries will also include an overall summary column.

Individual subject data obtained from the eCRFs and any derived data will generally be presented by subject in data listings. Additionally, a listing of the subjects requiring a narrative because they experience one of the following events: serious adverse event, death, discontinued the study due to an AE, hypertrophic or keloid scarring, unresolved scars, or LSR suggestive of allergic contact dermatitis will be presented.

A separate report for the results from the study exit interview will be created.

The analyses described in this plan are considered a priori, in that they have been defined prior to breaking the blind.

Any analyses performed subsequent to breaking the blind will be considered post hoc and exploratory. Post hoc analyses will be labeled as such on the output and identified in the clinical study report.

All analyses and tabulations will be performed using SAS® software Version 9.4 or higher. Tables, listings, and figures will be presented in RTF format.

The process for SAS program validation and quality control (QC) for programs and outputs is documented in the Synteract working instruction "SAS programming quality control." Study-specific QC requirements can be found in Appendix B: SAS Programming QC Requirements.

The study treatment period is defined from randomization through completion of Week 12/ET1 visit and the safety follow-up period is from the completion of Week 12 through the completion of Week 24/ET2 visit.

After all subjects have completed their Week 12/ET1 visit, thus completing the Treatment Period of the study, the database through Week 12/ET1 will be frozen and unblinded for purposes of the primary analysis of efficacy and safety. While this analysis is being prepared, the subjects will continue through Week 24/ET2. After all subjects have completed their Week 24/ET2 visit, the database will be locked and the follow-up data will be analyzed.

7. NOTATION OF TREATMENT GROUPS AND VISITS

Analysis visits

Baseline is defined as the last non-missing value recorded prior to the first application of study drug. If time is not recorded and the assessment was on the same day as the first application of study drug, then it will be assumed the assessment occurred prior to the application of study drug. The baseline record will have an analysis visit of "Baseline". If a subject was randomized but not treated, then baseline is defined as the last non-missing value on or before the date of randomization.

For analysis of lesion counts, if there are multiple assessments of differing assessment types (remote vs in-clinic) within a given window, then the in-clinic lesion count will be used for analysis. For all evaluations of all parameters, assessments will be analyzed according to the visit at which they occurred (i.e., per the eCRF visit label). In order to account for the fact that the Week 12 and Early Termination visits are collected on the same eCRF in the database, a check against the Study Exit form will be made to determine if the visit is a Week 12 or an Early Termination Visit. If the visit is an Early Termination visit, then the visit will be mapped according to the table below. If there is no corresponding Study Exit form on the same date, then it will be considered a Week 12 visit. If there is no assessment in a given visit, then the following visit windows will be applied to determine if an unscheduled visit may be used:

Visit	Analysis Visit	Target Study Day	Study Day Analysis Window
Week 2 (±3 days)	Week 2	15	Day 12 to 22
Week 4 (±5 days)	Week 4	29	Day 23 to 43
Week 8 (±5 days)	Week 8	57	Day 44 to 71
Week 12 (±5 days)	Week 12	85	Day 72 to 98
Week 16 (±7 days)	Week 16	113	Day 99 to 126

Visit	Analysis Visit	Target Study Day	Study Day Analysis Window
Week 20 (±7 days)	Week 20	141	Day 127 to 154
Week 24 (±7 days)	Week 24	169	Day 155 to 182

Days are measured from the date of randomization. Study days corresponding to measurements are calculated as:

- Assessment date – date of randomization + 1, if assessment date is on or after the date of randomization.
- Assessment date – date of randomization, if measurement date is before the date of randomization.

In the event of multiple unscheduled records falling in the same analysis window, the assessment which is closest to the target study day will be chosen for analysis.

8. ANALYSIS POPULATIONS

The following subject population will be used for disposition analyses:

- The Enrolled Population will consist of all subjects who have a signed informed consent or assent. If a subject is randomly assigned to a treatment, then treatment assignment will be based on randomized treatment; otherwise, they will be deemed a screen failure and will only appear in the overall summary columns of output.

The following subject population will be used for safety analyses:

- The Safety Population will consist of all subjects who receive at least 1 application of study treatment. Treatment assignment will be based on the treatment actually received. If a subject receives any amount of SB206, then they will be summarized in the SB206 10.3% QD arm.

The following subject populations will be used for efficacy analyses:

- The ITT Population will consist of all subjects who are randomized who have a signed informed consent or assent as applicable. Treatment assignment will be based on the randomized treatment.
- The Per-Protocol (PP) Population will consist of all subjects in the ITT Population who had no significant protocol deviations that impacted the analyses of efficacy endpoints. Final determination of subject inclusion in the PP Population will be made prior to unblinding. Treatment assignment will be based on the randomized treatment.

9. STUDY POPULATION

9.1 Subject Disposition

Subject disposition information will be summarized for all subjects by treatment group. Summaries will include: the number of subjects screened, the number of screen failures, the number of subjects in each analysis population, the number of subjects where study treatment stopped, primary reason for study treatment stopped, the number of subjects completing 12 weeks of the study (defined as having answered Yes to the question "Did the subject complete treatment through Week 12?" on the end of treatment eCRF page or an end of treatment reason of "Complete Clearance Confirmed by Investigator Prior to Week 12" and does not have a Study Exit date prior to the subject's Week 12 visit date), the number of subjects completed the study, and the primary reason for discontinuation (including summaries for prior to Week 12 and overall).

A summary of the number of subjects in each center and each population will be presented.

9.2 Eligibility

A listing of subjects not fulfilling any eligibility criteria will be created.

9.3 Demographic and Baseline Characteristics

Demographic variables include age, sex, ethnicity, and race. Age will be calculated in years relative to the informed consent date.

Other baseline characteristics include lesion counts at baseline, site type (dermatologist vs other), number of sites and a breakdown of site type, number of randomly assigned subjects in household (1 subject vs 2 subjects), number of households and breakdown of household size, BOTE score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]), and age at and time since onset of symptoms of current molluscum episode. Descriptive statistics will be presented for ages and other continuous variables. Frequency counts and percentages will be presented for sex, ethnicity, race, and other categorical variables. Demographic and baseline characteristics will be summarized for the Safety, ITT, and PP Populations.

9.4 Extent of Exposure

Study drug exposure will be summarized for each treatment using the total number of applications, the duration of treatment, and the number of subjects requiring a treatment interruption and modification. Duration of treatment is defined as the last application date minus the first application date plus 1. A subject will be deemed to have a treatment interruption if there is at least 1 dose not applied due to investigator's instruction and to have a treatment modification if there is at least 1 dose which was modified due to investigator instruction.

Study drug compliance will be summarized and calculated as follows:

Compliance [%] = (Actual applications applied)/(Planned applications) ×100, where

- Actual applications applied = Planned applications – the number of applications not applied by the subject which were not investigator instructed.
- Planned applications = Number of applications (days) planned up to the point of treatment discontinuation or date of Week 12 visit, whichever is later.

This implies that if a subject discontinues treatment prior to Week 12 due to complete clearance, then the number of applications takes this information into account and the subject is not penalized for having complete clearance.

Compliance will be further summarized into 2 groups:

1. Subjects who did not have any interruptions or modifications using the same formula as above, and
2. Subjects with modifications or interruptions. For these subjects, an adjusted compliance will be calculated in the following manner: the actual applications applied will be the same as above but the planned applications will exclude the number of doses which were not applied due to investigator's instruction.

Subjects who do not have any interruptions or modifications will have no doses not applied due to investigator's instruction and no doses modified due to investigator's instruction.

9.5 Protocol Deviations

Significant protocol deviations that could potentially affect the efficacy or safety conclusions of the study will be identified prior to database lock and unblinding of individual subject treatment information. Significant protocol deviations may include, but are not limited to:

- Randomly assigned subjects who did not satisfy selected inclusion and exclusion criteria
- Randomly assigned subjects who developed withdrawal criteria during the study but were not withdrawn
- Subjects who were randomized incorrectly including subjects of the same household who are randomized to different treatments
- Subjects who received the wrong treatment
- Subjects where the subject/site staff were unproperly unblinded
- Week 12 Lesion Count not performed
- Subjects who received an excluded concomitant treatment.

The decision on the criteria for whether a subject is excluded from the PP Population will be made during the data review meeting prior to unblinding and database lock. Reasons for exclusion of a subject from the analysis will be listed.

A listing of all protocol deviations including the deviation designation (major or minor and significant or not), category, and indication of whether the deviation led to an exclusion of a subject from the PP Population will be presented in a data listing. Additionally, a listing of missed visits due to COVID-19 will be presented.

Major and significant protocol deviations will be summarized by deviation category and treatment group.

9.6 Medical History

Medical history verbatim terms in the eCRFs will be mapped to system organ classes and preferred terms using the Medical Dictionary for Regulatory Activities (MedDRA) Version 23.0 or later. Subject incidence of unique medical history terms by MedDRA system organ class and preferred term will be presented. The summary will be ordered by descending order of incidence in the SB206 treatment group of system organ class and preferred term within each system organ class.

9.7 Prior and Concomitant Medications

Prior and concomitant medication verbatim terms in the eCRFs will be mapped to anatomical/therapeutic/chemical (ATC) class and preferred names using the WHODrug Global B3 (version March 2020 or later). Prior medications are those medications started prior to the first application of study drug. Concomitant medications are those medications started on or after the date of first application of study drug or medications started prior to initial application of study drug and continued during the study. A medication can be classified as both prior and concomitant. If it cannot be determined whether the medication was a prior (or concomitant) medication due to a partial start or stop date, then it will be counted as both prior and concomitant; see Appendix A for the imputation of missing dates algorithm.

Prior and concomitant medications will be summarized for each treatment by WHODrug Global ATC class Level 3 and preferred name. These summaries will present the number and percentage of subjects using each medication. Subjects may have more than 1 medication per ATC class and preferred name. At each level of subject summarization, a subject is counted once if he/she reported 1 or more medications at that level. Each summary will be ordered by descending order of incidence in the SB206 treatment group of ATC class and preferred name within each ATC class.

10. EFFICACY ANALYSES

The primary efficacy analysis will be based on the ITT Population. Additional supportive efficacy analyses will be performed using the PP Population.

10.1 Efficacy Endpoints

The primary efficacy endpoint is the proportion of subjects with complete clearance of all treatable MC at Week 12.

The secondary efficacy endpoints are the following:

1) Proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Week 12.
2) Proportion of subjects achieving at least a 90% reduction from Baseline in the number of all treatable MC at Week 12.
3) Proportion of subjects with complete clearance of all treatable MC at Week 8.
4) Percent change from Baseline in the number of all treatable MC at Week 4.

Complete clearance is defined as having a total number of lesions count of 0.

The exploratory endpoints are the following:

- Percent change from Baseline in number of treatable MC at each visit (Weeks 2, 8, 12)
- Proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Weeks 2, 4, and 8
- Proportion of subjects achieving at least a 90% reduction from Baseline in the number of all treatable MC at Weeks 2, 4, and 8
- Proportion of subjects achieving at least a 75% reduction from Baseline in the number of all treatable MC at Weeks 2, 4, 8, and 12
- Absolute change from Baseline in number of treatable MC at Weeks 2, 4, 8, 12
- Proportion of subjects with complete clearance of all treatable MC at Week 2 and Week 4 visits
- Time to complete clearance of all treatable MC, defined as the days between the date of first dose and the first date of clearance. Subjects who do not achieve complete clearance will be censored at the date of the last lesion count assessment.
- Proportion of subjects who have a recurrence of MC after the first visit at which complete clearance was observed, defined as the total lesion count being greater than 0 at any point after achieving a count of 0
- Subject-reported spread to household members as measured by any new occurrence of MC in household members of subjects at each visit (Weeks 2, 4, 8, 12)
- Investigator Global Severity Assessment at Baseline, Week 12, and Week 24
- Subject Global Severity Assessment at Baseline, Week 12, and Week 24
- Investigator Global Impression of Change at Week 12 and Week 24

- Subject Global Impression of Change at Week 12 and Week 24

10.2 Baseline Values

Unless otherwise noted, baseline is defined as the last non-missing value recorded prior to the first application of study drug. If time is not recorded and the assessment was on the same day as the first application of study drug, then it will be assumed the assessment occurred prior to the application of study drug. If a subject was randomized but not treated, then baseline is defined as the last non-missing value on or before the date of randomization. For lesion counts, the in-clinic count will be used as baseline and sensitivity analyses on type of assessment may be performed.

10.3 Adjustments for Covariates

The model for the primary efficacy will include adjustments for the following covariates: investigator type (dermatologist vs other), household number of randomized subjects (1 subject per household vs 2 subjects per household), BOTE score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]), age (0 to <3, 3 to <4, 4 to <5, 5 to <6, 6 to <7, 8 to <9, 9 to <12, and ≥12 years old), and baseline lesion count. If there are not at least 3 responders and non-responders at each level of the stratification factors, then that stratification factor will be removed from the model. If there are fewer than 10 responders/non-responders, then neither investigator type nor household number of randomized subjects nor BOTE score at Baseline will be included in the model.

10.4 Handling of Dropouts or Missing Data

In analyses based on response, a missing response will be considered a non-responder.

For a sensitivity analysis, subjects with missing lesion count at Week 12 but who demonstrated complete clearance at the last collected lesion assessment will be counted as responders.

An additional sensitivity analysis using multiple imputation will be performed to address subjects with missing lesion count at Week 12. A monotone imputation model will be utilized in the following steps:

1. The lesion count data will be imputed to follow a monotone missing data pattern using the Markov chain Monte Carlo method for 10 imputations.
2. Each of the 10 imputations will then be imputed a further 10 times using a monotone regression method to impute the remaining missing data and the Week 12 lesion count using lesion count assessments at Weeks 2, 4, and 8, treatment group, investigator type (dermatologist vs other), BOTE score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]), age (0 to <3, 3 to <4, 4 to <5, 5 to <6, 6 to <7, 8 to <9, 9 to <12, and ≥12 years old), and baseline lesion count as covariates.

3. Complete Clearance at Week 12 will be re-derived based on the imputed values and then analyzed using the primary analysis model. The results of the 100 analyses will be transformed into a normal statistic and combined into a single analysis using PROC MIANALYZE.

The ROUND and MINIMUM options will be utilized in Steps 1 and 2 to ensure imputed values are non-negative integers. The seed to be used is 20200727.

For subjects with a missing lesion count at Week 12 and identified as "Lost to Follow-up" by the investigator, the following sensitivity analysis will be performed. These lost to follow-up subjects will be contacted to obtain a patient reported outcome of MC disease status at the end of their study participation of either 'Resolved (no lesions left)', 'Resolving (number of lesions decreasing)', 'Unchanged (number of lesions similar to previous clinic visit', or 'Worsening (number of lesions increased)'. A reported response endpoint will be derived in the following manner:

- If a subject has a Week 12 lesion count assessment, then the subject will be deemed a responder if the count was 0.
- If the subject did not have a Week 12 lesion count assessment and the response to the disease status was 'Resolved (no lesions left)' or 'Resolving (number of lesions decreasing)', then the subject will be deemed a responder.
- Otherwise, the subject will be deemed a non-responder.

The sensitivity analysis of this patient reported response will be analyzed in the same manner as the primary endpoint. A list of the patient reported outcome of MC disease status will be provided.

A sensitivity analysis on the manner of assessment of lesion counts at Week 12 (in-clinic vs. remote) may be performed where subjects with only a remote assessment will be counted as non-responders.

10.5 Interim Analysis and Data Monitoring

When approximately 200 subjects are randomized, a DSMB will review all available unblinded safety data (including patch testing). All responsibilities of the DSMB and details of the analysis and data to be reviewed is detailed in the DSMB charter.

10.6 Examination of Subgroups

Subgroup analyses of complete clearance at Week 12 for investigator type (dermatologist vs other), household number of randomized subjects (1 subject per household vs 2 subjects per household), and BOTE score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]) will be presented in a forest plot for the ITT Population. An additional subgroup analysis of complete clearance at Week 12 for age

(<1 year old, ≥1 to <2 years old, ≥2 to <6 years old, ≥6 to <12 years old, ≥12 to <18 years old, ≥18 years old) and onset of MC symptoms duration prior to baseline (≤6 months vs >6 months) will be summarized descriptively for the ITT Population.

10.7 Multiple Comparison/Multiplicity

The familywise error rate with respect to the primary endpoint and secondary endpoints will be strongly controlled at the alpha=0.05 level using a hierarchical fixed sequence method testing strategy in the order listed in Section 10.1. If the primary endpoint is not statistically significant at the alpha=0.05 level, the secondary efficacy endpoints will be considered not significant. If the primary endpoint is statistically significant at the alpha=0.05 level, then the secondary efficacy endpoints will be tested in the following hierarchical fixed sequence:

1) Proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Week 12.
2) Proportion of subjects achieving at least a 90% reduction from Baseline in the number of all treatable MC at Week 12.
3) Proportion of subjects with complete clearance of all treatable MC at Week 8.
4) Percent change from Baseline in the number of all treatable MC at Week 4.

At each subsequent test of the secondary efficacy endpoints, if the secondary efficacy endpoint is not statistically significant at the alpha=0.05 level, the remaining secondary efficacy endpoints will be considered not statistically significant. If the secondary endpoint is statistically significant at the alpha=0.05 level, then the next subsequent secondary efficacy endpoint will be tested following the sequence.

10.8 Multicenter Studies

This is a multicenter study, having approximately 55 centers participating in the study. The center effects will be investigated in the primary statistical analysis model by including the treatment by investigator type interaction. As exploratory analyses to examine the variability of the treatment effect across centers, the following analyses will be performed:

- Pool the centers into self-standing center pools of at least 18 subjects. The pooling will sort by investigator type and number of randomized subjects. Then, the centers with fewer than 18 subjects will be pooled until the self-standing center pool is at least 18 subjects. The treatment effect will be explored using a logistic model including treatment, household number of randomly assigned subjects, age, baseline lesion counts, BOTE score at Baseline, and a strata statement which will include the pooling identified and examining the residual score statistics. To address treatment by center interaction, each pooling center will have a corresponding variable coded as 1 if on SB206 and 0 otherwise. Some pooled centers may have a p-value of less than 0.05 by chance since there will be multiple pooled centers.

- Pool the centers using the method above but the self-standing center pools will contain at least 12 subjects. This analysis should be interpreted with caution due to the small sample size within pooled centers.

11. METHODS OF EFFICACY ANALYSIS

11.1 Primary Efficacy Analysis

The primary efficacy comparison will test the following hypotheses:

H_0: The proportion of subjects with complete clearance is equal between SB206 10.3% QD and Vehicle;

H_1: The proportion of subjects with complete clearance is different between SB206 10.3% QD and Vehicle.

The primary efficacy model will be the following for π_{H_i} as the probability of complete clearance for the i^{th} subject in household H:

$$\pi_{H_i} = e^{(\beta_0+\beta_1*X_{H_{i1}}+\beta_2*X_{H_{i2}}+\beta_3*X_{H_{i3}}+\beta_4*X_{H_{i4}}+\beta_5*X_{H_{i5}}+\beta_6*X_{H_{i6}}+\beta_7*X_{H_{i7}}+\beta_8*X_{H_{i8}}+\beta_9*X_{H_{i9}}+\beta_{10}*X_{H_{i10}}+\beta_{11}*X_{H_{i11}}+\beta_{12}*X_{H_{i12}}+\beta_{13}*X_{H_{i13}})} \div [1 + e^{(\beta_0+\beta_1*X_{H_{i1}}+\beta_2*X_{H_{i2}}+\beta_3*X_{H_{i3}}+\beta_4*X_{H_{i4}}+\beta_5*X_{H_{i5}}+\beta_6*X_{H_{i6}}+\beta_7*X_{H_{i7}}+\beta_8*X_{H_{i8}}+\beta_9*X_{H_{i9}}+\beta_{10}*X_{H_{i10}}+\beta_{11}*X_{H_{i11}}+\beta_{12}*X_{H_{i12}}+\beta_{13}*X_{H_{i13}})}]$$

where H = Household, i = subject within the household, β_0 is the intercept,

$$X_{H_{i1}} = \begin{cases} 1, if\ treatment\ is\ SB206 \\ 0, if\ treatment\ is\ Vehicle \end{cases}$$

$$X_{H_{i2}} = \begin{cases} 1, if\ subject\ is\ from\ a\ 2\ subject\ household \\ 0, if\ subject\ is\ from\ a\ 1\ subject\ household \end{cases}$$

$$X_{H_{i3}} = \begin{cases} 1, if\ subject\ is\ from\ a\ Dermatology\ site \\ 0, if\ subject\ is\ from\ an\ Other\ site \end{cases}$$

$$X_{H_{i4}} = \begin{cases} 1, \text{if subject has a Baseline BOTE Score of 0} \\ 0, \text{if subject has a Baseline BOTE Score of} \geq 1 \end{cases}$$

$$X_{H_{i6}} = \begin{cases} 1, \text{if subject has a baseline age 0 to} < 3 \\ 0, \text{otherwise} \end{cases}$$

$$X_{H_{i7}} = \begin{cases} 1, \text{if subject has a baseline age 3 to} < 4 \\ 0, \text{otherwise} \end{cases}$$

$$X_{H_{i8}} = \begin{cases} 1, \text{if subject has a baseline age 4 to} < 5 \\ 0, \text{otherwise} \end{cases}$$

$$X_{H_{i9}} = \begin{cases} 1, \text{ if subject has a baseline age 5 to <6} \\ 0, \text{ otherwise} \end{cases}$$

$$X_{H_{i10}} = \begin{cases} 1, \text{ if subject has a baseline age 6 to <7} \\ 0, \text{ otherwise} \end{cases}$$

$$X_{H_{i11}} = \begin{cases} 1, \text{ if subject has a baseline age 7 to <8} \\ 0, \text{ otherwise} \end{cases}$$

$$X_{H_{i12}} = \begin{cases} 1, \text{ if subject has a baseline age 8 to <9} \\ 0, \text{ otherwise} \end{cases}$$

$$X_{H_{i13}} = \begin{cases} 1, \text{ if subject has a baseline age 9 to <12} \\ 0, \text{ otherwise} \end{cases}$$

$X_{H_{i5}}$ is the subject's baseline lesion count; treatment will be included in the class statement with PARAM=REF and REF=Vehicle. The working correlation will have an exchangeable structure. In the event of issues with convergence, an independent structure will be utilized. The model will include a repeated statement for subject household with household ID in the class statement. Also, the data structure will be ordered according to household ID and subject ID in household.

Treatment groups will be compared using a generalized estimating equation for logistic regression with an exchangeable working correlation structure. The model will include treatment, investigator type (dermatologist vs other), household number of randomly assigned subjects (1 subject per household vs 2 subjects per household), BOTE score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]), age (0 to <3, 3 to <4, 4 to <5, 5 to <6, 6 to <7, 8 to <9, 9 to <12, and ≥12 years old), and baseline lesion count as factors. For subjects from a household with 2 subjects, the actual values of BOTE score at Baseline and investigator type will be utilized with the recognition that such subjects did not have further stratification according to investigator type and BOTE score at Baseline. The odds ratio between SB206 10.3% and vehicle gel, 95% confidence intervals for the odds ratio, and *P*-value for the covariate-adjusted treatment comparison will be presented together with predicted proportions along with their associated 95% confidence interval. The difference in proportion confidence interval will be calculated using the following formula:

$$(p_{SB206} - p_{Vehicle}) \pm z_{0.025}{}^{*} * SE$$

Where the SE =

$$\sqrt{\begin{array}{c} p_{SB206}^2(1-p_{SB206})^2\, s_{SB206}^2 + p_{Vehicle}^2(1-p_{Vehicle})^2\, s_{Vehicle}^2 - 2p_{SB206}(1-p_{SB206})p_{Vehicle} \\ (1-p_{Vehicle}) * \left(\frac{s_{SB206}^2 + s_{Vehicle}^2 - s_{Odds\ Ratio}^2}{2}\right) \end{array}}$$

The p_{SB206} and $p_{Vehicle}$ are the transformed predicted log odds at the mean of the covariates, s^2's are the standard errors of the predicted log odds and the odds ratio.

As a sensitivity analysis, the above primary analysis model will be applied to the PP Population. An additional sensitivity analysis counting subjects who discontinued prior to Week 12 or have a missing lesion count assessment but who demonstrated complete clearance at the last collected lesion assessment as responders will be presented. The impact of mis-stratifications may be explored via a sensitivity analysis.

An analysis where the complete clearance response probability for each dropout prior to Week 12 will be independently generated across the following scenarios (Vehicle response probability, SB206 response probability):

(0.1, 0.1)	(0.1, 0.0)			
(0.2, 0.2)	(0.2, 0.1)	(0.2, 0.0)		
(0.3, 0.3)	(0.3, 0.2)	(0.3, 0.1)	(0.3, 0.0)	
(0.4, 0.4)	(0.4, 0.3)	(0.4, 0.2)	(0.4, 0.1)	(0.4, 0.0)

The analysis will be produced 100 times and then the results for the estimated log odds ratio for treatment and its estimated standard error from each of the imputations will be combined using PROC MIANALYZE.

An additional sensitivity analysis will be performed in a subset of the ITT population where only 1 subject from each household is chosen to contribute to the model based on the subject with the highest number of baseline treatable MC and the oldest to break ties, if needed. A logistic regression model including treatment, investigator type, BOTE score at Baseline. age, and baseline lesion counts will be utilized.

11.2 Secondary Efficacy Analyses

The secondary endpoints of the proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Week 12, proportion of subjects achieving at least a 90% reduction from Baseline in the number of all treatable MC at Week 12, and the proportion of subjects with complete clearance of all treatable MC at Week 8 will be analyzed in the same manner as the primary endpoint.

Since the within-household correlation is expected to be small, the percent change from Baseline in the number of all treatable MC at Week 4 will be analyzed using a repeated measures mixed model for the respective visits with the same covariates as the primary model together with visits and treatment by visit interaction; an unstructured covariance matrix will be utilized. If the calculation of the percent change from baseline is influenced by outliers with calculated values >100%, then the influence of outliers will be avoided for analysis by censoring them so that all values were in the range of -100% to 100%. If the model fails to converge, then other correlation structures will be explored.

A sensitivity analysis using the PP Population will be provided for the proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Week 12 and the proportion of subjects achieving at least a 90% reduction from Baseline in the number of all treatable MC at Week 12.

11.3 Exploratory Analyses

Since the within-household correlation is expected to be small, the change and percent change from Baseline in the number of treatable MC at Weeks 2, 8 and 12 will be analyzed using a repeated measures mixed model for the respective visits with the same covariates as the primary model together with visits and treatment by visit; an unstructured covariance matrix will be utilized. If the calculation of the percent change from baseline is influenced by outliers with calculated values >100%, then the influence of outliers will be avoided for analysis by censoring them so that all values were in the range of -100% to 100%. If the model fails to converge, then other correlation structures will be explored.

The exploratory endpoints based on the proportion of subjects achieving complete clearance (at Weeks 2 and 4) or 75% (at Weeks 2, 4, 8, and 12) or 90% (at Weeks 2, 4, and 8) reductions will be analyzed in the same manner as the primary endpoint.

Likewise, the proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Weeks 2, 4 and 8 will be analyzed in the same manner as the primary endpoint.

Since the within-household correlation is expected to be small and time to first complete clearance of all MC is exploratory, it will be analyzed using Kaplan-Meier methods. The number and percentage of subjects achieving complete clearance, number and percentage of censored subjects, and Kaplan-Meier estimates of first quartile, median, and third quartile will be summarized by treatment group. Differences in Kaplan-Meier curves between the treatments will be tested for significance using a stratified log-rank test.

The proportion of subjects who have a recurrence of MC after the first visit at which complete clearance was observed will be summarized.

The subject-reported spread of MC to household members not in the study will be summarized descriptively including a breakdown of whether or not there was any spread and then a breakdown of the amount of spread within the household at each visit at the household level.

A shift table comparing the baseline global severity assessment to each scheduled postbaseline assessment will be presented by subject and investigator. The Investigator and Subject Global Impression of change will be summarized descriptively.

Separate plots of the cumulative distribution function of the percentage change from Baseline in lesion count at Week 12 across each of the scores for the Subject Global Impression of Change and the Subject Global Severity Assessment will be presented for all subjects. These plots will also be replicated separately for subjects in the SB206 10.3% QD group and for subjects in the Vehicle group.

A plot comparing Improved vs Not Improved, including the correlation between percentage change from Baseline in lesion count at Week 12 and improvement category, in each treatment group and overall will be presented separately for the Subject Global Impression of Change at Week 12. Improved will be defined as selecting either 'Very Much Improved', 'Much Improved', or 'Minimally Improved' and not improved as selecting either 'No Change', 'Minimally Worse', 'Much Worse', or 'Very Much Worse'.

A plot comparing No Severity vs Severity, including the correlation between percentage change from Baseline in lesion count at Week 12 and severity category, in each treatment group and overall will be presented separately for the Subject Global Severity Assessment at Week 12. No Severity will be defined as having a response of 'None' and Severity will be defined as having a response of 'Mild', 'Moderate', 'Severe', or 'Very Severe'.

The correlation between the following will be presented.

- Baseline lesion count and Subject Global Severity Assessment at Baseline
- Baseline lesion count and Subject Global Severity Assessment at Week 12
- Baseline lesion count and Subject Global Impression of Change at Week 12
- Baseline lesion count and Investigator Global Severity Assessment at Baseline
- Baseline lesion count and Investigator Global Severity Assessment at Week 12
- Baseline lesion count and Investigator Global Impression of Change at Week 12

The following efficacy endpoints will be descriptively summarized for the Improved vs Not Improved categories of the Subject Global Impression of Change at Week 12.

- Proportion of subjects with complete clearance of all treatable MC at Week 12.
- Proportion of subjects achieving a lesion count of 0 or 1 of all treatable MC at Week 12.
- Proportion of subjects achieving at least a 90% reduction from Baseline in the number of all treatable MC at Week 12.
- Proportion of subjects achieving at least a 75% reduction from Baseline in the number of all treatable MC at Week 12.
- Percent change from Baseline in the number of treatable MC at Week 12.

12. OTHER ANALYSES

The following analyses of BOTE vs lesion count will be presented:

- Percent change from Baseline in the number of treatable MC at each visit in relation to the Baseline BOTE score comparing No Inflammation vs Inflammation within each treatment group through a repeated measures mixed model with number of randomly assigned subjects in household (1 subject vs 2 subjects), investigator type (dermatologist

vs other), BOTE score at Baseline (no inflammation [BOTE=0] vs mild/moderate/severe/very severe inflammation [BOTE≥1]), BOTE score at Baseline by visit interaction, age, and baseline lesion count as covariates with a compound symmetry covariance matrix.

- Percent change from Baseline in the number of treatable MC at Week 12 in relation to the highest post-baseline BOTE score during treatment will be summarized descriptively.
- Percent change from Baseline in the number of treatable MC at Week 12 in relation to the highest BOTE score at any visit (including Baseline) will be summarized descriptively.
- Percent change from Baseline in the number of treatable MC at each visit in relation to the BOTE score at Week 2 will be summarized descriptively.

In these analyses, the BOTE scores will be analyzed in the ITT Population as follows:

1. Dichotomized:
    a. Score of 0: No Inflammation
    b. Score of 1, 2, 3, or 4: Mild to Very Severe

A shift table comparing the baseline BOTE score to each scheduled post-baseline assessment will be presented for the ITT Population.

13. SAFETY ANALYSES

All safety analyses will be based on the Safety Population.

A listing of all deaths will be presented.

13.1 Adverse Events

Adverse event summaries restricted to treatment-emergent AEs (TEAEs) will include those AEs that occurred any time on or after the first in-clinical application of study drug through the last application of study drug and those existing AEs that worsened during this same period. If it cannot be determined whether the AE is treatment emergent due to a partial onset date, then it will be counted as such; see Appendix A for the imputation of missing dates algorithm. Post treatment AEs will include those AEs that occurred after the last application of study drug. Verbatim terms in the eCRFs will be mapped to preferred terms and system organ classes using the MedDRA Version 23.0 or later.

Each AE summary will be displayed by treatment group. Summaries that are displayed by system organ classes and preferred terms will be ordered by descending order of incidence in SB206 treatment group of system organ class and preferred term within each system organ class. Summaries of the following types will be presented:

- Overall summary of TEAEs and post treatment AEs that contain an overview of each item below.

- Subject incidence of TEAEs with total number of unique TEAEs and post treatment AEs with total number of unique post treatment AEs by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs and post treatment AEs by MedDRA system organ class, preferred term, and maximum severity. At each level of subject summarization, a subject is classified according to the maximum severity if the subject reported 1 or more events. Adverse events with missing severity will be considered severe for this summary.
- Subject incidence of TEAEs by MedDRA system organ class, preferred term, and closest relationship to study drug (Related/Not Related). Related AEs are those reported as "Definite," "Probable," or "Possible," and unrelated AEs are those reported as "Unlikely" or "Unrelated." At each level of subject summarization, a subject is classified according to the closest relationship if the subject reported 1 or more events. Adverse events with a missing relationship will be considered related for this summary.
- Subject incidence of serious TEAEs with total number of unique serious TEAEs and post treatment serious AEs with total number of unique post treatment serious AEs by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs leading to study drug discontinuation by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs leading to study discontinuation and post treatment AEs leading to study discontinuation by MedDRA system organ class and preferred term.
- Subject incidence of TEAEs leading to death as an outcome by MedDRA system organ class and preferred term.
- Subject incidence of Transient TEAEs, identified based on a search of verbatim terms containing 'TRANSIENT'.

Separate listings of all AEs, all SAEs, all TEAEs leading to study drug discontinuation, all AEs leading to study discontinuation, ongoing treatment related AEs, and all post treatment AEs will be provided.

The proportion of subjects with scarring and the proportion of subjects with keloid or hypertrophic scarring will be summarized descriptively at each visit.

13.2 Local Skin Reaction

The LSR composite score will be calculated by summing up all the numerical responses (0-4) to each individual parameter for a composite score that ranges between 0 and 24. The observed and change from Baseline values of the LSR composite score will be summarized descriptively at each visit. This analysis will be repeated based on the dichotomized baseline BOTE score described in Section 12.

A table summarizing each LSR parameter (erythema, flaking/scaling, crusting, swelling, vesiculation/pustulation, and erosion/ulceration) score at each scheduled post-baseline assessment will be presented. This analysis will be repeated based on the dichotomized baseline BOTE score described in Section 12.

Additionally, a shift table comparing the baseline score for each LSR parameter to each scheduled postbaseline assessment will be presented for the Safety Population.

13.3 Events of Special Interest

Events of Special Interest will include subjects with any TEAE where the preferred term contains "application site", subjects with at least 1 post-baseline occurrence of moderate BOTE inflammation score, or subjects with at least 1 post-baseline LSR component score ≥1. A summary including the number and percentage of subjects with at least 1 event of special interest, AEs of interest by MedDRA system organ class and preferred term, at least 1 moderate BOTE inflammation score, and any LSR component score ≥1 will be provided for the Safety Population. The analysis will also be repeated with the LSR criteria revised to include any post-baseline LSR component (excluding erythema) score ≥1 and erythema score ≥2.

13.4 Urine Pregnancy Test

Urine pregnancy test results will be included in a data listing only.

13.5 Physical Examination

A shift table summarizing the shift from Baseline to end of treatment in normal/abnormal will be presented.

13.6 Patch Testing

A table summarizing any patch testing results by time point may be presented.

14. CHANGES TO PROTOCOL-SPECIFIED ANALYSES

There were no modifications and/or clarifications to the methodology specified in the protocol.

15. REFERENCES

US Department of Health and Human Services (DHHS), Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Guidance for industry ICH E9: Statistical principles for clinical trials. September 1998 [cited 2019 May 20]. Available from: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073137.pdf

APPENDICES

Appendix A: Presentation of Data and Programming Specifications

General

- Specialized text styles, such as bold, italics, borders, and shading will not be used in tables, figures, and data listings unless they add significant value to the table, figure, or data listing.
- Only standard keyboard characters are to be used in tables and data listings.
- Special characters, such as nonprintable control characters, printer-specific, or font-specific characters, will not be used on a table, figure, or data listing.
- Hexadecimal character representations are allowed (e.g., μ, α, ß).
- All footnotes will be left justified and at the bottom of a page. Footnotes must be used sparingly and must add value to the table, figure, or data listing.

Tables

- Formal organization of tabulations may be changed during programming, if appropriate, e.g., tables for the different variables may be combined into a single table, or tables with more than 1 variable may be split into several tables.
- Means and medians will be presented to 1 more decimal place than the raw data. Standard deviations will be presented to 2 more decimal places than the raw data. Minimums and maximums will be reported with the same number of decimal places as the raw data.
- Percentages will be presented to the tenths place.
- For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness. For example, if none of the subjects discontinue due to "lost to follow-up," this reason will be included in the table with a count of 0. Categories with zero counts will not have zero percentages displayed.
- Lower and upper confidence interval values must be presented to 1 decimal place more than the raw/derived data (i.e., to the same number of decimal places as the mean).
- Percentiles (e.g., 25%, 75%) must be presented to 1 decimal place more than the raw/derived data.
- For all inferential analyses, *P* values will be rounded to 4 decimal places (or at the highest level of precision) with a leading zero (0.0001). *P* values less than 0.0001 will be presented as "<0.0001."
- The last footnotes will be
  - "Source: xxx", where xxx indicates the source **table number**(s) if applicable (in case aggregated results like mean or median are plotted) or the source listing(s) (in case individual responses are plotted) and/or source dataset(s) (e.g., AdaM).
  - "PROGRAM SOURCE: …\\xx.sas, DATA CUT OFF DATE: DDMMMYYYY, RUN DATE: DDMMYY hh:mm".

  where extract date (e.g., data cut off, database lock) is the date stamp of the data snapshot used.

Figures

- Legends will be used for all figures with more than 1 variable or item displayed. Treatment group sizes (n=xx) will be included, as appropriate.
- Figures will be in black and white but can be in color to add value to the clarity and readability of a figure. Lines must be wide enough to see the line after being copied.
- For box plots, the horizontal line will represent the median, + represents the group mean, the length of the box represents the interquartile range (25^{th}-75^{th} percentiles), and the whiskers will represent the minimum and maximum.
- The last footnotes will be
  - "Source: xxx", where xxx indicates the source listing number(s) and/or source dataset(s) (e.g., AdaM).
  - "PROGRAM SOURCE: …\\xx.sas, DATA CUT OFF DATE: DDMMMYYYY, RUN DATE: DDMMYY hh:mm".

  where extract date (e.g., data cut off, database lock) is the date stamp of the data snapshot used.

Listings

- Formal organization of the listing may be changed during programming, if appropriate, e.g., additional variables may be included, change in the column order, or the listing may be split into multiple parts due to space constraints.
- If not otherwise specified, all data listings will be sorted by sequence/treatment, center, subject number, visit, and date/time, as appropriate.
- All date values will be presented in a SAS date (e.g., 29AUG2001) format.
- All observed time values will be presented using a 24-hour clock HH:MM:SS format (e.g., 01:35:45 or 11:26). Seconds will only be reported if they were measured as part of the study.
- The last footnote will be
  - "PROGRAM SOURCE: …\\xx.sas, DATA CUT OFF DATE: DDMMMYYYY, RUN DATE: DDMMYY hh:mm".

  where extract date (e.g., data cut off, database lock) is the date stamp of the data snapshot used.

Missing or incomplete dates (i.e., AEs, concomitant medications, and start dates of current molluscum episode and initial diagnosis)
The most conservative approach will be systematically considered. If the AE onset date is missing/incomplete, it is assumed to have occurred during the study treatment phase (i.e., considered a TEAE) except if the partial onset date or other data, such as the stop date, indicates differently. Similarly, a medication with partial start and stop dates could be considered as both a prior and concomitant treatment.

The following algorithms will be applied to missing and incomplete start and stop dates:

Start Dates

- If the day portion of the start date is missing, then the start date will be estimated to be equal to the date of first application of study drug, provided the start month and year are the same as the first application of study drug and the stop date is either after the first application of study drug or completely missing. Otherwise, the missing day portion will be estimated as "01."
- If both the day and month portions of the start date are missing, then the start date will be estimated to be equal to the date of first application of study drug, provided the start year is the same as the first application of study drug and the stop date is either after the first application of study drug or completely missing. Otherwise, the event will be assumed to start on the first day of the given year (e.g., ??-???-2013 is estimated as 01-JAN-2013) or the subject's date of birth, whichever is latest.
- If the start date is completely missing and the stop date is either after the application of study drug or completely missing, the start date will be estimated to be the first day of study drug application. Otherwise, the start date will be estimated to be the first day of the same year as the stop date or the subject's date of birth, whichever is latest. All other non-AE and non-concomitant medication day calculations where only partial dates are available will be handled as follows: the first day of the month will be used in the calculations if the day part of a start date is missing while January 1 will be employed if both the month and day parts of a start date are missing.

Stop Dates

- If only the day of resolution is unknown, the day will be assumed to be the last day of the month (e.g., ??-JAN-2013 will be treated as 31-JAN-2013).
- If both the day and month of resolution are unknown, the event will be assumed to have ceased on the last day of the year (e.g., ??-???-2013 will be treated as 31-DEC-2013).
- If the stop date is completely missing and the event is not continuing, the event will be assumed to be after first application of study drug and will be imputed using the last known date on the study.

For the start dates of current molluscum episode and initial diagnosis, the day will be estimated as "01" if only day is missing. If both month and day are missing, then it will be estimated as "01-JUN-YYYY" or the subject's date of birth, whichever is latest.

If the start date of current molluscum episode is partial, then the following imputation will be made:

- If only the day is unknown, the day will be assumed to be the first day of the month or the date of birth, whichever is later.
- If the day and month are missing, then the start date will be estimated to be June 1st or the date of birth, whichever is later.

Standard Calculations

Variables requiring calculation will be derived using the following formulas:

- **Days** – A duration expressed in days between 1 date (date1) and another later date (date2) is calculated using the formulas noted below:
  Duration in days = date2 – date1 + 1.
- **Months –** A duration expressed in months will be calculated as (later date – earlier date + 1)/(30.4167).
- **Years –** A duration expressed in years will be calculated as (later date – earlier date + 1)/(365).
- **Change from Baseline** – Change from Baseline will be calculated as follows:
  Change from Baseline = postbaseline value – baseline value.
- **Percent change from Baseline** – Change from Baseline will be calculated as follows:
  Percent change from Baseline = (postbaseline value – baseline value)/baseline value × 100.

Appendix B: SAS Programming QC Requirements

Derived datasets are independently programmed by two programmers. The separate datasets produced by the 2 programmers must match 100%. Detailed specifications for the derived datasets are documented in the study analysis dataset specifications provided to the client at study conclusion.

Tables are independently reprogrammed by a second programmer for numeric results.
Listings are checked for consistency against corresponding tables, figures, and derived datasets.
Figures are checked for consistency against corresponding tables and listings, or independently reprogrammed if there are no corresponding tables or listings.

The entire set of TLFs is checked for completeness and consistency prior to its delivery to the client by the lead biostatistician and a senior level, or above, reviewer.

Appendix C: List of Tables, Figures, and Listings

The following proposal for section 14 and 16.2 is completed according to ICH E3 guidelines. The ICH heading numbers and description are in **bold**. Minor changes from this planned index do not need to be amended in the SAP.

Formal organization of tabulations may be changed during programming, if appropriate, e.g., tables for the different variables may be combined into a single table, or tables with more than 1 variable may be split into several tables.

The shells for the outputs below are contained in the file titled "Novan NI-MC302 SAP TLFs Version 2.0".

TABLES, FIGURES, AND GRAPHS

Table Number	Table Title	Analysis Population
14	**TABLES, FIGURES, AND GRAPHS REFERRED TO BUT NOT INCLUDED IN THE TEXT**	
14.1	**DEMOGRAPHIC DATA**	
14.1.1.1	Subject Disposition	Enrolled
14.1.1.2	Enrollment by Site	ITT
14.1.2.1	Demographic and Baseline Characteristics	Safety
14.1.2.2	Demographic and Baseline Characteristics	ITT
14.1.2.3	Demographic and Baseline Characteristics	PP
14.1.3.1	Study Drug Exposure	Safety
14.1.3.2	Study Drug Compliance	Safety
14.1.4	Significant and Major Protocol Deviations	ITT
14.1.5	Medical History	ITT
14.1.6.1	Prior Medications	ITT
14.1.6.2	Concomitant Medications	ITT
14.2	**Efficacy data**	
14.2.1.1	Complete Clearance Lesion Count Response at Week 12	ITT
14.2.1.2	Complete Clearance Lesion Count Response at Week 12	PP

Table Number	Table Title	Analysis Population
14.2.1.3	Sensitivity Analysis: Complete Clearance Lesion Count Response at Week 12	ITT
14.2.1.4	Sensitivity Analysis: Complete Clearance Lesion Count Response at Week 12	PP
14.2.1.6	Proportion of Subjects with Complete Clearance at Week 12 by Subgroup	ITT
14.2.1.7	Dropout Sensitivity Analysis: Complete Clearance Lesion Count Response at Week 12	ITT
14.2.1.8	One Subject Per Household Sensitivity Analysis: Complete Clearance Lesion Count Response at Week 12	ITT
14.2.1.9	Exploratory Analysis of the Variability of Treatment Effect of Complete Clearance Lesion Count Response at Week 12 Across Pools of at Least 18 Subjects	ITT
14.2.1.10	Exploratory Analysis of the Variability of Treatment Effect of Complete Clearance Lesion Count Response at Week 12 Across Pools of at Least 12 Subjects	ITT
14.2.1.11	Multiple Imputation Sensitivity Analysis: Complete Clearance Lesion Count Response at Week 12	ITT
14.2.1.12	Patient Reported Response Sensitivity Analysis: Complete Clearance Lesion Count Response at Week 12	ITT
14.2.2.1.1	Complete Clearance Lesion Count of 0 or 1 Response at Week 12	ITT
14.2.2.1.2	Complete Clearance Lesion Count of 0 or 1 Response at Week 12	PP
14.2.2.2.1	Summary of a 90% Reduction in Lesion Counts from Baseline Response at Week 12	ITT
14.2.2.2.2	Summary of a 90% Reduction in Lesion Counts from Baseline Response at Week 12	PP
14.2.2.3	Complete Clearance Lesion Count Response at Week 8	ITT
14.2.2.4.1	Percent Change from Baseline in Lesion Counts by Visit	ITT
14.2.3.1	Complete Clearance Lesion Count of 0 or 1 Response at Weeks 2, 4 and 8	ITT
14.2.3.2	Summary of a 90% Reduction in Lesion Counts from Baseline at Weeks 2, 4 and 8	ITT
14.2.3.3	Summary of a 75% Reduction in Lesion Counts from Baseline Response by Visit	ITT
14.2.3.4.1	Change from Baseline in Lesion Count by Visit	ITT
14.2.3.5	Complete Clearance Lesion Count Response at Weeks 2 and 4	ITT
14.2.3.6.1	Kaplan-Meier Estimates of Time to Complete Clearance (Days) from Start of Dosing	ITT
14.2.3.7	Summary of Recurrence after Complete Clearance	ITT
14.2.3.8	Summary of Increase in Subject-Reported Spread to Household Members by Visit	ITT
14.2.3.9.1	Summary of Global Severity Assessment by Assessor and Visit	ITT

Table Number	Table Title	Analysis Population
14.2.3.10.1	Summary of Global Impression of Change by Assessor and Visit	ITT
14.2.3.10.4	Summary of Efficacy by Subject Global Impression of Change Improvement at Week 12	ITT
14.2.3.11	Analysis of Correlation of Lesion Count with Global Severity Assessment and Global Impression of Change	ITT
14.2.3.12.1	Analysis of Lesion Counts by Dichotomized BOTE Score	ITT
14.2.3.12.2	Summary of Lesion Counts by Highest Post-Baseline Dichotomized BOTE Score	ITT
14.2.3.12.3	Summary of Lesion Counts by Highest Dichotomized BOTE Score	ITT
14.2.3.12.4	Summary of Percent Change from Baseline in Lesion Count over Time by Dichotomized Week 2 BOTE Score	ITT
14.2.3.12.5	Summary of the Beginning of the End (BOTE) Inflammation Score Results by Visit	ITT
14.3	**Safety data**	
14.3.1	**Displays of Adverse Events**	
14.3.1.1.1	Overall Summary of Treatment-Emergent Adverse Events	Safety
14.3.1.1.2	Overall Summary of Post Treatment Adverse Events	Safety
14.3.1.2.1	Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety
14.3.1.2.2	Summary of Post Treatment Adverse Events by System Organ Class and Preferred Term	Safety
14.3.1.3.1	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity	Safety
14.3.1.3.2	Summary of Post Treatment Adverse Events by System Organ Class, Preferred Term, and Maximum Severity	Safety
14.3.1.4	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug	Safety
14.3.1.5.1	Summary of Serious Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety
14.3.1.5.2	Summary of Serious Post Treatment Adverse Events by System Organ Class and Preferred Term	Safety
14.3.1.6	Summary of Treatment-Emergent Adverse Events Leading to Study Drug Discontinuation by System Organ Class and Preferred Term	Safety

Table Number	Table Title	Analysis Population
14.3.1.7	Summary of Treatment-Emergent Adverse Events Leading to Death by System Organ Class and Preferred Term	Safety
14.3.1.8.1	Summary of Treatment-Emergent Adverse Events Leading to Study Discontinuation by System Organ Class and Preferred Term	Safety
14.3.1.8.2	Summary of Post Treatment Adverse Events Leading to Study Discontinuation by System Organ Class and Preferred Term	Safety
14.3.1.9	Summary of Transient Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety
14.3.2	**Listings of deaths, other serious and significant adverse events**	
14.3.2.1	Listing of Serious Adverse Events	Safety
14.3.2.2	Listing of Adverse Events Leading to Study Drug Discontinuation	Safety
14.3.2.3	Listing of Deaths	Safety
14.3.6	**Other safety data**	
14.3.6.1.1	Summary of Local Skin Reaction (LSR) Composite Score by Visit	Safety
14.3.6.1.2	Summary of Local Skin Reaction (LSR) Score Results by Visit	Safety
14.3.6.1.3	Shift Summary of Local Skin Reaction (LSR) Score Results by Visit	Safety
14.3.6.1.4	Summary of Local Skin Reaction (LSR) Composite Score by Dichotomized Baseline BOTE Score and Visit	Safety
14.3.6.1.5	Summary of Local Skin Reaction (LSR) Individual Score by Dichotomized Baseline BOTE Score and Visit	Safety
14.3.6.2	Summary of Events of Special Interest	Safety
14.3.6.3	Summary of Physical Examination Findings	Safety
14.3.6.4	Summary of Scarring and Keloid by Visit	Safety
14.3.6.5	Summary of Patch Testing Results	Safety

Figure Number	Figure Title	Analysis Population
14.2.1.5	Forest Plot: Odds Ratio (95% CI) of the Proportion of Subjects with Complete Clearance at Week 12	ITT
14.2.2.4.2	Percent Change from Baseline in Lesion Count over Time by Treatment Group	ITT
14.2.3.4.2	Mean Change from Baseline in Lesion Count over Time by Treatment Group	ITT
14.2.3.6.2	Kaplan-Meier Plot of Time to Complete Clearance (Days) from Start of Dosing	ITT
14.2.3.9.2	Cumulative Distribution Function Plot Across Subject Global Severity Assessment Responses	ITT
14.2.3.9.3	Cumulative Distribution Function Plot by Subject Global Severity Assessment Severity Category	ITT
14.2.3.10.2	Cumulative Distribution Function Plot Across Subject Global Impression of Change Responses	ITT
14.2.3.10.3	Cumulative Distribution Function Plot by Subject Global Impression of Change Improvement Category	ITT

Section 16.2: **List of Data Listings**

ICH Listing Number	Listing Title	Analysis Population
16.1.7	Subject Randomization	ITT
16.2	**SUBJECT DATA LISTINGS**	
16.2.1	**Discontinued subjects**	
16.2.1.1	Subject disposition	Enrolled
16.2.2	**Protocol deviations**	
16.2.2.1	Protocol Deviations	ITT
16.2.2.2	Inclusion/Exclusion Criteria	Enrolled
16.2.3	**Subjects excluded from the efficacy analysis**	
16.2.3.1	Analysis Populations	Enrolled
16.2.4	**Demographic data**	
16.2.4.1	Demographic and Baseline Characteristics	ITT
16.2.4.2	Patch Testing Consent/Assent	Enrolled
16.2.4.3	Medical History	ITT
16.2.4.4	Prior and Concomitant Medications	ITT
16.2.4.5	Study Exit Interview Consent/Assent	Enrolled
16.2.5	**Compliance and/or drug concentration data**	
16.2.5.1	Study Drug Administration	Safety
16.2.5.2	Study Drug Accountability	Safety
16.2.5.3	Dosing Plan and Treatment Modification	Safety
16.2.5.4	Treatment Compliance	Safety
16.2.6	**Individual efficacy response data**	
16.2.6.1	Lesion Counts	ITT
16.2.6.2	Lesion Count Derived Efficacy Variables and Time to Complete Clearance	ITT
16.2.6.3	Household Transmission	ITT
16.2.6.4	Global Severity Assessment	ITT
16.2.6.5	Global Impression of Change	ITT
16.2.6.6	Patient Reported Outcome of Molluscum Contagiosum Disease Status based on Lost to Follow-up Contact	ITT

ICH Listing Number	Listing Title	Analysis Population
16.2.7	**Adverse events listings**	
16.2.7.1	All Adverse Events	Safety
16.2.7.2	Adverse Events Leading to Study Discontinuation	Safety
16.2.7.3	Ongoing Treatment Related Adverse Events	Safety
16.2.7.4	Post Treatment Adverse Events	Safety
16.2.8	**Listing of individual laboratory measurements by subject, when required by regulatory authorities**	
16.2.8.1	Urine Pregnancy Test	Safety
16.2.9	**Other data**	
16.2.9.1	Beginning-of-the-End (BOTE) Inflammation Score	Safety
16.2.9.2	Local Skin Reactions (LSR)	Safety
16.2.9.3	Physical Examination	Safety
16.2.9.4	Scarring/Keloid Assessment	Safety
16.2.9.5	Patch Testing	Safety
16.2.9.6	Telephone Contact	ITT
16.2.9.7	Subjects who Require a Narrative	ITT
16.2.9.8	COVDI-19 Impacted Visits	ITT