Statistical Analysis Plan

Study code	TRIF-2101
Version	Final
Date	Dec 22, 2022

A MULTI-CENTER, DOUBLE-BLIND, RANDOMIZED, THREE-ARM, PLACEBO-CONTROLLED, PARALLEL-DESIGN GROUP STUDY TO EVALUATE THE THERAPEUTIC EQUIVALENCE COMPARING TRIFAROTENE CREAM, 0.005% (TEVA PHARMACEUTICALS, INC. TO AKLIEF ® (TRIFAROTENE 0.005% CREAM) (GALDERMA LABORATORIES, L.P., USA), IN THE TREATMENT OF ACNE VULGARIS

Study Statistician

[redacted]

Sponsor Representative

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse Event
ANOVA	Analysis of Variance
ATC	Anatomic-Therapeutic-Chemical
BMI	Body Mass Index
CI	Confidence Interval
COVID-19	Coronavirus Disease 2019
CRF	Case Report Form
CRO	Contract Research Organization
FDA	Food and Drug Administration
hCG	Human Chorionic Gonadotropin
HEENT	Head, Eyes, Ears, Nose and Throat
ICF	Informed Consent Form
IGA	Investigator's Global Assessment
IMP	Investigational Medicinal Product
IWRS	Interactive Web Response System
LOCF	Last Observation Carried Forward
LS	Least square
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent to Treat
PP	Per Protocol
PT	Preferred Term
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol "A Multi-Center, Double-Blind, Randomized, Three-Arm, Placebo-Controlled, Parallel-Design Group Study to Evaluate the Therapeutic Equivalence Comparing Trifarotene Cream, 0.005% (Teva Pharmaceuticals, Inc. to Aklief ® (Trifarotene 0.005% Cream) (Galderma Laboratories, L.P., USA), in the Treatment of Acne Vulgaris" version 2.0, dated June 30, 2022.

AKLIEF® (trifarotene) cream, for topical use was approved by the FDA in 2019 as a prescription product for the safe and effective topical therapy for acne vulgaris. Teva Pharmaceuticals, Inc. has developed a generic formulation of trifarotene 0.005% cream. The current study is designed to evaluate the safety and efficacy of this new generic formulation.

This document will give a description of the planned methods of the analysis.

2. OBJECTIVES

The objectives of this study are:

- To evaluate the therapeutic equivalence of TRIFAROTENE CREAM, 0.005% (Teva Pharmaceuticals, Inc.) and the Reference Listed Product, AKLIEF® (TRIFAROTENE 0.005% CREAM) (GALDERMA LABORATORIES, L.P., USA) in the treatment of acne vulgaris.
- To demonstrate the superiority of the efficacy of the test and reference products over that of the placebo control in the treatment of acne vulgaris.

3. STUDY OVERVIEW

3.1 Study Design

Approximately [redacted] will be assigned [redacted] to treatment with the test product, TRIFAROTENE 0.005% CREAM, the reference product, AKLIEF® (TRIFAROTENE 0.005% CREAM, and placebo (vehicle cream of the test product) in this multiple-center, double-blind, randomized, three-arm, placebo controlled, parallel-design study.

The duration of each Subject's participation in the study will be 84 days. Scheduled study visits will include:

- Visit 1 (Day 0) Screening/Baseline
- Visit 2 (First Interim Visit, Day 28)
- Visit 3 (Second Interim Visit, Day 56)
- Visit 4 (End of Treatment/Early Discontinuation Visit, Day 84).

A window period [redacted] will be considered acceptable for each scheduled visit following the screening/baseline visit.

At Visit 1, an informed consent/assent will be obtained from the potential study Subject prior to any study procedures taking place. After the Subject has been consented/assented, the Subject's medical history will then be documented, including the Subject's concomitant medications. A urine pregnancy test with a sensitivity to at least 25 mIU/mL hCG will be performed for female subjects of childbearing potential. A negative result of this test should be obtained.

The Subject will undergo a physical examination, including the recording of vital signs and COVID-19 symptom (such as stuffy or runny nose, sore throat, shortness of breath (difficulty breathing), cough, low energy or tiredness, muscle or body aches, headache, chills or shivering, feeling hot or feverish (temperature check), vomiting, diarrhea, loss of sense of taste and smell) screening. The Subject will be reviewed against the inclusion/exclusion criteria and a screening/baseline acne grade will be assigned to the Subject using the Investigator's Global Assessment (IGA) and a screening/baseline lesion count will be performed.

The Subject will be evaluated for signs and symptoms of local irritation. Blinded Investigational Medicinal Product (IMP) will be dispensed by Independent Dispenser to subjects who meet all of the inclusion and none of exclusion criteria. Eligible subjects will be assigned to one of the three study treatment groups as per the computer-generated randomization schedule, and the assigned study medication will be dispensed by the Independent Dispenser. Since this is a double-blind study, neither the study team other than

Independent Dispenser at the site nor the subjects will know the treatment the subject is assigned.

Subjects will be instructed by an Independent Dispenser on the application of IMP and completion of Subject diaries. The Independent Dispenser will not be involved in the assessments/evaluations of acne lesion counts, Investigator's Global Assessment (IGA), evaluation of signs and symptoms of local irritation and safety assessments.

The Subject will also be instructed not to open the study medication pump while at the study site. The application of Investigational Medicinal Products will be performed by the Subject at home.

Subjects will be instructed to apply enough study investigational drug to lightly cover the entire affected areas of the face for 84 consecutive days. The IMP should be applied once daily at bedtime after the Subject's face has been washed with a non-medicated cleanser provided by the Sponsor, and gently patted dry. The subjects should wait 20 to 30 minutes before applying the study medication to allow the skin to be completely dry in order to minimize possible irritation. For the purposes of this study, the face is considered to start at the hairline and end at the jaw line and excludes the eyes, the lips/mouth, angles/corners of the nose, open wounds and all mucous membranes.

The subject will be instructed to wash their hands before and after application of study drug. Subjects will be required to use diaries to document the date and time of study treatments, any missed treatments, any concomitant medication and the occurrence of all adverse events.

The subjects will be instructed to bring their used IMP and their study diaries to each study visit.

Subjects will return to the study site for Visit 2, Visit 3 and Visit 4.

The subjects are instructed to bring their used IMP and their study diaries to each study visit.

Compliance with drug applications will be assessed at each visit. Additional pumps of IMP will be dispensed at the visit if required, [redacted]

[redacted] In addition, all IMP and diaries will be collected from the Subject during each scheduled visit or the Early Discontinuation Visit.

At every study visit, all the procedures mentioned in the Table of events (section 11.1) will be carried out. The treatment area of enrolled subjects will be assessed using IGA and lesion count. Additionally, adverse events (AEs) will be monitored, and the treatment area will be examined to assess application site reactions. Any change in concomitant medications will be noted.

3.2 Sample Size

[redacted]

The following assumptions were used in computer simulations to determine the sample size.

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

ratio of mean percent reduction from baseline, for both lesion types, are contained within the interval [0.80, 1.25] for the PP population.

3.3 Randomization and Unblinding Procedures

This is a Randomized, Double-blind, Three-Arm, Placebo Controlled, Parallel design, multicenter study. Subjects will be randomly assigned [redacted] to receive the test product, the reference product and the vehicle (placebo) control. The randomization assignment will be generated using SAS® by an independent statistician (SyMetric), not involved with the study. The treatment each subject will receive will not be disclosed to the Investigators, study center personnel, subject, sponsor, or CRO involved in the study conduct, monitoring, data review or analysis.

SyMetric will hold the randomization code throughout the conduct of the study in order to minimize bias. If necessary, for the safety and proper treatment of the subjects, the Investigator can unblind the subject's treatment assignment to conduct appropriate follow-up care. Whenever possible, the Sponsor or medical monitor will be notified before unblinding the study medication. The date and signature of the person breaking the code as well as the reason for breaking the code and any associated AEs will be recorded in the subject's source documentation.

The Investigator, staff at the study site, study monitors, and data analysis/management personnel will be blinded to the treatment assigned to the subject, with the exception of the Independent Dispenser. The study site will have at least one Independent Dispenser. The role of the unblinded Independent Dispenser will be to dispense and collect study medication to/from the subjects, maintain dispensing records, and ensure the study medication logs are complete and accurate. The subject will be requested not to discuss the appearance of the study medication with the Investigator or study staff outside of the Independent Dispenser. Independent dispenser should not participate in any activities other than the activities defined in this section.

In case of an emergency, if the details of the study drug are required for management of an emergency as per the opinion of Investigator, the Investigator can unblind the product that is

received by the subject during the study. In case of non-emergency condition that requires study drug information for management of condition as per Investigator's opinion, the Investigator should obtain Sponsor or medical monitor approval in writing prior to breaking of the blind. [redacted] It is recommended that all attempts should be made to maintain the study blind. However, in case that unblinding is performed, the reason for breaking the blind must be clearly documented in the source documentation and Case Report Form (CRF) and the subject must be discontinued from the study.

The date on which the code was broken together with the identity of the person responsible must also be documented.

The treatment assignments will remain blinded until the final database is locked.

4. STUDY ENDPOINTS/OUTCOMES

The co-primary endpoints of this study are:

1. Mean percent change from screening/baseline to week 12 in the inflammatory (papules and pustules) lesion counts;

2. Mean percent change from screening/baseline to week 12 in the non-inflammatory (open and closed comedones) lesion counts.

5. HYPOTHESES TESTING

Hypothesis of Equivalence

A two-sided, 90% confidence intervals on the test/reference ratio for mean percent change from Baseline in the inflammatory and non-inflammatory lesion counts will be constructed using Fieller's method. The estimates of treatment means and standard errors will be obtained from a two-way Analysis of Variance (ANOVA) of the Test and Reference results, using a statistical model containing terms for treatment and center. Non-parametric methods will be considered if the skewness factor for the residuals from the ANOVA model is outside the range -2 to +2.

Bioequivalence will be established if the 90% confidence intervals for the ratios of test/reference means for both the inflammatory and non-inflammatory lesions are contained within the interval [0.80, 1.25].

Hypothesis of Superiority

The null hypothesis to be tested is that there is no difference in the mean percent change (reduction) from baseline in the inflammatory and non-inflammatory lesion counts between the active treatment and the Vehicle treatment. The hypothesis testing will be performed separately for the Test treatment versus the Vehicle treatment and for the Reference treatment versus the Vehicle treatment using ANOVA under assumption of normal error and homogeneity of variance. Only data for the relevant two treatment arms will be included in each ANOVA. Similarly to bioequivalence, non-parametric methods will be considered if the skewness factor for the residuals from the ANOVA model is outside the range -2 to +2.

Superiority will be established if the mean percent change (reduction) from Baseline in both the inflammatory and non-inflammatory lesion count for each active treatment is greater than, and statistically different from ($p < 0.05$, two-sided), that for the Vehicle.

6. ANALYSIS SUBSETS

6.1 Safety Population

The safety population will include all randomized subjects who use at least one dose of product (as recorded at the "Investigational Medicinal Product Use Compliance" CRF based on the subject diary).

The Safety population will be the primary population for the safety analysis. Subjects will be analyzed according to the treatment they actually received.

6.2 Modified Intent to Treat Population (mITT Population)

[redacted]

This population will be considered as supportive for testing the clinical equivalence and as definitive while testing the superiority. Subjects will be analyzed according to the treatment they were randomized to receive.

6.3 Per Protocol Population (PP Population)

[redacted]

Subjects who are discontinued early from the study due to lack of treatment [redacted] will be included in the PP population as treatment failures (i.e., non-responders) even if they do not have Week 12 evaluation within the visit window or at all. A "last observation carried forward" (LOCF) approach will be used for imputing missing lesion counts in these PP subjects.

Per protocol, subjects whose condition or symptoms worsens and require alternate or supplemental or rescue therapy for the treatment of their condition during the treatment phase of the study should be discontinued, included in the mITT and PP population analysis using LOCF, and provided with effective treatment. The reason for study discontinuation for these subjects should be set to "Lack of treatment effect", therefore for practical purposes this group of subjects is a subgroup of subjects who are discontinued early from the study due to lack of treatment effect.

[redacted]

The PP population will be considered as definitive for testing the clinical equivalence and as supportive while testing the superiority. Subjects will be analyzed according to the treatment they were randomized to receive.

7. STATISTICAL METHODS OF ANALYSIS

7.1 General Principles

[redacted]

All tables, figures, and listings will be produced in the landscape format.

In general, all data will be listed by treatment group, subject and visit where appropriate. The summary tables will also be stratified by, or have columns corresponding to, treatment group.

The total number of subjects in the treatment group (N) under the stated population will be displayed in the header of summary tables.

Data will be summarized using descriptive statistics for continuous variables. Unless otherwise specified, descriptive statistics will include number of subjects, mean, standard deviation, minimum, median, and maximum. The minimum and maximum statistics will be presented to the same number of decimal places as the original data. The mean and median will be presented to one more decimal place than the original data. The standard deviation will be presented to two more decimal places than the original data. The number of missing observations will be presented only if non-zero.

In summary tables of categorical variables counts and percentages will be used. The count [n] indicates the actual number of subjects in a particular category, which should always be less than or equal to the total number of subjects in the respective treatment group/population with known (non-missing) category [N]. Percentage will be obtained by: % = n/N*100. Unless otherwise specified, all percentages will be expressed to one decimal place.

All statistical tests will be two-sided at a significance level of $\alpha = 0.05$, unless otherwise specified. No adjustment will be made for multiplicity.

Relative days will be calculated relative to the date of the first dose of study medication. Relative days will be calculated as follows only when the full assessment date is known (i.e., partial dates will have missing relative days).

For assessment on or after the day of the first dose of study drug:

Relative Day = Date of Assessment – Date of the First Dose of study Drug+1.

For assessment before the day of the first dose of study drug:

Relative Day = Date of Assessment – Date of the First Dose of study Drug.

All dates will be displayed in DDMMMYYYY format.

7.2 Subject Disposition

The number of subjects enrolled in the study, randomized to treatment, included in the Safety, mITT and PP populations, prematurely discontinued from the study (along with the reasons for discontinuation) will be calculated. The percentages will be based on the number of subjects randomized to each treatment group. Percentages for discontinuation reasons will be based on the sub-population of subjects who discontinued from the study.

Number and percentage of subjects enrolled by site will be tabulated for all enrolled subjects, Safety, mITT, and PP populations.

Number and percentage of subjects excluded from each of the study populations will be summarized by reason for exclusion.

Details of subject eligibility, informed consent and disposition will be listed.

7.3 Demographic and Baseline Characteristics

Demographic characteristics will include:

- age;
- gender;
- race;
- ethnicity;

Baseline characteristics include:

- Time since acne vulgaris diagnosis (years)

- Baseline inflammatory lesion count (i.e papules and pustules)
- Baseline non-inflammatory lesion count (i.e. open and closed comedones)
- Baseline nodulocystic count (i.e. nodules and cysts)
- Baseline IGA score
- Baseline height, weight and body mass index (BMI)

Descriptive statistics will be presented for age (years), height, weight, BMI, time since diagnosis (years) and lesion counts. Frequency counts and percentages will be presented for race, ethnicity and IGA. Height will be reported in centimeters and weight in kilograms.

Age will be derived from Informed Consent Signed Date and Date of Birth as the number of whole years between those two dates.

Demographic and baseline characteristics will be evaluated for comparability across treatment groups in the following manner. Continuous variables will be analyzed with an ANOVA with factors of treatment and investigational site. Overall p-value for the global null hypothesis of all groups being equal will be displayed. Categorical variables will be analyzed with a Cochran-Mantel-Haenszel stratified by investigational site. For IGA score "row mean scores differ" CMH statistic will be used, while for all other parameters "general association" CMH statistic will be used.

These analyses will be performed for Safety, mITT, and PP populations.

All demographic and baseline characteristics will be listed.

7.4 Medical History

Medical history (other than acne history) will be summarized by Medical Dictionary for the Drug Regulatory Activities (MedDRA) System Organ Class and Preferred Term for the Safety population. One subject will be counted once for each applicable Preferred Term and System Organ Class. MedDRA version 25.0 will be used.

Medical history data will also be listed.

7.5 Protocol Deviations

Protocol deviations will be recorded in the eCRF during the study and identified programmatically. The deviations that are identified programmatically and the algorithms for their identification are defined in a separate Programmed Protocol Deviations Specification.

[redacted]

All major protocol deviations will be summarized by deviation category and treatment group. This analysis will be performed for the mITT population.

All protocol deviations will be listed.

7.6 Efficacy Analyses

7.6.1 Center pooling

To eliminate potential effect of random fluctuations at small site on the primary endpoints small centers will be pooled. [redacted]

- [redacted]
- [redacted]

[redacted]

Pooled centers will be used in all ANOVA models, including the efficacy analyses and the comparison of the treatment group in the summary of the demographic and baseline characteristics.

7.6.2 Analyses of Primary Endpoint

At each study visit, beginning at Visit 1, an Investigator will count the number of open and closed comedones (non-inflammatory lesions), pustules and papules (inflammatory lesions) and nodulo-cystic lesions. Lesion counts will be recorded in the CRF.

Missing Week 12 assessments will be imputed as follows. In the mITT analysis, if a subject discontinues the study prior to Week 12 or misses the Week 12 assessment for any other reason, the LOCF rule will be used to impute the number of lesions. In the PP analysis, if a subject discontinues the study prior to Week 12 due to lack of treatment effect [redacted], the LOCF rule will be used to impute the number of lesions. If a subject is missing the Week 12 assessment for any other reason, the subject will be excluded from the PP population.

The LOCF rule will be applied as follows: the last available post-baseline lesion count of the given type will be carried forward to impute the missing Week 12 assessment of the same type. Subjects who do not have any post-baseline lesion counts of a certain type will be excluded from the analysis of that type.

The baseline lesion count will be defined as the results from the latest examination prior to the first application of the study medication. For each post-baseline visit the change from baseline and percent change from baseline in the number of inflammatory and non-inflammatory lesions will be calculated. The baseline value, visit value and percent change from baseline at the visit will be analyzed using descriptive statistics and will be tabulated by visit and treatment group.

7.6.2.1 Analysis of clinical equivalence of test and reference treatments

To show the clinical equivalence, estimates of mean percent change from baseline in the inflammatory and non-inflammatory lesion counts will be calculated for the Test and Reference treatment, and then the 90% CI for the mean ratio will be constructed using Fieller's method. Bioequivalence will be established if the 90% confidence intervals for the

ratios of Test/Reference means are contained within the interval [0.80, 1.25] for the PP population for both lesion types (inflammatory and non-inflammatory).

To this end, first an ANOVA model will be fit (separately for each lesion type) with percent change from baseline in lesion count as outcome and treatment, center and treatment-by-center interaction as factors on the data from Test and Reference treatments only (excluding Vehicle subjects). If the treatment-by-center interaction factor is not significant at the 0.05 level, the model will be rerun without the interaction term. Treatment means and standard errors will be estimated from this model. Then Fieller's formula will be applied; covariance between the treatment means will be assumed to be 0. See Appendix 11.2 for complete description of the Fieller's formula.

A non-parametric rank based ANOVA will be considered when the data is highly skewed. The evaluation of skewness, (using SAS® PROC UNIVARIATE) will be performed using the residuals from ANOVA and if the skewness statistic is less than -2 or greater than +2, the analysis will be performed on the ranks of the percent change in the inflammatory lesion count values.

Analysis of bioequivalence will be performed on the PP population

7.6.2.2 Analysis of superiority to vehicle control

The analysis of superiority will be performed separately for the Test treatment versus the Vehicle treatment and for the Reference treatment versus the Vehicle treatment and separately for the inflammatory and non-inflammatory lesions. Each of these analyses will be performed using an ANOVA model with percent change from baseline in the lesion count as outcome and treatment, center and treatment-by-center interaction as factors. If the treatment-by-center interaction factor is not significant at the 0.05 level, the model will be rerun without the interaction term. The model will be fit on data from the Test and Vehicle treatment for analysis of superiority of the Test treatment and separately on data from the Reference and Vehicle treatment for analysis of superiority of the Reference treatment. From this model, the least square (LS) mean estimate for each treatment group with the 95% CI will be calculated; further, an estimate of the LS mean difference between the active treatment (Test or

Reference) and Vehicle with 95% CI and the p-value for test of no difference will be calculated.

Superiority will be established if the mean percent change from Baseline (reduction) for each active treatment and each lesion type is estimated to be greater than, and statistically significantly different from ($p < 0.05$ for test of no difference) that for the Vehicle, for the mITT population.

Analysis of superiority will be performed on the mITT population.

A non-parametric rank-based ANOVA will be considered under the same circumstances as for the analysis of bioequivalence.

7.6.3 Analyses of IGA

At each study visit, beginning at Visit 1, an Investigator will assess the overall status of the subject's face for acne vulgaris using the IGA on a scale from 0=Clear to 4=Severe. To be included in the study, subjects must have a definite clinical diagnosis of acne vulgaris of severity grade 2, 3, or 4 at screening/baseline.

Number and percentage of subjects at each IGA severity grade will be tabulated by visit. In addition, a shift table for changes in IGA grades from Baseline to each post-baseline visit will be created. This analysis will be performed for both mITT and PP populations.

7.7 Safety Analyses

7.7.1 Adverse Events

Adverse Events will be coded using the MedDRA version 25.0 AE coding system for purposes of summarization.

Only treatment-emergent adverse events (TEAEs) will be used for the summary analysis. An AE will be considered as treatment-emergent if the date of onset is on or after the date of the first study medication administration or if it increased in severity during the study treatment period. AEs with unknown start dates will be counted as treatment-emergent unless the AE resolution date is prior to the study drug start date. If the AE start date is partially missing, the

AE will be considered treatment-emergent, unless the month and year (when available) rule out the possibility that the event occurred post start of the study drug.

AE will be considered treatment-related if its relationship to study product is recorded as Possible, Probable or Definitely related. AEs with missing relationship will also be considered related.

An overall summary will include, by treatment group and overall, the number of TEAEs and the number and percentage of subjects reporting at least 1 TEAE in the following categories:

- Any AE
- Any TEAE,
- Treatment-related TEAE,
- Serious TEAE,
- TEAE leading to discontinuation of the study medication.
- Fatal AE

Count of AEs by severity and relationship to the study medication will also be presented.

The following TEAE frequency tables will be prepared summarizing the number and percentage of subjects reporting at least one TEAE by MedDRA system organ class (SOC) and preferred term (PT):

- All TEAEs
- Treatment-related TEAEs
- Serious TEAEs
- TEAE leading to discontinuation of the study medication
- TEAEs by maximum severity

A subject experiencing the same AE multiple times will only be counted once for that preferred term. Similarly, if a subject experiences multiple AEs within the same system organ class, that subject will be counted only once in that system organ class. In the summary by severity only the highest applicable severity will be chosen for each preferred term and system organ class.

Summaries of AEs will be presented in alphabetical order of SOC and preferred terms. Additionally, TEAEs will be summarized by the preferred terms in the descending order of frequency in the total treatment group. In this table a p-value from Fisher's exact test comparing event rates between the Test and the Reference treatment groups will be provided for those preferred terms that have frequency > 1% in either Test or Reference group.

All information pertaining to adverse events noted during the study will be listed by subject, detailing verbatim, preferred term, system organ class, start date, stop date, severity, outcome, action taken, and causal relationship to the study drug. The adverse event onset will also be shown relative (in number of days) to the date of first administration of the study medication. In addition, the adverse event duration (if AE Stop Date is available) will be evaluated as below and presented (in number of days).

$$\text{AE Duration} = \text{AE Stop Date} - \text{AE Start Date} + 1$$

7.7.2 Vital Signs

Vital signs, including blood pressure, pulse rate, respiratory rate and body temperature will be documented at Visit 1, Visit 2, Visit 3 and Visit 4. Vital signs will be measured after the subject has rested in a seated position for about 5 minutes. Weight will be measured at Visits 1 and 4 as part of the physical examination.

Actual values and changes from baseline of each vital sign (including weight) will be summarized by visit and treatment group using descriptive statistics. All results will be listed.

7.7.3 Physical Examination

A brief physical examination will be performed at screening/baseline, Visit 4 End of Treatment/Early Discontinuation and during unscheduled visits where applicable. The physical examination will include, at a minimum, examination of the Subject's general appearance, skin, HEENT (head, eyes, ears, nose and throat), heart, lungs, musculoskeletal system, neurological system, lymph nodes, abdomen and extremities. Each body system will be classified as normal, abnormal not clinically significant or abnormal clinically significant.

Number and percentage of subjects with each assessment result will be summarized by body system, visit and treatment group. All results will be listed.

7.7.4 Signs/Symptoms of Local Irritation

At each study visit, beginning with Visit 1, subjects will be evaluated for any signs and symptoms of local irritation (application site reactions), including erythema, dryness, burning/stinging, erosion, edema, pain and itching. Each subject will be assigned a severity score by an Investigator based on the scale ranging from 0=Absent to 3=Severe.

Number and percentage of subjects with each severity will be presented by visit and by sign/symptom. Additionally, the shifts from baseline in irritation score will be tabulated by treatment, symptoms and scheduled post-baseline visits. The denominator for the proportions will be the number of subjects with any assessment of the sign/symptom both at baseline and at given visit.

7.7.5 Exposure to Product

The subjects will be instructed to use the diary to document all treatments administered, including the date and time and all treatments missed.

[redacted]

Number of missed doses will be calculated as follows. The number of doses missed in the dosing period (between the first and the last study drug application) will be taken directly from the CRF. [redacted] The compliance will be analyzed using

the descriptive statistics by treatment group. The proportion of compliant vs. non-compliant subjects will be tabulated for each treatment.

Duration of exposure will be calculated as Date of last use of study medication – Date of first use of study medication + 1. Duration of exposure will be summarized descriptively by treatment group.

Compliance and duration of exposure will be summarized for the Safety and mITT populations.

7.7.6 Exposure to Concomitant Medication

Concomitant medications and therapeutic treatments, including the use of sunscreen, OTC products, herbal products, topical and systemic medicines, dietary supplements and any current non-drug treatments/therapies will be assessed at the screening/baseline visit and at each subsequent study visit.

Medication or non-drug therapy will be classified as prior if the end date is known and is prior to the first use of the study medication. Medications and non-drug therapies that are ongoing or ended after the first use of the study medication will be classified as concomitant. If the end date of the medication or non-drug therapy is partially known so that there is an ambiguity whether it ended before or after the start of the study medication, the medication will be considered concomitant. If the end date is unknown, it will also be considered concomitant.

The concomitant medications will be summarized by treatment group, ATC class level 3 and preferred term as coded by the World Health Organization Drug dictionary March-2022 B3 and be listed in a by-subject listing. Prior medications will be listed only.

7.7.7 COVID-19 Screening

Temperature check and COVID-19 symptom (such as Stuffy or runny nose, Sore throat, Shortness of breath (difficulty breathing), Cough, Low energy or tiredness, Muscle or body aches, Headache, Chills or shivering, feeling hot or feverish (temperature check), vomiting, diarrhea, loss of sense of taste and smell) screening will be done for each subject at visit 1, 2, 3, 4 and Unscheduled / Early Discontinuation visit. During the screening, if the subject found to have any one of the above said symptoms, then it will be judged by the investigator whether

it is related or not-related to COVID-19 and appropriate procedures will be followed by the investigator and site. Results of the COVID-19 screening will be listed.

8. CHANGES FROM PROTOCOL-SPECIFIED ANALYSES

The definition of the mITT population was expanded in the SAP.

Otherwise, there are no changes from the protocol-specified analyses.

9. LIST OF PLANNED TABLES, FIGURES, AND LISTINGS

See separate document with the table, figure and listing shells.

10. LITERATURE CITATIONS / REFERENCES

1. Study Protocol: A Multi-Center, Double-Blind, Randomized, Three-Arm, Placebo-Controlled, Parallel-Design Group Study to Evaluate the Therapeutic Equivalence Comparing Trifarotene Cream, 0.005% (Teva Pharmaceuticals, Inc. to Aklief ® (Trifarotene 0.005% Cream) (Galderma Laboratories, L.P., USA), in the Treatment of Acne Vulgaris.

11. APPENDICES

11.1 Study visit Schedule

Visit Number	Visit 1	Visit 2	Visit 3	Visit 4	Unscheduled/ Early Termination
Visit Day	Day 0	Day 28[redacted]	Day 56[redacted]	Day 84[redacted]	
Visit Name	Baseline	Interim	Interim	End of Treatment	
COVID-19 symptom screening	X	X	X	X	X
Informed Consent	X				
Inclusion/Exclusion Criteria	X				
Demographics	X				
Medical History	X				
Physical Exam	X			X	X as needed
Vital Signs	X	X[3]	X[3]	X[3]	X[3]
Urine Pregnancy Test [1]	X	X	X	X	X
Clinical Assessment of Acne: Investigator's Global Assessment (IGA)	X	X	X	X	X
Investigator's Lesion Counts Inflammatory and Non-Inflammatory	X	X	X	X	X
Randomization	X				
Dispense IMP and Supplies	X	X[3]	X[3]		
Dispense Diary/Instructions	X	X	X		
IMP Review and Collection[4]		X	X	X	X as needed
Subject Compliance/ Diary Review and collection		X	X	X	X as needed
Evaluation of Signs/Symptoms of Local Irritation (Application Site Reactions)	X	X	X	X	X
Adverse Event Assessment[2]	X	X	X	X	X
Concomitant Medication Review	X	X	X	X	X

1 Tubal ligation is not considered equivalent to female sterilization. Women with a history of tubal ligation are still considered females of childbearing potential and must have a negative urine pregnancy test with a sensitivity of at least 25 mIU/ml hCG.

2 Any AEs reported after signing Informed Consent should be reported.

3 Only Temperature check at visit 2, 3, 4 and Unscheduled / Early Discontinuation as part of COVID-19 symptom screening.

4 Additional IMP and supplies will be provided to the Subject as required.

11.2 Code Fragments

ANOVA model for superiority analysis in the primary endpoint

```
proc glm data=<datasets name>;
   by lesion_type;
   class <treatment> <center>;
   model <Percent change from baseline in lesion count at week 12> = <treatment> <center>
   <treatment>*< center>/ ss3;
   lsmeans <treatment> / pdiff cl;
   output out=residuals residual=residual;
run;
quit;
```

Note: this analysis needs to be performed separately for test and reference treatments on a dataset containing only test and vehicle or reference and vehicle treatment subjects. If treatment-by-center interaction term is not significant at 0.05 level, the model will be rerun without this term.

Analysis of clinical equivalence in the primary endpoint

ANOVA model:

```
proc glm data=<datasets name>;
   by lesion_type;
   class <treatment> <center>;
```

model <Percent change from baseline in lesion count at week 12> = <treatment>

<treatment>* / ss3;

lsmeans <treatment> / stderr;

output out=residuals residual=residual;

run;

quit;

Here dataset contains test and reference treatment subjects only.

Note: if the treatment-by-center interaction term is not significant at the 0.05 level, the model will be rerun without this term.

Fieller's method.

Generally, Fieller's formula allows to calculate the confidence interval for the ratio of two (possibly correlated) means of two samples a and b with expectations μ_a and μ_b, and variances $\nu_{11}\sigma^2$ and $\nu_{22}\sigma^2$ and covariance $\nu_{12}\sigma^2$. If $\nu_{11}, \nu_{12}, \nu_{22}$ are all known, then a $(1-2\alpha)$ confidence interval (mL, mU) for μ_a/μ_b is given by

$$(m_L, m_U) = \frac{1}{(1-g)}\left[\frac{a}{b} - \frac{g\nu_{12}}{\nu_{22}} \mp \frac{t_{r,\alpha}s}{b}\sqrt{\nu_{11} - 2\frac{a}{b}\nu_{12} + \frac{a^2}{b^2}\nu_{22} - g\left(\nu_{11} - \frac{\nu_{12}^2}{\nu_{22}}\right)}\right]$$

where

$$g = \frac{t_{r,\alpha}^2 s^2 \nu_{22}}{b^2}.$$

Here s^2 is an unbiased estimator of σ^2 based on r degrees of freedom, and $t_{r,\alpha}$ is the α-level deviate from the Student's t-distribution based on r degrees of freedom.

In the case of this study the two samples are independent (different subjects), thus covariance can be assumed to be zero, and the formula simplifies to:

$$(m_L, m_U) = \frac{1}{1-g}\left[\frac{m_t}{m_r} \mp \frac{t_{r,\alpha}}{m_r}\sqrt{se_t^2 + \frac{m_t^2}{m_r^2}se_r^2 - g \cdot se_t^2}\right]$$

where

$$g = \frac{t_{r,\alpha}^2 se_r^2}{m_r^2}$$

Here m_t and se_t are mean and standard error estimate for test treatment and m_r and se_r for reference treatment correspondingly obtained from the above model. The degrees of freedom r can be obtained from the model as degrees of freedom for the error term in the overall ANOVA table, $\alpha = 0.05$ (for the 90% confidence interval). The term $t_{r,\alpha}$ can be calculated in SAS as tinv(α, r).

Evaluation of skewness

This evaluation will be performed on the residuals output by the "output" statement in PROC GLM above:

```
proc univariate data=residuals;
   var residual;
run;
```

Non-parameteric rank ANOVA

If decision to use the rank ANOVA is taken, the data will be ranked first:

```
proc rank data=<dataset>
   var <change_from_baseline>;
   ranks rank;
run;
```

After that the analysis will proceed using the same SAS code as described above but using rank variable instead of the original change from baseline.