Statistical Analysis Plan

Sponsor:	Cutanea Life Sciences, Inc.
Protocol No:	CLS006-CO-PR002
[redacted] **Project Id:**	[redacted]
Seizure frequency	
Version No.:	1.0
Title:	A Phase 3, Randomized, Double-Blind, Vehicle-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of CLS006 versus Vehicle in Subjects 2 years of age or older with Cutaneous Common Warts
CRF Version No./Date:	Version 4.0 /25-Oct-2017
SAP No.	1.0

1.0 Approvals

Sponsor	
Sponsor Name:	Cutanea Life Sciences, Inc.
Representative/ Title:	[redacted]
Signature /Date:	
[redacted]	
Project Manager/Title:	[redacted]
Signature /Date:	
Biostatistician / Title (Owner):	[redacted]
Signature /Date:	

**If using eTMF3, this section should be removed from the document as document approvals are recorded electronically through the eApproval workflow.*

(NOTE: Electronic Signatures should only be used if all parties have the ability to eSign.)

Table of Contents

2.0 Purpose

The statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Cutanea Life Sciences, Inc. Protocol CLS006-CO-PR-002.

3.0 Scope

This plan is a living document that will be created during the trial start-up. The Statistical Analysis Plan will be finalized prior to database lock. SAP will require sign off from the Project Manager and the sponsor.

The Statistical Analysis Plan outlines the following:

- Study objectives
- Study design
- Variables analyzed and analysis sets
- Applicable study definitions
- Statistical methods regarding important protocol deviations, study drug exposure, efficacy analysis, concomitant medications, adverse events handling, laboratory data and physical examinations

4.0 Introduction

This statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Cutaneous Life Sciences, Inc. Protocol CLS006-CO-PR-002.

This SAP should be read in conjunction with the study protocol and case report form (CRF). This version of the plan has been developed using the protocol dated 23May2017 and CRF dated 02Aug2017. Any further changes to the protocol or CRF may necessitate updates to the SAP.

4.1 Changes from Protocol

No changes as of this version of the SAP.

5.0 Study Objectives

The objective of this study is to evaluate the efficacy and safety of six (6) weeks of once daily application of Furosemide Topical Gel 0.125% (CLS006) compared to vehicle in subjects ≥2 years of age with nongenital cutaneous common warts (verruca vulgaris).

6.0 Study Design

The study will be conducted as a randomized, double-blind, vehicle-controlled, parallel group, multicenter study at approximately 40 sites involving approximately 480 subjects (~240 per treatment group) with nongenital cutaneous common warts. Approximately 160 subjects enrolled will be pediatric subjects (~80 will be ≥ 2 to <12 years of age and ~80 will be ≥12 to <18 years of age). Subjects may be in the study up to 34 weeks.

Table 1 Schedule of Visits and Study Assessments

Study Period	Screening	Double-Blind Treatment (DBT)				Post-treatment Efficacy Evaluation (PTEE)		Follow-up[1]	
Visit	Screening	Baseline	Week 2	Week 4	Week 6 / End of Treatment (or Early Term during DBT Period)	Week 12	Week 18 / End of Post-tx Efficacy Evaluation (or Early Term during PTEE Period)	Follow-up #1	Follow-up #2 (Final)
Time	Day -28 to -1	Day 1	Day 14	Day 28	Day 42	Day 84	Day 126 (3 mon post-treatment)	Week 24	Week 30 (6 mo. post-treatment)
Procedures (Visit Window)			(± 3 days)			(± 5 days)		(± 7 days)	
Informed Consent	X								
Demographics & Medical/Surgical History	X	X[2]							
Prior & Concomitant Medication/Treatment	X								
Limited Physical Exam[3]	X				X				
Vital Signs (BP, HR)	X	X	X	X	X	X	X		
ECG	X								
Urine Pregnancy Test[4]	X	X			X				
Clinical Safety Lab Testing (Hematology, Chemistry)	X				X				
Wart Assessments	X	X	X[5]	X[5]	X[5]	X[5]	X[5]	X[5]	X[5]
Clinical Photography (select sites only)[6]		X			x		x		
Inclusion/Exclusion Criteria	X	X							
Enroll/Randomization		X							
Dispense Study Drug & Diary		X	X	X					
Treatment Compliance / Collect and Review Study Drug & Diary			X	X	X				
Apply Study Drug (Once Daily)		X[7]							
Concomitant Medication/Treatment Review		X	X	X	X	X	X	X[8]	X[8]
Adverse Events		X[2]	X	X	X	X	X		

1 Only required for subjects who achieved clearance of at least one treated wart as determined by the PI or an appropriately qualified Sub-I at week 18 / End of Post-tx Efficacy Evaluation.

2 Adverse event assessment will be performed following the first study drug application (any changes in subject's health before first study drug application should be recorded as medical history).

3 A limited physical exam (refer to Section 7.5.2 of the protocol) including height & weight measurements at Screening and weight

measurement only at the Week 6/EOT (or Early Termination during DBT Visit).

4 Urine pregnancy test for women of childbearing potential only. Urine pregnancy testing should also be conducted at any visit for any female who is suspected of being pregnant.

5 New common warts that appear after the Baseline visit will be identified, characterized and location recorded/mapped **only**, but not treated.

6 Clinical Photography will be optional for subjects < 12 years of age.

7 The first application of study drug will be performed under supervision at the study center.

8 Concomitant medication/treatment review at Follow-up visit will focus on review of restricted concomitant medication/treatment only (refer to Section 7.6.10.1 of the protocol).

6.1 Sample Size Considerations

The sample size estimates for the current study are based upon the rates of clearance of warts in the previous Phase 2 study [redacted]. This was a randomized, double blind, vehicle controlled, Phase II study in 80 subjects who were treated with topical [redacted], furosemide, [redacted], or vehicle for 42 days daily with an 8-week follow-up period. For the furosemide treatment group, the proportion of subjects with all cleared warts was 0.1 (10%) versus 0 in the vehicle.

The difference in the proportion of subjects with all treated warts cleared (complete clearance) at the Week 18/End of Post-Treatment Efficacy Evaluation period (EOS) is the primary endpoint of the current study. For conservative purposes, assuming 10% and 2% complete clearance rates in the furosemide and vehicle treated groups, respectively, 200 subjects randomized per treatment group will provide for 92% power at a two-sided α of 0.05 to detect a difference between active and vehicle treated groups.

With the inclusion of an additional 80 subjects in the ≥ 2 to < 12 cohort, the total sample size of 480 will provide 96% power with the assumption unchanged.

6.2 Randomization

Subjects that meet all inclusion and exclusion criteria will be randomized to one of two treatment groups in a 1:1 ratio for each study center according to a predetermined computer-generated randomization code.

Subjects will be randomized to treatment using an Interactive Web Response System (IWRS). The randomization scheme will include age stratum at 3 levels (≥ 2 to <12, ≥ 12 to <18 or ≥ 18) and investigative site.

7.0 Study Variables and Covariates

7.1 Primary Efficacy Variable

Proportion of subjects with complete clearance of all treated warts (i.e., Baseline Treated warts) between the active and vehicle at Week 18 /End of Post-treatment Efficacy Evaluation.

7.2 Secondary Efficacy Variables

- The ratio of cleared warts to all treated warts for each subject at Week 18/End of Post-treatment Efficacy Evaluation.
- Proportion of subjects with complete clearance of all treated warts at Week 12, then Week 6/EOT.
- The ratio of cleared warts for each subject to all treated warts at Week 12, then Week 6/EOT.
- Reduction from baseline in wart size of treated warts at Week 18/End of Post-treatment Efficacy Evaluation

7.3 Tertiary Efficacy Variables

- All endpoints at all other time-points.
- Evaluation of durability of response and incidence of recurrence of treated warts from Week 18 end of post-treatment efficacy evaluation period to Week 30 the end of the follow-up period.

7.4 Other Variables

- New common warts that appear after Baseline by visit

7.5 Safety Variables

Safety variables including AEs, vital signs, and laboratory parameters will be collected according to the Schedule of Visits and Study Assessments (refer to Table 1).

7.6 Predetermined Covariates and Prognostic Factors

Subject age ($\geq$2 - < 12, $\geq$12 - < 18, and $\geq$18) and investigator site will be used as covariates for Cochran-Mantel-Haenszel (CMH) analyses.

Baseline Treatment Wart size (i.e., diameter) will also be included as a covariate in the analysis of covariance (ANCOVA).

8.0 Definitions

Baseline and Change from Baseline

Baseline is defined as the last available pre-dose value.

Change from Baseline = (post-Baseline value – Baseline value). For the purpose of tabulations, the unscheduled post-Baseline values generally will be excluded.

Completion of study

The completion of study is defined as an answer of "Yes" to the question "Did the subject complete the study at Week 18?" on the form of Study Completion in the CRF.

Prior and Concomitant Medications

Prior medications will be any medication with a start date before the first day of treatment.

Concomitant medications are defined as any medications ongoing at the start of treatment or with a start date on or after the first dose date. Consequently, medications ongoing at the start of treatment will be reported as both prior and concomitant medications.

Treatment-Emergent Adverse Events

Treatment emergent AEs (TEAEs) are defined as AEs whose onset occurs, severity worsens or intensity increases after receiving the study medication. Events which occur more than thirty days after the last dose of study medication will not be considered treatment emergent. Adverse Events will be considered treatment-emergent if they have missing or have partial start dates for which it cannot be determined whether the AE started before or after the first dose of study medication.

Adverse Events with Outcome of Death

Any AE with an outcome of fatal will be considered as AE with an outcome of death.

Imputation of AE (for determination of TEAE only) and concomitant medication start and stop dates:

Start Date: If only 'day' is missing, and the month and year are not the same as the month of first dose, then impute day with '01'. Otherwise, if the month and year are the same as the first dose date, use the first dose date. If 'day' and 'month' are missing, and 'year' is not missing, then impute month and day with month and day of the first dose date (assuming same 'year'). If 'day' and 'month' are missing and 'year' is not missing and is not the same year as first dose date, then impute with '01' for both 'day' and 'month'. If the start date is completely missing, it will be set to the first dose date.

Stop Date: If only 'day' is missing, impute day with last day of the month. If 'day' and 'month' are missing, and 'year' is not missing, then impute month with '12' and day with '31' (or date of study discontinuation/completion if earlier than 12-31 and year is the same as the year of discontinuation). If the stop date is completely missing, it will be set to the date of study discontinuation/completion. A stop date will not be applied to ongoing AEs.

Treatment-related Adverse Events

Any AE with a relationship to study treatment of "Possibly Related" or "Definitely Related" or missing will be considered a treatment-related AE as determined by the Investigator.

Study Day 1

The first day of medication administered.

Study Day

Study day is defined as the number of days from Study Day 1.

- Before Study Day 1, Study Day = (Date of Interest – Date of Study Day 1)

- On or After Study Day 1, Study Day = (Date of Interest – Date of Study Day 1) + 1

Therefore, the day prior to Study Day 1 is -1.

Analysis Visit Window

Analysis window won't be required for this study. The unscheduled post-Baseline values will be excluded from the summary table, but only listed.

9.0 Analysis Sets

9.1 Intention-to-Treat

The 'Intent-to-Treat' (ITT) analysis set will include all randomized subjects. The ITT population will be the primary population for all efficacy analyses.

9.2 All Treated

The "All-treated" analysis set will consist of all subjects receiving at least one application of study drug. All safety analyses will be performed on the all-treated analysis set.

9.3 Per Protocol

The Per-Protocol (PP) analysis set will include all subjects in the ITT population who complete 6 weeks of treatment and 12 weeks of post-treatment efficacy evaluation without any major deviations from the protocol. The subjects to be included in the PP analysis population will be determined by the Sponsor prior to the unblinding of the study. The PP population will be used as secondary for the primary endpoint only.

Subjects potentially excluded from the PP population will be based on the following criteria and reviewed by the clinical team:

1 - Subjects who have not completed 6 weeks of treatment and 12 weeks of post-treatment efficacy evaluations
2- Subjects less than 2 years of age
3 - Subjects with more than 6 common warts at baseline
4 - Subjects without a pre-treatment wart assessment
5 - Subject who have taken an excluded prior and concomitant medications / therapies as listed in the protocol
6 - Subjects whose warts are not present for at least 4 weeks before treatment (baseline) as reported in Medical History
7 - Subjects with clinically significant dermatological disorders in the treatment area
8 - Subjects who require ongoing treatment with furosemide
9 - Subjects whose study drug compliance is as follows: missing more than 12 doses in total throughout the study
10 - Subjects with week 18 visits out of window more than 14 days outside the target visit date

11 - Subjects who received the wrong assigned treatment medication
12 – Subjects who violated key inclusion/exclusion criteria

Subject listings will be provided to the Sponsor based on the above criteria for review and final determination of the PP population. Subjects to be excluded from the PP will be determined before the study is unblinded and listed in a separate document.

10.0 Interim Analyses

No interim analyses are planned for this study.

11.0 Data Review

11.1 Data Handling and Transfer

Data will be entered and exported as SAS® version 9.4 or higher datasets. Converted datasets will be created using SAS and following standard Clinical Data Interchange Standards Consortium Standard Data Tabulation Model (CDISC SDTM, version 1.3, Implementation Guide version v3.1.3) conventions. Analysis datasets will be created using SAS and following CDISC Analysis Data Model (ADaM, version 2.1, Implementation Guide 1.0) standards.

Medical history and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 21.0 at the time of the analysis to assign a system organ class (SOC) and preferred term (PT) to each event. Prior and concomitant medications and steroids will be coded using the World Health Organization Drug Dictionary 2018MAR01 DDE + HD at the time of the analysis.

Additional details can be found in the [redacted] Data Management Plan for this study.

11.2 Data Screening

Beyond the data screening built into the [redacted] Data Management Plan, the [redacted] programming of analysis datasets, tables, figures, and listings (TFL) provides additional data screening. Presumed data issues will be output into SAS logs identified by the word “Problem” and extracted from the logs by a SAS macro and sent to Data Management.

Review of a pre-freeze TFL run on clean subjects and a post- freeze TFL run on the frozen database allow for further data screening prior to lock. The post- freeze TFL will be discussed with the sponsor in a data review meeting to identify any final data issues and seek corrections prior to database lock. The [redacted] statistician and the sponsor must approve database lock.

12.0 Statistical Methods

All statistical analyses will be performed using SAS version 9.4 or higher.

The final analysis will be performed at the end of the study (end of trial) after all subjects have completed the Week 18 assessments and the Week 30 follow-up assessments or terminated early. Please note that only the subset of subjects who achieved clearance of at least one treated wart at

Week 18/End of Post-treatment Efficacy Evaluation period will continue to the Follow-up period.

Unless otherwise specified, descriptive data summaries will be tabulated for all endpoints. Categorical data will be summarized using number of subjects (n), frequency and percentages of subjects falling into each category, with the denominator for percentages being the number of subjects in the study population, unless otherwise noted. Percentages will be rounded to one decimal place except for 100%, which will have no decimal place. Counts of zero will not display a percentage.

All continuous variables will be summarized using mean, standard deviation, median, minimum, maximum, and number of subjects with observations. The mean and median will be presented to one decimal place greater than the original data, standard deviation will be to two decimal places greater than the original data, and the minimum and maximum will have the same number of decimal places as the original data.

12.1 Subject Disposition

The number and percentage of subjects in ITT analysis set will be summarized. A listing by subject will be provided to indicate which subjects are included in each analysis set.

The number and percentage of subjects enrolled, ITT analysis set, All-Treated analysis set, Per-Protocol analysis set, who complete the study and who are in Follow-up period will be presented by treatment, together with the number and percentage of subjects who withdrew from the study prematurely and a breakdown of the corresponding reasons for withdrawal.

A listing of all subjects' analysis set inclusions, disposition, and subjects who discontinued will be presented.

12.2 Important Protocol Deviations

Important protocol deviations (IPD) are a subset of protocol deviations that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a subject's rights, safety, or well-being.

Protocol deviations will be determined based on programmed checks of the study database. These Protocol deviations are based on the following categories deemed to be important:

1. Subjects that violate key inclusion/exclusion criteria
2. Subjects who received the wrong treatment
3. Subjects who received prohibited medications/treatment during the trial
4. Subjects who met withdrawal criteria but were not withdrawn
5. Subjects who did not have necessary data collected to interpret the primary endpoint
    a. Missed Week 18 visit
    b. Week 18 more than 14 days outside target visit date
6. Subjects with a deviation that significantly affects safety or well-being
    a. Physical exam, labs or ECG not performed at screening prior to randomization
    b. Week 6 physical exam or labs not performed
    c. Pregnancy test not preformed prior to randomization
    d. Pregnancy test not performed at Week 6

For reporting purposed Protocol Deviations will be assigned to the following categories: Inclusion Criteria, Exclusion Criteria, Study Drug, Assessment Safety, Visit Window, and Prohibited concomitant medications.

The protocol deviation list will be finalized prior to the database lock, and important protocol deviation data will be listed and tabulated using incidence by deviation type.

12.3 Treatments

12.3.1 Extent of Study Drug Exposure

Exposure to study drug (e.g. at least 1 dose etc.), duration of exposure, and number of missed doses (Day 1-Week 2, Week 2-Week 4, Week 4-Week 6) will be summarized by group for the All-Treated analysis set. Summary statistics will be provided for total duration (days) of study drug exposure and treatment compliance for the All-Treated analysis set.

Duration of exposure will be defined as the total number of days a subject was exposed to study drug, calculated as (last dose date - first dose date+1).

Study drug dosing information will be presented by subject in data listings

12.3.2 Prior/Concomitant Medications

Prior/concomitant medications will be coded by World Health Organization Drug Dictionary 2018Mar01 DDE + HD and will be summarized by Anatomic Therapeutic Classification (ATC) class and preferred name. The analysis will be performed using the ITT population.

The numbers and percentages of subjects using each medication will be displayed for within double-blind treatment period and throughout the study. Subjects taking more than one medication in the same ATC class or preferred name will be counted once for the number of subjects taking that preferred name.

12.4 Demographic and Baseline Characteristics

Demographic and baseline characteristics (sex, ethnicity, race, age, age group ($\geqslant$2 to <12, $\geqslant$12 to <18 or $\geqslant$18), weight [kg], height [cm], body mass index [BMI; kg/m^2] will be summarized by treatment using descriptive statistics for subjects in ITT Analysis Set. Qualitative variables (age group, sex, ethnicity, race, wart characterization category, and wart location) will be summarized using frequencies while quantitative variables (age, weight, height, BMI, wart size, and number of warts) will be summarized using mean, SD, median, minimum, and maximum. Demographic and baseline data will also be listed for subjects for the ITT Analysis Set.

General Medical/Surgical History will be coded by MedDRA version 19.1 or later, as applicable at time of analysis and will be presented by subject listing.

Cutaneous Common Wart History will be listed by subject.

12.5 Efficacy Analyses

The primary and secondary efficacy analyses are planned analyses and the tertiary analyses are exploratory only in order to gain insight regarding the pharmacodynamic nature of the treatment response.

Wart assessments are conducted at Screening, Baseline, Week 2, Week 4, Week 6/EOT, Week 12, Week 18/End of Post-treatment Efficacy Evaluation, and Week 30 Follow up (for the subset of subjects). Wart characterization categories (solitary and cluster/splitting/satellite), the total number of warts, and the location (hands: dorsal, periungual, and palmar; foot: dorsal, limbs: arm and leg; and trunk) will be tabulated for the baseline visit for the ITT analysis set, using descriptive statistics.

The changes from baseline in total wart sizes (mm) will be summarized by group and by visit. Clinical clearance of a wart, as determined by the PI or an appropriately qualified Sub-I, should be recorded with a size of 0 mm in the CRF.

Wart assessment data will be listed for all subjects.

New common warts that appear after the Baseline visit will be identified, characterized and location recorded/mapped only, but not treated in this study. Any new common warts will be summarized in ITT analysis set for descriptive purpose and listed by subject.

Clinical Photography information collected in the CRF will be listed.

12.5.1 Primary Variable

The Primary efficacy analyses will be performed on the ITT analysis set to test the difference in the proportion of subjects with complete clearance of all treated warts (i.e., Baseline Treated warts) between the active and vehicle at Week 18 /End of Post-treatment Efficacy Evaluation.

- Primary efficacy endpoint
    - Proportion of subjects with complete clearance of all treated warts at Week 18 End of Post-treatment Efficacy Evaluation.

Proportion of subjects with complete clearance of all treated warts is defined as the percentage of the number of subjects who have all baseline treated warts as "clinical clearance" at Week 18 out of the total subjects in the ITT Analysis Set.

Statistical comparisons of proportions of subjects with all treated warts cleared will be made using Cochran-Mantel-Haenszel (CMH) test for treatment stratified by age group (≥2 to <12, ≥12 to <18 or ≥18) and investigative site at $\alpha = 0.05$. The number and percentage of subjects achieving complete clearance of all treated warts (i.e., Baseline Treated warts) will be presented by group. The 95% confidence interval will be based on the asymptotic Wald CI and will be constructed for the difference of the proportions of 2 treatments. P-value from CMH test will be presented.

The Breslow-Day chi-square test for homogeneity of the odds ratios across strata will be used to check the assumption. If the assumption test fails, the nature of the difference may be further explored, e.g. subgroup analysis by site and by age group may be explored.

The above primary efficacy analysis will be repeated on the PP analysis set as supportive.

12.5.2 Methods for Handling Dropouts and Missing Data

The frequency and percentage of missing values for each variable will be reported.

The primary method of dealing with missing data will be fully conditional specification (FCS) multiple imputations (MI; also called multiple imputation by chained equation [MICE]).

FCS MI relaxes the assumption of joint multivariate normality of all variables. FCS MI specifies the multivariate imputation model on a variable-by-variable basis on a set of conditional densities, one for each incomplete variable. Buuren, et al (2011) and Lee and Carlin (2012) provide results from simulation studies that suggest that FCS MI generally yields estimates that are unbiased with proper standard errors and provide appropriate coverage when data are missing at random (MAR).

For the multiple imputations, twenty sets of imputed data will be generated. These will be imputed using SAS Proc MI as shown below.

```
Proc MI NIMPUTE=20 SEED=861879112 ROUND=. . . . 1 1 1 1 1 1
                MINIMUM=. . . . 1 0 0 0 0 0 MAXIMUM=. . . .  6 6 6 6 6 6;
  class treat site sex agegrp;
  var treat site sex agegrp w0 w2 w4 w6 w12 w18;
  fcs logistic(treat) logistic(site) logistic(sex) logistic(agegrp) reg(w0) reg(w2)
        reg(w4) reg(w6) reg(w12) reg(w18);
run;
```

Where:

treat = study treatment (CLS006, Vehicle)
site = investigative site

sex =sex (male, female)

agegrp = age group (≥2 to <12, ≥12 to <18 or ≥18)

w0 = number of warts at baseline (total number to be treated during the study)

w2 = number of remaining treated warts at Week 2

w4 = number of remaining treated warts at Week 4

w6 = number of remaining treated warts at Week 6

w12 = number of remaining treated warts at Week 12

w18 = number of remaining treated warts at Week 18

Note: Variable names may differ from those used in the actual programs, but the methodology will be the same.

If an imputed value is not within the appropriate range for each variable in the model after the default of 100 tries, the SAS Proc MI option MINMAXITER may be increased in order to impute appropriate values.

The PROC FREQ procedure will be used to perform the *Cochran–Mantel–Haenszel test* and

will take the output dataset created from the MI procedure described above.

```
*** Obtain Mantel-Haenszel estimates ***;
PROC FREQ data=imputed;
  table site*agegrp*treat*response / cmh;
  ODS OUTPUT cmh=cmh;
  by _imputation_;
RUN;
```

The chi-square distribution is highly skewed for smaller degrees of freedom, and thus obtaining a combined result of the CMH test from multiply-imputed data requires a transformation that would normalize the CMH statistic that have a mean = 0 and unity variance. Given that the chi-square distribution for the CMH test has 1 degree of freedom, a square root transformation will be used for this purpose.

```
*** Square root transformation of the CMH statistic resulting in a normalized variate ***;
DATA cmh_std;
  SET cmh(WHERE=(AltHypothesis="General Association"));
  cmh_value_std= value**.5;
  cmh_sterr_std = 1.0;
RUN;
```

The multiple results from the CMH will be combined using PROC MIANALYZE

```
*** Combine transformed estimates;
PROC MIANALYZE DATA=cmh_std;
  ODS OUTPUT PARAMETERESTIMATES=mian_cmh_std;
  MODELEFFECTS cmh_value_std;
  STDERR cmh_sterr_std;
RUN;
```

A one-sided p-value for the combined CMH will be obtained from the normal test produced by PROC MIANALYZE on the transformed statistic. This will be done using the SAS Code below.

```
*** Compute one-sided p-value;
DATA mian_cmh_std_p;
  SET mian_cmh_std;
  IF tValue > 0 THEN Probt_upper = Probt/2;

  ELSE Probt_upper = 1-Probt/2;
RUN;
```

Last Observation Carried Forward (LOCF) and Baseline Observation Carried Forward (BOCF) will be used as sensitivity analyses.

LOCF sensitivity analysis will be performed on the primary endpoint, where values from a

subject's prior non-missing observation (including baseline and post-baseline) will be imputed for any missing post baseline visit for which there is a prior non-missing value.

BCOF sensitivity analysis will also be performed on the primary endpoint, where the baseline observation is treated as the final response from the subject if it is missing.

In general, data will not be imputed for safety analysis.

12.5.3 Multiplicity

The primary and secondary endpoints will be tested at a two-sided 0.05 level of significance sequentially in the order presented in sections 7.1 and 7.2. However, to control type-1 error, the statistical testing will stop if non-significant superiority of active to vehicle ($p>0.025$ one-sided) is observed.

12.5.4 Pooling of Sites

Approximately 40 centers are planned in this study; whenever possible, approximately at least 12 subjects will be enrolled per center. In the event a center has a low number of subjects enrolled or an insufficient number of subjects with complete clearance for the stratified CMH analysis, sites will be pooled in a manner ensuring that each pooled site will have sufficient number of subjects cleared for a stratified CMH analysis.

Planning of site pooling of sites by Meta-group will be finalized prior to database lock by a statistician independent from the study team. The pooling strategy will be 1 complete clearance and 1 not complete clearance as well as 1 in each treatment group for the given age group and site. Note that each age group need not be present in each site.

12.5.5 Secondary Variables

Secondary efficacy analyses will be conducted sequentially as follows for the following endpoints:

- Difference in ratio of cleared warts to all treated warts for each subject at Week 18/End of Post-treatment Efficacy Evaluation
- Difference in the proportion of subjects with complete clearance of all treated warts at Week 12, then Week 6/EOT.
- Difference in ratio of cleared warts to all treated warts for each subject at Week 12, then Week 6/EOT.
- Comparisons of reduction from baseline in wart size of treated warts at Week 18/End of Post-treatment Efficacy Evaluation

(1) For secondary efficacy variables, including

- Ratio of cleared warts to all treated warts for each subject at

- Week 18/End of Post-Treatment Efficacy Evaluation
- Week 12
- Week 6/End of Treatment

Ratio of cleared warts to all treated warts for each subject are defined as the ratio of the number of cleared warts out of the number of treated warts.

These endpoints will be analyzed using two-sided Wilcoxon rank sum testing at $\alpha = 0.05$. P value from Wilcoxon rank sum test of no difference between the 2 groups will be produced. Summary statistics will be presented for the endpoints.

(2) For secondary efficacy variables, including

- Proportion of subjects with complete clearance of all treated warts at
  - Week 12
  - Week 6/End of Treatment

Proportion of subjects with complete clearance of all treated warts is defined as the percentage of the number of subjects who have all baseline treated warts as "clinical clearance" at Week 12/week 6 out of the total subjects

These endpoints will be analyzed using the Cochran-Mantel-Haenzsel (CMH) test as described in the primary efficacy analysis, including the use of age group (≥2 to <12, ≥12 to <18 or ≥18), and site serving as the stratifying variables. The number and percentage of subjects achieving complete clearance of all treated warts (i.e., Baseline Treated warts) will be presented by group. The 95% confidence interval based on the asymptotic Wald CI will be constructed for the difference of the proportions of 2 treatments. P-value from CMH test will be presented. The Breslow-Day chi-square test for homogeneity of the odds ratios across strata will also be used to check the assumption described in the primary efficacy analysis.

(3) For secondary efficacy variables, including

- Reduction from baseline in total wart size of treated warts at
  - Week 18/End of Post-treatment Efficacy Evaluation

Reduction from baseline in total wart size (mm) at Week 18= total wart size of treated warts at Week 18 – total wart size of treated warts at baseline.

This endpoint will be analyzed using the analysis of covariance (ANCOVA) model with the group (active drug and vehicle) as fixed effect factors and the baseline total wart size, age group (≥2 to <12, ≥12 to <18 or ≥18), and site serving as the covariates. Descriptive statistics will include the number of observations, unadjusted mean, standard deviation, median, minimum and maximum, and least squares means. Treatment differences from vehicle will be presented and

estimated by least squares means from the analysis model along with 95% confidence intervals and associated 2-sided p-values.

Interactions between the main effects and each of the covariates will be tested in a separate model that includes the interaction terms. If the interaction terms are significant at 0.05 level, they will be further explored using descriptive statistics for each significant interaction.

The assumptions of normality and homogeneity of variance from the ANCOVA model will be tested at 0.05 level of significance. If any failure of these assumptions is observed, a non-parametric test will be used to corroborate the results of the ANCOVA. The Shapiro-Wilk test for normality and Levene's test for the homogeneity of variance will be implemented as follows:

```
*** obtain residuals from ANCOVA;
proc glm;
  class site agegrp treat;
  model change = baseline site agegrp treat / ss3;
  output r=resid out=residual;
run;quit;

*** perform Shapiro-Wilk test;
ods select TestsForNormality;
proc univariate data=residual normal;
  var resid;
run;

*** perform Levenes test;
ods select hovftest;
proc glm data=residual;
  class treat;
  model resid = treat / ss3;
means treat / hovtest=levene;
run;quit;
```

If necessary, the Non-parametric ANCOVA will be performed based on Stokes et al. as follows

1) Separately for each site and age group, Baseline and change from Baseline will be ranked across treatments using fractional ranks with tied values being assigned the mean of the corresponding ranks;

2) Separate ANCOVAs will be conducted for each site and age group for the ranked change from Baseline data with the ranked Baseline result as a covariate;

3) The residuals from the ANCOVA models will be tested using a single CMH test for treatment effect stratified by site and age group.

If necessary, the Non-parametric ANCOVA will be implemented as follows:

```
*** rank baseline and change from baseline within site and age group;
```

```
proc rank nplus1 ties=mean out=ranks;
  by site agegrp;
  var change baseline;
  ranks rchg rbase;
run;

*** get residuals within each site and age group;
proc glm data=ranks;
  by site agegrp;
  model rchg = rbase;
  output r=resid out=residual;
run;quit;

*** perform stratified CMH test;
proc freq data=residual;
  table site*agegrp*treat*resid / cmh noprint;
run;
```

Missing data will also be imputed for the secondary analyses on the same endpoint as the primary endpoint, but at different visits by multiple imputations.

12.5.6 Tertiary Variables

The following tertiary endpoints will be explored:

- All endpoints will be compared at all other time-points.
- Evaluation of durability of response and incidence of recurrence of treated warts from Week 18 end of post-treatment efficacy evaluation period to Week 30 the end of the follow-up period.

A. All endpoints at all other time-points, including

1) Proportion of subjects with complete clearance of all treated warts at
    - Week 2
    - Week 4

Comparison of the active group and vehicle will be done by the same method employed for the primary analysis on the same endpoint (i.e. CMH test).

2) Ratio of cleared warts to all treated warts for each subject at
    - Week 2
    - Week 4

Comparison of the active group and vehicle will be done by the same method employed for secondary analysis on the same endpoints (i.e. Wilcoxon rank sum test).

3) Reduction from baseline in wart size of total treated warts at

- Week 2
- Week 4
- Week 6
- Week 12

Comparison of the active group and vehicle will be done by the same method employed for secondary analysis on the same endpoints (i.e. ANCOVA).

B. Durability/Recurrence
   - At Week 18, 3 groups of responders (non-responder, complete responder, and partial responder) will be summarized.

The summary of durability and recurrence will be presented for the subjects who are complete or partial responders at Week 18 and continue to Week 30 Follow up. For the complete responders at Week 18, the number and percentage of subjects with durable response, with non-durable response from Week 18 to Week 30 will be summarized. For partial responders at Week 18, new complete responder, partial response, and non-responder at week 30 will be summarized by group. Complete responders are defined as all treated warts cleared. Partial responders are defined as at least 1 but not all treated warts cleared. Non-responders are defined as no treated warts cleared. Recurrence is defined as any treated wart not cleared at Week 30 for a complete responder at Week 18. Total complete responders at Week 30 are defined as maintenance of complete response at Week 30 plus new complete responders at Week 30.

No formal hypothesis testing is planned for these tertiary analyses. Therefore, any statistical testing done will be considered to be descriptive and not inferential.

12.5.7 Other Variables

The number of new common warts will be presented by visit.

12.6 Safety Analyses

In general, missing data will not be imputed for safety analyses.

12.6.1 Adverse Events

Any condition, event and/or signs and symptoms occurring after signing the informed consent form and before starting study drug treatment should be recorded as medical history in the eCRF. Medical conditions/diseases are considered AEs only if they worsen after starting study drug treatment (first application of study drug).

Therefore, An AE includes any condition (including a pre-existing condition) that: 1) was not

present prior to study treatment, but appeared or reappeared following initiation of study treatment, or 2) was present prior to study treatment, but worsened during study treatment.

A treatment-emergent adverse event (TEAE) is defined as an AE with onset date at or after the first dose administration of treatment, but no more than 30 days after the last dose of study medication. Only TEAEs will be included in the summary tables.

Adverse events will be coded using MedDRA version 21.0, as applicable at time of analysis. All reported AEs will be coded to the appropriate SOC and PT according to MedDRA Version 21.0.

The relationship of an AE to study drug is to be assessed according to the following definitions:

- Not Related – no temporal association or the cause of the event has been identified, or the drug cannot be implicated based upon available information.
- Possibly Related – temporal association, but other etiologies are likely to be the cause. However, involvement of the drug cannot be excluded, based upon available information.
- Definitely Related – established temporal or other association (e.g., re-challenge) and event is not reasonably explained by the subject's known clinical state or any other factor, based on available information.

The severity of an AE is to be scored according to the following scale:

- Mild - Awareness of sign or symptom, but easily tolerated
- Moderate - Moderate Discomfort enough to cause interference with usual activity
- Severe - Severe Incapacitation with inability to work or perform usual activity

A high level overall summary of TEAEs will be presented to summarize AEs in any category.

Subject incidence of the following AEs will be tabulated by SOC (descending order of frequency) and PT (descending order of frequency). The number and percentage of each TEAE will be calculated.

- TEAEs
- severe TEAEs
- treatment-related TEAEs
- maximum severity of TEAEs
- TEAE with an outcome of death
- TEAE leading to discontinuation of study treatment
- TEAE leading to discontinuation of the study

Subject incidence of the following AEs will be tabulated by PT in descending order of frequency:

- TEAEs

- TEAEs experienced by ≥ 2 subjects (for any treatment)
- treatment-related TEAEs by ≥ 2 subjects (for any treatment)

Subjects will be counted only once within each SOC or PT. For tables categorized by severity, subjects with multiple events within a particular SOC or PT will be counted under the category of their most severe event for that SOC or PT. Treatment-related TEAEs included AEs from AE CRF page where relationship to study treatment is checked with either 'Possibly Related' or 'Probably Related'. Missing values for severity will be assigned as 'Severe' and missing value for relationship to treatment will be considered as 'Related' in the summary.

AEs leading to discontinuation of treatment are those with an action taken with study treatment on the AE CRF page of "Drug Withdrawn".

AEs leading to discontinuation from the study are those on the AE CRF page of "YES" to the question "Did the adverse event cause the subject to be discontinued from the study?".

By-subject listings will be provided for the following: All AEs, All TEAE, TEAEs leading to discontinuation of treatment and TEAEs leading to discontinuation of the study.

The duration of AEs will be derived as the AE end date (including imputation date for incomplete AE end date) – AE onset date (imputed date for incomplete AE onset date) +1.

AEs of Special Interest (AEOSI)

Study personnel should question and observe subjects for evidence of AEs paying particular attention to local application site reactions and any potential cardiovascular or hematologic AEs.

Subject incidence of the following AEs will be tabulated by SOC (descending order of frequency) and PT (descending order of frequency). The number and percentage of each TEAE will be calculated.

- TEAE of special interest
- Serious TEAE of special interest
- TEAE of special interest leading to discontinuation of study treatment
- TEAE of special interest related to study drug
- Preferred terms used to identify adverse events of special interest (AEOSIs) will be provided by the medical team after a review of AEs appearing in the following system organ classes: general disorders and administration site conditions, cardiac disorders, blood and lymphatic system disorders, renal and urinary disorders, investigations, skin and subcutaneous tissue disorders, and metabolism and nutrition disorders. All AEOSIs will also be provided in a by-subject listing. Preferred terms used to identify adverse events of special interest (AESIs) will be determined by the medical team before database freeze. All AEOSIs will be provided by-subject listing.

12.6.2 Deaths and Serious Adverse Events

Serious TEAEs will be summarized by SOC and PT for the All-Treated analysis set.

A further summary of treatment-related SAEs will be provided for the All-Treated analysis set.

A table presenting the number and subjects who died during the study will be presented for the All-Treated analysis set.

By-subject listings will be provided for the following: All Deaths, SAEs.

12.6.3 Laboratory Data

Laboratory test results are reported in International System of Units (SI).

Descriptive summaries of changes from baseline for each laboratory test listed below will be provided for numerical laboratory assessments by visit.

Laboratory shift tables from baseline to each post baseline visit for lab values evaluation of 'High' 'Low' or 'Normal', when available, will be planned when the baseline results are available. The categorical lab results will be summarized by frequency and percentage and by visit.

The clinical safety laboratory tests will include the following parameters:

- Hematology: hemoglobin, hematocrit, red blood cell (RBC) including mean corpuscular volume (MCV) and red cell distribution with (RDW), white blood cell (WBC) count with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils), and platelet count.
- Chemistry/Lipids:
    - General: calcium, glucose, lactate dehydrogenase (LDH)
    - Electrolytes: carbon dioxide (bicarbonate), chloride, magnesium, sodium, potassium
    - Kidney/Liver: blood urea nitrogen (BUN), creatinine, uric acid, total protein, aspartate aminotransferase (AST), alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), total bilirubin
    - Lipids: total cholesterol, high-density lipoproteins (HDL), low-density lipoproteins (LDL), triglycerides

12.6.4 Vital Signs

Vital signs (blood pressure [BP], heart rate [HR]) will be measured at each visit during Screening, Double-blind Treatment and Post-treatment Efficacy Evaluation Periods.

Observed values and changes from baseline in vital signs will be summarized by vital sign parameter and visit. Descriptive statistics will be shown for baseline, each post-baseline time point, and the change from baseline to each post-baseline time point.

12.6.5 Physical Examinations, ECGs, and Other Observations Related to Safety

A limited physical exam and only baseline ECG are conducted in this study.

Physical exam, ECG, and urine pregnancy test data will be listed.

13.0 Validation

PRA's goal is to ensure that each TFL delivery is submitted to the highest level of quality. Our quality control procedures will be documented separately in the study specific quality control plan.

14.0 References

Breslow NE, Day NE: The analysis of case–control studies. Statistical Methods in Cancer Research, Volume 1. Lyons: IARC Scientific Publications; 1980.

Buuren, S.V., Brand, J.P., Groothuis-Oudshoom, C.G., & Rubin, D.B. (2006). Fully conditional specification in multivariate imputation, Journal of Statistical Computation and Simulation. 76(12), 1049-1064.

Goria, M.N. (1992). On the fourth root transformation of chi-square. Australian Journal of Statistics, 34 (1), 55-64.

Lee, K.J. & Carlin, J.B. (2012). Recovery of information from multiple imputation: a simulation study, Emerging Themes in Epidemiology. 9(3), 1-10.

Wilson, E.B. & Hilferty, M.M. (1931). The distribution of chi-squared. Proceedings of the National Academy of Sciences, Washington, 17, 684–688.

Stokes, Maura E., Davis, Charles S., and Koch, Gary G. (2000), Categorical Data Analysis Using the SAS® System, Second Edition (pp. 174-179), Cary, NC: SAS Institute Inc.

Appendix 1 Glossary of Abbreviations

Glossary of Abbreviations:	
AE	Adverse Event
AEOSI	Adverse Event of Special Interest
ALT	Alanine aminotransferase
ANCOVA	Analysis Of Covariance
AST	Aspartate aminotransferase
BOCF	Baseline Observation Carried Forward
BP	Blood pressure
BUN	Blood urea nitrogen
CBC	Complete blood count
CLS	Cutanea Life Sciences, Inc.
CLS006	Furosemide Topical Gel 0.125% w/w
CMH	Cochran-Mantel-Haenszel
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EOS	End of Study
EOT	End of Treatment
ET	Early Termination
GGT	Gamma glutamyl transferase
HDL	High-density lipoproteins
HR	Heart rate
ITT	Intent-to-Treat
IWRS	Interactive Web Response System
LDH	Lactic dehydrogenase
LDL	Low-density lipoproteins
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
PI	Principal investigator
PP	Per Protocol
RBC	Red blood cell
SAE	Serious Adverse Event
SD	Standard Deviation
TEAE	Treatment emergent adverse events
WBC	White blood cell
w/w	Weight for weight