STATISTICAL ANALYSIS PLAN - TEXT

Title: **THWART 1: A PHASE 3 RANDOMIZED, DOUBLE-BLIND, VEHICLE-CONTROLLED, PARALLEL GROUP STUDY OF A-101 TOPICAL SOLUTION APPLIED TWICE A WEEK IN SUBJECTS WITH COMMON WARTS**
Protocol: A-101-WART-302
Study Drug: A-101
Sponsors: Aclaris Therapeutics,Inc.
Version (Date): Text 1.2 (26 September 2019)
Status: ***Final***

Prepared by:

John Ilgenfritz: Statistical Consultant — Date: 10 OCT 2019

Approved by:

Aclaris: Rebecca DeMoor — Date: 09OCT2019

Rebecca DeMoor: Principal Statistician

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Term
AE	Adverse Event
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
CBC	Complete Blood Count
CMH	Cochran-Mantel-Haenszel
CPH	Cox Proportional Hazards
CRF	Case Report Form
FCS	Fully Conditionally Specified
ID	Identification
ITT	Intent to Treat
LSR	Local Skin Reactions
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
PHREG	SAS Procedure for Proportional Hazards Regression
PP	Per Protocol
PT	Preferred Term
PWA	Physician's Wart Assessment
SAE	Serious Adverse Event
SAP	Statistical Analysis Alan
SOC	System Organ Class
US	United States

INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide a description of the statistical analyses performed for the Phase 3 protocol, A-101-WART-302, Amendment 2, Version 4 (12Oct2018).

1. STUDY OBJECTIVES

1.1 Primary Objective

The primary objective of this study is to evaluate the effectiveness of A-101 45% compared to Vehicle when applied twice weekly to common warts.

1.2 Secondary Objectives

The secondary objectives of this study include:

- Duration of response to A-101 45% compared to Vehicle.
- Onset of action of A-101 45% compared to Vehicle.
- Safety of A-101 45%.

2. STUDY DESIGN

This is a phase 3, multicenter, randomized vehicle controlled, double blind parallel group study to evaluate the safety and efficacy of A-101 45% vs Vehicle in subjects with common warts. Investigators will be required to identify at least 1 and up to 6 clearly identifiable common warts for treatment with A-101 study medication. All identified common warts will be treated twice a week for up to 8 weeks (maximum of 16 treatment applications). **S**ubjects will be required to complete a total of 13 study visits. The protocol defined study visits are:

- Visit 1 (Day -13 to 0) screening
- Visit 2 (Day 1)* randomization; study medication treatment
- Visit 3 (Day 8)* study medication treatment
- Visit 4 (Day 15)* study medication treatment
- Visit 5 (Day 22)* study medication treatment
- Visit 6 (Day 29)* study medication treatment
- Visit 7 (Day 36)* study medication treatment
- Visit 8 (Day 43)* study medication treatment
- Visit 9 (Day 50)* study medication treatment
- Visit 10 (Day 60) follow up evaluations, no identified Wart retreatment
- Visit 11 (Day 78) follow up evaluations, no identified Wart retreatment
- Visit 12 (Day 106) follow up evaluations, no identified Wart retreatment
- Visit 13 (Day 137) follow up evaluations, no identified Wart retreatment; end of study

*The second weekly application will be applied by the subject/parent/legal guardian at home (Day 4, 11, 18, 25, 32, 39, 46, and 53).

Approximately 500 subjects will be randomized to one of 2 treatment arms in a 1:1 ratio

at approximately 25 investigational centers in the US. Of the 500 subjects randomized, approximately 100 subjects will be between 1 and 17years of age.

The study schedule of assessments can be found in Table 1.A.

Schedule of Assessments

Visit	V1 Screening	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13
Treatment Day	-13 to 0	1	8	15	22	29	36	43	50	60	78	106	137
Treatment Window	N/A	N/A	±1 day	±1 day	±1 day	±1 day	±1 day	±1 day	±1 day	+4 days	±7 days	±7 days	+7 days
Study Procedures													
Informed Consent	▲												
Inclusion Criteria/Exclusion Criteria	▲	▲[1]											
Subject Identifier	▲[2]												
Medical history/demographics	▲												
Fitzpatrick Skin Type Assessment	▲[3]												
Vital Signs	▲[4]								▲				▲
Prior Medications/Therapies	▲[5]												
Clinical Chemistry and CBC[6]	▲												▲
Common Wart Identification[7]	▲												
Physician's Wart Assessment[8]	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Common Wart Dimensions[9]	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Standardized Photography[10]	▲	▲[11]		▲		▲			▲	▲	▲	▲	▲
Subject Randomization		▲[12]											
Local Skin Reactions		▲[13]	▲[13]	▲[13]	▲[13]	▲[13]	▲[13]	▲[13]	▲	▲	▲	▲	▲
Study Medication Application[14]		▲	▲	▲	▲	▲	▲	▲	▲				
Study Medication Application at Home[15]			A-101 Study Medication will be applied at home on Days 4, 11, 18, 25, 32, 39, 46, 53. A window of +1 day is allowed.										
Subject Instructions	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
Concomitant therapies[16]		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
Adverse Events[17]		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲		

See Footnotes starting on next page.

[1]Subject inclusion/exclusion criteria will be re-assessed prior to randomization during Visit 2.
[2]Investigational sites will assign a unique five-digit subject identifier to each subject at Visit 1. This subject identifier will be used in all study documentation for the duration of the study.
[3]Each subject's skin must be assessed during Visit 1 using the Fitzpatrick Skin Type Assessment.
[4]Vital signs [including temperature, pulse, respiratory rate, blood pressure, height and weight (Visit 1 only)] will be measured by a qualified staff member at Visit 1, Visit 9 and Visit 13.
[5]Prior medications/therapies will be collected for a time-period of 14 days prior to Visit 2.
[6]A complete blood count (including hematocrit, hemoglobin, platelet count, red blood cell count and morphology, white blood cell count and differential (absolute and %) including basophils, eosinophils, lymphocytes, monocytes and neutrophils and a clinical chemistry panel including albumin, alkaline phosphatase, alanine aminotransferase (ALT), aspartate aminotransferase (AST), blood urea nitrogen (BUN), bicarbonate, calcium, chloride, creatinine, glucose, lactate dehydrogenase (LDH), phosphorus, potassium, sodium, total bilirubin, total protein, and uric acid.
[7]The treating investigator will identify at least 1 and up to 6 common warts located on the trunk or extremities that meet the inclusion criteria.
[8] The investigator will use the Physician Wart Assessment (PWA) scale to assess each identified common wart. The investigator must assess the identified common warts prior to application of the study medication at Visit 2-Visit 9. In order to be eligible for randomization at Visit 2, each common wart must have a PWA grade ≥ 2.
[9] The investigator will measure the dimensions (longest axis and thickness of the common wart) at Visit 1 and prior to randomization at Visit 2. The investigator will be required to measure the longest axis of each common wart at Visits 2-13.
[10] At Visits 1, Visit 2, 4, 6, and 9 (prior to study medication application), and Visit 10-13, a qualified investigational center staff member will take a photograph of each identified common wart using the Aclaris supplied camera. All photographs will be sent to a central imaging laboratory.
[11]At a sub-set of investigational sites (up to 3 sites), subjects randomized at these sites will have additional standardized photographs taken of their common warts at the following time-points: V2-10 minutes post application, 1-hour post application and 24 hours post application.
[12]Subjects will be randomized at Visit 2 prior to application of study medication. Investigational study staff will re-confirm subject eligibility prior to randomization.
[13] The investigator and subject will assess each identified common wart for local skin reactions associated with irritation at Visits 2-13. At Visit 2-Visit 9, the investigator and subject will assess each identified common wart for LSRs prior to the application of the study medication.
[14] During in-office visits, subjects 18 years of age and older will apply their A-101 study medication in the presence of the treating physician or a member of the investigational study staff that has been trained on the protocol. Subjects ages 1 to 17 will have their A-101 study medication applied by a parent/legal guardian in front of the treating physician or a member of the investigational study staff during in-office visits. If the identified common wart meets the criteria for re- treatment the lesion will be re-treated at Visits 3-9. Following application of study medication, subjects must NOT wash/submerge the treated wart for at least 6 hours and they must NOT apply any topical products to the treated common warts for at least 6 hours. Parents/legal guardians of children must ensure that the child does not put the treated area in their mouth, or eyes, until completely dry after the application. If the treated area is not completely dry 10 minutes after the application, the area should be blotted dry.
[15]All subjects will receive their second weekly treatment of the A-101 study medication at home on Days 4, 11, 18, 25, 25, 32, 39, 46, 53 as determined by the PWA assessment performed during the in-office visit. Subjects ages 1 to 17 will have their A-101 study medication applied by a parent or legal guardian. All subjects will need to document the application of A-101 study medication in a subject diary. The second weekly treatment will depend on the PWA performed during the in-office visit that week. If the identified common wart meets the criteria for re-treatment during the in-office visit, then the subject/parent/guardian will perform the second weekly treatment at home.
[16]All concomitant therapies including (topical and oral) prescription medications, over the counter medications and natural supplements and non-drug therapies including chiropractic, physical therapy, and energy-based therapy must be documented in the subject CRF. Subjects must not apply any topical products (*e.g.*, moisturizers, sunscreen, etc.) to the identified common wart within 12 hours prior to any study visit.
[17]The reporting period for Serious Adverse Events (SAEs) begins when the subject signs the informed consent and continues through Visit 11. Non-serious clinical adverse events will be collected following the application of the study medication at Visit 2 and through Visit 11. Non-serious adverse events that occur between the time of consent and study medication application will be documented as medical history. All safety reporting (AEs and SAEs) will conclude at Visit 11 (approximately 28 days after last study medication application) except for clinical adverse events related to local skin reactions.

3. STATISTICAL METHODS

The Intent-to-Treat Population (ITT) will be used for primary and secondary efficacy analyses and summaries. The Per-Protocol (PP) population will be used for sensitivity analyses for the primary efficacy analysis and all secondary efficacy analyses included in the hierarchical testing procedure (Figure 1). Subjects will be summarized and analyzed according to randomized treatment group for efficacy. The Safety Population will be used for safety analyses. Subjects will be summarized according to actual treatment received at first application of study medication for safety.

The baseline scores for PWA will be the last non-missing value collected ≤ the date of the first application of double-blind study medication.

For the purpose of analysis, data will be assigned to an analysis visit based on relative study day as defined in Appendix B.

All Data listings will be sorted by randomized treatment group and subject ID.

3.1 Study Populations, Disposition, Baseline Characteristics, Medical History Concomitant Medications, Exposure and Protocol Violations

3.1.1 Analysis Populations

The populations used for analyses will be defined as follows:

- Intent-to-Treat Population (ITT): The ITT population includes all randomized subjects.
- Per-Protocol Population (PP): The PP Population includes all subjects who received at least one application of study medication and had no protocol violations as defined in Section 3.1.6.
- Safety Population: The Safety population includes all subjects who received at least one application of study medication.

3.1.2 Subject Disposition and Discontinuation from Study

The number of subjects screened, randomized and treated will be presented together with the study completion status and the reasons for discontinuation grouped by treatment and by major reason (lost to follow-up, adverse event, poor compliance etc.).

A listing will be provided of all patients discontinued from the study after enrolment, broken down by center and treatment group, giving a patient identifier, the specific reason for discontinuation, the treatment (drug and dose), and the duration of treatment before discontinuation. Whether or not the blind for the patient was broken at the time of discontinuation will be noted. Demographic and Baseline Characteristics
Demographic and baseline characteristics will be presented in the data listings.

Descriptive summary statistics (n, mean, standard deviation, median, minimum, and maximum) and/or frequency distributions, as appropriate, will be generated by randomized treatment group and combined.

3.1.3 Wart Diagnosis and Characteristics at Baseline

Descriptive summary statistics (n, mean, standard deviation, median, minimum, and maximum) and/or frequency distributions, as appropriate, will be generated by randomized treatment group and combined. Summarized variables include total warts treated, time since first appearance of most and least recent treated warts, longest axis of largest wart treated, and thickness of largest wart treated.

3.1.4 Prior and Concomitant Medications, Therapies and Procedures

Medications are coded using the WHO Drug (March 2019). Prior and concomitant medications are listed only. The period(s) in which the medications were administered will be flagged.

3.1.5 Protocol Violations

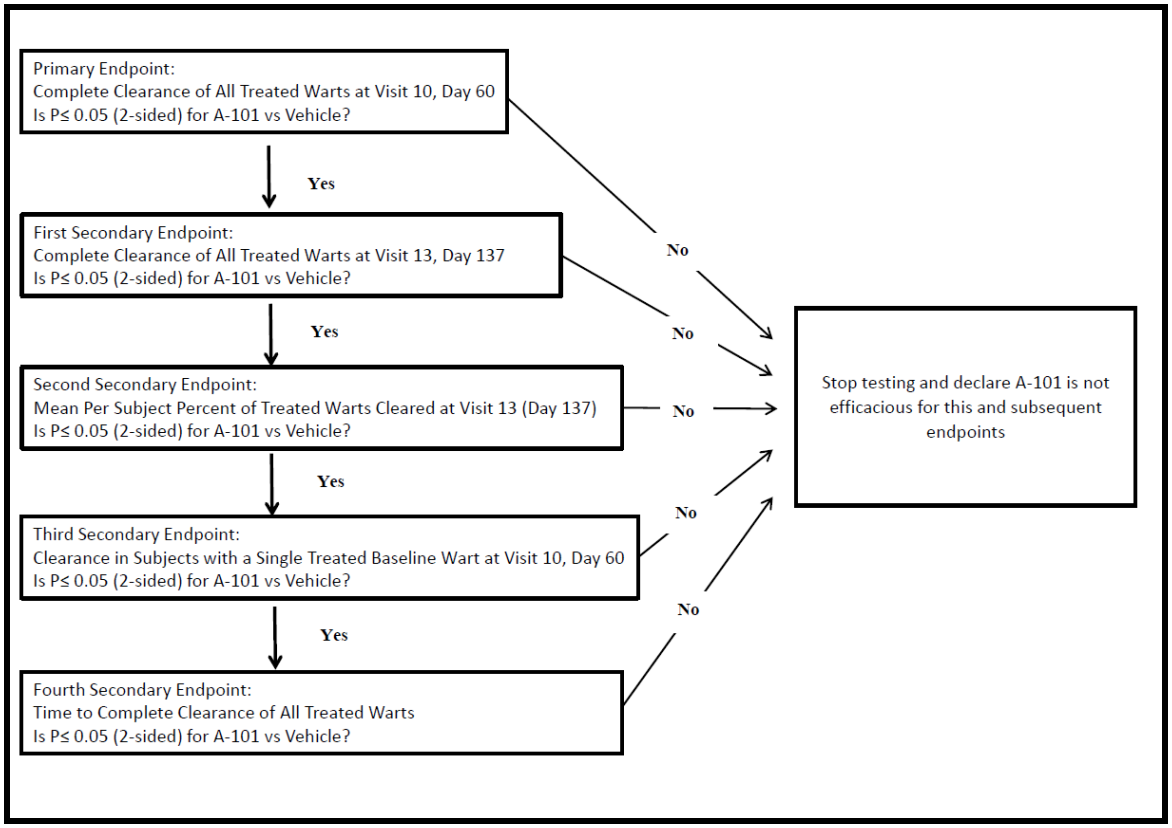
Protocol violations will be identified prior to database lock to measure adherence to key aspects the protocol (see section 16.2 of the protocol for full definitions). Specific data fields that will be examined to identify protocol violations include entrance inclusion/exclusion criteria and prohibited prior and concomitant medications as well as all deviations identified by the investigator. All protocol violations will be listed by patient number and summarized (n and percent) by randomized treatment group.

3.2 Efficacy Endpoints

The primary efficacy analysis will evaluate the effectiveness of A-101 45% compared to Vehicle at the endpoint visit, Visit 10 (Day 60). It will be based on the proportion of subjects who achieve complete clearance (PWA=0) of all treated common warts.

Four secondary efficacy analyses will be performed and tested in a hierarchical Fixed-Sequence step-down method to control alpha among the pre-specified secondary efficacy variables. The predefined order of the tests is specified to be the order of these analyses as presented below (Figure 1). All analyses will be conducted at the same alpha level (2-tail alpha = 0.05). The results of the second analysis will only be reported if the first analysis reaches statistical significance at alpha = 0.05. The results of the third analysis will only be reported if the first and second analyses reach statistical significance, each at alpha = 0.05. The results of the fourth analysis will only be reported if the first, second and third analyses reach statistical significance, each at alpha = 0.05.

Figure 1 Hierarchical Testing of Primary and Secondary Endpoints to Maintain Study-Wise Type I error at α=0.05 (2-Sided)



For regulatory purposes, testing for will halt at the first endpoint with P> 0.05, however all endpoints will be analyzed irrespective of the results. Supportive and sensitivity analyses of the primary and secondary endpoints will also be performed.

Should sparse enrollment at one or more sites occur, the following algorithm will be employed to "pool" sites for analysis. Sites with fewer than 5 subjects randomized will be pooled by ordering the sites based on the number of subjects randomized and the site number (in the event multiple sites have the same number randomized) and then combining the smallest with the next smallest site in the sort order in a stepwise fashion until all the sites/pooled sites have more than 4 subjects. A pooled site can be pooled again if, after an initial pooling, it remained part of the smallest or next smallest site or pooled site. In the sort order when there are ties in count, a "pooled site" comes after non-pooled sites in the pooling hierarchy. The pooled site(s) will be used in all analyses that are stratified by site and/or include site in the statistical model.

3.2.1 Complete Clearance of All Warts

The primary efficacy analysis will evaluate the effectiveness of A-101 45% compared to Vehicle at the endpoint visit, Visit 10 (Day 60). It will be based on the proportion of subjects who achieve complete clearance (PWA=0) of all treated common warts. A

Cochran-Mantel-Haenszel (CMH) test stratified by the baseline number of treated warts (1 wart, 2 warts, 3 warts, >3 warts) will be used to perform the primary efficacy comparison between treatment groups, with 2-tail alpha set at 0.05. Any treated common wart for a randomized subject that has missing data at the endpoint visit will be treated as not clear for the purpose of the primary efficacy analysis. Corresponding estimates of the common Mantel-Haenszel odds ratio will be provided along with 95% confidence intervals (CIs). Stratum specific odds ratio estimates and 95% confidence intervals will also be provided.

Syntax for the SAS statements used to perform the analysis will be consistent with respect to the statistics calculated in the following:

```
PROC FREQ;
TABLE WARTS*TREATMENT*RESPONSE/CMH;
OUTPUT OUT=STATS (KEEP=P_CMHGA  _MHOR_ L_MHOR  U_MHOR  P_BDCHI) CMH;
```

Where WARTS represents the number of warts assessed at baseline (1 wart, 2 warts, 3 warts, >3 warts); TREATMENT is 1 for A-101 and 2 for Vehicle; and RESPONSE = 1 for those with all warts cleared and 2 for those without all warts cleared. P_CMHGA is the CMH general association test, _MHOR_ is the Mantel-Haenszel estimated common odds ratio, and L_MHOR and U_MHOR are the lower and upper limits, respectively, of the corresponding 95% confidence interval. P_BDCHI is the p-value from the Breslow-Day test of the homogeneity of the odds ratios across strata and is provided as a supportive statistic to assess the model. A statistically significant finding ($P<0.05$) of the Breslow-Day test indicates that the treatment effect differs by stratum and these differences should be investigated by stratum-specific odds ratios estimates which will be included in the output for the primary endpoint.

3.2.2 Sensitivity Analyses for the Primary Endpoint

The first sensitivity analysis for the primary endpoint will be an analysis based upon the PP population. This analysis will use the same CMH model described for the primary analysis.

A second sensitivity analysis will be conducted for the primary endpoint using the same CMH model for the ITT population that includes the imputation of missing PWA data using a fully conditionally specified (FCS) regression-based Multiple Imputation (MI) procedure. This regression-based MI will be conducted on the worst (maximum) PWA score for all treated warts at a given visit for each subject. The utilization of the FCS methodology accounts for the correlation within a subject and models mean changes over time in the maximum PWA score. For each subject, only data selected as representative of an analysis visit (i.e. within the defined relative day range and closest to the “target”) as documented in Appendix B are included in the imputation process. The imputed data will be in the form of the maximum PWA score and will be rounded to an integer and bounded by 0 below and 3 above. All imputed maximum PWA scores that are assigned a 0 will correspond to complete clearance of all warts at that visit. The complete clearance at visit 10 using this MI data will be analyzed using the same CMH model described for

the primary analysis. Results from the CMH on the MI data will then be summarized using the Wilson-Hilferty[3] transformation to obtain the sensitivity results primary efficacy endpoint. The SAS syntax used to perform the MI and the MI summary is shown in Appendix A

To assess the impact of missing data imputation on primary analysis results a tipping point analysis will be conducted. This analysis will shift the treatment coefficient in the regression model used in the MI procedure to impose a worse mean maximum PWA for the A-101 45% arm and a better mean maximum PWA for the vehicle arm. The shifts will be from 0 to 3 in increments of 0.5 for the A-101 45% arm and from 0 to -3 in increments of 0.5 for the vehicle arm. Apart from these tipping point shifts, the analysis will follow the same methodology described for the MI procedure. The odds ratio and CMH Chi-square P-value are provided for each of the 7 x 7 =49 shift combinations for the two arms.

In addition, a model similar to the primary analysis model but stratified instead by Study Center will be performed as a sensitivity analysis including the Breslow-Day test to evaluate heterogeneity across sites. The SAS syntax differs only by replacing the WARTS in the above example with SITE where site may be a "pooled site" as detailed in section 3.2.

3.2.3 Secondary Efficacy Endpoints

Visit 13 (Day 137) will be analyzed using an identical model as the primary endpoint at Visit 10 (Day 60). This will be the first secondary endpoint tested in the hierarchy of endpoints described in Section 3.2. This analysis will be run on the ITT population using randomized treatment group.

Sensitivity analyses performed for the primary endpoint will be repeated for the Visit 13 analysis. These sensitivity analyses will include the PP analysis, missing data imputation impact analyses, and the tipping point analysis. The site heterogeneity analysis will not be performed.

3.2.4 Mean Per Subject Percent of Warts Cleared

Mean per-subject percent of warts cleared at Visit 13 (Day 137) will be evaluated as the second secondary endpoint using analysis of covariance (ANCOVA) with treatment group as a class variable and the baseline number of warts assessed as a continuous covariate. Patents without a week Visit 13 measurement according to the assignment to analysis visit in Appendix B will be assigned a score of 0% warts cleared. Ninety-five percent confidence intervals will be constructed about the estimated treatment difference based on the model estimates and associated variability.

Syntax for the SAS statements used to perform the analysis will be consistent with respect to the statistics calculated in the following:

PROC GLM:
CLASS TREATMENT;

```
MODEL PER_CLEAR = TREATMENT WARTS / SOLUTION CLPARM;
LSMEANS TREATMENT / PDIFF STDERR ALPHA=0.05 CL E;
ESTIMATE 'A-101 vs Vehicle' TREATMENT 1 -1 / CL;
```

Where WARTS represents the number of warts assessed at baseline; TREATMENT is 1 for A-101 and 2 for Vehicle; and PER_CLEAR is the mean per subject percent of warts cleared.

A separate sensitivity analysis for the mean per-subject percent of warts cleared will be conducted using the PP population.

3.2.5 Complete Clearance in Subjects with a Single Wart at Baseline

A Cochran Mantel Haenszel (CMH) test stratified by investigative site will be used to compare treatment groups with respect to the proportion whose wart is Clear (PWA=0) at Visit 10 in the subset of ITT subjects with a single wart at baseline. This is the third secondary endpoint in hierarchy summarized in section 3.2.

Syntax for the SAS statements used to perform the analysis will be consistent with respect to the statistics calculated to the following:

```
PROC FREQ;
TABLE SITE*TREATMENT*RESPONSE/CMH;
OUTPUT OUT=STATS (KEEP=P_CMHGA _MHOR_ L_MHOR  U_MHOR  P_BDCHI) CMH;
```

Where SITE may be a "pooled site" as detailed in Section 3.2 and the other terms are as detailed in Section 3.2.1.

A sensitivity analysis for the complete clearance in subjects with a single wart at baseline will be conducted in the subset of the PP subjects with a single wart at baseline.

3.2.6 Time to Clearance of All Warts

The fourth secondary endpoint tested, the difference in time to clearance of all warts from randomization between the two treatment arms, is assessed in the ITT population using the stratified log-rank test (Score statistic from PHREG and ties=Breslow) from a Cox proportional hazards (CPH) model including treatment and the number of warts assessed at baseline in the model. Survival for each arm is summarized using Kaplan Meier curves and is further characterized in terms of the median and survival probability at selected timepoints, along with the corresponding 2-sided 95% confidence intervals for the estimates. Confidence intervals for median survival are based upon the methods of Brookmeyer and Crowley[1]. Confidence intervals for survivorship estimates are calculated using the log-log transformation[2].

Syntax for the SAS statements used to perform the analysis will be consistent with respect to the statistics calculated to the following:

```
PROC PHREG;
MODEL DAYSFOL*RESPONSE(2) = TREATMENT / RISKLIMITS TIES=BRESLOW;
```

```
STRATA  WARTS;
```

Where WARTS represents the number of warts assessed at baseline (1 wart, 2 warts, 3 warts, >3 warts); TREATMENT is 1 for A-101 and 0 for Vehicle; and RESPONSE = 1 for those with all warts cleared and 2 for those without all warts cleared. Subjects without complete clearance (i.e. RESPONSE = 2) are censored on the date of the last wart assessment. DAYSFOL = Date of last wart(s) assessment – date randomized +1. Patients without a post baseline assessment are censored on date of randomization. The Wald Chi-Square P-value is used.

Kaplan-Meier estimates for plotting and for estimates of median and of the survivorship curve at selected times (Days from Randomization) will be obtained using SAS statements consistent with the following syntax:

```
PROC LIFETEST ALPHA=.05;
TIME DAYSFOL*RESPONSE(2);
SURVIVAL OUT=SURV CONFTYPE=LOGLOG;
STRATA TREATMENT;
```

A sensitivity analysis for the time to clearance of all warts will be conducted on the PP population.

3.2.7 Subgroup Analyses of Efficacy

Exploratory subgroup analyses of the primary and secondary endpoints will be performed. Descriptive statistics will be provided by treatment group. Odd Ratios for wart clearance and hazard ratio for time to wart clearance will also be presented. The following subgroups will be assessed:

- Age (<18, ≥18)
- Sex (Male, Female)
- Race (White, Non-White)
- Maximum PWA at Baseline (2,3)
- Fitzpatrick Skin Type (I, II, III, IV, V, VI)
- Number of treated warts (1 wart, 2 warts, 3 warts, >3 warts)
- Subjects <18 years old with 4 or more treated warts

Additionally, an analysis of wart clearance according to wart location will be performed using the individual warts as the experimental unit (i.e. the denominator). Five location categories are defined as follows based upon the grid locations as recorded on the CRF:

Subgroup	Grid Locations
Hands	'A4', 'D4', 'E4', 'H4'
Feet	'B6', 'C6', 'F6', 'G6'
Elbows	'A3', 'D3', 'E3', 'H3'
Knees	'B5', 'C5', 'F5', 'G5'
Rest of Body	All locations not included in above 4 categories

3.3 Safety Analyses

Safety summaries will be conducted on the Safety Population and use actual treatment received.

The following sections detail the summaries performed on the safety data. Additional data handling rules including those for imputation of partially missing dates are provided in a separate SAP document pertaining to data handling and programming specifications.

The baseline scores for vital signs, laboratory assessments and local skin reactions will be the last non-missing value collected ≤ the date of the first application of double-blind study medication.

3.3.1 Extent of Exposure

The duration of treatment, the number of application days (i.e. the actual count of days where an application occurred), the mean applications per wart (for each subject the sum of all applications across all treated warts and all application days divided by the total number of treatments for all treated warts), and the total applications will be summarized using descriptive statistics (n, mean, SD, median, minimum and maximum). The percentage of total wart treatments that were performed in 1, 2 or 3 applications is also provided.

Overall compliance is calculated as the number total number of warts treated for the treatment period (through date of last application to any wart) divided by the total number of warts expected to be treated. The total number of warts treated is the sum of the retreatment and home application diary records completed times the number of warts assessed at baseline minus the number of warts scored as PWA=0 or having an LSR grade 3 at the most recent assessment (<= the date of the application).

3.3.2 Adverse Events

Adverse events are coded using the MedDRA (Version 22) dictionary and are categorized by system organ class (SOC) and preferred term (PT). Only treatment emergent adverse events are included in summary tabulations (i.e. adverse events occurring on or after the first application of study medication and within 30 days following the last application of study medication). All adverse events are included in the data listings including those that are not considered treatment emergent.

To allow differentiation as to which study period an Adverse Event occurred, three categories are defined based upon onset date. Adverse Events that had onset dates prior to the first application of study medication are considered "prior". Adverse events with onset dates on or after the first application of double-blind study medication and within 30 days following the last application of study medication are considered "treatment emergent". Events with onset dates more than 30 days after the last application of study medication are considered as "post-therapy".

Frequency tabulations are presented by MedDRA SOC and preferred term, for all adverse events; study treatment-related, adverse events resulting in discontinuation of study treatment, serious adverse events and adverse events by maximum severity. Adverse events resulting in discontinuation are those with 'action taken' recorded as 'drug withdrawn'.

Treatment emergent adverse event summaries by MedDRA SOC and preferred term will also be provided according to the following subgroups:

- Age (<18, 18 to 65, and >65)
- Sex (Male, Female)
- Race (White, Non-White)

3.3.3 Local Skin Reactions

The number of subjects with local skin reaction signs that worsened by at least 1 grade from baseline any time after the first application of study medication will be tabulated according to treatment group for each sign as well any sign. A separate summary limited to LSRs that worsen and which occur within 30 days of the last application of study medication will be provided. The number of subjects with local skin reaction signs that worsened by at least 1 grade from baseline any time after the first application of study medication will be tabulated by treatment group and Fitzpatrick Skin Type for each sign as well as any sign.

3.3.4 Vital Signs

Systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate and temperature are summarized using descriptive statistics for each visit. Both actual values and changes from baseline will be summarized.

3.3.5 Laboratory Evaluations

Shift tables showing changes in relationship to the normal reference range grade from baseline to the "worst" relationship recorded post-baseline are tabulated. In the event a subject had both "low" and "high" post-baseline values the subject will be counted under both "low" and "high". As such, the percentages of the cells in the shift table may add to more than 100%.

Chemistry and hematology parameters are summarized using descriptive statistics for

each visit. Both actual values and changes from baseline will be summarized.

The number and percentage of subjects meeting criteria for hepatobiliary abnormalities will be provided.

3.4 Interim Analyses

No interim analyses are planned for this study.

3.5 Sample Size

Based on efficacy results from study A-101-WART-203 with the same active study medication and treatment plan specified in the current protocol for a comparable subject population and indication, an overall per-wart clearance rate of at least 26% for Active compared with 9% for Vehicle is anticipated. Assuming a distribution of treated warts per subject comparable to the Phase 2 study results, statistical modeling provided a forecast of an advantage in subject responder rates (all treated warts with PWA=0) for Active over Vehicle groups of at least 11 percentage points. This leads to a required sample size for the current study of approximately 500 subjects to provide greater than 90% power to achieve statistical significance in the primary efficacy analysis. This projection takes into account the impact of the anticipated subject dropout rate on power for the primary efficacy analysis, based on the Intent-to-Treat (ITT) population and planned missing data imputation procedures.

Additionally, a simulation was conducted in a manner that accounted for the distribution of multiple warts, the treatment effect in clearance between the different warts and the correlation in wart clearance among multiple warts. Specific details of the simulation are described in Appendix C. Results of this simulation indicate that data from 500 subjects (250 per group) will provide 97% power to demonstrate the superiority of 45% A-101 topical solution over vehicle for the primary endpoint.

3.6 Changes in/ Clarifications to the Conduct of the Study or Planned Analysis

The secondary endpoint of mean per subject percent of warts decreased will be analyzed by an analysis of covariance (ANCOVA) including, in addition to the treatment, terms for baseline number of warts as a continuous covariate and study site as a class variable. The protocol mentioned only analysis of variance (ANOVA) with specifying the treatment term to be included in the model.

A Breslow-Day test of the homogeneity of the odds rations across strata for the analysis of the primary endpoint – i.e. across the strata defined by the number of warts assessed at baseline – is added to assess the model.

A PP Population has been added to test the sensitivity of the primary and select secondary analyses to major violations in the protocol.

A “Safety Population” has been added that will be used for all safety summaries as

opposed to the ITT Population mentioned in the protocol. Safety analyses will be performed according to actual treatment received at the first application in the event of dosing error(s).

An error in Section 3.2.4 in version 1.1 of the SAP was corrected where the text in the first sentence incorrectly indicated that site was class variable in the model. The SAS syntax in this same section did not include site and reflects the intended analysis. Site has been removed from the first sentence in Version 1.2. This correction was performed after unblinding the study. An Ad Hoc analysis with the inclusion of site confirmed that there is no meaningful difference between the two models.

4. STATISTICAL SOFTWARE

All data summaries and listings will be performed using SAS® Version 9.4 or higher, under Windows operating system.

5. REFERENCES

1. Brookmeyer and Crowley, A Confidence Interval for the Median Survival Time, Biometrics 38, 29-41, March 1982.

2. Kalbfleisch, J. D. and Prentice, R. L. (1980), The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons, Inc.

3. Wilson, E. H. (1931). The distribution of chi-squared. *National Academy of Sciences*, (pp. 684-688). Washington.

APPENDIX A: MULTIPLE IMPUTATION AND TIPPING POINT ANALYSES

This appendix details the syntax used to perform the sensitivity analyses described in section 3.2.1 of the SAP. The syntax is consistent with SAS version 9.4.

```
proc mi data=PWAtran seed=123 nimpute=20 out=MIout maximum=3 minimum=0
round=1;
   class TREATMENT;
   fcs reg(PWAV2= PWAV8 PWAV9 PWAV10 PWAV13 WartN TREATMENT /details);
   fcs reg(PWAV8= PWAV2 PWAV9 PWAV10 PWAV13 WartN TREATMENT /details);
   fcs reg(PWAV9= PWAV8 PWAV2 PWAV10 PWAV13 WartN TREATMENT /details);
   fcs reg(PWAV10= PWAV8 PWAV9 PWAV2 PWAV13 WartN TREATMENT /details);
   fcs reg(PWAV13= PWAV8 PWAV9 PWAV10 PWAV2 WartN TREATMENT /details);
   var PWAV2 PWAV8 PWAV9 PWAV10 PWAV13 TREATMENT WARTN;
run;

*** PWAV2 PAWV8 PWAV9 PWAV10 PWAV10 PWAV11 PWAV12 and PWAV13 ***
*** represent the maximum of PWA scores from all warts assessed ***
*** at baseline and analysis visits 8 through 13 as defined in ***
*** Appendix B. WARTN is the number of warts assessed at baseline ***;

**** Assign Response and Strata ****;

data MIresp;
  set MIout;
  if PWAV10 =0 then response=1;
  else response=2;
  WARTS=wartN;
  if WARTS gt 3 then WARTS=4;  ** Greater than 3 warts **;

proc sort;
  by _imputation_;

*** Obtain Mantel-Haenszel estimate of the common odds ratio and General
Association Statistic adjusted for baseline score category ***;
PROC FREQ DATA=MIresp;
TABLES WARTS*TREATMENT*RESPONSE / CMH;
output out=stats cmh;
BY _Imputation_;

DATA OR;
  SET stats;
  *** Log-transform odds ratio estimates and obtain standard error from
confidence intervals ***;
  log_or_mh_value=log(_MHOR_);
  log_or_mh_se=(log(U_MHOR)-log(L_MHOR))/(2*1.96);
  *** Apply Wilson-Hilferty transformation to the CMH statistic and
standardize the resulting normal variable;
  cmh_value_wh=
  ((_CMHGA_/DF_CMHGA)**(1/3) - (1-2/(9*DF_CMHGA)))/SQRT(2/(9*DF_CMHGA));
  cmh_sterr_wh = 1.0;
  *** Apply Wilson-Hilferty transformation to the Breslow Day statistic
and standardize the resulting normal variable ****;
```

```
  BD_value_wh=
  ((_BDCHI_/DF_BDCHI)**(1/3) - (1-2/(9*DF_BDCHI)))/SQRT(2/(9*DF_BDCHI));
  BD_sterr_wh = 1.0;

*** Combine transformed OR estimates;
PROC MIANALYZE DATA=OR;
ODS OUTPUT PARAMETERESTIMATES=LOG_EST;
MODELEFFECTS log_or_mh_value;
STDERR log_or_mh_se;

*** Back-transform OR combined values;
DATA BACKTRANS;
   SET LOG_EST;
   Estimate_back = EXP(ESTIMATE); *Pooled odds ratio;
   LCL_back=Estimate_back*EXP(-1.96*STDERR); *Pooled lower limit;
   UCL_back=Estimate_back*EXP(+1.96*STDERR); *Pooled upper limit;

*** Combine CMH General Association Results;
PROC MIANALYZE DATA=OR;
ODS OUTPUT PARAMETERESTIMATES=cmh_wh;
MODELEFFECTS cmh_value_wh;
STDERR cmh_sterr_wh;

*** Compute combined CMH P-value as one-sided p-value from normal test
from MIANALYZE ***;
DATA cmh_wh_p;
SET cmh_wh;
IF tValue > 0 THEN ProbCMH = Probt/2;
ELSE ProbCMH = 1-Probt/2;

*** Combine Breslow-Day Results;
PROC MIANALYZE DATA=OR;
ODS OUTPUT PARAMETERESTIMATES=bd_wh;
MODELEFFECTS bd_value_wh;
STDERR bd_sterr_wh;

*** Compute combined Breslow-Day P-value as one-sided p-value from
normal test from MIANALYZE ***;
DATA bd_wh_p;
SET bd_wh;
IF tValue > 0 THEN ProbBD = Probt/2;
ELSE ProbBD = 1-Probt/2;

****** Calculate Binomial Response Rates *****;

data MIresp2;
  set MIresp;
  output; *** For each Stata ***;
  Warts=9;
  output; *** All Strata combined ***;

** Create Dummy Data that has all both responder and non-responders **;
***** for each strata and treatment but weight will be zero **********;
proc sort data=miresp2 nodupkey out=trt_strat(keep=treatment warts
_imputation_);
```

```
by treatment warts _imputation_;

data zeros;
  set trt_strat;
  wt=0;
  Response=1; output;
  Response=2; output;

**** Add the Zero Weight Records to the Data ***;
data MIrespWT;
  set MIresp2 zeros;
  if wt=. then wt=1;

proc sort;
  by treatment warts _imputation_;

PROC FREQ DATA=MIrespWT;
TABLES RESPONSE / BINOMIAL;
output out=Bin_prop binomial;
BY treatment warts _Imputation_;
*** Use zeros option in weight statement ***;
weight wt/zeros;

*** Combine proportion estimates ***;
proc means data=bin_prop;
  by treatment warts;
  var N _bin_;
  output out=Prop mean=N Prop;

data Props;
   set Prop;
   *** Calculate Estimated Number of Responders (Does not need to be an
integer) ***;
   N_Resps=N*Prop;

data All;
  merge props(keep=TREATMENT WARTS N Prop N_Resps)
        backtrans(keep=Estimate_back lcl_back ucl_back)
        cmh_wh_p(keep=ProbCMH)
        bd_wh_p (keep=ProbBD);
            label
            TREATMENT='Treatment Arm'
        WARTS='Number of Warts at Baseline (9=All Levels)'
        N='Subjects'
        N_Resps='n Responders'
        Prop='Proportion Responders'
        estimate_back='MH OR'
        LCL_back='Lower Limit MH OR'
        UCL_back='Upper Limit MH OR'
        ProbCMH='CMH P-Value'
        ProbBD='Breslow-Day P-Value';
run;
```

APPENDIX B: ASSIGNMENT TO ANALYSIS VISIT

Due to the possibility of deviation from the protocol schedule as well as to incorporate data from unscheduled and/or early termination visits, analysis visits are defined below based on relative study day. These analysis visits are based on the nominal, scheduled visits in the protocol but allow for wider windows (more days around the target study day). Analysis visit will be used in analyses performed by or on specific visit as well as for the identifiers on data listings. Both nominal visits as collected on the CRF and analysis visits will be retained in the SAS ADaM datasets.

Analysis Visit	Study Day Range	Target Study Day
V3 – Day 8	2 - 11	8
V4 – Day 15	12 - 18	15
V5 – Day 22	19 - 25	22
V6 – Day 29	26 - 32	29
V7 – Day 36	33 - 39	36
V8 – Day 43	40 - 46	43
V9 – Day50	47 - 53	50
V10 – Day 60	54 - 69	60
V11 – Day 78	70 - 97	78
V12 – Day 106	98 - 121	106
V13 – Day 137	≥122	137

In the event that more than one observation falls into the same study day range, the value closest to the "target day" is selected. If there's a tie in the absolute distance from the target day, the observation with the greater cycle day is selected.

APPENDIX C: SIMULATION DETAILS FOR POWER COMPUTATION

A 27-parameter multinomial distribution was used to simulate the primary endpoint in the Phase 3 WART studies. The 27 parameters corresponding to the 27 possibilities for the number of warts that are cleared at Visit 10 is shown in Table 2.

Table 2: Grid of 27-Parameters Corresponding to the Number of Possible Warts Cleared

Number Cleared	Number of Treated Warts					
	1	2	3	4	5	6
0	X1	X3	X6	X10	X15	X21
1	X2	X4	X7	X11	X16	X22
2		X5	X8	X12	X17	X23
3			X9	X13	X18	X24
4				X14	X19	X25
5					X20	X26
6						X27

The probability estimates for the multinomial parameters were derived from a Bayesian posterior Dirichlet distribution that resulted from updating a non-informative Jeffreys conjugate prior with the data from the A101-WART-203 Visit 10 data. Random probability estimates were drawn from the Bayesian posterior distribution for each of the 10,000 simulations. Then a random event was drawn from the 27-parameter multinomial using simulation specific probabilities for each of the 10,000 simulations.

In each individual simulation, the number of warts cleared in the first (1 baseline wart), second (2 baseline warts) and third (3 baseline warts) strata were obtained from X2, X5 and X9, respectively, as defined in the grid in Table 2. The number of cleared warts in the 4th striatum was obtained from the sum of X14, X20 and X27. The odds ratio between A101 45% and vehicle was tested using a CMH stratified by these 4 strata. The number of positive simulation results divided by the total number of simulations was used to compute power.