Statistical Analysis Plan

A Phase 3, Multicenter, Randomized, Double–Masked and Placebo–Controlled Study Evaluating the Efficacy of a 5.0% Concentration of Lifitegrast Ophthalmic Solution Compared to Placebo in Subjects with Dry Eye Currently Using Artificial Tears (OPUS-2)

1118–DRY–300

Sponsored by:

SARcode Bioscience, Inc. (a wholly owned entity of Shire Pharmaceuticals)

1000 Marina Blvd, Suite 250
Brisbane, CA 94005

Version: 1.0

Date Approved: 08 Aug 2013

Table of Contents

List of Tables

List of Abbreviations

AE	adverse event
ANOVA	Analysis of Variance
BCVA	best corrected visual acuity
BID	bis in die (two times daily)
CRF	case report form
ET	early termination
ICH	International Conference on Harmonisation
ICS	inferior corneal staining (Also inferior corneal fluorescein staining)
ITT	Intent-to-Treat
LOCF	last observation carried forward
LogMAR	minimum angle of resolution
MedDRA	Medical Dictionary for Regulatory Authorities
OD	right eye
OSDI	Ocular Surface Disease Index
OU	oculus uterque (each eye or both eyes)
SAE	serious adverse event
SD	standard deviation
SOC	system organ class
STT	Schirmer Tear Test
TEAE	treatment-emergent adverse event
VAS	Visual Analogue Scale
WHO	World Health Organization

1 Study Description

The purpose of this study is to confirm the efficacy and safety of a 5.0% concentration of lifitegrast ophthalmic solution as compared to placebo in the treatment of the signs and symptoms of dry eye following 12 weeks of BID dosing. Lifitegrast ophthalmic solution is an antagonist of lymphocyte function antigen-1 (LFA–1) and is under investigation for the treatment of the signs and symptoms of dry eye in subjects with dry eye disease.

1.1 Objectives

1.1.1 Efficacy

Primary Efficacy Objectives

- To evaluate the efficacy of lifitegrast ophthalmic solution (5.0%) compared to placebo in the treatment of dry eye in subjects currently using artificial tears as assessed by the co-primary endpoints of:
  - Sign: inferior corneal fluorescein staining (ICS) score (0–4 point scale) measured by mean change from baseline to Day 84 in the designated study eye
  - Symptom: eye dryness score (0–100 point visual analogue scale (VAS), OU) measured by mean change from baseline to Day 84

Secondary Efficacy Objectives

- To evaluate the efficacy of lifitegrast ophthalmic solution (5.0%) compared to placebo in the treatment of dry eye in subjects currently using artificial tears as assessed by the secondary endpoints of:
  - Sign: total corneal fluorescein staining score (0–12 point scale) measured by mean change from baseline to Day 84 in the designated study eye
  - Sign: nasal lissamine staining score (0–4 point scale) measured by mean change from baseline to Day 84 in the designated study eye
  - Symptom: eye discomfort score (0–100 point VAS, OU) measured by mean change from baseline to Day 84
  - Symptom: ocular discomfort score (0–4 point scale) measured by mean change from baseline to Day 84 in the designated study eye

1.1.2 Safety Objectives

To evaluate the safety and tolerability of lifitegrast ophthalmic solution (5.0%) compared to placebo in subjects with dry eye when administered BID for 84 days.

1.2 Study Design

Multicenter, randomized, prospective, double–masked, placebo–controlled, parallel–arm design with block enrollment stratified by ICS score (≤ / > 1.5) and eye dryness score (< / ≥ 60). Approximately 700 subjects will be randomly assigned to one of two treatment groups (1:1) to receive either lifitegrast ophthalmic solution (5.0%) or placebo solution as topical ophthalmic drops administered bilaterally BID for 84 days (12 weeks). Subjects, sponsor, CROs and site personnel will be masked to treatment assignments.

The screening period consists of two visits, Visit 1 (Day -14 ± 3) and a confirmatory visit (Visit 2, Day 0). For all endpoints where measurement was scheduled for Visit 2, that visit will serve as the baseline assessment. Subjects must have an eye dryness score ≥ 40 (0-100 point VAS), ICS score of ≥ 0.5 point (0–4 point scale with allowance for 0.5 point increments), and Schirmer Tear Test (STT) ≥ 1 and ≤ 10 mm in at least one eye at Visit 1 (screening Visit 1) and replicate these scores in the same eye at Visit 2 (confirmatory screening Visit 2) in order to be eligible for the study. The eye that meets these criteria will be the designated study eye. If both eyes meet the criteria above, the eye with the greater ICS score at Visit 2 will be designated as the study eye. If both eyes have equal ICS, then the eye with the lowest STT value at Visit 2 will be designated as the study eye. If both eyes have equal ICS and STT scores at Visit 2, the right eye (OD) will be designated as the study eye.

At Visit 1, qualifying subjects will be dispensed sufficient placebo supply (single-use vials) to last until Visit 2 and will be educated on self-administration of placebo. Subjects will be instructed to self–administer one drop BID in each eye in the morning and the evening just prior to bedtime until screening Visit 2.

Randomized subjects will receive their initial dose of study drug (a single drop, approximately 50µL/drop volume, OU, from the same vial) at the study site by trained study personnel 30 ± 15 minutes following the last dose of placebo drops administered by trained study personnel. Only a single dose of randomized study drug will be administered OU from the same vial on Day 0. Prior to discharge from the study site on Visit 2 (Day 0), randomized subjects will be educated on self–administration of study drug and provided a supply of drug to cover the period until their next visit.

A Schedule of Events is displayed in Table 3.

1.3 Method of Assigning Subjects to Treatment Groups

Subjects are randomly assigned to receive 5.0% lifitegrast ophthalmic solution or placebo solution based on a 1:1 ratio [lifitegrast (5.0%): placebo] within the randomization strata using permuted blocks.

Randomization is centralized across study centers, stratified by Visit 2 ICS score (≤ / > 1.5) and by eye dryness score (< / ≥ 60) in the study eye.

These strata are designed such that approximately equal numbers of subjects are expected within each stratum. However, the number of subjects randomized in the study is not forced to be equal in each stratum.

1.4 Masking

This is a double-masked study. Study personnel, including sponsor staff, will be masked with regard to treatment assignments.

1.5 Sample Size

For the primary ocular sign, a 0.25 unit difference between treatment arms has been assumed for the mean change from baseline to Day 84 in ICS, with a common SD of 0.95 units. Under these assumptions, a sample size of 350 per group will yield approximately 93% power to show a significant difference at $\alpha = 0.05$ level under a two-sample t-test.

For the primary ocular symptom, a 10.0 unit difference between the treatment arms at Day 84 in the mean eye dryness score has been assumed, with a common SD of 40 units. Under these assumptions, a sample size of 350 per group will yield approximately 91% power to show a significant difference at $\alpha = 0.05$ level under a two-sample t-test.

It is expected that no subjects will be excluded from the primary analysis due to missing data given the proposed primary analysis method.

2 Statistical Methods

2.1 Populations Analyzed

Three analysis populations will be defined. These populations are:

- Randomized population – All randomized subjects.
- Intent-to-Treat (ITT) – The ITT population is defined as all randomized subjects who receive at least one dose of study medication. This will be the primary efficacy analysis population. This definition is consistent with the full analysis population defined following the intention-to-treatment principle as described in ICH-E9.
- Safety Population – The Safety population includes all randomized subjects who receive at least one dose of study medication.

The ITT and Safety populations are identical for this study. Analyses conducted using the ITT population will be based upon treatment assigned while analyses conducted using the Safety population will be based upon the treatment received.

2.2 Study Drug Dosing and Compliance

Vials will be counted at each follow-up visit with respect to: number dispensed at last visit, number returned unused, number returned used and number missing. Compliance will be calculated as the number of used vials returned divided by two times the number of days of dosing since the last visit. Compliance will also be calculated over the trial as a whole. Subjects will be classified as compliant if the compliance rate is between 0.8 and 1.2, inclusive.

2.3 Study Endpoints

See Table 3 for a Schedule of Events. For signs, measurements will be made for each eye and the analysis will be restricted to data captured for the study eye defined prior to randomization. For symptoms, an overall response may only be captured (eg, VAS and OSDI) rather than an eye-specific response. When the analysis is restricted to the study eye, this will be called out as each measure is discussed in the following.

2.3.1 Primary Efficacy Outcome Measures

- Ocular sign – change from baseline in ICS score (0-4 point scale) to Day 84 in the designated study eye
- Ocular symptom – change from baseline in eye dryness score (0–100 point VAS, OU) to Day 84

2.3.2 Secondary Efficacy Outcome Measures

- Ocular signs
  - Change from baseline in total corneal fluorescein staining score (0–12 point scale derived from the sum of the regions) to Day 84 in the designated study eye
  - Change from baseline in nasal lissamine score (0–4 point scale) to Day 84 in the designated study eye
- Ocular symptoms
  - Change from baseline in eye discomfort score (0–100 point VAS, OU) to Day 84
  - Change from baseline in ocular discomfort score (0–4 point scale) to Day 84 in the designated study eye

2.3.3 Tertiary Efficacy Outcome Measures

- Ocular signs – measures are made at each study visit. The designated study eye will be used for endpoint value in analyses
  - Corneal fluorescein staining score for each of the following regions:

- Inferior, superior, central, (0–4 point scale for each region) and total (0–12 point scale derived from the sum of the regions)

- Lissamine staining score for each of the following regions:
  - Nasal, temporal (0–4 point scale for each region), and total, (0-8 point scale derived from the sum of the regions)
- Conjunctival redness score (0-4 point scale)
- Schirmer Tear Test without anesthesia (mm/5 min)

- Ocular symptoms – measures are made at each study visit.
  - 7-item VAS score (0–100 point scale, OU) for each of the following items: burning/stinging, itching, foreign body sensation, eye discomfort, eye dryness, photophobia, pain
  - Ocular discomfort score (0–4 point scale) in the designated study eye
  - Ocular Surface Disease Index scores – 12 items with each individual item scored on a 0–4 point scale or "N/A" (Not Applicable). The scores used for data analysis will be derived as follows:
    - Total OSDI score (0–100 point scale) is derived as the sum of the non-missing items multiplied by 25 divided by the number of non-missing items
    - Symptom subscale OSDI score (0–4 point scale) derived as the mean of non-missing items #1 through #5
    - Visual-related function subscale OSDI score (0–4 point scale) derived as the mean of the non-missing items #6 through #9
    - Environmental trigger subscale OSDI score (0–4 point scale) derived as the mean of the non-missing items #10 through #12

    The OSDI scoring is described in Schiffman (2000). The scoring for the subscales differs from Schiffman as they are not multiplied by 25 to keep the resulting scores on the same scale as the individual components; this will not affect the resulting statistical tests.

2.3.4 Safety assessments by visit

The following measures are made at study visits according to the study schedule of events:

- Drop comfort (0–10 point scale; 0=very comfortable, 10=very uncomfortable) in each eye is assessed at the time of drop instillation (time=0) and then at 1, 2 and 3 minutes.
- Best corrected visual acuity (BCVA) is assessed for each eye before drop instillation. At Visit 5 only, a post-instillation assessment is also measured. The logMAR scoring

system will be used for the BCVA (a score of 0 corresponds to 20/20 vision). Vision is to be assessed using current correction. If current correction is not available, vision will be assessed using the pinhole occluder.

- Slit lamp biomicroscopy is used to assess six (6) anatomic anterior segment regions for each eye before drop instillation. At Visit 5 only, a post-instillation assessment is also measured. The six anatomic regions examined are as follows:
    - Cornea, conjunctiva, iris, anterior chamber, lens, lid

  Each anatomic region of the eye is rated by one of the following terms:

    - Normal
    - Abnormal - Not Clinically Significant
    - Abnormal - Clinically Significant
- Dilated fundoscopy (Visits 1 and 5 only) is used to assess five (5) anatomic posterior segment regions following drop instillation. The five anatomic regions are as follows:
    - Vitreous, retina, macula, choroid, optic nerve

  Each anatomic region of the eye is rated by one of the following terms:

    - Normal
    - Abnormal - Not Clinically Significant
    - Abnormal - Clinically Significant
- The incidence and severity of ocular adverse events (AEs) and the incidence and severity of non-ocular AEs. Ocular and non-ocular AEs will be also recorded at each visit.

2.4 Statistical Assessment of the Trial Objectives

2.4.1 Primary Analysis

Inferior Corneal Fluorescein Staining Score - Sign

The primary analysis of change from baseline to Day 84 in ICS will be performed using a stratified two-sample t-test (ie, analysis of variance) comparing lifitegrast ophthalmic solution to placebo in the ITT population with Last Observation Carried Forward (LOCF).

Eye Dryness - Symptom

The primary analysis of change from baseline to Day 84 in the eye dryness score will be performed using a stratified two-sample t-test (ie, analysis of variance) comparing lifitegrast ophthalmic solution to placebo in the ITT population with LOCF.

The stratified two-sample t-test will be referred to as analysis of variance (ANOVA) in this document. The stratification factors used for randomization will be used for this analysis. The individual strata will contribute to the overall analysis proportionate to their size as suggested by Anello, et al (2005). The ANOVA model used to conduct the protocol specified primary treatment comparison will include treatment, strata and the interaction between treatment and strata.

The treatment effects will be estimated from the ANOVA model as described below.

The analyses for both endpoints will reflect the stratification of randomization through the use of a cell means model. The cells will be formed by the 8 combinations of the 4 strata and two treatments. In the cell means approach, the following mean estimates and differences are calculated (Table 1).

Table 1: Stratum Level Estimators*

Strata		Treatment Group		
Inferior Staining Score	Dry Eye Score	Lifitegrast (5.0%)	Placebo	Difference
≤ 1.5	< 60	$\bar{x}_{11}$	$\bar{x}_{12}$	$d_1=\bar{x}_{11}-\bar{x}_{12}$
	≥ 60	$\bar{x}_{21}$	$\bar{x}_{22}$	$d_2=\bar{x}_{21}-\bar{x}_{22}$
> 1.5	< 60	$\bar{x}_{31}$	$\bar{x}_{32}$	$d_3=\bar{x}_{31}-\bar{x}_{32}$
	≥ 60	$\bar{x}_{41}$	$\bar{x}_{42}$	$d_4=\bar{x}_{41}-\bar{x}_{42}$

*Estimates are the change from baseline for the ICS score and eye dryness score.
ICS=inferior corneal staining

For the ij-th mean, there is a corresponding sample size (n_{ij}) and estimated SD s_{ij}. For combining the treatment differences over strata, the weight for the i-th stratum is $w_i=n_i/n$ where n_i is the total sample size for the i-th stratum ($n_i=n_{i1}+n_{i2}$, i=1, 2, 3, 4) and n is the sum of the n_i over the four strata. Refer to Anello, et al (2005) for a discussion of the issues associated with the analysis of stratified (multicenter) trials and Lin (1999) for a description of using the total sample size within each stratum as the weight to combine stratum-specific estimators. This estimated treatment effect will be approximately the same as the observed mean for each treatment group ignoring stratification.

The overall difference and variance for the overall difference will be calculated as shown below.

Overall treatment effect

$$d=\bar{x}_{.1}-\bar{x}_{.2}=\sum_{i=1}^{4} w_i[\bar{x}_{i1}-\bar{x}_{i2}]$$

Estimated variance overall treatment effect

$$s_{\hat{d}}^2 = \sum_{i=1}^{4} w_i^2 s_p^2 \left[\frac{1}{n_{i1}} + \frac{1}{n_{i1}}\right]$$

The overall estimated variance, s_p^2, is the mean square error calculated as the pooled variance over the 8 cells formed by the combination of treatment and stratum (ie, mean square error taken from a simple cell means ANOVA).

Calculations for the ANOVA estimates (differences, standard errors and confidence intervals) will be performed using PROC MIXED in SAS via the LSMEANS statement with the OM option and weights proportionate to stratum sample size. The model will be the cell means model (means estimated by treatment within stratum).

The primary inference will be based on the p-values from the ANOVA model for the change from baseline to Day 84 in ICS staining and for the change from baseline to Day 84 in eye dryness.

The interaction between treatment and strata is included in the model to allow for inconsistency across strata to be examined. The study is not powered to detect the effect of interaction between treatment and strata.

2.4.2 Secondary Endpoint Analyses

The secondary efficacy endpoints will be analyzed using the same ANOVA model as for the co-primary efficacy endpoints comparing lifitegrast ophthalmic solution to placebo in the ITT population with LOCF. Hochberg's (1988) procedure will be applied to control the type I error rate at 5% level across the two symptom secondary endpoints (ie, eye discomfort score and ocular discomfort score). Additionally, Hochberg's procedure will be applied to the two secondary sign endpoints, total corneal staining and nasal lissamine staining score at the type I error rate of 5% level. In applying Hochberg's procedure, nominal p-values will be produced and the adjustment applied at the time the p-values are interpreted for statistical significance.

To apply Hochberg's procedure of multiple testing for the two secondary endpoints related to sign, the higher p-value will be compared with 5% level. If this p-value is less 5%, then both the secondary sign endpoints will be declared significant. If the higher p-value is not less than 5%, then the smaller p-value will be compared with 2.5%. If the smaller p-value is less than 2.5%, then the secondary sign endpoint corresponding to this smaller p-value will be declared significant. Similarly Hochberg's procedure of multiple endpoints will be applied to the two symptoms endpoint.

2.4.3 Sensitivity Analyses for the Co-Primary and Secondary Endpoints

The co-primary efficacy endpoints and secondary endpoints will also be analyzed using additional statistical methods as sensitivity analyses. The planned sensitivity analyses will consist of repeating the primary analysis using observed data, a stratified rank-based test (ie, Wilcoxon) with LOCF and repeated measures ANOVA (no imputation). The stratified rank-based test will consist of repeating the primary analysis (LOCF), using the overall ranks rather than the observed data. The repeated measures analysis will model the outcome as a function of the randomization strata, treatment and time. In this model, all the model terms will be treated as categorical variates with a common treatment effect assumed over time and the randomization strata (ie, main effects model). An unstructured covariance matrix will be used for this analysis.

Subjects assigned to the incorrect strata during randomization will be analyzed using the stratification used for the randomization. The co-primary efficacy endpoints will also be analyzed using the strata which would have been the correct strata based on the baseline characteristics of the subjects.

2.4.4 Tertiary Endpoint Analyses

Tertiary efficacy endpoints will be analyzed using the same statistical methods as for the primary endpoint (ie, ANOVA using LOCF). Tests of significance will be limited to the change from baseline to Day 84 time point. As the tertiary endpoint analyses are viewed as supportive, the Wilcoxon rank sum test and the repeated measures ANOVA will not be performed. There will be no adjustments for multiplicity across the tertiary endpoints.

2.5 Study Day and Visit Windows

Following the protocol, Study Day is defined below:

Study Day = event date-first dosing date (Day 0 allowed under protocol)

Assessments will be assigned to visits based upon the date the assessment took place regardless of the CRF page completed. Assessments will be mapped to visits as outlined in Table 2. Should more than one assessment exist within a given visit window, the value closest to the scheduled visit should be used (choose last if equally close or scheduled if both scheduled and unscheduled visits are present within the window).

For subjects terminating the trial early, data is collected on the unscheduled/early termination CRF. The analysis visit window and analysis flag will be assigned for data collected both at the scheduled and unscheduled visits.

Table 2: Visit Windows (Study Day Based)

Visit	Start of Window	End of Window
Screening – Visit 1	Informed consent signed	-1
Baseline – Visit 2 Pre-randomization	0	Time of randomization
Visit 2 Post-randomization	Time of Randomization	0
Day 14 – Visit 3	1	27
Day 42 – Visit 4	28	62
Day 84 – Visit 5	63	Last recorded data

For subjects with a missing reference date, the study day will also be missing.

Table 3: Schedule of Events

Procedure	Visit 1 Day – 14 ± 3	Days –13 – –1	Visit 2 Day 0			Days 1–13	Visit 3 Day 14 ± 3	Days 15–41	Visit 4 Day 42 ± 4	Days 43–83	Visit 5 Day 84 ± 8	ET/ UNS[1]
			Pre-Random-ization	Random-ization	Post-Random-ization							
Informed Consent	X											
Demographic data	X											
Height and Weight (Subject-reported)	X											
Medical History/Medication History[2]	X											
Concomitant Medication Review			X				X		X		X	X
Inclusion/Exclusion Criteria[3]	X		X									
Urine Pregnancy Test[4]	X		X				X		X		X	X
Subjective Measures												
Visual Analogue Scale (VAS)[5]	X		X				X		X		X	X
Ocular Discomfort Score (ODS)[6]	X		X				X		X		X	X
Ocular Surface Disease Index (OSDI)	X		X				X		X		X	X
Drop Comfort[7]					X		X		X		X	
Objective Measures												
Best Corrected Visual Acuity (BCVA)[8]	X		X				X		X		2X	X
Slit lamp Biomicroscopy[9]	X		X				X		X		2X	X
Conjunctival Redness Score[10]	X		X				X		X		X	X
Corneal Staining (fluorescein)[11]	X		X				X		X		X	X
Conjunctival Staining (lissamine)[12]	X		X				X		X		X	X
Schirmer Tear Test (w/o anesthesia)	X		X				X		X		X	X
Dilated Fundoscopy[13]	X										X	X

Procedure	Visit 1 Day – 14 ± 3	Days –13 – –1	Visit 2 Day 0			Days 1–13	Visit 3 Day 14 ± 3	Days 15–41	Visit 4 Day 42 ± 4	Days 43–83	Visit 5 Day 84 ± 8	ET/ UNS[1]
			Pre-Random-ization	Random-ization	Post-Random-ization							
Study Therapy												
Open-label Placebo Administration at Study Site[14]	X		X									
Placebo Dispensation (Open–label) [15]	X											
Open-label Placebo Administration at Home [14]		X										
Placebo Vial Collection[16]			X									X
Randomization[17]				X								
Study Drug Administration at Study Site[14]					X		X		X		X	
Study Drug Dispensation[18]					X		X		X			
Study Drug Administration at Home [18]						X	X	X	X	X		
Study Drug Collection[18]							X		X		X	X
Adverse Event Assessment[19]	X	X	X		X	X	X	X	X	X	X	X
Study Exit											X	X

[1] ET/UNS: Early Termination and Unscheduled Visit Assessments.

[2] Only significant non–ocular medical history during the past year; only medications taken within the past 60 days. Artificial Tear Use should be within the past 30 days prior to Visit 1.

[3] Subjects must replicate the following findings in the *same eye* at Visits 1 and 2 in order to be considered for further study eligibility:

(1) Inferior corneal fluorescein staining (ICS) score ≥ 0.5 points (0–4 point scale with allowance for 0.5 point increments), and

(2) Schirmer Tear Test (STT) without anesthesia ≥ 1 and ≤ 10 mm, at Visits 1 and 2

If both eyes meet the two criteria above, the eye with the greater in ICS at Visit 2 will be selected as the study eye. If both eyes have an equal ICS score at Visit 2, the eye with the lowest STT value at Visit 2 will be designated as the study eye. If both eyes have equal score in ICS and equal STT scores at Visit 2, the right eye (OD) will be selected as the study eye.

[4] Women of childbearing potential only.

[5] 7-item Visual Analogue Scale (0–100 point scale, OU): burning/stinging, itching, foreign body sensation, eye discomfort, eye dryness, photophobia, pain.

[6] Ocular discomfort score (0–4 point scale).

[7] Drop comfort assessments (0–10 scale) are obtained in each eye immediately, then at 1, 2, and 3 minutes following instillation of the study drug. At Visits 2, 3, 4, and 5 – drop comfort assessments will be conducted following the initial dosing of the day. For all drop comfort assessments, subjects will be administered study drug by trained study personnel.

[8] A Best Corrected Visual Acuity (BCVA) assessment will be measured prior to open-label placebo administration at Visit 1 and prior to study drug administration at Visits 2, 3, 4 and 5. At Visit 5, a second BCVA assessment will also be measured after the final study drug dose has been administered at the site by trained study personnel.

[9] A slit lamp examination will be performed prior to open-label placebo administration at Visit 1 and prior to study drug administration at Visits 2, 3, 4 & 5. At Visit 5, a second slit lamp examination will be performed after the final study drug dose has been administered at the site by trained study personnel.

[10] 0–4 point scale with allowance for 0.5 point increments.

[11] 0–4 point scale with allowance for 0.5 point increments using 3 corneal regions (superior, central, inferior).

[12] 0–4 point scale with allowance for 0.5 point increments using 2 conjunctival regions (nasal, temporal).

[13] Dilated fundoscopy should be performed at the end of Visits 1 and 5.

[14] At Visit 1, after achieving a positive response, as defined by the protocol and noted above in footnote #4, subjects will self-administer open–label placebo drops in both eyes, for training purposes, at the study site under the supervision of trained study personnel **30 ± 15 minutes** following the last study assessment. **Only a single dose of placebo drops will be administered OU, from the same vial, on Day –14 (Visit 1).**

Following the screening procedures at Visit 2, all subjects will receive their last dose of placebo drops (open–label, 2 drops *each* eye, approximately 50 µL/drop volume, OU, from the same vial) at the study site by trained study personnel **20 ± 15 minutes** following the last study assessment. All subjects having a positive response (as defined above) and meeting all other screening eligibility criteria for Visit 2 will be randomized to one of two treatment arms.

At Visit 2, following randomization, subjects will be administered the initial dose of randomized study drug **30 ± 15 minutes** after the last dose of placebo drops administered by trained study personnel. **Only a single dose of randomized study drug will be administered OU on Day 0 (Visit 2).**

At Visits 3 and 4, subjects will receive their first dose of the day by trained study personnel following a **15 ± 15 minutes** after **the last study assessment**; the second dose of the day will be administered at home, by the subject, in the evening just prior to bedtime.

At Visit 5, subjects will receive their final study drug dose (a single drop, approximately 50 µL/drop volume, OU, from the same vial) at the study site by trained study personnel 15 ± 15 minutes following the last study assessment.

Subjects will then perform a drop comfort assessment immediately and then at 1, 2 and 3 minutes following initial dosing at Visit 2 (post-randomization study drug dosing) and Visit 3, 4 and 5.

[15] Open-label placebo vials will be dispensed at Visit 1 and will be self–administered by the subjects at home on Days –13 – –1. The open-label placebo vials will be collected at Visit 2. Upon return of the open-label placebo vials, site staff must confirm subjects have NOT administered their morning dose at home for Visit 2.

[16] Placebo vials will be collected at Visit 2. Site staff must confirm subjects have NOT administered their morning placebo dose at home for Visit 2.

[17] All subjects having a positive response (as defined above in footnote #3) and meeting all other screening eligibility criteria after Visit 2 will be randomized.

[18] Randomized study drug vials will be dispensed post-randomization at Visits 2, 3 and 4. The study drug will be collected at the visit following the visit when they were dispensed. Upon study drug return, site staff must confirm subjects have NOT administered their morning dose at home for Visits 3, 4 and 5. Study drug will be self–administered by the subjects at home on Days 1–13, Days 15–41 and Days 43–83.

[19] Adverse event reporting will begin after each subject signs an Informed Consent Form and continue through the end of the last study visit, Visit 5 (Day 84), or early termination assessment.

2.6 Handling of Missing Data

For the efficacy data, subjects will be analyzed either based upon observed data or LOCF. Other data collected, including missing dates, will, in general, not be imputed and will be displayed as observed.

For analysis based on LOCF, data will be taken from the last date collected without regard to analysis window flag.

For imputation of derived variables for LOCF, missing derived variables at a visit will be carried forward rather than carrying forward individual items and then calculating the derived variable. This will ensure that all components for a derived variable reflect data collected at the same visit.

For AEs and concomitant medications, partial dates will be used to classify events as before or after treatment. Where the partial dates do not allow such a classification, the event will be assumed to be after treatment.

2.7 Sensitivity Analyses, Subgroups and Covariates

Descriptive statistics (n, mean and SD based upon observed data) will be provided within the following subgroups for the primary and secondary endpoints: randomization strata, age groups (≥65, ≥75), sex and race.

2.8 Safety Analyses

2.8.1 Safety Endpoints

- Drop comfort
- Best corrected visual acuity
- Slit lamp biomicroscopy
- Dilated fundoscopy

Descriptive statistics will be presented for these measures (by component if multiple measures) at each visit.

2.8.2 Adverse Events

Statistical analyses will be descriptive in nature and will not be routinely tested for statistical significance. Safety analyses will be based upon the clinical database only. Adverse events recorded in the safety database will be reconciled with clinical database prior to database lock. Generally, AEs will be presented by whether or not the AE was an ocular or non-ocular AE.

2.9 Interim Analysis

No interim analyses are planned.

3 Summary Tables Listings and Figures

3.1 General Conventions

Continuous variables will be summarized using descriptive statistics including the number of observations, mean, SD, median, minimum, and maximum values. Categorical variables will be summarized using frequencies and percentages. Hypothesis testing will be performed using two-sided tests at $\alpha = 0.05$ significance level.

For analyses based upon subjects included in either efficacy or safety analyses, age will be calculated relative to the date of randomization. For analyses of screened subjects, age will be calculated using the time informed consent was signed.

For AEs reported on a per-subject basis, medical history and concomitant medications, the denominator for the percentage calculation will be the number of subjects in the analysis set and in the subgroup of interest at risk in each treatment arm.

All p-values will be rounded to 4 decimal places. If a p-value is less than 0.0001 it will be displayed as "< 0.0001".

All non-efficacy summary tables will use the Safety population and all efficacy summary tables will use the ITT population unless specified otherwise. Summaries will be provided by treatment group.

3.2 Subject Disposition and Treatment

The number of subjects will be summarized by treatment arm with respect to the following:

- Randomized
- Treated
- Completed treatment
- Early termination of treatment and the reason

A summary of subjects present at each visit will be presented for the Safety population by treatment arm, treatment status and reason for lost to follow-up.

3.3 Protocol Exemptions and Deviations

3.3.1 Protocol Exemptions

Subjects allowed to enroll in the trial without meeting the inclusion/exclusion criteria will be summarized for the randomized population with the following information: treatment assigned and inclusion/exclusion criteria violated. These exemptions will also be listed.

3.3.2 Violations of Inclusion/Exclusion Criteria

Subjects unintentionally enrolled in the trial without meeting the inclusion/exclusion criteria will be summarized for the randomized population with the following information: treatment assigned and inclusion/exclusion criteria violated.

Subjects assigned to incorrect strata for randomization will be listed along with the treatment assigned and the baseline scores for eye dryness VAS and baseline ICS.

3.3.3 Prohibited Concomitant Medications

Prohibited concomitant medications will be summarized for the Safety population with the following information: treatment assigned and concomitant medication used. This information will be derived from prohibited concomitant medications recorded as protocol deviations in the clinical database.

3.4 Study Treatment

The following information will be presented by treatment arm and overall:

- Duration of treatment calculated as the study day of the last day of randomized study drug,
- Disposition of dispensed vials by visit,
- Overall compliance with dosing by pre-randomization and post-randomization (mean and percent within range 80%-120%), and
- Percent doses instilled at study visits by pre-randomization and post-randomization.

3.5 Demographics and Baseline Characteristics

Descriptive statistics will be provided for the following subject characteristics at baseline.

Demographic variables:

- Age as a continuous variable
- Age by age groups
- Race

- Iris color (study eye)
- Sex
- Ethnicity

Statistical testing between treatment groups will not be performed.

3.5.1 Concomitant Medications

Concomitant Medications will be descriptively summarized based upon the coded values (dictionary WHO Drug March 2012) by treatment for the Safety population. Data summaries will be sorted by the overall frequency of recorded use. Ocular and non-ocular medications will be summarized separately. Medications with end dates prior to dosing will be summarized separately. A table of medications used prior to dosing will also be provided.

3.5.2 Medical History

Medical history with respect to ocular conditions will be descriptively summarized based upon the system organ class (SOC) and coded terms (MedDRA Version 14.1) by treatment and overall for the Safety population. Data summaries will be sorted by the overall frequency of reporting within SOC. Non-ocular conditions will be summarized separately.

3.6 Analysis of Efficacy Endpoints

3.6.1 Primary Efficacy Analysis

The primary analysis results (ANOVA) will be presented for each of the co-primary endpoints. The sensitivity analyses will then be presented: observed data (ie, no LOCF), rank-based test and repeated measures analysis described in Section 2.4.3.

Descriptive statistics for the co-primary endpoints will be presented over time using the observed data (ie, no LOCF). These statistics will be presented both with respect to the data at each time point as well as the change from baseline. The means over time will also be presented graphically.

3.6.2 Secondary Efficacy Analyses

The secondary endpoint analyses will be presented in a similar fashion to the primary endpoint analyses. The p-values displayed will be the nominal p-values. The Hochberg procedure will be applied as part of the interpretation of the results.

3.6.3 Analysis of Tertiary Endpoints

The tests of significance for the tertiary endpoints will be presented for the LOCF data and the observed data for Day 84 only. The confidence interval for the treatment effect for each

tertiary endpoint will be presented graphically (ie, forest plot) for the tertiary endpoints. Where the endpoints are defined with respect to scores and subscores, only the scores will be summarized rather than the individual components.

Descriptive statistics for the tertiary endpoints will be presented over time using the observed data (ie, no LOCF). These descriptive statistics (limited to sample size, mean and SD) will be presented both with respect to the data at each time point as well as the change from baseline. Data will be presented for both the study eye and the non-study eye where the measure was collected for each eye.

3.6.4 Subgroup Analyses

Descriptive statistics, including a graphical display of confidence intervals, will be presented for the subgroups defined in Section 2.7 using the Day 84 data (observed data only) as well as the change from baseline.

3.7 Analysis of Safety Endpoints

For reporting purposes, safety events will be divided into two groups: (1) those that occur prior to treatment (pre-treatment events) and (2) those that occur after the start of randomized treatment (treatment-emergent adverse events - TEAEs). Events will typically be summarized separately based upon whether or not they are considered ocular events.

In the event of multiple occurrences of the same AE with the same preferred term in one subject, the AE will be counted once as the occurrence with the highest severity. All of the tables described in the following are conducted using the Safety population. In listings, if the treatment received is different than the treatment assigned, this will be noted.

This section primarily describes information to be presented for events that occur after the start of treatment. Section 3.7.6 describes data summaries for events recorded prior to treatment.

3.7.1 Protocol Specific Safety Endpoints

The safety endpoints will be summarized descriptively by treatment arm at each time point where an assessment was made. Where the measure is repeated for each eye, results will be presented for each eye.

3.7.2 Brief Summary of AEs

A brief summary of TEAEs, serious TEAEs, deaths, TEAEs leading to drug discontinuation and summary by severity will be presented by arm and overall. In addition to the overall TEAE summary, this information will be repeated for ocular and non-ocular AEs.

3.7.3 Clinical Adverse Events

The incidence of ocular TEAEs will be tabulated by treatment group, SOC, and preferred term. Treatment-emergent AEs will be classified by SOC and preferred term using MedDRA Version 14.1. These summaries of TEAEs will be produced separately for ocular and non-ocular AEs in three ways:

- Overall (without regard to severity),
- By severity, and
- By relationship to study drug (probably or possibly related or not related).

3.7.4 Serious Adverse Events and Deaths

Treatment-emergent serious adverse events (SAEs) will be summarized separately for ocular and non-ocular AEs by treatment group, SOC, and preferred term. These tables will be repeated with events classified by relationship to study drug (probably or possibly related or not related). Treatment-emergent SAEs will also be listed with the following information: subject ID, treatment assigned, last treatment date, SOC, preferred term, start and stop date of SAE and SAE outcome.

Events with a fatal outcome will be summarized by treatment arm and cause of death.

3.7.5 Premature Discontinuations Due to Adverse Events and/or Laboratory Abnormalities

Treatment discontinuation due to TEAEs or laboratory abnormalities will be summarized by treatment arm and the reason for discontinuation. These discontinuations will also be listed with the following information: subject ID, treatment assigned, first treatment date, last treatment date, and reason for discontinuation.

3.7.6 Events Recorded Prior to Treatment

Pre-treatment events (captured on the AE form with a date/time that is before the first dose of randomized treatment) will be listed. The following information will be presented: subject ID, treatment assigned, coded event, date of event, date of randomization, date of treatment.

3.7.7 Clinical Laboratory Result Analysis

Clinical laboratory samples were not collected for this protocol.

3.7.8 Vital Signs

Vital Signs assessments were not performed as part of this protocol.

3.7.9 Physical Examination

Physical exam data were not collected as part of this protocol.

3.7.9.1 Pregnancies

Pregnancies and outcomes will be listed.

4 References

Anello C, O'Neill RT, Dubey S 2005. Multicentre trials: a US regulatory perspective. *Stat Methods Med Res*; 14(3): 303-18.

Hochberg, Y 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*; 75(4): 800-2.

Lin Z 1999. An issue of statistical analysis in controlled multi-centre studies: how shall we weight the centres? *Statist Med*; 18: 365-73.

Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL 2000. Reliability and validity of the ocular surface disease index. *Arch Ophthalmol*; 118(5): 615-21.